EXAMINATION PAPER SOLUTIONS

Two Hours

Statistical tables are attached

UNIVERSITY OF MANCHESTER

MEDICAL STATISTICS

23 May 2017

09:45-11:45

Answer **ALL** five questions in **SECTION A** (40 marks in total).

Answer **TWO** of the three questions in **SECTION B** (40 marks in total). If more than two questions from Section B are attempted, then credit will be given for the two best answers.

The total number of marks for the paper is 80.

Electronic calculators may be used provided that they cannot store text/transmit or receive information/display graphics

---

A1.

(i) In the context of randomised controlled trials, what is meant by the term bias?

Solution

In the context of a clinical trial, bias is a factor that deviates the estimate of the treatment effect systematically away from the true estimate.

[2 marks]

(ii) Describe two possible sources of bias in clinical trials.

Solution

A brief description of any two of the following (i) selection (ii) allocation (iii) performance (iv) follow-up (v) outcome assessment or (vi) analyses biases

[2 marks]

(iii) Explain what is meant by the term double-blind.

Solution

This is method to reduce bias in a randomised controlled trial, where neither the study participant nor the experimenters knows which of two treatments the participant is receiving

[2 marks]

(iv) Describe two ways in which a trial being double blind might reduce bias?

Solution

Two reasons from: It is advantageous for a trial to be double blind as knowledge of treatment allocation may influence (i) the behaviour of the patient, (ii) the treating health professional or (iii) the assessor of outcome. (i) For example if the patient know which treatment they are receiving it may motivate them to default from treatment or seek alternative treatments. (ii) If the treating health professional know the allocation it may influence choice of secondary treatments. (iii) If the outcome assessor is aware of treatment, allocation there judgement may be bias. For example it may affect a patient’s self-assessment if they know that they have received a placebo or standard treatment.

[2 marks]

[Book work]

[Total 8 marks]

A2.

A randomised controlled trial is being planned to compare a new treatment (T) and a control treatment (C)**.** Suppose the primary outcome measure is continuous and normally distributed. The power to demonstrate a treatment effect τ with a two-sided two sample t-test is given by the expression $1-\Phi\left(z_{\alpha/2}-\frac{\tau\sqrt{n}}{\sigma\sqrt{2}}\right)$ where σ is the known within treatment group standard deviation, n is the sample size of each of two equal size groups, α is the significance level, and Φ is the cumulative density function of a standardised normal distribution.

Suppose one wishes to detect a treatment effect of 5 units and the within treatment group standard deviation has been estimated to be 20 units, estimate the power of a trial with 160 subjects in each treatment group.

Solution

τ= 5 σ= 20 n=160

$$1-\Phi\left(z_{\alpha/2}-\frac{\tau\sqrt{n}}{\sigma\sqrt{2}}\right)=1-\Phi\left(1.96-\frac{5\sqrt{160}}{20\sqrt{2}}\right)=1-\Phi\left(1.96-\frac{\sqrt{80}}{4}\right)=1-\Phi\left(-0.276\right)=\Phi\left(0.276\right)=0.6088$$

because $\Phi\left(-0.276\right)=1-\Phi\left(0.276\right)$.

The power of the study is 60.1% [3 minutes calculation]

[Total 6 marks]

A3.

(i) Illustrate how you might prepare a randomisation list for the first twenty patients in a trial with two treatments using *Block randomisation* with a block size of 4.

Solution

With two treatments, say A and B, one could choose a block size of 4. With this block size there are 6 possible blocks (1) AABB (2) ABAB (3) ABBA (4) BBAA (5) BABA (6) BAAB

To assemble a randomisation list for twenty subjects one would select 5 random numbers between 1- 6 with replacement in sequence, say the numbers 2, 6, 3, 1, 3 from which one could assemble the following list for the first 20 allocations

A,B,A,B | B,A,A,B | A,B,B,A | A,A,B,B | A,B,B,A [4 marks]

(ii) How might you use block randomisation to improve balance between two randomly allocated treatment groups for a dichotomous prognostic factor?

Solution

Block randomisation can be used in conjunction with stratification to obtain balance in a categorical prognostic factor. Separate block randomisation lists are used for each prognostic stratum.

[2 marks]

[Bookwork]

[Total 6 marks]

A4.

(i) Explain what is meant by an equivalence trial.

Solution

Usually the aim of a trial is to detect a difference between the treatments under study, testing whether a new treatment is superior to the existing standard treatment or a placebo. Such trials are called Superiority Trials. In such a trial the null hypothesis is that the average outcome is the same. Equivalence trial are designed to establish that the efficacy of two or more treatments is the same. Therefore in such a trial the null hypothesis is that the average outcome is different.

[3 marks]

(ii) Outline the statistical analysis one could use in a parallel group trial to establish whether a new treatment T is equivalent to a control treatment C for a continuous normally distributed outcome measure *Y*.

Solution

Rather than using a formal significance test, statistical analysis of equivalence trials is usually based on the confidence interval of difference between treatments. Equivalence is established by demonstrating that the confidence interval of the difference lies in the specified range $(-\delta_E, +\delta_E)$. Suppose outcome measure *Y* is continuous and normally distributed with means μ_C and μ_T for the control and new treatment respectively. Rejection of the null hypothesis that H_0: $|\mu_T-\mu_C|> \delta_E$ against the alternative hypothesis H_1: $|\mu_T-\mu_C|\leq \delta_E$ where the $(1-2\alpha)$ confidence interval is within the interval $(-\delta_E, +\delta_E)$, will have a type I error of less than α.

[4 marks]

(iii) A randomised controlled equivalence trial is carried out to test whether a new *generic drug* is as effective as a current *standard drug* for controlling pain. At follow-up this is measured by a 100 mm analogue scale with higher scores representing greater pain. Forty-two patients are randomised to the standard treatment and forty-one to the new generic treatment. The statistical computer package output is given below. A difference of 5 mm was considered by researchers to be the minimum that was clinically important. Using the results in the output test whether the new generic drug is equivalent to the current standard drug specifying the significance level of the test.

```
Two-sample t test with equal variances
------------------------------------------------------------------------------
         |     Obs        Mean    Std. Err.   Std. Dev.   [90% Conf. Interval]
---------+--------------------------------------------------------------------
Standard |      42        35.2     2.79289        18.1    30.49991    39.90009
Generic  |      41        34.1     2.79551        17.9    29.39278    38.80722
---------+--------------------------------------------------------------------
    diff |                 1.1    3.952132               -5.475889    7.675889
------------------------------------------------------------------------------
   diff = mean(Standard) - mean(Generic)                         t =   0.2783
Ho: diff = 0                                     degrees of freedom =       81

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.6093         Pr(|T| > |t|) = 0.7815          Pr(T > t) = 0.3907
```

Solution

The output gives a 90% confidence interval. This can be used to carry out a 5% level test of equivalence. The question states that a 5mm difference on the visual analogue scale was considered to be the minimum clinically important difference. Therefore ± 5 mm is used as the limits of equivalence. From the printout the 90% confidence interval is (-5.48 to + 7.68) which overlaps both limits of equivalence. Therefore it is not possible to reject the null hypothesis in a 5% level test.

[Reading time 1 minute]

[3 minutes calculation time]

[3 marks]

[Total 10marks]

A5.

In meta-analysis suppose $\hat{\theta}_i$ is an estimate of the treatment effect for the i^{th} study and let $Var\left[\hat{\theta}_i\right]$ be its sampling variance.

(i) For the weighted estimate of the overall effect, defined by $\hat{\theta} = \dfrac{\sum_i^k w_i \hat{\theta}_i}{\sum_i^k w_i}$ where w_i are weights, show that $Var\left[\hat{\theta}\right] = \dfrac{\sum_i^k w_i^{\,2}.Var\left[\hat{\theta}_i\right]}{\left(\sum_i^k w_i\right)^2}$.

Solution

$$Var\left[\hat{\theta}\right] = Var\left[\frac{\sum_i^k w_i.\hat{\theta}_i}{\sum_i^k w_i}\right] = \frac{1}{\left(\sum_i^k w_i\right)^2} Var\left[\sum_i^k w_i.\hat{\theta}_i\right].$$

Since the studies are independent, it follows that $Var\left[\sum_i^k w_i.\hat{\theta}_i\right] = \sum_i^k w_i^2.Var\left[\hat{\theta}_i\right]$.

Hence $Var\left[\hat{\theta}\right] = \dfrac{\sum_i^k w_i^2.Var\left[\hat{\theta}_i\right]}{\left(\sum_i^k w_i\right)^2}$. [Book work]

[3 marks]

The table below summarizes the outcome of three randomised trials of a *new antipsychotic drug* compared to a *standard antipsychotic drug* for people with schizophrenia. The treatment effect for each study ($\hat{\theta}_i$, i = 1,2,3) is the difference in mean symptom scores for the two treatments, with lower values representing a benefit of treatment. $\widehat{Var}\left[\hat{\theta}_i\right]$ is the sample variance estimate of the i^{th} study.

Study	*Mean difference in symptoms*,* $\hat{\theta}_i$	$\widehat{Var}\left[\hat{\theta}_i\right]$
Holmes (2000)	2.4	0.5
Bourne (2002)	-2.8	2.5
Bond (2006)	3.0	2.0

* *new drug – standard drug.*

(ii) The minimum variance estimator, $\hat{\theta}_{MV}$, is obtained when $w_i \propto 1/Var\left[\hat{\theta}_i\right]$. For the data in the table above, compute the minimum variance estimate of the overall treatment effect.

Solution

Study	*Mean difference in symptoms*,* $\hat{\theta}_i$	$Var\left[\hat{\theta}_i\right]$	$w_i = 1/Var\left[\hat{\theta}_i\right]$
Holmes (2000)	2.4	0.5	2.0
Bourne (2002)	-2.8	2.5	0.4
Bond (2006)	3.0	2.0	0.5

$$\hat{\theta}_{MV} = \frac{\sum_i^k w_i \hat{\theta}_i}{\sum_i^k w_i} = \frac{(2.0\times 2.4)+(0.4\times -2.8)+(0.5\times 3.0)}{2.0+0.4+0.5} = \frac{5.18}{2.9} = 1.786$$

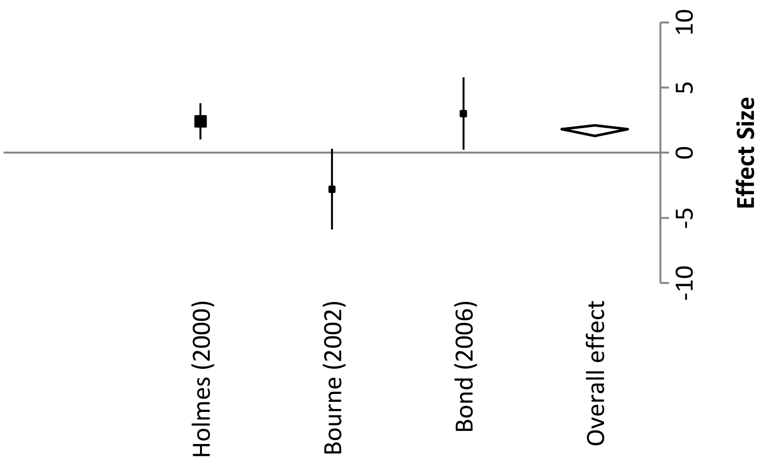
[Calculation time 5 mins]

[4 marks]

(iii) Illustrate the three studies and the overall treatment effect using a sketch of a forest plot.

Solution

The forest plot should look as follows:



[Time 3 minutes]

[Total 10 marks]

B1.

For a parallel group randomised controlled trial comparing a control treatment (C) with a new treatment (T) suppose Y is a continuous normally distributed outcome variable and X is the value of the same variable recorded prior to randomisation. Suppose that τ is the treatment effect such that:

$$Y = \mu_y + \varepsilon_y \quad \text{and} \quad X = \mu_x + \varepsilon_x \quad \text{for} \quad \text{treatment C}$$

$$Y = \mu_y + \tau + \varepsilon_y \quad \text{and} \quad X = \mu_x + \varepsilon_x \quad \text{for} \quad \text{treatment T}$$

with $E\left[\varepsilon_x\right] = E\left[\varepsilon_y\right] = 0$, $Var\left[\varepsilon_y\right] = \sigma_y^2$, $Var\left[\varepsilon_x\right] = \sigma_x^2$, and $Cov\left[\varepsilon_x, \varepsilon_y\right] = \sigma_{xy}$.

Define $D = Y - X$ and suppose that $\bar{x}_T$, $\bar{x}_C$, $\bar{y}_T$, $\bar{y}_C$, $\bar{d}_C$ and $\bar{d}_T$ are the sample means of X , Y and D for each treatment respectively. Define $\hat{\tau}(\theta) = (\bar{y}_T - \theta\bar{x}_T) - (\bar{y}_C - \theta\bar{x}_C)$.

(i) Show that $E\left[\hat{\tau}(\theta)\right] = \tau$.

Solution

Define $\hat{\tau}(\theta) = (\bar{y}_T - \theta\bar{x}_T) - (\bar{y}_C - \theta\bar{x}_C)$.

Now $E\left[\hat{\tau}(\theta)\right] = E\left[\bar{y}_T - \bar{y}_C\right] - \theta E\left[\bar{x}_T - \bar{x}_C\right]$

Since $E\left[\bar{y}_T - \bar{y}_C\right] = \tau + \beta E\left[\bar{x}_T - \bar{x}_C\right]$, it follows that

$$E\left[\hat{\tau}(\theta)\right] = \tau + (\beta - \vartheta) E\left[\bar{x}_T - \bar{x}_C\right]$$

Randomisation means that $E\left[\bar{x}_T\right] = E\left[\bar{x}_C\right]$.

Therefore $E\left[\hat{\tau}(\theta)\right] = \tau$.

[4 marks]

(ii) Show that $Var\left[\hat{\tau}(\theta)\right] = \lambda^2\left(\sigma_y^2 + \theta^2\sigma_x^2 - 2\theta\sigma_{xy}\right)$ where $\lambda = \sqrt{\frac{1}{n_T} + \frac{1}{n_C}}$, n_T is the numbers of patients allocated to the new treatment and n_C is the number allocated to the control treatment.

Solution

Consider $\hat{\tau}(\theta) = (\bar{y}_T - \theta \bar{x}_T) - (\bar{y}_{\boldsymbol{C}} - \theta \bar{x}_C)$

$$Var\left[\hat{\tau}(\theta)\right] = Var\left[(\bar{y}_T - \bar{y}_C) - \theta(\bar{x}_T - \bar{x}_C)\right]$$

$$= Var[\bar{y}_T - \bar{y}_C] + Var\left[\theta\left((\bar{x}_T - \bar{x}_C)\right)\right] - 2Cov\left[\bar{y}_T - \bar{y}_C, \theta(\bar{x}_T - \bar{x}_C)\right]$$

$$Var[\bar{y}_T - \bar{y}_C] + \theta^2 Var[\bar{x}_T - \bar{x}_C] - 2.\theta.Cov[\bar{y}_T - \bar{y}_C, \bar{x}_T - \bar{x}_C] \ [1]$$

Considering the first term

$$Var[\bar{y}_T - \bar{y}_C] = Var[\bar{y}_T] + Var[\bar{y}_C] - 2Cov[\bar{y}_T, \bar{y}_C]$$

Since treatment groups are independent $Cov\left[\bar{Y}_T, \bar{Y}_C\right] = 0$.

Therefore $Var[\bar{y}_T - \bar{y}_C] = Var[\bar{y}_T] + Var[\bar{y}_C]$.

Since observations are independent $Var[\bar{y}_T] = \dfrac{\sum_{i \in T} Var[y_T]}{n_T^2} = \dfrac{\sum_{i \in T} \sigma_Y^2}{n_T^2} = \dfrac{\sigma_Y^2}{n_T}$.

Similarly $Var[\bar{y}_C] = \dfrac{\sigma_Y^2}{n_C}$.

Therefore $Var[\bar{y}_T - \bar{y}_C] = \lambda^2 \sigma_Y^2$ where $\lambda = \sqrt{\dfrac{1}{n_T} + \dfrac{1}{n_C}}$.

Similarly $Var[\bar{x}_T - \bar{x}_C] = \lambda^2 \sigma_X^2$ and $Cov[\bar{y}_T - \bar{y}_C, \bar{x}_T - \bar{x}_C] = \lambda^2 \sigma_{XY}$.

Substitution into [1] gives $Var\left[\hat{\tau}(\theta)\right] = \lambda^2 \left(\sigma_y^2 + \theta^2 \sigma_X^2 - 2\theta \sigma_{XY}\right)$

[7 marks]

(iii) Show that $Var\left[\hat{\tau}(\theta)\right]$ has a minimum when $\theta = \beta$, where β is the regression coefficient of Y on X.

Solution

Differentiation with respect to θ gives

$$\frac{\partial}{\partial\theta}Var\left[\hat{\tau}(\theta)\right]=\lambda^2\left(2\theta\sigma_x^2-2\sigma_{xy}\right).$$

This equals zero when $\theta = \sigma_{xy}/\sigma_x^2$.

The second derivative $\frac{\partial^2}{\partial\theta^2}Var\left[\hat{\tau}(\theta)\right]=2\lambda^2\sigma_x^2$.

As this is positive, it follows that $Var\left[\hat{\tau}(\theta)\right]$ has a minimum when $\theta = \sigma_{xy}/\sigma_x^2$.

[4 marks]

(iv) Three statistical analyses might be used to estimate and test the treatment effect:

a) an unadjusted analysis using just the outcome variable *Y*

b) an analysis based on the change score *Y*-*X* or

c) a linear model of the outcome variable *Y* with treatment group and *X* as covariates.

What are the implications of the results in (i) and (iii) for the choice between the three analyses in terms of bias and precision of the treatment effect?

Solution

Values of θ equal to 0, 1 and β correspond to the treatment effect in an unadjusted, change and linear adjusted model analyses. All three estimates are unbiased, but an estimate of the treatment effect based on a linear model smaller variance compared to an unadjusted analysis or a change analysis. Reducing the variance of the treatment effect estimate increases the power of the analysis. As a consequence if a baseline variable is thought to be correlated with outcome, an analysis adjusting for baseline is recommended, and where the baseline value of the outcome is recorded a linear model analysis is superior to an analysis based on change.

[3 marks]

(v) Why is it important for randomised controlled trials to have a statistical analysis plan?

Solution

A statistical analysis should be prepared prior to beginning the analysis. There can be many different ways in which the outcome from a randomised controlled trial can be analysed. For example 3 possible analysis were considered in (iv). These analyses may give results that are more or less supportive of the investigators' opinion. Unless the analysis is pre-specified the investigators may present that which is closest to their opinion, instead of the best analysis.

[2 marks]

[Book work]

[Total 20 marks]

B2.

For an *AB*/*BA* crossover trial a standard model for a continuous outcome y_{ij} for the i^{th} patient in the j^{th} period is:

$y_{i1} = \mu + \tau + \xi_i + \varepsilon_{i1}$	for a patient in sequence *AB* in period 1
$y_{i2} = \mu + \phi + \xi_i + \varepsilon_{i2}$	for a patient in sequence *AB* in period 2
$y_{i1} = \mu + \xi_i + \varepsilon_{i1}$	for a patient in sequence *BA* in period 1
$y_{i2} = \mu + \tau + \phi + \xi_i + \varepsilon_{i2}$	for a patient in sequence *BA* in period 2

where μ is the mean for the sequence *BA* in period 1, τ is the treatment effect of A compared to B, and ϕ is the period effect, with ξ_i and ε_{ij} being two independent random variables with $\xi_i \sim N\left[0, \sigma_B^2\right]$ and $\varepsilon_{ij} \sim N\left[0, \sigma_\varepsilon^2\right]$. Defining $d_i = y_{i2} - y_{i1}$ let $\bar{d}_{AB}$, $\bar{d}_{BA}$, μ_{AB}^d, μ_{BA}^d be the sample and population means of d_i for sequences *AB* and *BA* respectively.

(i) Show that a test of the null hypothesis $H_0 : \mu_{AB}^d = \mu_{BA}^d$ is the same as a test of no treatment effect, $H_0 : \tau = 0$

Solution

$\mu_{AB}^d = E\left[y_{i2} - y_{i1}\right] = E\left[\left(\mu + \phi + \varepsilon_{i2}\right) - \left(\mu + \tau + \varepsilon_{i1}\right)\right] = \phi - \tau$ since $E\left[\varepsilon_{i2}\right] = E\left[\varepsilon_{i1}\right] = 0$

Similarly

$\mu_{BA}^d = E\left[y_{i2} - y_{i1}\right] = E\left[\left(\mu + \phi + \tau + \varepsilon_{i2}\right) - \left(\mu + \varepsilon_{i1}\right)\right] = \phi + \tau$

Hence $\mu_{BA}^d - \mu_{AB}^d = 2\tau$

$H_0 : \mu_{AB}^d = \mu_{BA}^d$ iff $H_0 : \tau = 0$.

[Book work]

[3 marks]

(ii) Two anti-cholesterol lowering drugs were compared in a randomised controlled crossover trial. Ten patients were randomly allocated to sequence drug A then drug B and eight patients were randomly allocated to sequence drug B then drug A. The table below summarizes the sample mean and standard deviation for each sequence and period interval of the treatment effect. Test the hypothesis $H_0 : \tau = 0$ vs. $H_1 : \tau \neq 0$ using a 5% significance level, and compute the 95% confidence interval.

	Period 1		*Period 2*		*Period 2 – Period 1*		
Sequence	*Mean*	*s.d.*	*mean*	*s.d.*	*Mean*	*s.d*	*N*
AB	6.42	0.81	6.01	0.72	-0.41	0.52	10
BA	6.23	0.63	6.11	0.71	-0.12	0.43	8

Solution

The hypothesis H_0:τ=0 vs. H_1:τ≠0 can be tested using a two-sample t-test of the means of the differences $H_0 : \mu_{AB}^d = \mu_{BA}^d$. The test statistic T_d is defined as

$$T_d = \frac{\bar{d}_{BA} - \bar{d}_{AB}}{\hat{S}E\left[\bar{d}_{BA} - \bar{d}_{AB}\right]} \text{ where } \hat{S}E\left[\bar{d}_{BA} - \bar{d}_{AB}\right] = s_d\sqrt{\frac{1}{n_{AB}} + \frac{1}{n_{BA}}} \text{ and } s_d = \sqrt{\frac{(n_{AB}-1)s_{d_{AB}}^2 + (n_{BA}-1)s_{d_{BA}}^2}{n_{AB} + n_{BA} - 2}}.$$

Substitution gives $s_d = \sqrt{\dfrac{9 \times 0.52^2 + 7 \times 0.43^2}{16}} = 0.483$ Hence

$$\hat{S}E\left[\bar{d}_{BA} - \bar{d}_{AB}\right] = 0.483\sqrt{\frac{1}{10} + \frac{1}{8}} = 0.229$$

Hence $T_d = \dfrac{-.41-(-0.12)}{0.229} = -1.266$ Under the null hypothesis this will have a t distribution with 16 d.f.

From tables p>0.05. Hence the null hypothesis is not rejected.

A (1-α)-size confidence interval for the treatment effect τ is defined by

$$\frac{1}{2}\left(\bar{d}_{BA} - \bar{d}_{AB}\right) \pm \frac{1}{2} t_{\alpha/2}\left(n_1+n_2-2\right)\hat{S}E\left[\bar{d}_{BA} - \bar{d}_{AB}\right].$$

From tables $t_{\alpha/2}\left(n_1+n_2-2\right) = t_{0.05}(16) = 2.12$

Therefore the 95% confidence interval

$\frac{1}{2}\left(\bar{d}_{BA} - \bar{d}_{AB}\right) \pm \frac{1}{2} t_{\alpha/2}\left(n_1+n_2-2\right)\hat{S}E\left[\bar{d}_{BA} - \bar{d}_{AB}\right]$ which is $\frac{1}{2}(-0.29) \pm \frac{1}{2} 2.12 \times .229$ giving

[-0.388 , 0.098] or [-0.39 , 0.10]

Based on the hypothesis test there is no evidence that there is any difference between the two treatment. The 95% confidence interval is [-0.388 , 0.098]

[5 minutes calculation time]

[7 marks]

(iii) Define $c_i = y_{i1} - y_{i2}$ for sequence AB and $c_i = y_{i2} - y_{i1}$ for sequence BA. Let μ_{AB}^c μ_{BA}^c, $\bar{c}_{AB}$ and $\bar{c}_{BA}$ be the population and sample means of these for sequences AB and BA respectively. Show that a test of the null hypothesis $H_0 : \mu_{AB}^c = \mu_{BA}^c$ is the same as a test of the period effect, $H_0 : \phi = 0$.

Solution

$$\mu_{AB}^c = E\left[y_{i1} - y_{i2}\right] = E\left[\left(\mu + \tau + \varepsilon_{i2}\right) - \left(\mu + \phi + \varepsilon_{i1}\right)\right] = \tau - \phi$$

$$\mu_{BA}^c = E\left[y_{i2} - y_{i1}\right] = E\left[\left(\mu + \phi + \tau + \varepsilon_{i2}\right) - \left(\mu + \varepsilon_{i1}\right)\right] = \phi + \tau$$

Therefore $\mu_{BA}^C - \mu_{AB}^C = -2\phi$.

Hence the test H_0: $\mu_{AB}^c = \mu_{BA}^c$ *is equivalent to a test of the period effect* $H_0 : \phi = 0$.

[Book work]

[3 marks]

(iv) From the data in the table above, test the null hypothesis $H_0 : \phi = 0$ vs. $H_1 : \phi \neq 0$ using a 5% significance level.

Solution

The hypothesis H_0: ϕ=0 vs. H_1: $\phi \neq 0$ can be tested using a two-sample t-test of the means of the differences $H_0 : \mu_{AB}^c = \mu_{BA}^c$. An appropriate test statistic T_C is defined as

$$T_C = \frac{\bar{c}_{BA} - \bar{c}_{AB}}{\hat{S}E\left[\bar{c}_{BA} - \bar{c}_{AB}\right]} \text{ where } \hat{S}E\left[\bar{c}_{BA} - \bar{c}_{AB}\right] = s_C\sqrt{\frac{1}{n_{AB}} + \frac{1}{n_{BA}}} \text{ and } s_C = \sqrt{\frac{\left(n_{AB} - 1\right)s_{C_{AB}}^2 + \left(n_{BA} - 1\right)s_{C_{BA}}^2}{n_{AB} + n_{BA} - 2}}$$

Where s_C is the pooled standard deviation and $s_{C_{AB}}, s_{C_{BA}}$ are the standard deviations for each sequence.

Now $\mu_{AB}^c = -\mu_{AB}^d$=0.41 and $\mu_{BA}^c = \mu_{BA}^d$=-0.12

$s_{C_{AB}} = s_{D_{AB}} = s_{C_{BA}} = s_{D_{AB}}$. Therefore $\hat{S}E\left[\bar{c}_{BA} - \bar{c}_{AB}\right] = \hat{S}E\left[\bar{d}_{BA} - \bar{d}_{AB}\right] = 0.229$

Therefore $T_C = \dfrac{0.41 + 0.12}{0.229} = 2.31$

From tables $t_{\alpha/2}\left(n_1 + n_2 - 2\right) = t_{0.05}\left(16\right) = 2.12$ Therefore one can reject the null hypothesis that H_0: ϕ=0

[3 minutes calculation time]

[5 marks]

(v) Briefly comment on the result of the trial

Solution

From the test of the hypothesis $H_0{:}\tau = 0$ there is no evidence of a treatment effect. In contrast there is evidence of a period effect. From inspection of the data in the table one can see that cholesterol levels reduce for both sequences.

[2 marks]

[Total marks 20]

B3.

Consider a randomised controlled trial. Suppose the patient population can be divided into three sub-groups as follows:

Compliers:	patients who will comply with the allocated treatment,
Always control treatment:	patients who will receive control treatment regardless of allocation
Always new treatment:	patients who will receive the new treatment regardless of allocation.

This division assumes that there are no defiers, that is patients who will always receive the opposite of the treatment to which they are randomised. Assume that the proportion and characteristics of *Compliers*, *Always control treatment*, *Always new treatment* is the same in both arms and that randomization can only affect the outcome through the receipt of treatment.

(i) Show that an *Intention-To-Treat* estimate of the treatment effect is biased towards the null hypothesis of no treatment effect.

Solution

Table of expected means under assumptions of model

	Type	**Control Group**	**New Treatment Group**	**Proportion In Latent Class**
As Randomized	A	μ	$\mu+\tau$	$\theta_A = 1-\theta_B-\theta_C$
Always Control	B	$\mu+\gamma_B$	$\mu+\gamma_B$	θ_B
Always New Treatment	C	$\mu+\gamma_C+\tau$	$\mu+\gamma_C+\tau$	θ_C

τ is the causal effect of treatment

For Intention-to-Treat Estimate

$$\tau_{ITT} = \left[\theta_A(\mu+\tau)+\theta_B(\mu+\gamma_B)+\theta_C(\mu+\gamma_C+\tau)\right]-\left[\theta_A\mu+\theta_B(\mu+\gamma_B)+\theta_C(\mu+\gamma_C+\tau)\right]$$
$$= \theta_A\tau$$

as second and third terms in each bracket cancel.

Hence $\left|\hat{\tau}_{ITT}\right| \le \tau$ which means $\hat{\tau}_{ITT}$ is biased towards zero if $\boldsymbol{\theta_A} < 1$ i.e. if some patients do not comply with treatment. [Book work]

[4 marks]

(ii) Show that an *As-Treated* estimate of the treatment effect may be biased either towards or away from the null hypothesis of no treatment effect.

Solution

For the As-Treated Estimate

$$\tau_{AT} = \left[\frac{\theta_A(\mu+\tau)+2\theta_C(\mu+\gamma_C+\tau)}{\theta_A+2\theta_C}\right]-\left[\frac{\theta_A\mu+2\theta_B(\mu+\gamma_B)}{\theta_A+2\theta_B}\right]$$

$$=\left[\frac{(\theta_A+2\theta_C)\mu+2\theta_C\gamma_C+(\theta_A+2\theta_C)\tau}{\theta_A+2\theta_C}\right]-\left[\frac{(\theta_A+2\theta_B)\mu+2\theta_B\gamma_B}{\theta_A+2\theta_B}\right]$$

$$=\tau+\mu+\left[\frac{2\theta_C\gamma_C}{\theta_A+2\theta_C}\right]-\mu-\left[\frac{2\theta_B\gamma_B}{\theta_A+2\theta_B}\right]$$

$$=\tau+\left[\frac{2\theta_C\gamma_C}{1-\theta_B-\theta_C+2\theta_C}\right]-\left[\frac{2\theta_B\gamma_B}{1-\theta_B-\theta_C+2\theta_B}\right]$$

$$=\tau+\left[\frac{\theta_C\gamma_C}{1-\theta_B+\theta_C}\right]-\left[\frac{\theta_B\gamma_B}{1-\theta_C+\theta_B}\right]$$

The values of γ_C and γ_B can be positive or negative so that the second and third terms can be either positive or negative. Hence the bias can be away from or towards the null hypothesis.

[Book work]
[5 marks]

The table below summarizes the outcome of patients from randomised controlled trial comparing two treatments according to randomized group and treatment received. Some patients allocated to the *New* treatment received the *Control* treatment and some patients allocated to *Control* treatment received the *New* treatment.

	Randomised Group			
	New Treatment		*Control*	
Recovered after 12 weeks	*Received New*	*Received Control*	*Received New*	*Received Control*
Yes	360	72	48	360
No	120	48	12	180
Total	480	120	60	540

(iii) For the outcome measure *Recovered after 12 weeks*, calculate the point estimate of the difference in proportions of the *New* treatment compared to the *Control* treatment assuming

(a) an *Intention-To-Treat* analysis, (b) an *As-Treated analysis*.

Solution

(a) Intention-To-Treat = 432/600 – 408/600= 0.72- 0.68 = 0.04

(b) As-treated = 408/540-432/660 = 0.75555 - 0.654545= 0.1010101

[2 minutes calculation time] [3 marks]

(iv) Explain why an *Intention-To-Treat* estimate is preferable to an *As-Treated* estimate in a superiority trial.

Solution

As we have seen in (i) an intention-to-treat analyses will bias an estimate of the treatment effect towards the estimate of no effect. This means that any effect will be biased towards the null hypothesis of a superiority trial. Hence if we reject the null hypothesis using an intention-to-treat analysis, we can be more confident that the true treatment effect is at least as large as that observed. In contrast a per-protocol analyses may bias the estimate of the treatment effect either away or towards the null hypothesis of no effect as seen in (ii). [3 marks]

(v) Calculate the point estimate of the C*omplier Average Causal Effect* of *New* treatment compared to *Control*.

Solution

From above $\tau_{ITT} = \theta_A \tau$ where θ_A is the proportion of subjects that accept randomized treatment and τ is the causal effect of treatment or the C*omplier Average Causal Effect* . θ_{A} =1-θ_{B} -θ_{C} where θ_{B} is the proportion of patients that will always receive the control and θ_C the proportion who will always receive the new active treatment.

θ_{B} can be estimated from the new treatment arm and θ_{C} from the control arm.

θ_{B} = 120/600. θ_{C} = 60/600. Hence θ_{A} =1 - (60/600) – (120/600) = 0.7

Hence the C*omplier Average Causal Effect* of *New* treatment compared to *Control* treatment,

τ=0.04/0.7=0.057 [3 minutes calculation time]

[5 marks]

[Total 20 marks]

END OF EXAMINATION PAPER

Math31042 Fractal Geometry: Summer 2017 Exam Solutions

March 10, 2017

A1) a) A δ-cover of E is a collection $\{U_j\}_{j\in J}$ of sets $U_j \subset X$ satisfying that $E \subset \bigcup_{j\in J} U_j$ and that each U_j has $|U_j| \leq \delta$.

$$\mathcal{H}^s_\delta(E) := \inf\{\sum_{j\in J} |U_j|^s : \{U_j\}_{j\in J} \text{ is a } \delta\text{-cover of } E\}.$$

[3 marks, bookwork]

b)

$$\mathcal{H}^s(E) := \lim_{\delta\to 0} \mathcal{H}^s_\delta(E).$$

For any $\delta' < \delta$ we have that any δ'-cover of E is also a δ-cover of E. Thus

$$\mathcal{H}^s_{\delta'}(E) \geq \mathcal{H}^s_\delta(E).$$

Since $\mathcal{H}^s_\delta(E)$ is monotone in δ we conclude that the limit as $\delta \to 0$ exists. [4 marks, bookwork]

c) For $0 < s < t$, given a δ-cover $\{U_j\}_{j\in J}$ of E we see that

$$|U_j|^{s-t} \geq \delta^{s-t}.$$

Hence

$$\begin{aligned} \sum_{j\in J} |U_j|^s &= \sum_{j\in J} |U_j|^t |U_j|^{s-t} \\ &\geq \delta^{s-t} \sum_{j\in J} |U_j|^t. \end{aligned}$$

Taking the infimum over all δ-covers gives that

$$\mathcal{H}^s_\delta(E) \geq \delta^{s-t} \mathcal{H}^t_\delta(E).$$

Letting $\delta \to 0$ gives that if $s < t$ and $\mathcal{H}^t(E) > 0$ then $\mathcal{H}^s(E) = \infty$.

Conversely, if $s < t$ and $\mathcal{H}^s(E) < \infty$ then $\mathcal{H}^t(E) = 0$.

Thus we can define

$$\dim_H(E) = \inf\{s : \mathcal{H}^s(E) = 0\} = \sup\{t : \mathcal{H}^t(E) = \infty\}.$$

[bookwork, 6 marks]

A2) If $A \subset B$ then any δ-cover of B is a δ-cover of A. Thus $\mathcal{H}^s_\delta(A) \leq \mathcal{H}^s_\delta(B)$.

Taking limits as $\delta \to 0$ gives $\mathcal{H}^s(A) \leq \mathcal{H}^s(B)$.

[Alternatively, this last line could be concluded directly from the properties of outer measure, using the fact (proved in the course) that $\mathcal{H}^s$ is an outer measure].

Thus $\mathcal{H}^s(B) = 0 \implies \mathcal{H}^s(A) = 0$, and so

$$\inf\{s : \mathcal{H}^s(A) = 0\} \leq \inf\{s : \mathcal{H}^s(B) = 0\}$$

giving

$$\dim_H(A) \leq \dim_H(B).$$

[4 marks, bookwork]

A3 a) An outer measure is a mass-distribution if it has bounded support and if

$$0 < \mu(X) < \infty.$$

[Alternative statement for last part, $0 < \mu(supp(\mu)) < \infty$.]

[3 marks, bookwork]

b) Let μ be a mass distribution on E, and suppose that for some s there are numbers $c > 0$ and $\epsilon > 0$ such that

$$\mu(u) \leq c|U|^s$$

for all sets $U \subset E$ with $|U| < \epsilon$. Then

$$\mathcal{H}^s(E) \geq \frac{\mu(E)}{c}$$

and hence $s \leq \dim_H(E)$.

[4 marks, bookwork]

c) Consider two dimensional Lebesgue measure λ_2 on $[0,1]^2$. If $|U| < \delta$ then

$$\lambda_2(U) \leq \pi \left(\frac{\delta}{2}\right)^2 = \frac{\pi}{4}\delta^2.$$

Thus $2 \leq \dim_H([0,1]^2)$ by the mass distribution principle.

[5 marks, example sheet]

A4) $f : (X, d) \to (X', d')$ is Lipschitz if there exists a constant $K > 0$ such that

$$d'(f(x), f(y)) \leq K d(x, y)$$

for all $x, y \in X$.

If $f:(X,d)\to(X',d')$ is Lipschitz then for any set $U\subset X$ with $|U|\leq\delta$ we have $|f(U)|\leq K\delta$.

Furthermore, if $\{U_j\}_{j\in J}$ is a δ-cover of E then $\{f(U_j)\}_{j\in J}$ is a $K\delta$-cover of $f(E)$. Thus if $\mathcal{H}^s_\delta(E)<C$ then there exists a δ cover $\{U_j\}_{j\in J}$ of E with

$$\sum_{j\in J}|U_j|^s\leq C$$

and hence

$$\mathcal{H}^s_{K\delta}(f(E))\leq\sum_{j\in J}(K|U_j|)^s\leq K^sC.$$

Hence $\mathcal{H}^s_{K\delta}(f(E))\leq K^s\mathcal{H}^s(E)$. Taking limits as $\delta\to 0$ gives

$$\mathcal{H}^s(f(E))\leq K^s\mathcal{H}^s(E)$$

and hence

$$\dim_H(f(E))\leq\dim_H(E).$$

[6 marks, bookwork]

A5) Let $\pi_1:\mathbb{R}^2\to\mathbb{R}$ be given by

$$\pi_1(x,y)=x.$$

Then

$$\begin{aligned}
d((x_1,y_1),(x_2,y_2)) &= \sqrt{(x_2-x_1)^2+(y_2-y_1)^2}\\
&\geq \sqrt{(x_2-x_1)^2}\\
&= |x_2-x_1|\\
&= d(\pi_1(x_1,y_1),\pi_1(x_2,y_2)).
\end{aligned}$$

Thus π_1 is Lipschitz (with Lipschitz constant 1).

Now given $g:[0,1]\to\mathbb{R}$ we see that

$$\pi_1\{(x,g(x)):0\leq x\leq 1\}=[0,1].$$

Thus, since π_1 is Lipschitz,

$$\dim_H\{(x,g(x)):x\in[0,1]\}\geq\dim_H(\pi_1\{(x,g(x)):x\in[0,1]\})=\dim_h([0,1])=1.$$

[unseen, although we have used the fact that π_1 is Lipschitz in example sheets, 5 marks]

B6) a) Let $\delta > 0$ and choose $k \in \mathbb{N}$ such that

$$10^{-k} \leq \delta < 10^{-(k-1)}.$$

Then

$$\mathcal{U} := \{\phi_{a_1} \circ \cdots \circ \phi_{a_k}([0,1]) : \text{ each } a_i \in \{1,3,5\}\}$$

is a δ-cover of E. Thus

$$\mathcal{N}_\delta(E) \leq |\mathcal{U}| = 3^k.$$

Hence

$$\begin{aligned} \frac{\log(\mathcal{N}_\delta(E))}{-\log \delta} &\leq \frac{\log(3^k)}{-log(10^{-(k-1)})} \\ &= \frac{klog3}{(k-1)\log 10} \\ &\to \frac{\log 3}{\log 10} \end{aligned}$$

as $k \to \infty, \delta \to 0$. Thus $\overline{dim}_B(E) leq \frac{\log 3}{\log 10}$.
[similar to example sheet, 5 marks]

b) Put a mass distribution μ on E by letting

$$\mu(\phi_{a_1} \circ \cdots \circ \phi_{a_k}([0,1]) = 3^{-k}$$

and extending this to an outer measure on all sets A by letting

$$\mu(A) := \inf\{\sum_{j=1}^{\infty} \mu(\phi_{a_1} \circ \cdots \circ \phi_{a_{k_j}}([0,1]) : A \subset \bigcup_{j=1}^{\infty} \phi_{a_1} \circ \cdots \circ \phi_{a_{k_j}}([0,1])\}.$$

Now suppose that $|U| < \delta$. Choose $k \in \mathbb{N}$ such that

$$10^{-k} \leq \delta < 10^{-(k-1)}.$$

Then U intersects at most one basic interval $\phi_{a_1} \circ \cdots \phi_{a_{k-1}}([0,1])$ of depth $(k-1)$, and so

$$\mu(U) \leq 3^{-(k-1)} < (10|U|)^{\frac{\log 3}{\log 10}}.$$

Hence by the mass distribution principle,

$$\mathcal{H}^s(E) \geq \frac{\log 3}{\log 10}.$$

Then

$$\frac{\log 3}{\log 10} \leq \dim_H(E) \leq \underline{\dim}_B(E) \leq \overline{dim}_B(E) \leq \frac{\log 3}{\log 10},$$

hence $\dim_B(E)$ exists and equals $\frac{\log 3}{\log 10}$. [7 marks, similar to example sheet (only the self similar set changed)]

c) E is the set of points in $[0,1]$ with only digits $1,3$ and 5 in their decimal expansions, i.e.

$$E = \{0.x_1x_2x_3\cdots : \text{ each} x_i \in \{1,3,5\}\}$$

[2 marks, similar to example sheet]

d) If $A \subset \bigcup_{k=1}^{\infty} A_k$ and each A_k has

$$\dim_H(A_k) \leq d$$

then for any $s > d$, for all $kin\mathbb{N}$, $\mathcal{H}^s(A_k) = 0$.

Thus, since $\mathcal{H}^s$ is an outer measure,

$$\begin{aligned}\mathcal{H}^s(A) &\leq \mathcal{H}^s(\bigcup_{i=1}^{\infty} A_k) \\ &\leq \sum_{k=1}^{\infty} \mathcal{H}^s(A_k) = 0.\end{aligned}$$

Then $\dim_H(A) \leq s$.

Since $s > d$ was arbitrary, we conclude $\dim_H(A) \leq d$.

[5 marks, bookwork]

e) The map $\pi_{a_1\cdots a_k} : \mathbb{R} \to \mathbb{R}$ given by

$$\pi_{a_1\cdots a_k}(0.x_1x_2x_3\cdots) = (0.a_1\cdots a_kx_1x_2x_3\cdots)$$

is Lipschitz and maps E onto

$$A_{a_1\cdots a_k} = \{0.a_1\cdots a_kx_1x_2\cdots : x_i \in \{1,3,5\}\}.$$

Thus $\dim_H(A_{a_1\cdots a_k}) \leq \dim_H(E)$.

Now we can write

$$A = \bigcup_{k=1}^{\infty} \bigcup_{a_1,\cdots,a_k \in \{0,1,\cdots,9\}^k} A_{a_1\cdots a_k}.$$

This is a countable union, and so by part d),

$$\dim_H(A) \leq \max\{\dim_H(A_{a_1\cdots a_k})\} = \dim_H(E) = \frac{\log 3}{\log 10}$$

[6 marks, unseen]

B7 a) $\mathcal{S}(X)$ is the space of non-empty compact subsets of X.

Define

$$d_H(A,B) = \inf\{\delta : \forall a \in A \exists b \in B \text{ with} d(a,b) \leq \delta, \forall b \in B \exists a \in A \text{ with } d(a,b) \leq \delta\}$$

If $\Phi = \{\phi_1 \cdots \phi_k$ is an iterated fuction system where each ϕ_i is a contraction then there exists a unique non-empty compact set E satisfying

$$E = \bigcup_{i=1}^{k} \phi_i(E).$$

[6 marks, bookwork]

b) $\dim_{sim}(\Phi)$ is the unique value of s satisfying

$$sum_{i=1}^{k} c_i^s = 1.$$

[2 marks, bookwork]

c) Let $Y \subset X$ satisfy that $\phi_i(Y) \subset Y$ for each $i \in \{1, \cdots k\}$.

Let $c = \max\{c_i\} < 1$.

Given $\delta > 0$ let $n \in \mathbb{N}$ be such that $c^n|Y| \leq \delta$.

Then $\{\phi_{a_1} \circ \cdots \phi_{a_n}([0,1]) : a_i \in \{1, \cdots, k\}\}$ is a δ-cover of E.

Then

$$\begin{aligned}
\mathcal{H}^s_\delta(E) &\leq \sum_{a_1,\cdots,a_n \in \{1,\cdots,k\}^n\}} |\phi_{a_1} \circ \cdots \phi_{a_n}(Y)|^s \\
&= \sum_{a_1,cdots,a_n \in \{1,\cdots,k\}^n} c^s_{a_1} \cdots c^s_{a_n} |Y|^s \\
&= (c_1^s + \cdots c_k^s)^n |Y|^s
\end{aligned}$$

But if $s = \dim_{sim}(Y)$ then $c_1^s + \cdots c_k^s = 1$.

