Assigment 4 Comp 233

Q1) $\boxed{\text{Sample mean}}$

$$\bar{X} = \frac{1}{12}\left(9.2 + 14.1 + 9.8 + 12.4 + 16.0 + 12.6 + 22.7 + 18.9 + 21.0 + 14.5 + 20.4 + 16.9\right)$$

$$= \boxed{15.7}$$

Sample median

n = even

$$\text{md values: } \{12.6,\ 22.7\} = \frac{12.6 + 22.7}{2} = \boxed{17.65}$$

$\boxed{\text{Sample standard deviation}}$

$$\sqrt{\frac{1}{n}\sum_{k=1}\left(x_k - \bar{X}\right)^2} = \boxed{4.40}$$

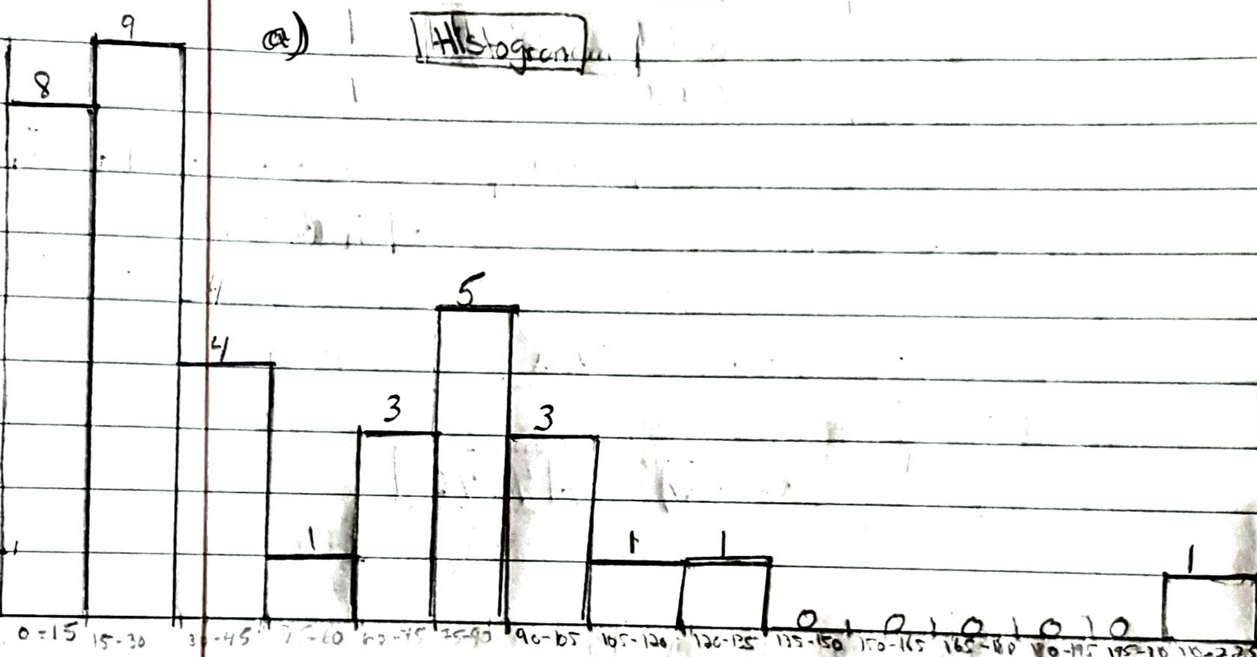
Q2)

| # of particles | micrograms | # of particles | microgram |
|---|---|---|---|
| 1 | 5 | 1 | 28 |
| 1 | 6 | 1 | 29 |
| 2 | 7 | 1 | 33 |
| 0 | 8 | 1 | 37 |
| 1 | 13 | 1 | 44 |
| 1 | 14 | 1 | 45 |
| 2 | 15 | 1 | 53 |
| 1 | 17 | 2 | 65 |
| 1 | 18 | 1 | 70 |
| | 23 | 1 | 77 |
| 2 | 24 | 1 | 80 |
| 2 | 25 | 1 | 81 |

0-15 = 8
15-30 = 9
30-45 = 4
45-60 = 1
60-75 = 3
75-90 = 5
90-105 = 3
105-120 = 1
120-135 = 1
135-150 = 0
150-165 = 0
165-180 = 0
180-195 = 0
195-210 = 0
210-275 = 1

| # of particles | micrograms | # of particles | micrograms |
|---|---|---|---|
| 1 | 82 | 1 | 130 |
| 1 | 85 | 2 | 220 |
| 1 | 92 | | |
| 1 | 95 | | |
| 1 | 103 | | |
| 1 | 110 | | |

a) Histogram




0,72

b) no

c) $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} x_n = 51,72$

d) $S_x = \sqrt{\frac{1}{n} \sum_{i=1} (x_i - \bar{x})^2} = 44.96$

e) $1,5\, S_x = 1,5(S_x) \approx 67,44$

Boundaries $(-15,72 < X_k < 119,16)$.

