9, 13)

A

Section 6.2 (20, 21, 29, 32)

1. The accompanying data on flexural strength (MPa) for concrete beams of a certain type was introduced in Example 1.2.

5.9	7.2	7.3	6.3	8.1	6.8	7.0
7.6	6.8	6.5	7.0	6.3	7.9	9.0
8.2	8.7	7.8	9.7	7.4	7.7	9.7
7.8	7.7	11.6	11.3	11.8	10.7	

a. Calculate a point estimate of the mean value of strength for the conceptual population of all beams manufactured in this fashion, and state which estimator you used. [*Hint*: $\Sigma x_i = 219.8$.]

b. Calculate a point estimate of the $\boxed{\text{strength value}}$ that separates the weakest 50% of all such beams from the strongest 50%, and state which estimator you used.

c. Calculate and interpret a point estimate of the population standard deviation σ. Which estimator did you use? [*Hint*: $\Sigma x_i^2 = 1860.94$.]

d. Calculate a point estimate of the proportion of all such beams whose flexural strength exceeds 10 MPa. [*Hint*: Think of an observation as a "success" if it exceeds 10.]

e. Calculate a point estimate of the population coefficient of variation σ/μ, and state which estimator you used.

1.

a) use the estimator $= \bar{X}$

estimate $= \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{219.8}{27} \approx 8.14$

b) use the estimator $= \tilde{X}$

since the median can divide the data into two part, one part larger than it, the other part smaller than it.

since the median $\tilde{x} = 7.7$

c) use the estimator that the sample standard deviation

the population standard deviation can be estimated:

$$\sigma \approx \sqrt{\frac{\Sigma(x_i-\bar{x})^2}{n-1}} = \sqrt{\frac{\Sigma x_i^2 - (\Sigma x_i)^2/n}{n-1}}$$

$$= \sqrt{\frac{1860.94 - 1789.33}{26}}$$

$$\approx 1.659$$

d) X = the number of the beams whose flexural strength exceeds 10MPa.

(it can be seen as a binomial rv) estimator $\hat{P} = \frac{X}{n} = \frac{4}{27} \approx 0.148$

e) estimator the sample standard deviation and the sample mean.

$$\frac{\sigma}{\mu} \approx \frac{1.659}{8.14} = 0.204$$

8. In a random sample of 80 components of a certain type, 12 are found to be defective.

a. Give a point estimate of the proportion of all such components that are *not* defective.

b. A system is to be constructed by randomly selecting two of these components and connecting them in series, as shown here.

The series connection implies that the system will function if and only if neither component is defective (i.e., both components work properly). Estimate the proportion of all such systems that work properly. [*Hint:* If p denotes the probability that a component works properly, how can P(system works) be expressed in terms of p?]

the proportion of all components that are not defective:

a). $\hat{p} = \dfrac{80-12}{80} = 0.85$

b) $\hat{p}$ denotes the probability that a component work properly.

$P(\text{system works}) = p^2 = 0.7225$.

9. Each of 150 newly manufactured items is examined and the ... r item is recorded (the items are sup- ... atches), yielding the following data:

Number of scratches per item	0	1	2	3	4	5	6	7
Observed frequency	18	37	42	30	13	7	2	1

Let X = the number of scratches on a randomly chosen item, and assume that X has a Poisson distribution with parameter μ.

$$P(x,\lambda)=\frac{e^{-\lambda}\lambda^x}{x!}$$

a. Find an unbiased estimator of μ and compute the estimate for the data. [*Hint*: $E(X)=\mu$ for X Poisson, so $E(\bar{X})=?$]

b. What is the standard deviation (standard error) of your estimator? Compute the estimated standard error. [*Hint*: $\sigma_X^2=\mu$ for X Poisson.]

a.)

$E(X)=\mu$ for Poisson distribution.

since I want to find the unbiased estimator of μ.

and we know that the sample mean is the unbiased estimator of the population mean.

the sample mean $\hat{\mu}=\frac{1}{n}\sum_{i=1}^{n}X_i = 0\times\frac{18}{150}+1\times\frac{37}{150}+2\times\frac{42}{150}+3\times\frac{30}{150}+4\times\frac{13}{150}+5\times\frac{7}{150}+6\times\frac{2}{150}+7\times\frac{1}{150}=2.113.$

b) since for possion distribution.

