PRIFYSGOL ABERYSTWYTH UNIVERSITY

ADRAN MATHEMATEG / DEPARTMENT OF MATHEMATICS

ARHOLIADAU SEMESTER 2 / SEMESTER 2 EXAMINATIONS

MAI / MAY 2020

MA11310 - Statistics

The questions on this paper are written in English.

If you have questions about the paper during the exam, contact the module co-ordinator, Dr Jukka Kiukas, on jek20@aber.ac.uk.

Amser a ganiateir - 2 awr
Maen rhaid cyflwyno eich atebion erbyn 11:30 (amser y DU).

- Rhoddir marciau llawn am atebion cyflawn i bob cwestiwn.
- Dylai myfyrwyr roi cynnig ar bob cwestiwn **ar bapur**.
- Dylai myfyrwyr **yna** gyflwyno eu hatebion ar safle Blackboard y modiwl hwn.
- Mae tablau ystadegol ar gael ar Blackboard.

Time allowed - 2 hours
Submission must be completed by 11:30 (UK time).

- Full marks will be given for complete answers to all questions.
- Students should attempt all questions **on paper**.
- Students should **then** submit their answers on the Blackboard site for this module.
- Statistical tables are available on Blackboard.

The blackboard test will present blank text boxes and multiple choice options for answer input.
The type of each answer is indicated in the corresponding text box in this paper.

Questions

---

1. The random variable X has mean $-\frac{1}{2}$ and standard deviation 3.

(a) (i) $\mathbb{E}[6X+5] =$ [*integer*] and (ii) $\mathrm{Var}(6X+5) =$ [*integer*] .

(b) The bulk of values of $6X+5$ lies between [*integer*] and [*integer*] .
(Note: the first number should be smaller than the second.) [5 marks]

---

2. The random variable X has a Pascal, $\mathrm{Pasc}(r,p)$, distribution with mean 8 and variance 24. Then $r =$ [*integer*] , $p =$ [*fraction*] , and $\mathbb{P}(X=4) =$ [*value to 3dp*] . [5 marks]

---

3. In each of the following cases, decide whether or not the Binomial distribution model is appropriate for the random variable X. If you think it is, identify the distribution as $\mathrm{Bin}(n,p)$ with appropriate values for n and p; if not, suggest a sensible alternative.

(a) 98% of electronic components assembled by a production line are functional. The line assembles components one by one until 800 components are obtained; X is the number of functional components assembled in the process.
Answer: [*Bin(n,p) / alternative*] .

(b) A fair coin is flipped, and placed on a table if it lands "tails" up. The same is repeated for 5 identical coins; X is the number of coins on the table at the end.
Answer: [*Bin(n,p) / alternative*] .

(c) A fair coin is flipped, and placed on a table if it lands "tails" up. The same is repeated for further identical coins until there are 5 coins on the table; X is the number of coin flips in this process.
Answer: [*Bin(n,p) / alternative*] . [5 marks]

---

4. The random variable X has the Binomial distribution $\mathrm{Bin}(50, \frac{1}{5})$.

(a) Then $\mathbb{P}(9 < X \leq 14) = \mathbb{P}(X \geq a) - \mathbb{P}(X \geq b)$, where
$a =$ [*integer*] and $b =$ [*integer*] .
Hence $\mathbb{P}(9 < X \leq 14) =$ [*value to 3dp*] .

(b) Let $Y = 50 - X$. Then $Y \sim \mathrm{Bin}(50,p)$, where $p =$ [*fraction*] . Hence
$\mathbb{P}(Y < 40) = \mathbb{P}(X \geq c) =$ [*value to 3dp*] , where $c =$ [*integer*] . [5 marks]

---

5. Questions 5-7 have the following context:

A statistic S has probability density function (pdf)

$$f(s) = \frac{3s^2}{\theta^3}, \quad 0 < s < \theta,$$

where $\theta > 0$ is an unknown parameter.

Which of the following integrals is equal to $\mathbb{P}(\frac{\theta}{2} < S < \theta)$?

