Point Estimator

Academy Year :2022-2023

1). Data on pull-off force (pounds) for connectors used in an automobile engine application are follows

79.5	75.1	78.2	74.1	73.9	75.0	77.6	77.3	73.8	74.6	75.5	74.0	74.7
75.9	72.9	73.8	74.2	78.1	75.4	76.3	75.3	76.2	74.9	78.0	75.1	76.8

(a). Calculate the point estimate of the mean pull-force of all connectors.

we get $\bar{x} = \frac{1}{n}\sum_{i=1}^{n}(x_i) = \frac{1}{26}\sum_{i=1}^{26} x_i = 75.615$

Therefore point estimate of data is $\bar{x} = 75.615$ ■

(b). Calculate a point estimate of the pull-off force value that separates the weakest 50% of the connectors in the population from the strongest 50%.

It means we find the point estimate of the median. (We need to set order the data from smallest to largest or from largest to smallest).

So median $MD = \frac{x_{n/2} + x_{(n+1)/2}}{2} = \frac{75.1 + 75.3}{2} = 75.2$

Therefore the point estimate of median is $\hat{MD} = 75.2$ ■

(c). Calculate point estimate of the population variance and the population standard deviation.

We have $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i^2 - 2x_i\bar{x} + \bar{x}^2) = \frac{1}{n-1}(\sum_{i=1}^{n} x_i^2 - 2\bar{x}\sum_{i=1}^{n} x_i + n\bar{x}^2) = \frac{1}{n-1}(\sum_{i=1}^{n} x_i^2 - 2n\bar{x}^2 + n\bar{x}^2) = \frac{1}{n-1}(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2) = \frac{1}{25}(\sum_{i=1}^{26} x_i^2 - 26(75.2)^2 = 1.654$

and population standard deviation is $s = \sqrt{s^2} = \sqrt{1.654} = 1.286$

Therefore the point estimate of population variance is $s^2 = 1.654$ and population standard deviation is $s = 1.286$ ■

(d): Calculate the standard error of the point estimate found in part (a)

we have $\hat{\sigma_x} = \sqrt{V(\bar{x})} = \frac{s}{\sqrt{n}} = \frac{1.286}{\sqrt{26}} = 0.252$

Therefore the standard error of mean is $\hat{\sigma_x} = 0.252$ (pound) ■

(e). Calculate a point estimate of the population of all connectors in the population whose pull-off force is less than 73 pound

we obtain $\hat{p} = \frac{x}{n} = \frac{1}{26} = 0.03846$

Therefore $\hat{p} = 0.03846 = 3.846\%$ ■

---

2. (a): A random sample of 10 houses in a particular area, each of which is heated with natural gas, is selected and the amount of gas (therms) used during the month of January is determined for each house. The resulting observations are

103 156 118 89 125 147 122 109 138 99

Let μ denote the average gas usage during January by all houses in this area . Compute a point estimate of μ

we have $\mu = \bar{x} = \frac{1}{16}\sum_{i=1}^{16} x_i = 120.7$

Therefore $\hat{\mu} = 120.7$ ■

b). Suppose there are 10,000 houses in this area that use natural gas for heating. Let t denote the total amount of gas used by all of these houses during January. Estimate t using the data of part(a). What estimator did you use in computing your estimate ? We know that t is the total amount of gas used by all of those houses during January.

$$\hat{t} = 10,000 \cdot \hat{\mu} = 120700 terms ■$$

In this computing , we use the point estimate of the average gas usage during January by 10 houses in

particular area.
c). Use the data in part (a) to estimate p , the proportion of all houses that used least 100 therms.
We have $\hat{p} = \dfrac{x}{n} = \dfrac{8}{10} = 0.8$
Therefore $\hat{p} = 0.8$ ■
d). Given a pint estimate of the population median usage (the middle value in the population of all houses) based on the sample of part (a). What estimator did you use?
we have $\hat{\mu} = MD = \dfrac{x_5 + x_6}{2} = 120$ ■

