Curtin University

MATH1019 Linear Algebra and Statistics for Engineers

Test, S2 2018; Time Allowed: **1** Hour

This paper contains x pages (including this cover sheet) and 5 questions

Write your answers in the spaces provided. Rough working can be put on the backs of the previous page. Write your name and student number on this cover sheet. If pages become separated write your name on all separated sheets. If you need more paper, please ask.

NAME:____

STUDENT NUMBER:________________________________

Question 1. Ten students have just had a statistics test and their results are

75	66	49	23	89	55	60	72	80	62

(a) Find the mean and variance of their grades. (2 marks)

Solution.

$$\text{mean} = y = (75 + 66 + 49 + 23 + 89 + 55 + 60 + 72 + 80 + 62)/10 = 63.1$$

$\boxed{\text{1 mark}}$

$$\text{variance} = ((75 - y)^2 + (66 - y)^2 + (49 - y)^2 + (23 - y)^2 + (89 - y)^2 + (55 - y)^2$$
$$+(60 - y)^2 + (72 - y)^2 + (80 - y)^2 + (62 - y)^2)/9 = 340.9889.$$

$$\sigma = \sqrt{340.9889} = 18.46588$$

$\boxed{\text{1 mark}}$

(b) Find the lower quartile Q_1, median Q_2 and upper quartile Q_3 of the above data set.

(6 marks)

Solution. First we rank the data in ascending order 23, 49, 55, 60, 62, 66, 72, 75, 80, 89. $n = 10$. $\boxed{\text{1 mark}}$

Since $(n + 1)/4 = 11/4 = 2\frac{3}{4}$, $Q_1 = 49 + \frac{3}{4}(55 - 49) = 53.5$ $\boxed{\text{2 mark}}$

Median $= (62 + 66)/2 = 64$. $\boxed{\text{1 mark}}$

Since $(n + 1) \times 3/4 = 33/4 = 8\frac{1}{4}$, $Q_3 = 75 + \frac{1}{4}(80 - 75) = 76.25$. $\boxed{\text{2 mark}}$

Alternative solution.

Rank the data in ascending order

23, 49, 55, 60, 62, 66, 72, 75, 80, 89. $n = 10$. $\boxed{\text{1 mark}}$

median $= (62 + 66)/2 = 64$. This divides the data set into two halves. $\boxed{\text{1 mark}}$

$Q_1 =$ median of the lower half, i.e. $Q_1 = 55$. $\boxed{\text{2 mark}}$

$Q_3 =$ median of the upper half, i.e. $Q_3 = 75$. $\boxed{\text{2 mark}}$

(c) Draw the boxplot of the above data set and indicating clearly the whiskers and outlier(s), if any. (4 marks)

Solution.

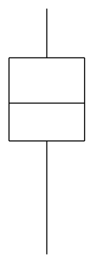
$IQR = 76.25 - 53.5 = 22.75$. $\boxed{\text{1 mark}}$

Find any outliers. $Q_1 - 1.5IQR = 19.375$; $Q_3 + 1.5IQR = 110.375$. $\boxed{\text{1 mark}}$

Thus, no outliers. $\boxed{\text{1 mark}}$

min $= 23$; max $= 89$.

Boxplot: $\boxed{\text{1 mark}}$



Alternative solution. If the alternative method for Q_1, Q_2, Q_3 is used in (b), then

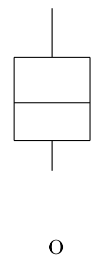
$IQR = 75 - 55 = 20$. $\boxed{\text{1 mark}}$

Find any outliers. $Q_1 - 1.5IQR = 25$; $Q_3 + 1.5IQR = 105$. $\boxed{\text{1 mark}}$

Thus, there is an outlier 23. $\boxed{\text{1 mark}}$

min $= 23$; max $= 89$

Boxplot: $\boxed{\text{1 mark}}$



Question 2. A survey result shows that hotline service receives on average 4 incoming phone calls in one period.

(a) Let X be a random number representing the number of incoming phone calls. What probability distribution does X satisfy? (1 mark)

Solution.
X satisfies a Poisson distribution with $\lambda = 4$. $\boxed{\text{1 mark}}$

(b) Find the probability that the hotline service receives 2 or fewer incoming calls in one period, i.e. $P(X \leq 2)$. (3 marks)

Solution.

$$\begin{aligned} P(X \leq 2) &= P(X = 0) + P(X = 1) + P(X = 2) \qquad \boxed{\text{1 mark}} \\ &= e^{-4}\left(\frac{4^0}{0!} + \frac{4^1}{1!} + \frac{4^2}{2!}\right) \qquad \boxed{\text{1 mark}} \\ &\approx 0.0183(1 + 4 + 8) = 0.2379. \qquad \boxed{\text{1 mark}} \end{aligned}$$

(c) What is the probability that the service receives 2 or fewer incoming calls in 2 consecutive periods? (4 marks)

Solution.
In this case X satisfies a Poisson distribution with $\lambda = 4 \times 2 = 8$. $\boxed{\text{1 mark}}$

$$\begin{aligned} P(X \leq 2) &= P(X = 0) + P(X = 1) + P(X = 2) \qquad \boxed{\text{1 mark}} \\ &= e^{-8}\left(\frac{8^0}{0!} + \frac{8^1}{1!} + \frac{8^2}{2!}\right) \qquad \boxed{\text{1 mark}} \\ &= e^{-8}(1 + 8 + 32) = 0.0138. \qquad \boxed{\text{1 mark}} \end{aligned}$$

(d) Using the probability found in part (b), find the probability that any three of the five chosen periods have 2 or less incoming calls. (4 marks)

Solution.
Let X be the number of periods with 2 or fewer incoming calls. Then, we see that