○ $\int_S^{2S} \frac{3s^2}{\theta^3} d\theta$ ○ $\int_\theta^{2\theta} \frac{3s^2}{\theta^3} ds$ ○ $\int_{\theta/2}^{\theta} \frac{3s^2}{\theta^3} ds$ ○ $\int_{S/2}^{S} \frac{3s^2}{\theta^3} d\theta$. [1 mark]

6. The value of $\mathbb{P}(\frac{\theta}{2} < S < \theta)$ is [*value to 3dp*] . [2 marks]

7. The value of $\mathbb{P}(\frac{\theta}{2} < S < \theta)$ gives the confidence coefficient of the confidence interval

○ $(\theta, 2\theta)$ ○ $(S, 2S)$ ○ $(\frac{S}{2}, S)$ ○ $(\frac{\theta}{2}, \theta)$. [2 marks]

8. Let $\mathrm{Var}(U) = 12$, $\mathrm{Var}(V) = 14$ and $\mathrm{Var}(U + V) = 20$. Then the covariance between U and V is [*integer*] , and $\mathrm{Cov}(U + 2V, U - 3V) =$ [*integer*] . [5 marks]

9. A construction project consists of 3 independent stages. The costs (in £1000) of the stages are independent and Normally distributed with the respective means 8, 22 and 15, and standard deviations 3, 7 and 5.

(a) The total cost of the project has expected value [*integer*] and standard deviation [*value to 3dp*] .

(b) The probability that the total cost exceeds £50000 is equal to [*value to 2dp*] .

[5 marks]

10. A customer service receives calls randomly at an average rate of λ per hour.

(a) Let N be the number of calls received in 2 hours. If $\lambda = 25$, then the distribution of N is [*identify distribution*] , and $\mathbb{P}(N \geq 62) =$ [*value to 3dp*] .

(b) Consider a test of hypotheses $H_0 : \lambda = 25$ versus $H_1 : \lambda > 25$. Given that 62 calls were received in a 2-hour period, the P-value is [*value to 3dp*] . Hence we [*can/cannot*] reject the null hypothesis at 5% significance level. [5 marks]

11. The remaining questions 11-17 (total 10 marks) have the following context:

It is suspected that more than 60% of a large population is infected with a certain disease. In order to find evidence for this (and to exclude the possibility that only 60% is infected), two thousand randomly selected people were tested, and 1238 of them were found infected. A hypothesis test was then carried out.

The relevant *parameter* p is

- ◯ the proportion of infected among the 2000 tested
- ◯ the number of infected among the 2000 tested
- ◯ the proportion of infected in the population
- ◯ the number of people tested for infection. [1 mark]

12. The appropriate test statistic X for a Binomial hypothesis test is [*description in a few words*], and its observed value is [*integer*]. [1 mark]

13. The alternative hypothesis is

◯ $X = 0.6$ ◯ $p = 0.6$ ◯ $X < 1238$ ◯ $X \geq 1238$ ◯ $p \geq 0.6$
◯ $p < 0.6$ ◯ $p > 0.6$ ◯ $X = 1238$ ◯ $p \leq 1238$. [1 mark]

14. Assuming the null hypothesis, the test statistic X has mean [*integer*] and variance [*integer*]. [1 mark]

15. The P-value is the probability that

◯ $X = 0.6$ ◯ $p = 0.6$ ◯ $X < 1238$ ◯ $X \geq 1238$ ◯ $p \geq 0.6$
◯ $p < 0.6$ ◯ $p > 0.6$ ◯ $X = 1238$ ◯ $p \leq 1238$. [1 mark]

16. The distribution of the appropriate large sample approximation of X is

◯ Pascal ◯ Binomial ◯ Normal ◯ Poisson ◯ Geometric. [1 mark]

17. Using the large sample approximation, the P-value is found to be [*value to 3dp*], and hence the null hypothesis [*can/cannot*] be rejected at 5% significance level. [4 marks]