---

3. let $X_1, X_2, \cdots, X_n$ be a random sample from a distribution having finite variance σ^2.
Show that $S^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ is an unbiased estimator of σ^2.
We have to prove $E(S^2) = \sigma^2$
we have $S^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 \Longrightarrow S^2 = \dfrac{1}{n-1}(\sum_{i=1}^{n} X_i^2 - n\bar{X}^2)$
We get $E(S^2) = E(\dfrac{1}{n-1}(\sum_{i=1}^{n} X_i^2 - n\bar{X}^2) = \dfrac{1}{n-1}(\sum_{i=1}^{n} E(X_i^2) - nE(\bar{X}^2))$
we know that $X_1, X_2, \cdots, X_n \sim^{iid} N(\mu, \sigma^2) \Longrightarrow E(X_i) = \mu$ and $V(X_i) = \sigma^2$. Moreover
$V(X) = E(X^2) - [E(X)]^2 \Longrightarrow E(X^2) = V(X) + [E(X)]^2$
$\Longleftrightarrow E(X_i^2) = V(X_i) + [E(x)]^2 = \sigma^2 + \mu^2$. similarly
$E(\bar{X}^2) = V(\bar{X}) + [E(\bar{X})]^2 = V(\dfrac{1}{n}\sum_{i=1}^{n} X_i) + [E(\dfrac{1}{n}\sum_{i=1}^{n} X_i)]^2 = \dfrac{1}{n^2}\sum_{i=1}^{n} V(X_i) + [\dfrac{1}{n}\sum_{i=1}^{n} E(X_i)]^2 = \dfrac{\sigma^2}{n} + \mu^2$
Then we get $E(S^2) = \dfrac{1}{n-1}(\sum_{i=1}^{2}(\sigma^2 + \mu^2) - n(\dfrac{\sigma^2}{n} + \mu^2)) = \dfrac{1}{n-1}(n\sigma^2 + n\mu^2 - \sigma 2 - n\mu^2)$
$\Longleftrightarrow E(S^2) = \dfrac{1}{n-1}((n-1)\sigma^2) = \sigma^2$ (True)
Therefore S^2 is an unbiased of σ^2 ■

---

4. Suppose that X is the number of observed " Successes" in a sample of n observations where p is the probability of success on each observation.
(a). Show that $\hat{p} = \dfrac{x}{n}$ is an unbiased estimator of p.
we have $X_1, X_2, \cdots, X_n \sim^{iid} B(n, p) \Longrightarrow E(X) = np$ and $V(X) = npq = np(1-p)$
consider $E(\hat{p}) = E(\dfrac{x}{n}) = \dfrac{1}{n}E(x) = \dfrac{1}{n}np = p$
Therefore $\hat{p}$ is an unbiased estimator of p ■.
(b). Show that the standard error of $\hat{p} = \sqrt{p(1-p)/n}$
consider $sig\hat{m}ap = \sqrt{V(\hat{p})} = \sqrt{V(\dfrac{x}{n})} = \sqrt{\dfrac{1}{n^2}V(X)} = \sqrt{\dfrac{1}{n^2}np(1-p)}$
Therefore the standard error of $\hat{p}$ is $\sqrt{\dfrac{(1-p)p}{n}}$ ■

---

5. let $X_1, X_2, \cdots, X_n$ be a random sample drawn from a distribution with mean μ and variance σ^2 and let $a_1, a_2, \cdots, a_n$ be real numbers such that $\sum_{i=1}^{n} a_i = 1$
Define $\hat{X} = \sum_{i=1}^{n} a_i X_i$.
(a). show that $\bar{X}$ is an unbiased estimator of μ.
We have $\hat{X} = \sum_{i=1}^{n} a_i X_i \Longrightarrow E(\hat{X}) = E(\sum_{i=1}^{n} a_i X_i) = sum_{i=1}^{n} a_i E(X_i)$
Since $X_1, X_2, \cdots, X_n \sim^{iid} N(\mu, \sigma^2)$
$\Longrightarrow E(X_i) = \mu \Longrightarrow E(\hat{X}) = \mu\sum_{i=1}^{n} = \mu$

Therefore $\hat{X}$ is an unbiased estimator of μ. ■

b). Show that $V(\bar{X}) \leq V(\hat{X})$ (hence among all estimator of μ of the form $\sum_{i=1}^{n} a_i X_i, \bar{X}$ is the MVUE).

We have $V(\hat{X}) = \sigma^2 \sum_{i=1}^{n} a_i^2$ and $V(\bar{X}) = \frac{1}{n^2} V(\sum_{i=1}^{n} X_i) = \frac{\sigma^2}{n}$

By using Cauchy-Schwarz inequality

$(a_1 b_1 + \cdots + a_n b_n)^2 \leq (a_1^2 + \cdots + a_n^2 (b_1^2 + \cdots + b_n^2)$ for a_i, b_i are positive for all $i \in \mathbb{N}$

Take $b_1 = \cdots = b_n = 1$, then $(a_1^2 + \cdots + a_n^2) \geq \frac{1}{n}$

Therefore $V(\bar{X}) \leq V(\hat{X})$ ■

---

6). Let $X_1, X_2, \cdots, X_n$ be a random sample from distribution with unknown mean $-\infty < \mu < +\infty$, and unknown variance $\sigma^2 > 0$. show that the statistic $\bar{X}$ and $Y = \frac{X_1 + 2X_2 + \cdots + nX_n}{\frac{n(n+1)}{2}}$ are both unbiased estimators of μ. Further,show that $V(\bar{X}) < V(Y)$

