University of Ottawa

MAT 2377 A – Midterm II -Solution, Winter 2018

Multiple Choice Questions (6 marks)

Please Record your answer to each multiple choice question in the table on the first page.

1. A random sample of 50 suspension helmets used by a motorcycle riders and automobile race-car drivers was subjected to an impact test. On 18 of these helmets, some damage was observed. Find a 90% confidence interval for the true proportion of this type of helmets that would show damage on this test.

   A) [0.27, 0.49] B) [0.25, 0.47] C) [0.19, 0.50] D) [0.23, 0.49]
   E) [0.21, 0.49]
   Answer B

   Solution:

   A point estimate for the true proportion of helmets of this type that would show damage on this test is

   $$\widehat{p} = \frac{x}{n} = \frac{18}{50} = 0.36.$$

   The estimated standard error for the estimate in part (a) is

   $$\widehat{\sigma}_{\widehat{P}} = \sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}} = \sqrt{\frac{.36(1-.36)}{50}} = 0.0679$$

A 90% confidence interval for the true proportion of helmets of this type that would show damage on this test is

$$\begin{aligned} & \widehat{p} \pm z_{0.05}\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}} \\ = \; & .36 \pm 1.645\,(0.0679) \\ = \; & .36 \pm 0.1117 \\ = \; & [0.25, 0.47] \end{aligned}$$

2. The Greenland ice sheet covers roughly 80% of the surface of Greenland. As the arctic climate is rapidly warming, the Greenland ice sheet has experienced record melting in the recent years. The following data gives the depth of the ice sheet (in m) measured at various locations during the summer months in the Northeast Greenland National Park. The data are given as follows.

3125	3115	3118	3145	3123	3131	3124	3133	3127	3120

Compute the sample median of this dataset. Is the sample median a measure of central tendency or dispersion?

A) median = 3124.5, measure of central tendency
B) median = 3123, measure of central tendency
C) median = 3124.5, measure of dispersion
D) median = 3126.1, measure of central tendency
E) median = 3126.1, measure of dispersion
F) None of the answers

Answer A

Solution: First we sort the data in ascending order:

3115	3118	3120	3123	3124	3125	3127	3131	3133	3145

$n = 10$, therefore:

$$(1-0.5)y_5 + 0.5\,y_6 = 0.5(3124) + 0.5(3125) = 3124.5.$$

The median is a measure of central tendency.

3. The carbon dioxide (CO2) emission of the latest model of Toyota Prius is a random variable following a normal distribution with mean $\mu = 70$ g/km and standard deviation $\sigma = 1.5$ g/km. Find x_0 such that 93% of these vehicles have an emission level smaller than x_0.

   A) 71.4 B) 73.5 C) 72.2 D)72.9 E)70.8

   Answer: C

   Solution: By standardization, we have:

$$0.93 = P(X < x_0) = P\left(\frac{X-70}{1.5} < \frac{x_0-70}{1.5}\right) = P\left(Z < \frac{x_0-70}{1.5}\right).$$

   From Z-Table, we find $z_0 = 1.475$ such that $P(Z < z_0) = 0.93$.

$$\frac{x_0-70}{1.5} = 1.475.$$

   Hence $x_0 = 70 + (1.5)(1.475) = 72.2125$. The answer is C.

4. A machine produces metal rods used in an automobile suspension system. A random sample of 8 rods is selected, and their diameters are measured. The resulting data (in millimeters) are as follows:

8.24	8.25	8.20	8.23
8.21	8.26	8.26	8.28

   The sample mean is $\overline{x} = 8.241$ and the sample standard deviation is $s = 0.027$. Assume that the diameter of this type of metal rod follows a Normal distribution. Calculate a 95% confidence interval for the mean rod diameter.

A)[8.15, 8.36] B) [7.98, 8.26] C) [8.19, 8.27] D) [8.21, 8.23]
E) [8.22, 8.26]

Answer E

Solution:

A 95% confidence interval for μ is

$$\begin{aligned} & \overline{x} \pm t_{.025,8-1}\frac{s}{\sqrt{n}} \\ = & \ 8.24125 \pm 2.365\left(\frac{0.02696}{\sqrt{8}}\right) \\ = & \ 8.24125 \pm 0.02254 \\ = & \ [8.22, 8.26] \end{aligned}$$

5. Let $X_1 \sim N(2, 1^2)$ and $X_2 \sim N(1, 2^2)$ be two normal random variables. Assume that X_1 and X_2 are independent. Let $Y = 3X_1 - X_2$. What is the value of $sd(Y)$?

   A)2.236 B) 1 C) 5 D) -1 E) 2.646 F) 3.606

   Answer F

   Solution:

   Since $Y = 3X_1 - X_2$, we have

$$Var(Y) = 3^2 Var(X_1) + (-1)^2 Var(X_2) = 9 * 1 + 4 = 13.$$

   Thus, $sd(Y) = \sqrt{Var(Y)} = \sqrt{13} = 3.606$.

6. The following data show the length (in cm) of 12 randomly selected white shrimps in a grocery store.

$$x_1 = 4.2 \quad x_2 = 4.6 \quad x_3 = 4.3 \quad x_4 = 4.9 \quad x_5 = 4.6 \quad x_6 = 4.5$$

$$x_7 = 4.2 \quad x_8 = 4.7 \quad x_9 = 4.9 \quad x_{10} = 4.4 \quad x_{11} = 5.1 \quad x_{12} = 5.6$$

Which of the following statements is true?