So $\mathcal{H}^s_\delta(E) \leq |Y|^s$. Letting $\delta \to 0$ gives $\mathcal{H}^s(E) \leq |Y|^s$ and hence

$$\dim_H(E) \leq dim_{sim}(\Phi).$$

[bookwork, 8 marks]

d) $\Phi = (phi_1, \cdots, \phi_k)$ satisfies the open set condition if there exists a non-empty, bounded open subset V of $(\mathbb{R}^n, d)$ with

$$V \supset \bigcup_{i=1}^{k} \phi_i(V)$$

and with

$$\phi_i(V) \cap \phi_j(V) = \phi$$

for $i \neq j$.

[bookwork, 3 marks]

e) Let $c_1 \cdots c_k$ be the contraction ratios of $phi_1, \cdots, \phi_k$ respectively. Then the contraction ratio of the map $\phi_i \circ \phi_j = c_i c_j$.

Then the similarity dimension s_2 of Φ_2 is the solution to

$$
\begin{aligned}
1 = \sum_{j=1}^{k} \sum_{i=1}^{k} (c_i c_j)^{s_2} &= \sum_{j=1}^{k} \sum_{i=1}^{k} (c_i)^{s_2} (c_j)^{s_2} \\
&= \left(\sum_{i=1}^{k} c_i^{s_2} \right) \left(\sum_{j=1}^{k} c_j^{s_2} \right).
\end{aligned}
$$

But this equation is satisfied by $s_2 = s$, the similarity dimension of Φ_1. Thus

$$
\dim_{sim}(\Phi_1) = \dim_{sim}(\Phi_2).
$$

If E is the unique non-empty compact set with $E = \bigcup_{i=1}^{k} \phi_i(E)$, then iterating gives

$$
\begin{aligned}
E &= \bigcup_{i=1}^{k} \phi_i \left(\bigcup_{j=1}^{k} \phi_j(E) \right) \\
&= \bigcup_{i=1}^{k} \bigcup_{j=1}^{k} \phi_i \circ \phi_j(E) \\
&= \bigcup_{\phi \in \Phi_2} \phi(E).
\end{aligned}
$$

Thus E is also the (unique) non-empty compact set with $E = \bigcup_{\phi \in \Phi_2} \phi(E)$, so the attractor of Φ_1 is the same as the attractor of Φ_2.

[6 marks, unseen]

B8 a) Let

$$\mathcal{N}_\delta(E) := \min\{k : \exists \text{ a } \delta - cover \mathcal{U} = \{U_j\}_{j=1}^k \text{ of } E\}$$

Then

$$\overline{\dim}_B(E) := \limsup_{\delta \to 0} \frac{\log \mathcal{N}_\delta(E)}{-log\delta}.$$

[bookwork, 3 marks]

b) Let $\{U_i\}_{i=1}^k$ be a δ-cover of E.

Then $\{U_i \times U_j\}_{i\in\{1,\cdots,k\},j\in\{1,\cdots,k\}}$ is a cover of $E \times E$.

$$|U_i \times U_j| = \sqrt{|U_i|^2 + |U_j|^2} \leq \sqrt{2}\delta.$$

Thus

$$\mathcal{N}_{\sqrt{2}\delta}(E \times E) \leq k^2 = (\mathcal{N}_\delta(E))^2.$$

Therefore

$$\frac{\log(\mathcal{N}_{\sqrt{2}\delta}(E \times E)}{-\log \delta} \leq \frac{\log((\mathcal{N}_\delta(E))^2)}{-\log \delta}$$

Therefore

$$\frac{\log(\sqrt{2}\delta}{\log \delta} \frac{\log(\mathcal{N}_{\sqrt{2}\delta}(E \times E)}{-\log(\sqrt{2}\delta)} \leq \frac{2\log(\mathcal{N}_\delta(E))}{-\log \delta}$$

Taking $\limsup$ as $\delta \to 0$, and observing that $\frac{\log(\sqrt{2}\delta)}{\log \delta} \to 1$, gives

$$\overline{\dim}_B(E \times E) \leq 2\overline{\dim}_B(E)$$

[exercise sheet, 6 marks]

c) Given $\delta > 0$ let $k \in \mathbb{N}$ satisfy

$$\frac{1}{k} - frac1k+1 = \frac{1}{k(k+1)} \leq \delta < \frac{1}{k(k-1)} = \frac{1}{k-1} - \frac{1}{k}.$$

Then each member of the set $\{1, \frac{1}{2}, \cdots \frac{1}{k}\}$ is separated by at least δ. So

$$\mathcal{N}_{delta}(E) \geq \mathcal{N}_\delta(\{1, \cdots, \frac{1}{k}\}) = k.$$

On the other hand, since $\frac{1}{k+1} \leq k\delta$ we have

$$\mathcal{N}_\delta[0, \frac{1}{k-1}] \leq k,$$

and hence

$$\begin{aligned} \mathcal{N}_\delta(E) &\leq \mathcal{N}_\delta[0, frac1k+1] + \mathcal{N}_\delta\{1, \frac{1}{2}, \cdots, \frac{1}{k}\} \\ &= k + k = 2k. \end{aligned}$$

This gives

$$\frac{\log k}{-\log\delta} \le \frac{\log(\mathcal{N}_\delta(E))}{-\log\delta} \le \frac{\log 2k}{-\log\delta}.$$

Hence

$$\frac{log k}{log(k(k+1)} \le \frac{\log(\mathcal{N}_\delta(E))}{-\log\delta} \le \frac{log 2k}{\log(k(k-1)}.$$

Now

$$\frac{\log k}{log(k(k+1)} = \frac{1}{1_{\frac{\log(k+1)}{\log k}}} \to \frac{1}{2}.$$

Similarly

$$\frac{\log(2k)}{\log(k(k-1))} = \frac{1+\frac{\log 2}{\log k}}{1+\frac{\log(k-1)}{\log k}} \to \frac{1}{2}.$$

[previous two lines could be omitted]

Thus, letting $\delta \to 0$ ($k \to \infty$) gives

$$\lim_{\delta\to 0}\frac{\log(\mathcal{N}_\delta(E)}{-\log\delta} = \frac{1}{2}.$$

[bookwork, 8 marks]

d) $E = \mathcal{Q} \cap [0,1]$ is an example.

[combining facts from different bits of bookwork, 2 marks]

e) Let $\{U_j\}_{j\in J}$ be a δ-cover of G.

Let k be such that $k\delta < 1 \le (k+1)\delta$. Then

$$\{U_j \times [i\delta,(i+1)\delta]\}_{j\in J, 1\le i\le k}$$

is a $\sqrt{2}\delta$-cover of $G \times [0,1]$.

Hence $\mathcal{N}_{\sqrt{2}\delta}(G\times[0,1]) \le \mathcal{N}_\delta(G)$.

So

$$\frac{\log(\sqrt{2}\delta)}{\log delta}\frac{\log \mathcal{N}_{\sqrt{2}\delta}(G\times[0,1])}{-\log(\sqrt{2}\delta)} \le \frac{\log(\mathcal{N}_\delta(G)) + \log(\delta^{-1})}{-\log\delta}$$

Taking limsup gives

$$\overline{\dim}_B(G\times[0,1]) \le \overline{\dim}_B(G) + 1 = \frac{3}{2}.$$

[unseen, 6 marks]

Solutions of problems

About Questions

First five questions A1 – A5

These questions are essentially bookwork and they are easy.

Just to mention

A2 b) – easy but little bit unexpected

A4 b) – bookwork, but difficulties may arise if student go in a wrong direction. (calculating df straightforwardly)

B6, B7, B8

These questions more difficult

B1 – problem becomes easy and short if student will find suitable vector $\bar{a}_t = (L, t)$ (to check condition $|p| < L$)

B2 – a) needs some level of (degree of) technique of calculation

– b) not difficult but needs a geom. intuition.

B3 – essential part is bookwork, but not easy.

The last question easy, but never was discussed in class.

C9, C10

C9 – first part heavy bookwork second part – unseen, but not diff

C10 – heavy bookwork + very difficult and unseen Question.

	A	B	C
Bookwork –	8 + 6 + 10 + 8 + 8 = 42	4 + 6 + 10	8 + 10
Easy	6 + 4 + 4 + 6 + 4 = 24	2 + 2 + 2	0 + 5
Difficult	0 + 4 + 0 + 2 + 2 = 8	8 + 4 + 4	2 + 10

A1 [3+3+2+2]

a) n-dimensional Riemanian manifold (M, G) is locally Euclidean if for every point $p \in M$ there exists an open neighborhood (domain) D such that it is isometric to a domain in Euclidean plane. In other words in a vicinity of every point p there exist local coordinates $(u^1, \dots, u^n)$ such that metric G has appearence

$$G = (du^1)^2 + \dots + (du^n)^2$$

b) For cylinder $\begin{cases} x = R\cos\varphi \\ y = R\sin\varphi \\ z = h \end{cases}$ $G = \mathcal{C}^* (dx^2 + dy^2 + dz^2) =$

$$= (-R\sin\varphi \, d\varphi)^2 + (R\cos\varphi \, d\varphi)^2 + dh^2 = R^2 d\varphi^2 + dh^2.$$

[for points with $\varphi = 0$ one can consider $\varphi \to \varphi - \pi$].

$\begin{cases} u = R\varphi \\ v = h \end{cases}$ (local reparametrisation)

$$du^2 + dv^2 = R^2 d\varphi^2 + dh^2.$$

Coordinate φ is local. Cylinder $\overset{\circ}{\not\approx} \mathbb{E}^2$ not diffeomorphic since Cylinder $= S^1 \times \mathbb{R}^1$, $S^1 \neq \mathbb{R}$.

A2 [4+2+4]

a) Connection ∇ on a manifold is an operation which assigns to vector fields $\vec{X}, \vec{Y}$ a vector field $\nabla_{\vec{X}} \vec{Y}$ such that

i) $\nabla_{\vec{X}}(\lambda \vec{Y} + \mu \vec{Z}) = \lambda \nabla_{\vec{X}} \vec{Y} + \mu \nabla_{\vec{X}} \vec{Z}$, $\lambda, \mu \in \mathbb{R}$

ii) $\nabla_{f\vec{X}+g\vec{Y}}(\vec{Z}) = f \nabla_{\vec{X}} \vec{Z} + g \nabla_{\vec{Y}} \vec{Z}$

f, g - functions on M $(f, g \in C^\infty(M))$

iii) $\nabla_{\vec{X}}(f \vec{Y}) = \partial_{\vec{X}} f \vec{Y} + f \nabla_{\vec{X}} \vec{Y}$

$\partial_{\vec{X}} f$ - deriv. of f along vector field $\vec{X}$.

$$\nabla_{\frac{\partial}{\partial u}}\left(\frac{\partial}{\partial u}\right) = \frac{\partial}{\partial v} = \Gamma^u_{uu} \frac{\partial}{\partial u} + \Gamma^v_{uu} \frac{\partial}{\partial v} \Rightarrow \Gamma^u_{uu} = 0,\ \Gamma^v_{uu} = 1.$$

b) $\nabla_{\vec{X}} \vec{X} = \nabla_{(\partial_x + x\partial_y)}(\partial_x + x\partial_y) =$

$$= \underline{\Gamma^x_{xx}\partial_x + \Gamma^y_{xx}\partial_y} + x(\Gamma^x_{yx}\partial_x + \Gamma^y_{yx}\partial_y) + \underline{\partial_y} + x(\Gamma^x_{xy}\partial_x + \Gamma^y_{xy}\partial_y) + x^2(\Gamma^x_{yy}\partial_x + \underline{\Gamma^y_{yy}\partial_y}) =$$

$$= xy\partial_x - \partial_y + \partial_y + x^2 y \partial_y =$$

$$= xy(\partial_x + x\partial_y) = xy\vec{X}.$$

A3 (3+3+4)

Let M be a Riemannian manifold with $G = g_{ik}(x)dx^i dx^k$.

We say that connection ∇ is symmetric if $\Gamma^i_{km} = \Gamma^i_{mk}$.

We say that symmetric connection ∇ is Levi-Civita connection if it preserves the scalar product, i.e. for arbitrary vector fields $\vec{X}, \vec{Y}, \vec{Z}$

$$\nabla_{\vec{X}} \langle \vec{Y}, \vec{Z} \rangle = \langle \nabla_{\vec{X}} \vec{Y}, \vec{Z} \rangle + \langle \vec{Y}, \nabla_{\vec{X}} \vec{Z} \rangle$$

$$\partial_{\vec{X}} \langle \vec{Y}, \vec{Z} \rangle =$$

Levi-Civita Theorem claims that on (M, G) there exists uniquely defined Levi-Civita connection. In local coordinates:

$$\Gamma^i_{km} = \frac{1}{2} g^{ij} \left(\frac{\partial g_{jk}}{\partial x^m} + \frac{\partial g_{jm}}{\partial x^k} - \frac{\partial g_{km}}{\partial x^j} \right)$$

b) $G = dr^2 + r^2 d\varphi^2$ $\quad G = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}$, $\quad g_{rr} = 1,\ g_{r\varphi} = 0$, $g_{\varphi r} = 0,\ g_{\varphi\varphi} = r^2$

$$G^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{r^2} \end{pmatrix}$$

$$\Gamma^r_{\varphi\varphi} = \frac{1}{2} g^{rr} \left(\frac{\partial g_{r\varphi}}{\partial \varphi} + \frac{\partial g_{r\varphi}}{\partial \varphi} - \frac{\partial g_{\varphi\varphi}}{\partial r} \right) =$$

$$= \frac{1}{2} \cdot 1 \cdot \left(0 + 0 - \frac{\partial r^2}{\partial r} \right) = -r$$

$$\boxed{\Gamma^r_{\varphi\varphi} = -r}$$

3 – formulation of Th
3 – formula for Γ^i_{km}
4 – calculations

Calculation of $\vec e, \vec f, \vec n$ — 3

Calcul. a, b, c — 7

A4 (3+7)

$$\vec r(\theta,\varphi): \begin{cases} x = R\sin\theta\cos\varphi \\ y = R\sin\theta\sin\varphi \\ z = R\cos\theta \end{cases}$$

$$\vec r_\theta = \frac{\partial \vec r}{\partial \theta} = \begin{pmatrix} R\cos\theta\cos\varphi \\ R\cos\theta\sin\varphi \\ -R\sin\theta \end{pmatrix}, \quad \vec r_\varphi = \frac{\partial \vec r}{\partial \varphi} = \begin{pmatrix} -R\sin\theta\sin\varphi \\ R\sin\theta\cos\varphi \\ 0 \end{pmatrix}$$

$|\vec r_\theta| = R$, $|\vec r_\varphi| = R\sin\theta$. These vectors are orthogonal to each other. Hence choose

$$\vec e = \frac{\vec r_\theta}{R} = \begin{pmatrix} \cos\theta\cos\varphi \\ \cos\theta\sin\varphi \\ -\sin\theta \end{pmatrix}, \quad \cancel{\vec f = \begin{pmatrix} -\sin\theta\cos\varphi \\ \sin \end{pmatrix}} \quad \vec f =$$

$$\vec f = \frac{\vec r_\varphi}{R\sin\theta} = \begin{pmatrix} -\sin\varphi \\ \cos\varphi \\ 0 \end{pmatrix}, \quad \vec e \perp \vec f, \quad |\vec e| = |\vec f| = 1$$

Vector $\vec r \perp$ sphere. Hence $\vec n = \frac{\vec r}{R} = \begin{pmatrix} \sin\theta\cos\varphi \\ \sin\theta\sin\varphi \\ \cos\theta \end{pmatrix}$

$$d\vec e = \begin{pmatrix} -\sin\theta\cos\varphi \\ -\sin\theta\sin\varphi \\ -\cos\theta \end{pmatrix} d\theta + \cos\theta \begin{pmatrix} -\sin\varphi \\ \cos\varphi \\ 0 \end{pmatrix} d\varphi$$

$$d\vec e = -\vec n\, d\theta + \vec f \cos\theta\, \cancel{\cot\theta}\, d\varphi$$

$$d\vec n = \begin{pmatrix} \cos\theta\cos\varphi \\ \cos\theta\sin\varphi \\ -\sin\theta \end{pmatrix} d\theta + \sin\theta \begin{pmatrix} -\sin\varphi \\ \cos\varphi \\ 0 \end{pmatrix} d\varphi = \vec e\, d\theta + \vec f \sin\theta\, d\varphi$$

$$d\begin{pmatrix} \vec e \\ \vec f \\ \vec n \end{pmatrix} = \begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{pmatrix} \begin{pmatrix} \vec e \\ \vec f \\ \vec n \end{pmatrix} = \begin{pmatrix} 0 & \cos\theta\, d\varphi & -d\theta \\ * & * & * \\ d\theta & \sin\theta\, d\varphi & 0 \end{pmatrix}$$

We do not need to calculate second row.

$$\boxed{a = \cos\theta\, d\varphi} \quad \boxed{b = -d\theta}$$

$$\boxed{c = -\sin\theta\, d\varphi}$$

A 5 (3+4+3)

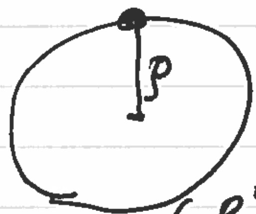
a) Let M be a surface in $\mathbb{E}^3$. Let K be its Gaussian curvature,

$\{R_{ikmn}\}$ – Riemann curvature tensor of Levi-Civita connection, R – scalar curvature of this tensor. Then

3

$$\frac{R}{2} = \frac{R_{1212}}{\det g} = K.$$

b) For sphere of radius ρ



$$G = \rho^2(d\theta^2 + \sin^2\theta)d\varphi^2.$$

$$\left(G = \begin{pmatrix} \rho^2 & 0 \\ 0 & \rho^2\sin^2\theta \end{pmatrix}\right)$$

$$\cancel{\det G|_{\substack{\text{north pole} \\ \theta=}} = \rho^2 \cdot \rho^2 = \rho^4}$$

$$\det G|_{\text{equator}} = \rho^2 \cdot \rho^2\sin^2\theta = \rho^4.$$

4

Gaussian curvature $K = \frac{1}{\rho^2}$

$$R_{1212} = K \det g = \frac{1}{\rho^2} \cdot \rho^4 = \rho^2$$

$R_{1111} = 0$, $R_{1112} = 0$, $-R_{1121} = R_{1122}$

$R_{12ik} = 0$ all component except

$$R_{1212} = \rho^2, \quad R_{1221} = -\rho^2$$

R_{21ik} – all component vanish except $R_{2112} = -\rho^2$, $R_{2121} = \rho^2$

R_{22ik} – all components vanish.

Suppose sphere is locally Euclidean $\Rightarrow G = du^2 + dv^2 \Rightarrow$

$\cancel{\Gamma^i_{km} = 0}$ $\Gamma^i_{km} = 0$ (in coord. u, v) $\Rightarrow R_{iklm} \equiv 0 \Rightarrow K = 0$

3

But $K = \frac{1}{\rho^2}$. Contradiction.

B 6 $(8+7)$

8

Consider vector $\vec{a}_t = \partial_x + t\partial_y$. Then

$(\vec{a}_t, \vec{a}_t) = (\partial_x + t\partial_y, \partial_x + t\partial_y) = 1 + 2pt + t^2$

If $t = -\frac{1}{p}$ then $(\vec{a}_t, \vec{a}_t) = 1 - 2 + \frac{1}{p^2} = \frac{1}{p^2} - 1 = \frac{1-p^2}{1} > 0 \Rightarrow |p| < 1$.

since $\vec{a}_t \neq 0$

7

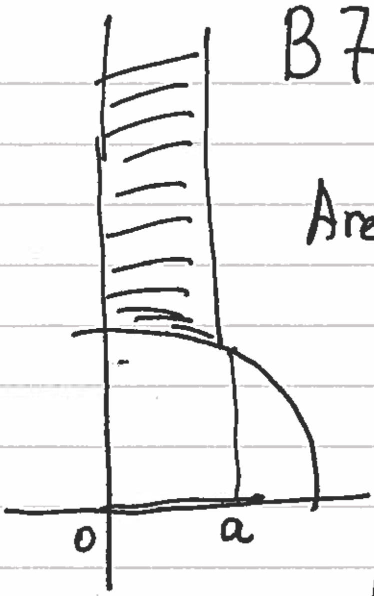
$\|g_{ik}\| = \begin{pmatrix} 1 & p \\ p & 1 \end{pmatrix}$ $\det g = 1 - p^2$.

Area (R) $= \int\limits_{x^2+y^2 \le R^2} \sqrt{\det g}\, dx\, dy =$

$= \det g \cdot \pi R^2 = \pi (1-p^2) R^2$

$(8+7)$

B7



$G = \frac{dx^2+dy^2}{y^2}$, $\sqrt{\det g} = \frac{1}{y^2}$

8

$$\text{Area } D = \int\limits_{\substack{x^2+y^2\geq 1\\ 0\leq x\leq a}} \frac{dx\,dy}{y^2} =$$

$$= \int_0^a dx \int_{\sqrt{1-x^2}}^{\infty} \frac{dy}{y^2} = \int_0^a \frac{dx}{\sqrt{1-x^2}} = \qquad a = \sin t$$

$$\left(\int_r^{\infty} \frac{dy}{y^2} = \frac{1}{r}\right)$$

$$= \int_0^{t_0} \frac{\cos t\,dt}{\cos t} = t_0 = \frac{\pi}{6} \text{ ye} \qquad a = \sin t_0 = \frac{1}{2}.$$

The length of the segment $|A_tB_t| = \sqrt{\frac{a}{t^2}} = \frac{\sqrt{a}}{t}$

7 It means that distance between these points is less (or equal) to $\frac{\sqrt{a}}{t}$. $t \to \infty$ $\frac{\sqrt{a}}{t} \to 0$

(In fact distance is less: $\vdash\!\dashv$)

$d(A_t, B_t) \leq$ length of $[A_t, B_t]$.

Marking scheme

8+7.

B8-L $\qquad$ B8-[7+8]

a) Let $\vec{r} = \vec{r}(u,v)$ be a surface M in $\mathbb{E}^3$

Induced connection

$$\nabla^{ind.}: \nabla^{ind}_{\vec{X}} \vec{Y} = \left(\nabla^{\mathbb{E}^3}_{\vec{X}} \vec{Y}\right)_{tang.}$$

∇^{ind} is symmetric since

$$\nabla^{ind}_{\partial_\alpha} \partial_\beta = (\vec{r}_{\alpha\beta})_{tang.} = (\vec{r}_{\beta\alpha})_{tang.} = \nabla^{ind}_{\partial_\beta} \partial_\alpha$$

Show that it preserves induced metric

$\vec{X}, \vec{Y}, \vec{Z}$ - fields tangent to M

in ambient space

$$\nabla^{\mathbb{E}^3}_{\vec{X}} \langle \vec{Y}, \vec{Z} \rangle = \partial_{\vec{X}} \langle \vec{Y}, \vec{Z} \rangle = \langle \nabla^{\mathbb{E}^3}_{\vec{X}} Y, Z \rangle + \langle Y, \nabla^{\mathbb{E}^3}_{\vec{X}} \vec{Z} \rangle$$

Project this relation on M:

$$\nabla^{\mathbb{E}^3}_{\vec{X}} \vec{Y} = \nabla^{ind}_{\vec{X}} \vec{Y} + \text{orthogonal to } M$$

$$\nabla^{\mathbb{E}^3}_{\vec{X}} \vec{Z} = \nabla^{ind}_{\vec{X}} \vec{Z} + \text{"} \quad \text{"} \quad \Rightarrow$$

$$\partial_{\vec{X}} \langle \vec{Y}, \vec{Z} \rangle_M = \langle \nabla^{ind}_{\vec{X}} \vec{Y}, \vec{Z} \rangle_G + \langle \vec{Y}, \nabla_{\vec{X}} \vec{Z} \rangle_G$$

Thus we proved that induced connection is Levi-Civita connection. Uniqueness of Levi Civita implies that induced = Levi-Civita.

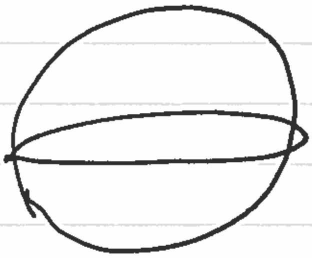
b) Let $\vec{r}(t)$ be parameterised geodesic on M

$$\Downarrow$$

$$\nabla^{ind}_{\vec{v}} \vec{V} = 0, \quad \nabla^{ind}_{\vec{v}} \vec{V} = \left(\nabla^{\mathbb{E}^3}_{\vec{v}} \vec{V}\right)_{tang.} = 0$$

$$\nabla^{\mathbb{E}^3}_{\vec{v}} \vec{V} = \frac{d^2\vec{r}(t)}{dt^2} \Rightarrow \left(\frac{d^2\vec{r}(t)}{dt^2}\right)_{tang.} = 0 \Rightarrow$$

$$\frac{d^2\vec{r}(t)}{dt^2} = \left(\frac{d^2\vec{r}(t)}{dt^2}\right)_{normal}$$



Let $\vec{r}(t)$ be a geodesic on sphere

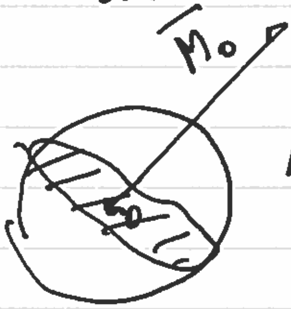
Consider

$$\vec{M}(t) = \vec{r}(t) \times \vec{v}(t)$$

$$\frac{d\vec{M}(t)}{dt} = \vec{v} \times \vec{v} + \vec{r} \times \vec{a} = 0 \text{ since } \vec{a} = \vec{a}_{\perp}$$

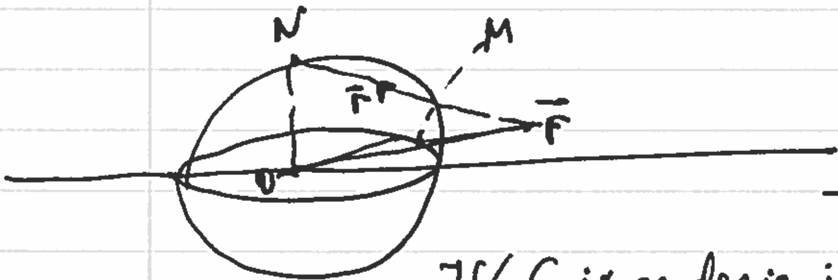
$$\Downarrow$$

$$\vec{M} = \vec{M}_0 = \text{const.}$$



Let α be plane orthogonal to $\vec{M}_0$ passing through centre of sphere.

$\vec{M}_0 = \vec{r}(t) \times v(t) \Rightarrow \vec{r}(t)$ belongs to this plane $\Rightarrow \vec{r}(t)$ is great circle.



Stereogr. projection maps:
N, $\vec{r}$, $\vec{r}(u,v)$ are on the line

If C is geodesic it belongs to the plane which passes via north

Let geodesic C intersects equator at the point M. Then C belong to the plane NMO i.e. stereogr. projection is line OM.

Formulation + Proof of Theorem — 7
Great circles geodesics — 8

— Marking scheme

C9 ~~(8+7)~~ (6+4+5)

Let (M, G) be Riemannian manifold.
Vector field $\vec{K}$ is Killing vector field
if $\mathcal{L}_{\vec{K}} G = 0$ (*)

$\vec{K}$ defines infinitesimal isometry
$\tilde{x}^i = x^i + \varepsilon K^i \quad (\varepsilon^2 = 0)$

(*) in components:

$$(\mathcal{L}_{\vec{K}} g)_{ij} = K^m \partial_m g_{ij} + \partial_i K^m g_{mj} + \partial_j K^m g_{mi} = 0$$

(b) if $\vec{K}$ Killing then $\forall X, Y$

$$\mathcal{L}_{\vec{K}} (G(X, Y)) = \partial_{\vec{K}} G(X, Y) = \mathcal{L}_{\vec{K}} G\underbrace{(\vec{X}, \vec{Y})}_{0} +$$

$$+ G(\mathcal{L}_{\vec{K}} \vec{X}, \vec{Y}) + G(\vec{X}, \mathcal{L}_{\vec{K}} \vec{Y}) =$$

$$= G([K, X], Y) + G(X, [K, Y]) =$$

$[K, X] = \nabla_K X - \nabla_X K$
$[K, Y] = \nabla_K Y - \nabla_Y K$

$$= G(\nabla_K X - \nabla_X K, Y) + G(X, \nabla_K Y - \nabla_Y K) =$$

$$= \underbrace{G(\nabla_K X, Y) + G(X, \nabla_K Y)}_{\partial_K \langle X, Y \rangle} - G(\nabla_X K, Y) - G(X, \nabla_Y K)$$

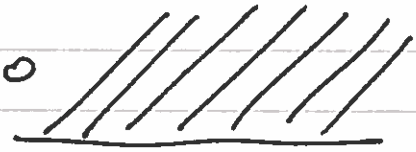
Levi - Civita
We see that

$$\partial_{\vec{K}} \langle \vec{X}, \vec{Y} \rangle = \partial_{\vec{K}} \langle X, Y \rangle - S(X, Y) - S(Y, X), \text{ i.e.}$$

$$S(X, Y) + S(Y, X) = 0$$

— E 9 (continuation)

Let (H, G) – Lobachevsky plane

$y > 0$  $G = \dfrac{dx^2 + dy^2}{y^2}$

$\begin{cases} x' = \lambda x \\ y' = \lambda y \end{cases}$ isometry since $\dfrac{d(\lambda x)^2 + d(\lambda y)^2}{(\lambda y)^2} = \dfrac{dx^2 + dy^2}{y^2}$

infinitesimally $\lambda = 1 + \varepsilon \ (\varepsilon^2 = 0$

$x' = x + \varepsilon x$
$y' = y + \varepsilon y$ $\to$ vector field $K = x\partial_x + y\partial_y$

Hence $K = x\partial_x + y\partial_y$ Killing vector field

Calculate $S_K(\vec{X}, \vec{Y}) = \langle \nabla_{\vec{X}} K, \vec{Y} \rangle$

$$S_K(\vec{X}, \vec{Y}) = S_{mn} X^m Y^n = S_K(\partial_m, \partial_n) X^m Y^n$$

$$S_{mn} = \begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix} \text{ antisymm. metrix}$$

$$S_{12} = a = S_K(\partial_x, \partial_y) = \langle \partial_x \vec{K}, \partial_y \rangle =$$

$$= \langle \partial_x (x\partial_x + y\partial_y), \partial_y \rangle =$$

since $\langle \partial_x, \partial_y \rangle = 0$, $\langle \partial_y, \partial_y \rangle = \dfrac{1}{y^2}$

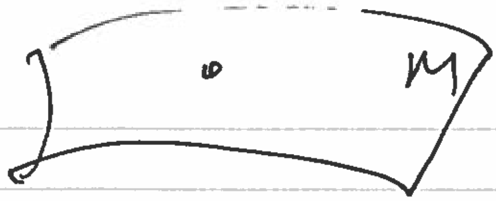
$$= \langle \partial_x + x \Gamma^y_{xx} \partial_y + y \Gamma^y_{xy} \partial_y, \partial_y \rangle =$$

$$= \frac{x \Gamma^y_{xx}}{y^2} + \frac{y \Gamma^y_{xy}}{y^2} = \frac{x}{y^3} \ V.$$

$$S_{21} = -\frac{x}{y^3};$$

Theoretical part $(S_K(\vec{X}, \vec{Y}) = -S_K(\vec{Y}, \vec{X}))$ — 6
Justification of Killing — 4
Calculation S_{1m} — 5

C10 − (10 + 5 + 10)



Let $\{\vec{e}, \vec{f}, \vec{n}\}$ be an orth. basis in $T_p\mathbb{E}^3$ such that at every point $\{\vec{e}, \vec{f}\}$ is an orthonormal basis at T_pM and $\vec{n}$ is normal to M. Then for Gaussian curvature

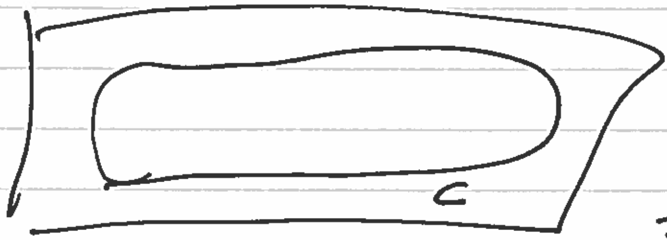
$$K = (b \wedge c)(\vec{e}, \vec{f}) = -da(\vec{e}, \vec{f})$$

where a, b, c are 1-forms in derivation formulae

$$d\begin{pmatrix}\vec{e}\\ \vec{f}\\ \vec{n}\end{pmatrix} = \begin{pmatrix} 0 & a & b\\ -a & 0 & c\\ -b & -c & 0\end{pmatrix}\begin{pmatrix}\vec{e}\\ \vec{f}\\ \vec{n}\end{pmatrix}$$

$$da + b \wedge c = 0.$$

We use it to calculate parallel transport.



$\vec{r} = \vec{r}(u, v)$
u, v coord. on M
$u = u(t),\ v = v(t)$
$\vec{X}(t)$ vector attached at the point $\vec{r}(u(t), v(t))$

$\vec{X}(t)$ is covariantly constant:

$$\frac{\nabla \vec{X}(t)}{dt} = 0. \qquad (*)$$

We will expand $(*)$ with respect to basis $\{\vec{e}, \vec{f}, \vec{n}\}$

Let $\vec{X}(t) = X^1(t)\,\vec{e}\,(u(t), v(t)) + X^2(t)\,\vec{f}\,(u(t), v(t))$.

$$= X^i(t)\,\vec{e}_i\,(u(t), v(t))$$

$$\vec{e}_1 = \vec{e},\quad \vec{e}_2 = \vec{f}$$

C10 continuation 2

C10-2

$$\frac{\nabla X(t)}{dt} = 0 = \nabla_{\vec{v}}\left(X^1(t)\vec{e}(t) + X^2(t)\vec{f}(t)\right) =$$

$$= \left(\dot{X}^1\vec{e} + X^1\nabla_{\vec{v}}\vec{e}\right) + \left(\dot{X}^2\vec{f} + X^2\nabla_{\vec{v}}\vec{f}\right)$$

∇ - induced connection on M:

$$\nabla_{\vec{v}}\vec{X} = \left(\nabla^{\text{const}}_{\vec{v}}\vec{X}\right)_{\text{tang.}}$$

$$\nabla_{\vec{v}}\vec{e} = \left(\nabla^{E^3}_{\vec{v}}\vec{e}\right)_{\text{tang.}} = \left(d\vec{e}(\vec{v})\right)_{\text{tang.}} =$$

$$= \left(a\vec{f} + b\vec{n}\right)_{\text{tang.}} = a\vec{f}$$

$$\nabla_{\vec{v}}\vec{f} = \left(\nabla^{E^3}_{\vec{v}}\vec{f}\right)_{\text{tang.}} = \left(d\vec{f}(\vec{v})\right)_{\text{tang.}} =$$

$$= \left(-a\vec{e} + c\vec{n}\right)_{\text{tang.}} = -a\vec{e}$$

Hence

$$\frac{\nabla\vec{X}(t)}{dt} = 0 = \left(\dot{X}^1 \vec{e} + a(\vec{v})X^1\vec{f}\right) + \left(\dot{X}^2 - a(\vec{v})X^2\vec{e}\right) =$$

$$= \left(\dot{X}^1 - a(\vec{v})X^2\right)\vec{e} + \left(\dot{X}^2 + a(\vec{v})X^1\right)\vec{f} = 0$$

$$\begin{cases} \dot{X}^1 - aX^2 = 0 \\ \dot{X}^2 + aX^1 = 0 \end{cases} \quad \left(a = a(\vec{v})\right)$$

$$\dot{Z}^t = -iaZ \quad \left(Z = X^1 + iX^2\right)$$

$$Z'(t) = e^{-i\int_0^t a(\vec{v}(t))dt} Z(0)$$

If curve is closed then

$$Z_{\text{final}} = Z(0)\,e^{-i\int_0^T a(\vec{v}(t))}$$

$$\int a(\vec{v}(t))dt = \int_C a = \int_{\partial D} a = \int_D da \qquad V$$

C10 − 3

$$Z(t) = e^{-i\int_c a} Z(0) = e^{-i\int_D da} Z(0)$$

$da + b\wedge c = 0$
(Gauss deriv. formula)

$$d\begin{pmatrix}\vec e\\ \vec f\\ \vec n\end{pmatrix} = \begin{pmatrix}0 & a & b\\ -a & 0 & c\\ -b & -c & 0\end{pmatrix}\begin{pmatrix}\vec e\\ \vec f\\ \vec n\end{pmatrix}$$

$$Z(t_f) = e^{i\phi} Z(0)$$

$$\phi = -\int_D da = \int_D b\wedge c = \int_D K\, d\sigma$$

$\Big(b\wedge c = K\,d\sigma$ since $b(\vec e)c(\vec f) - b(\vec f)c(\vec e) = K$
and $d\sigma(\vec e, \vec f) = 1.\Big)$

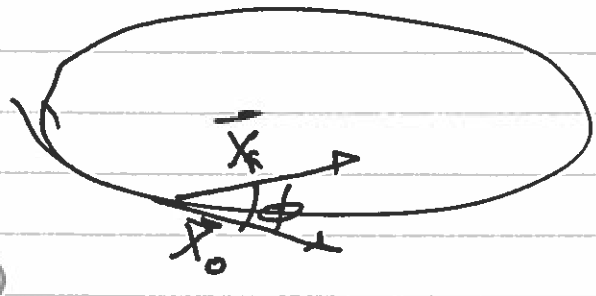
$\underline{Z = X^1 + iX^2}$

$$\cancel{\begin{pmatrix}X^1(t_0)\\ X^2(t_0)\end{pmatrix} = e^{i\phi}(\cos}$$

$$\left(X^1 + iX^2\right)\Big|_{t_{end}} = e^{i\phi}\left(X^1 + iX^2\right)\Big|_{t_{starting}} =$$

$$= (\cos\phi + i\sin\phi)\left(X_0^1 + iX_0^2\right)$$

$$\begin{pmatrix}X^1\\ X^2\end{pmatrix}_f = \underbrace{\begin{pmatrix}\cos\phi & -\sin\phi\\ \sin\phi & \cos\phi\end{pmatrix}}_{\text{rotation on angle } \phi}\begin{pmatrix}X^1\\ X^2\end{pmatrix}$$



C-10-4

$$M: \begin{cases} x = 3h\cos\varphi \\ y = 3h\sin\varphi \\ z = 4h\sin\varphi \end{cases} \qquad C_1 = \partial D. \qquad h > 0$$

upper-sheet of cone is locally Euclidean

(Indeed $G_{M} = dx^2 + dy^2 + dz^2 \Big|_{\substack{x = 3h\cos\varphi \\ y = 3h\sin\varphi \\ z = 4h\cos}} =$

$$= (3dh\cos\varphi - 3h\sin\varphi d\varphi)^2 + (3dh\sin\varphi + 3h\cos\varphi d\varphi) + (4dh)^2 = 25dh^2 + 9h^2 d\varphi^2$$

Changing of coordinates

$$\begin{cases} u = \alpha h\cos\beta\varphi \\ v = \alpha h\sin\beta\varphi \end{cases} \quad du^2 + dv^2 = \alpha^2 dh^2 + \alpha^2\beta^2 h^2 d\varphi^2 \quad \boxed{\begin{matrix} \alpha = 5 \\ \beta = 3/5 \end{matrix}}$$

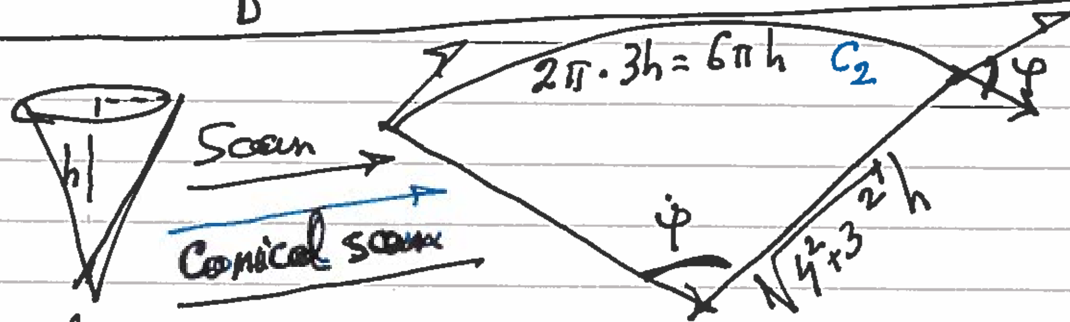
$\downarrow$

Locally Euclidean metric $\Rightarrow R = 0$;

Gaussian curvature $= 0 \Rightarrow$

$$\Delta\phi = \int_D K d\sigma = 0 .$$

C_2



We consider

$$\boxed{\varphi = \frac{6\pi h}{5h} = \frac{6\pi}{5}}$$

Theoretical part — 10.

~~Question about Killing vector~~ —

Integral over C_1 — 5

Integral over C_2 — 10

SECTION A

Answer **ALL** questions in this section (40 marks in total)

A1. (a) Define what is meant by:

Answer. [bookwork — 1+1+1+1+1=5]

— a word $\underline{x}$ of length n in a finite alphabet F;

[**Answer:** an element of $F^n = F \times \cdots \times F$ (*or equivalent*)]

— a code C of length n in a finite alphabet F;

[**Answer:** a non-empty subset of F^n]

— the Hamming distance $d(\underline{x}, \underline{y})$ between two words $\underline{x}$, $\underline{y}$ of length n;

[**Answer:** the number of positions where the two words differ]

— the minimum distance $d(C)$ of a code C;

[**Answer:** $\min\{d(\underline{x}, \underline{y}) : \underline{x}, \underline{y} \in C, \underline{x} \neq \underline{y}\}$]

— a binary code.

[**Answer:** a code in the alphabet $\mathbb{F}_2 = \{0, 1\}$]

(b) Define the trivial binary code of length n and the binary repetition code of length n. For each code, write down the parameters $[n, k, d]_q$ and explain briefly why k, d, q are as you state.

Answer. [seen — 5]

Both codes are binary, meaning that $q = 2$.
The trivial binary code of length n is the set $\mathbb{F}_2^n$ where $\mathbb{F}_2 = \{0, 1\}$. The set $\mathbb{F}_2^n$ has cardinality $M = (\#\mathbb{F}_2)^n = 2^n$ (*this does not require explanation*); by definition, $k = \log_q(M) = \log_2(2^n) = n$. The minimum distance of the trivial code is 1 and is attained at, say, $000\ldots0$ and $100\ldots0$. An $[n, n, 1]_2$-code.
The binary repetition code consists of the two codewords $000\ldots0$ and $111\ldots1$; $k = \log_2(2) = 1$ and the distance between the two codewords is n which gives the minimum distance of the code. Hence it is an $[n, 1, n]_2$-code.

[10 marks]

A2. In this question, C is the binary linear code with generator matrix $\begin{bmatrix} 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \end{bmatrix}$.

(a) List all the codewords of C and check that $d(C) = 3$. How many bit errors can C detect? How many bit errors can C correct? Write down the generator matrix of C in standard form.

Answer. [standard; similar codes were considered in class for which this was done — 5]

$C = \{00000, 01011, 11101, 10110\}$ generated by $\begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{bmatrix}$ in standard form. By inspection $d(C) = w(C) = 3$ so C can detect up to $d - 1 = 2$ errors and correct up to $[(d-1)/2] = 1$ error.

(b) Explain what is meant by a binary symmetric channel with bit error rate p, $BSC(p)$. If $0 \le i \le n$, state without proof the formula for the probability that i errors occur in a message of length n, transmitted down this channel.

Answer. [bookwork — 5]

$BSC(p)$ is a binary channel, that is, it takes one binary symbol $x = 0$ or 1 as input and outputs the symbol x with probability $1-p$ and the symbol $1-x$ with probability p. (*A diagram can be drawn to represent this.*) The channel is memoryless, that is, the probability of an error occurring in a given symbol does not depend on previous errors. The probability of i errors occurring in a message of length n is $\binom{n}{i}p^i(1-p)^{n-i}$.

(c) If the code C is transmitted via $BSC(p)$, calculate $P_{\text{undetect}}(C)$, the probability that the received vector contains an error that is not detected. Set $p = 0.01$ and show that the probability that an unencoded two-bit message sent via $BSC(0.01)$ is received with one or more errors is at least 5,000 times greater than $P_{\text{undetect}}(C)$.

Answer. [routine; such examples were done in class — 5]

$P_{\text{undetect}}(C) = 2p^3(1-p)^2 + p^4(1-p)$. (*The general formula is* $W_C(p, 1-p) - (1-p)^n$ *where* $W_C(x, y) = \sum_{i=0}^{n} A_i x^i y^{n-i}$ *is the weight enumerator.*)
By (b), the probability of a two-bit message arriving with errors is $2p(1-p)+p^2 = 2p-p^2 = 0.0199$. A crude upper bound for $P_{\text{undetect}}(C)$ is $3p^3 = 3 \times 10^{-6}$; clearly $5000 \times 3 \times 10^{-6} = 0.015 < 0.0199$.

(d) Explain why two vectors of weight 1 cannot lie in the same coset of C. Conclude that there are five cosets with coset leaders of weight 1. You are given that coset leaders have weight at most 2; use this information to find $P_{\text{corr}}(C)$, the probability that a codeword transmitted via $BSC(p)$ is decoded correctly (leave your answer as a polynomial in p).

Answer. [easy rider + standard formula — 5]

If $\underline{u}, \underline{v}$ are of weight 1, $\underline{u} \neq \underline{v}$, then $\underline{u} - \underline{v}$ has weight 2 hence cannot lie in C as $w(C) = 3$. So $\underline{u}, \underline{v}$ cannot be in the same coset (*alternatively*: it was shown in the course that each vector of weight $\leq t = [(w(C)-1)/2]$ is a unique coset leader).
A vector of weight 1 must be a coset leader (the only coset containing the vector of weight 0 is C and has no vectors of weight 1), hence there are five coset leaders of weight 1, one coset leader of weight 0, and the remaining $2^{5-2} - 6 = 2$ cosets have coset leaders of weight 2. This leads to $P_{\text{corr}}(C) = (1-p)^5 + 5p(1-p)^4 + 2p^2(1-p)^3$.