We have $E(\bar{X}) = \frac{1}{n} \sum_{i=1}^{n} E(X_i) = \frac{1}{n} n\mu = \mu$

and $E(Y) = E(\frac{X_1 + 2X_2 + \cdots + nX_n}{1 + 2 + 3 + \cdots + n}) = \frac{\frac{n(n+1)}{2}}{\frac{n(n+1)}{2}} \sum_{i=1}^{n} E(X_i) = \mu$

Therefore They are unbiased estimators of μ

$blacksquare$

Show that $V(\bar{X}) \leq V(Y)$

We have $V(\bar{X}) = \frac{\sigma^2}{n}$ and $V(Y) = \frac{4\sigma^2}{n^2(1+n)^2} \sum_{i=1}^{n} i^2$

By Cauchy-Schwarz inequality:

$(a_1 b_1 + \cdots + a_n b_n)^2 \leq (a_1^2 + \cdots + a_n^2)$, for a_i, b_i positive for all $i \in \mathbb{N}$

Take $b_1 = \cdots = b_n = 1$ and $a_i = 1 \forall i \in [1, n]$

We got $(\frac{n(n+1)2)^2 \leq n(1^2 + 2^[2}{+} \cdots + n^2)$

Then

$$\frac{1}{n} \leq (\frac{4 \sum_{i=1}^{n} i^2}{n^2(1+n)^2})$$

Multiply by σ^2 we get

$$V(\bar{X}) \leq V(Y)$$

Therefore $V(\bar{X}) \leq V(Y)$ ■

---

7)Using a long rod that has length μ, you are going to lay out of square plot in which the length of each side is μ. Thus the area of the plot will be μ^2. However , you do not know the value of μ, so you decide to make n independent measurements $X - 1, X - 2, \cdots, X_n$ of the length. Assume that each X_i has mean μ (unbiased measurements) and variance σ^2.

(a): Show that $\bar{X}^2$ is not an unbiased estimator for μ^2.[Hint: For any rv $Y, E(Y^2) = V(Y) + [E(Y)]^2$]

consider $E(\bar{X}^2) = V(\bar{X}) + [E(\bar{X})]^2$

Suppose $X_1, X_2, \cdots, X_n \sim^{iid} X \qquad with \qquad E(\bar{X}) = \mu, V(\bar{X}) = \frac{\sigma^2}{n}$

$\Longrightarrow E(\bar{X}^2) = \frac{\sigma^2}{n} + \mu^2 \neq \mu^2$

Therefore $E(\bar{X}^2)$ does not an unbiased estimator of μ^2 ■

(b): For what value of k is the estimator $\bar{X}^2 - kS^2$ unbiased for μ^2[Hint: Compute $E(\bar{X} - kS^2)$]

We have $E(\bar{X} - kS^2) = E(\bar{X}) - kE(S^2)$

for $X_1, X_2, \cdots, X_n \sim^{iid} X$ with $E(X) = \mu, V(X) = \sigma^2$ then the $E(S^2) = \sigma^2$

where S^2 is the sample variance

$\Longrightarrow E(\bar{X}-kS^2)=\dfrac{\sigma^2}{n}+\mu^2-k\sigma^2$
It's an unbiased if $k=\dfrac{1}{n}$
Therefore $k=\dfrac{1}{n}$ that is $E(\bar{X}-kS^2)$ is an unbiased estimator for μ^2 ■

---

8). Let $X_!, X_2, \cdots, X_n$ be uniformly distribution on the interval $[0,\theta]$. Recall that the maximum likelihood estimator of θ is $\theta = max(X_i)$
a). Argue intuitively why $\hat{\theta}$ cannot be an unbiased estimator for θ
we knew that $\hat{\theta}$ will be always less than θ in uniform distribution on interval $[0,\theta]$, that is why $\hat{\theta}$ can not be an unbiased estimator for θ. b). Suppose that $E(\hat{\theta})=n\theta(n+1)$, Is it reasonable that $\hat{\theta}$ consistently underestimate θ? . Show that the bias in the estimators approaches zero as n gets large .
The $\hat{\theta}$ is consistently underestimate, because it is not unbiased estimator.
Note: A statistic is positively biased if it tends to overestimate the parameter. A statistic is negatively biased if it tends to underestimate the parameter.
$B(\hat{\theta}=E(\hat{\theta})-\theta=\dfrac{n}{n+1}\theta-\theta=-\dfrac{\theta}{n+1}$
Therefore $B(\hat{\theta})\to 0$ when $n\to\infty$ ■
c). Propose an unbiased estimator of θ. we have $E(\hat{\theta})=\dfrac{n}{n+1}\theta$ it is unbiased if $E(\dfrac{n+1}{n}\hat{\theta})=\theta$
so we propose $\tilde{\theta}=\dfrac{n+1}{n}\hat{\theta}$ as an unbiased estimator of θ.■
d). Let $Y=max(x_i)$. Use the fact that $Y\le y$ if and only if each $X_i\le y$ to derive the cumulative distribution function of Y. Then show that the probability density function of Y is

$$f(y)=\begin{cases}\dfrac{ny^{n-1}}{\theta^n} & 0\le y\le\theta\\ 0 & otherwise\end{cases}.$$

