Aditya Das
1001675762
axd 5763

Assignment-2

4.24 Installation of some package requires downloading 82 files.

Time taken to download file = 15 sec (μ)

Variance of download time = 16 sec^2 (σ^2)

Thus the standard deviation σ = 4 sec

The probability that the software can be installed in under 20 mins can be found using the central limit theorem.

$$\cancel{S_n = \sum_{i=1}^{n} X_i}$$

$$F_{Z_n}(z) = \left\{ P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \le z \right) \right\} \longrightarrow \phi(z)$$

Now 20 mins = 1200 sec.

So, we can find $P(S_n(z))$ where $S_n \le 1200$

$$\Rightarrow P(S_n < 1200) = \left\{ \frac{1200 - 82 \times 15}{4 \times \sqrt{82}} \right\}$$

$$= \phi\left(\frac{1200 - 1230}{\sqrt{1312}} \right)$$

$$P(S_n < 1200) = \phi(-0.8282)$$

$$= 0.2033$$

(we can approximiate $\phi(-0.8282) \Rightarrow (-0.83)$)

Scanned by CamScanner

4.28 No. of messages sent = 70

Messages are processed sequentially one after another.

Transmission time of each component $\lambda = \frac{5}{min}$

Probability that all 70 messages are transmitted in less than 12 mins
$\downarrow$
720 secs.

Using the given data we find the average $(\mu) = \frac{1}{\lambda} = \frac{1}{5\ min^{-1}} = 0.2\ min$

and the variance $(\sigma^2) = \frac{1}{\lambda^2} = \frac{1}{25\ min^{-2}} = 0.04\ min^2$.

Thus using the Central limit theorem

$$P(S_n < 12) = \left\{ \frac{12-(70*0.2)}{0.2\sqrt{70}} \right\}$$

$$= \phi(-1.1952)$$

$$= 0.1151$$

Scanned by CamScanner

Q8-a Given data set = {43, 37, 50, 51, 58, 105, 52, 45, 45, 10}

a) (i) Sample mean (μ) = $\frac{43+37+50+51+58+105+52+45+45+10}{10}$

= 49.6

(ii) Median

The sorted data is = {10, 37, 43, 45, 45, 50, 51, 52, 58, 105}

Median of even = $\frac{\left(\frac{n}{2} + \frac{n+2}{2}\right)}{2} = \frac{45+50}{2}$

= 47.5

Quartiles

We know that the lower Quartile is the median of the lower half divided by the median and the higher Quartile is the median of the upper half.

lower half = {10, 37, 43, 45, 45} median = $\frac{n}{2}+1$

Q1

$Q_1 = 43$

$Q_2 = 47.5$

upper half = {50, 51, 52, 58, 105}

Q3

$Q_3 = 52$

b) Standard deviation

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n-1}} = \sqrt{\frac{\cancel{4046.4}\ 4960.4}{9}} = 23.4767$$

b) Checking for outliers using the 1.5 (IQR) rul.

$\hat{Q}_3 = IQR = \hat{Q}_3 - Q_1 \Rightarrow 52 - 43 = 9.$

$\hat{Q}_1 - 1.5[IQR]$

$\Rightarrow 43 - 13.5 = 29.5.$

$Q_3 + 1.5[IQR] = 52 + 13.5 = 65.5.$

Thus from the 1.5(IQR) rules 10 and 105 from the sample data set can be considered as outliers.

c) Sample mean = $\frac{37+43+45+45+50+51+52+58}{8} = 47.625$

$$\text{Median} = \frac{S(4)+S(5)}{2} = \frac{45+50}{2} = 47.5.$$

Quartiles:

$Q_1 = \{37, 43, 45, 45\}$. median of lower half

$$Q_1 = \frac{43+45}{2} = 44.$$

Upper half = $\{50, 51, 52, 58\}$ $\qquad Q_2 = 47.5$ (median)

$$Q_3 = \frac{51+52}{2} = \frac{103}{2} = 51.5$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{8}(x_i - \mu)^2}{n-1}} = \cancel{\sqrt{\frac{\quad}{7}}} = \cancel{6.79}$$

$$\sigma = \sqrt{\frac{291.875}{7}} = 6.4572$$

d) After the removal of the outliers:

- There was a slight change in the mean.
- The median remained the same
- A slight change in the quartiles.
- A huge change in the standard deviation.

Q.4 Sample of 3 observations

$X_1 = 0.4,\quad X_2 = 0.7,\quad X_3 = 0.9$

$$f(x) = \begin{cases} \theta x^{\theta-1} & \text{if } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

a) Method of moments

We know $\mu_1 = E(x) = \int_{-\infty}^{\infty} x * f(x)\,dx$

$$\Rightarrow \int_{-\infty}^{0} x * \cancel{f(x)}\,dx + \int_{0}^{1} x * f(x)\,dx + \int_{1}^{\infty} x * \cancel{f(x)}\,dx$$

$$\Rightarrow \int_{0}^{1} x * \theta x^{\theta-1}\,dx$$

$$\Rightarrow \theta \int_{0}^{1} x^{\theta}\,dx \quad \text{—— eqn ①}$$

$$\Rightarrow \theta \left[\frac{x^{\theta+1}}{\theta+1}\right]_0^1 = \frac{\theta}{\theta+1}$$

We know $m_1 = \bar{X} = \sum_i^n \frac{X_i}{n}$ $\quad n = 3$ and $X = [0.4, 0.7, 0.9]$

$$\therefore \Rightarrow m_1 = \left[\frac{0.4 + 0.7 + 0.9}{3}\right] = 0.6667 \longrightarrow \text{eqn ②}$$

Equating 1 and 2

$$\frac{\theta}{\theta+1} = \frac{2}{3} \quad = 3\theta = 2\theta + 2$$

$$= \underline{\theta = 2}$$

$\Rightarrow$

b) Method of Maximum Liklihood

$$\prod_{i=1}^{3} \theta x_i^{\theta-1} = \theta^3 \left(\prod_{i=1} x_i\right)^{\theta-1}$$

Applying log ~~on both sides~~ .

$$\Rightarrow f(\ln f x) = 3\log\theta + \theta - 1 \left(\log \prod_{i=1} x_i\right)$$

Differentiating =

$$\Rightarrow 0 = \frac{3}{\theta} + \log \prod_{i=1}^{3} x_i$$

$$\Rightarrow \theta = \frac{-3}{\sum_{i=1}^{3} \log x_i}$$

$$\sum_{i=1}^{3} \log x_i . (e$$

$$\log_e(0.4) + \log_e(0.7) + \log_e(0.9)$$

$$= -0.91629 - 0.35667$$

$$- 0.105360 .$$

$$= -1.283496 .$$

$$\Rightarrow \theta = \frac{-3}{-1.283496}$$

$$\theta = 2.337366$$

Q5 From the ~~word~~ excel file provided.

1, sample mean = 62.2.
Standard deviation = 25.005378
Variance = 625.26846

$\bar{x} = 62.2$
$s^2 = 625.26896$
$s = 25.005978$

2. $\mu = 62.2 = \bar{x}$
$\sigma = 25.0\cdot = s$

3. Outliers were 175 and 16.

a) The recalculated mean is
~~$\mu = 5$~~ $\bar{x} = 59.8214$.
$s^2 = 115.4113$
$s = 10.74296$.

b). $\mu = \bar{x} = 59.8214$
$\sigma = s = 10.74296$

c). Yes, removing the outliers improved accuracy of the estimate if a Normal of (60,10) is to be considered.

Scanned by CamScanner