Solution to Exercises

Chapter 6 Introduction to Statistical Inference

Section 6.1 Point Estimation

6.1. Let $X_1, X_2, \cdots, X_n$ represent a random sample from each of the distributions having the following probability density functions：

(a） $f(x,\theta) = \theta^x e^{-\theta} / x!$, $x = 0, 1, 2, \cdots, 0 \le \theta < \infty$, zero elsewhere, where $f(0,0) = 1$.

(b) $f(x,\theta) = \theta x^{\theta-1}$, $0 < x < 1, 0 < \theta < \infty$, zero elsewhere.

(c) $f(x,\theta) = \dfrac{1}{\theta} e^{-x/\theta}$, $0 < x < \infty, 0 < \theta < \infty$, zero elsewhere.

In each case find the m. l. e. $\hat{\theta}$ of θ.

Solution

(a) The likelihood function of the sample is

$$L(x;\theta) = \theta^{\sum x_i} e^{-n\theta} / (x_1! x_2! \cdots x_n!$$

Here

$$\ln L(x;\theta) = \sum x_i \ln\theta - n\theta - \sum \ln x_i!$$

So we have

$$\frac{d \ln L(\theta)}{d\theta} = \frac{\sum x_i}{\theta} - n = 0.$$

whose solution for θ is $\bar{x}$ which is the desired m .l. e. of the unknown parameter θ.

(b) The likelihood function of the sample is

$$L(x;\theta) = \theta^n (x_1 x_2 \cdots x_n)^{\theta-1}$$

Here

$$\ln L(x;\theta) = n \ln\theta + (\theta - 1)(\sum \ln x_i)$$

So we have

$$\frac{d \ln L(\theta)}{d\theta} = \frac{n}{\theta} + \sum \ln x_i = 0.$$

whose solution for θ is $-n / \sum \ln x_i$ which is the desired m. l. e. of the unknown parameter θ.

(c) The likelihood function of the sample is

$$L(x;\theta) = \frac{1}{\theta^n} e^{-\sum x_i / \theta}$$

Here

$$\ln L(x;\theta) = -n \ln\theta - \sum x_i / \theta$$

So we have

$$\frac{d \ln L(\theta)}{d\theta} = -\frac{n}{\theta} + \sum x_i / \theta^2 = 0.$$

whose solution for θ is $\bar{x}$ which is the desired m. l. e. of the unknown parameter θ.

6.2. Let $X_1, X_2, \cdots, X_n$ be i. i. d., each with the distribution having p. d. f.

$f(x;\theta_1,\theta_2) = (1/\theta_2) e^{-(x-\theta_1)/\theta_2}$, $\theta_1 \le x < \infty, -\infty < \theta_1 < \infty, 0 < \theta_2 < \infty$, zero elsewhere. Find the m. l. e. of θ_1 and θ_2.

Solution

Given θ_2**,** it is easily verify that the first order statistic can maximize the likelihood function, so the m. l. e. of θ_1 is the first order statistic Y_1.

The likelihood function of the sample is

$$L(x;\theta_1,\theta_2) = (1/\theta_2{}^n)e^{-\Sigma(x_i-\theta_1)/\theta_2},\ \theta_1 \le x < \infty, -\infty < \theta_1 < \infty, 0 < \theta_2 < \infty.$$

$$\ln L(x;\theta_1,\theta_2) = -n\ln\theta_2 - \Sigma(x_i-\theta_1)/\theta_2$$

We observe that we may maximize by differentiation. We have

$$\frac{\partial \ln L}{\partial\theta_2} = -\frac{n}{\theta_2} + \frac{\Sigma(x_i-\theta_1)}{\theta_2{}^2} = 0$$

whose solution is $\Sigma(X_i - Y_1)/n$ which is the m. l. e. of the unknown parameter θ_2**.**

6.3. Let $Y_1 < Y_2 < \cdots < Y_n$ be the order statistics of a random sample from a distribution with p. d. f. $f(x;\theta) = 1,\ \theta - \frac{1}{2} \le x \le \theta + \frac{1}{2}, -\infty < \theta < \infty$, zero elsewhere. Show that every statistic $u(X_1, X_2, \cdots, X_n)$ such that

$$Y_n - \frac{1}{2} \le u(X_1, X_2, \cdots, X_n) \le Y_n + \frac{1}{2}$$

is a m. l. e. of θ. In particular, $(4Y_1 + 2Y_n + 1)/6, (Y_1 + Y_n)/2$ and $(2Y_1 + 4Y_n - 1)/6$ are three such statistics. Thus the uniqueness is not in general a property of a m. l. e.

Solution

According to the definition of the order statistic, we have

$$\theta - \frac{1}{2} \le Y_1 < X_2 < \cdots < Y_n) \le \theta + \frac{1}{2}.$$

From the inequality, we obtain

$$Y_n - \frac{1}{2} \le \theta \le Y_1 + \frac{1}{2}$$

which means that any statistic $u(X_1, X_2, \cdots, X_n)$ such that

$$Y_n - \frac{1}{2} \le u(X_1, X_2, \cdots, X_n) \le Y_1 + \frac{1}{2}$$

is a m. l. e. of θ. Particularly, the statistics $Y_n - \frac{1}{2}$ and $Y_1 + \frac{1}{2}$ are both m. l. e. of θ. Furthermore, any weighty average of the two statistics is m. l. e.

Since the statistics can be formulated as

$$(4Y_1 + 2Y_n + 1)/6 = \frac{2}{6}(Y_n - \frac{1}{2}) + \frac{4}{6}(Y_1 + \frac{1}{2}),$$
$$(Y_1 + Y_n)/2 = \frac{1}{2}(Y_n - \frac{1}{2}) + \frac{1}{2}(Y_1 + \frac{1}{2}),$$
$$(2Y_1 + 4Y_n - 1)/6 = \frac{4}{6}(Y_n - \frac{1}{2}) + \frac{2}{6}(Y_1 + \frac{1}{2}).$$

So $(4Y_1 + 2Y_n + 1)/6, (Y_1 + Y_n)/2$ and $(2Y_1 + 4Y_n - 1)/6$ are three m. l. e. of θ.

6.4. Let X_1, X_2 and X_3 have the multinomial distribution in which $n = 25, k = 4$, and the unknown probabilities are θ_1, θ_2 and θ_3,respectively. Here we can, for convenience, let $X_4 = 25 - X_1 - X_2 - X_3$ and $\theta_4 = 1 - \theta_1 - \theta_2 - \theta_3$. If the observed values of the random variables are $x_1 = 4, x_2 = 11$, and $x_3 = 7$, find the m. l. e. of θ_1, θ_2 and θ_3.

Solution

It is easily to understand that

$$X_i \sim b(25, \theta_i), i = 1,2,3$$

So the m. l. e. of the unknown parameters is $\overline{X}_1,\ \overline{X}_2, \overline{X}_3$, respectively.

Thus the m. l. e. of θ_1, θ_2 and θ_3 is $\frac{4}{25}, \frac{11}{25}, \frac{7}{25}$, respectively.

6.5. The Pareto distribution is frequently used as a model in study of incomes and has the distribution function

$$F(x;\theta_1,\theta_2) = 1-(\theta_1/x)^{\theta_2},\ \ \theta_1 \le x, \text{zero elsewhere},$$
$$\text{where } \theta_1 > 0 \quad \text{and } \theta_2 > 0$$

If $X_1, X_2, \cdots, X_n$ is a random sample from this distribution, find the m. l. e. of θ_1 and θ_2.

Solution

The p. d. f. of the population is

$$f(x;\theta_1,\theta_2) = \frac{\theta_2 {\theta_1}^{\theta_2}}{x^{\theta_2+1}}, \quad \theta_1 \le x .$$

Obviously, the m. l. e. of θ_1 is the first order statistic Y_1.