[20 marks]

A3.

(a) Define the ISBN-10 code. State, with justification, the number of codewords and the weight of the code. Show that when two unequal symbols are swapped in a codeword, the resulting word does not belong to ISBN-10. (You may use the fact that ISBN-10 is a linear code and you do not have to prove this.)

Answer. [bookwork; all covered in lectures or examples classes — 7]

The code is $C = \{(x_1, x_2, \ldots, x_{10}) \in \mathbb{F}_{11}^{10} \mid 1x_1 + 2x_2 + \ldots + 10x_{10} = 0$ in the field $\mathbb{F}_{11}\}$.
The first nine symbols of a codevector are arbitrary elements of $\mathbb{F}_{11}$, giving 11^9 combinations, and the last symbol is uniquely determined by the formula $x_{10} = x_1 + 2x_2 + 3x_3 + \cdots + 9x_9$. Hence there are 11^9 codevectors. (*There are alternative ways to work that out, e.g. by determining the dimension from a one-row check matrix. They are also acceptable.*)

The weight of the code is 2 and is attained at, e.g., 1000000001. A vector of weight 1 is not a codevector: if $x_i \neq 0$ and $x_j = 0$ ($j \neq i$), the checksum $\sum_{j=1}^{10} jx_j$ is ix_i which is not zero as $i \neq 0$ and $x_i \neq 0$ in the field $\mathbb{F}_{11}$.
If x_i, x_j are swapped in a vector, the checksum changes by $(ix_j + jx_i) - (ix_i + jx_j) = (i - j)(x_j - x_i)$ which is not zero in $\mathbb{F}_{11}$ as long as $i \neq j$ and $x_i \neq x_j$. Hence a codevector (with checksum zero) becomes a non-codevector.

(b) Alice writes down all the codewords of ISBN-10 which are made up of symbols 0 and 1. Bob then counts all the symbols that were written down and claims that there are more 0s than 1s. Do you agree with Bob? Justify your answer.

Answer. [unseen — 3]

Bob is wrong because there is an equal number of 0s and 1s. The shortest way to work that out is probably to observe that 1111111111 is a codevector (the checksum is 55 which is 0 in $\mathbb{F}_{11}$) hence the map $\underline{v} \mapsto 1111111111 - \underline{v}$ is a permutation of Alice's set which replaces all 0s by 1s and all 1s by 0s.

[10 marks]

SECTION B

Answer **TWO** of the three questions in this section (40 marks in total).

If more than TWO questions from this section are attempted, then credit will be given for the best TWO answers.

B4.

(a) Define the inner product on $\mathbb{F}_q^n$. Define the dual code of a linear $[n,k,d]_q$-code C. Define a check matrix of C and state how many rows and columns it has.

Answer. [bookwork — 5]

If $\underline{x} = (x_1, \ldots, x_n)$, $\underline{y} = (y_1, \ldots, y_n)$, then $\underline{x} \cdot \underline{y} = \sum_{i=1}^n x_i y_i$, *or* $\underline{x} \cdot \underline{y} = \underline{x}\underline{y}^T$ (*accept either answer*); $C^\perp = \{\underline{x} \in \mathbb{F}_q^n \mid \underline{x} \cdot \underline{c} = 0 \ \forall \underline{c} \in C\}$. A check matrix of C is a generator matrix of $C^\perp$. It has $n-k$ rows and n columns.

(b) State and prove the Distance Theorem for linear codes.

Answer. [bookwork — 5]

Theorem. Let H be a check matrix of C. Then $d(C) = d$ if and only if every set of $d-1$ columns of H is linearly independent and some set of d columns of H is linearly dependent.
Proof. Recall that a vector $\underline{x} = (x_1, \ldots, x_n) \in \mathbb{F}_q^n$ is a codeword if and only if $\underline{x}H^T = \underline{0}$. This can be written as $x_1\overline{h}_1 + x_2\overline{h}_2 + \ldots + x_n\overline{h}_n = \overline{0}$ where $\overline{h}_1, \ldots, \overline{h}_n$ are the columns of H. The number of columns of H that appear in the above linear dependency with non-zero coefficient is the weight of $\underline{x}$. Therefore, $d(C)$, which is the minimum possible weight of a non-zero codeword of C, is the smallest possible number of columns of H that form a linearly dependent set.

(c) Show that the 7-ary code C with generator matrix $\begin{bmatrix} 1 & 2 & 3 & 0 \\ 0 & 2 & 1 & 3 \end{bmatrix}$ is self-dual, that is, $C = C^\perp$. Find $d(C)$ and justify your answer.

Answer. [standard; both parts appeared in examples done in class — 5]

It was shown in the course that a code with a $k \times n$ generator matrix G is self-dual, if and only if $n = 2k$ and $GG^T = 0$. We have $4 = 2 \times 2$ here, and can calculate $GG^T = \begin{bmatrix} 14 & 7 \\ 7 & 14 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, hence C is self-dual. Then G is also a check matrix for C; G has no zero columns and no two columns are proportional (by inspection) so $d(C) \geq 3$; rows of G which are codevectors have weight 3 so $d(C) = 3$.

(d) Prove: if $D \subseteq \mathbb{F}_7^n$ is a self-dual code, then $2 < d(D) < (n+3)/2$. Any results from the course can be used without particular comment.

Answer. [unseen but main ideas appeared in the course — 5]

It was shown in the course that $k = n/2$ for a self-dual code. By the Singleton bound $d(D) \leq n - k + 1 = (n/2) + 1 < (n+3)/2$.
To show that $2 < d(D)$, equivalently $2 < w(D)$, observe that $\underline{c} \cdot \underline{c} = 0$ for any codevector $\underline{c} \in D$. It is enough to check that non-zero vectors with this property have weight at least 3. Indeed,

$\underline{c} \cdot \underline{c} = c_1^2 + \ldots + c_n^2$; the square of a non-zero $c_i \in \mathbb{F}_7$ can be 1, 2 or 4, and the sum of two squares is never zero in $\mathbb{F}_7$.

[20 marks]

B5.

(a) State without proof the Hamming bound and the Singleton bound for codes in $\mathbb{F}_q^n$ of minimum distance d. Define what is meant by a perfect code.

Answer. [bookwork — 3]

Hamming bound: $M|S_t(\underline{0})| \le q^n$ where $t = \left[\frac{d-1}{2}\right]$ (*or any of the equivalent statements*).
Singleton bound: $M \le q^{n-d+1}$ (*or any of the equivalent statements*).
A code for which the Hamming bound is attained (i.e., holds with an equality) is a perfect code.

(b) Write down the length, dimension and weight of a Hamming code $\mathrm{Ham}(r, q)$. Prove that Hamming codes are perfect.

Answer. [bookwork — 5]

Length $n = (q^r - 1)/(q - 1)$, dimension $k = n - r$, weight 3. To prove that a code with these parameters is perfect, note that $t = [(d - 1)/2] = 1$ and write the Hamming bound in log-form for a one-error-correcting code: $k \le n - \log_q(1 + n(q - 1))$ which is $n - r \le n - \log_q(1 + (q^r - 1))$ — the equality holds.

(c) Describe the decoding algorithm for a Hamming code with a check matrix H over $\mathbb{F}_q$. Explain briefly why the algorithm is valid.

Answer. [bookwork — 5]

We fix an $r \times n$ check matrix H for $\mathrm{Ham}(r, q)$. When a vector $\underline{y} \in \mathbb{F}_q^n$ is received, calculate its syndrome $S(\underline{y}) = \underline{y}H^T$. If it is zero, return $\underline{y}$, otherwise find $\lambda \in \mathbb{F}_q$ and $i \in \{1, \ldots, n\}$ such that $S(\underline{y}) = \lambda \times i$th column of H. Decode $\underline{y}$ by subtracting λ from the ith symbol of $\underline{y}$.
Justification: it was shown in the course that for a perfect linear code with $t = 1$, such as $\mathrm{Ham}(r, q)$, the coset leaders are $\underline{0}$ and vectors of weight 1. The syndrome of the coset leader λe_i of weight 1 is equal to $\lambda \times i$th column of H, so λe_i is the coset leader of the coset of $\underline{y}$. Hence $\underline{y}$ is decoded to $\underline{y} - \lambda e_i$ as per general syndrome decoding.

(d) Prove that all 2017-ary perfect codes with fewer than 2017^{2017} codewords attain the Singleton bound. You may use any results from the course without giving a proof.

Answer. [unseen — 7]

$q = 2017$ is a prime number so the classification theorem for perfect codes applies (*it requires the size of the alphabet to be a prime power*). Since $q \ne 2, 3$ (ruling out repetition codes and Golay codes), by the classification theorem C must have the same parameters $[n, k, d]_q$ as a trivial code or a Hamming code.
Trivial codes attain the Singleton bound: $k = n$, $d = 1$ so $k = n - d + 1$.
A code with parameters of $\mathrm{Ham}(r, q)$, $r \ge 2$, has q^k codewords where $n = (q^r - 1)/(q - 1)$ and $k = n - r$. The question says that $q^k < q^q$ so $k < q$. This is only possible if $r = 2$: e.g., write

$k = (q^{r-1} - 1) + \ldots + (q - 1)$ which is at least $(r - 1)(q - 1)$, greater than q if $r > 2$. Hence $k = n - 2 = n - d + 1$ since $d = 3$ for Hamming codes; the Singleton bound is attained.

[20 marks]

B6.

(a) What is a cyclic code? What is meant by an ideal of a commutative ring R? Briefly explain how a cyclic code of length n over $\mathbb{F}_q$ can be viewed as an ideal of the ring $R_n = \mathbb{F}_q[x]/(x^n - 1)$.

Answer. [bookwork — 5]

A cyclic code is a <u>linear</u> code $C \subseteq \mathbb{F}_q^n$ with the property that $(a_0, a_1, \ldots, a_{n-1}) \in C \implies (a_1, \ldots, a_{n-1}, a_0) \in C$. An ideal is an additive subgroup $I \subseteq R$ such that $RI \subseteq I$. Under identification $(a_0, \ldots, a_{n-1}) \mapsto a_0 + a_1x + \cdots + a_{n-1}x^{n-1}$, a cyclic code is linear hence is an additive subgroup of R_n; also, closure under cyclic shifts translates into $xC \subseteq C$, so by iterating this we obtain $x^iC \subseteq C$ for all i hence by linearity $R_nC \subseteq C$ as required.

You are given a factorisation of the polynomial $x^7 - 1$ into irreducible polynomials over the field $\mathbb{F}_2$:

$$x^7 - 1 = (x + 1)(x^3 + x + 1)(x^3 + x^2 + 1).$$

(b) Show that there exist two binary cyclic codes of length 7 and dimension 3. Write down the generator polynomial and a generator matrix for each of these codes. By comparing the two generator matrices, or otherwise, show that the two codes are linearly equivalent.

Answer. [standard; these cyclic codes were considered many times in the course — 5]

Binary cyclic codes of length 7 and dimension 3 are in one-to-one correspondence with generator polynomials of degree 4 which are factors of $x^7 - 1$ in $\mathbb{F}_2[x]$. From the above factorisation, there are two such polynomials,

$$g_1(x) = (x+1)(x^3+x+1) = x^4+x^3+x^2+1 \quad \text{and} \quad g_2(x) = (x+1)(x^3+x^2+1) = x^4+x^2+x+1,$$

giving rise to the generator matrices

$$G_1 = \begin{bmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix} \quad \text{and} \quad G_2 = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \end{bmatrix}$$

which generate the codes C_1 and C_2. Observing that G_2 is obtained from G_1 by permuting columns (e.g., permutation $2 \to 6 \to 4 \to 2$, $3 \leftrightarrow 5$ *or* notice that both matrices are made up of *all* possible non-zero columns of size 3 — they are parity check matrices for Hamming code $\mathrm{Ham}(3, 2)$), we conclude that C_1 is linearly equivalent to C_2.

(c) Does either of the codes constructed in (b) contain vectors of weight 5? Justify your answer.

Answer. [one of the codes, and the even weight property, was discussed at length in class — 3]

No. The rows of both generator matrices are of weight 4 hence lie in the even weight code E_7. Therefore, both C_1 and C_2 are subspaces of E_7, and all their codevectors are of even weight.

(*One can write down all the* 8 *codevectors of each code, but that will cost students time in the exam.*)

(d) Prove: if n is odd, a cyclic code $C \subseteq \mathbb{F}_2^n$ with $w(C) \geq 3$ has no codevectors of weight $n-2$.

Answer. [unseen — 7]

Assume for contradiction that $C \ni \underline{v}$ where $w(\underline{v}) = n-2$. Then $\underline{v}$ is of the form

$$11\ldots1\,0\,11\ldots1\,0\,11\ldots1,$$

with zeros in the ith and $(i+k)$th bits. The cyclic shift by k bits to the right gives codevector $\underline{v}'$ with zero in the $(i+k)$th bit. The other zero in $\underline{v}'$ is in position $(i+2k) \bmod n$; this cannot be i, otherwise one would have $2k = n$, impossible since n is odd. Therefore, $\underline{v}$ and $\underline{v}'$ differ in exactly two positions, i and $(i+2k) \bmod n$. Hence $\underline{v} + \underline{v}' \in C$ is of weight 2, contradicting $w(C) \geq 3$.

[20 marks]

END OF EXAMINATION PAPER

MATH32062 exam solutions 2016/17

1 (a) (i) $V(J) = \{ (a_1, a_2, \ldots, a_n) \in K^n \mid f(a_1, a_2, \ldots, a_n) = 0 \ \forall f \in J \}$ (1*)

(ii) $I(X) = \{ f \in K[x_1, \ldots, x_n] \mid f(a_1, a_2, \ldots, a_n) = 0 \ \forall (a_1, a_2, \ldots, a_n) \in X \}$ (1*)

(iii) Let $f \in \sqrt{J}$, then $f^n \in J$ for some $n \in \mathbb{N}$.
Hence $(f^n)(P) = 0$ for every $P \in V(J)$. (2*) Bookwork
$(f^n)(P) = (f(P))^n$ and since K is a field. Bookwork
$(f(P))^n = 0$ implies $f(P) = 0$, so $f(P) = 0$ for every $P \in V(J)$.
Therefore $f \in I(V(J))$. This holds for every $f \in \sqrt{J}$, so $\sqrt{J} \subseteq I(V(J))$, as required.

(b) Element of J are of the form $h = f \times (x+y-z^2)^3 + g \times (x^3 + xy^2 - xz^2 - 2xz + 2)$ for some $f, g \in \mathbb{C}[x, y, z]$. Hence

$$\frac{\partial h}{\partial z} = \frac{\partial f}{\partial x} \cdot (x+y-z^2)^3 + f \times 3(x+y-z^2)^2 + \frac{\partial g}{\partial x} \times (x^3 + xy^2 - xz^2 - 2xz + 2)$$
$$+ g \times (3x^2 + y^2 - z^2 - 2z)$$

and

$$\left.\frac{\partial h}{\partial z}\right|_{(1,0,1)} = \left.\frac{\partial f}{\partial x}\right|_{(1,0,1)} (1+0-1^2)^3 + f(1,0,1) \times 3 \times (1+0-1^2)^2 + \left.\frac{\partial g}{\partial x}\right|_{(1,0,1)} (1^3 + 1 \times 0^2 - 1 \times 1^2 - 2 \times 1 \times 1 + 2)$$
$$+ g(1,0,1) \times (3 \times 1^2 + 0^2 - 1^2 - 2 \times 1)$$
$$= \left.\frac{\partial f}{\partial x}\right|_{(1,0,1)} \times 0 + f(1,0,1) \times 0 + \left.\frac{\partial g}{\partial x}\right|_{(1,0,1)} \times 0 + g(1,0,1) \times 0 = 0,$$

so $\left.\frac{\partial h}{\partial z}\right|_{(1,0,1)} = 0$ for every $h \in J$, as required.

$(x+y-z^2)^3 \in J$, so $x+y-z^2 \in \sqrt{J} \subseteq I(V(J))$, but $\frac{\partial (x+y-z^2)}{\partial x} = 1$ every point of $\mathbb{C}^3$, so $x+y-z^2 \notin J$. (4*) (Similar to a question on a problem sheet)

(c) (i) W is irreducible iff there do not exist affine algebraic varieties W_1, W_2 such that $W_1 \neq W$, $W_2 \neq W$ and $W_1 \cup W_2 = W$. (1*)

(2) (c)(ii) We shall prove the contrapositive, W is reducible iff $I(W)$ is not prime.

Assume W is reducible, $W = W_1 \cup W_2$, $W_1 \neq W \neq W_2$, W_1, W_2 are affine algebraic varieties. $W_1 \subseteq W$ implies $I(W_1) \supseteq I(W)$. As W_1, W are affine algebraic varieties, $V(I(W_1)) = W_1$ and $V(I(W)) = W$, so $I(W_1) \neq I(W)$. The there exists $f_1 \in I(W_1) \setminus I(W)$. Similarly, $\exists f_2 \in I(W_2) \setminus I(W)$. For any $P \in W$, $(f_1 f_2)(P) = f_1(P) f_2(P)$, and if $P \in W_1$, then $f(P_1) = 0$, if $P \in W_2$ then $f_2(P) = 0$, so $(f_1 f_2)(P) = 0$, for every $P \in W$. Hence $f_1 f_2 \in I(W)$, but $f_1, f_2 \notin I(W)$, so $I(W)$ is not prime.

6
Bookwork

Assume now that $I(W)$ is not prime. Then there exist polynomials $f_1, f_2 \notin I(W)$ such that $f_1 f_2 \in I(W)$. Let $W_i := V(\langle I(W), f_i \rangle) = \{ P \in W \mid f_i(P) = 0 \}$ for $i = 1, 2$. W_1, W_2 are affine algebraic varieties, $W_1 \subseteq W$, $W_2 \subseteq W$. For any $P \in W$, $0 = (f_1 f_2)(P) = f_1(P) f_2(P)$, so $f_1(P) = 0$ or $f_2(P) = 0$. If $f_1(P) = 0$, then $P \in W_1$, if $f_2(P) = 0$, then $P \in W_2$, so $W = W_1 \cup W_2$. $W_1 \neq W$ and $W_2 \neq W$ as $f_1, f_2 \notin I(W)$. Therefore W is reducible.

(d) $xy^2 - xy = x \cdot y \cdot (y-1)$, so

$W = V(\langle x, x^2-y^2-z^2+1 \rangle) \cup V(\langle y, x^2-y^2-z^2+1 \rangle) \cup V(\langle y-1, x^2-y^2-z^2+1 \rangle)$

Substituting $x=0$ into $x^2-y^2-z^2+1=0$ gives $-y^2-z^2+1=0$, so $V(\langle x, x^2-y^2-z^2+1 \rangle)$ is the circle $y^2+z^2-1=0$ in the plane $x=0$.
Substituting $y=0$ into $x^2-y^2-z^2+1=0$ gives $x^2-z^2+1=0$, this is a hyperbola in the plane $y=0$.
Substituting $y=1$ into $x^2-y^2-z^2+1=0$ gives $x^2-z^2=0$, so $\cancel{x-z=0}$ or $x+z=0$, therefore
$V(\langle y-1, x^2-y^2-z^2+1 \rangle) = V(\langle y-1, x-z \rangle) \cup V(\langle y-1, x+z \rangle)$, and these varieties are lines.

5
(part of it is *)

Thus
$W = V(\langle x, -y^2-z^2+1 \rangle) \cup V(\langle y, x^2-z^2+1 \rangle) \cup V(\langle y-1, x-z \rangle) \cup V(\langle y-1, x+z \rangle)$
These varieties are a circle, a hyperbola and 2 lines, which are all known to be irreducible. None of these varieties contains any of the others, for example because the intersection of any 2 of them is a finite set, so they are the irreducible components of W.

3

2(a)(i). ℓ is tangent to V at P iff $(\nabla f)_P \cdot v = 0$ for every $f \in I(V)$. 1 ×

T_PV is the union of the lines tangent to V at P and P. 1 ×

(ii). The ith component of $J_P \underline{v}$ is $(\nabla f_i)_P \underline{v}$.

$f_i \in I(V)$ for every i, $1 \le i \le r$.

If ℓ is tangent to V at P, then $(\nabla f_i)_P \cdot \underline{v} = 0$ for every i, $1 \le i \le r$, so $J_P \underline{v} = 0$.

4 Bookwork

Assume now that $J_P \underline{v} = 0$. Let $f \in I(V)$

f can be written as $f = \sum_{i=1}^{r} g_i f_i$ for suitable polynomials $g_i \in k[x_1, \ldots, x_n]$. Then $\nabla f = \sum_{i=1}^{r} \nabla(g_i f_i) = \sum_{i=1}^{r} (\nabla g_i) f_i + g_i \nabla f_i)$

and $(\nabla f)_P = \sum^{r} \left((\nabla g_i)_P f_i(P) + g_i(P)(\nabla f_i)_P \right) = \sum_{i=1}^{r} g_i(P)(\nabla f_i)_P$,

since $f_i(P) = 0$ because $f_i \in I(V)$ and $P \in V$.

Hence $(\nabla f)_P \cdot \underline{v} = \sum_{i=1}^{r} g_i(P)(\nabla f_i)_P \cdot \underline{v} = \sum_{i=1}^{r} g_i(P) \times 0 = 0$,

as $J_P \underline{v} = 0$ implies $(\nabla f_i)_P v = 0$ for every i, $1 \le i \le r$.

(b) (i) Let $f = x^5 - (y - x^2)^2$. The Jacobian is

$J = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right) = (5x^4 + 4x(y - x^2), 2(x^2 - y))$

4 × (Standard problem type)

Now we consider if there are any points $P \in C$ where $\operatorname{rk} J_P = 0$, i.e. any point in $P \in \mathbb{A}^2$, where $f(P) = \frac{\partial f}{\partial x}\big|_P = \frac{\partial f}{\partial y}\big|_P = 0$. $\frac{\partial f}{\partial y} = 0$ implies $y = x^2$, substituting this into $\frac{\partial f}{\partial x} = 0$ we get $5x^4 = 0$, so $x = 0$ and then $y = 0$. $(0,0) \in C$, and $\operatorname{rk} J_{(0,0)} = 0$.

At all other points $P \in C$, $\operatorname{rk} J_P = 1$, so $(0,0)$ is the only singular point of C.

4 (b)(ii) $t^2, t^5+t^4 \in \mathbb{C}[\mathbb{A}^1] \cong \mathbb{C}[t]$, so φ is a morphism $\mathbb{A}^1 \to \mathbb{A}^2$. If we substitute $x=t^2,\ y=t^5+t^4$ into f, we get $(t^2)^5 - ((t^2)^2-(t^5+t^4))^2 = t^{10} - (-t^5)^2 = 0$, so φ is a morphism $\mathbb{A}^1 \to C$.

3 *

(b)(iii) $\psi(\varphi(t)) = \psi(t^2, t^5+t^4) = \dfrac{(t^5+t^4)-(t^2)^2}{(t^2)^2} = \dfrac{t^5}{t^4} = t \quad (t \neq 0)$

$$\varphi(\psi(x,y)) = \varphi\left(\frac{y-x^2}{x^2}\right) = \left(\left(\frac{y-x^2}{x^2}\right)^2, \left(\frac{y-x^2}{x^2}\right)^5 + \left(\frac{y-x^2}{x^2}\right)^4\right)$$

$x^5 - (y-x^2)^2 \in I(V)$, therefore

$$\left(\frac{y-x^2}{x^2}\right)^2 = \frac{(y-x^2)^2}{x^4} = \frac{x^5}{x^4} = x \quad \text{in } K(V),$$

4

$$\left(\frac{y-x^2}{x^2}\right)^5 + \left(\frac{y-x^2}{x^2}\right)^4 = \left(\frac{y-x^2}{x^2}\right)^4\left(1+\frac{y-x^2}{x^2}\right) = \left(\left(\frac{y-x^2}{x^2}\right)^2\right)^2 \cdot \frac{y}{x^2} = x^2 \cdot \frac{y}{x^2} = y$$

in $K(V)$.

Therefore $\psi \circ \varphi = id_{\mathbb{A}^1}$, $\varphi \circ \psi = id_C$ as rational maps, so φ, ψ are inverses of each other.

(b)(iv) The only point of C with $x=0$ is $(0,0)$ and the calculations in (iii) show that φ is a bijection between $\mathbb{A}^1 \setminus \{0\}$ and $C \setminus \{(0,0)\}$. $\varphi(0) = (0,0)$, so φ is indeed a bijection between $\mathbb{A}^1$ and C.

C has a singular point, whereas $\mathbb{A}^1$ has none, therefore they cannot be isomorphic.

3

(Unseen, but a bit similar to the cuspidal cubic done in the lectures)

(5) 3 (a) $\Phi(X_0:X_1:\ldots:X_n)=(Y_0:Y_1:\ldots:Y_n)$, where

$$\begin{pmatrix} Y_0 \\ \vdots \\ Y_n \end{pmatrix} = A \begin{pmatrix} X_0 \\ \vdots \\ X_n \end{pmatrix}. \qquad (1*)$$

By this definition, Φ is a rational map given by homogeneous degree 1 polynomials. As A is invertible, $A\begin{pmatrix} X_0 \\ \vdots \\ X_n \end{pmatrix}=\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ implies $\begin{pmatrix} X_0 \\ X_1 \\ \vdots \\ X_n \end{pmatrix}=\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$, so Φ is defined at every $(X_0:X_1:\ldots:X_n)\in\mathbb{P}^n$, therefore it is a morphism $\mathbb{P}^n\to\mathbb{P}^n$.

Let $B=A^{-1}$, then B determines a morphism $\Psi:\mathbb{P}^n\to\mathbb{P}^n$, too.

We claim that $\Psi\circ\Phi=\Phi\circ\Psi=id_{\mathbb{P}^n}$. Let $(X_0:\ldots:X_n)\in\mathbb{P}^n$, then $\Phi(X_0:X_1:\ldots:X_n)=(Y_0:Y_1:\ldots:Y_n)$, where $\begin{pmatrix} Y_0 \\ \vdots \\ Y_n \end{pmatrix}=A\begin{pmatrix} X_0 \\ \vdots \\ X_n \end{pmatrix}$, and

(4) Bookwork

$(\Psi\circ\Phi)(X_0:\ldots:X_n)=\Psi(Y_0:Y_1:\ldots:Y_n)=(Z_0:Z_1:\ldots:Z_n)$, where

$\begin{pmatrix} Z_0 \\ \vdots \\ Z_n \end{pmatrix}=B\begin{pmatrix} Y_0 \\ \vdots \\ Y_n \end{pmatrix}=BA\begin{pmatrix} X_0 \\ \vdots \\ X_n \end{pmatrix}=\begin{pmatrix} X_0 \\ \vdots \\ X_n \end{pmatrix}$, so $(\Psi\circ\Phi)(X_0:X_1:\ldots:X_n)=(X_0:X_1:\ldots:X_n)$

$\forall (X_0:X_1:\ldots:X_n)\in\mathbb{P}^n$. $(\Phi\circ\Psi)(X_0:X_1:\ldots:X_n)=(X_0:X_1:\ldots:X_n)$ is similar.

(b) Let $\varphi_1(z)=(z_1,z_2;z_3,z)=\dfrac{(z_1-z_3)(z_2-z)}{(z_1-z)(z_2-z_3)}$, then

$\varphi_1(z_1)=\infty$, $\varphi_1(z_2)=0$, $\varphi(z_3)=1$. $\varphi_1(z)=\dfrac{\alpha z+\beta}{\gamma z+\delta}$ for some $\alpha,\beta,\gamma,\delta\in K$, and as φ_1 is not constant, it follows that $\alpha\delta-\beta\gamma\neq 0$, so φ_1 is a projective transformation $\mathbb{P}^1\to\mathbb{P}^1$.

Let $\varphi_2(z)=(w_1,w_2,w_3,z)$, then φ_2 is also a projective transformation, and $\varphi_2(w_1)=\infty$, $\varphi_2(w_2)=0$, $\varphi_2(w_3)=1$.

(5*) Bookwork

The inverse of a projective transformation is also a projective transformation, it follows from the matrix description that the composition of 2 projective transformations is also a projective transformation, therefore $\varphi=\varphi_2^{-1}\circ\varphi_1$ is a projective transformation and it satisfies $\varphi(z_i)=w_i$ for $i=1,2,3$.

6

(c). Such a φ exists by (b), and it must satisfy $(-1,2;4,z) = (1,3,11,\varphi(z))$ as projective transformations preserve the cross ratio.

$$(-1,2;4,z) = \frac{(-5)(2-z)}{(-1-z)(-2)} = \frac{5(z-2)}{2(z+1)}$$

$$(1,3,11,\varphi(z)) = \frac{(-10)(3-\varphi(z))}{(1-\varphi(z))(-8)} = \frac{5(\varphi(z)-3)}{4(\varphi(z)-1)}.$$

5*

(Standard problem type)

We get the equation

$$\frac{5(z-2)}{2(z+1)} = \frac{5(\varphi(z)-3)}{4(\varphi(z)-1)} \qquad \Big| \times \tfrac{4}{5}(z+1)(\varphi(z)-1)$$

$$2(z-2)(\varphi(z)-1) = (z+1)(\varphi(z)-3)$$

$$2z\varphi(z) - 4\varphi(z) - 2z + 4 = z\varphi(z) + \varphi(z) - 3z - 3$$

$$\varphi(z)(z-5) = -z-7$$

$$\varphi(z) = \frac{-z-7}{z-5}$$

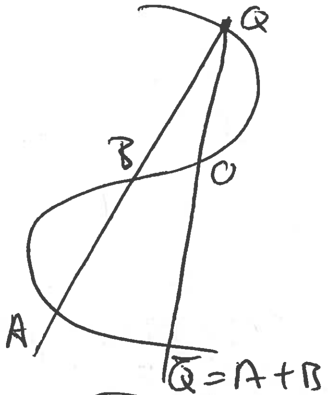
$a=-1$, $b=-7$, $c=1$, $d=-5$ is a solution. (Any others are multiples of this one)

(d). Let $\Phi: \mathbb{P}^1 \to \mathbb{P}^1$ be an isomorphism. By (b) there exists a projective transformation ψ such that if we set $\Theta = \psi \circ \Phi$, then $\Theta(1:0)=(1:0)$, $\Theta(0:1)=(0:1)$ and $\Theta(1:1)=(1:1)$.

$\Theta(X_0,X_1) = (\Theta_0(X_0,X_1):\Theta_1(X_0,X_1))$, where Θ_0, Θ_1 are coprime homogeneous polynomials of the same degree d. Θ is injective, so the only point $(X_0:X_1)$ with $\Theta(X_0,X_1)=(0:1)$ is $(X_0:X_1)=(0:1)$, this implies that $\Theta_0(X_0,X_1) = \alpha X_0^d$ for some $\alpha \in \mathbb{C}\setminus\{0\}$. Similarly, $\Theta_1(X_0,X_1) = \beta X_1^d$ for some $\beta \in \mathbb{C}\setminus\{0\}$. Now $\Theta(1:1)=(1:1)$ implies $\alpha=\beta$. If ε is a dth root of 1 then $\Theta(\varepsilon:1) = (1:1) = \Theta(1:1)$, so the injectivity of Θ implies $d=1$, $\Theta = id_{\mathbb{P}^1}$. Therefore $\Phi = \psi^{-1}$ is a projective transformation.

5 Bookwork

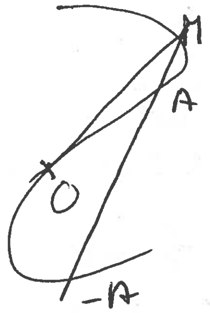
7. 4(a)(i) Let $A, B \in E$. Take the line AB (the tangent line to E at A if $A=B$) and let Q be its 3rd intersection point with E. (If $A \neq B$ and the line is tangent to E at A or B, then that point is Q, if $A=B$ and A is an inflection point, then $Q=A$.). Now take the line OQ (the tangent line at O if $O=Q$) let $\bar{Q}$ be its 3rd intersection point with E (special cases are dealt with as above), then $A+B=\bar{Q}$.



3*

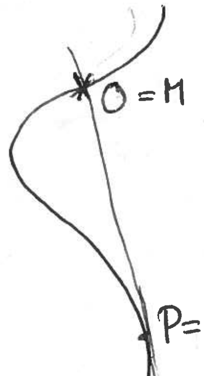
(The case $A=B$ has to be mentioned but full details of all special cases are not needed)

(ii) Let M be the "3rd" point of intersection of the tangent line to E at O with E. (If O is an inflection point, then $O=M$)
$-A$ is the 3rd intersection point of the line AM with E (special case dealt with as above, e.g. if $A=M$, then we take the tangent line at M instead of the line AM.)



2*

(iii) $2P=O$ is equivalent to $P=-P$.
Now $M=O$, so $-P$ is simply the 3rd intersection point of the line OP with E.
This 3rd intersection point is defined to be P exactly when the line OP is tangent to E at P, i.e. when the tangent line to E at P passes through O.



3
Bookwork

(b)(i) The equation of the line AB is $y=2-x$.
By substituting $2-x$ for y in the equation of F we get

$$(2-x)^2 = x^3-2x^2-8x+16.$$

(8)

$$x^2-4x+4 = x^3-2x^2-8x+16$$
$$0 = x^3-x^2-4x+12$$

We know that $(x-2)$ and $(x+2)$ are factors of the RHS, therefore it must factorise as $(x-2)(x+2)(x-3)$, so the x-coordinate of the 3rd intersection point of the line AB with F is 3, and therefore its y-coordinate is $2-3=-1$. Hence $A+B=(3,1)$

4 ×

(Standard calculation)

To calculate 2B, we first have to calculate the equation of the tangent line to F at B.
Implicit differentiation of the equation of F gives

$$2y\frac{dy}{dx} = 3x^2-4x-8$$
$$\frac{dy}{dx} = \frac{3x^2-4x-8}{2y}.$$

By evaluating this at $B=(-2,4)$, we get $\frac{3}{2}$, this is the slope of the tangent line at B, so the equation of the tangent line is $y-4=\frac{3}{2}(x+2)$, $y=\frac{3}{2}x+7$.
By substituting $\frac{3}{2}x+7$ in the equation of F we get.

$$\left(\frac{3}{2}x+7\right)^2 = x^3-2x^2-8x+16$$
$$\frac{9}{4}x^2+21x+49 = x^3-2x^2-8x+16$$
$$0 = x^3-\frac{17}{4}x^2-29x-33.$$

5 ×

(Standard calculation)

We know that $(x+2)^2$ is a factor of the RHS, so it must factorise as $(x+2)^2(x-\frac{33}{4})$, so the x coordinate of the "3rd" intersection point of the tangent line at B with F is $\frac{33}{4}$, and its y-coordinate is $\frac{3}{2}\times\frac{33}{4}+7=\frac{155}{8}$,
so $2B=\left(\frac{33}{4},-\frac{155}{8}\right)$

(ii) $O=(0:1:0)$ is an inflection point, and the lines passing through it are the parallel to the y-axis, so the points of order 2 are those at which the tangent line is parallel to the y-axis. The implicit differentiation in the previous part shows that this happens when $y=0$, so we need to find the roots of $x^3-2x^2-8x+16=0$.
$A=(2,0)\in F$, so $x=2$ is a root, $x^3-2x^2-8x+16=(x-2)(x^2-8)$, so the other roots are $\pm\sqrt{8}=\pm2\sqrt{2}$, so the points of order 2 are $(2,0)$, $(2\sqrt{2},0)$ and $(-2\sqrt{2},0)$.

3

(Mentioned in the lectures, but not stated as a theorem.)

The best three answers count.

1.

(a) **Bookwork (Stating a definition.)**
A finite simple continued fraction is an expression of the form

$$[x_0 : x_1, x_2, \ldots, x_n] := x_0 + \cfrac{1}{x_1 + \cfrac{1}{x_2 + \cfrac{1}{\ddots \, \cfrac{1}{x_n}}}}$$

where $n \geq 0$, $x_0 \in \mathbb{Z}$ and $x_1, \ldots, x_n$ are positive integers. [2 marks]
An infinite simple continued fraction $[x_0 : x_1, x_2, \ldots]$ is the limit of the sequence of finite simple continued fractions $[x_0 : x_1, x_2, \ldots, x_n]$, as n tends to infinity. [2 marks]

(b) **Bookwork (Describing a method.)**
A real number has an infinite simple continued fraction representation if and only if it is irrational. [2 marks]
To find an infinite simple continued fraction representation of an irrational α:

Step 1: Set $\alpha_0 = \alpha$, $\quad x_0 = \lfloor \alpha_0 \rfloor$, $\quad \varepsilon_1 = \alpha_0 - x_0$.

$\vdots$

Step k: Set $\alpha_{k-1} = \dfrac{1}{\varepsilon_{k-1}}$, $\quad x_{k-1} = \lfloor \alpha_{k-1} \rfloor$, $\quad \varepsilon_k = \alpha_{k-1} - x_{k-1}$.

$\vdots$

The algorithm produces an infinite sequence of positive integers $\{x_k\}_{k>0}$, yielding the infinite simple continued fraction representation $\alpha = [x_0 : x_1, x_2, \ldots]$ (i.e. the corresponding sequence of finite simple continued fractions converges to our original irrational number α). [3 marks]

(c) **Standard exercise**.

(i) We apply the Euclidean algorithm:

$$\begin{aligned} 24 &= (1 \times 19) + 5 \\ 19 &= (3 \times 5) + 4 \\ 5 &= (1 \times 4) + 1 \\ 4 &= (4 \times 1) + 0. \end{aligned}$$

To find the solution, we work through the Euclidean algorithm backwards:

$$\begin{aligned} 1 &= 5 - (1 \times 4) \\ &= 5 - (19 - (3 \times 5)) \\ &= (4 \times 5) - 19 \\ &= (4 \times (24 - (1 \times 19))) - 19 \\ &= (4 \times 24) - (5 \times 19). \end{aligned}$$

Thus $(x, y) = (4, -5)$ is a solution. [4 marks]

(ii) The continued fraction expansion can be read off from the numbers m_i given in the Euclidean algorithm ($a_i = m_i b_i + r_i$). This gives

$$\frac{24}{19} = [1 : 3, 1, 4].$$

[4 marks]

(d) **Bookwork (These proofs were covered in lectures.)**

(i) Base case: When $n = -2$ we have

$$p_{-1}q_{-2} - p_{-2}q_{-1} = 1 - 0 = 1 = (-1)^{-2}.$$

[1 marks]

Now suppose that $p_{n+1}q_n - p_n q_{n+1} = (-1)^n$ holds for some $n \geq -2$. Then

$$\begin{aligned} p_{n+2}q_{n+1} - p_{n+1}q_{n+2} &= (x_{n+2}p_{n+1} + p_n)q_{n+1} - p_{n+1}(x_{n+2}q_{n+1} + q_n) \\ &= x_{n+2}(p_{n+1}q_{n+1} - p_{n+1}q_{n+1}) + (p_n q_{n+1} - p_{n+1}q_n) \\ &= -(p_{n+1}q_n - p_n q_{n+1}) \\ &= -(-1)^n = (-1)^{n+1}. \end{aligned}$$

Finally we note that [3 marks]

$$\begin{aligned} p_{n+2}q_n - p_n q_{n+2} &= (x_{n+2}p_{n+1} + p_n)q_n - p_n(x_{n+2}q_{n+1} + q_n) \\ &= x_{n+2}(p_{n+1}q_n - p_n q_{n+1}) + (p_n q_n - p_n q_n) \\ &= x_{n+2}(p_{n+1}q_n - p_n q_{n+1}) \\ &= x_{n+2}(-1)^n. \end{aligned}$$

[3 marks]

(ii) By (i) we have

$$\frac{p_{n+2}}{q_{n+2}} - \frac{p_n}{q_n} = \frac{(-1)^n x_{n+2}}{q_n q_{n+2}} = \begin{cases} \frac{-x_{n+2}}{q_n q_{n+1}}, & \text{if } n \text{ is odd,} \\ \frac{x_{n+2}}{q_n q_{n+1}}, & \text{if } n \text{ is even.} \end{cases}$$

As $x_{n+2} > 0$ and q_n is a sequence of positive increasing integers, the result follows.

[3 marks]

(iii) By (i) we have

$$\frac{p_{n+1}}{q_{n+1}} - \frac{p_n}{q_n} = \frac{(-1)^n}{q_n q_{n+1}} = \begin{cases} \frac{-1}{q_n q_{n+1}}, & \text{if } n \text{ is odd,} \\ \frac{1}{q_n q_{n+1}}, & \text{if } n \text{ is even.} \end{cases}$$

Thus

$$\frac{p_{2n+1}}{q_{2n+1}} - \frac{p_{2n}}{q_{2n}} = \frac{1}{q_n q_{n+1}} > 0,$$

as required.

[3 marks]

2.

(a) **Bookwork (Stating a definition and a result.)**
A periodic simple continued fraction is an infinite simple continued fraction of the form

$$[\underbrace{x_0 : x_1, \ldots, x_{k-1}}_{\text{initial block}}, \underbrace{y_0, \ldots, y_r}_{\text{repeats}}, \underbrace{y_0, \ldots, y_r}_{\text{repeats}}, \ldots],$$

where $x_0 \in \mathbb{Z}$, $x_i, y_i \in \mathbb{N}$, $k \geq 0$ and $r \geq 0$.

[2 marks]

A real number α can be represented by a periodic simple continued fraction if and only if it is a quadratic irrational (that is, if α is an irrational root of a quadratic polynomial with integer coefficients).

[2 marks]

(b) **Bookwork (Describing a method.)**
To find all solutions: We can use the continued fraction expansion of $\sqrt{d}$. This has the form $\sqrt{d} = [x_0 : \overline{x_1, \ldots, x_r}]$ for some $r \geq 1$. The positive solutions to Pell's equation are $x = p_{kr-1}$, $y = q_{kr-1}$ where k ranges over all positive integers such that kr is even. The full set of solutions is therefore $(\pm 1, 0), (\pm p_{kr-1}, \pm q_{kr-1}), (\pm p_{kr-1}, \mp q_{kr-1})$.

[4 marks]

(c) **Standard exercise**.
We apply the continued fraction algorithm. This gives
$\alpha_0 = \sqrt{26}$, $x_0 = 5$, $\varepsilon_1 = \sqrt{26} - 5$,
$\alpha_1 = \dfrac{1}{\sqrt{26} - 5} = \dfrac{\sqrt{26} + 5}{(\sqrt{26} + 5)(\sqrt{26} - 5)} = \sqrt{26} + 5$, $x_1 = 10$, $\varepsilon_2 = \sqrt{26} - 5$.
Since $\varepsilon_2 = \varepsilon_1$, step 3 will produce the same results as step 2, step 4 will produce the same results as step 3, and so on. Thus we conclude that $\sqrt{26} = [5 : \overline{10}]$.

[3 marks]

Using the recurrence relations, we find that

i	-2	-1	0	1	2	3
x_i			5	10	10	10
p_i	0	1	5	51	515	5201
q_i	1	0	1	10	101	1020

Therefore the first four convergents are

$$5, \frac{51}{10}, \frac{515}{101}, \frac{5201}{1020}.$$

[4 marks]

(d)(i) is an exercise. The rest are variants of exam questions from previous years

(d) (i) $x^2 - 26y^2 = 1$. This is Pell's equation, so a positive solution can be found using the continued fraction expansion of $\sqrt{26}$. From above we have $\sqrt{26} = [5 : \overline{10}]$. Since the period is 1 the minimal positive solution will be $(x, y) = (p_1, q_1) = (51, 10)$.

[4 marks]

(ii) $20x^2 - 26y^2 = 1$. The left hand side is divisible by 2, but the right hand side is not. Therefore, there are no integer solutions.

[2 marks]

(iii) $x^2 - 26y^2 = -1$. There is the solution $(x, y) = (5, 1)$.

[2 marks]

(iv) $4x^2 - 9y^2 = 1$. Factorising gives $(2x - 3y)(2x + 3y) = 1$. The only way this can happen is if both factors are equal to ± 1. But $2x + 3y > 1$ for all positive integers x, y, showing that there is no positive integer solution.

[2 marks]

(e) **Difficult exercise (A couple of similar problems appeared on the exercise sheets.)** Let $\omega = [\overline{a : b}]$. Then we have

$$\omega = a + \frac{1}{b + \frac{1}{\omega}} = a + \frac{1}{\frac{b\omega + 1}{\omega}} = a + \frac{\omega}{b\omega + 1} = \frac{ab\omega + a + \omega}{b\omega + 1}.$$

Rearranging gives

$$b\omega^2 - ab\omega - a = 0.$$

Applying the quadratic formula gives

$$\omega = \frac{ab \pm \sqrt{ab(ab + 4)}}{2b}.$$

However, as $\omega > 0$, we need to take the $+$ sign above. This completes the proof.

[5 marks]

3.

(a) **Bookwork (Stating a definition.)**
Let $[a] \in \mathbb{Z}_n$. We say that $[a]$ is a unit if there exists $[b] \in \mathbb{Z}_n$ such that $[a][b] = [b][a] = [1]$. [2 marks]

Bookwork (Standard proof seen in class.)

$[a] \in \mathbb{Z}_n$ is a unit if and only if there exists $b \in \mathbb{Z}$ such that $ab \equiv 1 \bmod n$. By definition of congruence modulo n, this is the case if and only if there exist $b, k \in \mathbb{Z}$ satisfying $ab - kn = 1$. It is clear that the latter implies that $\gcd(a, n) = 1$. On the other hand, if $\gcd(a, n) = 1$, then (e.g. by using the Euclidean algorithm) one can write $ab - kn = 1$ for some integers b, k. Thus $[a]$ is a unit modulo n if and only if a and n are coprime. [3 marks]

(b) **Bookwork (Stating a definition.)**
We say that a unit a is a primitive root modulo n if the cyclic group generated by a is equal to the group of units U_n (i.e. if every unit can be represented as a power of a modulo n.) [2 marks]

(c) **Standard exercise**

(i) As 7 is prime, we have $|U_7| = \varphi(7) = 6$. The possible orders of elements in U_7 are 1,2,3, and 6. By trial and error we find that $3^2 \equiv 2 \bmod 7$ and $3^3 \equiv 6 \bmod 7$. Hence 3 does not have order $1, 2$, or 3, thus has order 6 and so is a primitive root.

