UNIVERSITY OF WARWICK

Summer Examinations 2015/16

Statistical Techniques B

Time Allowed: 1.5 Hours, plus 15 minutes reading time during which notes may be made (on the question paper) BUT NO ANSWERS MAY BE BEGUN.

Answer ALL SEVEN questions. Answer questions 1-4 in one booklet and questions 5-7 in a separate booklet.

Statistical Tables and a Formula Sheet are provided. Approved pocket calculators are allowed.

Read carefully the instructions on the answer book provided and make sure that the particulars required are entered on each answer book. If you answer more questions than are required and do not indicate which answers should be ignored, we will mark the requisite number of answers in the order in which they appear in the answer book(s): answers beyond that number will not be considered.

---

1. In a survey regarding congestion charges, 85% of taxi/bus drivers and 78% of public transport commuters were in favour of congestion charges. In contrast only 45% of regular private car users were in favour of the charges. In the survey 15% of interviewees were taxi/bus drivers, 50% were public transport commuters and the remainder were private car users.

(a) What is the probability of being a taxi/bus driver and in favour of congestion charges? **(2 marks)**

(b) What proportion of interviewees are in favour of congestion charges? **(3 marks)**

(c) What is the probability an interviewee in favour of congestion charges is either a taxi/bus driver or a public transport user? **(3 marks)**

(d) Given that half of people in each type of transport group are female and that females are 20% more likely to be in favour of congestion charges than males, calculate the probability of being a female given you are in favour of congestion charges. **(4 marks)**

(Continued overleaf)

2. Consider the following bivariate probability distribution for the variables X and Y:

		X		
		1	3	5
	-1	0.1	0.0	0.1
Y	0	0.1	0.0	0.1
	1	0.2	0.2	0.2

(a) Calculate $V(X)$ and $V(Y)$. **(4 marks)**

(b) Calculate $cov(X, Y)$. **(4 marks)**

(c) Calculate $E(X|X \neq 3)$, $E(Y|X \neq 3)$. **(4 marks)**

(d) Calculate $cov(X, Y|X \neq 3)$. **(4 marks)**

---

3. Scores in a certain statistics test are known to follow a normal distribution. The test is given to a very large group of students and the mean in the test is 54.2% and the standard deviation is 15%.

(a) If students scoring below 40% are required to attend remedial lessons, what proportion of the class will attend these remedial lessons? **(4 marks)**

(b) A random sample of two individuals are taken from the class. What is the probability the difference in their marks is more than 5% points? **(5 marks)**

(c) In actual fact the lecturer concerned is told that ideally around 25% of students should get above 70% and only 10% should score below 40%. Suggest a linear adjustment of the marks in the test (of the form $Newmark = a + bOldMark$) to ensure the lecturer gets the expected distribution. **(8 marks)**

(Continued overleaf)

4. A random sample of 28 individuals who had been on a speed awareness course had an average car speed, down a certain stretch of the M69, of 74mph, with a sample standard deviation of 8mph. An independent random sample of 62 individuals who had not been on the course had (on the same stretch of motorway) an average speed of 77mph and a sample standard deviation of 12mph. It is reasonable to assume that motorway speeds follow a normal distribution.

(a) At the 5% significance level test the hypothesis that there is no difference in the variance of speed of the car along the M69 between those who had been on the speed awareness programme and those who had not. **(5 marks)**

(b) At the 5% significance level test the hypothesis that there is no difference in the average speed in cars down the M69 according to whether the individual went on a speed awareness course, against a suitable 1-sided alternative. **(5 marks)**

(c) At the 5% significance level, calculate the approximate power of the test in (b), given that the true difference in the speed between individuals who went on the speed awareness programme and those who did not is -5mph. **(5 marks)**

(d) On a suitable diagram represent the power, significance level and probability of a type II error based on part (c). **(5 marks)**

.....

5. In a fairground there is a roulette wheel game with 20 numbers, labelled 1-20. The pay-off from playing this game is as follows: if the ball falls in the numbers: 1-4: you get nothing; 5-8: you get nothing; but you can roll the ball one more time (but get nothing for numbers 1-8); 9-12: you get 50p; 13-16: you get £1; and 17-20: you get £4.

(a) Write out the probability function for this game. **(4 marks)**

(b) What price would you charge to play the game to ensure you had a 90% probability of making at least £0.5 profit per player, assuming that you get 100 players playing the game? **(8 marks)**

(Continued overleaf)

6. A survey of 82 males found that support for pulling out of Europe was 42%. An independent random sample of 105 females found that support for pulling out of Europe was 38%.

(a) Construct the 95% confidence interval for the difference between males and females in the population proportion support for pulling out of Europe. **(4 marks)**

(b) Construct the 95% confidence interval for the population proportion support for pulling out of Europe for males and females combined. **(3 marks)**

(c) Calculate the sample size needed to ensure that the sample mean and true mean for support for pulling out of Europe for males and females combined differ by no more than 2 percentage points with 90% confidence. **(5 marks)**

---

7. (a) The unemployment rate in a particular country is believed to be 5%. A random sample of 100 villages each with 40 people found the following distribution, in terms of the number unemployed in each village:

	0	1	2	3	4	≥ 5
No. of villages	20	20	20	20	10	10

At the 1% significance level, test the hypothesis of that the distribution of unemployment across the villages is as expected. **(6 marks)**

(b) There is a speed reading course available for free on the internet. Six individuals are tested (and scored out of 20) on their understanding of a text passage before and after taking the speed reading course and the results are given in the table below:

	1	2	3	4	5	6
Before	18	14	15	10	12	9
After	18	12	14	6	10	8

At the 5% significance level use the sign test to determine whether the speed reading affected understanding of the text. **(5 marks)**

(End)