The likelihood function of the sample is

$$L(x;\theta_1,\theta_2) = \frac{{\theta_2}^n {\theta_1}^{n\theta_2}}{{x_1}^{\theta_2+1}{x_2}^{\theta_2+1}\cdots {x_n}^{\theta_2+1}},\ \ \frac{\partial \ln L}{\partial \theta_2} = \frac{n}{\theta_2} + n\ln\theta_1 - \sum \ln x_i = 0$$

Thus we obtain the m. l. e. of θ_2

$$\hat{\theta}_2 = \frac{n}{\sum \ln X_i - n\ln Y_1}.$$

6.6. Let Y_n be a statistic such that $\lim_{n\to\infty} E(Y_n) = \theta$ and $\lim_{n\to\infty} {\sigma_{Y_n}}^2 = 0$. Prove that Y_n is consistent estimator of θ.

Proof

Since

$$E[(Y_n-\theta)^2] = E[(Y_n - E(Y_n) + E(Y_n) - \theta)^2] = {\sigma_{Y_n}}^2 + [E(Y_n - \theta)]^2,$$

So, in accordance with Chebyshev's inequality, we have

$$\Pr(|Y_n - \theta| \ge \varepsilon) \le \frac{E[(Y_n-\theta)^2]}{\varepsilon^2} = \frac{{\sigma_{Y_n}}^2 + [E(Y_n-\theta)]^2}{\varepsilon^2} \to 0, \qquad \text{as } n\to\infty$$

for every $\varepsilon > 0$.

Thus according to the definition of consistent estimator, we complete the proof.

6.7. For each of the distributions in Exercise 6.1, find an estimator of θ by the method of moments and show that it is consistent.

Solution

(1) It is obvious that the population is Poisson distribution with parameter θ.

So $E(X) = \theta$. Let $\theta = \overline{X}$. We get the estimator of θ by the method of moments is the sample mean $\overline{X}$.

$E(\overline{X}) = \theta, V(\overline{X}) = \frac{\theta}{n}$. For any $\varepsilon > 0$, we have

$$\Pr(|\overline{X}-\theta|\geq\varepsilon)\leq\frac{V(\overline{X})}{\varepsilon^2}=\frac{\theta}{n\varepsilon^2}\to 0,\qquad \text{as } n\to\infty$$

Thus the sample mean $\overline{X}$ is a consistent estimator of the population mean θ.

(2) The population mean is

$E(X)=\int_0^1 \theta x^{\theta}\mathrm{d}x=\frac{\theta}{\theta+1}$. In accordance with the idea of method of moments, let

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

We have $\hat{\theta}=\frac{\overline{X}}{1-\overline{X}}$ which is the moment estimator of θ.

(3) In fact, the population is Gamma distribution with parameters 1 and θ. So $E(X)=\theta$.

Thus the estimator of θ by method of moments is the sample mean $\overline{X}$. It is easily to verify that the sample $\overline{X}$ converges in probability to the population mean θ, so $\overline{X}$ is a consistent estimator of θ.

6.8. If a random sample of size n is taken from a distribution having p. d. f. $f(x;\theta)=2x/\theta^2, 0<x\leq\theta,$ zero elsewhere, find

(a) The m. l. e. $\hat{\theta}$ for θ.

(b) The constant c so that $E(c\hat{\theta})=\theta$.

(c) The m. l. e. for the median of the distribution.

Solution

(a) The likelihood function of the sample is

$$L(x;\theta)=2^n x_1x_2\cdots x_n/\theta^{2n}.$$

It is obvious that the nth order statistic Y_n can maximize the likelihood function, so the m. l. e. $\hat{\theta}$ for θ is the nth order statistic Y_n.

b) (b) since the p. d. f. of Y_n is $f(y_n)=2ny_n{}^{2n-1}/\theta^{2n}, 0<y_n\leq\theta$, thus

$$E(Y_n)=\frac{2n}{2n+1}\theta$$

So $c=\frac{2n+1}{2n}$.

(c) Since $\frac{1}{2}=\int_0^m 2x/\theta^2dx=m^2/\theta^2,\qquad m=\theta/\sqrt{2}$, In accordance with the invariant property of m. l. e. we have

$$\hat{m}=Y_n/\sqrt{2}.$$

6.9. Let $X_1,X_2,\cdots,X_n$ be i. i. d., each with a distribution with p. d. f. $f(x;\theta)=(1/\theta)e^{-x/\theta}, 0<x<\infty$, zero elsewhere. Find the m. l. e. of $\Pr(X\leq 2)$.

Solution

It is not difficult to find that the m. l. e. of θ is the sample mean $\overline{X}$.

$$\Pr(X \le 2) = \int_0^2 (1/\theta) e^{-x/\theta} dx = 1 - e^{-2/\theta}.$$

In accordance with the invariance property of m. l. e., the m. l. e. of $\Pr(X \le 2)$ is $1 - e^{-2/\bar{X}}$.

6.10. Let X have a binomial distribution with parameters n and p. The variance of X/n is $p(1-p)/n$; This is sometimes estimated by the m. l. e. $\frac{X}{n}(1-\frac{X}{n})/n$. Is this an unbiased estimator of $p(1-p)/n$? If not, can you construct one by multiplying this one by a constant?

Solution

Since

$$E[\frac{X}{n}(1-\frac{X}{n})/n] = E(X/n)/n - E(X^2)/n^3 = p/n - \{V(X)+[E(X)]^2\}/n^3 = p/n - p^2/n - pq/n^2 = \frac{(n-1)pq}{n^2}.$$

So $\frac{X}{n}(1-\frac{X}{n})/n$ multiplied by $n/(n-1)$ becomes an unbiased estimator of $p(1-p)/n$.

6.11. Let the table

x	0	1	2	3	4	5
Frequency	6	10	14	13	6	1

represent a summary of a sample of size 50 from a binomial distribution having $n=5$. Find the m. l. e. of $\Pr(X \ge 3)$.

Solution

Since the population $X \sim b(n,\theta)$ .So the m. l. e. of $\Pr(X \ge 3)$ is (13+6+1)/50=2/5.

6.12. Let $Y_1 < Y_2 < \cdots < Y_n$ be the order statistics of a random sample of size n from the uniform distribution of the continuous type over the closed interval $[\theta - p, \theta + p]$. Find the maximum likelihood estimators of θ and p. Are these two unbiased estimators?

Solution

Since

$$\theta - p \le Y_1 < Y_2 < \cdots < Y_n \le \theta + p$$

The p. d. f. of the distribution is

$$f(x;\theta,p) = \frac{1}{2p}, \theta - p \le x \le \theta + p,$$

it is obvious that $f(x;\theta,p)$ is a decreasing function of the parameter p,

however,

$$Y_n - Y_1 \le 2p,$$

so the m. l. e. of p is $(Y_n - Y_1)/2$.

On the other hand,

$$Y_n - p \le \theta \le Y_1 + p$$

Obviously, the weighty average $\frac{1}{2}(Y_n - \hat{p}) + \frac{1}{2}(Y_1 + \hat{p}) = (Y_n + Y_1)/2$ of $Y_n - \hat{p}$ and $Y_1 + p$ is a m. l. e. of the parameter θ.

It is not difficult to compute the following

$$E[(Y_n + Y_1)/2] = \theta,\ E[(Y_n - Y_1)/2] = p(n-1)/(n+1).$$

So $(Y_n + Y_1)/2$ is an unbiased estimator of θ while $(Y_n - Y_1)/2$ is not.

Section 6.2 Confidence Intervals for Means

6.14. Let the observed value of the mean $\overline{X}$ of a random sample of size 20 from a distribution that is $N(\mu,80)$ be 81.2. Find a 95 percent confidence interval for μ.

Solution

Since $\overline{X} \sim N(\mu,4)$, so $\frac{\overline{X}-\mu}{2} \sim N(0,1)$ .And the 97.5 percent quantile of the distribution from Table III in Appendix B is 1.96. Thus we have

$$\Pr(-1.96 \le \frac{\overline{X}-\mu}{2} \le 1.96) = \Pr(\overline{X}-3.92 \le \mu \le \overline{X}+3.92) = 0.95 .$$

If the observed value of the mean $\overline{X}$ is 81.2, then a 95 percent confidence interval for μ is

$(81.2-3.92, 81.2+3.92) = (77.28, 85.12)$.