[2 marks]

(ii) We have $U_8 = \{[\pm 1], [\pm 3]\}$. We find that $(-1)^2 \equiv 1 \bmod 8$ and $(-3)^2 \equiv 1 \bmod 8$. Thus every element has order 2, so there is no primitive root.

[2 marks]

(d) **Bookwork (Stating results.)**

Lagrange's polynomial congruence theorem: Let $f(x)$ be a polynomial of degree d with integer coefficients and let p be a prime. If some coefficient of f is not divisible by p, then the congruence $f(x) \equiv 0 \bmod p$ has at most d solutions modulo p.

[2 marks]

Fermat's little theorem: Let p be a prime and a an integer not divisible by p. Then $a^{p-1} \equiv 1 \bmod p$.

[2 marks]

(e) **Exercises**

(i) As 13 is prime, by Fermat's little theorem the congruence is equivalent to $x^2 - 1 \equiv 0 \bmod 13$. By Lagrange's congruence theorem there are at most 2 solutions. However $x \equiv \pm 1 \bmod 13$ are solutions, thus these are all of them.

[2 marks]

(ii) Since $169 \equiv 0 \bmod 13$, the congruence can be equivalently written as $12x^4 \equiv 0 \bmod 13$. Since 12 is invertible modulo 13, this is equivalent to $x^4 \equiv 0 \bmod 13$. Thus $13 \mid x^4$ and hence, as 13 is prime, we have $13 \mid x$, so the congruence has the unique solution $x \equiv 0 \bmod 13$.

[2 marks]

(iii) We note that x is a solution to the congruence if and only if it is a solution to the pair of simultaneous congruences $x^{122}+39x^{44}+12 \equiv 0 \bmod 2$ and $x^{122}+39x^{44}+12 \equiv 0 \bmod 13$, or equivalently (by reducing the coefficients) $x^{122}+x^{44} \equiv 0 \bmod 2$ and $x^{122}-1 \equiv 0 \bmod 13$.

It is clear that the first congruence is satisfied for all x (e.g. by Fermat's little theorem).

For the second congruence, we use Fermat's little theorem to see that it is equivalent to $x^2-1 \equiv 0 \bmod 13$. We already solved this in (i). Thus the solutions to the original congruence are $x \equiv \pm 1 \bmod 13$, which gives $x \equiv 1, 12, 14, 25 \bmod 26$.

[3 marks]

(f) (i) **Bookwork (Stating a definition and a result.)**

We define $\varphi : \mathbb{N} \to \mathbb{N}$ by setting $\varphi(n)$ equal to the number of positive integers less than or equal to n which are coprime to n.

[2 marks]

If $n = p_1^{a_1} \cdots p_k^{a_k}$ where the p_i are distinct primes, then

$$\begin{aligned} \varphi(n) &= \varphi(p_1^{a_1})\varphi(p_2^{a_2}) \cdots \varphi(p_k^{a_k}) \\ &= (p_1^{a_1} - p_1^{a_1-1})(p_2^{a_2} - p_2^{a_2-1}) \cdots (p_k^{a_k} - p_k^{a_k-1}). \end{aligned}$$

[2 marks]

(ii) **Difficult variant on previous exam questions where it asked to show that $\varphi(n)$ is even for all $n > 2$.**

One easily checks that $\varphi(3), \varphi(4)$, and $\varphi(6)$ are all prime.

[1 marks]

Suppose that $\varphi(n)$ is prime. Then from the formula we see that

$$(p_1^{a_1} - p_1^{a_1-1})(p_2^{a_2} - p_2^{a_2-1}) \cdots (p_k^{a_k} - p_k^{a_k-1})$$

is prime. The only possibilily is that one term is prime and all others are equal to 1. If

$$p^a - p^{a-1} = 1$$

for some prime p and some $a \geq 1$ then we find that

$$a = 1, \quad p = 2.$$

Similarly, assume that

$$p^a - p^{a-1}$$

is prime. If $a = 1$ then $p-1$ is prime if and only if $p = 3$. If $a \geq 2$ then $p^a - p^{a-1}$ is prime if and only if $p^{a-1} - p^{a-2} = 1$, which from the above gives $p = 2$ and $a = 2$. Bringing this all together gives

$$\{n \in \mathbb{N} : \varphi(n) \text{ is a prime number}\} = \{3,4,6\}.$$

[3 marks]

4.

(a) **Bookwork (Stating two definitions.)**
We say that a unit (invertible element) u of $\mathbb{Z}_p$ is a quadratic residue modulo p if $x^2 \equiv u \bmod p$ for some unit (invertible element) $x \in \mathbb{Z}_p$.

[2 marks]

Let $a \in \mathbb{Z}$. We define

$$\left(\frac{a}{p}\right) = \begin{cases} 0, & \text{if } p \mid a, \\ 1, & \text{if } [a] \text{ is a quadratic residue,} \\ -1, & \text{if } [a] \text{ is a quadratic non-residue.} \end{cases}$$

[3 marks]

(b) **Bookwork (Stating a theorem.)**
Euler's criterion: Let p be an odd prime and let $a \in \mathbb{Z}$. Then

$$\left(\frac{a}{p}\right) \equiv a^{(p-1)/2} \bmod p.$$

[2 marks]

Bookwork (Simple proof from the lectures.)

For $a = -1$ we see that $\left(\frac{-1}{p}\right) \equiv (-1)^{(p-1)/2} \bmod p$ and thus $\left(\frac{-1}{p}\right) = (-1)^{(p-1)/2}$.
Since p is odd, $p-1$ is even, and $(p-1)/2$ will be even precisely when $p \equiv 1 \bmod 4$. The result follows.

[3 marks]

(c) (i) **Bookwork (Stating a theorem.)** Let p and q be distinct odd primes. Then

$$\left(\frac{q}{p}\right)\left(\frac{p}{q}\right) = (-1)^{(p-1)(q-1)/4}.$$

[2 marks]

(ii) **Standard exercises.** Since 7 is a prime, we have $U_7 = \{1, 2, \ldots, 6\}$. The set of quadratic residues can be found by squaring the elements of U_7.
[Using the fact that $(p-x)^2 \equiv x^2 \bmod p$, we only need to compute the squares of $1, 2, 3$.]
$1^2 \equiv 1 \bmod 7, \quad 2^2 \equiv 4 \bmod 7, \quad 3^2 \equiv 2 \bmod 7.$
Thus $Q_7 = \{1, 2, 4\}$.

[3 marks]

(iii) Since $20 \equiv 6 \bmod 7$ and 6 is a quadratic non-residue modulo 7, we see that $\left(\frac{20}{7}\right) = -1$.

[1 marks]

Using the fact that $\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$ gives

$$\left(\frac{15}{31}\right) = \left(\frac{3}{31}\right)\left(\frac{5}{31}\right).$$

Now by quadratic reciprocity

$$\left(\frac{15}{31}\right) = -\left(\frac{31}{3}\right)\left(\frac{31}{5}\right) = -\left(\frac{1}{3}\right)\left(\frac{1}{5}\right) = -1.$$

[3 marks]

(d) **Difficult exercise.**

One easily checks that 1 is the only quadratic residue modulo 3. [1 marks]

If $p \equiv 1 \bmod 4$, then by quadratic reciprocity we have

$$\left(\frac{3}{p}\right) = \left(\frac{p}{3}\right).$$

Thus

$$p \equiv 1 \bmod 3 \quad \text{and} \quad p \equiv 1 \bmod 4.$$

The Chinese remainder theorem then gives the possibilities

$$p \equiv 1 \bmod 12.$$

If $p \equiv 3 \bmod 4$, then by quadratic reciprocity we have

$$\left(\frac{3}{p}\right) = -\left(\frac{p}{3}\right).$$

We find the possibilities

$$p \equiv 2 \bmod 3 \quad \text{and} \quad p \equiv 3 \bmod 4.$$

Thus by the Chinese remainder theorem we have

$$p \equiv 11 \bmod 12.$$

Combining these, we see that 3 is a quadratic residue modulo an odd prime p if and only if

$$p \equiv \pm 1 \bmod 12.$$

[4 marks]

(e) (i) **Bookwork (the proof was seen in the lectures)** Since $[a]$ is a unit each $[a^i]$ is a unit. Hence $\left(\frac{a^i}{p}\right) = \pm 1$. The even powers of $[a]$ are clearly elements of Q_n, since $[a]^{2k} = ([a]^k)^2$. In fact the even powers of $[a]$ are *precisely* the elements of Q_n: if $[b] \in Q_n$, then $[b] = [x]^2$ for some $[x] \in U_n$. Thus $[b] = ([a]^k)^2 = [a]^{2k}$ for some k, since $[a]$ is a primitive root. We have shown that $[a]^i \in Q_p$ if and only if i is even, and the result follows.

[3 marks]

(ii) **Similar to some exercises and already observed in class in some special cases**. We have $|U_p| = p - 1$. As $[a]$ is a primitive root, every element of U_p has the form $[a^i]$ for some $1 \leq i \leq p - 1$. So exactly half of the elements of U_p are an even power of the primitive root $[a]$. The result now follows from (i).

[3 marks]

END OF EXAMINATION PAPER

Two hours

UNIVERSITY OF MANCHESTER

GREEN'S FUNCTIONS, INTEGRAL EQUATIONS AND APPLICATIONS : **SOLUTIONS**

Answer ALL six questions (100 marks in total)

University approved calculators may be used.

1.

(a) The adjoint operator $\mathcal{L}^*$ and boundary conditions $\mathcal{B}^*$ are defined by the requirement that whenever u satisfies $\mathcal{B}$, and v satisfies $\mathcal{B}^*$

$$\langle v, \mathcal{L}u \rangle = \langle \mathcal{L}^* v, u \rangle.$$

Marking guide: This is bookwork. Two marks for the formula involving the inner-products, and one mark for the boundary conditions.

(b) The derivative is

$$p'(x)H(x) + p(0)\delta(x)$$

where δ is the Dirac delta function.

Marking guide: This is the basis for the method of finding Green's functions in one dimension that we studied in class. If p rather than $p(0)$ is given in front of δ then one mark will be deducted. One mark for an indication they know that $H' = \delta$.

(c) The free-space Green's function for the Laplacian in two dimensions is

$$G_{2\infty}(\mathbf{x}, \mathbf{x}_0) = \frac{1}{2\pi} \ln(|\mathbf{x} - \mathbf{x}_0|).$$

Marking guide: This is bookwork. One mark if they give the three dimensional Green's function.

(d) A Fredholm integral equation of the second kind has the form

$$f(x) = u(x) - \lambda \int_a^b K(x, y)u(y) \, \mathrm{d}y.$$

The kernel is $K(x, y)$, and the kernel is called degenerate if it can be written in the form

$$K(x, y) = \sum_{j=1}^{n} w_j(x)v_j(y).$$

Marking guide: This is book work. Two marks for equation, one mark for kernel, one mark for degenerate.

2. First find the complementary solution by solving the characteristic equation

$$m^2 - m - 2 = 0 \Leftrightarrow m = 2, \ -1.$$

Thus a complementary solution is

$$u_c(x) = c_1 e^{2x} + c_2 e^{-x}.$$

Now we find u_1 satisfying the left boundary condition and u_2 satisfying the right boundary condition which are

$$u_1(x) = e^{2x} - e^{-x}, \quad u_2(x) = e^{2x} - 8e^{-x}.$$

Now calculate the Wronskian:

$$\begin{aligned} W(x) &= u_1(x)u_2'(x) - u_1'(x)u_2(x) \\ &= 21e^{x} \end{aligned}$$

Thus the Green's function is

$$\begin{aligned} G(x,x_0) &= \frac{u_1(x)u_2(x_0)}{W(x_0)}H(x_0-x) + \frac{u_1(x_0)u_2(x)}{W(x_0)}H(x-x_0) \\ &= \frac{(e^{2x}-e^{-x})(e^{2x_0}-8e^{-x_0})}{21e^{x_0}}H(x_0-x) + \frac{(e^{2x_0}-e^{-x_0})(e^{2x}-8e^{-x})}{21e^{x_0}}H(x-x_0). \end{aligned}$$

Marking guide: Students will have seen a number of very similar problems. A rough guide for allocation of marks will be:

- Complementary solution: 3 marks.
- Modifying complementary solution to satisfy appropriate BCs: 3 marks.
- Wronskian: 2 marks.
- Correct general formula for G: 1 mark
- Correct final formula/no minor errors: 1 marks.

3. (a) This is a regular Sturm-Liouville eigenvalue problem, and so there are infinitely many eigenvalues.
Marking guide: This is book work. It is not necessary to say this is a regular Sturm-Liouville problem, only that there are infinitely many eigenvalues.

(b) Let λ be a fixed eigenvalue, ϕ a corresponding eigenfunction, and let $\mathcal{L}$ be the differential operator in the eigenvalue problem so that

$$\mathcal{L}\phi(x) = -\lambda x^2\phi(x).$$

Since the boundary value problem is self-adjoint

$$\langle \phi, \mathcal{L}\phi\rangle = \langle \mathcal{L}\phi, \phi\rangle \Rightarrow \langle \phi, -\lambda x^2\phi\rangle = \langle -\lambda x^2\phi, \phi\rangle.$$

Therefore

$$-\lambda\langle \phi, x^2\phi\rangle = -\overline{\lambda}\langle x^2\phi, \phi\rangle \Rightarrow (\overline{\lambda}-\lambda)\langle x^2\phi, \phi\rangle = 0.$$

Since ϕ is not identically zero, the weighted inner product is not zero, and so this implies

$$\overline{\lambda} = \lambda,$$

which is equivalent to saying that λ is real.
Marking guide: This problem is book work. A rough guide for marking:

- Using self-adjointness: 3 marks.
- Replacing $\mathcal{L}\phi$ by $-\lambda x^2\phi$: 2 marks.
- Bringing λ in front of inner products with complex conjugation: 1 mark.
- Explaining that ϕ is not zero, and so division is possible: 1 mark.

(c) The associated boundary value problem would be

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}x}\left(x^2\frac{\mathrm{d}u}{\mathrm{d}x}(x)\right) = f(x), \quad x \in (1,2) \\ u(1) = 0, \quad u(2) = 0. \end{cases}$$

The condition to ensure that a unique solution exists is that zero is not an eigenvalue. If $\{\lambda_n\}_{n=1}^{\infty}$ are the eigenvalues, and $\{\phi_n\}_{n=1}^{\infty}$ the eigenfunctions, then the Green's function for this problem is

$$G(x, x_0) = -\sum_{n=1}^{\infty} \frac{\overline{\phi_n(x_0)}\phi_n(x)}{\lambda_n \int_1^2 |\phi_n(x_1)|^2 x_1^2 \, \mathrm{d}x_1}$$

Marking guide: This problem is bookwork. Two marks for the boundary value problem, one mark for the condition, and three marks for the Green's function.

(d) There are two methods that could work. First, if ϕ satisfies the boundary value problem with $\lambda = 0$, then using integration-by-parts

$$0 = \int_1^2 \frac{\mathrm{d}}{\mathrm{d}x}\left(x^2\frac{\mathrm{d}\phi}{\mathrm{d}x}(x)\right)\overline{\phi(x)} \, \mathrm{d}x = -\int_1^2 x^2 \left|\frac{\mathrm{d}\phi}{\mathrm{d}x}(x)\right|^2 \, \mathrm{d}x.$$

Since $x^2 > 0$ on $[1,2]$, this is only possible if ϕ is identically zero, and so $\lambda = 0$ cannot be an eigenvalue.

The second method involves solving the problem with $\lambda = 0$ explicitly. This is

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}x}\left(x^2\frac{\mathrm{d}\phi}{\mathrm{d}x}(x)\right) = 0, \quad x \in (1,2) \\ \phi(1) = 0, \quad \phi(2) = 0. \end{cases}$$

The ODE is

$$x^2\phi'' + 2x\phi' = 0.$$

This is an Euler equation, and the indicial equation is

$$m(m-1) + 2m = 0 \Rightarrow m = 0, -1.$$

So the general solution is

$$\phi(x) = c_1 + c_2 x^{-1}.$$

Applying the boundary conditions we have

$$0 = \phi(1) = c_1 + c_2, \quad 0 = \phi(2) = c_1 + \frac{c_2}{2} \Rightarrow c_1 = c_2 = 0.$$

Therefore, $\lambda = 0$ is not an eigenvalue.

Marking guide: Students will have found eigenvalues, although not in the case of an Euler equation. Thus this is likely to be a difficult problem. The marking will depend on the method used, but they will be given three marks for work that shows they understand they need to show the only solution of the problem with $\lambda = 0$ is $\phi = 0$.

4. (a) Suppose that $C > 0$ is sufficiently large so that $f(x) = 0$ and $k(x) = k_0$ when $x > C$. Then on the interval $x > C$

$$u''(x) + k(x)^2 u(x) = 0 \Leftrightarrow u(x) = b_1 e^{ik_0 x} + b_2 e^{-ik_0 x}.$$

The boundary condition as $x \to \infty$ is thus equivalent to

$$u'(x) - ik_0 u(x) = -2ik_0 b_2 e^{-ik_0 x} = 0 \quad \text{for } x > C \Leftrightarrow b_2 = 0.$$

Therefore the time-dependent waves on $x > C$ are given by the real part of

$$U(x) = b_1 e^{i(k_0 x - \omega t)},$$

which are waves moving to the right. Therefore all waves sufficiently far to the right are moving to the right, which means no waves are coming from positive infinity.

On the other hand, suppose that $C' < 0$ is sufficiently negative so that $f(x) = 0$ and $k(x) = k_0$ when $x < C'$. Then on the interval $x < C'$

$$u''(x) + k(x)^2 u(x) = 0 \Leftrightarrow u(x) = a_1 e^{ik_0 x} + a_2 e^{-ik_0 x}.$$

The boundary condition $x \to -\infty$ is thus equivalent to

$$u'(x) + ik_0 u(x) = 2ik_0 a_1 e^{ik_0 x} = 0 \quad \text{for } x < C' \Leftrightarrow a_1 = 0.$$

Therefore the time-dependent waves on $x < C'$ are given by the real part of

$$U(x) = a_2 e^{-i(k_0 x + \omega t)},$$

which are waves moving to the left. Therefore all waves sufficiently far to the left are moving to the left, which means no waves are coming from negative infinity.

Marking guide: This question is bookwork. Rough marking guide:

- Formulas for u when $|x|$ is sufficiently large: 2 marks.
- Applying boundary conditions to these formulas: 1 mark.
- Connecting this to the time-dependent waves: 2 marks.

(b) When $k(x) = k_0$ is constant, the complementary solution is

$$u_c(x) = c_1 e^{-ik_0 x} + c_2 e^{ik_0 x}.$$

Note that the boundary conditions are separated, and so we can use the general formula for the Green's function in this case. The Wronskian is

$$W = 2ik_0,$$

and so

$$G(x, x_0) = \frac{e^{ik(x_0 - x)}}{2ik_0} H(x_0 - x) + \frac{e^{ik(x - x_0)}}{2ik_0} H(x - x_0) = \frac{e^{ik_0|x - x_0|}}{2ik_0}$$

Marking guide: This problem is bookwork. Rough marking guide

- Recognizing boundary conditions are separated, and general formula can be used: 2 marks.
- Wronskian: 1 mark.

- Correct final formula for G: 2 marks.

(c) We have
$$u''(x) + k(x)^2 u(x) = 0.$$
Putting in $u(x) = e^{ik_0 x} + u_{sc}(x)$ we have
$$u_{sc}''(x) + k(x)^2 u_{sc}(x) = (k_0^2 - k(x)^2)e^{ik_0 x} \Rightarrow u_{sc}''(x) + k_0^2 u_{sc}(x) = (k_0^2 - k(x)^2)u(x).$$
Since u_{sc} also satisfies the radiation conditions we can apply the Green's function from part (b) to get
$$u_{sc}(x) = \int_{-\infty}^{\infty} e^{ik_0|x-x_0|}\frac{k_0^2 - k(x_0)^2}{2ik_0}u(x_0)\ \mathrm{d}x_0.$$
Adding $e^{ik_0 x}$ we get that u satisfies the Fredolm integral equation
$$u(x) = e^{ik_0 x} + \int_0^1 e^{ik_0|x-x_0|}\frac{k_0^2 - k(x_0)^2}{2ik_0}u(x_0)\ \mathrm{d}x_0.$$
Marking guide: This problem is bookwork. Rough guide for marking.

- Putting form for u into ODE: 1 mark.
- Arriving at non-homogeneous equation for u_{sc}: 2 marks.
- Applying Green's function: 2 marks.
- Adding $e^{ik_0 x}$ to get final equation for u: 1 mark.

(d) Define the integral operator $\mathcal{K}$ by
$$\mathcal{K}u(x) = \int_0^1 e^{ik_0|x-x_0|}\frac{k_0^2 - k(x_0)^2}{2ik_0}u(x_0)\ \mathrm{d}x_0.$$
Then the Neumann series solution for the equation in (c) is
$$u(x) = \sum_{j=0}^{\infty} \mathcal{K}^n e^{ik_0 x}.$$
Marking guide: This problem is bookwork. One mark for definition of $\mathcal{K}$ and two marks for series formula.

(e) The Born approximation consists of that the first two terms in the Neumann series which are in this case
$$u(x) \approx e^{ik_0 x} + \int_0^1 e^{ik_0|x-x_0|}\frac{k_0^2 - k(x_0)^2}{2ik_0}e^{ik_0 x_0}\ \mathrm{d}x_0.$$
Marking guide: This problem is bookwork. Two marks for knowing what the Born approximation is, and one mark for writing down the formula in this case.

5. (a) Start by setting
$$c = \int_1^2 \ln(y)u(y)\mathrm{d}y$$

so that the equation becomes

$$u(x) = \frac{\lambda}{\ln(2^8) - 3} cx + f(x).$$

Multiplying this by $\ln(x)$ and integrating from 1 to 2 we have

$$c = \frac{\lambda}{\ln(2^8) - 3} c \left(\int_1^2 x \ln(x) \, \mathrm{d}x \right) + \int_1^2 \ln(x) f(x) \, \mathrm{d}x,$$

or, after evaluating the integral,

$$c \left(1 - \frac{\lambda}{4} \right) = \int_1^2 \ln(x) f(x) \, \mathrm{d}x.$$

This will have a unique solution c if and only if

$$1 - \frac{\lambda}{4} \neq 0 \Leftrightarrow \lambda \neq 4.$$

In the case that $\lambda \neq 4$ we solve for c to get

$$c = \frac{1}{1 - \frac{\lambda}{4}} \int_1^2 \ln(x) f(x) \, \mathrm{d}x,$$

and putting this back into the original equation we find that

$$u(x) = \frac{\lambda x}{(1 - \frac{\lambda}{4})(\ln(2^8) - 3)} \int_1^2 \ln(y) f(y) \, \mathrm{d}y + f(x)$$

must be the unique solution of the integral equation. To answer (a) plainly then, there is a unique solution for every continuous f if and only if $\lambda \neq 4$.

Marking guide: Note that the marking for parts (a) and (b) overlaps somewhat.

- Define c (2 marks)
- Find equation for c from integral equation (2 marks)
- Correct conclusion. (1 mark)

(b) The answer to part (b) has already been found in the work above for part (a). The formula for the solution is

$$u(x) = \frac{\lambda x}{(1 - \frac{\lambda}{4})(\ln(2^8) - 3)} \int_1^2 \ln(y) f(y) \, \mathrm{d}y + f(x)$$

Marking guide: Note that the marking for parts (a) and (b) overlaps somewhat.

- Find equation for c from integral equation (1 mark)
- Solve for c (2 marks)
- Write down equation for u correctly (2 marks)

(c) Looking back at the work on part (a), we see that when $\lambda = 4$ we must have

$$\int_1^2 \ln(x) f(x)\ \mathrm{d}x = 0$$

for there to be a solution. In this case c can be anything, and so putting c back into the original equation yields the general solution

$$u(x) = cx + f(x)$$

where we have absorbed $\lambda/(\ln(2^8) - 3)$ into the arbitrary constant c.
Marking guide:

- Correct condition (2 marks)
- Correct general solution (2 marks)

6. a) First find the gradient $\nabla_{\mathbf{x}} G(\mathbf{x}, \mathbf{x}_0)$ when $\mathbf{x} \neq \mathbf{x}_0$ by the chain rule:

$$\nabla_{\mathbf{x}} G(\mathbf{x}, \mathbf{x}_0) = \frac{1}{4\pi} \frac{\mathbf{x} - \mathbf{x}_0}{|\mathbf{x} - \mathbf{x}_0|^3}.$$

Now, $\nabla^2_{\mathbf{x}} G(\mathbf{x}, \mathbf{x}_0)$ is the divergence of this gradient, which we calculate when $\mathbf{x} \neq \mathbf{x}_0$ by applying the chain and product rules

$$\nabla^2_{\mathbf{x}} G(\mathbf{x}, \mathbf{x}_0) = \frac{1}{4\pi} \left(\frac{3}{|\mathbf{x} - \mathbf{x}_0|^3} - 3\frac{(\mathbf{x} - \mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)}{|\mathbf{x} - \mathbf{x}_0|^5} \right) = 0.$$

Marking guide: We will have gone over this calculation in lecture.

- Correct gradient: 3 marks.
- Correct calculation of divergence: 3 marks.

(b) We apply Green's formula from the cover of the exam with $D = \mathbb{R}^3 \setminus B_\epsilon(\mathbf{x}_0)$, and $g(\mathbf{x}) = G(\mathbf{x}, \mathbf{x}_0)$ to get

$$\int_{\mathbb{R}^3 \setminus B_\epsilon(\mathbf{x}_0)} (f(\mathbf{x})\nabla^2 G(\mathbf{x}, \mathbf{x}_0) - G(\mathbf{x}, \mathbf{x}_0)\nabla^2 f(\mathbf{x}))\ d\mathbf{x} = \int_{\partial B_\epsilon(\mathbf{x}_0)} (f(\mathbf{x})\nabla_{\mathbf{x}} G(\mathbf{x}, \mathbf{x}_0) - G(\mathbf{x}, \mathbf{x}_0)\nabla f(\mathbf{x})) \cdot \mathbf{n}(\mathbf{x})\ ds(\mathbf{x})$$

where $\mathbf{n}(\mathbf{x}) = (\mathbf{x}_0 - \mathbf{x})/\epsilon$ is the unit normal vector pointing into $B_\epsilon(\mathbf{x}_0)$. Note that this is allowed, and there is no contribution from any other boundary since $f(x) = 0$ for $|x|$ sufficiently large. Next we use part (a) which gives

$$\int_{\mathbb{R}^3 \setminus B_\epsilon(\mathbf{x}_0)} G(\mathbf{x}, \mathbf{x}_0)\nabla^2 f(\mathbf{x})\ d\mathbf{x} = -\int_{\partial B_\epsilon(\mathbf{x}_0)} (f(\mathbf{x})\nabla_{\mathbf{x}} G(\mathbf{x}, \mathbf{x}_0) - G(\mathbf{x}, \mathbf{x}_0)\nabla f(\mathbf{x})) \cdot \mathbf{n}(\mathbf{x})\ ds(\mathbf{x}).$$

Next, using the formula we found in part (a) for $\nabla_{\mathbf{x}} G(\mathbf{x}, \mathbf{x}_0)$ we have

$$\int_{\mathbb{R}^3 \setminus B_\epsilon(\mathbf{x}_0)} G(\mathbf{x}, \mathbf{x}_0)\nabla^2 f(\mathbf{x})\ d\mathbf{x} = -\frac{1}{4\pi\epsilon} \int_{\partial B_\epsilon(\mathbf{x}_0)} \left(\frac{f(\mathbf{x})}{|\mathbf{x} - \mathbf{x}_0|^3}(\mathbf{x} - \mathbf{x}_0) + \frac{\nabla f(\mathbf{x})}{|\mathbf{x} - \mathbf{x}_0|} \right) \cdot (\mathbf{x}_0 - \mathbf{x})\ ds(\mathbf{x}).$$

Finally, using the fact that on $B_\epsilon(\mathbf{x}_0)$, $|\mathbf{x}-\mathbf{x}_0| = \epsilon$, gives

$$\int_{\mathbb{R}^3 \setminus B_\epsilon(\mathbf{x}_0)} G(\mathbf{x},\mathbf{x}_0)\nabla^2 f(\mathbf{x})\ d\mathbf{x} = \frac{1}{4\pi\epsilon^2}\int_{\partial B_\epsilon(\mathbf{x}_0)} f(\mathbf{x}) + \nabla f(\mathbf{x})\cdot(\mathbf{x}-\mathbf{x}_0)\ ds(\mathbf{x}).$$

Notes for marking: This will be covered in lecture.

- Use of Green's formula: 4 marks.
- Correct identification of normal vector: 2 marks.
- Use of formula for $\nabla_{\mathbf{x}} G$: 2 marks
- Use of fact $|\mathbf{x}-\mathbf{x}_0| = \epsilon$ on ball: 2 marks.
- Final formula/no small errors: 1 mark.

(c) First we make the change of variables suggested, $\mathbf{r} = (\mathbf{x}-\mathbf{x}_0)/\epsilon$ on the right hand side of the formula from (b) to get

$$\int_{\mathbb{R}^3 \setminus B_\epsilon(\mathbf{x}_0)} G(\mathbf{x},\mathbf{x}_0)\nabla^2 f(\mathbf{x})\ \mathrm{d}\mathbf{x} = \frac{1}{4\pi}\int_{\mathbb{S}^2} f(\mathbf{x}_0+\epsilon\mathbf{r}) + \epsilon\nabla f(\mathbf{x}_0+\epsilon\mathbf{r})\cdot\mathbf{r}\ ds(\mathbf{r})$$

Now, since f has continuous derivatives, $|\nabla f|$ is bounded, and so upon taking the limit the second term on the right goes to zero, and for the other terms we get in the limit $\epsilon \to 0^+$

$$\int_{\mathbb{R}^3} G(\mathbf{x},\mathbf{x}_0)\nabla^2 f(\mathbf{x})\ \mathrm{d}\mathbf{x} = f(\mathbf{x}_0)\frac{1}{4\pi}\int_{\mathbb{S}^2}\ \mathrm{d}r = f(\mathbf{x}_0),$$

where we have used the fact that the area of the unit sphere is 4π as given on the front of the exam.
Marking guide: This problem will be covered in lecture.

- Making change of variables: 3 marks.
- Arguing why ∇f term goes to zero: 1 mark.
- Using volume of sphere: 1 mark.

Q1. [seen similar but w/o damping]

$$y_{tt} + 2\lambda y_t = c^2 y_{xx}. \qquad y(0) = y(l) = 0.$$

$y = X(x)T(t)$

$$XT'' + 2\lambda XT' = c^2 X''T \Rightarrow \frac{T'' + 2\lambda T'}{c^2 T} = \frac{X''}{X} = -k^2.$$

$k \in \mathbb{R}$ (not required to show).

$$X'' + k^2 X = 0 \Rightarrow X = A\cos(kx) + B\sin(kx)$$

$X(0) = 0 \Rightarrow A = 0$, $X(l) = 0 \Rightarrow k = \frac{n\pi}{l}$ for $B \neq 0$ $n \in \mathbb{N}$.

$$X(x) = B\sin\left(\frac{n\pi x}{l}\right)$$

— (4)

Then $T'' + 2\lambda T' = -k^2c^2 T = -\frac{n^2\pi^2}{l^2}Tc^2$ char. poly. $\mu^2 + 2\lambda\mu + \frac{n^2\pi^2c^2}{l^2} = 0$

So $\mu = \frac{1}{2}\left(-2\lambda \pm \sqrt{4\lambda^2 - \frac{4n^2\pi^2c^2}{l^2}}\right) = -\lambda \pm i\sqrt{\left[\frac{n^2\pi^2c^2}{l^2} - \lambda^2\right]}.$

So $T = \alpha e^{\mu_+ t} + \beta e^{\mu_- t}$

$$= \hat{\alpha} e^{-\lambda t}\cos(\omega_n t + \psi) \qquad \omega_n = \sqrt{\frac{n^2\pi^2c^2}{l^2} - \lambda^2}$$

— (3)

using superposition: — (1)

$$y(x,t) = e^{-\lambda t}\sum_{n=1}^{\infty} A_n \sin\left(\frac{n\pi x}{l}\right)\cos(\omega_n t + \psi_n).$$

(8)

Q2. [unseen].

$$f_t - R^2 (\nabla^2 f)_t - \beta R^2 f_x = 0. \qquad R, \beta > 0.$$

$$-i\omega A - R^2(-i\omega)(-k^2 - l^2) A - \beta R^2 (+ik) A = 0.$$

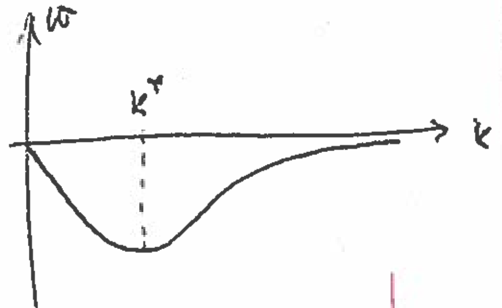
$$\omega - R^2 \omega (-k^2 - l^2) + \beta R^2 k = 0$$

$$\omega \left(1 + R^2(k^2 + l^2)\right) = -\beta R^2 k$$

$$\boxed{\omega = \frac{-\beta R^2 k}{1 + R^2(k^2 + l^2)}}$$ dispersion relation — (3)

For $l = 1$: $$\omega = \frac{-\beta R^2 k}{1 + R^2(k^2 + 1)}$$

$k \to 0$ $\quad \omega \sim -\dfrac{\beta R^2 k}{1 + R^2}$ $\qquad$ $k \to \infty$ $\quad \omega \sim -\dfrac{\beta \cancel{R^2} \cancel{k}}{\cancel{R^2} k^{\cancel{2}}}$ — (3)



$$c_g = \frac{d\omega}{dk}$$

$$= \frac{-\beta R^2}{1 + R^2(k^2+1)} - \frac{\beta R^2 k (-1) . 2kR^2}{\left(1 + R^2(k^2+1)\right)^2}$$

$$= \frac{-\beta R^2}{\left(1 + R^2(k^2+1)\right)^2} \left\{ 1 + R^2(k^2+1) - 2k^2 R^2 \right\}$$

$$= \underbrace{\frac{-\beta R^2}{\left(1 + R^2(k^2+1)\right)^2}}_{-ve} \left\{ \underbrace{1 + R^2(1 - k^2)}_{-ve \text{ if } k > k^* = \left(\frac{1+R^2}{R^2}\right)^{1/2}} \right\}$$ — (3)

$\Rightarrow c_g > 0$ for $k > k^*$.

(9)

Q3 [seen similar apart from (iii)].

$$\eta_t = \phi_z \qquad -\rho\phi_t = -\rho g\eta + T\eta_{xx}. \qquad \text{on } z=0$$

(i) $\nabla^2\phi = 0$ in the fluid. —(1)

(ii) So $\Phi'' - k^2\Phi = 0 \qquad \Phi = Ae^{kz} + Be^{-kz} \qquad k>0.$

$A=0$ for decay as $z\to\infty \qquad \boxed{\Phi = Be^{-kz}}$ —(3)

So $\boxed{\phi = Be^{-kz}\cos(kx-\omega t)}$

Apply the surface condition:

$$-\rho\phi_{tt} = -\rho g\eta_t + T\eta_{txx} \quad \& \quad \eta_t = \phi_z \qquad \text{on } z=0$$

$$-\rho\phi_{tt} = -\rho g\phi_z + \phi_{zxx}.T$$

—(3)

$$\Rightarrow \quad -\rho(-\omega^2) = -\rho g(-k) + (-k)(-k^2).T$$

$$\boxed{\omega^2 = gk + \frac{Tk^3}{\rho}}$$ is the disp. relation.

(iii) $c = \dfrac{\omega}{k} = \varpi \qquad \dfrac{\omega^2}{k^2} = \varpi^2 = g/k + \dfrac{Tk}{\rho}$

or $\dfrac{Tk^2}{\rho} - \varpi^2 k + g = 0 \qquad k = \dfrac{\varpi^2 \pm\sqrt{\varpi^4 - \frac{4T}{\rho}g}}{2T/\rho} = k_+, k_-$

if $\varpi > \varpi_c = \left(\dfrac{4Tg}{\rho}\right)^{1/4}$ then wavelengths are $\lambda_{1,2} = \dfrac{2\pi}{k_\pm}$ associated with wavenumbers k_+, k_-.

if $\varpi < \varpi_c$ there is no such solution. —(4)

(11)

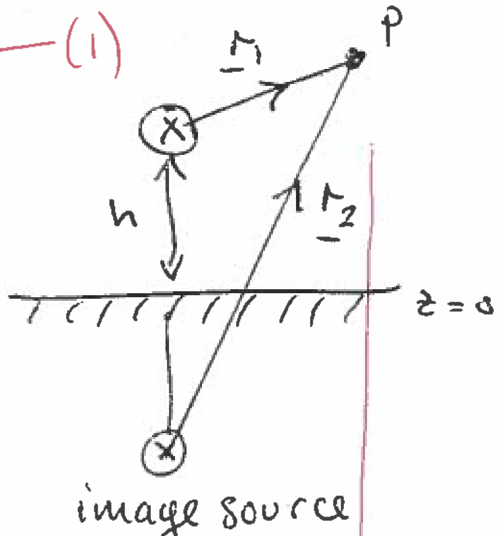
Q4. [in the lecture notes].

(i) $\phi = -\frac{1}{4\pi r} m(t - r/c)$ $\phi_{tt} = -\frac{1}{4\pi r} m''(t - r/c)$

$\phi_r = \frac{1}{4\pi r^2} m(\cdot) + \frac{1}{4\pi r c} m'(\cdot)$ $(r^2\phi_r)_r = \cancel{\frac{m'(\cdot)}{4\pi c}} + \frac{r}{4\pi c} m''(\cdot)\left(-\frac{1}{c}\right)$
$\cancel{-\frac{m'(\cdot)}{4\pi c}}$

So $\phi_{tt} = \frac{c^2}{r^2}(r^2\phi_r)_r$ (3)

(ii) impermeability: $\phi_z = 0$ on $z = 0$. (1)



(iii) $\phi = \phi_1 + \phi_2$ $\phi_1 = -\frac{1}{4\pi r_1} m(t - r_1/c)$

$r_1^2 = x^2 + y^2 + (z-h)^2$

So $r_2^2 = x^2 + y^2 + (z+h)^2$

ϕ_1 satisfies the wave eq^n from (i) (4)

$\phi_2 = -\frac{1}{4\pi r_2} m(t - r_2/c)$ $\therefore$ also satisfies the wave eq^n.

Superposition $\Rightarrow$ $\phi_1 + \phi_2$ satisfies the wave eq^n.

So we need to show that $\frac{\partial}{\partial z}(\phi_1 + \phi_2) = 0$ on $z = 0$.

$$\phi_{1z}\Big|_{z=0} = \phi_{1r_1}\frac{\partial r_1}{\partial z}\Big|_{z=0} = \left\{\frac{1}{4\pi r_1^2} m\left(t - \frac{r_1}{c}\right) + \frac{1}{4\pi r_1 c} m'\left(t - \frac{r_1}{c}\right)\right\}\frac{\partial r_1}{\partial z}\Big|_{z=0}.$$

Q4 [cont.]

$$\varphi_{2z}\Big|_{z=0} = \phi_{2r_2}\frac{\partial r_2}{\partial z}\Big|_{z=0} = \left\{\frac{1}{4\pi r_2^2}m\left(t-\frac{r_2}{c}\right)+\frac{1}{4\pi r_2 c}m'\left(t-\frac{r_2}{c}\right)\right\}\frac{\partial r_2}{\partial z}\Big|_{z=0}$$

however, on $z=0$ $r_1=r_2$ and

$$2r_1\frac{\partial r_1}{\partial z} = 2(z-h) \Rightarrow \frac{\partial r_1}{\partial z}\Big|_{z=0} = -\frac{\cancel{2}h}{\cancel{2}r_1}\Big|_{z=0}$$

$$2r_2\frac{\partial r_2}{\partial z} = 2(z+h) \Rightarrow \frac{\partial r_2}{\partial z}\Big|_{z=0} = +\frac{h}{r_2}\Big|_{z=0}$$

(4)

$\therefore \frac{\partial}{\partial z}(\phi_1+\varphi_2) = 0$ on $z=0$ and imperm. condition is satisfied.

(12)

Q5. [unseen – but we cover the internal wave analogue].

$$\underline{\Omega} \times \underline{u} = \begin{pmatrix} -\Omega v \\ +\Omega u \\ 0 \end{pmatrix} \quad \text{so} \quad \begin{aligned} u_t - 2\Omega v &= -\frac{1}{\rho} p_x \qquad (1) \\ v_t + 2\Omega u &= -\frac{1}{\rho} p_y \qquad (2) \\ w_t &= -\frac{1}{\rho} p_z . \qquad (3) \\ u_x + v_y + w_z &= 0 . \qquad (4) \end{aligned}$$

—(2)

(i) $(1)_x + (2)_y + (3)_z$ leads to

$$-\frac{1}{\rho}\nabla^2 p = \underbrace{(u_x + v_y + w_z)_t}_{=0} + 2\Omega(u_y - v_x) \quad \text{—(5)}$$

—(2)

(i) $(1)_y - (2)_x$ leads to

$$\frac{\partial}{\partial t}(-v_x + u_y) - 2\Omega(u_x + v_y) = 0. \quad \text{—(6)}$$

—(2)

[obvious]

(ii) $(5)_t$ leads to $(\nabla^2 p)_t + \rho 2\Omega(u_y - v_x)_t = 0$

From (6) $(u_y - v_x)_t = +2\Omega(u_x + v_y)$

$$\text{So} \quad (\nabla^2 p)_t + 4\Omega^2 \rho\,(\underbrace{\cancel{u_x + v_y}}_{-w_z}) = 0$$

—(4)

$u_x + v_y = -w_z$ & $(3)_t$ leads to $w_{zt} = -\frac{1}{\rho} p_{zz}$

$$\Rightarrow (\nabla^2 p)_{tt} + 4\Omega^2 p_{zz} = 0.$$

[less obvious!]

(iii) $p = A\exp(ikx + ily + imz - \omega t)$

[Q^n can be picked up again from here]

$$\text{So} \quad (-i\omega)^2(-1)(k^2 + l^2 + m^2) + 4\Omega^2 im.(+im) = 0$$

$$\omega^2 = \frac{4\Omega^2 m^2}{k^2 + l^2 + m^2}.$$

—(2)

Q5 [cont.]

$$\frac{4\Omega^2 m^2}{k^2+l^2+m^2} \leq 4\Omega^2 \quad \text{So} \quad \omega^2 \leq 4\Omega^2 \quad \text{period} \geq \frac{\pi}{\Omega}$$ (1)

(iv) $\underline{c}_g = \left(\partial\omega/\partial k, \partial\omega/\partial l, \partial\omega/\partial m\right) \qquad \underline{k} = (k, l, m)$ (1)

$$2\omega\frac{\partial\omega}{\partial k} = -\frac{4\Omega^2 m^2}{|\underline{k}|^4}\cdot 2k = -\frac{\omega^2}{|\underline{k}|^2}\cdot 2k.$$

$$2\omega\frac{\partial\omega}{\partial l} = -\frac{\omega^2}{|\underline{k}|^2}\cdot 2l \quad (\text{using } k\to l \text{ in above})$$ (3)

$$2\omega\frac{\partial\omega}{\partial m} = \frac{4\Omega^2}{|\underline{k}|^2}\cdot 2m - \frac{4\Omega^2 m^2}{|\underline{k}|^4}\cdot 2m = \omega^2\cdot 2m^{-1} - \frac{\omega^2}{|\underline{k}|^2}\cdot 2m.$$

$$\therefore \quad \frac{\partial\omega}{\partial k} = -\frac{\omega}{|\underline{k}|^2}\cdot k \qquad \frac{\partial\omega}{\partial l} = -\frac{\omega}{|\underline{k}|^2}\cdot l \qquad \frac{\partial\omega}{\partial m} = \frac{\omega}{m} - \frac{\omega m}{|\underline{k}|^2}$$

$$\underline{c}_g\cdot\underline{k} = \frac{\omega}{|\underline{k}|^2}\begin{pmatrix} -k \\ -l \\ \frac{|\underline{k}|^2}{m} - m \end{pmatrix}\cdot\begin{pmatrix} k \\ l \\ m \end{pmatrix}$$

$$= \frac{\omega}{|\underline{k}|^2}\left(-k^2 - l^2 + |\underline{k}|^2 - m^2\right) = 0$$ (2)

energy propagates perpendicularly to the wavenumber vector. (1)

(20)

Q6. [unseen — but have seen planar waveguides].

(i) $f = r\phi \qquad f_{tt} = c^2 f_{rr}$.

Volume flux through a sphere of radius R is

$$Q = \int_{S^r} \frac{\partial \phi}{\partial r}\, dS = 4\pi R^2 \frac{\partial}{\partial r}\left(\frac{f}{r}\right)\Bigg|_{r=R} \quad (3)$$

$$= 4\pi R^2 \left\{ \frac{f_R(R,t)}{R} - \frac{f(R,t)}{R^2} \right\}. = 4\pi f_R \cdot R - 4\pi f$$

on $r = R$.

(ii) For $r = \varepsilon$, $Q(R=\varepsilon) = 0$

$$\Rightarrow -4\pi f(\varepsilon,t) + 4\pi\varepsilon f_R(\varepsilon,t) = 0 \quad (2)$$

As $\varepsilon \to 0$ this reduces to $f(0,t) = 0$ or $F(0) = 0$

provided that $\varepsilon F'(\varepsilon) = o(1)$ as $\varepsilon \to 0$.

ie $\frac{\varepsilon}{\lambda} << 1$ where λ is the wavelength. (1)

(iii) If $r = L$ is impermeable, then $\frac{\partial \phi}{\partial r}\Big|_{r=L} = 0$ } or again using (i).