Use this result to show that the maximum likelihood estimator for θ is biased.
Find cdf of Y such that
$P(Y<y)=P(max(X_i)<y)$ since X_i are independent for $i=1,\cdots,n$
we obtain $P(Y<y)=P\big[X<y\big]^n$
We have $X_i\sim U[0,\theta]$, so $P(Y<y)=\dfrac{y}{\theta}$ for $0\le y\le\theta$
Therefore $P(Y<y)=(\dfrac{y}{\theta})^n$ ■
so the pdf is given by $f_Y(y)=n\dfrac{y^{n-1}}{\theta^n}$ for $0<y<\theta$ and 0 otherwise
Use this to show that the maximum likelihood estimator for θ is biased.
we have $E(\hat{\theta})=\displaystyle\int_0^\theta\frac{ny^n}{\theta^n}=\frac{n}{n+1}\theta$
Therefore , It is biased estimator. ■

e). We have two unbiased estimator for θ, the moment estimator $\hat{\theta_1}=2\hat{X}$ and $\hat{\theta_2}=[(n+1)/n]max(X_i,,$ where $max(X_i)$ is the largest observation in a random sample of size n. It can be shown that $V(\hat{\theta_1}=\theta^2/(3n)$ and that $V(\theta_2^2=\dfrac{\theta^2}{n(n+2)})$. Show that if $n>1,\hat{\theta}_2$ is a better estimator than $\hat{\theta_1}$. In what sense is it a better estimator of θ?
we have $V(\hat{\theta}_1)=\dfrac{\theta^2}{3^n}$ and $V(\hat{\theta_2})=\dfrac{\theta^2}{n(n+1)}$
For $n>1$ we have $V(\hat{\theta}_1)\ge V(\hat{\theta_2})$
Therefore $\hat{\theta_2}$ is better estimator of θ.■

9.). A random sample $X_1,X_2,\cdots,X_n$ of size n is taken from Poison distribution with a mean of $\lambda, 0<\lambda<\infty$
a). Show that the maximum likelihood estimator of λ is $\hat{\lambda}=\bar{X}$
We have $X_1,X_2,\cdots,X_n\sim^{iid}Poi(\lambda)\Longrightarrow f(x;\lambda)=\dfrac{e^{-\lambda}\lambda^{x_i}}{x_i!}$

we obtain likelihood function $L(x;\lambda) = \prod_{i=1}^{n} f(x;\lambda) = e^{-n\lambda} \prod_{i=1}^{n} \frac{\lambda^{x_i}}{x_i!}$

$\Longrightarrow \ln[L(x;\lambda)] = -n\lambda + \ln(\lambda)\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} x_i!$

$\Longrightarrow \frac{\partial \ln[L(x;\lambda)]}{\partial \lambda} = -n + \frac{1}{\lambda}\sum_{i=1}^{n} x_i \quad \& \quad \frac{\partial^2 \ln[L(x;\lambda)]}{\partial \lambda^2} = -\frac{1}{\lambda^2}\sum_{i=1}^{n} x_i < 0(max)$

Set $\frac{\partial \ln[L(x;\lambda)]}{\partial \lambda} = 0 \Longleftrightarrow -n + \frac{1}{\lambda}\sum_{i=1}^{n} x_i = 0 \Longrightarrow \lambda = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{X}$

Therefore $\hat{\lambda} = \bar{X}$ ■

b). Let X equal the numbers of flaws per 100 feet of a used computer tape. Assume that X has a Poison distribution with a mean of λ . If 40 observation of X yielded 5 zeros, 7 one , 12 twos , 9 threes , 5 fours, 1 five and 1 six. find the maximum likelihood estimator of λ.

we have $\hat{\lambda} = \bar{X} = \frac{5(0)+7(1)+12(2)+9(3)+5(4)+1(5)+1(6)}{40} = 2.225$