6.15. Let $\overline{X}$ be the mean of a random sample of size n from a distribution that is $N(\mu,9)$. Find n such that

$\Pr(\overline{X}-1 < \mu < \overline{X}+1) = 0.90$, approximately.

Solution

Since $\overline{X} \sim N(\mu,9/n)$, we have

$$\Pr(\overline{X}-1 < \mu < \overline{X}+1) = \Pr(-1 < \overline{X}-\mu < 1) = \Pr(-\frac{1}{\sqrt{9/n}} < \frac{\overline{X}-\mu}{\sqrt{9/n}} < \frac{1}{\sqrt{9/n}}) = 090 .$$

From Table III in Appendix B , we have

$$\frac{1}{\sqrt{9/n}} = 1.645 ,$$

Approximately, we obtain $n = 24$ or $n = 25$.

6.16. Let a random sample of size 17 from the normal distribution $N(\mu,\sigma^2)$ yield $\bar{x} = 4.7$ and $s^2 = 5.76$. Determine a 90 percent confidence interval for μ.

Solution

Since $\frac{\sqrt{n-1}(\overline{X}-\mu)}{S} \sim t(n-1)$, we get the 95 percent quantile from Table IV in Appendix A being 1.746.

The events

$$-1.746 < \frac{\sqrt{n-1}(\overline{X}-\mu)}{S} < 1.746$$

and

$$\overline{X}-1.746S/\sqrt{n-1} < \mu < \overline{X}+1.746S/\sqrt{n-1}$$

are equivalent to each other.

So if $\bar{x} = 4.7$ and $s^2 = 5.76$, a 90 percent confidence interval for μ is

$$(\bar{x}-1.746s/\sqrt{17-1}, \bar{x}-1.746s/\sqrt{17-1}) = (3.6524, 5.7476) .$$

6.17. Let $\overline{X}$ denote the mean of a random sample of size n from a distribution that has mean μ and variance $\sigma^2 = 10$. Find n so that the probability is approximately 0.954 that the random interval $(\overline{X} - \frac{1}{2}, \overline{X} + \frac{1}{2})$ includes μ.

Solution

In accordance with the central limit theorem, approximately, X is normally distributed with mean μ and variance $10/n$.

So approximately,

$$\Pr(\overline{X} - \frac{1}{2} < \mu < \overline{X} + \frac{1}{2}) = \Pr(-\frac{1}{2\sqrt{10/n}} < \frac{\overline{X} - \mu}{\sqrt{10/n}} < \frac{1}{2\sqrt{10/n}}) = 0.954.$$

Let $\frac{1}{2\sqrt{10/n}} = 2$, we obtain $n = 160$.

6.18. Let $X_1, X_2, \cdots, X_n$ be a random sample of size 9 from a distribution that is $N(\mu, \sigma^2)$.

(a) If σ is known, find the length of a 95 percent confidence interval for μ if this interval is based on the random variable $\sqrt{9}(\overline{X} - \mu)/\sigma$.

(b) If σ is unknown, find the expected value of the length of a 95 percent confidence interval for μ if this interval is based on the random variable $\sqrt{8}(\overline{X} - \mu)/S$.

Solution

(a) If σ is known, the statistic $\sqrt{9}(\overline{X} - \mu)/\sigma \sim N(0,1)$, so we have

$$\Pr(-1.96 < \sqrt{9}(\overline{X} - \mu)/\sigma < 1.96) = \Pr(\overline{X} - 1.96\sigma/\sqrt{9} < \mu < \overline{X} + 1.96\sigma/\sqrt{9}) = 0.95.$$

The length of the interval is

$$(\overline{X} + 1.96\sigma/\sqrt{9}) - (\overline{X} - 1.96\sigma/\sqrt{9}) = 2 \times 1.96\sigma/3 = 1.3\sigma.$$

(b) If σ is unknown, the statistic $\sqrt{8}(\overline{X} - \mu)/S \sim t(8)$, so we have

$$\Pr(-2.306 < \sqrt{8}(\overline{X} - \mu)/S < 2.306) = \Pr(\overline{X} - 2.306S/\sqrt{8} < \mu < \overline{X} + 2.306S/\sqrt{8}) = 0.95.$$

The length of the interval is

$$(\overline{X} + 2.306S/\sqrt{8}) - (\overline{X} - 2.306S/\sqrt{8}) = 2 \times 2.306S/\sqrt{8} = 1.6S.$$

6.19. Let $X_1, X_2, \cdots, X_n, X_{n+1}$ be a random sample of size $n+1, n>1$, from a distribution that is $N(\mu, \sigma^2)$. Let $\overline{X} = \sum_{i=1}^{n} X_i$ and $S^2 = \sum_{i=1}^{n}(X_i - \overline{X})/n$. Find the constant c so that the statistic $c(\overline{X} - X_{n+1})/S$ has a t-distribution. If $n = 8$, determine k such that $\Pr(\overline{X} - kS < X_9 < \overline{X} + kS) = 0.80$. The observed interval $(\overline{x} - ks,\ \overline{x} + ks)$ is often called an 80 percent *prediction interval* for X_9.

Solution

Since $\bar{X} \sim N(\mu, \sigma^2/n), nS^2/\sigma^2 \sim \chi^2(n)$, so $\bar{X} - X_{n+1} \sim N(0, \frac{n+1}{n}\sigma^2), \frac{\bar{X} - X_{n+1}}{S\sqrt{n+1}} \sim t(n-1)$,

So $c = 1/\sqrt{n+1}$.

If $n = 8$, $\Pr(\bar{X} - 1.415 \times 3S < X_9 < \bar{X} + 1.415 \times 3S) = 0.80$, thus

$$k = 1.415 \times 3 = 4.248.$$

6.20. Let Y be $b(300, p)$ .If the observed value of Y is $y = 75$, find an approximate 90 percent confidence interval for p .

Solution

In accordance with the central limit theorem, we approximately have

$$\frac{Y - np}{\sqrt{n(Y/n)(1 - Y/n)}} \sim N(0,1) .$$

Thus approximately,

$$\Pr(-1.645 < \frac{Y - np}{\sqrt{n(Y/n)(1 - Y/n)}} < 1.645) = \Pr(\frac{Y - 1.645\sqrt{n(Y/n)(1 - Y/n)}}{n} < p < \frac{Y + 1.645\sqrt{n(Y/n)(1 - Y/n)}}{n}) = 0.9$$

If the observed value of Y is $y = 75$, an approximate 90 percent confidence interval for p is

$$(0.2088, 0.2911) .$$

6.24. Let $\bar{x}$ be the observed mean of a random sample of size n from a distribution having mean μ and known variance σ^2 . Find n so that $\bar{x} - \sigma/4$ to $\bar{x} + \sigma/4$ is an approximate 95 percent confidence interval for μ .

Solution

In accordance with the central limit theorem, we approximately have

$$\frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \sim N(0,1) .$$

Thus approximately,

$$\Pr(-1.96 < \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} < 1.96) = \Pr(\bar{X} - 1.96\sigma/\sqrt{n} < \mu < \bar{X} - 1.96\sigma/\sqrt{n}) = 0.95.$$

Let $1.96/\sqrt{n} = 1/4$, we have $n = 61$ or $n = 62$.

6.25. Assume a binomial model for a certain random variable. If we desire a 90 percent confidence interval for p that is at most 0.02 in length, find n .

Solution

In accordance with the central limit theorem, we approximately have

$$\frac{Y - np}{\sqrt{n(Y/n)(1 - Y/n)}} \sim N(0,1) .$$

and

$$\sqrt{y/n(1 - y/n)} \le \sqrt{1/2(1 - 1/2)} = 1/2 .$$

According to the Exercise 6.20 in the preceding, A 90 percent confidence interval for p in length is

$$\frac{2\times 1.645\sqrt{n(Y/n)(1-Y/n)}}{n} \le \frac{1.645}{\sqrt{n}},$$

Let $\frac{1.645}{\sqrt{n}} \le 0.02$, we obtain $n = 6766$.