So $(f/r)_r = 0$ on $r = L$, $Lf_r - f = 0$ on $r = L$. (2)

We must solve $f_{tt} = c^2 f_{rr} \qquad f = e^{i\omega t} F(r)$

so $F'' + \frac{\omega^2}{c^2} F = 0$

$$F = \cancel{A\cos\left(\frac{\omega}{c} r\right)} + B\sin\left(\frac{\omega}{c} r\right)$$

from (ii) $F(0) = 0 \Rightarrow A = 0$

and $LF'(L) - F(L) = 0$

Q6 [cont.]

So L.B. $\frac{\omega}{c}\cos\left(\frac{\omega L}{c}\right) - B\sin\left(\frac{\omega L}{c}\right) = 0$

$$\frac{L\omega}{c} = \tan\left(\frac{\omega L}{c}\right)$$

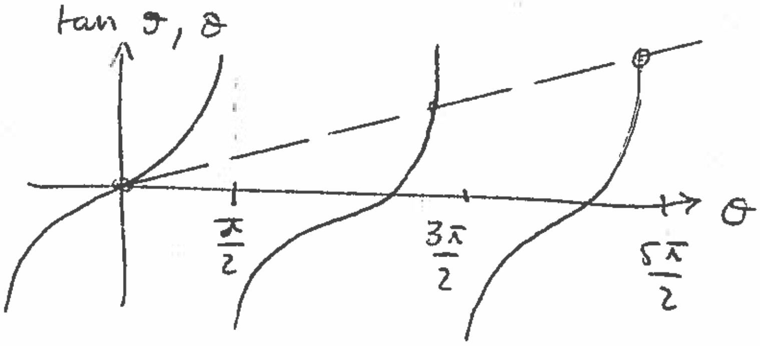
—(4)

time taken to propagate a distance L is $\tau = \frac{L}{c}$

time for one period of the wave $T = \frac{2\pi}{\omega}$

so we need solutions of

$$\theta = \tan\theta \quad \text{where} \quad \theta = \frac{L\omega}{c} = \frac{\tau}{T}.2\pi \gg 1$$



—(4)

for large θ solutions are $\theta \simeq (2n+1)\pi/2$

$$\omega \simeq \left(n+\tfrac{1}{2}\right)\frac{\pi c}{L} \quad \text{for large } n \in \mathbb{N}$$

(iv) if $r=L$ is open s.t $p=0$

$$\frac{\partial\phi}{\partial t} = \frac{\partial}{\partial t}\left(r^{-1}f\right) = 0 \text{ on } r=L \Rightarrow F(L)=0$$

in this case $F = B\sin\frac{\omega r}{c}$ with

$$\frac{\omega L}{c} = n\pi \quad n\in\mathbb{N}, \text{ so } \omega = \frac{n\pi c}{L}.$$

—(4)

(20)

MATH35032: Mathematical Biology
Solutions to the June exam, 2017

The marking scheme in these solutions is provisional: I may revise it slightly after I've reviewed the scripts, which is why I haven't included the same level of detail on the exam paper itself. The remarks below sometimes refer to the previous exam papers for the years 2014–2016: these are the ones readily available to current students.

A1. *This problem should be new to the students, but we've studied many similar things.*

(a) *[3 marks]* The ODE
$$\frac{dN}{dt} = rN^2\left(1 - \frac{N}{K}\right)$$
is reminiscent of the logistic growth law and the parameters are similar too. If we think of $N(t)$ as the number of animals (as opposed to the population density)

r the intrinsic, per-pair *growth rate* of the population: it has units of $1/(\text{pandas} \times \text{time})$.

K the *carrying capacity* of the environment. In this problem it has the same units as N, pandas.

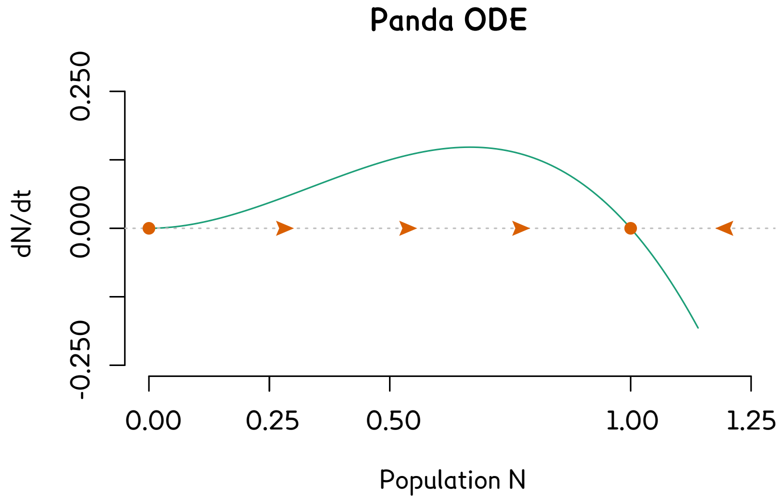
(b) *[3 marks]* A suitable change of variables is
$$x = \frac{N}{K} \qquad \text{and} \qquad \tau = rKt \;\text{ or }\; t = \frac{\tau}{rK}.$$
Then one can compute
$$\begin{aligned}
\frac{du}{d\tau} &= \frac{du}{dN}\frac{dt}{d\tau}\frac{dN}{dt} \\
&= \frac{1}{K}\left(\frac{1}{rK}\right)\left[rN^2\left(1 - \frac{N}{K}\right)\right] \\
&= \frac{N^2}{K^2}\left(1 - \frac{N}{K}\right) \\
&= x^2\left(1 - x\right),
\end{aligned}$$
where the last line is the expression we were aiming for.

(c) *[4 marks]* One can sketch $dx/d\tau$ (see below) and then use topological arguments to say that if $x_0 > 0$ then one of the following three things must happen:

(i) $0 < x_0 < 1$: then $x(t)$ is monotone increasing and $\lim_{t\to\infty} x(t) = 1$;

(ii) $x_0 = 1$: this is a stable equilibrium, so $x(t) = 1$ for all time and, trivially, $\lim_{t\to\infty} x(t) = 1$;

(iii) $x_0 > 1$: then $x(t)$ is monotone decreasing and $\lim_{t\to\infty} x(t) = 1$.



One could, alternatively, integrate the ODE directly,

$$\begin{aligned}
\int_{x_0}^{x(t)} \frac{dx}{x^2(1-x)} &= \int_0^t ds \\
\int_{x_0}^{x(t)} \left(\frac{1+x}{x^2} + \frac{1}{1-x} \right) dx &= t \\
\int_{x_0}^{x(t)} \left(\frac{1}{x^2} + \frac{1}{x} + \frac{1}{1-x} \right) dx &= t \\
\frac{1}{x_0} - \frac{1}{x(t)} + \log\left(\frac{x(t)}{x_0}\right) + \log\left(\frac{1-x_0}{1-x(t)}\right) &= t,
\end{aligned}$$

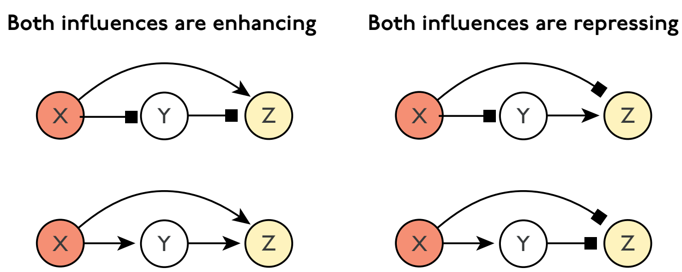
but the result for $x(t)$ is only implicit. All the same, it is clear that as t increases, the right hand side of the equation above grows without bound and so the left hand side must do so as well. The only ways for this to happen are for $x(t)$ to tend to zero or one and, as $x = 0$ is unstable, the only real possibility is that $\lim_{t\to\infty} x(t) = 1$.

A2. *A similar, but not identical problem appeared on 2014's exam and the relevant material is discussed both in lecture and in the problem sets.*

In the language of Uri Alon and his collaborators:

- *[2 marks]* A *motif* is a small, directed, weakly connected subgraph of a regulatory network that has no parallel edges and no loops: they are interesting if they appear more often that one would expect by chance.

- *[3 marks]* A *3-node feed-forward loop* (FFL) is a regulatory motif in which one gene, say X, controls the expression of another, Z in two ways: directly, and indirectly, through the expression of a third gene Y. Such a loop is *coherent* if the both the direct and indirect influences of X on Y have the same sign (that is, both are repressing or both are enhancing). There are four coherent three-node FFLs (all illustrated below) but a correct answer to the question need only include one of them.

  Here a blunt-ended arrow indicates repression while a sharp-ended arrow indicates an enhancing regulatory interaction.

The remaining 5 marks are for the following argument.
To find the desired probability, consider assembling the adjacency matrix of the network at random. We need to scatter E 1's (representing the E regulatory interactions) among the $N \times N$ elements of the matrix. Further, as we exclude parallel edges, we need to place the 1's in distinct positions. There are thus

$$\begin{pmatrix} N^2 \\ E \end{pmatrix} = \frac{(N^2)!}{E!\,(N^2-E)!}$$

possible random regulatory networks, all equally likely.

To count the networks that contain the two-node feedback loop pictured in the exam, note that one first needs to choose the two genes that participate in the loop. Installing the two interactions that define the loop then leaves a further $(E-2)$ interactions to be accounted for, which suggests that there are

$$\begin{pmatrix} N \\ 2 \end{pmatrix} \begin{pmatrix} N^2-2 \\ E-2 \end{pmatrix}$$

networks containing the two-node motif. If $E \geq 4$ this formula will over-count networks that include more than one examples of the motif and so one must account for the possibility that the graph contains two or more such mutually-influencing pairs (one needs to do an inclusion/exclusion sum), but as the problem asks about the case $E = 3$ there can be at most one such pair, no over-counting is possible, and the desired probability is just

$$P = \frac{\binom{7}{2}\binom{47}{1}}{\binom{49}{3}} = \frac{21 \times 47}{18424} = \frac{3}{56}.$$

A3. *We studied the Lotka-Volterra system extensively, both in lecture and in example classes, though mainly in dimensionless form, so the conserved quantity in part (b) may thus look slightly unfamiliar. The calculation in part (c) comes directly from a problem set.*

(a) *[2 marks]* At an equilibrium $(N_\star, P_\star)$ one has $dN/dt = dP/dt = 0$ or

$$0 = N_\star(a - bP_\star) \qquad \text{and} \qquad 0 = P_\star(cN_\star - d).$$

The first implies that either $N_\star = 0$ or $P_\star = a/b$, while the second says either $P_\star = 0$ or $N_\star = d/c$. Thus the only equilibrium with both populations positive is

$$N_\star = \frac{d}{c} \qquad \text{and} \qquad P_\star = \frac{a}{b}.$$

(b) *[5 marks in total, distributed as indicated below]*

- *[1 mark]* A *constant of the motion* is a function $H : \mathbb{R}^2 \to \mathbb{R}$ such that, for any solution to the ODEs $(N(t), P(t))$ we have

$$H(N(t), P(t)) = \text{a constant} = H(N_0, P_0)$$

where $(N_0, P_0) = (N(0), P(0))$ is the initial data.

- *[1 mark]* The constancy of H along trajectories implies that H satisfies the PDE

$$\frac{\partial H}{\partial N}\frac{dN}{dt} + \frac{\partial H}{\partial P}\frac{dP}{dt} = 0 \tag{3.1}$$

- *[2 marks]* For the proposed function

$$\frac{\partial H}{\partial N} = \left(c - \frac{d}{N}\right) H \qquad \text{and} \qquad \frac{\partial H}{\partial P} = \left(b - \frac{a}{P}\right) H.$$

- *[1 mark]* Substituting these into the PDE (3.1) leads to:

$$\begin{aligned} \left(c - \frac{d}{N}\right) HN(a - bP) + \left(b - \frac{a}{P}\right) HP(cN - d) &\stackrel{?}{=} 0 \\ (cN - d)\, H(a - bP) + (bP - a)\, H(cN - d) &\stackrel{?}{=} 0 \\ 0 &= 0 \end{aligned}$$

(c) *[3 marks]* Suppose $(N(t), P(t))$ is periodic with period T. From the original ODEs, we know

$$\frac{dP}{dt} = P(cN - d) \qquad \text{or} \qquad \frac{1}{P}\left(\frac{dP}{dt}\right) = cN - d.$$

If we integrate both sides of this equation over one period we find

$$\int_0^T \frac{1}{P}\left(\frac{dP}{dt}\right) dt \;=\; \int_0^T (cN - d)\; dt. \tag{3.2}$$

Treating the left-hand side first, and changing the variable of integration to P, we find

$$\begin{aligned}\int_0^T \frac{1}{P}\left(\frac{dP}{dt}\right)\,dt &= \int_{P(0)}^{P(T)}\left(\frac{1}{P}\right)\,dP \\ &= \ln(P(T)) - \ln(P(0)) \\ &= 0 \end{aligned} \tag{3.3}$$

where the last equality follows as $P(T) = P(0)$ by the T-periodicity of $P(t)$. On the other hand, we can also integrate the right-hand side of (3.2) to get

$$\begin{aligned}\int_0^T (cN - d)\,dt &= c\left(\int_0^T N(t)\,dt\right) - dT \\ &= cT\,\overline{N} - dT. \end{aligned} \tag{3.4}$$

Putting (3.3) and (3.4) together yields

$$0 = cT\,\overline{N} - dT \qquad \text{or} \qquad \overline{N} = \frac{d}{c} = N_\star,$$

which is the result we sought.

B4. *We studied the Principle of Competitive Exclusion using a model similar to the one here, but with $m_1 = m_2 = 0$. These added terms complicate matters and make the problem more interesting and challenging.*

(a) *[2 marks]* The m_j are dimensionless numbers that specify the rate at which animals are removed. For example, the ODE for N_1 is the same as in a standard two-species competition model, but with an added term $-r_1 m_1 N_1$. If both m_j exceed one, then $(1 - m_j) < 0$ in both equations and so the factors

$$\left(1 - m_1 - \frac{N_1 + a_1 N_2}{K_1}\right) \qquad \text{and} \qquad \left(1 - m_2 - \frac{N_2 + a_2 N_1}{K_2}\right)$$

are both strictly negative. In this case the only biologically sensible way to make the right-hand-sides of the ODEs vanish is to set $N_1 = N_2 = 0$. Any other equilibria would involve negative populations for one or both species.

(b) *[5 marks]* Suitable choices are

$$u = \frac{N_1}{K_1}, \qquad v = \frac{N_2}{K_2} \text{ and } \tau = r_1 t.$$

Then standard calculations lead to

$$\begin{aligned}
\frac{du}{d\tau} &= \frac{dt}{d\tau}\left(\frac{du}{dN_1}\right)\frac{dN_1}{dt} \\
&= \frac{1}{r_1}\left(\frac{1}{K_1}\right) r_1 N_1 \left(1 - m_1 - \frac{N_1 + a_1 N_2}{K_1}\right) \\
&= \frac{N_1}{K_1}\left(1 - m_1 - \frac{N_1}{K_1} + \left(a_1 \frac{K_2}{K_1}\right)\frac{N_2}{K_2}\right) \\
&= u\left((1 - m_1) - u + \left(a_1 \frac{K_2}{K_1}\right) v\right)
\end{aligned}$$

which has the desired form provided we define

$$\gamma_1 = (1 - m_1) \text{ and } \alpha = a_1 \left(\frac{K_2}{K_1}\right).$$

Similar calculations for $dv/d\tau$ lead to

$$\rho = \frac{r_2}{r_1}, \qquad \gamma_2 = (1 - m_2) \text{ and } \beta = a_2 \left(\frac{K_1}{K_2}\right).$$

(c) *[5 marks]* Null clines are curves on which one of the two derivatives vanish. Figure B4.1 provides the desired sketches and shows that when $\alpha < \gamma_1 < 1/\beta < 1$ and $\gamma_2 = 1$, the equilibria $(u_\star, v_\star)$ are $(u_\star, v_\star) = (0, 0)$, $(0, 1)$, $(\gamma_1, 0)$ and

$$(u_\star, v_\star) = \left(\frac{\gamma_1 - \alpha}{1 - \alpha\beta}, \frac{1 - \beta\gamma_1}{1 - \alpha\beta}\right)$$

The conditions $\alpha < \gamma_1 < 1/\beta < 1$ ensure that this last equilibrium has positive coordinates, as they imply that $(1 - \alpha\beta) > 0$, $(\gamma_1 - \alpha) > 0$ and $(1 - \beta\gamma_1) > 0$.

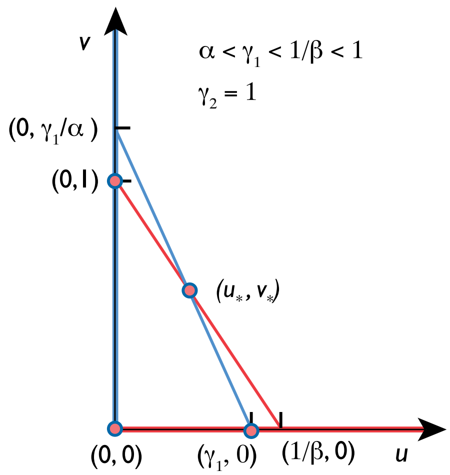


Figure B4.1: *Null clines and equilibria for the case* $\alpha < \gamma_1 < 1/\beta < 1$ *and* $\gamma_2 = 1$. *Here the curves on which* $du/d\tau = 0$ *are shown in blue, while those on which* $dv/d\tau = 0$ *are in red.*

(d) *[4 marks]* The general linearisation is

$$A = \begin{bmatrix} (\gamma_1 - 2u - \alpha v) & -\alpha u \\ -\rho\beta v & \rho(\gamma_2 - 2v - \beta u) \end{bmatrix}$$

and, at an equilibrium $(u_\star, v_\star)$ where both populations are nonzero, we also know

$$(\gamma_1 - u_\star - \alpha v_\star) = 0 = (\gamma_2 - v_\star - \beta u_\star).$$

But then the diagonal entries in A simplify:

$$\begin{aligned} (\gamma_1 - 2u_\star - \alpha v_\star) &= -u_\star + (\gamma_1 - u_\star - \alpha v_\star) \\ &= -u_\star \\ \rho(\gamma_2 - 2u_\star - \alpha v_\star) &= -\rho v_\star + \rho(\gamma_2 - u_\star - \alpha v_\star) \\ &= -\rho v_\star \end{aligned}$$

(e) *[7 marks]* The table below summarises the linear stabilities of the equilibria for the case $\alpha < \gamma_1 < 1/\beta < 1$ and $\gamma_2 = 1$.

$(u_\star, v_\star)$	A	Stability
$(0,\ 0)$	$\begin{bmatrix} \gamma_1 & 0 \\ 0 & \rho \end{bmatrix}$	unstable node
$(0,\ 1)$	$\begin{bmatrix} \gamma_1 - \alpha & 0 \\ -\rho\beta & -\rho \end{bmatrix}$	saddle
$(\gamma_1, 0)$	$\begin{bmatrix} -\gamma_1 & -\alpha\gamma_1 \\ 0 & \rho(1 - \gamma_1\beta) \end{bmatrix}$	saddle
$\left(\dfrac{\gamma_1 - \alpha}{1 - \alpha\beta}, \dfrac{1 - \beta\gamma_1}{1 - \alpha\beta}\right)$	$\begin{bmatrix} -u_\star & -\alpha u_\star \\ -\rho\beta v_\star & -\rho v_\star \end{bmatrix}$	stable node

The stability for the node with coexisting populations follows because

$$\mathrm{Tr}(A) = -(u_\star + \rho v_\star) < 0 \qquad \text{and} \qquad \det(A) = \rho u_\star v_\star (1 - \alpha\beta) > 0.$$

(f) *[2 marks]* When we studied this example in lecture, we had $\gamma_1 = 1$, in which case $(1, 0)$ is the sole stable equilibrium. Reducing γ_1, which corresponds to removing the species that would otherwise win the battle of competitive exclusion, stabilises coexistence.

B5. *Most of my old exams include some question about chemical reaction systems and we study them both in lecture and in problem sets, but this example should be new to the students.*

(a) *[4 marks]* The stoichiometric matrix N has one row for each chemical species and one column for each reaction. The entry N_{ij} is then the number of molecules of species i produced when reaction j occurs: if the reaction consumes species i—if species i appears on the left-hand side of the reaction—this number may be negative. If we arrange the reactants in the order

$$\{D_{\alpha\alpha}, D_{\beta\beta}, D_{\alpha\beta}, A, B\}$$

and the reactions in the order $\{k_1, k_2, k_3, k_{-3}\}$ (where k_j indicates the reaction with rate constant k_j) the stoichiometric matrix is

$$N = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ -2 & 0 & -1 & 1 \\ 0 & -2 & -1 & 1 \end{bmatrix}$$

Of course, other orderings on the reactions and reactants will give rise to permuted forms of N, any one of which will be correct if explained clearly.

(b) *[5 marks]* The desired conserved quantities are

$$2D_{\alpha\alpha} + D_{\alpha\beta} + A \qquad \text{and} \qquad 2D_{\beta\beta} + D_{\alpha\beta} + B.$$

One can find them by performing row-reduction on a copy of N that is augmented at right with a copy of the identity,

$$\left[\begin{array}{cccc|ccccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 1 & 0 & 0 \\ -2 & 0 & -1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & -2 & -1 & 1 & 0 & 0 & 0 & 0 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{cccc|ccccc} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 1 & 0 & 1 \end{array}\right].$$

Some students may simply write down the the quantity $2D_{\alpha\alpha} + D_{\alpha\beta} + A$ and observe that α-monomers are neither created nor destroyed and thus are either free or tied up in dimers and provide a similar argument about the β-monomers. These sorts of arguments are also acceptable.

(c) *[4 marks]* The rank is the number of linearly independent rows, here $r = 3$, and a suitable decomposition is

$$N \;=\; LN_R \;=\; \left[\begin{array}{c} I_r \\ \hline L_0 \end{array}\right] N_R \;=\; \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \hline -2 & 0 & -1 \\ 0 & -2 & -1 \end{array}\right] \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}$$

where N_R consists of three of linearly independent rows from N. Here I've used the first three rows, but students will receive full credit for any answer in which N_R consists of a set of linearly independent rows from N, L has the specified form and $N = LN_R$.

(d) *[3 marks]* The requisite ODEs may be represented conveniently using the stoichiometric matrix from above. If one writes c_α for the concentration of the α-monomer, c_β for the concentration of the β-monomer, $c_{\alpha\alpha}$ for the concentration of $D_{\alpha\alpha}$ with similar expressions for the other two dimers, then

$$\begin{pmatrix} \dot{c}_{\alpha\alpha} \\ \dot{c}_{\beta\beta} \\ \dot{c}_{\alpha\beta} \\ \dot{c}_{\alpha} \\ \dot{c}_{\beta} \end{pmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ -2 & 0 & -1 & 1 \\ 0 & -2 & -1 & 1 \end{bmatrix} \begin{bmatrix} k_1 c_\alpha^2 \\ k_2 c_\beta^2 \\ k_3 c_\alpha c_\beta \\ k_{-3} c_{\alpha\beta} \end{bmatrix},$$

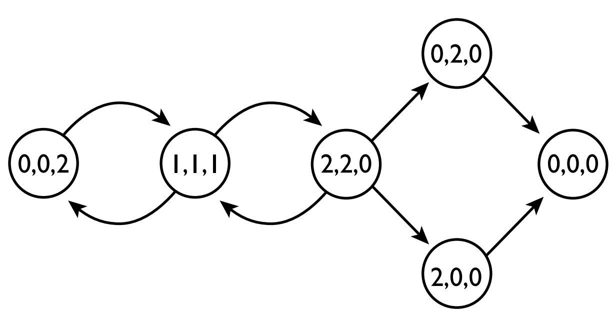
but any equivalent set of ODEs would also receive full marks.

(e) *[4 marks]* The allowed states are

A	B	$D_{\alpha\alpha}$	$D_{\alpha\beta}$	$D_{\beta\beta}$
2	2	0	0	0
1	1	0	1	0
0	0	0	2	0
0	2	1	0	0
2	0	0	0	1
0	0	1	0	1

where the rows correspond to states and the columns to chemical species. The j, k-entry in the table is the number of molecules of species k when the system is in state j.

(f) *[3 marks]* The desired graph is



(g) *[2 marks]* The reactions that form the homodimers $D_{\alpha\alpha}$ and $D_{\beta\beta}$ are irreversible, so the state that includes two homodimers—the rightmost one in the digraph above—is an absorbing state. Thus $\lim_{t\to\infty} p_{000}(t) = 1$ while for all other states $\lim_{t\to\infty} p_{jkl}(t) = 0$.

B6. *I lecture about a few models with spatial structure, but the bulk of the course is about systems of ODEs. The reduction in part (b) should be familiar from the problem sets, as should much of the rest of the analysis, but the example is unseen.*

The students needn't reproduce the PDE in their answers, but it is convenient to have it here. The system under discussion is

$$\partial_t u = D\partial_{xx} u + \mu u(1 - u^3) \tag{6.1}$$

with boundary conditions $u(0, t) = u(L, t) = 0$.

(a) *[3 marks]* If $u(x)$ is a steady-state solution to Eqn. (6.1) it is a function of x alone and hence $\partial_t u = 0$ and the PDE above becomes the ODE

$$0 = D\frac{d^2u}{dx^2} + \mu u(1 - u^3)$$

with the boundary conditions $u(0) = u(L) = 0$.

(b) *[3 marks]* The desired system of ODEs is

$$\frac{du}{dx} = v \qquad \text{and} \qquad \frac{dv}{dx} = \frac{d}{dx}\left(\frac{du}{dx}\right) = \frac{d^2u}{dx^2} = -\left(\frac{\mu}{D}\right) u(1 - u^3). \tag{6.2}$$

(c) *[4 marks]* The null clines are curves on which either $du/dt = 0$ or $dv/dt = 0$. The former are given by the horizontal line $v = 0$ while the latter are the union of the vertical lines $u = 0$ and $u = 1$. Intersections of these curves represent constant solutions $(u(x), v(x)) = (u_\star, v_\star)$ and in this problem the only such solutions are $(u_\star, v_\star) = (0, 0)$ and $(u_\star, v_\star) = (1, 0)$. Only the first of these matches the boundary conditions, but it's a trivial solution.

(d) *[5 marks]* To be a constant of the motion for a system of ODEs like (6.2) a function $H(u, v)$ should satisfy the following PDE

$$\frac{d}{dx}H(u(x),\, v(x)) = \frac{du}{dx}\,\partial_u H + \frac{dv}{dx}\,\partial_v H = 0. \tag{6.3}$$

For the proposed function we thus need to compute

$$\partial_u H = \frac{\mu}{2D}\left[2u - 2u^4\right] = \left(\frac{\mu}{D}\right) u(1 - u^3) \qquad \text{and} \qquad \partial_v H = v.$$

Putting these into the PDE in Eqn. (6.3) produces

$$\begin{aligned}\frac{du}{dx}\,\partial_u H + \frac{dv}{dx}\,\partial_v H &= v\left[\left(\frac{\mu}{D}\right) u(1 - u^3)\right] + \left[-\left(\frac{\mu}{D}\right) u(1 - u^3)\right] v \\ &= 0\end{aligned}$$

which establishes that $H(u, v)$ is indeed a constant of the motion.

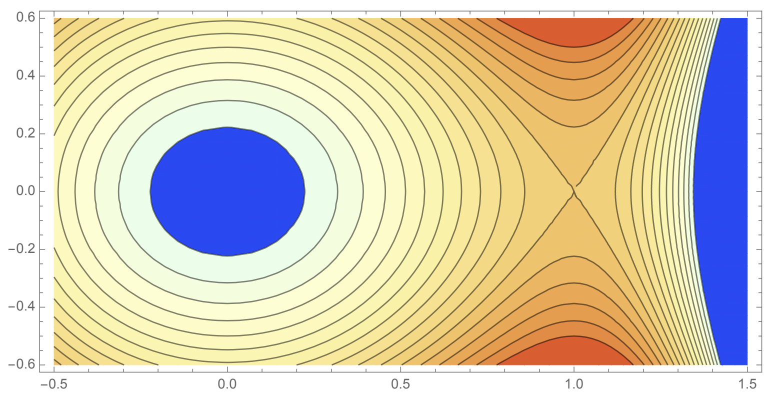
(e) *[6 marks]* The critical points of the function $H(u, v)$ are places where both $\partial_u H$ and $\partial_v H$ vanish. Thus we want

$$\partial_u H = 0 \Rightarrow u(1 - u^3) = 0 \qquad \text{and} \qquad \partial_v H = 0 \Rightarrow v = 0.$$

The only such points are $(0, 0)$ and $(1, 0)$. The Hessain of H is

$$\begin{bmatrix} \partial_{uu} H & \partial_{uv} H \\ \partial_{vu} H & \partial_{vv} H \end{bmatrix} = \begin{bmatrix} 1 - 4u^3 & 0 \\ 0 & 1 \end{bmatrix}.$$

Evaluating this at the critical points leads to the conclusions that $(0, 0)$ is a local minimum while $(1, 0)$ is a saddle. This, in turn, implies that the contour map looks like this.



Students needn't provide the same level of detail for full marks.

(f) *[4 marks]* Solutions to the ODE from part (a) or, equivalently, to the system in part (b), appear on the contour map as curves that begin on the positive v axis (which is the set $u = 0$ with $v = du/dx > 0$) and then arc through the region $u > 0$ before finishing on the negative v axis (which is the set $u = 0$ with $v = du/dx < 0$). There is only a single curve that corresponds to a domain of a given length.

SECTION A

The material in this section is considered 'core' and is essentially all bookwork - upto the specific choice of $f(x)$ in A1, the choice of k and m in A2(b), and the choice of interval for the quadrature scheme in A4 (in class, we did $[-1, 1]$ - so the equations are now slightly more complicated).

A1. The leading term of $T_{n+1}(x)$ is $2^n x^{n+1}$ so for $n = 2$, it is $4x^3$. **1**

The extremal values of $T_3(x)$ in $[-1, 1]$ are attained at

$$x_i = \cos\left(\frac{i\pi}{3}\right), \quad i = 0, 1, 2, 3$$

and we have $T_3(x_0) = T_3(1) = +1$, $T_3(x_2) = T_3(1/2) = -1$, $T_3(x_3) = T_3(-1/2) = +1$ and $T_3(x_3) = T_3(-1) = -1$ (the sign oscillates) **2**

The polynomial $p_n(x)$ of degree $\leq n$ is a best L_∞ approximation to $f \in C[-1, 1]$ if and only if $\exists$ at least $n+2$ points $x_0, x_1, \ldots, x_{n+1}$, with $-1 \leq x_0 < x_1 < \cdots < x_{n+1} \leq 1$, such that

$$|f(x_i) - p_n(x_i)| = \|f - p_n\|_\infty, \qquad i = 0\colon n+1$$

and

$$f(x_i) - p_n(x_i) = -\big(f(x_{i+1}) - p_n(x_{i+1})\big), \qquad i = 0\colon n.$$

3

The best approximation from the set of polynomials of degree $n \leq 2$ is

$$p_2(x) = f(x) - \frac{3}{4}T_3(x) = 1 - 2x^2 + 3x^3 - \frac{3}{4}T_3(x).$$

Since the leading term of $T_3(x)$ is $4x^3$, there is no x^3 term in $p_2(x)$. To see p_2 is a best approximation, note that $f(x) - p_2(x) = \frac{3}{4}T_3(x)$. We know that $\| f - p_2 \|_\infty = \pm\frac{3}{4}$ and this is achieved at the four points x_0, x_1, x_2, x_3 found in part (a). Since the extremal error alternates in sign, the conditions of the Chebyshev Equioscillation Theorem are satisfied by p_2. **4**

A2. (a) We require that $f(x) - r_{km}(x) = O(x^{k+m+1})$, or equivalently that the coefficients in front of the terms $1, x, x^2, \ldots, x^{k+m}$ in the expansion of $f(x)q_{km}(x) - p_{km}(x)$ are zero. **2**

(b) Rearranging, we have

$$r_3(x) = \frac{x}{1 + \frac{x/2}{1+x/6}} = \frac{x + x^2/6}{1 + 2x/3}.$$

2

The Taylor series expansion of $f(x) = \log(1 + x)$ about $x = 0$ is

$$f(x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots$$

We see that $r_3(x)$ is the ratio of a quadratic polynomial $p_{21}(x) = x + x^2/6$ and the linear polynomial $q_{21} = 1 + 2x/3$, with $q_{21}(0) = 1$. We also have

$$\begin{aligned} f(x)q_{21}(x) - p_{21}(x) &= \left(x - \frac{x^2}{2} + \frac{x^3}{3} + O(x^4)\right)(1 + 2x/3) - (x + x^2/6) \\ &= x + 2x^2/3 - x^2/2 - 2x^3/6 + x^3/3 + O(x^4) - x - x^2/6 \\ &= O(x^4). \end{aligned}$$

Hence $r_3(x)$ is the [2/1] Padé approximant, as claimed. **5**

A3. (a) To make the rule exact for polynomials of degree up to $n-1$ we replace $f(x)$ by $p_{n-1}(x)$, the polynomial of degree $n-1$ or less that interpolates $f(x)$ at the nodes. Then, expressing $p_{n-1}(x)$ in its Lagrange interpolating form, we have (with $f_i = p_{n-1}(x_i)$)

$$\int_a^b w(x)p_{n-1}(x)\,dx = \int_a^b w(x)\sum_{i=1}^n f_i l_i(x)\,dx = \sum_{i=1}^n \left(\int_a^b w(x)l_i(x)\,dx\right) f_i,$$

where $l_i(x)$ satisfies $l_i(x_j) = \delta_{ij}$. This shows that we should take

$$w_i = \int_a^b w(x)l_i(x)\,dx.$$

4

(b) To make the rule exact for all polynomials of degree $\le 2n-1$, we choose the nodes to be the roots of the polynomials $\phi_n(x)$ drawn from the family of polynomials that are orthgonal with respect to $w(x)$ on the interval $[a, b]$. **2**

A4. We have

$f(x)$	$I(f)$	$G_2(f)$	
1	1	$w_1 + w_2$	(1)
x	$1/2$	$w_1x_1 + w_2x_2$	(2)
x^2	$1/3$	$w_1x_1^2 + w_2x_2^2$	(3)
x^3	$1/4$	$w_1x_1^3 + w_2x_2^3$	(4)

3

Let $\phi_2(x) = (x - x_1)(x - x_2) \equiv x^2 + ax + b$. Then the linear combination $(3) + a(2) + b(1)$ gives

$$\frac{1}{3} + \frac{a}{2} + b = w_1\Phi(x_1) + w_2\Phi(x_2) = 0$$

Similarly, $(4) + a(3) + b(2)$ gives $\frac{1}{4} + \frac{a}{3} + \frac{b}{2}a = w_1x_1\phi(x_1) + w_2x_2\phi(x_2) = 0$.

Combining these two equations (eliminate b) gives $a = -1$ and hence $b = 1/6$. Thus, $\Phi(x) = x^2 - x + 1/6$ and the roots are:

$$x_1 = \frac{1}{2} + \frac{1}{2}\sqrt{1/3}, \quad x_2 = \frac{1}{2} - \frac{1}{2}\sqrt{1/3}.$$

5

To find the weights, use the first equation to give $w_2 = 1 - w_1$ and then the second equation gives

$$\frac{1}{2} = w_1(x_1 - x_2) + x2 = w_1\sqrt{1/3} + \frac{1}{2} - \frac{1}{2}\sqrt{1/3}$$

Hence, $w_1 = 1/2$ and $w_2 = 1/2$. **2**

[Note - if students repeat the more familiar calculation associated with the interval $[-1, 1]$ and then use an appropriate mapping from $[-1, 1]$ to $[0, 1]$ to obtain the correct nodes then I will accept. I will mark generously - I suspect many students will get in a mess with solving the four equations].

A5. To derive the ℓ-step A–B method, we replace the integrand in

$$y(x_{n+1}) = y(x_n) + \int_{x_n}^{x_{n+1}} f(x, y(x))\, dx,$$

by the polynomial $p(x)$ of degree $\ell - 1$ which interpolates f at the ℓ points $x_n, x_{n-1}, \ldots, x_{n+1-\ell}$. Similarly, to derive the ℓ-step A–M method, we replace the integrand by the polynomial $p(x)$ of degree ℓ which interpolates f at the $\ell + 1$ values $x_{n+1}, x_n, x_{n-1}, \ldots, x_{n+1-\ell}$. **2**

The Adams-Bashforth method is explicit (has $b_0 = 1$) and has order ℓ. The Adams-Moulton method is implicit but has order $\ell + 1$. Hence the A–B methods are easier to implement but less accurate; the A–M methods are harder to implement (as a nonlinear equation needs to be solved for y_{n+1}) but have better accuracy.

3

SECTION B

B6. This is all bookwork. Parts (a),(c) and (d) should be accessible to all students (pass level marks). Strong hints have been given about how to approach part (b) so all students should be able to start - even if they can't complete it.

(a) We have $\| f \|_{2,w} := \left(\int_a^b w(x) f(x)^2 dx \right)^{1/2}$ where $w(x)$ is non-negative and continuous on (a, b) with $\int_a^b w(x)dx > 0$. **2**

We say g is a best $L_{2,w}$ approximation to f from the set G if

$$\| f - g \|_{2,w} \leq \| f - h \|_{2,w}, \qquad \forall h \in G. \qquad \boxed{\mathbf{1}}$$

(b) Since $\phi_0 = 1$, we have

$$\langle \phi_1, \phi_0 \rangle_w = \langle (x - \alpha_0)\phi_0, \phi_0 \rangle_w = \langle x\phi_0, \phi_0 \rangle_w - \alpha_0 \langle \phi_0, \phi_0 \rangle_w = 0$$

(by the definition of α_0). **2**

We use this first result as the base case. If we assume the stated assertion is true for (up to) n, we now need to show that

$$\langle \phi_{n+1}, \phi_i \rangle_w = 0 \qquad \text{for } i = 0, 1, \ldots, n.$$

For $i = n$, we have

$$\langle \phi_{n+1}, \phi_n \rangle_w = \langle x\phi_n, \phi_n \rangle_w - \alpha_n \langle \phi_n, \phi_n \rangle_w - \beta_n \langle \phi_{n-1}, \phi_n \rangle_w = 0,$$

by definition of α_n and the inductive assumption. Similarly, for $i = n - 1$,

$$\langle \phi_{n+1}, \phi_{n-1} \rangle_w = \langle x\phi_n, \phi_{n-1} \rangle_w - \alpha_n \langle \phi_n, \phi_{n-1} \rangle_w - \beta_n \langle \phi_{n-1}, \phi_{n-1} \rangle_w = 0,$$

by definition of β_n and the inductive assumption. **4**

For $i = 0\colon n - 2$, using the three-term recurrence and the inductive assumption gives

$$\begin{aligned} \langle \phi_{n+1}, \phi_i \rangle_w &= \langle x\phi_n, \phi_i \rangle_w - \alpha_n \langle \phi_n, \phi_i \rangle_w - \beta_n \langle \phi_{n-1}, \phi_i \rangle_w \\ &= \langle \phi_n, x\phi_i \rangle_w \\ &= \langle \phi_n, \phi_{i+1} + \alpha_i \phi_i + \beta_i \phi_{i-1} \rangle_w = 0 \end{aligned}$$

Note that if $i = 0$ we should adjust the last line to read $\langle \phi_n, \phi_1 + \alpha_0 \phi_0 \rangle_w = 0$. Hence the result is proved by induction. **4+1**

(c) The best $L_{2,w}$ approximation of this form is found by solving the normal equations $A\mathbf{c} = \mathbf{f}$ (for the coefficients $\mathbf{c} = [c_0, \ldots, c_n]^T$) where

$$A_{i,j} = \langle \phi_i, \phi_j \rangle_w, \quad f_j = \langle f, \phi_j \rangle_w, \qquad i, j = 0, 1, \ldots, n.$$

When the polynomials are orthogonal, the matrix A is diagonal and the system is very easy to solve. **2+2**

(d) (We only met three examples in class - will accept others if they know them). If $w(x) = 1$, we need Legendre polynomials, with $w(x) = 1/\sqrt{1-x^2}$, we need Chebyshev polynomials, and if $w(x) = e^{-x^2}$ we need Hermite polynomials (any two). **2**

B7. This is all bookwork. Part (c) is taken from an examples sheet and parts (a) and (b) are from the lecture notes. (b) and (c) should be accessible to all students.

(a) We need the mean value for sums which says: Let $g(x)$ be continuous on $[a, b]$, let $x_1, \ldots, x_n \in [a, b]$, and let $w_1, \ldots, w_n$ all be of the same sign. Then

$$\sum_{i=1}^{n} w_i g(x_i) = g(\xi) \sum_{i=1}^{n} w_i \qquad \text{for some } \xi \in (a, b).$$

3

For the error (using the given result on one interval), we have

$$\begin{aligned} E_{RT}(f) &= \sum_{i=0}^{n-1} \left[\int_{x_i}^{x_{i+1}} f(x)\,dx - \frac{h}{2}\big(f_i + f_{i+1}\big) \right] = \sum_{i=0}^{n-1} -\frac{h^3}{12} f''(\theta_i), \qquad \theta_i \in (x_i, x_{i+1}), \\ &= -\frac{h^3}{12} f''(\theta) n, \qquad \theta \in (x_0, x_n) = (a, b), \qquad \text{using M.V.T, with } g = f'' \text{ and } w_i = -h^3/12 \\ &= -\frac{h^2(b-a)}{12} f''(\theta) \qquad (\text{since } nh = b - a). \end{aligned}$$

4

(b) Doubling the number of intervals corresponds to reducing h by a factor of 2. So, if we replace h by $h/2$ in the asymptotic formula, we get

$$I(f) - T(h/2) = a_2 \frac{h^2}{4} + a_4 \frac{h^4}{16} + \mathcal{O}(h^6).$$

The new scheme $T(h/2)$ still has an error that is $\mathcal{O}(h^2)$ but the error is reduced by a factor of four. **3**

We have

$$4I(f) - 4T(h/2) = a_2 h^2 + a_4 \frac{h^4}{4} + \mathcal{O}(h^6)$$

and

$$I(f) - T(h) = a_2 h^2 + a_4 h^4 + \mathcal{O}(h^6).$$

Hence,

$$I(f) - \left(\frac{4T(h/2) - T(h)}{3} \right) = -\frac{a_4 h^4}{4} + \mathcal{O}(h^6).$$

Hence the approximation $\frac{4}{3}T(h/2) - \frac{1}{3}T(h)$ gives an error that is $\mathcal{O}(h^4)$. **4**

(c) The Romberg scheme is best described in terms of the following table.

$$\begin{array}{rcccccc} & & \longrightarrow & \text{error} & \times h^2 & & \\ & T_{11} & & & & & \\ & T_{21} & T_{22} & & & & \\ h \text{ halved} \ \downarrow & T_{31} & T_{32} & T_{33} & & & \\ & T_{41} & T_{42} & T_{43} & T_{44} & & \\ & \vdots & \vdots & \vdots & \vdots & \ddots & \end{array}$$

2

Starting with T_{11} (the standard Trapezium rule with $h = 1$ (1 interval)), the first column contains repeated trapezium rules with the step halved each time. The error for this sequence is $O(h^2)$ (see part (a)). Each pair of approximations in the first column can be combined to obtain a new scheme in the second column whose error is $\mathcal{O}(h^4)$ (as in part (b)). These schemes can be combined to give new schemes whose error is $\mathcal{O}(h^6)$ etc. **2**

Starting at T_{11} we double the number of intervals to form T_{21} and then use extrapolation to form T_{22}. We then compute T_{31} and again use extrapolation with T_{21} to form $T_{32}\ldots$ The table is computed row by row: $T_{11}, T_{21}, T_{22}, T_{31}, T_{32}, T_{33}, \ldots$. We can stop when $|T_{kk} - T_{k-1,k-1}| \leq \epsilon$ for the desired tolerance ϵ. **2**

B8. Parts (a) and (b) are bookwork (from the lecture notes) and were set as questions on the 2015 exam. Parts (c) and (d), however, are taken from an examples sheet - but phrased differently. On the examples sheet, the method was presented as a predictor-corrector method. Here, it is presented as an RK method.