Therefore $\bar{X} = 2.225$ ■

---

10. Let $f(x) = (1/\theta)x^{(1-\theta)/\theta}, 0 < x < 1, 0 < \theta < \infty$

a). Show that the maximum likelihood estimator of θ is $\hat{\theta} = -(1/n)\sum_{i=1}^{n} \ln X_i$

we have $L(x;\theta) = \prod_{i=1}^{n} f(x) = \prod_{i=1}^{n} \frac{1}{\theta} x_i^{\frac{1-\theta}{\theta}} = \frac{1}{\theta^n}\prod x_i^{(1-\theta)/\theta)}$

$\Longrightarrow \ln[L(x;\theta)] = -n\ln\theta + \sum_{i=1}^{n} \ln[x_i^{\frac{(1-\theta)}{\theta}}] = -n\ln\theta + (\frac{1}{\theta} - 1)\sum_{i=1}^{n} \ln x_i$

$\Longrightarrow \frac{\partial \ln[L(x;\theta)]}{\partial \theta} = -\frac{n}{\theta} - \frac{1}{\theta^2}\sum_{i=1}^{n} \ln x_i$

Set $\frac{\partial \ln[L(x;\theta)]}{\partial \theta} = 0 \Longleftrightarrow \frac{n}{\theta} = -\frac{1}{\theta^2}\sum_{i=1}^{n} \ln x_i \Longrightarrow \theta = -\frac{1}{n}\sum_{i=1}^{n} \ln x_i$

Therefore $\hat{\theta} = -\frac{1}{n}\sum_{i=1}^{n} \ln x_i$ ■

b). show that $E(\hat{\theta}) = \theta$ and thus that $\hat{\theta}$ is an unbiased estimator of θ

Consider $E(\hat{\theta}) = E(-\frac{1}{n}\sum_{i=1}^{n} \ln x_i) = -\frac{1}{n}\sum_{i=1}^{n} E(\ln x_i)$

find $E(\ln x) = \int_0^1 \ln x(\frac{1}{\theta}x^{(1-\theta)/\theta})dx = \frac{1}{\theta}\int_0^1 x^{(1-\theta)/\theta)}(\ln x)dx$

By using integral by part we get $E(\ln x) = -\theta$ then we obtain

$E(\hat{\theta}) = -\frac{1}{n}\sum_{i=1}^{n} -\theta = \theta$

Therefore $\hat{\theta}$ is an unbiased estimator of θ.

---

11. Let $X_1, X_2, \cdots, X_n$ be a random sample of size n from the exponential distribution whose pdf is $f(x;\theta) = \frac{1}{\theta}e^{-x/\theta}, 0 < x < \infty, 0 < \theta < \infty$

a). show that $\bar{X}$ is an unbiased estimator of θ

We have $E(\bar{X}) = \frac{1}{n}\sum_{i=1}^{n} E(X_i)$

Since $X_1, X_2, \cdots, X_n \sim^{iid} Exp(\theta)$ then we have $E(X) = \theta$ and $V(X) = \theta^2$

We obtain $E(\bar{X}) = \frac{1}{n} n\theta = \theta$

Therefore $\bar{X}$ is an unbiased estimator of θ ■

b). Show that the variance of $\bar{X} = \frac{\theta^2}{n}$

We have $V(\bar{X}) = (\frac{1}{n}\sum_{i=1}^{n} X_i) = \frac{1}{n^2}\sum_{i=1}^{n} V(X_i) = \frac{1}{n^2}n\theta^2 = \frac{\theta^2}{n}$

Therefore $V(\bar{X}) = \frac{\theta^2}{n}$

c). What is a good estimator of θ if a random sample of size 5 yielded the sample values 3.5, 8.1, 0.9, 4.4 , 0.5?

Since $\bar{X}$ is unbiased ,so it is a good estimator such that $\bar{x} = \frac{3.5+8.1+0.9+4.4+0.5}{5} = 2.02$

Therefore $\bar{X} = 2.02$ ■

---

12). A diagnostic test for a certain disease is applied to n individuals known to not have the disease. Let X = the number among the n test results that are positive (indicating presence of the disease, so X is the number of false positive) and p = the probability that a disease-free individual's test result is positive (i.e. p is the true proportion of test results from disease from disease-free individuals that are positive). Assume that only X is available rather that the actual sequence of test results.

a). Derive the maximum likelihood estimator of p. If $n = 20$ and $x = 3$, what is the estimator?