6.26. It is known that a random variable X has a Poisson distribution with parameter μ. A sample of 200 observations from this population has a mean equal to 3.4. Compute an approximate 90 percent confidence interval for μ.

Solution

6.27. Let $Y_1 < Y_2 < \cdots < Y_n$ denote the order statistics of a random sample of size n from a distribution that has p.d.f. $f(x) = 3x^2/\theta^3, 0 < x < \theta$, zero elsewhere.

(a) Show that $\Pr(c < Y_n/\theta < 1) = 1 - c^{3n}$, where $0 < c < 1$.

(b) If n is 4 and if the observed value of Y_n is 2.3, what is a 95 percent confidence interval for θ?

Solution

(a) The distribution function of the population is

$$F(x) = \int_0^x 3t^2/\theta^3 \mathrm{d}t = t^3/\theta^3,\ 0 < x < \theta,$$

So the p.d.f. of the nth order statistic is

$$f(y) = n[F(y)]^{n-1} f(y) = \frac{3ny^{3n-1}}{\theta^{3n}},\ 0 < y < \theta.$$

Thus

$$\Pr(c < Y_n/\theta < 1) = \Pr(\theta c < Y_n < \theta) = \int_{c\theta}^{\theta} \frac{3ny^{3n-1}}{\theta^{3n}} \mathrm{d}y = 1 - c^{3n}.$$

(b) In accordance with the preceding discussion, let $1 - c^{12} = 0.95$, we have $c = \sqrt[12]{0.05}$, thus a 95 percent confidence interval for θ is

$$(y_4, y_4/c) = (2.3, 2.3/c).$$

6.28. Let $X_1, X_2, \cdots, X_n$ be a random sample from $N(\mu, \sigma^2)$, where both parameters μ and σ^2 are unknown. A *confidence interval* for σ^2 can be found as follows. We know that nS^2/σ^2 is $\chi^2(n-1)$. Thus we can find constants a and b so that $\Pr(nS^2/\sigma^2 < b) = 0.975$ and $\Pr(a < nS^2/\sigma^2 < b) = 0.95$.

(a) Show that this second probability statement can be written as $\Pr(nS^2/b < \sigma^2 < nS^2/a) = 0.95$.

(b) If $n = 9$ and $S^2 = 7.63$, find a 95 percent confidence interval for σ^2.

(c) If μ is known, how would you modify the preceding procedure for finding a confidence interval for σ^2?]

Solution

(a) Since the events

$$a < nS^2/\sigma^2 < b \quad \text{and} \quad nS^2/b < \sigma^2 < nS^2/a$$

are equivalent.
So we have

$$\Pr(a < nS^2/\sigma^2 < b) = \Pr(nS^2/b^2 < \sigma^2 < nS^2/a^2) = 0.95 .$$

(b)) If $n = 9$ and $S^2 = 7.63$, we can be get $a = 2.18, b = 17.5$ from the Table II in the Appendix A. Thus

According to the part (a), we have a 95 percent confidence interval for σ^2 is the interval $(3.924, 31.5)$.

6.29. Let $X_1, X_2, \cdots, X_n$ be a random sample from a gamma distribution with known parameter $\alpha = 3$ and unknown $\beta > 0$. Discuss the construction of a confidence interval for β.

Solution

It is easy to verify the sample mean $\overline{X} \sim \Gamma(3n, \beta/n)$.

Section 6.3 Confidence Intervals for differences for Means

6.30. Let two independent random samples, each of size 10, from two normal distributions $N(\mu_1, \sigma^2)$ and $N(\mu_2, \sigma^2)$ yield $\bar{x} = 4.8, s_1{}^2 = 8.64, \bar{y} = 5.6, s_2{}^2 = 7.88$. Find a 95 confidence interval for $\mu_1 - \mu_2$.

Solution

From the Table IV in the Appendix A we get $b = 2.101$. And the observed value of R in the text of section 6.3 is $R = 1.355$. Thus a 95 confidence interval for $\mu_1 - \mu_2$ can be

$(\bar{x} - \bar{y} - 2.101 \times 1.355, \bar{x} - \bar{y} + 2.101 \times 1.355) = (-3.646, 2.047)$.

6.31. Let two independent random variables Y_1 and Y_2, with binomial distributions that have parameters $n_1 = n_2 = 100$, p_1 and p_2, respectively, be observed to be equal to $y_1 = 50$ and $y_2 = 40$. Determine an approximate 90 percent confidence interval for $p_1 - p_2$.

Solution

In accordance with the central limiting theorem, we approximately have

$$\Pr\left(\frac{Y_1}{n_1} - \frac{Y_2}{n_2} - 1.645U < p_1 - p_2 < \frac{Y_1}{n_1} - \frac{Y_2}{n_2} + 1.645U\right) = 0.95.$$

From the given data, the observed value of the statistic U is

$$\sqrt{y_1 / n_1(1 - y_1 / n_1) / n_1 + y_2 / n_2(1 - y_2 / n_2) / n_2} = 0.07.$$

Thus an approximate 90 percent confidence interval for $p_1 - p_2$ is

$$(0.5 - 0.4 - 2u, 0.5 - 0.4 + 2u) = (-0.04, 0.24).$$

6.32. Discuss the problem of finding a confidence interval for the difference $\mu_1 - \mu_2$ between the two means of two normal distribution if the variances $\sigma_1{}^2$ and $\sigma_2{}^2$ are known but not necessarily equal.

Solution

If the variances $\sigma_1{}^2$ and $\sigma_2{}^2$ are known but not necessarily equal, the sampling theorem of the sample mean is

$$\bar{X} \sim N\left(\mu_1, \frac{\sigma_1{}^2}{n_1}\right), \bar{Y} \sim N\left(\mu_2, \frac{\sigma_2{}^2}{n_2}\right),$$

So we have

$$\bar{X} - \bar{Y} \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1{}^2}{n_1} + \frac{\sigma_2{}^2}{n_2}\right).$$

Thus for given confidence level α, we can obtain the number a from the Table III in the Appendix A such that

$$\Pr\left((\bar{X} - \bar{Y}) - b\sqrt{\frac{\sigma_1{}^2}{n_1} + \frac{\sigma_2{}^2}{n_2}} < \mu_1 - \mu_2 < (\bar{X} - \bar{Y}) + b\sqrt{\frac{\sigma_1{}^2}{n_1} + \frac{\sigma_2{}^2}{n_2}}\right) = \alpha .$$

6.33. Discuss Exercise 6.32 when it is assumed that the variances are unknown and unequal. This is a very difficult problem, and the discussion should point out exactly where the difficulty lies. If, however, the variances are unknown but their ratio $\sigma_1{}^2 / \sigma_2{}^2$ is a known constant k, then a statistic that is a T random variable can again be used . Why?

Solution

According to the sampling theorem, we have

$$\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_1}}} \sim N(0,1)\text{, and } \frac{n_1 S_1^2}{\sigma_1^2}+\frac{n_2 S_2^2}{\sigma_2^2} \sim \chi^2(n_1+n_2-2)$$

and these two are independent each other, furthermore $\sigma_1^2/\sigma_2^2 = k$, that is $\sigma_1^2 = k\sigma_2^2$ thus we can construct the statistic

$$T=\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{n_1 S_1^2+kn_2 S_2^2}}\sqrt{\frac{kn_1 n_2(n_1+n_2-2)}{n_1+kn_2}} \sim t(n_1+n_2-2).\,.$$

Then we can apply the static T to obtain the confidence interval of $\mu_1-\mu_2$.

6.34. As an illustration of Exercise 6.33, one can let $X_1, X_2, \cdots, X_9$ and $Y_1, Y_2, \cdots, Y_{12}$ represent two independent random samples from the respective normal distributions $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$. It is given that $\sigma_1^2 = 3\sigma_2^2$, but σ_2^2 is unknown. Define a random variable which has a $t-$distribution that can be used to find a 95 percent interval for $\mu_1-\mu_2$.

Solution

Since

$$T=\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{n_1 S_1^2+kn_2 S_2^2}}\sqrt{\frac{kn_1 n_2(n_1+n_2-2)}{n_1+kn_2}} \sim t(n_1+n_2-2).$$

From the Table IV in the Appendix A, $b = 2.093$.

