# Curtin University

## MATH1019 Linear Algebra and Statistics for Engineers

Test, S1 2019; Time Allowed: **1** Hour

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

*Write your answers in the spaces provided. Write your name and student number on this cover sheet. If pages become separated write your name on all separated sheets. A blank page is attached should you require additional space, however if you need more paper than this, please ask.*

**NAME:**________________________________________

**STUDENT NUMBER:**________________________________

---

**Question 1.** A shop would like to estimate its average number of customers per hour. After a simple random sample process, the shop has found the following sample of custom number per hour for 10 hours.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 109 | 66 | 49 | 23 | 89 | 99 | 70 | 88 | 92 | 85 |

(a) Find the mean and standard deviation of the sample. (3 marks)

**Solution.**

$$\text{mean} = y = (109 + 66 + 49 + 23 + 89 + 99 + 70 + 88 + 92 + 85)/10 = 77.$$

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

$$\text{variance} = ((109 - y)^2 + (66 - y)^2 + (49 - y)^2 + (23 - y)^2 + (89 - y)^2 + (99 - y)^2$$
$$+ (70 - y)^2 + (88 - y)^2 + (92 - y)^2 + (85 - y)^2)/9 = 659.1111.$$

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

$$\sigma = \sqrt{659.1111} = 25.67316$$

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

(b) Find the five number presentation of the above data set.

(7 marks)

**Solution.** First we rank the data in ascending order 23, 49, 66, 70, 85, 88, 89, 92, 99, 109. $n = 10$. $\boxed{\text{1 mark}}$

Since $(n+1)/4 = 11/4 = 2\frac{3}{4}$, $Q_1 = 49 + \frac{3}{4}(66-49) = 61.75$ $\boxed{\text{2 mark}}$

Median $= (85+88)/2 = 86.5$. $\boxed{\text{1 mark}}$

Since $(n+1) \times 3/4 = 33/4 = 8\frac{1}{4}$, $Q_3 = 92 + \frac{1}{4}(99-92) = 93.75$. $\boxed{\text{2 mark}}$
So, the five number presentation is 23, 61.75, 86.5, 93.75, 109. $\boxed{\text{1 mark}}$

**Alternative solution.**
Rank the data in ascending order
23, 49, 66, 70, 85, 88, 89, 92, 99, 109. $n = 10$. $\boxed{\text{1 mark}}$
Median $= (85+88)/2 = 86.5$. This divides the data set into two halves. $\boxed{\text{1 mark}}$
$Q_1 =$ median of the lower half, i.e. $Q_1 = 66$. $\boxed{\text{2 mark}}$
$Q_3 =$ median of the upper half, i.e. $Q_3 = 92$. $\boxed{\text{2 mark}}$
So, the five number presentation is 23, 61.75, 86.5, 93.75, 109. $\boxed{\text{1 mark}}$

(c) Use the results in part (b) to draw the boxplot of the above sample, indicating clearly the whiskers and outlier(s), if any. (4 marks)

**Solution.**
$IQR = 93.75 - 61.75 = 32$. $\boxed{\text{1 mark}}$
Find any outliers. $Q_1 - 1.5IQR = 61.75 - 1.5 \times 32 = 13.75$; $Q_3 + 1.5IQR = 93.75 + 1.5 \times 32 = 141.75$. $\boxed{\text{1 mark}}$
Thus, no outliers, since min $= 23$ and max $= 109$. $\boxed{\text{1 mark}}$
Boxplot: draw reasonably a boxplot indicating the box from Q1 to Q3, the median and the whiskers from min to Q1 and from Q3 to max. $\boxed{\text{1 mark}}$

**Alternative solution.** If the alternative method for $Q_1, Q_2, Q_3$ is used in (b), then
$IQR = 92 - 66 = 26$. $\boxed{\text{1 mark}}$
Find any outliers. $Q_1 - 1.5IQR = 66 - 1.5 \times 26 = 27$; $Q_3 + 1.5IQR = 92 + 1.5 \times 26 = 131$. $\boxed{\text{1 mark}}$
Thus, there is an outlier 23. $\boxed{\text{1 mark}}$
Boxplot: draw a reasonable box-plot with whiskers and the outlier 23. $\boxed{\text{1 mark}}$

**Question 2.** A company offers you the job to design a payoff function for a casino game machine. It is known that the game has three outcomes A, B and C and their respective probabilities are 0.2, 0.5 and 0.3. The payoff function for the game player equals $P_1$ for A, 0 for B and $P_2$ for C. The company requires that the expected value of the player's payoff is -\$0.1 (unfair game). If it is also required that $P_1 + P_2 = -0.04$. Help the company to determine $P_1$ and $P_2$. Also, find the variance of the payoff function. (4 marks)

**Solution.** The expected payoff and

$$0.2P_1 + 0.5 \times 0 + 0.3P_2 = -0.1, \qquad \boxed{\text{1 mark}}$$
$$P_1 + P_2 = -0.04$$

Solving this this system gives $P_1 = 0.88$ and $P_2 = -0.92$. $\boxed{\text{2 mark}}$
Using this solution we find $Var = 0.2 * 0.88^2 + 0.3 * (-0.92)^2 - 0.1^2 = 0.3988$. $\boxed{\text{1 mark}}$

**Question 3.** Kellogg's produces boxed breakfast cereals. The weight of cereals $w$ in each box is a random variable satisfying the normal distribution with the mean $\mu = 500g$ and standard deviation $\sigma$. From time to time The Australian Competition & Consumer Commission (ACCC) randomly chooses 20 boxes of the product and calculate the average weight $\bar{w}$ of the box contents. If ACCC finds $P(\bar{w} < 485g) > 0.1$, the company will be fined. What is the maximum standard deviation $\sigma$ of $w$ so as for Kellogg's to avoid a fine by ACCC? (5 marks)

