Homework 2

GOUTAMI PADMANABHAN
1001669338

4.24) $n = 82$ files, For one file $\mu = 15$ sec, Var $= \sigma^2 = 16\ \text{sec}^2$

P{ software installed in less than 20 minutes } = ?

Let X_1 be a random variable representing download of 1 file

Let S_n be the sum of random variables

$$S_n = X_1 + X_2 + \cdots + X_n$$

$$S_n = X_1 + X_2 + \cdots X_{82}$$

20 minutes = 20 x 60 seconds = 1200 seconds

By central limit theorem,

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}, \quad \mu = 15, \ n = 82, \ \sigma = \sqrt{16} = 4$$

It follows Normal distribution $N(N\mu, \sigma\sqrt{n})$

$$P(S < 1200) = P\left(Z_n < \frac{1200 - n\mu}{\sigma\sqrt{n}}\right)$$

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

$$= P(Z_n < -0.83)$$

$$= \phi(-0.83)$$

$$= \underline{\underline{0.2033}}$$

4.28) $n = 70$, $\lambda = 5\ min^{-1}$, $t = 12\ min$

P(70 messages transmitted in less than 12 minutes) = ?

By central limit theorem, following normal distribution $N(n\mu, \sigma\sqrt{n})$

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

$$S_n = X_1 + X_2 + \dots X_n = X_1 + X_2 \dots + X_{70}$$

$$E(X_i) = \mu = \frac{1}{\lambda} = \frac{1}{5} = 0.2$$

$$\cancel{Var(X) = \frac{1}{\lambda^2}} \quad \sigma = \frac{1}{\lambda} = \frac{1}{5} = 0.2$$

$$P(T < 12) = P\left(Z < \frac{12 - n\mu}{\sigma\sqrt{n}}\right)$$

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

$$= P(Z < -1.20)$$

$$= \phi(-1.20)$$

$$= \underline{\underline{0.1151}}$$

Scanned by CamScanner

8.9) Sample data: 43, 37, 50, 51, 58, 105, 52, 45, 45, 10

a) Compute mean, median, quartile & standard deviation

Mean $= \bar{X} = \dfrac{x_1 + x_2 + \cdots x_n}{n}$

$$\bar{X} = \frac{43+37+50+51+58+105+52+45+45+10}{10}$$

$$\bar{X} = \frac{496}{10} = 49.6$$

$$\boxed{\bar{X} = 49.6}$$

Median: Sort the values and get the middle value

10, 37, 43, 45, 45, 50, 51, 52, 58, 105

Median $M = \dfrac{45+50}{2} = \dfrac{95}{2} = 47.5$

$$\boxed{M = 47.5}$$

Quartile: $Q_1 = ?$ $Q_3 = ?$

25% of sample $= np$, Here $p = 0.25$, $n = 10$

$= 0.25 \times 10$

$= 2.5 \Rightarrow$ 3rd element $\Rightarrow Q_1$

75% of sample $= n(1-p) = 10(1-0.25) = 7.5 \Rightarrow$ 8th element Q_3

$\boxed{Q_1 = 43}$ $\boxed{Q_3 = 52}$

50% of sample $= 0.5 \times 10 = 5 \Rightarrow$ 5th element Q_2

$$\boxed{Q_2 = 45}$$

Scanned by CamScanner

Standard deviation:

Sample variance $s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 = \frac{1}{n-1} \sum_{i=1}^{n} x_i^2 - n\bar{x}^2$

Here $n = 10$, $\bar{x} = 49.6$

$$s^2 = \frac{1}{10-1}\Big[(43-49.6)^2 + (37-49.6)^2 + (50-49.6)^2 + (51-49.6)^2 + (58-49.6)^2 + (105-49.6)^2 + (52-49.6)^2 + (45-49.6)^2 + (45-49.6)^2 + (10-49.6)^2\Big]$$

$$s^2 = \frac{1}{9}(551.155)$$

~~$s^2 = 23.5$~~

Standard deviation $= \sqrt{s^2} = \boxed{s = 23.5}$

b) Outliers using 1.5 (IQR) rule

$\widehat{IQR} = \hat{Q}_3 - \hat{Q}_1 = 52 - 43 = 9$

$\hat{Q}_1 - 1.5(\widehat{IQR}) = 43 - 1.5(9) = 29.5$

$\hat{Q}_3 + 1.5(\widehat{IQR}) = 52 + 1.5(9) = 65.5$

From the sample data given, the value that falls outside the interval $[29.5, 65.5]$ is 105

c) Delete outlier 105 and compute $\bar{x}$, M, Q_1, Q_3, Q_2, s

By ommiting 105 from the sample data.