Thus a 95 percent interval for $\mu_1-\mu_2$ is

$$((\bar{X}-\bar{Y})-2.093\sqrt{\frac{(n_1 S_1^2 kn_2 S_2^2)(n_1+kn_2)}{kn_1 n_2(n_1+n_2-2)}}\,,(\bar{X}-\bar{Y})+2.093\sqrt{\frac{(n_1 S_1^2 kn_2 S_2^2)(n_1+kn_2)}{kn_1 n_2(n_1+n_2-2)}}\,)\text{。}$$

6.35. Let $\bar{X}$ and $\bar{Y}$ be the means of two independent random samples, each of size n, from the respective distribution $N(\mu_1, \sigma^2)$ and $N(\mu_2, \sigma^2)$, where the common variance is known. Find n such that

$$\Pr(\bar{X}-\bar{Y}-\sigma/5<\mu_1-\mu_2<\bar{X}-\bar{Y}+\sigma/5)=0.90.$$

Solution

Since

$$\Pr(\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sigma\sqrt{2/n}}) = \Pr(\bar{X}-\bar{Y}-\sigma/5<\mu_1-\mu_2<\bar{X}-\bar{Y}+\sigma/5)$$
$$=\Pr(\bar{X}-\bar{Y}-1.645\sigma\sqrt{2/n}<\mu_1-\mu_2<\bar{X}-\bar{Y}+1.645\sigma\sqrt{2/n})=0.90.$$

So we have

$$1.645\sqrt{2/n}=1/5,$$

Thus

$$n = 135 \text{ or } n = 136.$$

6.37. Let $X_1, X_2, \cdots, X_n$ and $Y_1, Y_2, \cdots, Y_m$ be two independent random samples from the respective normal distribution $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$, where the four parameters are unknown. To construct a *confidence interval for the ratio,* σ_1^2 / σ_2^2, of the variances, form the quotient of the two independent chi-square variables, each divided by its degrees of freedom, namely

$$F = \frac{\dfrac{mS_2^2}{\sigma_2^2}/(m-1)}{\dfrac{nS_1^2}{\sigma_1^2}/(n-1)},$$

where S_1^2 and S_2^2 are the respective sample variances.

(a) what kind of distribution does F have?

(b) From the appropriate table, a and b can be found so that

$$\Pr(F < b) = 0.975 \text{ and } \Pr(a < F < b) = 0.95.$$

(c) Rewrite the second probability statement as

$$\Pr\left(a \frac{nS_1^2/(n-1)}{mS_2^2/(m-1)} < \frac{\sigma_1^2}{\sigma_2^2} < b \frac{nS_1^2/(n-1)}{mS_2^2/(m-1)}\right) = 0.95.$$

The observed values, s_1^2 and s_2^2, can be inserted in these inequalities to provide a 95 percent confidence interval for σ_1^2 / σ_2^2.

Solution

(a) It follows from the sampling theorem that

$$\frac{mS_2^2}{\sigma_2^2} \sim \chi^2(m-1) \text{ and } \frac{nS_1^2}{\sigma_1^2} \sim \chi^2(n-1),$$

and these two statistics are independent with each other. Thus

$$F = \frac{\dfrac{mS_2^2}{\sigma_2^2}/(m-1)}{\dfrac{nS_1^2}{\sigma_1^2}/(n-1)} \sim F(m, n).$$

(b) For given degrees of freedom m *and* n, we can obtain a and b from the Table V in the Appendix A such that $\Pr(F < b) = 0.975$ and $\Pr(a < F < b) = 0.95$.

(c) (c) Since $\Pr(a < F < b) = 0.95$, we Rewrite the probability statement as

$$\Pr(a < F < b) = \Pr\left(a \frac{nS_1^2/(n-1)}{mS_2^2/(m-1)} < \frac{\sigma_1^2}{\sigma_2^2} < b \frac{nS_1^2/(n-1)}{mS_2^2/(m-1)}\right) = 0.95,$$

It is easy to verify the formulation.

Section 6.4 Tests of Statistical Hypotheses

6.38. Let X have a p.d.f. of the form $f(x;\theta)=\theta x^{\theta-1},\ 0<x<1$, zero elsewhere, where $\theta\in\{\theta:\theta=1,2\}$. To test the simple hypothesis $H_0:\theta=1$ against the alternative simple hypothesis $H_1:\theta=2$, use a random sample X_1,X_2 of size $n=2$ and define the critical region to be $C=\{(x_1,x_2):\frac{3}{4}\le x_1x_2\}$. Find the power function of the test.

Solution

The joint p.d.f. of the sample X_1,X_2 is

$$f(x_1,x_2;\theta)=\theta^2(x_1x_2)^{\theta-1},\ 0<x_1,x_2<1.$$

The power function of the test is

$$k(\theta)=\Pr((X_1,X_2)\in C)=\Pr(X_1X_2\ge 3/4)=\int_{3/4}^{1}dx_1\int_{3/4x_1}^{1}f(x_1,x_2;\theta)dx_2.$$

Thus

$$k(1)=\Pr((X_1,X_2)\in C\,|\,\theta=1)=\Pr(X_1X_2\ge 3/4\,|\,\theta=1)=\int_{3/4}^{1}dx_1\int_{3/4x_1}^{1}1\,dx_2=\frac{1}{4}+\frac{3}{4}\ln\frac{3}{4},$$

$$k(2)=\Pr((X_1,X_2)\in C\,|\,\theta=2)=\Pr(X_1X_2\ge 3/4\,|\,\theta=2)=\int_{3/4}^{1}dx_1\int_{3/4x_1}^{1}4x_1x_2dx_2=\frac{7}{16}+\frac{9}{8}\ln\frac{3}{4}.$$

6.39. Let X have a binomial distribution with parameters $n=10$ and $p\in\{p:p=\frac{1}{4},\frac{1}{2}\}$. The simple hypothesis $H_0:p=\frac{1}{2}$ is rejected, and the alternative simple hypothesis $H_1:p=\frac{1}{4}$ is accepted, if the observed value of X_1, a random sample of size 1, is less than or equal to 3. Find the power function of the test.

Solution

The power function of the test is

$$k(\theta)=\Pr(X_1\in C)=\Pr(X_1\le 3)=\sum_{i=0}^{3}\Pr(X_1=i).$$

Thus

$$k(1/2)=\Pr(X_1\in C\,|\,1/2)=\Pr(X_1\le 3\,|\,1/2)=\sum_{i=0}^{3}\Pr(X_1=i)=(\binom{10}{0}+\binom{10}{1}+\binom{10}{2}+\binom{10}{3})\frac{1}{2^{10}}$$

$$=176/2^{10}=0.1718.$$

$$k(1/4)=\Pr(X_1\in C\,|\,1/4)=\Pr(X_1\le 3\,|\,1/4)=\sum_{i=0}^{3}\Pr(X_1=i)=(\binom{10}{0}(\frac{3}{4})^{10}+\binom{10}{1}(\frac{1}{4})(\frac{3}{4})^{9}+$$

$$\binom{10}{2}(\frac{1}{4})^2(\frac{3}{4})^8+\binom{10}{3}(\frac{1}{4})^3(\frac{3}{4})^7=(31)3^8/4^9.$$

6.40. Let X_1,X_2 be a random sample of size $=2$ from the distribution having p.d.f. $f(x;\theta)=(1/\theta)e^{-x/\theta},\ 0<x<\infty$, zero elsewhere. We reject $H_0:\theta=2$ and accept $H_1:\theta=1$ if the observed values of X_1,X_2, say x_1,x_2, are such that

$$\frac{f(x_1;2)f(x_2;2)}{f(x_1;1)f(x_2;1)}\le\frac{1}{2}.$$

Here $\Omega = \{\theta : \theta = 1, 2\}$. Find the significance level of the test and the power function of the test when H_0 is false.