**Solution.**
Using the CLT,

$$P(\bar{w} < 485) = P(\frac{\bar{w} - \mu}{\sigma/\sqrt{20}} < \frac{485 - 500}{\sigma/\sqrt{20}}) \qquad \boxed{\text{1 mark}}$$

$$= P(z < \frac{-15}{\sigma/\sqrt{20}}) < 0.1.$$

From the $N(0, 1)$ table we see that $P(z \le -1.29) = 0.0985 < 0.1$. $\boxed{\text{1 mark}}$
Thus, we set $\frac{-15}{\sigma/\sqrt{20}} \le -1.29$. $\boxed{\text{1 mark}}$
From this,

$$\sigma \le 15 * \sqrt{20}/1.29 = 52.001581.$$

$\boxed{\text{2 mark}}$

**Question 4.** This is a continuation of Question 3 above. ACCC wants to assess independently whether there is any significant evidence to support Kellogg's claim that each cereal box contains 500g of cereal. Using the randomly chosen 20 boxes of the product, ACCC finds that the average weight per box is $\bar{x} = 495$g with the sample standard deviation $s = 6$g.

(a) Perform a test of hypothesis at the 5% significance level with the intent to show that Kellogg over-estimates the average weight of cereal in a box. (6 marks)

**Solution.** We define hypotheses (in thousand dollars)

$H_0 : \mu = 500,$ $\boxed{\text{1 mark}}$

$H_A : \mu < 500.$ $\boxed{\text{1 mark}}$

Test statistic

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{495 - 500}{6/\sqrt{20}} = -3.7267.$$

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

Degree of freedom is $df = 19$. Since the t-table is for $P(T \ge t)$, we need to find $P(T \ge 3.7267)$ with $df = 19$. $\boxed{\text{1 mark}}$

Since $3.7267 > 3.579$, we see that $p-$ value $< 0.001 < 0.05$. $\boxed{\text{1 mark}}$

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

(b) If we accept Kellogg's claim of the average cereal weight per box is $\mu = 500$g with the standard deviation $\sigma = 10$g, how large a sample is required if we want a 98% confidence interval for the mean $\mu$ to have a margin of error of $\pm 5$g? (4 marks)

**Solution.**

$\alpha = 0.02$ and need to find $z_{\alpha/2} = z_{0.01}$. $\boxed{\text{1 mark}}$

Check the $N(0,1)$ table we see that $P(z \leq -2.33) \approx 0.01$. Thus, $z_{\alpha/2} = 2.33$.

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

So, set $\frac{5}{\sigma/\sqrt{n}} = 2.33$. $\boxed{\text{1 mark}}$

This give

$$n = \left(\frac{2.33 \times 10}{5}\right)^2 = 4.66^2 = 21.7156.$$

So $n = 22$. $\boxed{\text{1 mark}}$

**Question 5.** City Toyota sales department receives on average 5 customers per hour.

(a) Let $X$ be the number of customers visiting the department. What probability distribution does $X$ satisfy? (1 mark)

**Solution.**

$X$ satisfies a Poisson distribution with $\lambda = 5$ or $X \sim Poi(5)$. $\boxed{\text{1 mark}}$

(b) Find the probability that there are 6 to 8 customers inclusive in one hour, i.e. $P(6 \leq X \leq 8)$. (3 marks)

**Solution.**

$$P(6 \leq X \leq 8) = P(X = 6) + P(X = 7) + P(X = 8) \qquad \boxed{\text{1 mark}}$$

$$= e^{-5}\left(\frac{5^6}{6!} + \frac{5^7}{7!} + \frac{5^8}{8!}\right) \qquad \boxed{\text{1 mark}}$$

$$\approx 0.0067379(21.701389 + 15.500992 + 9.68812) = 0.3159457. \qquad \boxed{\text{1 mark}}$$

(c) What is the probability that the department has 6 to 8 customers inclusive in 2 hours? (4 marks)

**Solution.**

In this case $X$ satisfies a Poisson distribution with $\lambda = 5 \times 2 = 10$,or $X \sim Poi(10)$.

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

$$P(6 \leq X \leq 8) = P(X = 6) + P(X = 7) + P(X = 8) \qquad \boxed{\text{1 mark}}$$

$$= e^{-10}\left(\frac{10^6}{6!} + \frac{10^7}{7!} + \frac{10^8}{8!}\right) \qquad \boxed{\text{1 mark}}$$

$$= e^{-10}(1388.88888 + 1984.12698 + 2480.15873) = 0.2657341. \qquad \boxed{\text{1 mark}}$$

(d) Using the probability found in part (b), find the probability that any 2 of the 4 chosen hours the department has 6 to 8 customers. (4 marks)

**Solution.**

Let $X$ be the number of hours with 6 to 8 customers. Then, we see that $X \sim \text{Bin}(n = 4, p = 0.3159457)$, where $p = 0.3159457$ is from the result found in (b). $\boxed{\text{2 mark}}$

Thus, we have

$$P(X = 2) = \binom{4}{2} 0.3159457^2(1-0.3159457)^2 = \frac{4!}{2! \times 2!} 0.3159457^2 \times 0.684054^2 \quad \boxed{\text{1 mark}}$$

$$= 0.280257 \quad \boxed{\text{1 mark}}$$

**END OF TEST PAPER**