(a) Substituting $y(x_n)$ for y_n and subtracting the right-hand side from the left-hand side gives:

$$\tau(h) = y(x_{n+1}) - y(x_n) - hb_1 f(x_n, y(x_n)) - hb_2 f(x_n + c_2 h, y(x_n) + ha_{21} f(x_n, y(x_n))).$$

2

Noting that $f(x_n, y(x_n)) = y'(x_n)$ then gives

$$\tau(h) = y(x_{n+1}) - y(x_n) - hb_1 y'(x_n) - hb_2 f(x_n + c_2 h, y(x_n) + ha_{21} y'(x_n)).$$ **1**

(b) For an order two method, we need that $\tau(h)$ is $\mathcal{O}(h^3)$. Using a Taylor series expansion for $y(x_{n+1})$ and the hint gives

$$\begin{aligned} y(x_{n+1}) &= y(x_n) + hy'(x_n) + \frac{h^2}{2} y''(x_n) + \mathcal{O}(h^3) \\ &= y(x_n) + hy'(x_n) + \frac{h^2}{2} \left(\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} f \right) \Big|_{(x_n, y(x_n))} + \mathcal{O}(h^3). \end{aligned}$$

3

Next, use a Taylor series in two dimensions:

$$f(x_n + c_2h, y(x_n) + ha_{21}y'(x_n)) = y'(x_n) + c_2h\frac{\partial f}{\partial x}(x_n, y(x_n)) \\ + ha_{21}y'(x_n)\frac{\partial f}{\partial y}(x_n, y(x_n)) + \mathcal{O}(h^2).$$

3

Substituting both expansions into the expression for $\tau(h)$ in part (b) gives

$$\tau(h) = \left[y(x_n) + hy'(x_n) + \frac{h^2}{2}\left(\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}f\right)|_{(x_n,y(x_n))} + \mathcal{O}(h^3)\right] - y(x_n) - hb_1y'(x_n) \\ -hb_2\left(y'(x_n) + c_2h\frac{\partial f}{\partial x}(x_n, y(x_n)) + ha_{21}y'(x_n)\frac{\partial f}{\partial y}(x_n, y(x_n)) + \mathcal{O}(h^2)\right).$$

Rearranging then gives

$$\tau(h) = h\,(1 - b_1 - b_2)\,y'(x_n) + h^2\left(\frac{1}{2} - b_2c_2\right)\frac{\partial f}{\partial x}(x_n, y(x_n)) \\ + h^2\left(\frac{1}{2} - b_2a_{12}\right)\left(\frac{\partial f}{\partial y}f\right)|_{(x_n,y(x_n))} + \mathcal{O}(h^3)$$

Hence, we need $b_1 + b_2 = 1$, $b_2a_{21} = 1/2$ and $b_2c_2 = 1/2$. **4**

(c) With the stated coefficients, the method is

$$y_{n+1} = y_n + hf(x_{n+1}, y_n + hf(x_n, y_n)).$$

1

This is equivalent to using the (explicit) Euler method to 'predict' the value of y_{n+1} required by the implicit Euler method.

$$\begin{array}{ll} \text{Predict (Explicit Euler):} & y_{n+1}^{(0)} = y_n + hf_n \\ \text{Evaluate:} & f_{n+1}^{(0)} = f(x_{n+1}, y_{n+1}^{(0)}) \\ \text{Correct (Implicit Euler):} & y_{n+1}^{(1)} = y_n + hf_{n+1}^{(0)} \end{array}$$

2

(d) If $y' = \lambda y$, then applying the method gives

$$y_{n+1} = y_n + h\lambda(y_n + \lambda hy_n) = (1 + h\lambda + (h\lambda)^2)y_n.$$

Thus $|y_{n+1}| < |y_n|$ (the method is absolutely stable) if $|p(\lambda h)| < 1$ where

$$p(\lambda h) = 1 + \lambda h + (\lambda h)^2\,.$$

4

THE UNIVERSITY OF MANCHESTER

Time Series Analysis MATH48032

2016/2017

Solutions

Note for the External examiners for the Actuarial programmes: The examination paper addresses comprehensively the concepts specified in the document *Subject CT6, Statistical Methods, Core Technical Syllabus* (I have the 2010 edition), more specifically part (viii) *Define and apply the main concepts underlying the analysis of time series models.*

The three papers have identical A and B sections. Section C is identical for the MSc and level four papers.

SECTION A
Answer ALL four questions

A1.

A graph of a monthly time series, $\{x_t\}$, is shown below. Give concise answers to the following questions.

a) There is strong seasonality with period 1 year (12 months). There is also a slow trend. [2 marks]

b) Differencing can be expected to remove the trend. (Though a second differencing may be needed if the trend turns out to be of a quadratic type.) [2 marks]

c) Seasonal differencing with $s = 12$ should remove the seasonal trend which seems pretty stable. [2 marks]

d) For example, if the variability is larger for larger values of the series I might consider log-transformation. This seems not to be necessary for this series.
Another reason might be: Log-transformation is often used for financial data where "returns" are of primary interest. [2 marks]

Qu. Total
8 marks

A2. There are various ways to write down the requested forms, I give only the “major” ones.

a)

$$(1 + 0.1\mathbf{B} - 0.72\mathbf{B}^2)X_t = (1 - 0.5\mathbf{B}^{12})\varepsilon_t$$

[2 marks]

b) Taking expectation on both sides of the equation above and using stationarity we get

$$\mu - 0.3\mu = 4.2.$$

Hence, $\mu = 4.2/0.7 = 6$. The required mean-corrected form is

$$\begin{aligned} X_t - \mu &= 0.3(X_{t-1} - \mu) + \varepsilon_t + 0.7\varepsilon_{t-1} && \text{diff. eq. form} \\ (1 - 0.3\mathbf{B})(X_t - \mu) &= (1 + 0.7\mathbf{B})\varepsilon_t && \text{operator form} \end{aligned}$$

[2 marks]

c) The model for $\{Y_t\}$ is

$$(1 - \mathbf{B}^{12})Y_t = (1 - 0.5\mathbf{B})(1 + 0.1\mathbf{B}^{12})\varepsilon_t.$$

[2 marks]

d) $(1 - \mathbf{B})(1 + \mathbf{B} + \mathbf{B}^2 + \mathbf{B}^3)$ and $(1 - \mathbf{B}^2)(1 + \mathbf{B}^2)$. [2 marks]

Qu. Total

8 marks

A3.
A time series is analysed with R and a model is fit with the `arima` function. Here is a summary of the fitted model.

a)
$$(1 - 0.4016\mathbf{B})X_t = \varepsilon_t \qquad \text{or} \qquad X_t - 0.4016X_{t-1} = \varepsilon_t.$$

[1 mark]

b) $\hat{\sigma}^2 = 1.798$. [0.5 marks]

c) It is statistically significant since its standard error is more than 5 time smaller than the estimate. (Other appropriate justifications are acceptable, e.g. *"... since the approximate 95% CI is* $0.4016 \pm 1.96 * 0.0765$ *doesn't contain zero* [1 mark]

d) i) on visual inspection the residuals seem OK although on second thought (after looking at acf) they seem to show some periodicity.
ii) All residual autocorrelations are small except the one at lag 12 (1 year). This suggests that seasonality has not been accounted for completely by the model.
iii) The p-values of the Ljung-Box statistic from lag 12 on show definte lack of fit at the 5% level and are marginally not significant for smaller lags.
So, there is strong evidence for lack of fit. [0.5+1+2 mark

e) Since only the lag 12 residual autocorrelation seems significant, I would fit a model with a seasonal term, $(1 + \theta\mathbf{B}^{12})$, added to the original model. I expect this to be a good model but I will explore its diagnostics as well. [2 marks]

Qu. Total
8 marks

A4.

a) Both plots show that the residuals are not white noise and therefore the fit is not good. The acf decays gradually may be suggesting a low order AR model for the residuals. The pacf clearly indicates an AR(2) model for the residuals since only the first 2 partial autocorrelations seem significantly different from zero. [4 marks]

b) The model originally fitted is ARIMA$(0,1,0)(0,1,1)_{12}$, i.e.

$$(1-\mathbf{B})(1-\mathbf{B}^{12})X_t = (1+\theta B^{12})r_t$$

The above analysis suggest a model $(1-\phi_1\mathbf{B}-\phi_2\mathbf{B}^2)r_t = \varepsilon_t$. Combining the two models gives

$$(1-\mathbf{B})(1-\mathbf{B}^{12})(1-\phi_1\mathbf{B}-\phi_2\mathbf{B}^2)X_tt = (1+\theta B^{12})\varepsilon_t,$$

i.e. an ARIMA$(2,1,0)(0,1,1)_{12}$ model. [2 marks]

c) The standard errors of all parameters in Model 1 are small which coupled with what is said about the residual diagnostics suggests a very good fit.
Model 2 is clearly overfitted—the standard errors of most parameters are (very) large, e.g. the s.e. of ar1 is about 2.5 times the value of the estimated parameter.
The AIC statistic for Model 1 is also smaller.
Model 1 is clear winner. [2 marks]

Qu. Total
8 marks

SECTION B
Answer 2 of the 3 questions

B5.

a)
$$X_t = 0.8X_{t-12} + \varepsilon_t + 0.5\varepsilon_{t-1},$$
(Other correct ways to write this down are equally good.) [2 marks]

b) $s = 12$, $q = 1$, $p_s = 1$, the remaining are zeroes. [2 marks]

c) The root of $1 + 0.5z$ is -2 whose modulus is greater than 1, hence invertible. [2 marks]

d) We have
$$(1 - 0.8z^{12})^{-1} = \sum_{k=0}^{\infty} 0.8^k z^{12k}.$$
Hence,
$$\begin{aligned}(1 - 0.8z^{12})^{-1}(1 + 0.5z) &= (1 - 0.8z^{12})^{-1}(1 + 0.5z)\\ &= \left(\sum_{k=0}^{\infty} 0.8^k z^{12k} + 0.5z\sum_{k=0}^{\infty} 0.8^k z^{12k}\right)\\ &= \left(\sum_{k=0}^{\infty} 0.8^k z^{12k} + \sum_{k=0}^{\infty} 0.5 \times 0.8^k z^{12k+1}\right).\end{aligned}$$
Hence,
$$\begin{aligned}X_t &= (1 - 0.8z^{12})^{-1}(1 + 0.5\mathbf{B})\varepsilon_t\\ &= \sum_{k=0}^{\infty} 0.8^k \mathbf{B}^{12k}\varepsilon_t + \sum_{k=0}^{\infty} 0.5 \times 0.8^k \mathbf{B}^{12k+1}\varepsilon_t.\end{aligned}$$
[4 marks]

e) The coefficients are 1, 0.5, 0.8 at ε_t, ε_{t-1}, ε_{t-12}, respectively. The coefficients at ε_{t-i} for $i = 2, 3, \ldots, 11$ are all zeroes. [4 marks]

f) i) We have
$$\begin{aligned}X_{t+1} &= 0.8X_{t-11} + \varepsilon_{t+1} + 0.5\varepsilon_t\\ X_{t+12} &= 0.8X_t + \varepsilon_{t+12} + 0.5\varepsilon_{t+11}\\ X_{t+13} &= 0.8X_{t+1} + \varepsilon_{t+13} + 0.5\varepsilon_{t+12}\end{aligned}$$
Using the basic properties of predictors or directly the basic forecasting theorem we get
$$\begin{aligned}\hat{X}_{t+1} &= 0.8X_{t-11} + 0.5\varepsilon_t\\ \hat{X}_{t+12} &= 0.8X_t\\ \hat{X}_{t+13} &= 0.8\hat{X}_{t+1}\\ &= 0.8(0.8X_{t-11} + 0.5\varepsilon_t)\\ &= 0.64X_{t-11} + 0.4\varepsilon_t\end{aligned}$$

[4 marks]

ii) Subtracting the predictors from the corresponding predicted values and calculating the variances (alternatively, using the general formula) we obtain $\sigma^2 = 1$ for $h = 1$, $\sigma^2(1 + 0.5^2) = 1.25$ for $h = 12$, and for $h = 12$:

$$\begin{aligned}\operatorname{Var}(X_{t+13} - \hat{X}_{t+13}) &= \sigma^2(1^2 + 0.5^2 + 0.8^2)\\ &= 1 + 0.25 + 0^2 + \cdots + 0^2 + 0.64\\ &= 1.89\end{aligned}$$

[6 marks]

Qu. Total
24 marks

B6. Let $\{X_t\}$ be a stationary time series with $\mathrm{E}\, X_t = 0$.

a) For any integer $j \geq 1$ consider the best linear predictor of X_t in terms of $X_{t-1}, \ldots, X_{t-j}$. Denote by v_j the variance of the corresponding prediction error, and by $\phi_1^{(j)}, \ldots, \phi_j^{(j)}$ the coefficients at $X_{t-1}, \ldots, X_{t-j}$, respectively.

i) Bookwork.

ii) Bookwork.

iii) partial autocorrelations. [8+2+2 marks]

b) i) $X_t = 0.8X_{t-1} + \varepsilon_t$ — No, $\beta_2 = 0$. [1 mark]

ii) $X_t = -0.8X_{t-1} + \varepsilon_t$ — No, $\beta_2 = 0$. [1 mark]

iii) $X_t = \varepsilon_t + 0.8\varepsilon_{t-1}$ — No, $\beta_1 = \rho_1 \neq 0.8$. [2 marks]

iv) $X_t = \varepsilon_t - 0.8\varepsilon_{t-1}$ — No, $\beta_1 = \rho_1 \neq 0.8$. [2 marks]

v) $X_t = X_{t-1} - 0.5X_{t-2} + \varepsilon_t$ — No, $\beta_2 \neq 0.5$. [2 marks]

vi) $X_t = X_{t-1} - 0.25X_{t-2} + \varepsilon_t$ — Yes.
Indeed, $\beta_2 = -.25$. From the Yule-Walker equations for $j = 2$ we get

$$\begin{aligned}\rho_1 &= 1 - 0.25\rho_1 \\ \rho_2 &= \rho_1 - 0.25.\end{aligned}$$

We do not even need the second one since from the 1st equation we get $\rho_1 = 1/1.25 = 0.8$. Since $\beta_1 = \rho_1$, we get $\beta_1 = \rho_1 = 0.8$, as required for affirmative answer. [2 marks]

vii) random walk — No, not stationary. [2 marks]

Qu. Total
24 marks

B7. To answer this question a students needs to understand the definition of an ARMA model (the definition a gave postulates causality).

a) In operator form the model is

$$(1 - \mathbf{B} + \tfrac{1}{2}\mathbf{B}^2)X_t = (1 + \tfrac{1}{2}\mathbf{B})\varepsilon_t, \quad (1)$$

The root of the polynomial $1 + \frac{1}{2}z$ is -2, outside the unit circle, hence the process is invertible. [4 marks]

b) Multiply equation (1) by X_{t-k} and take expected values. Using causality, for $k \geq 2$ we get

$$\gamma_k - \gamma_{k-1} + \tfrac{1}{2}\gamma_{k-2} = 0.$$

Let $\sigma^2 = \mathrm{Var}(\varepsilon_t)$. Multiplying equation (1) by ε_t and taking expected values we get $R_{X\varepsilon}(0) = \sigma^2$.

Using this and multiplying (as above) by X_{t-k} but with $k = 0, 1$ we get

$$\begin{aligned} \gamma_0 &= \gamma_1 - \tfrac{1}{2}\gamma_2 + \sigma^2 + \tfrac{1}{2}R_{X\varepsilon}(1) \\ \tfrac{3}{2}\gamma_1 &= \gamma_0 + \tfrac{1}{2}\sigma^2. \end{aligned}$$

So, the system is

$$\begin{aligned} \gamma_0 = \gamma_1 - \tfrac{1}{2}\gamma_2 + 1 + \tfrac{1}{2}R_{X\varepsilon}(1) \\ \tfrac{3}{2}\gamma_1 = \gamma_0 + \tfrac{1}{2} \\ \gamma_1 - \gamma_0 + \tfrac{1}{2}\gamma_1 = 0 \\ \gamma_2 - \gamma_1 + \tfrac{1}{2}\gamma_0 = 0. \end{aligned}$$

[8 marks]

c) For $k \geq 2$ we get

$$R_{X\varepsilon}(k) - R_{X\varepsilon}(k-1) + \tfrac{1}{2}R_{X\varepsilon}(k-2) = 0.$$

So, for $k \geq 2$, $R_{X\varepsilon}(k)$ may be computed recursively by the formula

$$R_{X\varepsilon}(k) = R_{X\varepsilon}(k-1) - \tfrac{1}{2}R_{X\varepsilon}(k-2),$$

using $R_{X\varepsilon}(0) = \sigma^2 = 1$ and $R_{X\varepsilon}(1) = 3/2$ as initial values. [5 marks]

Causality implies that $R_{X\varepsilon}(k) = 0$ for negative k. [1 mark]

Plug the given values in the expressions above to get $R_{X\varepsilon}(0)$ and $R_{X\varepsilon}(1)$. To get $R_{X\varepsilon}(2)$, multiply equation (1) by ε_{t-2} and take expected value to get

$$R_{X\varepsilon}(2) = R_{X\varepsilon}(1) - \tfrac{1}{2}R_{X\varepsilon}(0) = R_{X\varepsilon}(1) - \tfrac{1}{2}\sigma^2,$$

plug in the previously obtained values to get the final result.

d) The fitted model is

$$X_t = 4.3474 + 0.9272X_{t-1} - 0.3879X_{t-2} + \varepsilon_t + 0.6286\varepsilon_{t-1},$$

The mean corrected form is:

$$\begin{aligned} X_t &= \mu + 0.9272(X_{t-1} - \mu) - 0.3879(X_{t-2} - \mu) + \varepsilon_t + 0.6286\varepsilon_{t-1} \\ &= \mu - 0.9272\mu + 0.3879\mu + 0.9272X_{t-1} - 0.3879X_{t-2} + \varepsilon_t + 0.6286\varepsilon_{t-1}. \end{aligned}$$

Hence, $4.3474 = \mu - 0.9272\mu + 0.3879\mu$ and so $\mu = 4.3474/(1 - 0.9272 + 0.3879) = 9.43$

[2 marks]

Qu. Total
24 marks

SECTION C
Answer ALL questions

C8. g

a) i) No, the mean of the returns is not zero. The 95% confidence interval is $[0.00039, 0.0055]$.

ii) The test statistic is $t = -0.119/\sqrt{6/1945} = -2.14$ with a p value 0.032. Thus, the log returns are skewed to the left.

iii) The test statistic is $t = 2.705/\sqrt{24/1945} = 24.35$, which is highly significant. Thus, the log returns have heavy tails.

iv) No, the log returns have no serial correlations. The Ljung-Box statistics give $Q(10) = 12.96$ with a p value 0.23.

v) Yes, because the Ljung-Box statistics of the squared deviations from the mean give $Q(10) = 338.96$ with a p value close to zero.

[5*2 marks]

b) Yes, the model is adequate because the model checking statistics all have high p values. The fitted model is

$$\begin{aligned} r_t &= 0.0017 + a_t \\ a_t &= \sigma_t \epsilon_t, \qquad \epsilon_t \sim t^*_{6.43} \\ \sigma_t^2 &= 0.00012 + 0.0896 a_{t-1}^2 + 0.875 \sigma_{t-1}^2 \end{aligned}$$

where t^*_ν denotes standardized Student-t distribution with ν degrees of freedom. [3 marks]

c) i) Yes, the model is adequate as all model checking statistics have high p values. The fitted model is

$$\begin{aligned} r_t &= 0.0015 + a_t \\ a_t &= \sigma_t \epsilon_t, \qquad \epsilon_t \sim t^*_{6.41,0.983} \\ \sigma_t^2 &= 0.00012 + 0.0892 a_{t-1}^2 + 0.876 \sigma_{t-1}^2 \end{aligned}$$

where $t^*_{\nu,\xi}$ denotes standardized Student-t distribution with ν degrees of freedom and skew parameter ξ. [3 marks]

ii) The test statistic is $t = (0.983 - 1)/0.0295 = -0.58$ with a p value 0.56. The null hypothesis of $\xi = 1$ cannot be rejected. [2 marks]

d) The ARMA(0,0)+GARCH(1,1) model with Student-t innovations. The model has a smaller AIC criterion -3.0594, whereas that of the skew innovation is -3.0586. This is consistent with the testing result for ξ. [2 marks]

Qu. Total

20 marks

C9.

a) We have

$$X_{t+1} = (1+\phi)X_t - \phi X_{t-1} + \varepsilon_{t+1}$$
$$X_{t+h} = (1+\phi)X_{t+h-1} - \phi X_{t+h-2} + \varepsilon_{t+h}$$

i) Setting $t = T = 100$ and using the observed values $x_{100} = 3$ and $x_{99} = 1$ we get

$$\hat{X}_{T+1|T,T-1,\ldots} = (1+\phi)x_T - \phi x_{T-1} = x_T + (x_T - x_{T-1})\phi = 3 + (3-1)\phi$$
$$= 3 + 2\phi$$

[1 mark]

ii) Similarly, setting $h = 2$, $t = T + h = 102$ and using the one-step predictor found above, we get

$$\hat{X}_{T+2|T,T-1,\ldots} = (1+\phi)\hat{X}_{T+1|T,T-1,\ldots} - \phi\hat{X}_{T|T,T-1,\ldots} = (1+\phi)(3+2\phi) - 3\phi$$
$$= 3 + 2\phi + 2\phi^2$$

[2 marks]

iii) $f(h) = (1+\phi)f(h-1) - \phi f(h-2)$. [2 marks]

iv) The solution of this equation is of the form $f(h) = a + b\phi^h$, where a and b can be determined by equating the formulae for $f(1)$ and $f(2)$ (the 1-step and 2-step predictions found above) to $a + b\phi^1$ and $a + b\phi^2$, respectively,

$$a + b\phi = 3 + 2\phi$$
$$a + b\phi^2 = 3 + 2\phi + 2\phi^2$$

and solving the resulting system of equations. For example, substracting the first equation from the second we get

$$b\phi(\phi - 1) = 2\phi^2.$$

Hence,

$$b = \frac{2\phi^2}{\phi(\phi-1)} = \frac{2\phi}{\phi-1}$$

and

$$a = 3 + 2\phi - b\phi = 3 + (2 - \frac{2\phi}{\phi-1})\phi = 3 - \frac{2\phi}{\phi-1}$$

So,

$$f(h) = 3 - \frac{2\phi}{\phi-1} + \frac{2\phi}{\phi-1}\phi^h$$

[4 marks]

v) $3 - \frac{2\phi}{\phi-1}$ [1 mark]

b) i) Yes, the augmented Dickey-Fuller tests fail to reject the null hypothesis of unit-root time series. [4 marks]

ii) Yes,the co-integration test shows that the two unemployment rate series are co-integrated.
A co-integrating vector is $(1, -18.377)$ and a cointegrating relation is $DM - 18.377CD$. [6 marks]

Qu. Total
20 marks

END OF EXAMINATION PAPER

MATH3/4/68052 Solutions

A1 (a) The mean and variance of the distribution are given by [Bookwork]

$$\mathrm{E}[Y] = b'(\theta), \quad Var\{Y\} = b''(\theta)a(\phi).$$ [2]

(b) For the geometric distribution,

$$f(y) = (1-p)^y p = \exp\left(\frac{y\log(1-p) - \log\frac{1}{p}}{1} + 0\right) \in \text{exponential family}$$

with $\theta = \log(1-p)$, $\phi = 1$, $a(\phi) = \phi$, $b(\theta) = -\log p = -\log(1-e^\theta)$, $c(y,\phi) = 0$. [4]

(c) Using the formulas in part (a),

$$\mathrm{E}[Y] = b'(\theta) = -\tfrac{1}{1-e^\theta}(-e^\theta) = \tfrac{e^\theta}{1-e^\theta} = \tfrac{1-p}{p},$$
$$Var\{Y\} = b''(\theta)a(\phi) = \left(\tfrac{e^\theta}{1-e^\theta}\right)' \times 1 = \tfrac{e^\theta(1-e^\theta)-e^\theta(-e^\theta)}{(1-e^\theta)^2} = \tfrac{e^\theta}{(1-e^\theta)^2} = \tfrac{1-p}{p^2}.$$ [2]

(d) The three components of a generalised linear model are [Bookwork]

(i) the error distribution, which is the distribution of the response variable;

(ii) the linear predictor, which is a linear combination of the explanatory variables, and

(iii) the link function, which is a function of the mean response.

The canonical link is the same function of μ as θ is (a function of μ). [4]

(e) For the geometric distribution, $\mu = \frac{1-p}{p}$, thus $p = \frac{1}{1+\mu}$ and $1-p = \frac{\mu}{1+\mu}$. Then

$$\theta = \log(1-p) = \log\left(\tfrac{\mu}{1+\mu}\right).$$

Therefore $g(\mu) = \log\left(\frac{\mu}{1+\mu}\right)$ is the canonical link. [2]

(f) The log-likelihood function is given by

$$\ell(\underset{\sim}{\beta}, \phi) = \sum_{i=1}^{n}\left(\frac{y_i\theta_i - b(\theta_i)}{a(\phi)} + c(y_i,\phi)\right).$$

Thus (using the chain rule) the Fisher scores are

$$\frac{\partial \ell}{\partial \beta_j} = \frac{1}{a(\phi)}\sum_{i=1}^{n}(y_i - b'(\theta_i))\frac{\partial \theta_i}{\partial \beta_j},\ j = 1,\ldots,p.$$

When the canonical link is used, $g(\mu_i) = \theta_i = \underset{\sim}{x}_i^T\underset{\sim}{\beta}$, so $\frac{\partial\theta_i}{\partial\beta_j} = \frac{\partial}{\partial\beta_j}\underset{\sim}{x}_i^T\underset{\sim}{\beta} = x_{ij}$.

From Property 1, $b'(\theta_i) = \mu_i$. Therefore

$$\frac{\partial \ell}{\partial \beta_j} = \frac{1}{a(\phi)}\sum_{i=1}^{n}(y_i - \mu_i)x_{ij},\ j = 1,\ldots,p.$$

[4]

A2 (a) Binomial response probit regression model with intercept:

$$y_i = n_i\pi_i + \varepsilon_i \sim B(n_i,\ \pi_i) \text{ independent,}$$

$$\Phi^{-1}(\pi_i) = \beta_0 + \beta_1 x_i, \quad i = 1, \dots, n,$$

where Φ^{-1} is the inverse CDF or quantile function of N(0,1). [3]

(b) (i) W is a diagonal matrix. The diagonal elements are

$$w_i = \frac{1}{V(\pi_i)g'(\mu_i)^2}, i = 1, \dots, n,$$

where $V(\pi_i) = \text{var}(y_i)/a(\phi) = n_i\pi_i(1-\pi_i)/1 = n_i\pi_i(1-\pi_i)$,

$$g'(\mu_i) = \left(\frac{\mathrm{d}}{\mathrm{d}\pi_i}\log\frac{\pi_i}{1-\pi_i}\right)\cdot\frac{\mathrm{d}\pi_i}{\mathrm{d}\mu_i} = \frac{1-\pi_i}{\pi_i}\cdot\frac{1-\pi_i-\pi_i\times(-1)}{(1-\pi_i)^2}\cdot\frac{1}{n_i} = \frac{1}{n_i\pi_i(1-\pi_i)}.$$

Therefore $w_i = n_i\pi_i(1-\pi_i),\ i = 1, \dots, n.$ [3]

(ii) The Fisher information matrix is

$$\begin{aligned} I(\underset{\sim}{\beta}) &= X^T W X \\ &= \begin{pmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_n \end{pmatrix}\begin{pmatrix} w_1 & & 0 \\ & \ddots & \\ 0 & & w_n \end{pmatrix}\begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \\ &= \sum_{i=1}^n w_i \begin{pmatrix} 1 \\ x_i \end{pmatrix}(1 \ \ x_i) \\ &= \sum_{i=1}^n n_i\pi_i(1-\pi_i)\begin{pmatrix} 1 & x_i \\ x_i & x_i^2 \end{pmatrix}. \end{aligned}$$

[4]

(iii) $I(\underset{\sim}{\beta})$ can be used to find the standard errors of $\hat{\underset{\sim}{\beta}}$. It is done by inverting $I(\hat{\underset{\sim}{\beta}})$ and taking square roots of the diagonal elements. [3]

(c) (i) The mean vector of $\underset{\sim}{\xi}$ is $X\underset{\sim}{\beta}$ and the variance-covariance matrix is W^{-1}. [3]

(ii) The updating equation is

$$\underset{\sim}{\beta}^{(k+1)} = (X'W^{(k)}X)^{-1}X'W^{(k)}\underset{\sim}{\xi}^{(k)},\ k = 0, 1, 2, \dots,$$

where $\underset{\sim}{\xi}^{(k)}$ and $W^{(k)}$ are calculated using $\underset{\sim}{\beta}^{(k)}$.

The role of $\underset{\sim}{\xi}$ is to be used as working response in a weighted least squares regression. [4]

B1 (a) Putting the Gamma density in exponential family form,

$$\begin{aligned} f(y;\ \alpha,\lambda) &= \exp\{-\lambda y + \alpha\log\lambda + (\alpha-1)\log y - \log\Gamma(\alpha)\} \\ &= \exp\left\{\frac{y(-\lambda/\alpha)+\log(\lambda/\alpha)}{1/\alpha} + \alpha\log\alpha + (\alpha-1)\log y - \log\Gamma(\alpha)\right\}. \end{aligned}$$

The natural parameter is $\theta = -\frac{\lambda}{\alpha}$. [3]

(b) From $b(\theta) = -\log(\lambda/\alpha) = -\log(-\theta)$, we have $\mu = b'(\theta) = -\frac{1}{-\theta} \times (-1) = -1/\theta = \alpha/\lambda$.

The variance function is $V(\mu) = b^{''}(\theta) = (-1/\theta)' = 1/\theta^2 = \mu^2$. [4]

(c) (i) From (b), $\theta = -1/\mu$. Thus the canonical link is $g(\mu) = -1/\mu$. The reciprocal link is just as good as (or equivalent to) the canonical link. [2]

(ii) The denominator $a(\phi) = 1/\alpha$ is estimated by the deviance divided by its df=35-3, which gives $1/\hat{\alpha} = 44.28/32$ and thus $\hat{\alpha} = 32/44.28 = 0.7227$. [4]

(iii) No, the deviance cannot be used to check the adequacy of the model. This is because its distribution depends on the unknown parameter α. [2]

(iv) When the x^2 term is removed, the change in deviance is $48.72 - 44.28 = 4.44$ on 1 df. The model with x^2 in it has deviance 44.28 on 32 df. The F-ratio is

$$\frac{4.44/1}{44.28/32} = 3.21$$

which is less than the critical value $F_{0.05;\ 1,32} > 4.17$ (from table). Thus we do not reject $H_0:\ \beta_2 = 0$ at the 5% level, and the model without x^2 is not significantly worse. [5]

B2 (a) The model fit1 is not saturated, because its deviance 19.29 is not zero. [2]

The saturated model has both Insecticide and Ldosage as factors:

$y_{ij} \sim B(n, \pi_{ij})$ independent,

$\log \frac{\pi_{ij}}{1-\pi_{ij}} = \mu + \alpha_i + \beta_j + \gamma_{ij},\ i = 1, 2, 3, j = 1, 2, \ldots, 6,$

with constraints $\alpha_1 = \beta_1 = \sum_j \gamma_{1j} = \sum_i \gamma_{i1} = 0.$ [4]

(b) Fit1 has deviance 19.29 on 12 df, which is less than the 5% upper tail critical value $\chi^2_{0.05;\,12} = 21.03$. This is not significant evidence at 5%. Thus the model is adequate as far as deviance is concerned. [4]

(c) When fit1 becomes fit2, the change in deviance is 3.39 on 2 df, which is not significant at 5% (P-value$= 0.18 > 0.05$). Yes, `fit1` can be simplified without causing a sig. change in deviance. [4]

(d) Calculating fitted probabilities of being killed by a 3mg dosage:

$$A:\ \eta = -4.4541 + 2.6938 \times 1.0986 = -1.4947,\quad \pi = \frac{1}{1 + e^{1.4947}} = 0.1832$$
$$B:\ \eta = -1.4947 + 0.6144 = -0.8803,\quad \pi = \frac{1}{1 + e^{0.8803}} = 0.2931$$
$$C:\ \eta = -1.4947 + 3.0314 = 1.5367,\quad \pi = \frac{1}{1 + e^{-1.5367}} = 0.8230$$

[3]

(e) Calculating LD50:

$$A:\ \text{LD50} = \exp\left(\frac{4.4541}{2.6938}\right) = 5.23 \text{ mg}$$
$$B:\ \text{LD50} = \exp\left(\frac{4.4541 - 0.6144}{2.6938}\right) = 4.16 \text{ mg}$$
$$C:\ \text{LD50} = \exp\left(\frac{4.4541 - 3.0314}{2.6938}\right) = 1.70 \text{ mg}$$

[3]

B3 (a) The multinomial distribution is given by the probability function [Bookwork]

$$P(Y_{11} = y_{11}, \ldots, Y_{IJ} = y_{IJ}) = \frac{n!}{y_{11}! \cdots y_{IJ}!} \pi_{11}^{y_{11}} \cdots \pi_{IJ}^{y_{IJ}},$$

where π_{ij} are cell probabilities with $\sum_{ij} \pi_{ij} = 1$.

[3]

(b) For independent $Y_{ij} \sim Pois(n\pi_{ij})$, $Y_{..} = Y_{11} + \cdots + Y_{IJ} \sim Pois(n)$.

$$\begin{aligned} P(Y_{11} = y_{11}, \ldots, Y_{IJ} = y_{IJ} | Y_{..} = n) &= \frac{P(Y_{11} = y_{11}, \ldots, Y_{IJ} = y_{IJ})}{P(Y_{..} = n)} \\ &= \frac{\prod_{ij} (n\pi_{ij})^{y_{ij}} \exp(-n\pi_{ij})/y_{ij}!}{n^n \exp(-n)/n!} \\ &= \frac{n!}{\prod_{ij} y_{ij}!} \prod_{ij} \pi_{ij}^{y_{ij}}, \end{aligned}$$

when $y_{11} + \cdots + y_{IJ} = n$. Thus difference between the Poisson log-likelihood and the multinomial is $\log(n^n \exp(-n)/n!)$, which is a constant.

[5]

(c) (i) The additive model has deviance 10.205 on 4 df, which is significant at the 5% sig. level as it is greater than $\chi^2_{0.05;\,4} = 9.488$. Thus significance evidence to reject independence between vehicle type and mechanical condition. [4]

(ii) Mechanical condition cannot be removed from the model because of its significant interaction with vehicle type. One can also say that because of the lack of fit of the additive model – it cannot be simplified further. [4]

(iii) When vehicle age is added, the deviance becomes 5.384 on 3 df. This is not significant at 5% since it is less than $\chi^2_{0.05;\,3} = 7.81$. Thus sufficient improvement for model to provide adequate fit. [4]

A3 (MATH4/68052 only)

(a) Definitions: [Bookwork]

$S(t) = P(T > t),\ t > 0.$

$h(t) = \lim_{\delta \to 0^+} \frac{P(T \leq t+\delta | T > t)}{\delta},\ t > 0.$

Calculation of $h(t)$ from $S(t)$: $h(t) = -\frac{S'(t)}{S(t)}$.

Calculation of $S(t)$ from $h(t)$: $S(t) = \exp\left(-\int_0^t h(t)\mathrm{d}t\right)$. [4]

(b) (i) $f(t) = 3t^2 e^{-t^3},\ t > 0.$

$F(t) = \int_0^t 3t^2 e^{-t^3}\mathrm{d}t = -e^{-t^3}|_0^t = 1 - e^{-t^3},\ t > 0.$

$S(t) = 1 - F(t) = e^{-t^3},\ t > 0.$ [2]

(ii) $h(t) = f(t)/S(t) = 3t^2$, $t > 0$.

Quadratic curve when plotted against t. [3]

(c) A semi-parametric model for the hazard function is

$$h(t) = h_0(t)e^{\beta x},$$

where $h_0(t)$ is an unspecified hazard function and β is a constant. The nonparametric part $h_0(t)$ represents the baseline corresponding to $x = 0$. The implication of this model is proportional hazards, i.e. $h(t;\ x)/h(t;\ x') = e^{\beta(x-x')}$ not dependent on t. [3+2+3]

(d) Let $t_1 < t_2 < \ldots < t_n$ be observed lifetimes, and i_j be the subject who failed at t_j. [Bookwork]

The partial likelihood is defined as

$$\begin{aligned} & P(I_1 = i_1, \ldots, I_n = i_n;\ x_1, \ldots, x_n) \\ = \ & P(I_1 = i_1)P(I_2 = i_2 | I_1 = i_1)P(I_3 = i_3 | I_2 = i_2, I_1 = i_1) \cdots P(I_n = i_n | I_{n-1} = i_{n-1}, \ldots, I_1 = i_1) \end{aligned}$$

where the conditional probability at t_j is

$$P(I_j = i_j | I_{j-1} = i_{j-1}, \ldots, I_1 = i_1) = \frac{\exp(x'_{i_j}\beta)}{\sum_{k \geq j} \exp(x'_{i_k}\beta)}$$

[3]

C1 (MATH4/68052 only)

(a) Calculating Kaplan-Meier estimate of the survival function:

At $t = 5$, $r = 10$, $d = 1$, $\hat{S}(t) = 1 - \frac{1}{10} = 0.9$

At $t = 7$, $r = 9$, $d = 1$, $\hat{S}(t) = 0.9 \times (1 - \frac{1}{9}) = 0.8$

At $t = 8$, $r = 8$, $d = 1$, $\hat{S}(t) = 0.8 \times (1 - \frac{1}{8}) = 0.7$

At $t = 11$, $r = 5$, $d = 1$, $\hat{S}(t) = 0.7 \times (1 - \frac{1}{5}) = 0.56$

At $t = 15$, $r = 2$, $d = 1$, $\hat{S}(t) = 0.56 \times (1 - \frac{1}{2}) = 0.28$

The estimated survival function is

$$\hat{S}(t) = \begin{cases} 1, & 0 < t < 5 \\ 0.9, & 5 \le t < 7 \\ 0.8, & 7 \le t < 8 \\ 0.7, & 8 \le t < 11 \\ 0.56, & 11 \le t < 15 \\ 0.28, & 15 \le t \le 16 \\ 0, & 16 < t \end{cases}$$ [5]

(b) 95% c. i. for $\log(-\log S(10))$:

$$\log(-\log(0.7)) \pm 1.96 \left(\tfrac{1}{10\times 9} + \tfrac{1}{9\times 8} + \tfrac{1}{8\times 7}\right)^{1/2} / |\log(0.7)|$$

$$= -1.0309 \pm 1.96 \times 0.2070/0.3567 = -1.0309 \pm 1.96 \times 0.5804 = (-2.1685,\ 0.1067).$$

95% c. i. for $S(10)$: $0.7^{e^{\pm 1.96\times 0.5848}}$ or

$$(\exp(-\exp(0.1067)),\ \exp(-\exp(-2.1685))) = (0.3287, 0.8919).$$ [5]

(c) The estimated mean survival time is

$$1 \times 5 + 0.9 \times 2 + 0.8 \times 1 + 0.7 \times 3 + 0.56 \times 4 + 0.28 \times 1 = 12.22.$$ [4]

The estimated median survival time is 15 using the minimum t such that $\hat{S}(t) < 0.5$. [2]

(d) Nelson-Aalen estimate of $H(10)$:

$$\frac{1}{10} + \frac{1}{9} + \frac{1}{8} = 0.3361,$$ [2]

and a 95% confidence interval for $H(10)$ is

$$(-\log(0.8919), -\log(0.3287)) = (e^{-2.1685},\ e^{0.1067}) = (0.1143, 1.1126)$$ [2]

using results in (b).

C2 (MATH4/68052 only)

(a) The log rank test statistic takes the value $\chi^2 = 1.1$ on 1 df which is not significant at 10% as P-value= $0.303 > 0.1$. Thus no significant difference between treatments A and B. [4]

(b) (i) The fitted model is Cox proportional hazards:

$$\hat{h}(t) = h_0(t)\exp(0.147 \times \text{age} - 0.804 \times rx2),$$

where $rx2$ is an indicator for treatment B. [3]

(ii) Question is whether treatment makes a difference taking into account year when diagnosed. Answer is No, as P-value corresponding to test statistic for rx2 is $0.2 > 0.05$. [3]

(iii) Hazard ratio is $e^{0.147\times 5} = 2.09$ if someone is diagnosed 5 years later. [4]

(c) From the summary of survreg,

$\hat{\lambda}_0 = e^{-11.0299} = 1.6210 \times 10^{-5}, \quad \hat{\alpha} = 1/0.551 = 1.8149.$

The fitted Weibull model is a proportional hazards model with coefficients

age: $-0.0791 \times (-1.8149) = 0.1436$

rx2: $0.5673 \times (-1.8149) = -1.0296$

that are similar to those in (b). Additionally it gives $h_0(t)$ in parametric form:

$\hat{h}_0(t) = 1.8149 \times e^{-11.0299\times 1.8149} \times t^{0.8149}, \ t > 0.$

[4]

(d) Testing H_0: Scale=1 vs H_1: Scale $\neq 1$,

Log(Scale)= -0.5967, se= 0.2352, $z = -0.5967/0.2352 = -2.54$ is significant at 5% (P-value= $0.0112 < 0.05$). Cannot reduce to exponential model. [2]

see similar

A1 $\Pi = 2C_{20} - 3C_{30} + C_{50}$ ✓

$\pi(t=T) = 2\max(S-20,0)$

$-3\max(S-30,0)$

when $t=T$ $+\max(S-50,0)$ 3

If $0 < S < 20$

$\pi = 0$

$P_{50} = \cancel{\ldots}\ 50 - S > \pi$ 3

If $20 < S < 30$

$s=20$ $s=30$

$\pi = 2S - 40$ $(0 < \pi < 20)$

$P_{50} = 50 - S$ $(20 < \pi < 30)$

$s=30$ $s=20$

$\Rightarrow P_{50} \geq \pi$ 3

If $30 < S < 50$

$P_{50} = 50 - S$, $(20 < P_{50} < 0$

$\pi = -S + 50$, $(20 > \pi > 0$

$\Rightarrow P_{50} = \pi$

3

If $S > 50$

$$P_{50} = 0$$

$$\pi = 0$$

$$\Rightarrow P_{50} = \pi$$

(2)

3

15

1

A~

(i) no arbitrage (seen in class)

$F = Se^{rT}$

$= 10 \cdot e^{+0.75 \times .05}$

$= \$10.382$ 5

(ii) After 6 month

$25 \cdot e^{0.5 \times .025} = 25.031$ 5

Then

$25.031 \cdot e^{0.5 \times .04} = 25.5366$

– extension of (i) – no arbitrage

(iii) $F = 31500 \cdot e^{0.5 \times (0.1 + 0.04)}$

$= 33784$

storage not seen before

– unseen (no arbitrage)