Solution

The inequality $\frac{f(x_1;2)f(x_2;2)}{f(x_1;1)f(x_2;1)} \le \frac{1}{2}$ is equivalent to the following

$$e^{(x_1+x_2/2)} \le 2 \Leftrightarrow \frac{x_1 + x_2}{2} \le \ln 2 .$$

This implies that the critical of the test is

$$C = \{(X_1, X_2) : \frac{X_1 + X_2}{2} \le \ln 2\} .$$

If H_0 is true, it follows from the sampling theorem that $X_1 + X_2$ is Chi-square distribution with 4 degrees of freedom, thus the significance level of the test is

$$\alpha = \Pr\{(X_1, X_2) \in C \mid \theta = 2\} = \Pr\{X_1 + X_2 \le 2\ln 2\} = F(2\ln 2) ,$$

where the function F denotes the distribution function of Chi-square distribution with 4 degrees of freedom.
when H_0 is false, it follows from the sampling theorem that $2X_1 + 2X_2$ is Chi-square distribution with 4 degrees of freedom, thus the power function of the test when H_0 is false is

$$\beta = \Pr\{(X_1, X_2) \in C \mid \theta = 1\} = \Pr\{2X_1 + 2X_2 \le 4\ln 2\} = F(4\ln 2)$$

where the function F denotes the distribution function of Chi-square distribution with 4 degrees of freedom.

6.42. Let us assuming that the life of a tire in miles, say X, is normally distributed with mean θ and standard deviation 5000. Past experience indicates that $\theta = 30000$. The manufacturer claims that the tires made by a new process have mean $\theta > 30000$, and it is very possible that $\theta = 35000$. Let us check his claim by testing $H_0 : \theta = 30000$ against $H_1 : \theta > 30000$. We shall observe n independent values of X, say $x_1, x_2, \cdots, x_n$, and we shall reject H_0 (thus accept H_1) if and only if $\bar{x} \ge c$. Determine n and c so that the power function $k(\theta)$ of the test has the values $k(30000) = 0.001$ and $k(35000) = 0.98$.

Solution

In accordance with the sampling theorem, we have that $\frac{\sqrt{n}(\bar{X} - \theta)}{\sigma}$ is standard normal distribution. If the critical region is of the form

$$C = \{(X_1, X_2, \cdots, X_n) \mid \bar{X} \ge c\}$$

and $k(30000) = 0.001$ and $k(35000) = 0.98$

We have

$$\frac{\sqrt{n}(c - 30000)}{5000} = 3, \quad \frac{\sqrt{n}(c - 35000)}{5000} = -2.05 ,$$

Whose solution is $n = 19$ or $\mathrm{n} = 20$, $c = 10323$ or $\mathrm{c} = 10062$.

6.43. Let X have a Poisson distribution with mean θ. Consider the simple hypothesis $H_0 : \theta = \frac{1}{2}$ and the alternative composite hypothesis $H_0 : \theta < \frac{1}{2}$. Thus $\Omega = \{\theta : 0 < \theta \le 1/2\}$. Let $X_1, X_2, \cdots, X_{12}$ denote a random samples of size 12 from this distribution. We reject H_0 if and only if the observed value of $Y = X_1 + X_2 + \cdots + X_{12} \le 2$. If $k(\theta)$ is the power function of the test, find the powers $k(1/2)$, $k(1/3)$, $k(1/4)$, $k(1/6)$ and $k(1/12)$. What is the significance level of the test?

Solution

It follows from the definition that the power function is of the form

$$k(\theta) = \Pr(X_1 + X_2 + \cdots + X_{12} \le 12) .$$

Since $X_1+X_2+\cdots+X_{12}$ is Poisson distribution with mean 12θ, thus

$$k(1/2)=\Pr(X_1+X_2+\cdots+X_{12}\le 2\,|\,\theta=1/2)=\Pr(X_1+X_2+\cdots+X_{12}=0\,|\,\theta=1/2)+\Pr(X_1+X_2+\cdots+X_{12}=1\,|\,\theta=1/2)+\Pr(X_1+X_2+\cdots+X_{12}=2\,|\,\theta=1/2)=0.062,$$

$$k(1/3)=\Pr(X_1+X_2+\cdots+X_{12}\le 2\,|\,\theta=1/3)=\Pr(X_1+X_2+\cdots+X_{12}=0\,|\,\theta=1/3)+\Pr(X_1+X_2+\cdots+X_{12}=1\,|\,\theta=1/3)+\Pr(X_1+X_2+\cdots+X_{12}=2\,|\,\theta=1/3)=0.238,$$

$$k(1/4)=\Pr(X_1+X_2+\cdots+X_{12}\le 2\,|\,\theta=1/4)=\Pr(X_1+X_2+\cdots+X_{12}=0\,|\,\theta=1/4)+\Pr(X_1+X_2+\cdots+X_{12}=1\,|\,\theta=1/4)+\Pr(X_1+X_2+\cdots+X_{12}=2\,|\,\theta=1/4)=0.423,$$

$$k(1/6)=\Pr(X_1+X_2+\cdots+X_{12}\le 2\,|\,\theta=1/6)=\Pr(X_1+X_2+\cdots+X_{12}=0\,|\,\theta=1/6)+\Pr(X_1+X_2+\cdots+X_{12}=1\,|\,\theta=1/6)+\Pr(X_1+X_2+\cdots+X_{12}=2\,|\,\theta=1/6)=0.677,$$

and

$$k(1/12)=\Pr(X_1+X_2+\cdots+X_{12}\le 2\,|\,\theta=1/12)=\Pr(X_1+X_2+\cdots+X_{12}=0\,|\,\theta=1/12)+\Pr(X_1+X_2+\cdots+X_{12}=1\,|\,\theta=1/12)+\Pr(X_1+X_2+\cdots+X_{12}=2\,|\,\theta=1/12)=0.920.$$

The significance level of the test is $k(1/2)=0.062$.

6.44. Let Y have a binomial distribution with parameters n and p. We reject $H_0: p=\frac{1}{2}$ and accept $H_1: p>\frac{1}{2}$ if $Y\ge c$. Find n and c to give a power function $k(p)$ which is such that $k(1/2)=0.10$ and $k(2/3)=0.95$, approximately.

Solution It follows from the sampling theorem that the sample mean $\frac{\bar{X}-np}{\sqrt{np(1-p)}}$ is approximately with mean 0 and variance 1. Thus approximately we have

$$k(p)=\Pr(Y\ge c)=\Pr\left(\frac{\bar{X}-np}{\sqrt{np(1-p)}}\ge\frac{c/n-np}{\sqrt{np(1-p)}}\right).$$

If $k(1/2)=0.10$ and $k(2/3)=0.95$, then we can obtain

$$\frac{2c/n-n}{\sqrt{n}}\approx 1.282,\ \frac{c/n-2n/3}{\sqrt{n\cdot\frac{2}{3}\cdot\frac{1}{3}}}=-1.645,$$

Whose solutions are

$$n\approx 73,\ c\approx 42.$$

6.45. Let $Y_1<Y_2<Y_3<Y_4$ be the order statistics of a random sample of size 4 from a distribution with p.d.f. $f(x;\theta)=1/\theta, 0<x<\theta$, zero elsewhere, where $0<\theta$. The hypothesis $H_0:\theta=1$ is rejected and $H_1:\theta>1$ accepted if the observed $Y_4\ge c$.

(a) Find the constant c so that the significance level is $\alpha=0.05$.

(b) Determine the power function of the test.

Solution

(a) If $H_0:\theta=1$ is true, the population is uniform distribution on the interval (0,1), so the p.d.f. of the 4th order statistic is

$$f(x;1)=4x,\ 0<x<1.$$

So $\alpha=\mathrm{pr}(Y_4\ge c)=\int_c^1 4x^3dx=1-c^4=0.05$, thus we can get $c=\sqrt[4]{0.95}.$

(b) The power function of the test is

$$k(\theta)=\int_c^1 4x^3/\theta^4\mathrm{d}x=1/\theta^4-c^4/\theta^4.$$

Section 6.5 Additional Comments about Statistical Tests

6.46. Assume that the weight of cereal in a "10-ounce box" is $N(\mu,\sigma^2)$. To test $H_0:\mu=10.1$ against $H_1:\mu>10.1$, we take a random sample of size $n=16$ and observe that $\bar{x}=10.4$ and $s=0.4$.