Let X = the number among the n test results that are positive.

and p = the probability that a disease-free individual's test results is positive.

we have $X \sim Bin(n, p)$

we have pmf $\Longrightarrow f(x; p) = p^x(1-p)^{n-x}, x = 0, 1$

the likelihood function is $L(x; p) = \prod_{i=1}^{n} p^{x_i}(1-p)^{n-x_i} = C_n^x p^x(1-p)^{n-x}$

The log likelihood function $\ln L(x; p) = \ln[C_n^x p^x(1-p)^{n-x}]$

$\Longrightarrow \frac{\partial \ln L(x; p)}{\partial p} = \frac{x}{p} - \frac{n-x}{1-p} = 0 \Longleftrightarrow p = \frac{x}{n}$

Therefore $\hat{p} = \frac{x}{n}$ is the estimator of p. ■

for the data above we get $\hat{p} = \frac{3}{20} = 0.15$ ■

b). Is the estimator of part (a) unbiased? consider $E(\hat{p}) = E(\frac{x}{n}) = \frac{1}{n}E(x) = \frac{1}{n}np = p$

Therefore $\hat{p}$ is an unbiased estimator of p ■

c). If $n = 20$ and $x = 3$, what is MLE of the probability$(1-p)^5$ that none of the next five tests done on disease free individuals are positive?

we have $n = 20, x = 3$; and $\hat{p} = 0.15$, so $(1-\hat{p})^5 = (1-0.15)^5 = 0.85^5 = 0.443$

Therefore $(1-p)^5 = 0.443$ ■

---

13). The shear strength of each of ten test spot welds is determined, yielding the following data (psi)

$$392 \quad 376 \quad 401 \quad 367 \quad 389 \quad 362 \quad 409 \quad 415 \quad 358 \quad 375 \tag{1}$$

a). Assume that shear strength is normally distributed, estimate the true average shear strength and standard deviation deviation of shear strength using the method of maximum likelihood.

We have $X_1, X_2, \cdots, X_n \sim N(\mu, \sigma^2)$

by previous exercise we have $\hat{\mu} = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{10}\sum_{i=1}^{10} x_i = 384.4$

for standard deviation

we have $L(x; \sigma) = (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}$

$\Longrightarrow \ln L(x; \sigma) = \frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$

$\Longrightarrow \frac{\partial L(x; \sigma)}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(x_i-\mu)^2 = 0 \Longrightarrow \sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2}$

Therefore the standard deviation is $\hat{\sigma} = \sqrt{\frac{1}{10}\sum_{i=1}^{10}(x_i-\mu)^2} = 3556.4$ ■

b). Again assuming a normal distribution, estimate the strength value below which 95% of all welds will have their strengths. [Hint: what is the 95th percentile in terms of μ and σ ? Now use the variance principle].
We have $P(X \leq c) = 0.95$
$\iff P(X \leq c) = \Phi(\frac{c-\mu}{\sigma}) = 0.95$
$\implies \frac{c-\mu}{\sigma} = 1.65 \implies c = 1.65\sigma + \mu$
then $\hat{c} = 1.65\hat{\sigma} + \hat{\mu}$ (By invariance principle)
Therefore $\hat{c} = 6252.46$ ■

---

14). At time $t = 0, 20$ identical components are tested. The lifetime distribution of each is exponential with parameter λ. The experimenter then leaves the test facility unmonitored. On his return 24 hours later, the experimenter immediately terminates the test after noticing that $y = 15$of the 20 components are all still in operation (so 5 have failed). Derive the MLE of λ . [Hint: Let $Y =$ the number that survive 24 hours. Then $Y \sim Bin(n, p)$. what is the of p? Now notice that $p = P(X_i \geq 24)$, where X_i is exponentially distributed. This relates λ to p, so the former can be estimates once the latter has been.]
Let $T_i =$ the life time of component $i-th$ and $T_i \sim Exp(\lambda)$
Derive MLE of λ
Let $Y =$ the number that survive 24hours.
$Y \sim Bin(n, p)$,
We have $p = P(T_i \geq 24) = e^{\frac{24}{\lambda}}$ then
$p(y) = C(n, y)p^y(1-p)^{n-y}$ and $\ln(L(y; p)) = \ln[c(n, y)p^y(1-p)^{n-y}]$
$= \ln C(n, y) + y \ln p + (n-y)\ln(1-p)$
set $\frac{\partial \ln L(y;p)}{\partial p} = 0 \iff \frac{y}{p} - \frac{y-n}{1-p} = 0 \iff p = \frac{y}{n}$
Therefore $\hat{p} = \frac{y}{n} = \frac{15}{20} = 0.75$ and $\hat{\lambda} = \frac{24}{\ln 0.75}$ ■