5

15

See Smith

A3 (i) $\Pi = 2S - P(S,X)$

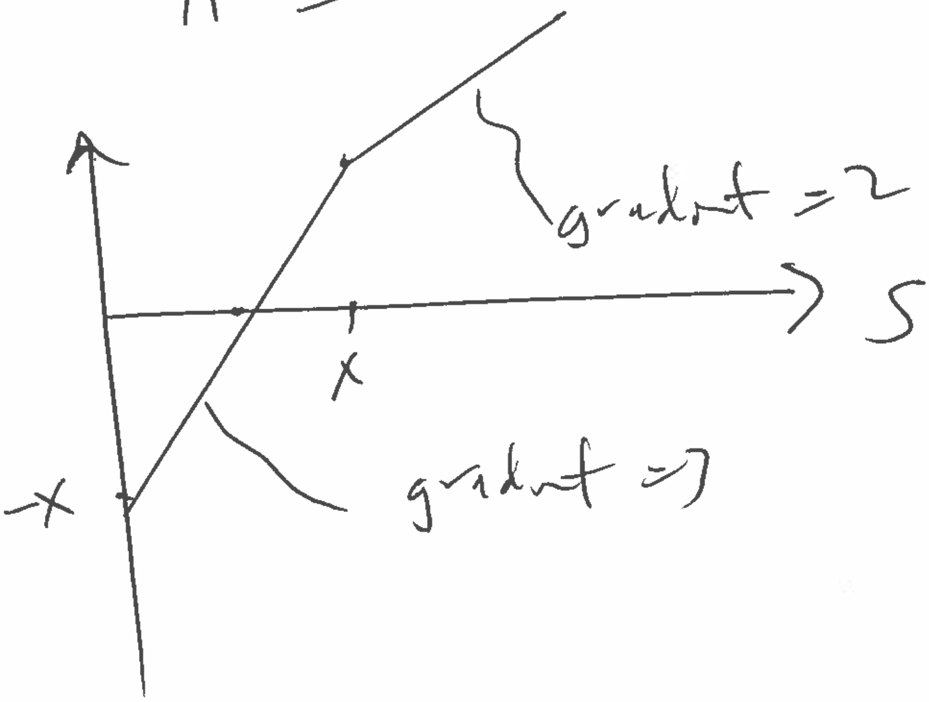
$\Pi(t=T) = 2S - \max(X-S, 0)$

If $0 < S < X$

$\Pi = 2S - X + S = 3S - X$

If $S > X$

$\Pi = 2S$



(ii) $\Pi = S - 2p(X_1) + c(X_2)$ 2

$\Pi(t=T) = S - 2\max(X_1 - S, 0)$

$+ \max(S - X_2, 0)$

If $X_1 > X_2$

If $0 < S < X_2$

$\Pi = S - 2(X_1 - S) = 3S - 2X_1$

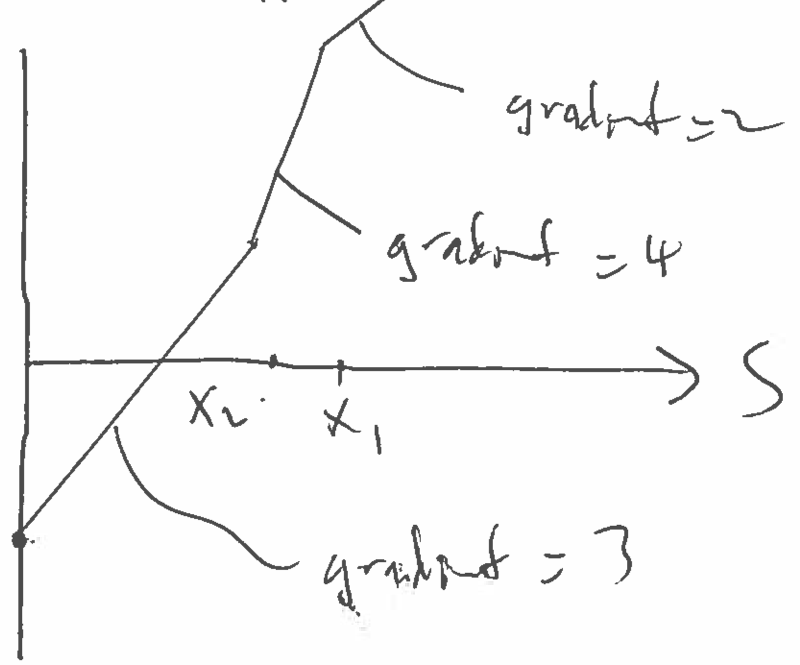
If $X_2 < S < X_1$

$\Pi = S - 2(X_1 - S) + (S - X_2)$

$= 4S - 2X_1 - X_2$

If $X_1 < S$

$\Pi = S + S - X_2 = 2S - X_2$



$\underline{\overline{4}}$

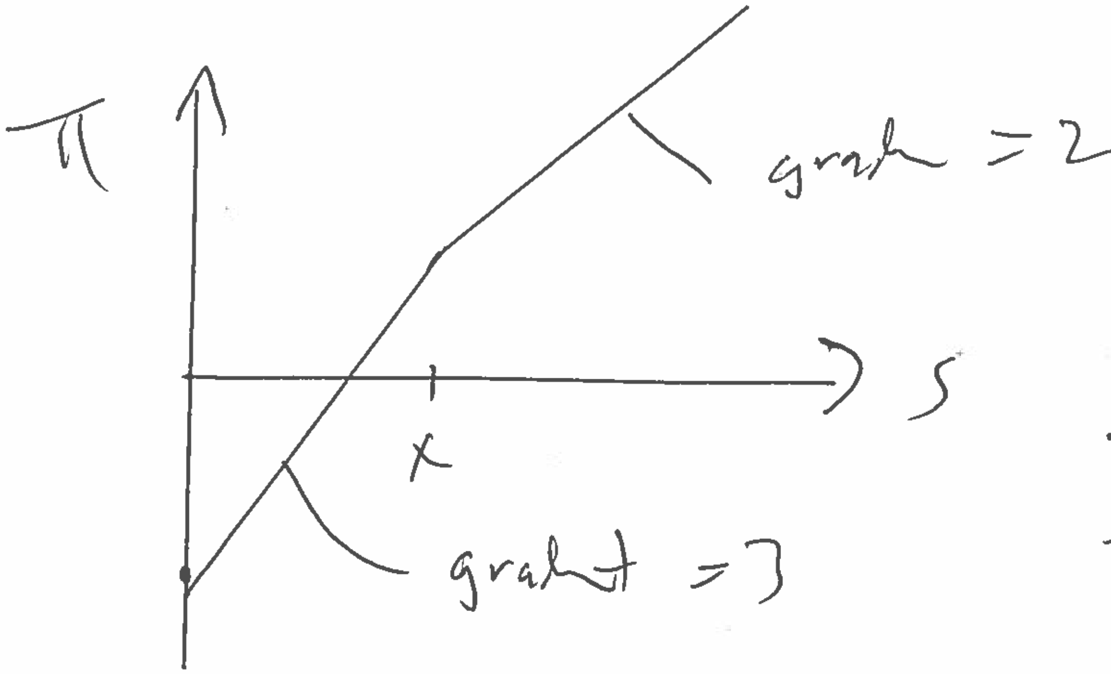
If $x_1 = x_2 = x$ B

If $0 < s < x$

$\pi = s - 2(x - s) = 3s - 2x$

If $x < s$

$\pi = s + (s - x) = 2s - x$



4

If $x_2 > x_1$

If $0 < s < x_1$

$\pi = s - 2(x_1 - s) = 3s - 2x_1$

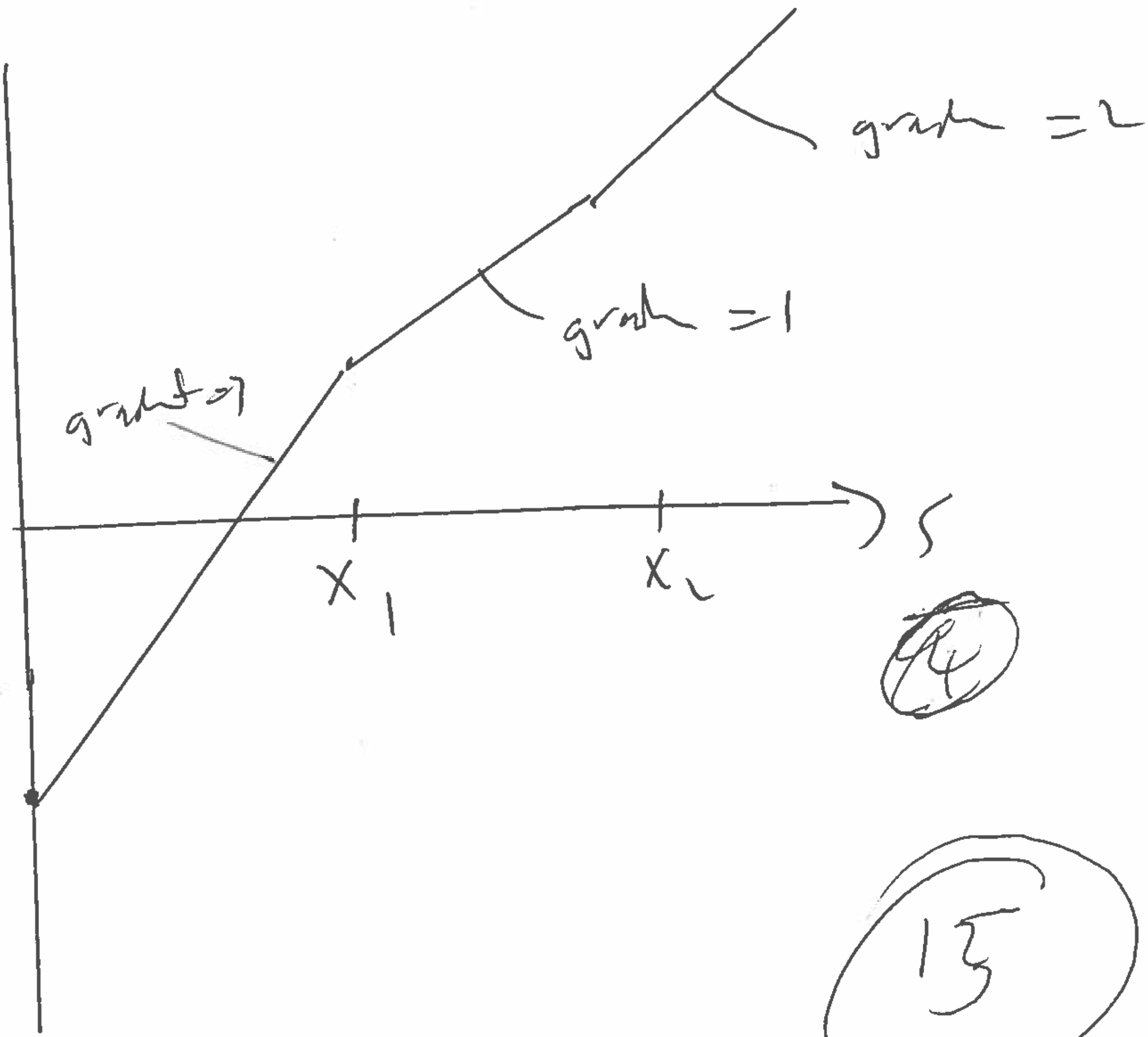
4

If $X_1 < S < X_2$

$\Pi = S$

If $S > X_2$

$\Pi = S + S - X_2 = 2S - X_2$



4

15

A4 1f $P = AS^{\alpha}$ [(i) seen similar (ii), (iii) unseen]

(i)

$$\frac{1}{2}\sigma^2\alpha(\alpha-1) + r\alpha - r = 0$$

$$\alpha^2\left(\frac{1}{2}\sigma^2\right) + \alpha\left(r - \frac{1}{2}\sigma^2\right) - r = 0$$

$$\alpha = 1, \; -2r/\sigma^2$$ 4

(ii)

On $S = S_f$

$$P(S = S_f) = e^{-\lambda S_f} \quad 2$$

$$\frac{dP}{dS}(S = S_f) = -\lambda e^{-\lambda S_f} \quad 2$$

smooth pasting

$$P(S \to \infty) \to 0 \quad 2$$

(iii) $AS_f^{\alpha} = e^{-\lambda S_f}$ $\alpha = \frac{-2r}{\sigma^2}$

$$\alpha A S_f^{\alpha - 1} = -\lambda e^{-\lambda S_f}$$

$$\frac{S_f}{\alpha} = -\frac{1}{\lambda} \Rightarrow S_f = -\frac{\alpha}{\lambda}$$

$$S_f = \frac{2r}{\sigma^2 \lambda}$$ 5

15

BS (i). $dV = \frac{\partial V}{\partial S} dS + \frac{1}{2}\frac{\partial^2 V}{\partial S^2} dS^2 + \frac{\partial V}{\partial t} dt + \ldots$ 1

$E[dS^2] = \sigma^2 dt$ 2

$$dV = \left[\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 V}{\partial S^2}\right]dt + \frac{\partial V}{\partial S} dS$$

see similar

$\Pi = V - \Delta S, \quad d\Pi = dV - \Delta dS$

Choose $\Delta = \frac{\partial V}{\partial S}$ 2

$$d\Pi = \left[\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 V}{\partial S^2}\right]dt$$ riskfree

If no arbitrage $d\Pi = [r\Pi]dt$

$$\Rightarrow \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 V}{\partial t^2} + rS\frac{\partial V}{\partial S} - rV = 0$$

3

(2

(ii) If $1 = \sigma S$

$$\frac{\partial V}{\partial t} + \frac{\sigma^2}{2} S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0$$

unseen

If $V = S^2 f(t)$

$$f' + \sigma^2 f + 2rf - rf = 0$$

$$f' + (\sigma^2 + r) f = 0 \qquad f(T) = 1$$

$$f = \exp\left[\int_t^T (\sigma^2 + r)\, dt\right]$$

$$V = S^2 \exp\left[\int_t^T (\sigma^2 + r)\, dt\right]$$

(iii) $V = S^2 f_1(t) + f_2(t)$

unseen

$$\frac{\partial V}{\partial t} + \frac{\sigma_1^2}{2} \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} - rV = 0$$

$O(S^2)$ $\quad f_1' + 2rf_1 - rf_1 = 0$

$$f_1(T) = 1 \Rightarrow f_1 = \exp[r(T-t)]$$

Q(5c) $f_2' + \sigma_1^2 f_1 - r f_2 = 0$

$$f_2 = A e^{rt} + \alpha e^{-rt}$$

Unseen

$$-2r e^{-rt} \alpha = -\sigma_1^2 e^{r(T-t)}$$

$$\alpha = \frac{\sigma_1^2}{2r} e^{rT}$$

$$f_2 = A e^{rt} + \frac{\sigma_1^2}{2r} e^{r(T-t)}$$

But $f_2(T) = 0$

$$A = -\frac{\sigma_1^2}{2r} e^{-rT}$$

$$f_2 = \frac{\sigma_1^2}{2r}\left[e^{r(T-t)} - e^{-r(T-t)}\right]$$

8

20

BL (i) If $V = e^{-D(T-t)} V_1(S,t)$ ✓

$$DV_1 + \frac{\partial V_1}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V_1}{\partial S^2} + (r-D)\frac{\partial V_1}{\partial S} - rV_1 = 0$$

Seen in class

$$\Rightarrow \frac{\partial V_1}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V_1}{\partial S^2} + (r-D)\frac{\partial V_1}{\partial S} - (r-D)V_1 = 0$$

$\Rightarrow$ BSE, with $r \rightarrow r-D$ $\frac{4}{}$

(ii)

$$C_d = e^{-D(T-t)} S N(d_{10}) - e^{-r(T-t)} N(d_{20})$$

(iii) $$d_{10} = \frac{\log(S/X) + (r-D+\frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$

$$d_{20} = \frac{\log(S/X) + (r-D-\frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$ $\frac{4}{}$

(iii). $T-t = 0.75$

$D = 0.05$

$\sigma = 0.2$

$r = 0.015$

$X = 51, \quad S = 50$

seen similar

$d_{10} = -0.17928, \quad d_{20} = -0.352$

$N(d_{10}) = 0.42885$

$N(d_{20}) = 0.36227$

$Cd = 2.786295$

10 minutes

$\frac{8}{8}$

(iv). $C_d = C_1 e^{-D(T-t)}$

$P_d = P_1 e^{-D(T-t)}$

seen similar

$C_1 + X e^{-(r-D)(T-t)} - P_1 - S = 0$

$e^{D(T-t)} C_d + X e^{-(r-D)(T-t)} - e^{D(T-t)} P_d - S = 0$

$C_d + X e^{-r(T-t)} - P_d - S e^{-D(T-t)} = 0$

$\frac{4}{4}$

20

B6 (i) $dS_i = \mu_i S_i dt + \sigma_i S_i dW_i$ 1

→ geometric Brownian motion

→ changes in S_i linked to magnitude of S_i $\frac{}{2}$

→ S_i cannot go negative $\frac{}{2}$

see similar

(iii) $d\Pi = dV - \Delta_1 dS_1 - \Delta_2 dS_2$

$$= \left[\frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial S_1} dS_1 + \frac{\partial V}{\partial S_2} dS_2 + \frac{1}{2} \frac{\partial^2 V}{\partial S_1^2} dS_1^2 + \frac{1}{2} \frac{\partial^2 V}{\partial S_2^2} dS_2^2 + \frac{\partial^2 V}{\partial S_1 \partial S_2} dS_1 dS_2 \right]$$

$$- \Delta_1 [\mu_1 S_1 dt + \sigma_1 S_1 dW_1]$$
$$- \Delta_2 [\mu_2 S_2 dt + \sigma_2 S_2 dW_2]$$

Similar on ex. sheets

$$= \frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial S_1}\left[\mu_1 S_1 dt + \sigma_1 S_1 dW_1\right]$$

$$+ \frac{\partial V}{\partial S_2}\left[\mu_2 S_2 dt + \sigma_2 S_2 dW_2\right] + \frac{1}{2}\sigma_1^2 S_1^2 \frac{\partial^2 V}{\partial S_1^2} dt$$

$$+ \frac{1}{2}\sigma_2^2 S_2^2 \frac{\partial^2 V}{\partial S_2^2} dt + \sigma_1\sigma_2\rho S_1 S_2 \frac{\partial^2 V}{\partial S_1 \partial S_2} dt$$

$$- \Delta_1\left[\mu_1 S_1 dt + \sigma_1 S_1 dW_1\right]$$

$$- \Delta_2\left[\mu_2 S_2 dt + \sigma_2 S_2 dW_2\right]$$

$$+ o(dt)$$

simila to et sheets

$$\Delta_i = \frac{\partial V}{\partial S_i} \quad \text{removes } dW_i$$

(iii) $$d\pi = \left[\frac{\partial V}{\partial t} + \frac{1}{2}\frac{\partial^2 V}{\partial S_1^2} S_1^2 \sigma_1^2 + \frac{1}{2}\frac{\partial^2 V}{\partial S_2^2} S_2^2 \sigma_2^2 + \frac{\partial^2 V}{\partial S_1 \partial S_2} S_1 S_2 \sigma_1 \sigma_2\right] dt$$

$$= r\pi \, dt$$

$$= dt\, r\left[V - S_1 \frac{\partial V}{\partial S_1} - S_2 \frac{\partial V}{\partial S_2}\right]$$

$$\underline{\underline{S}}\;\; \frac{\partial V}{\partial t} + \frac{1}{2}\sigma_1^2 \frac{\partial^2 V}{\partial S_1^2} + \frac{1}{2}\sigma_2^2 \frac{\partial^2 V}{\partial S_2^2} + \rho\sigma_1\sigma_2 \frac{\partial^2 V}{\partial S_1 \partial S_2}$$

$$+ r S_1 \frac{\partial V}{\partial S_1} + r S_2 \frac{\partial V}{\partial S_2} - rV = 0$$

3

(iv) If $V = a_1 S_2 H(\xi, t)$

$$V(t=T) = \max(a_1 S_1 - a_2 S_2, 0)$$

$$= a_1 S_2 \max\left(\xi - \frac{a_2}{a_1}, 0\right)$$

$$\Rightarrow H(t=T) = \max(\xi - K, 0)$$

As $\xi \to \infty$, $H \to \xi$

As $\xi \to 0$, $H \to 0$

used (v)

$$\frac{\partial V}{\partial S_1} = a_1 S_2 \frac{\partial H}{\partial \xi}\left(\frac{1}{S_2}\right) = a_1 \frac{\partial H}{\partial \xi}$$

$$\frac{\partial V}{\partial S_2} = a_1 H - a_1 \frac{S_1}{S_2} \frac{\partial H}{\partial \xi}$$

$$\frac{\partial^2 V}{\partial S_1^2} = \frac{a_1}{S_2} \frac{\partial^2 H}{\partial \xi^2}$$

$$\frac{\partial^2 V}{\partial S_2^2} = \cancel{a_1 \frac{\partial H}{\partial \xi}\left(-\frac{S_1}{S_2^2}\right)} + \cancel{\frac{a_1 S_1}{S_2^2}\frac{\partial H}{\partial \xi}}$$

$$- a_1 \frac{S_1}{S_2} \frac{\partial^2 H}{\partial \xi^2}\left(-\frac{S_1}{S_2^2}\right) = \frac{\xi^2}{S_2^2}\frac{\partial^2 H}{\partial \xi^2}$$

$$\frac{1}{2}\sigma_1^2 \frac{q_1 S_1^2}{S_2} \frac{\partial^2 H}{\partial \xi^2} + \frac{1}{2} q_1 \sigma_2^2 \xi^2 \frac{S_1}{S_2} \frac{\partial^2 H}{\partial \xi^2}$$ 4

$$+ q_1 S_2 \frac{\partial H}{\partial \tau} + r S_1 q_1 \frac{\partial H}{\partial \xi}$$

$$+ r S_2 \left[q_1 H - \frac{q_1 \xi}{S_2} \frac{\partial H}{\partial \xi} \right] - r q_1 S_1 H = 0$$

unseen

$$\frac{\partial H}{\partial \tau} + \left[\frac{1}{2}\sigma_1^2 + \frac{1}{2}\sigma_2^2 \right] \xi^2 \frac{\partial^2 H}{\partial \xi^2} = 0$$

(i) – (iii) similar to material on example sheet. Remainder unseen.

5

20

Solutions to exam MATH39512 Actuarial Models 2 2017

General remarks MATH39512 Actuarial Models 2 is a new course unit that covers units 5, 6 and 7 of the syllabus of CT4 of the Actuarial Profession on survival analysis which includes in particular non-parametric, parametric and semi-parametric estimation of survival times. The material covered in this course unit goes deeper than the CT4 syllabus as, for the purpose of non-parametric estimation, an introduction to the theory of counting processes and martingales is given. Also the concept of frailty is treated in this course unit.

Regarding the originality level of the questions:

- Question 1 on non-parametric estimation: part (a) is new, parts (b), (c) and (d) are fairly standard. Regarding parts (e) and (f), students have seen similar examples (e.g. the log-rank test) but not this specific one.

- Question 2 on parametric estimation: similar questions as part (a) have been dealt with in class but not with this particular hazard function. Part (b) has appeared in an exercise.

- Question 3 on Cox PH model: the occurrence of left-truncation in computing the partial likelihood is new though left-truncation has been seen in other contexts (like with computing the K-M or N-A estimators). The one-sided test in part (e) is new.

- Question 4 on frailty: similar examples will be covered in class but not with this form (of the pdf) of the frailty Z.

Answer to 1

(a) The survival data consist also of censored values which are lower bounds for the actual survival time. Hence a simple average of the given numbers would severely underestimate the true mean of the survival time.

(b) We first find the values for τ_i (the ith ordered observed genuine failure time), d_i (the observed number of failures at τ_i) and r_i (observed number of individuals at risk of failing just before τ_i). The values are given in the table below. Then we can compute the Kaplan-Meier estimate $\widehat{S}(t)$ at the failure times τ_i using $\widehat{S}(\tau_1) = (1 - d_1/r_1)$ and the recursion, $\widehat{S}(\tau_i) = \widehat{S}(\tau_{i-1})(1 - d_i/r_i)$, $i = 2, 3, 4, 5, 6$. This leads to the last column of the table below. Note that as $\widehat{S}(t)$ is a right-continuous step function, the value of $\widehat{S}(t)$ for $t \in [\tau_i, \tau_{i+1})$, $i = 1, 2, 3, 4, 5$ is given by $\widehat{S}(t) = \widehat{S}(\tau_i)$, whereas $\widehat{S}(t) = 1$ for $t \in [0, \tau_1)$ and $\widehat{S}(t) = \widehat{S}(\tau_6)$ for $t \geq \tau_6$. The plot of the Kaplan-Meier estimate is provided in Figure 1.

i	τ_i	d_i	r_i	$\widehat{S}(\tau_i)$
1	9	1	11	0.909
2	20	1	8	0.795
3	25	1	7	0.682
4	28	1	6	0.568
5	32	1	4	0.426
6	40	1	3	0.284

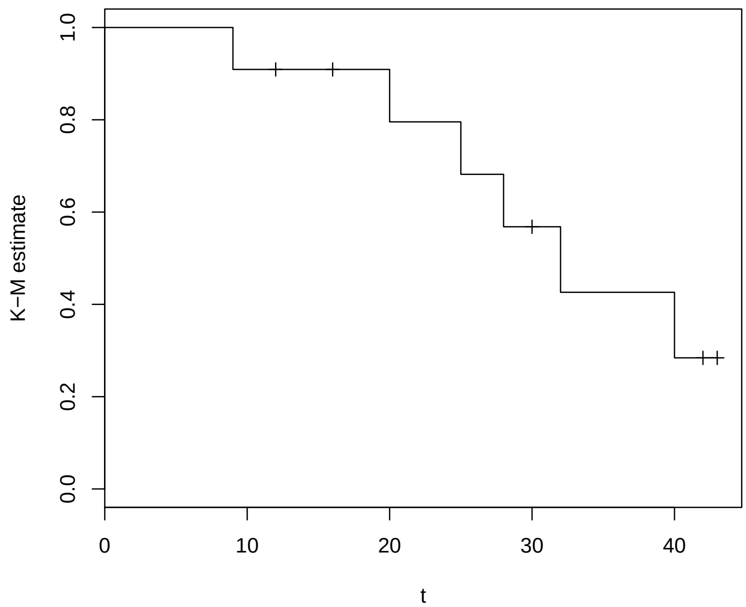


Figure 1: Plot of the Kaplan-Meier estimate.

(c) The median of the survival time T is, by definition, any value t such that both $\mathbb{P}(T \leq t) \geq 0.5$ and $\mathbb{P}(T \geq t) \geq 0.5$. Hence the median of T is any value t such that $S(t) = 1 - \mathbb{P}(T \leq t) \leq 0.5$ and $\lim_{u \uparrow t} S(u) = \mathbb{P}(T \geq t) \geq 0.5$. Using the Kaplan-Meier estimate we can estimate the median of T by finding a value of t such that $\widehat{S}(t) \leq 0.5$ and $\lim_{u \uparrow t} \widehat{S}(u) \geq 0.5$. Since $S(32) = 0.426$ and $\lim_{u \uparrow 32} S(u) = 0.568$, it follows that the estimate of the median of T is 32.

(d) We need to estimate the probability

$$\mathbb{P}(27 < T \leq 35 | T > 15) = \frac{\mathbb{P}(T > 27) - \mathbb{P}(T > 35)}{\mathbb{P}(T > 15)} = \frac{S(27) - S(35)}{S(15)}.$$

We can estimate this probability, using the Kaplan-Meier estimate, via

$$\frac{\widehat{S}(27) - \widehat{S}(35)}{\widehat{S}(15)} = \frac{\widehat{S}(25) - \widehat{S}(32)}{\widehat{S}(9)} = \frac{0.682 - 0.426}{0.909} = 0.281.$$

(e) It is now relevant to record from which group the individual died at a genuine failure time τ_i and how many were at risk from each group separately just before time τ_i. To this end, we denote by d_i^j the observed number of failures from group j at time τ_i and by r_i^j the observed number of individuals from group j at risk of failing just before time τ_i. The table below gives the numbers of these quantities.

i	τ_i	d_i^a	d_i^b	r_i^a	r_i^b
1	9	0	1	6	5
2	20	0	1	5	3
3	25	0	1	5	2
4	28	1	0	5	1
5	32	0	1	3	1
6	40	1	0	3	0

Note that $d_i^j = \Delta N_{\tau_i}^j$ and $R_{\tau_i}^j = r_i^j$. By definition of the Nelson-Aalen estimator, $\widehat{A}_j(t) = \int_0^t \frac{\mathrm{d}N_s^j}{R_s^j}$. Hence we deduce,

$$
\begin{aligned}
Z_{t_0} &= \int_0^{43} R_s^a R_s^b \mathrm{d}\left(\widehat{A}_a(s) - \widehat{A}_b(s)\right) \\
&= \int_0^{43} R_s^b \mathrm{d}N_s^a - \int_0^{43} R_s^a \mathrm{d}N_s^b \\
&= \sum_{i=1}^{6} r_i^b d_i^a - \sum_{i=1}^{6} r_i^a d_i^b \\
&= 1 - 19 = -18.
\end{aligned}
$$

Similarly,

$$
V_{t_0} = \int_0^{43} R_s^a R_s^b \mathrm{d}\left(N_s^a + N_s^b\right) = \sum_{i=1}^{6} r_i^a r_i^b (d_i^a + d_i^b) = 30 + 15 + 10 + 5 + 3 + 0 = 63.
$$

Hence

$$
|Z_{t_0}/\sqrt{V_{t_0}}| = \frac{18}{\sqrt{63}} = 2.27.
$$

Since $2.27 > z_{0.025} = 1.96$, we reject the null hypothesis. We conclude that there is a difference between the hazard functions of both groups.

(f) Using $\widehat{A}_j(t) = \int_0^t \frac{\mathrm{d}N_s^j}{R_s^j}$, $M_t^j = N_t^j - \int_0^t R_s^j \mu_j(s)\mathrm{d}s$ and assuming that the null hypothesis holds,

$$
\begin{aligned}
Z_t &= \int_0^t R_s^a R_s^b \mathrm{d}\left(\widehat{A}_a(s) - \widehat{A}_b(s)\right) \\
&= \int_0^t R_s^b \mathrm{d}N_s^a - \int_0^t R_s^a \mathrm{d}N_s^b \\
&= \int_0^t R_s^b \mathrm{d}M_s^a - \int_0^t R_s^a \mathrm{d}M_s^b + \int_0^t R_s^a R_s^b (\mu_a(s) - \mu_b(s))\mathrm{d}s \\
&\overset{H_0}{=} \int_0^t R_s^a \mathrm{d}M_s^a - \int_0^t R_s^a \mathrm{d}M_s^b.
\end{aligned}
$$

By the hint both $\{M_t^a : t \geq 0\}$ and $\{M_t^b : t \geq 0\}$ are martingales w.r.t. $\{\mathcal{F}_t : t \geq 0\}$. The processes $\{R_s^a : s \geq 0\}$ and $\{R_s^b : s \geq 0\}$ are bounded, adapted to the filtration $\{\mathcal{F}_t : t \geq 0\}$ and have left-continuous sample paths (this is because R_s^j gives the number at risk *just before* time s). Then by a lemma from the notes the stochastic integrals $\{\int_0^t R_s^b \mathrm{d}M_s^a : t \geq 0\}$ and $\{\int_0^t R_s^a \mathrm{d}M_s^b : t \geq 0\}$ are martingales as well. Therefore under H_0, $\{Z_t : t \geq 0\}$ is a martingale. Since martingales have a constant (in time) mean, it follows that $\mathbb{E}[Z_{t_0}] = \mathbb{E}[Z_0] = 0$.

Answer to 2

(a) Denote by T the generic survival time with hazard function $\mu(t)$. Since the hazard function is $\mu(t) = \beta \mathrm{e}^{\gamma t}$, the survival function is

$$
S(t) = \exp\left(-\int_0^t \beta \mathrm{e}^{\gamma s}\mathrm{d}s\right) = \exp\left(-\frac{\beta}{\gamma}(\mathrm{e}^{\gamma t} - 1)\right)
$$

and hence the pdf of T is

$$
f(t) = \mu(t)S(t) = \beta \mathrm{e}^{\gamma t} \exp\left(-\frac{\beta}{\gamma}(\mathrm{e}^{\gamma t} - 1)\right).
$$

Let $L_i(\beta, \gamma)$ be the likelihood of (β, γ) given the data of individual i only, where the individuals are labelled 1 to 7 from left to right. For a censored observation t_i+, the likelihood by the independence of the censoring and the survival time is $L_i(\beta, \gamma) = \Pr(T > t_i)c_i = S(t_i)c_i$, where c_i is some unimportant constant (i.e. does not depend on β or γ). Similarly, for an uncensored observation t_i, the likelihood is $L_i(\beta, \gamma) = \mathrm{Lik}(T = t_i)c_i = f(t_i)c_i$, where again c_i is some unimportant constant. Then given that individuals 1, 2 and 3 fail and the others are censored and that all individuals are independent, the likelihood of all individuals combined is,

$$L(\beta,\gamma) = \prod_{i=1}^{7} L_i(\beta,\gamma) = Cf(4)f(8)S(10)^2 = C\mu(4)\mu(8)S(4)S(8)S(10)^2,$$

where C is an unimportant constant and for the last equality we used $f(t) = \mu(t)S(t)$. Hence the log-likelihood is

$$\ell(\beta,\gamma) = \log(L(\beta,\gamma)) = \log(C) + \log(\mu(4)) + \log(\mu(8)) - A(4) - A(8) - 2A(10),$$

where

$$A(t) = -\log(S(t)) = \frac{\beta}{\gamma}(\mathrm{e}^{\gamma t} - 1)$$

is the cumulative hazard function. Recalling that $\log(\mu(t)) = \log(\beta) + \gamma t$, we get in the end,

$$\ell(\beta,\gamma) = \log(C) + 2\log(\beta) + 12\gamma - \frac{\beta}{\gamma}\left(\mathrm{e}^{4\gamma} + \mathrm{e}^{8\gamma} + 2\mathrm{e}^{10\gamma} - 4\right).$$

(b) From the expression for the cumulative hazard function

$$\log(A(t+1) - A(t)) = \log\left(\frac{\beta}{\gamma}\mathrm{e}^{\gamma t}(\mathrm{e}^{\gamma} - 1)\right) = \gamma t + \log\left(\frac{\beta}{\gamma}(\mathrm{e}^{\gamma} - 1)\right).$$

We can estimate the cumulative hazard function by the Nelson-Aalen estimator denoted by $\widehat{A}(t)$. Then a way for checking that the choice of the parametric form of the hazard function is an appropriate one is to plot $\log(\widehat{A}(t+1) - \widehat{A}(t))$ against t. If the resulting graph resembles closely a straight line, then this gives a good indication that the hazard function is indeed of the proposed form.

Answer to 3

(a) If a player has a left-truncated survival time (meaning that the left endpoint s of the survival data interval $(s,t]$ is strictly positive), then this means that he/she already played a few matches injury-free before he/she started to be observed as part of the study. A non-truncated survival time means that the player just got back from an injury.

(b) The hazard function of a player with training method B is given by

$$\mu_B(t) = \mu_A(t)\mathrm{e}^{\beta},$$

where $\mu_A(t)$ is the hazard function of a player with training method A. Since there are no ties with the genuine failures, we can use the following formula for the partial likelihood that one can find in the notes:

$$L_P(\beta) = \prod_{j=1}^{n_0} \frac{\exp(\beta x_{\mathtt{I}_j})}{\sum_{m \in R_j} \exp(\beta x_m)},$$

where $n_0 = 5$ is the number of genuine failures, $\mathtt{I}_j$, $j = 1, 2, 3, 4$ denotes the individual who failed at the ith ordered, genuine failure time (note that we only have to take into account the training method of the individual as players with the same training method have the same failure time distribution), R_j consists of the set of players who are at risk of failing/getting injured just before the jth ordered, genuine failure time and x_m equals zero if player m works with training method A and equals one if player m works with training method B.
Because of the one categorical covariate involved in the model, we can rewrite this partial likelihood as

$$L_P(\beta) = \frac{\exp(\beta \sum_{j=1}^{5} d_j^B)}{\prod_{j=1}^{5}(r_j^A + r_j^B \mathrm{e}^{\beta})},$$

where

- r_j^k denotes the number of players at risk in group $k \in \{A, B\}$ just before the jth genuine, ordered failure time, which we denote by τ_j;
- d_j^B is the number of players that got injured at the jth genuine, ordered failure time. Note that $d_j^B \in \{0, 1\}$.

These numbers are given by $\sum_{j=1}^{5} d_j^B = 3$ and

j	τ_j	r_j^A	r_j^B
1	4	2	1
2	8	3	2
3	10	2	3
4	14	3	1
5	25	1	0

Therefore,

$$
\begin{aligned}
L_P(\beta) =& \frac{e^{3\beta}}{(2+e^{\beta})(3+2e^{\beta})(2+3e^{\beta})(3+e^{\beta})(1+0e^{\beta})} \\
=& \frac{e^{3\beta}}{(2+e^{\beta})(3+2e^{\beta})(2+3e^{\beta})(3+e^{\beta})}.
\end{aligned}
$$

(c) This test checks whether or not there is a significant difference between the hazard of the number of matches played until injury with training method A and the one with training method B. In particular, if the null hypothesis $\beta = 0$ is not rejected, then there is no significant difference. If on the other hand, H_0 is rejected, then there is strong evidence that there is a difference between the two hazards.

(d) We see from the figures that the maximum likelihood estimate of β is $\hat{\beta} \approx 1.62$ as this is the point where both L_P and ℓ_P reaches its maximum and where the score function ℓ'_P is equal to 0. Hence the relative risk, which is the ratio of the hazard functions of each group, is estimated by

$$e^{\hat{\beta}} = 5.05.$$

(e) We want to test the hypothesis $H_0 : \beta = 0$ versus $H_A : \beta > 0$ as $\beta > 0$ indicates that training method A is better than method B for preventing/delaying injuries. With the Wald test, we reject H_0 at significance level 0.05 if

$$\left| \frac{\hat{\beta} - 0}{\sqrt{-1/\ell''_P(\hat{\beta})}} \right| = \left| \hat{\beta}\sqrt{\ell''_P(\hat{\beta})} \right| > z_{0.025} = 1.96.$$

In order to evaluate this test statistic, we need to compute $-\ell''_P(\hat{\beta})$. So we need to determine the negative of the slope of $\ell'_P(\beta)$ at $\beta = \hat{\beta} = 1.62$. As $\ell'_P(\beta)$ is close to linear, we can from the plot of ℓ'_P approximate the required quantity by,

$$-\ell''_P(\hat{\beta}_P) = -\frac{\ell'_P(1.6) - \ell'_P(1.7)}{1.6 - 1.7} \approx -\frac{0.02 - -0.055}{-0.1} = 0.75.$$

We conclude

$$\left| \hat{\beta}\sqrt{\ell''_P(\hat{\beta})} \right| \approx 1.62\sqrt{0.75} = 1.40 < z_{0.025} = 1.96.$$

Hence we do not reject H_0 at 0.05 significance level according to the Wald test. We conclude that there is not enoug evidence to say that training method A is better in preventing injuries than method B.

Answer to 4

(a) Let T be the "aggregate/collective" survival time of the population. By definition of the individual hazard function $\mu_1(t|Z)$, the individual survival function (given the frailty Z) is given by

$$\mathbb{P}(T > t|Z) = \mathrm{e}^{-\int_0^t \mu_1(s|Z)\mathrm{d}s} = \mathrm{e}^{-\int_0^t 2sZ\mathrm{d}s} = \mathrm{e}^{-t^2 Z}.$$

Hence the survival function of the population is given by

$$\mathbb{P}(T > t) = \mathbb{E}[\mathbb{P}(T > t|Z)] = \mathbb{E}[\mathrm{e}^{-t^2 Z}].$$

The Laplace transform of Z is given by

$$\begin{aligned}\mathbb{E}[\mathrm{e}^{-\lambda Z}] = \int_0^\infty \mathrm{e}^{-\lambda z} f_Z(z)\mathrm{d}z = \int_0^\infty \mathrm{e}^{-\lambda z} 3\left(\mathrm{e}^{-1.5z} - \mathrm{e}^{-3z}\right)\mathrm{d}z =& \frac{3}{\lambda + 1.5} - \frac{3}{\lambda + 3} \\ =& \frac{4.5}{(\lambda + 1.5)(\lambda + 3)}.\end{aligned}$$

This then leads to $\mathbb{P}(T > t) = \frac{4.5}{(t^2+1.5)(t^2+3)}$. The connection between the survival function and the hazard function is $\mathbb{P}(T > t) = \mathrm{e}^{-\int_0^t \mu_1(s)\mathrm{d}s}$. Using this, we then deduce that the population hazard function $\mu_1(t)$ is given by

$$\mu_1(t) = -\frac{\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{P}(T > t)}{\mathbb{P}(T > t)} = \frac{\frac{4.5\cdot 2t(2t^2+4.5)}{[(t^2+1.5)(t^2+3)]^2}}{\frac{4.5}{(t^2+1.5)(t^2+3)}} = 2t\frac{2t^2 + 4.5}{(t^2 + 1.5)(t^2 + 3)}.$$

(b) A crossover phenomenon is when, as a function of time t, the ratio of two population hazard functions (assumed to be continuous), say $\mu_1(t)$ and $\mu_2(t)$, crosses the line $y = 1$. So a crossover occurs if for some time t, $\mu_1(t) > \mu_2(t)$ and for some other time t', $\mu_1(t') < \mu_2(t')$.

(c) First we need to compute $\mu_2(t)$, the population hazard function of the second population. Let T_2 be the survival time with hazard function $\mu_2(t)$. We have

$$\begin{aligned}\mathrm{e}^{-\int_0^t \mu_2(s)\mathrm{d}s} = \mathbb{P}(T > t) =& \mathbb{E}[\mathbb{P}(T > t|Y)] \\ =& \mathbb{E}[\mathrm{e}^{-\int_0^t \mu_2(s|Y)\mathrm{d}s}] \\ =& \mathbb{E}[\mathrm{e}^{-\int_0^t 4sY\mathrm{d}s}] \\ =& \mathbb{E}[\mathrm{e}^{-2t^2 Y}] = \frac{1}{2t^2 + 1},\end{aligned}$$

where the last equality follows because the Laplace transform of Y, which is exponentially distributed with parameter 1, is given by

$$\mathbb{E}[\mathrm{e}^{-\lambda Y}] = \int_0^\infty \mathrm{e}^{-\lambda y}\mathrm{e}^{-y}\mathrm{d}y = \frac{1}{\lambda+1}.$$

We deduce

$$\mu_2(t) = -\frac{\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{P}(T_2 > t)}{\mathbb{P}(T_2 > t)} = \frac{\frac{4t}{(2t^2+1)^2}}{\frac{1}{2t^2+1}} = \frac{4t}{2t^2+1}.$$

Next, we show that there is a crossover. We see that $\mu_1(t) \sim \frac{4t^3}{t^4} = \frac{4}{t}$ as $t \to \infty$ (i.e. $\lim_{t\to\infty} \frac{\mu_1(t)}{4/t} = 1$). Similarly, $\mu_2(t) \sim \frac{2}{t}$ as $t \to \infty$. Hence $\lim_{t\to\infty} \frac{\mu_2(t)}{\mu_1(t)} = \frac{1}{2}$, so for large t, $\mu_1(t)$ is bigger than $\mu_2(t)$. On the other hand, $\mu_1(t) \sim \frac{9t}{4.5} = 2t$ as $t \downarrow 0$ (i.e. $\lim_{t\downarrow 0} \frac{\mu_1(t)}{2t} = 1$) and $\mu_2(t) \sim 4t$ as $t \downarrow 0$. Hence $\lim_{t\downarrow 0} \frac{\mu_2(t)}{\mu_1(t)} = 2$, so for small t, $\mu_1(t)$ is smaller than $\mu_2(t)$. Hence there is a crossover phenomenon.

(d) Consider an individual from each population with the same level of frailty $z > 0$. Then the hazard rate, at any time t, of the individual from population 1 is smaller than the hazard rate of the individual from population 2, since $\mu_1(t|Z = z) = 2tz < 4tz = \mu_2(t|Y = z)$. You might think then that, at any time t, the population hazard of population 1, i.e. $\mu_1(t)$, is also smaller than $\mu_2(t)$, the population hazard of population 2. However, part(c) shows that this is not the case; in particular $\mu_1(t) > \mu_2(t)$ for large t. This counterintuitive result is an illustration of Simpson's paradox (which is about a trend appearing on the individual level, but disappearing or even reversing on the aggregate level).

Contingencies 2, Exam May-June 2016-17 SOLUTIONS

1 Context

The examination follows on from the approach adopted last year and is similar to that used for Contingencies 1. A formula sheet is attached to the examination paper which is a subset of relevant information in the Orange Book. It is a shortened version of the sheet for Contingencies 1 with some added joint life functions described.

The exam will be sat in the computer cluster with internet access restricted to the manchester.ac.uk domain only and access to no software other than R, R Studio and Notepad. Supervision, including the presence of IT staff, will be arranged in the same way as last year.

As for the questions, they are set consistently with the material covered in the course notes. Questions are also set consistently with the tutorials. Question 1 is straightforward book work although does get progressively harder. The use of a same sex couple means that the student does not have to spend time changing the mortality table - this aspect is covered in the following question. Question 2 is similar to the questions from the exercise sheet although part b has an aspect not covered in recent years which is the calculation of assurances where the order of death is relevant. Question 3 is largely book work but with an extra ingredient in part b) which will be new but doable for the students. Question 4 is similar to exercise sheet questions but part b) may extend some students. Mortality rates come from AM92 and are given to save some time for the students. Question 5 tests basic use of the service table and is book work. It has been set separately from Question 6 to encourage more students to tackle a pensions based question than we have seen in earlier years. Question 6 reflects the approach adopted for the first time last year for pensions in R where the basic EPV commutation functions are placed in a vector and the benefits are put into a separate vector. This contrasts with the commutation functions in the Orange Book which mixes service periods (intrinsically part of the benefits calculation) with the EPV of £1 of benefit. Keeping benefits and basic EPV calculations separate makes more sense and reflects practice in industry where all the hard work is in defining the correct benefits when carrying out triennial actuarial valuations. Question 7 is a standard unit linked cash flow question but with the insertion of both a guaranteed sum assured but also a surrender value. The final two questions are book work based.

The material covered in Contingencies 2 is based primarily on objectives (vi) to (x) (but with reference back to (i) to (v)) of the CT5 course. In terms of the questions they are set based on:

- Question 1 - Based mainly on objective (vi), part 1.
- Question 2 - Based mainly on objective (vi), part 1.
- Question 3 - Based mainly on objective (vii), parts 1 and 3.
- Question 4 - Based mainly on objective (viii), part 2 .
- Question 5 - Based mainly on objective (viii), part 6 .
- Question 6 - Based mainly on objective (viii), part 6 .
- Question 7 - Based mainly on objective (ix), part 2 .
- Question 8 - Based mainly on objective (ix), part 5.
- Question 9 - Based mainly on objective (x), part 4.

2 Solutions

1. In PFA92C20, express the probabilities below in terms of single life probabilities and then calculate their value (in all cases both lives are female):

   (a) $p_{65:60}$

   (b) ${}_5q_{60:55}$

   (c) ${}_5p_{\overline{60:55}}$

   (d) ${}_{5|}q_{\overline{60:55}}$

[Total 13 marks]

SOLUTION Q1

R Code to produce the single life probabilities needed:

```
useLifeTable("PFA92C20")

x<-65
n<-1

px65<-getLives(x+n)/getLives(x)
px65
```

```
[1] 0.9953193

 x<-60
 n<-1

 px60<-getLives(x+n)/getLives(x)
 px60
[1] 0.9979422

 x<-60
 n<-5

 px560<-getLives(x+n)/getLives(x)
 px560
[1] 0.9853063

 x<-55
 n<-5

 px555<-getLives(x+n)/getLives(x)
 px555
[1] 0.9930238
```

(a) $p_{65} \cdot p_{60} = 0.9953 \times 0.9979 = 0.9932$

(b) $(1 - {}_5p_{60} \cdot {}_5p_{55}) = 1 - 0.9853 \times 0.9930 = 0.02160$

(c) ${}_5p_{60} + {}_5p_{55} - {}_5p_{60} \cdot {}_5p_{55} = 0.9853 + 0.9930 - 0.9853 \times 0.9930 = 0.99990$

(d) ${}_5p_{60}(1-p_{65}) + {}_5p_{55}(1-p_{60}) - {}_5p_{60} \cdot {}_5p_{55}((1-p_{65}) \cdot (1-p_{60})) = 0.00463 + 0.00209 - 0.00665 = 0.00007$

2. An insurance company issues a policy to two male lives both aged 30 exact. We will refer to the two people as *person 1* and *person 2* to distinguish between them. You have been asked to calculate the net premium for two possible policies.

   In both cases, assume interest at 4% p.a. and mortality of PMA92C20 for both lives. Also use the standard approximation $\bar{A}_u \approx (1+i)^{0.5} \cdot A_u$ to convert from an assurance payable at the end of the year in which the status u fails to an assurance that pays out immediately on the failure of status u.

   (a) Assume the policy pays out £20,000 immediately on the second death of the two lives. Premiums are payable annually in advance until the first death occurs.

Calculate the net premium

(b) In the second case assume the policy pays out a sum of £15,000 immediately on the death of *person 1* but only if he dies first or a payment of £5,000 immediately on the death of *person 2* if he dies second. Premiums continue to be paid annually in advance until the first death occurs.

Calculate the net premium.