(a) Do we accept or reject H_0 at the 5percent significance level?

(b) What is the approximate $p-value$ of the test?

Solution

(a) If the null hypothesis H_0 is true, $\bar{X}$ is normal distribution with mean 10.1 and variance $\sigma^2/16$, thus

We have

$$\Pr(\bar{X}\ge c)=\Pr(\frac{\sqrt{n-1}(\bar{X}-10.1)}{S}\ge\frac{\sqrt{n-1}(c-10.1)}{S})=0.05$$

From the Table IV in the Appendix A, we obtain $b=1.753$, let $\frac{\sqrt{n-1}(c-10.1)}{S}$ be 1.753, we get $c=10.28$ and the observed value of the sample mean $\bar{x}=10.4>10.28$, so we reject the null hypothesis H_0 at the 5 percent significance level.

(b) The $p-value$ of the test is

$$\Pr(\bar{X}\ge 10.4)=\Pr(\frac{\sqrt{15}(\bar{X}-10.1)}{S}\ge\frac{\sqrt{15}(10.4-10.1)}{0.4})=\Pr(\frac{\sqrt{15}(\bar{X}-10.1)}{S}\ge\frac{\sqrt{15}(10.4-10.1)}{0.4})$$
$$=\Pr(\frac{\sqrt{15}(\bar{X}-10.1)}{S}\ge 2.905)\approx 1-0.998=0.002.$$

Since 0.002<0.05, so we reject the null hypothesis H_0.

6.47. Each of 51 golfers hit three golf balls of brand X and three golf balls of brand Y in a random order. Let X_i and Y_i equal the averages of the distances traveled by the brand X and brand Y golf balls hit by the *ith* golfer, $i=1,2,\cdots,51$. Let $W_i=X_i-Y_i$, $i=1,2,\cdots,51$. Test $H_0:\mu_W=0$ against $H_1:\mu_W>0$, where μ_W is the mean of the differences. If $\mu_W=2.07$ and $s_W{}^2=84.63$, would H_0 be accepted at an $\alpha=0.05$ significance level? What is the $p-value$ of this test?

Solution

Let $W=X-Y$, in fact, W is $N(\mu_W,\sigma_W{}^2)$. It follows from the sampling theorem that

$$\frac{\sqrt{n-1}(\bar{W}-\mu_W)}{S_W}\sim t(n-1),$$

Given the significance level $\alpha=0.05$, we have $b=1.645$, and the observed value of $\frac{\sqrt{n-1}(\bar{W}-\mu_W)}{S_W}$ is 1.591 if $H_0:\mu_W=0$ is true. Since $1.591<1.645$, so we accept the null hypothesis.

Since the critical region of the test is of the form $\bar{W}\ge c$, then the $p-value$ of this test is

$$\Pr(\bar{W}\ge 2.07)=\Pr(\frac{\sqrt{50}\bar{W}}{S_W}\ge\frac{\sqrt{50}\cdot 2.07}{\sqrt{84.63}})=\Pr(\frac{\sqrt{50}\bar{W}}{S_W}\ge 1.591)>\Pr(\frac{\sqrt{50}\bar{W}}{S_W}\ge 1.6)\approx 0.005.$$

6.48. Among the data collected for the World Health Organization air quality monitoring project is a measure of suspended particles in $\mu g/m^3$. Let X and Y equal the concentration of suspended particles in $\mu g/m^3$ in the city center (commercial district) for Melbourne and Houston, respectively. Using $n=13$ observations of X and $m=16$ Observations of Y, we shall test $H_0:\mu_X=\mu_Y$ against $H_1:\mu_X<\mu_Y$.

(a) Define the test statistic and critical region, assuming that the variances are equal. Let $\alpha=0.05$.

(b) If $\bar{x}=72.9, s_X=25.6, \bar{y}=81.7,$ and $s_Y=28.3$, calculate the value of the test statistic and state your conclusion.

Solution

(a) It follows from the sampling theorem that

$$T=\frac{(\bar{X}-\bar{Y})-(\mu_X-\mu_Y)}{\sqrt{\frac{nS_X{}^2+mS_Y{}^2}{n+m-2}(\frac{1}{n}+\frac{1}{m})}}\sim t(m+n-2).$$

If we shall test $H_0:\mu_X=\mu_Y$ against $H_1:\mu_X<\mu_Y$, the critical region should be of the form

$$C=\{T\le c\}.$$

(b) If $\bar{x}=72.9, s_X=25.6, \bar{y}=81.7,$ and $s_Y=28.3$, the $p-value$ of this test is

$$\Pr(T\le\frac{(\bar{x}-\bar{y})}{\sqrt{\frac{ns_X{}^2+ms_Y{}^2}{n+m-2}(\frac{1}{n}+\frac{1}{m})}})=\Pr(T\le-0.838)=1-\Pr(T\le 0.838)>1-\Pr(T<1.703)=0.05.$$

Since the $p-value$ of this test is no less than 0.05, so we reject the null hypothesis.

6.49. Let p equal the proportion of drivers who use a seat belt in a state that does not have a mandatory seat belt law. It was claimed that $p=0.14$. An advertising campaign was conducted to increase this proportion. Two months after the campaign, $y=104$ out of a random sample of $n=590$ drivers were wearing their seat belts. Was the campaign successful?

(a) Defined the null and the alternative hypotheses.

(b) Define a critical region with an $\alpha=0.01$ significance level.

(c) Determine the approximate $p-value$ and state your conclusion.

Solution

(a) $H_0:p=0.14$ against $H_1:p>0.14$.

(b) The critical of this test should have the form $C=\{\bar{X}\ge c\}$. Approximately we have

$$\frac{\bar{X}-p}{\sqrt{\bar{X}(1-\bar{X})/n}}\dot{\to}N(0,1),$$

In accordance with the Table III in the Appendix A, we approximately have

$$\frac{c-0.14}{\sqrt{\bar{X}(1-\bar{X})/n}}=2.326\Leftrightarrow c=0.1765.$$

(c) The observed value of the statistic $\frac{\bar{X}-p}{\sqrt{\bar{X}(1-\bar{X})/n}}\approx 2.312$, so the approximate $p-value$ of the test is

$$\Pr(\frac{\bar{X}-p}{\sqrt{\bar{X}(1-\bar{X})/n}}\ge 2.312)\approx 1-0.989=0.011.$$

Since 0.011<0.05, so we reject the null hypothesis.

6.50. A machine shop that manufactures toggle levers has both a day and a night shift. A toggle lever is defective if a standard nut cannot be screwed onto the threads. Let p_1 and p_2 be the proportion of defective levers among test the null hypothesis, $H_0; p_1 = p_2$ against a two-sided alternative hypothesis based on two random samples, each of 1000 levers taken from the production of the respective shifts.

(a) Define the test statistic which has an approximate $N(0,1)$ distribution. Sketch a standard normal p.d.f. illustrating the critical region having $\alpha = 0.05$.

(b) If $y_1 = 37$ and $y_2 = 53$ defectives were observed for the day and night shifts, respectively, calculate the value of the statistic and the approximate $p-value$. Locate the calculated test statistic on your figure in part (a) and state your conclusion.

Solution

Additional Exercises

6.62. Let $Y_1 < Y_2 < \cdots < Y_n$ be the order statistics of a random sample of size n from the distribution having p.d.f. $f(x) = 2x/\theta^2,\ 0 < x < \theta,$ zero elsewhere.

(a) If $0 < c < 1$, show that $\Pr(c < Y_n/\theta < 1) = 1 - c^{2n}$.

(b) If $n = 5$ and if the observed value of Y_n is 1.8, find a 99 percent confidence level for θ.