(the standard error of the sample mean) $S(\hat{\mu})=\frac{\mu}{n}$ (since $\hat{\mu}$ is the mean of the random sample)

$\sigma(\hat{u})=\sqrt{\frac{\mu}{n}}=0.119.$

13. Consider a random sample $X_1,\ldots,X_n$ from the pdf

$$f(x;\theta)=.5(1+\theta x)\qquad -1\le x\le 1$$

where $-1\le\theta\le 1$ (this distribution arises in particle physics). Show that $\hat{\theta}=3\bar{X}$ is an unbiased estimator of θ. [*Hint*: First determine $\mu=E(X)=E(\bar{X})$.] $\rightarrow$ for random sample

since I want to show $\hat{\theta}=3\bar{X}$ is an unbiased estimator of θ

we need to prove $E(\bar{X})=\frac{\theta}{3}$

$f(x;\theta)=\frac{1}{2}(1+\theta x)$

$$E(X)=\int_{-1}^{1}x\cdot f(x;\theta)\,dx$$

$$=\frac{1}{2}\int_{-1}^{1}(x+\theta x^2)\,dx$$

$$=\frac{1}{2}\left[\frac{1}{2}x^2+\frac{\theta}{3}x^3\right]_{-1}^{1}$$

$$=\frac{\theta}{3}$$

since $X_1\ldots X_n$ are random sample,

the mean of them are $E(X)$, $E(\bar{X})=E(\frac{1}{n}\sum_{i=1}^{n}X_i)=\frac{1}{n}\sum_{i=1}^{n}E(X_i)=\frac{1}{n}\cdot n\cdot\frac{\theta}{3}=\frac{\theta}{3}$

$E(\theta=3\bar{X})=3E(\bar{X})=3\cdot\frac{\theta}{3}=\theta$ (equal!)

so $\theta=3\bar{X}$ is the unbiased estimator of θ

20. A 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 positives) 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-free individuals that are positive). Assume that only X is available rather than the actual sequence of test results.

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

b. Is the estimator of part (a) unbiased?

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

a. since the proportion of positive is p,
then $1-p$ is the proportion that is negative.

n disease-free individual.
X test is positive
that its probability is p^x
the $n-X$ are test negative.
that its probability is $(1-p)^{n-x}$.

that. $f(p) = p^x(1-p)^{n-x}$ (use the maximum likelihood estimator).

$$\ln f(p) = \ln p^x(1-p)^{n-x}$$
$$= x\ln p + (n-x)\ln(1-p)$$

to find the max of $f(p)$. (when the p is).

$$\frac{d\ln f(p)}{dp} = \frac{x}{p} - \frac{n-x}{1-p} = 0$$

that $\hat{p} = \frac{X}{n}$.

when $n=20$, $X=3$. $\hat{p} = 0.15$.

b) the unbiased estimation: if $E(\hat{p}) = p$

since $X \sim B(n,p)$

$E(X) = np$

$E(\hat{p}) = E(\frac{X}{n}) = \frac{E(X)}{n} = p$ so that $\hat{p}$ is unbiased.

c) to find the MLE of $(1-p)^5$

$\hat{p} = \frac{X}{n}$.

MLE of $(1-p)^5 = (1-\frac{X}{n})^5$

when $n=20$, $X=3$ $(1-p)^5 = (\frac{17}{20})^5$

21. Let X have a Weibull distribution with parameters α and β, so

$$E(X) = \beta \cdot \Gamma(1 + 1/\alpha)$$
$$V(X) = \beta^2\{\Gamma(1 + 2/\alpha) - [\Gamma(1 + 1/\alpha)]^2\}$$

a. Based on a random sample $X_1, \ldots, X_n$, write equations for the method of moments estimators of β and α. Show that, once the estimate of α has been obtained, the estimate of β can be found from a table of the gamma function and that the estimate of α is the solution to a complicated equation involving the gamma function.

b. If $n = 20$, $\bar{x} = 28.0$, and $\sum x_i^2 = 16{,}500$, compute the estimates. [*Hint*: $[\Gamma(1.2)]^2/\Gamma(1.4) = .95$.]

a) for the sample.

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad ①$$

$$S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1} \quad ②$$

by the method of Moment.

$$\bar{X} = \beta \cdot \Gamma(1+\tfrac{1}{a}) \quad ③$$

$$S^2 = \beta^2\{\Gamma(1+2/a) - [\Gamma(1+1/a)]^2\} \quad ④$$

by ③, $\beta = \dfrac{\bar{X}}{\Gamma(1+\frac{1}{a})}$ ⑤

by ④ ⑤

$$S^2 = \left(\frac{\bar{X}}{\Gamma(1+\frac{1}{a})}\right)^2 \{\Gamma(1+\tfrac{2}{a}) + [\Gamma(1+\tfrac{1}{a})]^2\}$$

here if we get the estimate of a $(\hat{a})$.

we can find $\hat{\beta} = \dfrac{\bar{X}}{\Gamma(1+\frac{1}{\hat{a}})}$

b)

$n=20$, $\bar{X}=28.0$ $\quad \sum x_i^2 = 16500$ $\quad$ Let $A = \Gamma(1+\frac{1}{a})$ $\quad B = \Gamma(1+\frac{2}{a})$

$$S^2 = \frac{\sum X_i^2 - (\sum X_i)^2/n}{n-1} = 43.16 = \frac{784}{A^2}(B + A^2)$$

it can represent the relation between $\Gamma(1+\frac{2}{a})$ and $\Gamma(1+\frac{1}{a})$.

since $\dfrac{\Gamma(1.2)}{\Gamma(1.4)} = 0.95$. $\quad \beta = \dfrac{\bar{X}}{\Gamma(1+\frac{1}{a})}$

$\Gamma(1+\frac{1}{a}) \approx -1.114$.