[Total 12 marks]

SOLUTION Q2

```
 #q2
x<-30
y<-30
i<-0.04
useLifeTable("PA92C20")
A3030<-getEPVJointWholeLifeAss(x,"M",y,"M",i)
A3030
[1] 0.1556561

a3030<- getEPVJointAnnAdvLife(x,"M",y,"M",i)
a3030
[1] 21.95294
useLifeTable("PMA92C20")

A30<-getEPVWholeLifeAss(x,i)
A30
[1] 0.1278313
```

(a) This is:

$$
\begin{aligned}
\text{Net premium} &= 20{,}000 \times 1.04^{0.5} A_{\overline{30:30}} / \ddot{a}_{30:30} \\
&= 20{,}396(A_{30} + A_{30} - A_{30:30})/21.953 \\
&= 20{,}396(0.12783 + 0.12783 - 0.15566)/21.953 \\
&= 92.91
\end{aligned}
$$

(b) This is:

$$\begin{aligned}
\text{Net premium} &= (15{,}000 \times 1.04^{0.5} A^{1}_{30:30} + 5{,}000 \times 1.04^{0.5} A^{2}_{30:30})/\ddot{a}_{30:30} \\
&= (15{,}000 \times 1.04^{0.5} \times 0.5 A_{30:30} + 5{,}000 \times 1.04^{0.5} \times 0.5 A_{\overline{30:30}})/21.953 \\
&= (15{,}297 \times 0.5 \times 0.15566 + 5{,}099 \times 0.5 \times 0.1)/21.953 \\
&= 64.57
\end{aligned}$$

3. Consider a 3 state state sickness model with states a, able, i, ill and d, dead. Transition intensities are:

- from a to i at age x : μ_x^{ai},
- from a to d at age x : μ_x^{ad},
- from i to a at age x : μ_x^{ia},
- from i to d at age x : μ_x^{id}.

(a) Write down the EPV for a whole of life lump sum death benefit of £10,000 payable immediately on death.

(b) Another policy also pays out a lump sum on death but the amount payable varies depending on the circumstances. In the event of death having never been ill the payment is £20,000. In all other circumstances the payment on death is £10,000. Write down the EPV of these revised death benefits, again in integral form.

[Total 10 marks]

SOLUTION Q3

(a) answer is $10{,}000 \int_0^\infty e^{-\delta t}({}_tP_x^{aa} \cdot \mu_{x+t}^{ad} + {}_tP_x^{ai} \cdot \mu_{x+t}^{id}) \cdot dt$.

(b) answer is $20{,}000 \int_0^\infty e^{-\delta t}{}_tP_x^{\overline{aa}} \cdot \mu_{x+t}^{ad} \cdot dt$

and

$10{,}000 \int_0^\infty e^{-\delta t}(({}_tP_x^{aa} - {}_tP_x^{\overline{aa}}) \cdot \mu_{x+t}^{ad} + {}_tP_x^{ai} \cdot \mu_{x+t}^{id}) \cdot dt$.

4. Consider a 3 state multiple decrement model with states a, active, r, retired and d dead.

   (a) Calculate the dependent probabilities of dying and of retiring when aged 60 and 61 where the independent probabilities of death are $q^d_{60} = 0.008022$ and $q^d_{61} = 0.009009$ and the force of retirement $\mu^{ar}_x = 0.1$ for all $x > 0$.

   (b) Now consider the introduction of a fourth state, withdrawal, in our model so that the model remains a multiple decrement model. If withdrawal is assumed to occur only at each year end and the independent probability of withdrawal is 0.1 for all $x > 0$, calculate the dependent probability of a person aged 60 exact dying when aged 61.

[Total 15 marks]

SOLUTION Q4

(a) $q^d_{60} = 0.008022$ and $q^d_{61} = 0.009009$.

Hence:

$$
\begin{aligned}
\mu^d_{60} &= -ln(1 - q^d_{60}) = 0.008054, \text{ and} \\
\mu^d_{61} &= -ln(1 - q^d_{61}) = 0.009050
\end{aligned}
$$

Therefore:

$$
\begin{aligned}
(aq)^d_{60} &= \frac{0.008054}{0.108054} \cdot (1 - e^{-0.108054}) \\
&= 0.007634 \\
(aq)^r_{60} &= 0.1 \times 0.007634/0.008054 \\
&= 0.09479 \\
(aq)^d_{61} &= \frac{0.009050}{0.109050} \cdot (1 - e^{-0.109050}) \\
&= 0.008574 \\
(aq)^r_{61} &= 0.1 \times 0.008574/0.009050 \\
&= 0.09474
\end{aligned}
$$

(b) The probability is:

$$(ap)_{60} \cdot (1 - 0.1) \cdot (aq)^d_{61} = (1 - (0.007634 + 0.09479)) \times 0.9 \times 0.008574 = 0.006926.$$

5. Use the pension scheme service table to calculate the following:

   (a) The probability that a life aged 40 exact withdraws from the pension scheme in the following year.

   (b) The probability that a life aged 50 exact retires in normal health at any time between the ages of 60 and 65 inclusive.

   (c) The estimated salary earned in the 12 month period through to retirement when aged 65 exact for a person now aged 30 exact with a salary in the previous year of £15,000.

[Total 9 marks]

SOLUTION Q5

```
useLifeTable("Pension")
getPensionWithdrawn(40)
[1] 413
```

```
 pw40<-getPensionWithdrawn(40)/getPensionActive(40)
 pw40
[1] 0.02682341

 rx<-getPensionRetired("all")
 rx
 [1]    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
[16]    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
[31]    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
[46]    0    0    0    0    0    0    0    0    0    0    0    0    0    0 3681
[61]  516  453  395  342 3757
 rx60<-rx[60:65]
 rx60
[1] 3681  516  453  395  342 3757
 getPensionActive(40)
[1] 15397
 pr<-sum(rx60)/getPensionActive(40)
 pr
[1] 0.5938819

 getSalScale(29)
[1] 4.991
 getSalScale(64)
[1] 11.328
```

Hence the answers are:

(a) 0.02682

(b) 0.59388

(c) 15,000 × 11.328/4.991 = 34,045

6. A pension scheme provides a lump sum benefit on retirement which is calculated as follows:

$$\text{Lump sum payable on retirement at age } x = 0.05 \times FPS \text{ at } x \times N.$$

where:

- FPS is an estimate of the average salary earned over the three years prior to retirement and
- N is the length of service in years and part years, and is calculated from the date of joining the pension scheme through to the estimated date of retirement.

Use the pension service table to:

(a) Write down a vector showing an estimate of the amount of lump sum payable on retirement when aged 62 and all subsequent ages through to normal retirement at age 65 inclusive for a person now aged 50 exact and who joined the pension scheme 25 years previously. The person has a salary of £20,000 earned in the previous year.

(b) Write down the vector for the EPV of a lump sum payment of £1 on retirement when aged 62 and all subsequent ages through to normal retirement at age 65 inclusive. Then calculate the EPV of the benefits calculated in part a). Use an interest rate of 4%p.a.

[Total 16 marks]

SOLUTION Q6

```
 N<-c(37.5,38.5,39.5,40)
 N
[1]  37.5 38.5 39.5 40.0
 zh<-getAveSalScale("allHalf")
 zhr<-zh[62:65]
 zhr
[1]  10.74417 10.90567 11.07183 11.15667
 benvec<-N*0.05*20000*zhr/getSalScale(49)
 benvec
[1]  44613.69 46491.88 48426.24 49414.98

 i<-0.04
 cr<-getCommCr("all",i)


 epvvec<-cr[60:65]/getCommD(50,i)
 epvvec
[1]  0.02173846 0.01822612 0.01517364 0.16345128
 epv<-benvec*epvvec
```

```
 sum(epv)
[1] 10628.85
```

Hence the answers are the vector *benvec* shown above for part a) and for part b) the vector *epvvec* shown above and the EPV is £10,629.

7. An insurer writes a 2 year unit linked endowment policy for a life aged 50. Level premiums are paid yearly in advance. Benefits are paid at the end of the year.

   (a) Construct a table setting out the development of the unit fund each year and then a table showing the cash flows each year for the non unit fund. Briefly show your workings so that the derivation of each item in the tables can be seen and state what the profit vector for the policy is.

   Details are as follows:

   - Premium = £ 2,000 p.a.
   - Allocation is 50% in year 1 and 105% in year 2
   - Bid offer spread = 5%
   - Management charges 0.5% deducted from the unit fund at the end of the year, before any claim payments are made.
   - Initial expenses are 15% of first premium and £ 500 on receipt of the first premium.
   - Renewal expenses are 1% of the second premium and £ 50 on payment of this premium.
   - Dependent probability of death = 0.01 in year 1 and 0.012 in year 2 and the dependent probability of withdrawal is 0.1 in each year.
   - A minimum sum assured of £ 3000 is payable on death or on maturity of the policy (both payable at the policy year end).
   - A surrender value of 75% of premiums paid is made at the policy year end.
   - Assumed rate of return of 5% p.a. on the unit fund and 3% p.a. on the non-unit fund.

[Total 15 marks]

SOLUTION Q7

Unit Fund	explanation	Year 1	Year 2
Allocation	50% then 105% of £2,000	1,000	2,100
less bid/offer	5% off	-50 =950	-105 =1,995
add brought forward fund		-	+992.51 =2,987.50
add interest	at 5%	+47.50 =997.50	+149.38 =3,136.88
less mangt charge	less 0.5%	-4.99	-15.68
closing UF value		=992.51	=3,121.20

Non Unit Fund	explanation	Year 1	Year 2
Allocation	50% then -5% of £2,000	1,000	-100
add bid/offer	from UF	+50	+105
less initial expenses	15% of 2,000 + 500	-800	-
less renewal expenses	1% of 2,000 + 50	- =250	-70 -65
add interest	at 3%	+7.50 =257.50	-1.95 =-66.95
less death cover	$(3000 - 992.51) \times 0.01$, 0 in yr 2	-20.07	-
less surrender cost	$(.75 \times 2000 - 992.51) \times 0.1$	-50.75	-
add mangt charge	from UF	+4.99	+15.68
closing NUF value		=191.67	=-51.27

Closing NUF is the profit vector.

8. Give two reasons why insurance companies zeroise their reserves.

[Total 4 marks]

SOLUTION Q8

Insurers zeroise their reserves so that:

(a) any negative cash flows that may occur in some future years of the policy's lifetime are removed so that there is no expected need for capital at a later date - capital is only needed at the outset of the policy, and

(b) Equally insurers do not want to over reserve and tie up too much capital in reserves.

Zeroising meets these objectives.

9. Define what temporary initial selection and class selection are and give an example of each in the area of life assurance.

[Total 6 marks]

SOLUTION Q9

Class selection refers to permanent differences in the expected mortality experience of two groups. For example in life assurance different premiums might be charged for men and women because men and women show different mortality experience across all future years.

Temporary initial selection refers to different mortality experience that exists in the early years of a policy and where these differences disappear after a period of time. An example is the impact of a medical on taking out an insurance policy. A group of people aged x who have just passed a medical would be expected to show a lower mortality rate than that of a another group also aged x but who passed a medical several years ago or indeed have never taken a medical.

Solutions to exam MATH39542 Risk Theory 2016/17

General remarks The syllabus of MATH39542 Risk Theory consists of (i) ruin theory (Questions 1&2), (ii) premium principles (Question 3), (iii) Bayesian statistics (Question 4 and part Question 5) and (iv) credibility theory (part Question 5).

Most of the questions are similar to questions on the example sheets (the degree of similarity varies). However none are identical, other models (not just other parameters) are used instead and/or the question is somewhat different. Notably:

- (partly) unseen are Question 1(c) (though a similar technique has been used in the lecture), Question 2(b) (bounding the Lundberg coefficient), both premium principles in Question 3, and Question 4 (though a similar technique has been used in the lecture to prove the important result when $g \equiv 1$)
- Question 3(a)(i) and 3(c)(i) are bookwork

Finally, last year's exam turned out to be a bit too straightforward (heavy scaling down had to be applied), this exam paper is hence on purpose more challenging — the students will be told this.

Answer to 1

(a) The requested probability is $\mathbb{P}(N_1 \leq 1)$, which using that $N_1 \sim \text{Poisson}(\lambda)$ can be computed as

$$\mathbb{P}(N_1 \leq 1) = \mathbb{P}(N_1 = 0) + \mathbb{P}(N_1 = 1) = e^{-\lambda}(1 + \lambda).$$

(b) The requested probability is $\mathbb{P}(U_{J_1} \geq 0)$, where J_1 denotes the arrival time of the first claim. Hence

$$\mathbb{P}(U_{J_1} \geq 0) = \mathbb{P}(u + cJ_1 - 1 \geq 0) = \mathbb{P}(J_1 \geq (1-u)/c).$$

If $u \geq 1$ then $(1-u)/c \leq 0$ and hence the probability equals 1. If $u \in [0,1)$ then using that $J_1 \sim \text{Exp}(\lambda)$ it follows that the probability equals $\exp(-\lambda(1-u)/c)$.

(c) Let A_k be the event that the first k claims do not lead to ruin given that the initial capital is u. For any $z \geq 0$ it holds that

$$\mathbb{P}(A_k \,|\, U_{J_1} = z) = p_{k-1}(z)$$

since by the conditioning the first claim did not lead to ruin and by homogeneity the probability of no ruin in the interval $(J_1, J_k]$ given that $U_{J_1} = z$ equals $p_{k-1}(z)$. For any $z < 0$ we have that

$$\mathbb{P}(A_k \,|\, U_{J_1} = z) = 0$$

since by the conditioning ruin happens due to the first claim. As given in the question we also have that $p_{k-1}(z) = 0$ for all $z < 0$. So we have that $\mathbb{P}(A_k \,|\, U_{J_1} = z) = p_{k-1}(z)$ for all $z \in \mathbb{R}$. Hence by the Tower Property

$$p_k(u) = \mathbb{P}(A_k) = \mathbb{E}\left[p_{k-1}(U_{J_1})\right] = \mathbb{E}\left[p_{k-1}(u + cJ_1 - 1)\right] = \int_0^\infty p_{k-1}(u+cs-1)\lambda e^{-\lambda s}\,\mathrm{d}s,$$

where in the final step we used that $J_1 \sim \text{Exp}(\lambda)$.

(d) Excess of loss reinsurance with retention level M entails that the insurer pays $X_k^{(M)} := \min\{X_k, M\}$ for the k-th claim. Since $X_k = 1$ and $M \in [0,1]$ this simplifies to $X_k^{(M)} = M$.

(i) Writing the capital process with reinsurance in force as

$$U_t^{(M)} = u + c^{(M)}t - \sum_{k=1}^{N_t} X_k^{(M)} \quad \text{for } t \geq 0$$

where $c^{(M)} = c - \beta c(1-M)$, the NPC entails that $c^{(M)} > \lambda \mathbb{E}[X_1^{(M)}]$ i.e.

$$c - \beta c(1-M) > \lambda M \iff M > \frac{c(1-\beta)}{\lambda - \beta c}.$$

(Note that both denominator and numerator in above fraction are negative due to $\beta > 1$ and the assumption that $c > \lambda$, the latter being the NPC for U).

(ii) Let $h(M)$ denote the ruin probability for $U^{(M)}$. The insurer aims to minimise this over $M \in [0,1]$. Denoting

$$M_0 := \frac{c(1-\beta)}{\lambda - \beta c} \in (0,1),$$

we know from part (i) that for $M \in [0, M_0]$ the NPC does not hold and hence $h(M) = 1$. For $M \in (M_0, 1]$, since the initial capital is 0 we know from the lecture (notes) that

$$h(M) = \frac{\lambda}{c^{(M)}} \mathbb{E}[X_1^{(M)}] = \frac{\lambda M}{c - \beta c(1-M)}.$$

Hence for such M

$$h'(M) = \frac{c\lambda(1-\beta)}{(c - \beta c(1-M))^2}$$

which is strictly negative since $\beta > 1$.
So $h(M) = 1$ for $M \in [0, M_0]$ and strictly decreasing on $(M_0, 1]$, hence it follows that $M = 1$ i.e. no reinsurance at all is the optimal choice.

Answer to 2

(a) The Laplace transform of the claim sizes is computed as

$$\mathbb{E}[e^{-\theta X_1}] = \int_0^1 e^{-\theta x} \frac{e^{-x}}{1-e^{-1}} \,\mathrm{d}x = \frac{1}{1-e^{-1}} \int_0^1 e^{-(\theta+1)x} \,\mathrm{d}x = \frac{1-e^{-(\theta+1)}}{(1-e^{-1})(\theta+1)} \quad \text{for } \theta \in \mathbb{R}$$

(where this fraction is understood to be equal to $1/(1-e^{-1})$ for $\theta = -1$).
Using a formula from the (lecture) notes we can compute the Laplace exponent of U as

$$\xi(\theta) = c\theta - \lambda + \lambda \mathbb{E}[e^{-\theta X_1}] = \theta - 1.5 + 1.5 \frac{1-e^{-(\theta+1)}}{(1-e^{-1})(\theta+1)} \quad \text{for } \theta \in \mathbb{R},$$

with (hence) domain $\mathbb{R}$.

(b) (i) We first compute the common mean of the claim sizes as

$$\mathbb{E}[X_1] = \int_0^1 x \frac{e^{-x}}{1-e^{-1}} \,\mathrm{d}x = \frac{1}{1-e^{-1}} \int_0^1 x e^{-x} \,\mathrm{d}x = \frac{1-2e^{-1}}{1-e^{-1}} = \frac{e-2}{e-1} \approx 0.4.$$

The Net Profit Condition $c > \lambda \mathbb{E}[X_1]$ i.e. $1 > 1.5\mathbb{E}[X_1]$ is hence satisfied.
We know from the lecture (notes) that the survival probability φ is non-decreasing as a function of u and that $\varphi(0) = 1 - \lambda \mathbb{E}[X_1]/c$, hence for all $u \geq 0$

$$\varphi(u) \geq \varphi(0) = 1 - 1.5 \frac{e-2}{e-1} \implies \psi(u) \leq 1.5 \frac{e-2}{e-1}.$$

(ii) To establish $\psi(u) \leq e^{-1.3u}$ we use the Lundberg inequality which states that $\psi(u) \leq e^{-Ru}$ where $-R$ is the unique negative root (provided it exists) of the Laplace exponent ξ of U. Now, we know from the lecture (notes) — or alternatively by direct analysis — that ξ is a strictly convex function with $\xi(0) = 0$ and $\xi'(0) > 0$ (the latter since the NPC holds). Since the domain of ξ is $\mathbb{R}$ as verified in part (a), we know from the lecture (notes) — or again direct analysis — that $\xi(-\infty) = \infty$. Hence it follows that ξ does indeed have a unique negative root $-R$ and furthermore that $\xi(\theta) < 0$ iff $\theta \in (-R, 0)$. Even though we can't find an explicit expression for R, since $\xi(-1.3) \approx -0.03 < 0$, the analysis of ξ shows that hence $R \geq 1.3$ and therefore by the Lundberg inequality

$$\psi(u) \leq e^{-Ru} \leq e^{-1.3u}$$

as required.

Answer to 3

(a) (i) The no rip-off property for a premium principle π entails that if a constant $C > 0$ exists such that $X \leq C$ (a.s.), then $\pi(X) \leq C$.

(ii) No. A counter example is for instance the trivial risk $X = 0$, or otherwise for instance any non-trivial rv X with range $\{0, a\}$ since then $X \leq a$ but $\pi(X) = \mathbb{E}[X] + a > a$.

(b) (i) Since $F_X(x) = 1 - e^{-\alpha x}$ for $x \geq 0$ it follows that

$$\pi_2(X) = \int_0^\infty \left(e^{-\alpha x}\right)^{1/b} \mathrm{d}x = \int_0^\infty e^{-\alpha x/b} \, \mathrm{d}x = \frac{b}{\alpha},$$

and since $\mathbb{E}[Y] = 1/\beta$ it follows that we should set $\beta = \alpha/b$.

(ii) Comparing the pdf's of X and Y, i.e. $f_X(x) = \alpha e^{-\alpha x}$ and $f_Y(x) = (\alpha/b)e^{-\alpha x/b}$ for $x > 0$, there exists an $x_0 > 0$ (specifically $x_0 = \log(b)/(\alpha(1-1/b))$) such that $f_Y(x) > f_X(x)$ iff $x > x_0$. Hence Y takes large values with a larger probability than X (Y has a fatter tail than X but the students haven't seen this term) and is therefore more risky.

(c) (i) The scale invariance property for a premium principle π entails that $\pi(cX) = c\pi(X)$ for all risks X and $c > 0$.

(ii) Yes. Setting $Y = cX$ we have that

$$\mathbb{P}(Y \leq x) = \mathbb{P}(X \leq x/c) \implies F_Y(x) = F_X(x/c).$$

Hence

$$\begin{aligned}\pi(cX) = \int_0^\infty \left(1 - F_X(x/c)\right)^{1/b} \mathrm{d}x &= c\int_0^\infty \left(1 - F_X(x/c)\right)^{1/b} \frac{1}{c} \, \mathrm{d}x \\ &= c\int_0^\infty \left(1 - F_X(y)\right)^{1/b} \mathrm{d}y = c\pi(X)\end{aligned}$$

where we used the substitution $y = x/c$.

Answer to 4

We know from the lecture (notes) that the Bayes estimate $\hat{\theta}_B(x)$ is a minimiser of the posterior risk

$$\phi_p : y \mapsto \mathbb{E}[l(\Theta, y) \,|\, X = x].$$

For any $y \in \mathbb{R}$ we may rewrite this as follows, using that g is a strictly positive function and hence $\mathbb{E}[g(\Theta) \,|\, X = x] > 0$:

$$\begin{aligned}
\phi_p(y) &= \mathbb{E}[g(\Theta)(\Theta - y)^2 \,|\, X = x] \\
&= \mathbb{E}[g(\Theta)\Theta^2 \,|\, X = x] - 2y\mathbb{E}[g(\Theta)\Theta \,|\, X = x] + y^2\mathbb{E}[g(\Theta) \,|\, X = x] \\
&= \mathbb{E}[g(\Theta) \,|\, X = x] \left(\frac{\mathbb{E}[g(\Theta)\Theta^2 \,|\, X = x]}{\mathbb{E}[g(\Theta) \,|\, X = x]} - 2y \frac{\mathbb{E}[g(\Theta)\Theta \,|\, X = x]}{\mathbb{E}[g(\Theta) \,|\, X = x]} + y^2 \right) \\
&= \mathbb{E}[g(\Theta) \,|\, X = x] \left(\left(y - \frac{\mathbb{E}[g(\Theta)\Theta \,|\, X = x]}{\mathbb{E}[g(\Theta) \,|\, X = x]} \right)^2 - \frac{\mathbb{E}[g(\Theta)\Theta \,|\, X = x]^2}{\mathbb{E}[g(\Theta) \,|\, X = x]^2} + \frac{\mathbb{E}[g(\Theta)\Theta^2 \,|\, X = x]}{\mathbb{E}[g(\Theta) \,|\, X = x]} \right).
\end{aligned}$$

Again using that $\mathbb{E}[g(\Theta) \,|\, X = x] > 0$ it readily follows that the (unique) minimiser is

$$\hat{\theta}_B(x) = \frac{\mathbb{E}[g(\Theta)\Theta \,|\, X = x]}{\mathbb{E}[g(\Theta) \,|\, X = x]}.$$

Answer to 5

We're given that $\Theta \sim \text{Exp}(2)$ and denoting by X the number of bite incidents in a year, given $\Theta = \theta$ we have that $X \sim \text{Poisson}(\theta)$.

(a) We are asked to compute the collective premium $\mu = \mathbb{E}[X]$, which can be computed using the Tower Property directly, or (equivalently) as we know from the lecture (notes) may also be computed as $\mu = \mathbb{E}[\mu(\Theta)]$, where

$$\mu(\theta) = \mathbb{E}[X \mid \Theta = \theta] \quad \text{for } \theta > 0.$$

Since given $\Theta = \theta$, $X \sim \text{Poisson}(\theta)$ we have that $\mu(\theta) = \theta$ and hence

$$\mu = \mathbb{E}[\mu(\Theta)] = \mathbb{E}[\Theta] = \frac{1}{2},$$

where we used that $\Theta \sim \text{Exp}(2)$.

(b) For the Bühlmann credibility estimate given the observed sample $x = 4$, i.e. $\hat{\mu}_{BM}(4)$, in addition to $\mu(\theta) = \theta$ and $\mu = 1/2$ as computed in part (a) we need

$$\nu(\theta) = \text{Var}(X \mid \Theta = \theta) = \theta, \quad \nu = \mathbb{E}[\nu(\Theta)] = \mathbb{E}[\Theta] = \frac{1}{2}$$

and

$$\kappa = \text{Var}(\mu(\Theta)) = \text{Var}(\Theta) = \frac{1}{4},$$

where we used that given $\Theta = \theta$, $X \sim \text{Poisson}(\theta)$ and that $\Theta \sim \text{Exp}(2)$. Now in general for an observed sample of length n denoted $\vec{x}$:

$$\hat{\mu}_{BM}(\vec{x}) = (1-w)\mu + w\bar{x} \quad \text{where } w = \frac{n\kappa}{\nu + n\kappa}.$$

In this case we hence have that

$$w = \frac{1 \cdot 1/4}{1/2 + 1 \cdot 1/4} = \frac{1}{3}$$

and

$$\hat{\mu}_{BM}(4) = \frac{2}{3} \cdot \frac{1}{2} + \frac{1}{3} \cdot 4 = \frac{5}{3}.$$

(c) We know from the (lecture) notes that Bayes estimate $\hat{\theta}_B(4)$ for Brutus' risk parameter value under squared error loss is the mean of the posterior distribution. We have that $f_\Theta(\theta) = 2e^{-2\theta}$ for $\theta > 0$ and $f_{X|\Theta}(x|\theta) = e^{-\theta}\theta^x/x!$ for $\theta > 0$ and $x \in \mathbb{N}$. Hence by Bayes Thm for any $\theta > 0$:

$$f_{\Theta|X}(\theta|4) = c f_{X|\Theta}(4|\theta) f_\Theta(\theta) = ce^{-\theta}\frac{\theta^4}{4!}2e^{-2\theta} = \tilde{c}\theta^4 e^{-3\theta}$$

where c and $\tilde{c}$ are constants independent of θ. We recognise this as a Gamma$(5, 3)$ distribution, with mean $5/3$. Hence $\hat{\theta}_B(4) = 5/3$.

Note: if a student does not recognise this distribution then $\tilde{c}$ could be determined using that the total mass is 1 and then the mean could be computed. Requires integration by parts and more work obviously.

SECTION A

A1.

(a) Define what is meant by a *topology* on a set X.

(b) Define what is meant by saying that a function $f\colon X \to Y$ between topological spaces is *continuous*. Define what is meant by saying that f is a *homeomorphism*.

(c) Prove that the closed disc $\mathbb{D}^2 = \{x \in \mathbb{R}^2 \mid |x| \leqslant 1\}$ with the usual topology is homeomorphic to the hemisphere $\{x = (x_1, x_2, x_3) \in S^2 \mid x_3 \geqslant 0\}$.

[Here S^2 denotes the unit sphere $\{x \in \mathbb{R}^3 \mid |x| = 1\}$ with the usual topology.]

[10 marks]

Solution

(a) Given a set X, a *topology* on X is a collection τ of subsets of X with the following properties:

(i) $\emptyset \in \tau$, $X \in \tau$;

(ii) the intersection of any two subsets in τ is in τ:

$$U_1, U_2 \in \tau \Rightarrow U_1 \cap U_2 \in \tau;$$

(iii) the union of any collection of subsets in τ is in τ:

$$U_\lambda \in \tau \text{ for all } \lambda \in \Lambda \Rightarrow \bigcup_{\lambda\in\Lambda} U_\lambda \in \tau.$$

[5 marks, bookwork]

(b) $f\colon X \to Y$ is *continuous* if

$$V \text{ is open in } Y \Rightarrow f^{-1}(V) \text{ is open in } X$$

[1 marks, bookwork]

A *homeomorphism* is a continuous bijection with continuous inverse.

[2 marks, bookwork]

(c) A homeomorphism $f\colon \{\, \mathbf{x} \in S^2 \mid x_3 \geqslant 0 \,\} \to D^2$ is given by $f(x_1, x_2, x_3) = (x_1, x_2)$ with inverse $f^{-1}(y_1, y_2) = (y_1, y_2, \sqrt{1 - y_1^2 - y_2^2})$.

[2 marks, question set]

[Total: 10 marks]

A2.

(a) Define what is meant by saying that a topological space X is *path-connected.*

(b) What is meant by saying the path-connectedness is a *topological property*?

(c) Prove that path-connectedness is a topological property.

(d) Prove that

$$\{x \in \mathbb{R}^2 \mid |x - (0,1)| \leqslant 1 \text{ or } |x + (0,1)| \leqslant 1\} \subset \mathbb{R}^2$$

(with the usual topology) is path-connected.

[10 marks]

Solution

(a) A *path* from x_0 to x_1 in X is a continuous function $\sigma\colon [0,1] \to X$ with $\sigma(0) = x_0$ and $\sigma(1) = x_1$. X is said to be *path-connected* if, for each pair of points x_0, $x_1 \in X$, there is a path in X from x_0 to x_1.

[3 marks, bookwork]

(b) Saying that path-connectedness is a *topological property* means that, if $X \cong Y$ are homeomorphic topological spaces, then X is path connected if and only if Y is path-connected.

[1 marks, bookwork]

(c) To prove this, suppose that X is path-connected. Then, given two points y_0, $y_1 \in Y$ let x_0 , $x_1 \in X$ be points such that $f(x_i) = y_i$ (these points exist since f is a bijection). Since X is path-connected there is a path $\sigma\colon I \to X$ such that $\sigma(0) = x_0$ and $\sigma(1) = x_1$. Then $f \circ \sigma\colon I \to Y$ is a path in Y from y_0 to y_1 (since the composition of continuous maps is continuous). Hence, Y is path-connected. Conversely, if Y is path-connected then so is X by the same argument (interchanging the roles of X and Y).

[3 marks, bookwork]

(d) One sees that $X_+ = \{x \in X \mid x \geqslant 0\}$ and $X_- = \{x \in X \mid x \leqslant 0\}$ are path-connected. For example by considering the straight line $\sigma(t) = tx_0 + (1-t)x_1$ between two given points x_0, $x_1 \in X_-$ or x_0, $x_1 \in X_+$.

Now, since $X = X_+ \cup X_-$ and $0 \in X_+ \cap X_-$ one can find a path σ_x^0 from every point in $x \in X$ to 0. By composition and inversion one obtains a path between two arbitrary points x_0, x_1:

$$\sigma = \overline{\sigma}_{x_1}^0 * \sigma_{x_0}^0.$$

[3 marks, new]

[Total: 10 marks]

A3.

(a) Define what is meant by saying that a topological space is *Hausdorff*.

(b) Determine whether the set $S = \{a, b, c\}$ with topology $\tau = \{\emptyset, \{a, c\}, \{b\}, \{a, b, c\}\}$ is Hausdorff.

(c) Suppose that X and Y are topological spaces. Define the *product topology* on the Cartesian product $X \times Y$. [It is not necessary to prove that this is a topology.]

(d) Prove that if $\Delta \subset X \times X$ is closed in the product topology, then X is Hausdorff.

[10 marks]

Solution

(a) The topological space X is *Hausdorff* if, for each distinct pair of points x, $y \in X$, there exist open sets U and V in X such that $x \in U$, $y \in V$ and $U \cap V = \emptyset$.

[2 marks, bookwork]

(b) This space is not Hausdorff because every open subset containing a also contains c and so open subsets as required cannot be found for $x = a$ and $y = c$.

[2 marks, bookwork]

(c) The product topology on $X \times Y$ has a basis

$$\{U \times V \mid U \text{ open in } X, V \text{ open in } Y\},$$

i.e. the open sets consist of all unions of such sets.

[2 marks, bookwork]

(d) Assume Δ is closed. Hence $X \times X \setminus \Delta$ is open. By definition of the product topology this means it is a union of open rectangles, i.e. sets of the form $U \times V \subset X \times X \setminus \Delta$ with U and V both open in X. Consider $x, y \in X$ with $x \neq y$ then (x, y) lies outside the diagonal. Hence, is has to be contained in such a set

$$U \times V \subset X \times X \setminus \Delta.$$

On the one hand this implies that $x \in U$ and $y \in V$. On the other hand $U \cap V = \emptyset$, since for $x \in U \cap V$ one would have $\Delta \ni (x, x) \in U \times V$. [4 marks, question set]

[Total: 10 marks]

A4.

(a) Suppose that X_1 is a subspace of a topological space X. Define what is meant by saying that X_1 is a *retract* of X.

(b) Use the functorial properties of the fundamental group to prove that, if X_1 is a retract of X, then, for any $x_0 \in X_1$, the homomorphism induced by the inclusion map

$$i_*\colon \pi_1(X_1, x_0) \to \pi_1(X, x_0)$$

is injective.

(c) Hence prove that S^1 is not a retract of the closed disc $\mathbb{D}^2$.

[You may quote any fundamental groups that you need, without proof.]

[10 marks]

Solution

(a) $X_1 \subset X$ is a retract of X when there is a continuous map $r\colon X \to X_1$, such that $r(x) = x$ for $x \in X_1$. [3 marks, bookwork]

(b) By the functorial properties we have

$$r_* \circ i_* = (r \circ i)_* = (\mathrm{id}_{X_1})_* = \mathrm{id}_{\pi_1(X_1,x_0)}\colon \pi_1(X_1, x_0) \to \pi_1(X_1, x_0).$$

Since the composition of r_* and i_* is bijective r_* must be surjective and i_* must be injective.

[4 marks, bookwork]

(c) We have $\pi_1(S^1, x_0) = \mathbb{Z}$ and $\pi_1(D^2, x_0) = 1$, the trivial group. But there is not injective map $\mathbb{Z} \to \{1\}$. Hence, S^1 cannot be a retract of D^2. [3 marks, bookwork]

[Total: 10 marks]

SECTION B

B5.

(a) Suppose that $q\colon X \to Y$ is a surjection from a topological space X to a set Y . Define the *quotient topology* on Y determined by q. State the *universal property* of the quotient topology.

(b) Suppose that $f\colon X \to Z$ is a continuous surjection from a compact topological space X to a Hausdorff topological space Z . Define an equivalence relation $\sim$ on X so that f induces a bijection $F\colon X/{\sim} \to Z$ from the identification space $X/{\sim}$ of this equivalence relation to Z. Prove that F is a homeomorphism. [State clearly any general results which you use.]

(c) Prove that the quotient space $[0,1]\times[0,1]/\sim$ with $(0,s)\sim(1,s)$ is homeomorphic to the cylinder $[0,1]\times S^1 \subset \mathbb{R}^3$.

[15 marks]

Solution

(a) Given a topological space (X,τ) and a surjection $q\colon X \to Y$ the quotient topology on Y is given by

$$\{V \subset Y \mid q^{-1}(V) \in \tau\}.$$

The *universal property* of the quotient topology is: $f\colon Y \to Z$ to a topological space Z is continuous if and only if the composition $f\circ q\colon X \to Z$ is continuous.

[4 marks, bookwork]

(b) Given a continuous surjection $f\colon X \to Z$, define an equivalence relation on X by $x \sim x' \Leftrightarrow f(x) = f(x')$. Then we may define $F\colon X/{\sim}\to Z$ by $F([x]) = f(x)$. Since $[x] = [x'] \Leftrightarrow x \sim x' \Leftrightarrow f(x) = f(x')$ (by the definition of the equivalence relation), the function F is well-defined. Since $F([x]) = F([x']) \Leftrightarrow f(x) = f(x') \Leftrightarrow x \sim x'$ (by the definition of the equivalence relation) it follows that $[x] = [x']$ and F is injective. Since f is a surjection, $y = f(x)$ for some $x \in X$ and so $y = F([x])$. Hence F is a surjection. This shows that $F\colon X/{\sim} \to Z$ is a bijection. The map $F : X/{\sim} \to Z$ is continuous by the universal property since $F \circ q = f$ which is given as continuous, where $q\colon X \to X/{\sim}$ is the quotient map given by $q(x) = [x]$.

The space $X/{\sim}= q(X)$ is compact since it is the continuous image of a compact set. Hence F is a homeomorphism since it is a continuous bijection from a compact space to a Hausdorff space.

[7 marks, bookwork]

(c) To see this, define a surjection $f\colon I^2 \to I \times S^1$ by $f(x,y) = \big(x, \exp(2\pi i y)\big)$ where we think of S^1 as $\{\, z \in \mathbb{C} \mid |z| = 1 \,\}$ using the standard identification $\mathbb{C} \cong \mathbb{R}^2$. This function is continuous by the universal property of the product topology since the component functions are continuous. Now, $I \times S^1$ is Hausdorff (a subset of Euclidean space) and $I \times I$ is compact (a closed and bounded subset of Euclidean space). Now the result follows from (b).

[4 marks, bookwork]

[Total: 15 marks]

B6.

(a) Define what is meant by a *compact subset* of a topological space and by a *compact topological space.*

(b) Prove that, if $f\colon X \to Y$ is a continuous function of topological spaces and $K \subset X$ is a compact subset, then $f(K)$ is a compact subset of Y.

(c) Given a non-compact Hausdorff space (X, τ) consider the set $X^* = X \sqcup \{\infty\}$ and the topology

$$\tau^* = \tau \cup \{(X \setminus C) \cup \{\infty\} \mid C \subset X \text{ compact}\}.$$

Show that (X^*, τ^*) is compact.

[It is not necessary to prove that τ^* is a topology.]

[15 marks]

Solution

(a) $K \subset X$ is compact if each cover of K by open subsets of X has a finite subcover.

If X itself is a compact subset then X is a compact space.

[3 marks, bookwork]

(b) Suppose that $\mathcal{F}$ is an open cover for $f(K)$. Let $f^{-1}(\mathcal{F}) = \{f^{-1}(V) \mid V \in \mathcal{F}\}$. Then $f^{-1}(\mathcal{F})$ is an over cover for K since, given $a \in K$, $f(a) \in f(K)$ so that $f(a) \in V$ for some $V \in \mathcal{F}$. Hence $a \in f^{-1}(V)$ for some $V \in \mathcal{F}$.

Now, since K is compact, $f^{-1}(\mathcal{F})$ has a finite subcover for K, $\{\, f^{-1}(V_1), f^{-1}(V_2), \ldots, f^{-1}(V_n)\,\}$. Thus, given $b \in f(K)$, $b = f(a)$ for some $a \in K$. Then $a \in f^{-1}(V_i)$ for some i, $1 \leqslant i \leqslant n$, so that $b = f(a) \in V_i$. Hence $\{\, V_1, V_2, \ldots, V_n\,\}$ is a finite subcover of $\mathcal{F}$ for $f(K)$.

Hence $f(K)$ is compact.

[6 marks, bookwork]

(c) Consider an open cover $\mathcal{F}$ of X^*. In order to contain ∞ it has to include at least one open subset U_∞ of the form $X \setminus C \cup \{\infty\}$ where $C \subset X$ is compact. Now, $\mathcal{F}' = \{U \cap X \mid U \in \mathcal{F}\}$ is an open cover of X (since U and X are open in X^*) and hence of C.

By compactness of C a finite subcover $\{U_1 \cap X, \ldots, U_m \cap X\} \subset \mathcal{F}$ suffices to cover C. But then one has the finite subcover $\{U_\infty, U_1, \ldots, U_m\} \subset \mathcal{F}$.

[6 marks, exercise set]

[Total: 15 marks]

B7.

(a) Prove that, if the product $\sigma_0 * \tau_0$ of two paths σ_0 and τ_0 in a topological space X is defined and the paths σ_1 and τ_1 are homotopic to σ_0 and τ_0 respectively, then the product $\sigma_1 * \tau_1$ is defined and is homotopic to $\sigma_0 * \tau_0$.

(b) Explain how a continuous function $f\colon X \to Y$ induces a homomorphism $f_*\colon \pi_1(X, x_0) \to \pi_1(X, f(x_0))$. You should indicate why f_* is well-defined and why it is a homomorphism.

(c) Prove that, for topological spaces X and Y with points $x_0 \in X$, $y_0 \in Y$, there is an isomorphism of groups

$$\pi_1(X \times Y, (x_0, y_0)) \cong \pi_1(X, x_0) \times \pi_1(Y, y_0).$$

[15 marks]

Solution

(a) Given homotopic paths $H\colon \sigma_0 \sim \sigma_1$ and $K\colon \tau_0 \sim \tau_1$ such that $\sigma_0 * \tau_0$ is defined. Then $_1(1) = \sigma_0(1) = \tau_0(0) = \tau_1(0)$ and so the product $\sigma_1 * \tau_1$ is defined.

Suppose that $H\colon \sigma_0 \sim \sigma_1$ and $K\colon \tau_0 \sim \tau_1$. Then we may define a homotopy $L\colon \sigma_0 * \tau_0 \sim \sigma_1 * \tau_1$ by

$$L(s,t) = \begin{cases} H(2s,t) & \text{for } 0 \leqslant s \leqslant 1/2 \text{ and } t \in I, \\ K(2s-1,t) & \text{for } 1/2 \leqslant s \leqslant 1 \text{ and } t \in I. \end{cases}$$

This is well defined since, for $s = 1/2$, $H(1,t) = x_1 = K(0,t)$. In addition, L is continuous by the Gluing Lemma since $[0,1/2] \times I$ and $[1/2,1] \times I$ are closed subsets of I^2

[5 marks, bookwork]

(b) The function f_* is defined by $f_*([\sigma]) = [f \circ \sigma]$. It is well-defined since, if $[\sigma_0] = [\sigma_1]$ then $\sigma_0 \sim \sigma_1$ and so there exists a homotopy $H\colon \sigma_0 \sim \sigma_1$. Then $f \circ H\colon I^2 \to Y$ gives a homotopy $f \circ \sigma_0 \sim f \circ \sigma_1$ and so $[f \circ \sigma_0] = [f \circ \sigma_1]$.

To see that f_* is a homomorphism suppose that $[\sigma]$, $[\tau] \in \pi_1(X, x_0)$. Then

$$f_*([\sigma][\tau]) = f_*([\sigma * \tau]) = [f \circ (\sigma * \tau)]$$

and

$$f_*([\sigma])f_*([\tau]) = [f \circ \sigma][f \circ \tau] = [(f \circ \sigma) * (f \circ \tau)]$$

and by writing out the formulae we see that $f \circ (\sigma * \tau) = (f \circ \sigma) * (f \circ \tau)\colon I \to Y$. Hence, $f_*([\sigma][\tau]) = f_*([\sigma])f_*([\tau])$.

[5 marks, bookwork]

(c) Let $p_1\colon X \times Y \to X$ and $p_2\colon X \times Y \to Y$ be the projection maps. The function

$$\pi_1\big(X \times Y, (x_0, y_0)\big) \to \pi_1(X, x_0) \times \pi_1(Y, y_0)$$

given by $\alpha \mapsto \big((p_1)_*(\alpha), (p_2)_*(\alpha)\big)$ is an isomorphism. To see this we write down the inverse. Given a loop σ_1 in X based at x_0 and a loop σ_2 in Y based at y_0 then we may define a loop σ

in $X \times Y$ based at (x_0, y_0) by $\sigma(s) = \big(\sigma_1(s), \sigma_2(s)\big)$. Then $([\sigma_1], [\sigma_2]) \mapsto [\sigma]$ is well-defined and provides the necessary inverse. Hence the given function is an isomorphism of groups.

[5 marks, question set]

[Total: 15 marks]

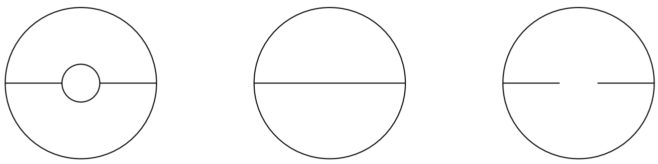
B8.

(a) Define what is meant by the *path-components* of a topological space. [You may assume the definition of a path and properties of paths.]

(b) Prove that a continuous map of topological spaces $f\colon X \to Y$ induces a map $f_*\colon \pi_0(X) \to \pi_0(Y)$ between the sets of path-components, taking care to prove that your function is well-defined. Prove that if f is a homeomorphism then f_* is a bijection.

(c) A pair of distinct points $\{p, q\}$ in a path-connected topological space X is called a cut-pair of type n when the subspace $X \setminus \{p, q\}$ has n path-components. Prove that a homeomorphism $f : X \to Y$ induces a bijection between the subsets of cut-pairs of type n for every $n \in \mathbb{N}$.

(d) Hence show, using cut-pairs of type 3 or otherwise, that no two of the following subspaces of $\mathbb{R}^2$ with the usual topology are homeomorphic.



(a) (b) (c)

[15 marks]

Solution

(a) Define an equivalence relation on X by $x \sim x'$ if and only if there is a path in X from x to x'. Then the path-components of X are the equivalence classes.

[2 marks, bookwork]

(b) Suppose that $f\colon X \to Y$ is a continuous map. Then this induces a function $f\colon \pi_0(X) \to \pi_0(Y)$ by $f([x]) = [f(x)]$. This is well-defined because $[x] = [x']$ implies that $x \sim x'$ so that there is a path $\sigma\colon [0,1] \to X$ in X from x to x'. Then $f \circ \sigma\colon [0,1] \to Y$ is a path in Y from $f(x)$ to $f(x')$ and so $[f(x)] = [f(x')]$.

[3 marks, bookwork]

If f is a homeomorphism then f_* is a bijection since the inverse $g = f^{-1}\colon Y \to X$ induces a function $g_* : \pi_0(Y) \to \pi_0(X)$ inverse to f_* since $g_*(f_*([x])) = [g(f(x))] = [x]$ and $f_*(g_*([y])) = [y]$.

[2 marks, bookwork]

(c) Suppose that $f\colon X \to Y$ is a homeomorphism and $\{p, q\}$ is a pair of distinct points in X. Then f induces a homeomorphism $X \setminus \{p, q\} \to Y \setminus \{f(p), f(q)\}$ and this induces a bijection $f_*\colon \pi_0(X \setminus \{p, q\}) \to \pi_0(Y \setminus \{f(p), f(q)\})$. Hence $\{p, q\}$ is a cut-pair of type n in X if and only if $\{f(p), f(q)\}$ is a cut-pair of type n in Y.

[3 marks, exercise set]

(d) In space (i) there are two cut-pairs of type 3 (the intersection points of the line segments and the inner or out circle respectively). In space (ii) there is a unique cut-pair of type 3 (the two points at the ends of the diameter). In space (iii) there are infinitely many cut-pairs of type 3 (picking two arbitrary points on the radial line segments). [5 marks, new]

[Total: 15 marks]

END OF EXAMINATION PAPER