Proportion of values: $34/36 = 0,94$

$\bar{X} - 1,5\,S_x = -15,72$

$\bar{X} + 1,5\,S = 119,16$

lower bound Chebychev's

$$P(|X-\mu| \geq k\sigma) \; 1 - P(|X-\mu| \leq k\sigma)$$

$$1 - P(-k\sigma \leq X - \mu \leq k\sigma) =$$

$$1 - P(-k\sigma + \mu \leq X \leq k\sigma + \mu) =$$

$$1 - P(-15.72 \leq X \leq 119.16) \leq \frac{1}{k^2}$$

$$\rightarrow 1 - P(-15.72 \leq X \leq 119.16) \leq \frac{4}{9}$$

$$\rightarrow 1 - \frac{4}{9} \leq P(-15.72 \leq X \leq 119.16)$$

$$P(-15.72 \leq X \leq 119.16) \geq 0.56$$

Comparison

we can see that Chebychev's inequality gives us a correct inequality such that in our sample the probability of data between $\bar{X} \pm 1.5\sigma$ is 0.94 and Chebychev's inequality tell us that the probability must be greater than 0.56.

f) $\bar{X} \pm 2S \rightarrow 2S = 2(44.96) = 89.92$

boundaries $(-38.2 \leq X_k \leq 141.64) -$

$$P(-38.2 \leq X_k \leq 141.64) = \frac{35}{36} = 0.97$$

Chebychev's

$$P(|X-\mu| \geq \sigma) = 1 - P(|X-\mu| \leq k\sigma)$$

$$1 - P(-38.2 \leq X_k \leq 141.64) \leq \frac{1}{4}$$

$$P(-38.2 \leq X_k \leq 141.64) \geq 3/4 = 0.75$$

Comparison

we can observe that by taking a larger range of boundary, the Chebychev's inequality become more accurate for obvious reasons.

Q3) a) $P(s^2/\sigma^2 \leq 1.5)$ -.

7 laptops

let the total sample of laptop have a Chi-Square distribution

Properties of Chi-square

$E[X_n^2] = n = 7$

$Var(X_n^2) = \delta^2 = 2(n) = 14$

I'm looking to find the probability that the ratio of sample variance to population variance is less than or equal 1.5

$$P\left(\frac{(n-1)s^2}{\sigma^2} \leq 1.5(7-1)\right)$$

$$P(X_6^2 \leq 9) = 0.825$$

b) $P(0.8 \leq \frac{s^2}{\sigma^2} \leq 1.1)$

$$P\left(0.8 \leq \frac{s^2}{\sigma^2} \leq 1.1\right)$$

$n = 7$

$$P\left(0.8(7-1) \leq \frac{(7-1)}{\sigma^2}s^2 \leq 1.1\right)$$

$$P\left(4.8 \leq \frac{(n-1)}{\sigma^2}s^2 \leq 6.6\right)$$

$P(4.8 \le X_6^2 \le 6.6) =$

$P(X_6^2 \le 6.6) - P(4.8 \le X_6^2)$

b)

Q4)

$n = 24$, $\bar{X} = 6.7229$ $\sum x_i = 161.35$

a) $\sum x_i^2 = 1084.81$

$$S^2 = \frac{\sum x_i^2 - n\bar{x}_i^2}{n-1} \approx 0.0029$$

$S = \sqrt{S^2} = 0.054$

a) $$\sigma^2 = \frac{\sum x_i^2 - n\bar{x}^2}{n} \approx \boxed{0.0028}$$

b) $\alpha = 0.05$

$1 - \alpha = 0.95$

$$X^2_{\alpha/2, n-1} = 38.076$$

$$X^2_{1-\alpha/2, n-1} = 11.689$$

Two sided confidence interval

$$\left( \frac{(n-1)S^2}{X^2_{\alpha/2, n-1}}, \frac{(n-1)S^2}{X^2_{1-\alpha/2, n-1}} \right) \approx \boxed{(0.0017, 0.0057)}$$