$\hat{\beta} = -25.13$.

a random sample $X_1, X_2, \ldots, X_n$ from the shifted exponential pdf

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

Taking $\theta = 0$ gives the pdf of the exponential distribution considered previously (with positive density to the right of zero). An example of the shifted exponential distribution appeared in Example 4.5, in which the variable of interest was time headway in traffic flow and $\theta = .5$ was the minimum possible time headway.

a. Obtain the maximum likelihood estimators of θ and λ.

b. If $n = 10$ time headway observations are made, resulting in the values 3.11, .64, 2.55, 2.20, 5.44, 3.42, 10.39, 8.93, 17.82, and 1.30, calculate the estimates of θ and λ.

a) maximum likelihood estimator.

taking $\theta = 0$

$$f(x_1, \ldots x_n; \lambda) = (\lambda e^{-\lambda x_1}) \ldots (\lambda e^{-\lambda x_n}) = \lambda^n e^{-\lambda \sum x_i}.$$

$$\frac{d f(x_1 \ldots x_n; \lambda)}{d\lambda} = \frac{n}{\lambda} - \sum x_i = 0$$

we get $\hat{\lambda} = \frac{1}{\bar{X}}$ the estimator.

if $\theta \neq 0$.

$$t = f(x_1 \ldots x_n; \theta, \lambda) = \prod_{i=1}^{n} \lambda e^{-\lambda(x_i - \theta)}$$

$$\ln t = n \log \lambda - \lambda \sum_{i=1}^{n} (x_i - \theta)$$

let $\frac{\partial \ln t}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} (x_i - \theta)$

let it equal to 0

$\hat{\lambda} = \frac{n}{\sum_{i=1}^{n} (x_i - \hat{\theta})}$. is the estimator.

b) since we first need to find $\hat{\theta}$. (since I want $\ln t$ be the max) $\hat{\theta} = \min(x_1 \ldots x_n)$

$\hat{\theta} = 0.64$ from the data

$$\hat{\lambda} = \frac{10}{\sum_{i=1}^{10} (x_i - 0.64)}$$

32. a. Let $X_1, \ldots, X_n$ be a random sample from a uniform distribution on $[0, \theta]$. Then the mle of θ is $\hat{\theta} = Y = \max(X_i)$. Use the fact that $Y \leq y$ iff each $X_i \leq y$ to derive the cdf of Y. Then show that the pdf of $Y = \max(X_i)$ is

$$f_Y(y) = \begin{cases} \dfrac{ny^{n-1}}{\theta^n} & 0 \leq y \leq \theta \\ 0 & \text{otherwise} \end{cases}$$

b. Use the result of part (a) to show that the mle is biased but that $(n + 1)\max(X_i)/n$ is unbiased.

累计分布函数.
体现变量小于某个值的概率.

relation between $f(x)$ and $F(x)$

① Given $f(x)$, find $F(x)$

$F(x) = \int_{-\infty}^{x} f(y)dy$

② Given $F(x)$, find $f(x)$

$f(x) = \frac{d}{dx} F(x)$

(PDF是随机变量在某个值附近的概率分布情况)

a) since X_i are independent.

$P(Y \leq y) = P(X_1 \leq y, X_2 \leq y \ldots X_n \leq y)$

since X_i is a uniform distribution.

that $P(X_i \leq y) = \frac{y}{\theta}$ (when $0 \leq y \leq \theta$)

so that $P(Y \leq y) = \left(\frac{y}{\theta}\right)^n$

then we derivate the CDF,

we get $f_Y(y) = \frac{d}{dy}\left(\frac{y}{\theta}\right)^n = \frac{ny^{n-1}}{\theta^n}$ $(0 \leq y \leq \theta)$

when $y < 0$ or $y > \theta$, $f_Y(y) = 0$.

b). the biase is $E[\hat{\theta}] - \theta$.

the PDF of Y is $f_Y(y) = \frac{ny^{n-1}}{\theta^n}$.

$$E(Y) = \int_0^\theta \frac{ny^{n-1}}{\theta^n} dy = \frac{n}{\theta^n} \cdot \frac{1}{n} y^n \Big|_0^\theta = \frac{n}{n+1}\theta \neq \theta$$

so that $\hat{\theta} = Y$ is biased.

but when we use $\frac{n+1}{n}\max(X_i)$. $E\left(\frac{n+1}{n}\max(X_i)\right) = \frac{n+1}{n} E(Y) = \theta$.

so that it is unbiased.