A) The IQR is 0.412 and 4.2 and 5.6 are the outliers.

B) The IQR is 0.575 and 5.6 is an outlier.

C) The IQR is 0.575 and 4.2 and 5.6 are the outliers.

D) The IQR is 0.575 and there are no outliers.

E) The IQR is 0.412 and 5.6 is an outlier.

F) The IQR is 0.412 and there are no outliers.

Solution
The answer is D).

The first quartile is

$$Q_1 = (0.75)y_3 + (0.25)y_4 = (0.75)(4.3) + (0.25)(4.4) = 4.325.$$

The third quartile is

$$Q_3 = (0.25)y_9 + (0.75)y_{10} = (0.25)(4.9) + (0.75)(4.9) = 4.9.$$

Thus, $IQR = Q_3 - Q_1 = 4.9 - 4.325 = 0.575$.

The fences are located at $Q_1 - 1.5(IQR) = 4.325 - 0.8625 = 3.4625$ and $Q_3 + 1.5(IQR) = 4.9 + 0.8625 = 5.7625$. Since all values are within the fences [3.4625, 5.7625], there is no outlier in this dataset.

Short answer questions (9 marks)

7. (4 marks) The amount of time that a customer spends on waiting at an airport check-in counter is a random variable with mean μ minutes and standard deviation $\sigma = 1.5$ minutes. Suppose that a random sample of n customers is selected and their waiting times are recorded. Let $\overline{X}$ be the averaged waiting time for those n customers.

   (a) Assume that $\mu = 8.5$ and $n = 49$. Compute $P(5 < \overline{X} < 10)$ approximately based on CLT.

   (b) Suppose that we do not know the true value of μ and the sample size n is not determined. We plan to use $\overline{X}$ to estimate μ. Suppose that we want our estimation to achieve the accuracy $P(|\overline{X} - \mu| < 0.5) = 0.95$. What sample size should be used? (Assume n is large)

Solution:
(a) Since $n = 49 > 30$ is large, CLT tells us

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} = \frac{\overline{X} - 8.5}{1.5/\sqrt{49}} \sim N(0,1)$$

approximately. Thus,

$$\begin{aligned} P(5 < \overline{X} < 10) &\approx P\left(\frac{5 - 8.5}{1.5/\sqrt{49}} < \frac{\overline{X} - 8.5}{1.5/\sqrt{49}} < \frac{10 - 8.5}{1.5/\sqrt{49}}\right) \\ &= P(-16.33 < Z < 7.0) \\ &= P(Z < 7.0) - P(Z < -16.33) \\ &\approx 1 - 0 = 1 \end{aligned}$$

(b) Assume n is large. CLT tells us

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

approximately. Thus,

$$\begin{aligned} P(|\overline{X} - \mu| < 0.5) &= P\left(\left|\frac{\overline{X} - \mu}{\sigma/\sqrt{n}}\right| < \frac{0.5}{\sigma/\sqrt{n}}\right) \\ &= P(|Z| < 0.333\sqrt{n}) \\ &= P(-0.333\sqrt{n} < Z < 0.333\sqrt{n}) = 0.95 \end{aligned}$$

This implies that $P(Z < 0.333\sqrt{n}) = 0.975$. Using Z-table, we find $P(Z < 1.96) = 0.975$. Thus, we have

$$0.333\sqrt{n} = 1.96$$

Solving the about equation, we find $n \approx 35$.

Marking Scheme 2 points for each part.

8. (5 marks) Assume that arrivals of small aircraft at an airport can be modeled as a Poisson process with a rate 1 aircraft per hour.

   (a) Determine the length of an interval of time (in hours) such that the probability that no arrivals occur during the interval is 0.10.

   (b) What is the probability that the waiting time for 3 arrivals will be more than 3 hours?

   (c) What is the mean and variance for the waiting time for 3 arrivals?

Solution:

(a) Let X be the waiting time for an arrival in hours. X follows an exponential distribution with $\lambda = 1$. We want to find x such

$$0.1 = P(X > x) = 1 - P(X \leq x) = 1 - (1 - e^{-\lambda x}) = e^{-x}.$$

Hence $x = -\ln(.1) = 2.3026$ hours.

(b) Let X be the waiting time (in hours) for 3 arrivals. X follows an Erlang distribution with $\lambda = 1$ and $r = 3$. We want

$$\begin{aligned} P(X > 3) &= 1 - P(X \leq 3) = 1 - \left[1 - \sum_{k=0}^{r-1} \frac{e^{-\lambda (3)}[\lambda (3)]^k}{k!}\right] \\ &= 1 - \left[1 - \left(\frac{e^{-3}3^0}{0!} + \frac{e^{-3}3^1}{1!} + \frac{e^{-3}3^2}{2!}\right)\right] = 1 - 0.5768 = 0.4232 \end{aligned}$$

(c) Let X be the waiting time (in hours) for 3 arrivals. X follows an Erlang distribution with $\lambda = 1$ and $r = 3$. We want

$$\mu_X = \frac{r}{\lambda} = \frac{3}{1} = 3 \quad \text{and} \quad \sigma_X^2 = \frac{r}{\lambda^2} = \frac{3}{1^2} = 3.$$

Marking Scheme *2 points each for part (a) and (b), 1 point for part (c)*