Q5)

$$f(x) = \begin{cases} c x^2 e^{-\lambda x} & x \geq 0 \\ 0 & , \text{ otherwise.} \end{cases}$$

$$1 = c\int_0^\infty x^2 e^{-\lambda x}\, dx$$ → integration by parts $u = x^2 \quad dv = e^{-\lambda x}$

$$1 = c\left[ \cancel{-\frac{1}{\lambda} x^2 e^{-\lambda x}}^{\,0} \Big|_0^\infty + \frac{2}{\lambda}\int_0^\infty x e^{-\lambda x}\, dx \right]$$

↳ by l'hopital

$$1 = c\left[ 0 + \frac{2}{\lambda}\int_0^\infty x e^{-\lambda x}\, dx \right]$$

$$1 = \frac{2c}{\lambda}\int_0^\infty x e^{-\lambda x}\, dx$$

$$1 = \frac{2c}{\lambda}\left[ \cancel{\frac{-x}{\lambda} e^{-\lambda x}}^{\,0} \Big|_0^\infty + \frac{1}{\lambda}\int_0^\infty e^{-\lambda x}\, dx \right]$$

$$1 = \frac{2c}{\lambda^2}\left[ \frac{1}{\lambda}\left(-e^{-\lambda x}\right)\Big|_0^\infty \right] = \frac{2c}{\lambda^3}$$

$$\boxed{c = \frac{\lambda^3}{2}}$$

Q5) continued

Joint density function

$$\prod_{i=1}^{n} f(x_i) = \left(\frac{\lambda^3}{2}\right)^n \left(\prod_{i=1}^{n} x_i^2 \left(e^{-\lambda \sum_{i=1}^{n} x_i}\right)\right)$$

$$\log f_x = 3n \log \lambda - n \log 2 + 2 \sum_{1}^{n} \log(x_i) - \lambda \sum x_i$$

$$\frac{d}{d\lambda} \log f_x = \frac{3n}{\lambda} - \sum x_i = 0$$

$$\sum x_i = \frac{3n}{\hat{\lambda}}$$

$$\boxed{\hat{\lambda} = \frac{3}{\bar{X}}}$$

$\frac{d}{dx} \log_{10} =$

$1/x$

e

Q6) $u = 8{,}20$

a) $\sigma = 0{,}02$

$\bar{x} = 8{,}179$

$H_0 : u = 8{,}20$

$H_1 : u \neq 8{,}20$

P-value $= 2(1 - \phi(|z_0|))$

$z_0 \pm 8{,}20$

$z_0 - 8{,}20 < z_0 < z_0 + 8{,}20$

$$z_0 = \frac{8{,}179 - 8{,}20}{0{,}02/\sqrt{10}} = -3{,}32$$

$2(1 - \phi(|-3{,}32|)$

$2(1 - (0{,}9995)$

Pvalue $= 0{,}001 < 0{,}10 \rightarrow H_0$ is rejected.

b) $H_0 : u = 8{,}20$

$H_1 : u \neq 8{,}20$

$\bar{x} = 8{,}179$, $\sigma = 0{,}02$ $u = 8{,}20$

confidence interval $\alpha_2 = 1 - 0{,}025 =$

$$u_x \in \left( \bar{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{10}}, \bar{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right) =$$

$z_\alpha = \Phi^{-1}(1 - \alpha)$ $\quad u_x \in (\bar{x} - 0{,}012, \bar{x} + 0{,}012)$

$z_{\frac{0.05}{2}} = \Phi^{-1}(1 - 0.025)$

$z_\alpha = \Phi^{-1}(0.975) = 1{,}96$

Q6 (b) continued

$u_\alpha \in (8,167 \ , \ 8,191)$

The claimed mean is not in this interval therefore the hypothesis is rejected.

Q6 (c)

as we did in a).

Pvalue = 0,001 which is smaller than 0,05 or 0,01.

Q7) $n = 64$ $\bar{x} = 756,4$ hours $S = 33,9$ h

a) For $n = 64$, $S \approx \sigma = 33,9$ h

$h_0 : u \leq 747,5$

$h_1 : u > 747,5$

upper boundary confidence interval

$$u_x \in \left(-\infty, \bar{X} + z_\alpha \frac{\sigma}{\sqrt{n}}\right) \qquad z_\alpha = \Phi^{-1}(1-\alpha)$$

$$= \Phi^{-1}(1-0,05)$$

$$u_x \in \left(-\infty, 756.4 + \frac{1,65(33,9)}{\sqrt{64}}\right) \qquad = \Phi^{-1}(0,95)$$

$$z_\alpha = 1,65$$

$$u_x \in (-\infty, 763,39)$$

The claim of the null hypothesis is rejected since $u$ has not 747,5 this is an upper boundary