$X \sim \text{Bin}(n = 5, p = 0.2379)$, where $p = 0.2379$ is from the result found in (b). $\boxed{\text{2 mark}}$

Thus, we have

$$P(X = 3) = \binom{5}{3} 0.2379^3 (1 - 0.2379)^2 = \frac{5!}{3! \times 2!} 0.2379^3 \times 0.7621^2 \qquad \boxed{\text{1 mark}}$$

$$= 0.07820009 \qquad \boxed{\text{1 mark}}$$

Question 3. You are asked by your boss to design a game having three possible outcomes -1 (loss), 0 (break-even) and 1 (win) for its player with probabilities p_1, p_2 and p_3 respectively. Your boss would like to see that the expected value and variance of the random outcome are respectively -0.1 and 0.8 (non-fair game). How do you choose p_1 and p_3 if $p_2 = 0.19$? (6 marks)

Solution. From the given conditions we see that $p_i, i = 1, 2, 3$ satisfy

$$p_1 + p_2 + p_3 = 1, \qquad \boxed{\text{1 mark}}$$
$$\mu = -p_1 + 0 \times p_2 + p_3 = -0.1, \qquad \boxed{\text{1 mark}}$$
$$\sigma^2 = (-1 + 0.1)^2 p_1 + (0 + 0.1)^2 p_2 + (1 + 0.1)^2 p_3 = 0.81 p_1 + 0.01 p_2 + 1.21 p_3 = 0.8.$$

$\boxed{\text{1 mark}}$

From the middle eq we have $p_3 = p_1 - 0.1$. So, using this and $p_2 = 0.19$, we have

$$0.81 p_1 + 0.01 p_2 + 1.21(p_1 - 0.1) = 1 \qquad \boxed{\text{1 mark}}$$

$$\Rightarrow \ (0.81 + 1.21) p_1 + 0.00019 - 0.121 = 0.8.$$

This gives $p_1 = (0.8 - 0.0019 + 0.121)/(0.81 + 1.21) = 0.455.$ $\boxed{\text{1 mark}}$
Therefore, $p_3 = 0.355.$ $\boxed{\text{1 mark}}$

Question 4. Curtin University claims that the mean salary of its graduates is \$67K (\$67,000). You think Curtin over-estimates the average salary with the aim of attracting more students. You then conduct a survey by randomly sampling 5 Curtin's graduates' salaries and find that the mean salary is $\bar{x} = \$55\text{K}$ with the standard deviation $s = \$3\text{K}$.

(a) Perform a test of hypothesis at the 5% significance level with the intent to prove that Curtin University over-estimates the average salary. (6 marks)

Solution. We define hypotheses (in thousand dollars)
$H_0 : \mu = 67,$ $\boxed{\text{1 mark}}$
$H_A : \mu < 67.$ $\boxed{\text{1 mark}}$
Test statistic

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{55 - 67}{3/\sqrt{5}} = -8.94427.$$

$\boxed{\text{1 mark}}$

Degree of freedom is $df = 4$. Since the t-table is for $P(T \geq t)$, we need to find $P(T \geq 8.94427)$ with $df = 4$. $\boxed{\text{1 mark}}$

Since $8.994427 > 8.610$, we see that $p-$ value $< 0.0005 < 0.05$. $\boxed{\text{1 mark}}$

Therefore, at the 5% significant level, we REJECT H_0. $\boxed{\text{1 mark}}$

(b) If we accept Curtin University's claim of the graduate's mean salary $\mu = \$67$K with the standard deviation $\sigma = \$1$K, how large a sample is required if we want a 95% confidence interval for the mean μ to have a margin of error of ± 0.05? (4 marks)

Solution.

$\alpha = 0.05$ and need to find $z_{\alpha/2} = z_{0.025}$. $\boxed{\text{1 mark}}$

Check the $N(0, 1)$ table we see that $P(z \le -1.96) \approx 0.025$. Thus, $z_{\alpha/2} = 1.96$.

$\boxed{\text{1 mark}}$

So, set $\frac{0.05}{\sigma/\sqrt{n}} = 1.96$. $\boxed{\text{1 mark}}$

This give

$$n = \left(\frac{1.96 \times 1}{0.05}\right)^2 = 1537.$$

$\boxed{\text{1 mark}}$

Question 5. You are an engineer of a bottled water supplier and you've been asked to set up a machine which fills 500ml water into a bottle. It is known that standard deviation of the volume filled into a bottle by the machine is $\sigma = 20$ml. From time to time The Australian Competition & Consumer Commission (ACCC) randomly chooses 30 bottles of your products and calculate the average volume $\bar{x}$ of the bottle contents. What us the minimum mean water volume μ in a bottle filled by the machine you need to set in order that $P(\bar{x} > 500\text{ml}) > 0.8$? (5 marks)

Solution.

Using the CLT,

$$P(\bar{x} > 500) = P(\frac{\bar{x} - \mu}{\sigma/\sqrt{30}} > \frac{500 - \mu}{\sigma/\sqrt{30}}) \quad \boxed{\text{1 mark}}$$

$$= 1 - P(z \le \frac{500 - \mu}{20/\sqrt{30}}) > 0.8.$$

From this we have

$$P(z \le \frac{500 - \mu}{20/\sqrt{30}}) < 0.2. \quad \boxed{\text{1 mark}}$$

From the $N(0, 1)$ table we see that $P(z \le -0.85) = 0.1977 < 0.2$. $\boxed{\text{1 mark}}$

Thus, we set $\frac{500-\mu}{20/\sqrt{30}} = -0.85$. $\boxed{\text{1 mark}}$

From this,

$$\mu = 500 + 0.85 \times 20/\sqrt{30} = 503.1038.$$

$\boxed{\text{1 mark}}$

END OF TEST PAPER