---

15). Let $X_1, X_2, \cdots, X_n$ be a random from Bin(1,p)(i.e., n Bernoulli trails). Thus ,

$$Y = \sum_{i=1}^{n} X_i \sim Bin(n, p)$$

a). show that $\bar{X} = Y/n$ is an unbiased estimator of p.
we have $E(\bar{X}) = E(Y/n) = \frac{1}{n}E(Y) = \frac{1}{n}\sum_{i=1}^{n} E(x_i)$,
since $X_i \sim Bernoulli(p) \implies E(x) = p$ and $V(x) = p(1-p)$
$\implies E(\bar{X}) = \frac{1}{n}np = p$
Therefore $\bar{X}$ is an unbiased estimator of p. ■
b). show that $Var(\bar{X}) = p(1-p)/n$
consider $Var(\bar{X}) = \frac{1}{n^2}Var(Y) = \frac{1}{n^2}\sum_{i=1}^{n} V(x_i) = \frac{1}{n^2}np(1-p)$
Therefore $Var(\bar{X}) = \frac{p(1-p)}{n}$ ■
c). Show that $E[\bar{X}(1-\bar{X})/n] = (n-1)[p(1-p)/n^2]$
Consider $E[\bar{X}(1-\bar{X})/n] = \frac{1}{n}[E(\bar{X}) - E(\bar{X}^2)]$
find $E(\bar{x}^2) = V(\bar{X}) + [E(\bar{X})]^2 = \frac{p(1-p)}{n} + p^2$
so $E[\bar{X}(1-\bar{X})/n] = \frac{1}{n}[p - \frac{p(1-p)}{n} - p^2] = \frac{p(1-p)}{n} - \frac{p(1-p)}{n^2} = \frac{p(1-p)}{n}(1-\frac{1}{n}) = \frac{(n-1)p(1-P)}{n^2}$
Therefore $E[\bar{X}(1-\bar{X})/n] = (n-1)\frac{p(1-p)}{n^2}$ ■
d). Find the value of c so that $c\bar{X}(1-\bar{X})$ is an unbiased estimator of $Var(\bar{X}) = \frac{p(1-p)}{n}$
from c question above we get $E(c\bar{X}(1-\bar{X})) = cE(\bar{X}(1-\bar{X})) = c(n-1)p(1-p)/n$

Therefore we get $c = \dfrac{1}{(n-1)}$ for the question d is hold.■

---

16). Assume that the number of defects in a car has a Poisson distribution with parameter λ.To estimate λ we obtain the random sample $X_1, X_2, \cdots, X_n$

a). Find the Fisher information in a single observation using two methods.

we have $X_i \sim Poi(\lambda) \longrightarrow pmf, f(x, \lambda) = \dfrac{e^{\lambda}\lambda^{x}}{x!}$

...First method $I(\lambda) = -E\Big(\dfrac{\partial^2}{\partial\lambda^2}\ln(L(f(x;\lambda)\Big)$

we have $\ln f(x;\lambda) = -\lambda + x\lambda - \ln x! \Longrightarrow \dfrac{\partial}{\partial\lambda}\ln f(x;\lambda) = -1 + \dfrac{x}{\lambda} \Longrightarrow \dfrac{\partial^2}{\partial\lambda^2}\ln f(x;\lambda) = -\dfrac{x}{\lambda^2}$

Thus, $I(\lambda) = \dfrac{1}{\lambda^2}E(x)$

we have $E(x) = \lambda$and $V(x) = \lambda$

Therefore $I(\lambda) = \dfrac{1}{\lambda}$

...Second method

$I(\lambda) = V\Big(\dfrac{\partial}{\partial\lambda}\ln f(x;\lambda)\Big) = V(-1 + \dfrac{x}{\lambda}) = \dfrac{1}{\lambda^2}V(x) = \dfrac{1}{\lambda}$

Therefore $I(\lambda) = \dfrac{1}{\lambda}$ ■

b). Find the Cramer-Rao lower bound for the variance of an unbiased estimator of λ.

The lower bound is $V(x) \geq \dfrac{1}{nI(\lambda)} = \dfrac{\lambda}{n}$

c). Find the MLE of λ and show that the MLE is an efficient estimator

we have $\dfrac{\partial \ln l(f(x;\theta))}{\partial\lambda} = -1 + \dfrac{x}{\lambda} = 0 \Longleftrightarrow \lambda = x$

thus $\hat{\lambda} = \hat{x}$

we have $V(\hat{\lambda}) = V(\hat{x}) = \dfrac{1}{n^2}V(x_i) = \dfrac{1}{n^2}n\lambda = \dfrac{\lambda}{n}$ (it's equal to the Cramer-Rao lower bound from question (b)

Therefore the efficient estimator $\hat{\lambda} = \bar{x}$ ■

17.Suppose the waiting time for a bus is uniformly distribution on $[0, \theta]$ and the results $x_1, x_2, \cdots, x_n$ of a random sample from this distribution have been observed.