7 (b)

$H_0: \mu \leq 747,5$

$H_1: \mu > 747,5$

$\alpha_1 = 0,01$

upper boundary interval

$z_\alpha = \Phi^{-1}(1-0,01) = \Phi^{-1}(0,99) = 2,33$

$$\mu \in \left(-\infty, 756,4 + \frac{2,33 \cdot (33,9)}{\sqrt{64}}\right)$$

$$\mu \in \left(-\infty, 766,27\right)$$

again we reject the null hypothesis because the upper boundary is greater than 747,5

c) $H_0: \mu \leq 747,5$

$H_1: \mu > 747,5$

$\bar{x} = 756,4\,h$, $n = 64$, $s = \sigma = 33,9\,h$

$H_1: \mu > 747,5$

$P\text{ value} = 1 - \Phi(z_0)$

$\alpha = 0,05$

$P\text{ value} = 1 - \Phi(1,65) = 1 - 0,9505 = 0,0495$

P value $< \alpha$ therefore we must reject the null hypothesis

$\alpha = 0,01$

$P\text{ value} = 1 - \Phi(2,33) = 1 - 0,9901 = 0,0099$

P value $< \alpha$, therefore we reject hypothesis

Q8) $n = 10$, $\bar{x} = 26.4$, $S = 3.502$

$H_0: \mu \geq 30$
$H_1: \mu < 30$

lower boundary confidence interval

$$\mu \in \left(\bar{x} - T_{\alpha, n-1} \frac{S}{\sqrt{n}}, \infty\right) \qquad T_{0.05, 9} = 1.833113$$

$\alpha = 0.05$ (level of significance)

$$\mu \in \left(26.4 - \left(1.833113 \cdot \frac{3.502}{\sqrt{10}}\right), \infty\right)$$

$$\mu \in (24.37, \infty)$$

I reject the hypothesis because the lower boundary is 24.37 miles/gallon and not 30 miles/gallon. I do not believe this advertisement.

Q9)

a) $H_0: \sigma < 0.10$
$H_1: \sigma \geq 0.10$

b) $S = 0.08$

$$\chi^2_{19} = \frac{(n-1)S^2}{\sigma^2} = \frac{19(0.08)^2}{(0.1)^2} \approx 12.16$$

c)