$\bar{X} = \frac{391}{9} = \underline{\underline{43.4}}$

$M = \underline{\underline{45}}$

25% of sample $= np = 9 \times 0.25 = 2.25 \Rightarrow$ 3rd element

$\boxed{Q_1 = 43}$

50% of sample $= 9 \times 0.5 = 4.5 \Rightarrow$ 5th element

$\boxed{Q_2 = 45}$

75% of sample $= 9 \times 0.75 = 6.75 \Rightarrow$ 7th element

$\boxed{Q_3 = 51}$

$$s = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

$$s^2 = \frac{1}{9-1} \Big[(43-49.6)^2 + (37-49.6)^2 + (50-49.6)^2 + (51-49.6)^2 + (58-49.6)^2 + (52-49.6)^2 + (45-49.6)^2 + (45-49.6)^2 + (10-49.6)^2 \Big]$$

$$s^2 = \frac{1891.24}{8} = 236.405$$

standard deviation $= \sqrt{s^2} = s = \underline{13.9}$

d) By removing the outliers, there is a significant change in the mean $\bar{x}$ and standard deviation s significantly. Quartile Q_1, Q_2, Q_3 and Median M do not seem to change significantly. Though there is slight variation in Q_3

Scanned by CamScanner

9.A) $X_1 = 0.4,\ X_2 = 0.7,\ X_3 = 0.9$

continuous distribution with density.

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

Estimate θ by method of moments:

$$\mu_1 = E(X) = \int_0^1 x f(x)\, dx$$

$$\mu_1 = E(X) = \int_0^1 x \theta x^{\theta-1} dx$$

$$= \int_0^1 \theta x^{\theta} dx$$

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

$$= \frac{\theta}{\theta+1}\left[1^{\theta+1} - 0^{\theta+1}\right]$$

$$= \frac{\theta}{\theta+1}$$

$\mu_1 = m_1$

$$m_1 = \bar{X} = \frac{0.4 + 0.7 + 0.9}{3} = \frac{2}{3}$$

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

Estimate θ by method of Maximum Likelihood:

$$\ln f(x) = \sum_{i=1}^{n} \ln(\theta x^{\theta-1}) \quad , \quad f(x) = \prod_{i=1}^{3} \theta x_i^{\theta-1}$$

$$= \ln\theta + (\theta-1)\ln x_1 + \ln\theta + (\theta-1)\ln x_2 + \ln\theta + (\theta-1)\ln x_3$$

$$= 3\ln\theta + (\theta-1)\sum_{1}^{3} \ln x_i$$

$$\frac{\partial}{\partial\theta} \ln(f(x)) = \frac{3}{\theta} + 1\sum_{1}^{3} \ln x_i$$

Maximizing θ,

$$\frac{3}{\theta} + \sum_{1}^{3} \ln x_i = 0$$

$$\theta = \frac{-3}{\ln 0.4 + \ln 0.7 + \ln 0.9} = \underline{\underline{2.1766}}$$

Q5)

(i) Find sample mean, variance and standard deviation

$$\bar{x} = (69+47+175+70+53+64+74+52+58+45+67+44+58+64+49+70+65+70+48+16+67+55+42+72+61+65+77+70+60+39)/30$$

$$\bar{x} = 1866/30 = \underline{\underline{62.2}}$$

Scanned by CamScanner

(i) Variance $= s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2 = \frac{1}{30-1}\sum_{i=1}^{30}(X_i-62.2)^2$

$s^2 = 625.269$

Std deviation $= s = 25.005$

(ii) Estimate the parameters of normal distribution (μ and σ)

sample mean $\bar{X}$ is close to μ

sample standard deviation s is far from σ

(iii) a) $Q_1 = 49$

$Q_3 = 70$

$\widehat{IQR} = Q_3 - Q_1 = 21$

$Q_1 - 1.5\,\widehat{IQR} = 17.5$

$Q_3 + 1.5\,\widehat{IQR} = 101.5$

The sample values 16 and 175 are outliers.

b) Eliminating them, mean, variance and standard deviation is

$\bar{X} = 59.821$

$s^2 = 115.411$

$s = 10.743$

c) By removing outliers, sample mean & standard deviation is close to μ and σ