(a). Find MLE $\hat{\theta}$ of θ

we have $x_1, x_2, \cdots, x_n \sim^{iid} U[0, \theta]$

$\Longrightarrow f(x;\theta) = \begin{cases} \dfrac{1}{\theta} & ,0 \leq x_i \leq \theta \\ 0 & ,\text{otherwise} \end{cases}$

The likelihood function $L(x;\theta) = \displaystyle\prod_{i=1}^{n}\dfrac{1}{\theta} = \dfrac{1}{\theta^n}$

$\Longrightarrow \ln(L(x;\theta)) = \ln(\dfrac{1}{\theta}) = -n\ln(\theta)$

$\Longrightarrow \dfrac{\partial \ln(L(x;\theta))}{\partial\theta} = \dfrac{-n}{\theta}$ and $\dfrac{\partial^2 \ln(L(x;\theta))}{\partial\theta^2} = \dfrac{n}{\theta^2} > 0$

Therefore the MLE $\hat{\theta} = max(x_i)$ ■

b). Letting $\tilde{\theta} = \dfrac{(n+1)\hat{\theta}}{n}$, show that $\tilde{\theta}$ is an unbiased and find its variance

we have

$$E(\tilde{\theta}) = E(\frac{(n+1)\hat{\theta}}{n}) = \frac{(n+1)}{n}E(\hat{\theta})$$

Find $E(\hat{\theta})$ Moreover, find cdf : consider $F(y) = P(\hat{\theta} \leq y) = P(max(x_i) \leq y)$

Since $x_1, x_2, \cdots, x_n$ are independent variables

$\Longrightarrow F(y) = P(\hat{\theta} \leq y) = [P(x \leq y)]^n$,

then we obtain cdf $F(y) = \begin{cases} 0 & ,y < 0 \\ \dfrac{y-0}{\theta-0} & ;0 \leq y \leq \theta \\ 1 & ,y > \theta \end{cases}$

Hence, $[P(x \leq y)]^n = \left[\frac{y}{\theta}\right]^n, 0 \leq y \leq \theta$, so we get pdf $f_{\hat{\theta}}(y) = \begin{cases} n\frac{y^{n-1}}{\theta^n} & ,0 \leq y \leq \theta \\ 0 & ,otherwise \end{cases}$

then $E(\hat{\theta}) = \int_0^\theta \frac{ny^n}{\theta^n} dy = \frac{n}{n+1}\theta$

from (1) $\Longrightarrow \frac{(n+1)}{n}E(\hat{\theta}) = \frac{(n+1)}{n}(\frac{n}{n+1})\theta$

Therefore $\tilde{\theta}$ is an unbiased estimator of θ ■

Find its variance

we have $V(\tilde{\theta}) = V(\frac{(n+1)}{n}\hat{\theta}) = \frac{(n+1)^2}{n^2}V(\hat{\theta})$

We know that $V(\hat{\theta}) = E(\hat{\theta}^2) - [E(\hat{\theta}]^2$

$\Longrightarrow E(\hat{\theta}^2) = \int_0^\theta n\frac{y^{n+1}}{\theta^n} dy = \frac{n}{n+2}\theta^2$

$\Longrightarrow V(\tilde{\theta}) = \frac{(n+1)^2}{n^2}\Big[\frac{n}{n+2}\theta^2 - [\frac{n+1}{n}\theta]^2\Big] = \frac{1}{n(n+2)}\theta^2$

Therefore $V(\tilde{\theta}) = \frac{1}{n(n+2)}\theta^2$ ■

c). find the Cramer-Rao lower bound for the variance of an unbiased estimator of θ

we have $V(\theta) \geq -\frac{1}{E\left[\frac{\partial^2 \ln(L(x;\theta))}{\partial\theta^2}\right]}$

$\Longrightarrow V(\theta) \geq \frac{-1}{E(\frac{n}{\theta^2}})) = -\frac{1}{\frac{n}{\theta}^2} = -\frac{\theta^2}{n}$

Therefore the Cramer-Rao lower bound of θ is $-\frac{\theta^2}{n}$ ■

18). An estimator $\hat{\theta}$ is said to be consistent if for any$\epsilon > 0, P(|\hat{\theta} - \theta| \geq \epsilon) \to 0 as, n \to \infty$. That is , $\hat{\theta}$is consistent if , as the sample size gets larger, it is less and less likely that $\hat{\theta}$ will be further than ϵ from the true value of θ. Show that $\bar{X}$ is a consistent estimator of μ when $\sigma^2 < \infty$ by using Chebyshev's inequality . Hint (Chebyshev's inequality) Let X be a random variable with finite exoected value μ finite non-zero variance σ^2. Then for any real number $k > 0$,

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$$

By using Chebyshev's inequality : $P(|\bar{X} - \mu| \geq \epsilon) \leq \frac{\sigma^2}{n\epsilon^2}$

for $n \to]infty$ then $p(|\bar{X} - \mu| \geq \epsilon) \to 0$

Therefore , $\bar{X}$ is a consistently estimator of μ ■