$P(s \geq 0{,}08) = P(s^2 \geq (0{,}08)^2) =$

$P((n-1)s^2 \geq 19(0{,}08)^2) =$

$P\left(\frac{(n-1)s^2}{\sigma^2} \geq 19(0{,}08)^2\right) =$

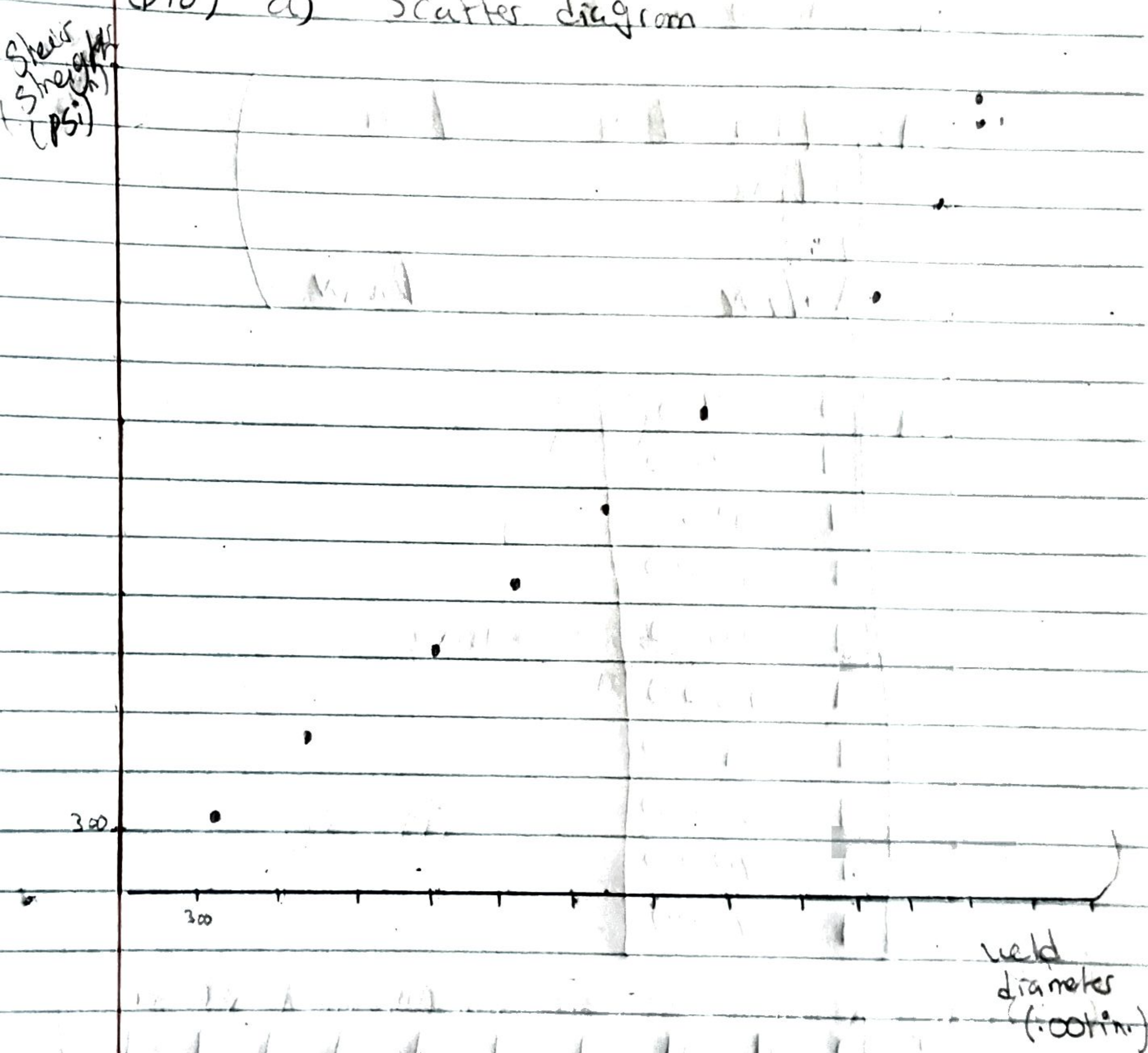
$P(\chi^2_{19} \geq 12{,}16) =$

$1 - P(\chi^2_{19} \leq 12{,}16) = 1 - \Phi\left(\frac{12.16 - 19}{\sqrt{38}}\right)$

$= 1 - \Phi(-1{,}11) = \Phi(1{,}11) \approx 87\%.$

d) For the significance level of 0,05, the claim of $\sigma \leq 0{,}10$ is accepted which means that the pacemaker is reliable.

Q10) a) Scatter diagram




B) $n = 10$,

model: $P(x) = C_1 + C_2 X$

$\phi_1(x) = 1$

$\phi_2(x) = X$

Question 10 (b) continued

$$A = \begin{pmatrix} \phi_1 x_1 & \phi_2 x_1 & \cdots & \phi_n x_1 \\ \phi_1 x_2 & & & \\ \vdots & & & \\ \phi_1 x_N & \cdots & \cdots & \phi_n x_N \end{pmatrix}$$

$$A = \begin{bmatrix} 1 & 400 \\ 1 & 800 \\ 1 & 1250 \\ 1 & 1600 \\ 1 & 2000 \\ 1 & 2500 \\ 1 & 3100 \\ 1 & 3600 \\ 1 & 4000 \\ 1 & 4000 \end{bmatrix} \rightarrow 10 \times 2$$

$$A^T = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 400 & 800 & 1250 & 1600 & 2000 & 2500 & 3100 & 3600 & 4000 & 4000 \end{bmatrix}$$

$2 \times 10$

$$A^T A = \begin{bmatrix} 10 & 19250 \\ 23250 & 53742500 \end{bmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = A^T y$$

$$A^T y = \begin{bmatrix} 18911 \\ 52926500 \end{bmatrix}$$

Question 10 (b) continued

$$\begin{bmatrix} 10 & 19250 \\ 23250 & 53742500 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 18911 \\ 52926500 \end{bmatrix}$$

$$10c_1 + 19250c_2 = 18911$$
$$23250c_1 + 53742500c_2 = 52926500$$

$$c_1 = 5.10112$$
$$c_2 = 0.9850$$

c) $P(x) = 0.9850x + 5.10112$

$x = 0.2250 \text{ in} = 2250$

$P(2250) = 0.9850(2250) + 5.10112 =$

$$\boxed{2221.35112 \text{ (psi)}}$$