Solution

(a) The distribution function of the population is

$$F(x) = \int_0^x 2t/\theta^2 \mathrm{dt} = x^2/\theta^2.$$

Thus the p.d.f. of the nth order statistic Y_n is

$$f(y) = 2ny^{2n-1}/\theta^{2n},\ 0 < y < \theta.$$

Then if $0 < c < 1$,

$$\Pr(c < Y_n/\theta < 1) = \Pr(c\theta < Y_n < \theta)\int_{c\theta}^{\theta} 2ny^{2n-1}/\theta^{2n} dy = 1 - c^{2n}.$$

(b) In Part (a), let $1 - c^{10} = 0.99,\ c = \sqrt[10]{0.01} = 0.954.$ So a 99 percent confidence level for θ is $(1.8,\ 1.885)$.

6.63. If $0.35, 0.92, 0.56,$ and 0.71 are the four observed values of a random sample from a distribution having p.d.f. $f(x;\theta) = \theta x^{\theta-1},\ 0 < x < 1$, zero elsewhere, find an estimate for θ.

Solution

The likelihood function of the population is

$$L(\theta) = \theta^n (x_1 x_2 \cdots x_n)^{\theta-1},\ 0 < x_i < 1.$$

The logarithm likelihood function is

$$\ln L(\theta) = n\ln\theta + (\theta - 1)\sum_{i=1}^{n} \ln x_i,$$

$$\frac{d\ln L(\theta)}{d\theta} = \frac{n}{\theta} + \sum_{i=1}^{n} \ln x_i = 0.$$

So the m.l.e. of the unknown parameter θ is

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

If $0.35, 0.92, 0.56,$ and 0.71 are the four observed values of a random sample, then the m.l.e. for θ is 1.945.

6.64. Let the table

x	0	1	2	3	4	5
Frequency	6	10	14	13	6	1

represents a summary of a random sample of size 50 from a Poisson distribution. Find the maximum likelihood estimate of $\Pr(X=2)$.

Solution

It is easy to verify that the m.l.e. of θ is the sample mean $\overline{X}$.

Since the population is Poisson distribution with parameter θ, so $\Pr(X=2)=\frac{\theta^2}{2}e^{-\theta}$, it follows from the variance property that the m.l.e. of $\Pr(X=2)$ is $\frac{\overline{X}^2}{2}e^{-\overline{X}}$. The observed value of $\overline{X}$ is 2.12, thus the m.l.e. of $\Pr(X=2)$ is

$$\frac{2.12^2}{2}e^{-2.12}=0.2697.$$

6.65. Let X be $N(\mu,100)$. To test $H_0:\mu=80$ against $H_1:\mu>80$, let the critical region be defined by $C=\{\bar{x}\geq 83\}$, where $\bar{x}$ is the sample mean of a random sample of size $n=25$ from this distribution.

(a) How is the power function $k(\mu)$ defined for this test?

(b) What is the significance level of this test?

(c) What are the values of $k(80)$, $k(83)$ and $k(86)$?

(d) What is the $p-value$ corresponding to $\bar{x}=83.41$?

Solution

(a) Since the sample mean $\overline{X}$ is $N(\mu,\sigma^2/n)=N(\mu,4)$. Thus the power function of the test is

$$k(\mu)=\Pr(\overline{X}\geq 83)=\Pr(\frac{5(\overline{X}-\mu)}{10}\geq\frac{5(83-\mu)}{10}),\text{ for any }\mu\in R.$$

(b) The significance level of the test is

$$\alpha=k(80)=1-\Phi(\frac{5\cdot 3}{10})=1-\Phi(1.5)=1-0.933=0.067.$$

(c) $k(80)=1-\Phi(1.5)=0.067,\quad k(83)=1-\Phi(0)=0.5,\quad$ and $k(86)=1-\Phi(\text{-}1.5)=0.933$.

(d) The $p-value$ corresponding to $\bar{x}=83.41$ is

$$p-value=1-\Phi(1.705)\approx 1-0.955=0.005.$$

6.66. Let X equal the yield of alfalfa in tons per acre per year. Assume that X is $N(1.5,0.09)$. It is hoped that new fertilizer will increase the average yield. We shall test the null hypothesis $H_0:\mu=1.5$ against the alternative hypothesis $H_1:\mu>1.5$. Assume that the variance continues to equal $\sigma^2=0.09$ with the new fertilizer. Using $\overline{X}$, the mean of a random sample of size n, as the test statistic, reject H_0 if $\bar{x}\geq c$. Find n and c so that the power function $k(\mu)=\Pr(\overline{X}\geq c:\mu)$ is such that $\alpha=k(1.5)=0.05$ and $k(1.7)=0.95$.

Solution

It follows from the sampling theorem that the distribution of the sample mean is

$$\overline{X}\sim N(\mu,0.09/n).$$

The power function of the test is

$$k(\mu)=\Pr(\overline{X}\geq c:\mu)=\Pr(\frac{\sqrt{n}(\overline{X}-\mu)}{0.3}\geq\frac{\sqrt{n}(c-\mu)}{0.3}).$$

If $\alpha=k(1.5)=0.05$ and $k(1.7)=0.95$, then we have

$$\frac{\sqrt{n}(c-1.5)}{0.3}=1.645,\qquad\frac{\sqrt{n}(c-1.7)}{0.3}=-1.645$$

whose solutions with respective to n and c are

$$n = 24,\ c = 1.6.$$

6.67. A random sample of 100 observations from a Poisson distribution has a mean equal to 6.25. Construct an approximate 95 percent confidence interval for the distribution.

Solution

It follows from Slutsky's theorem that

$$\frac{\sqrt{n}(\overline{X}-\theta)}{\sqrt{\overline{X}}} \approx N(0,1).$$

And we have an approximate 95 percent confidence interval for the distribution

$$-1.96 < \frac{\sqrt{n}(\overline{X}-\theta)}{\sqrt{\overline{X}}} < 1.96,$$

Equivalently, $(5.76, 6.74)$ when $n = 100,\ \bar{x} = 6.25$.

6.68. Say that a random sample of size 25 is taken from a binomial distribution with parameters $n = 5$ and p. These data are then lost, but we recall that the relative frequency of the value 5 was $\frac{6}{25}$. Under these conditions, how would you estimate p? Is this suggested estimate unbiased?

Solution

6.69. When 100 tacks were thrown on a table, 60 of them landed point up. Obtain a 95 percent confidence interval for the probability that a tack of this type will land point up. Assume independent.

Solution

In accordance with the central limit theorem, we approximately have

$$\frac{Y-np}{\sqrt{n(Y/n)(1-Y/n)}} \sim N(0,1).$$

Approximately we have

$$-1.96 < \frac{Y-np}{\sqrt{n(Y/n)(1-Y/n)}} < 1.96,$$

Here the observed value of Y is 60 and $n = 100$, so we obtain a 95 percent confidence interval for the probability

$$(0.5, 0.696).$$

6.70. Let $X_1, X_2, \cdots, X_8$ be a random sample of size $n = 8$ from a Poisson distribution with mean μ. Reject the simple null hypothesis $H_0: \mu = 0.5$ and accept $H_1: \mu > 0.5$ if the observed sum $\sum_{i=1}^{8} x_i \ge 8$.

(a) Compute the significance level α of the test.

(b) Find the power function $k(\mu)$ of the test as a sum of Poisson probabilities.

(c) Using the Appendix, determine $k(0.75), k(1),$ and $k(1.25)$.

Solution

We know that the sum of the sample $\sum_{i=1}^{8} X_i$ is Poisson distribution with mean 8μ.

(a) If the simple null hypothesis is true, $\sum_{i=1}^{8} X_i \sim \mathrm{P}(8\mu_0) = \mathrm{P}(4)$, thus the significance level α of the test is

$$\alpha = \Pr(\sum_{i=1}^{8} X_i \ge 8 : \mu = 0.5) = 1 - \Pr(\sum_{i=1}^{8} X_i \le 7 : \mu = 4) = 1 - 0.949 = 0.051 .$$

(b) The power function of the test is

$$k(\mu) = \Pr(\sum_{i=1}^{8} X_i \ge 8 : \mu) = 1 - \Pr(\sum_{i=1}^{8} X_i \le 7 : \mu) = 1 - \sum_{i=0}^{7} \frac{\mu^i}{i!} e^{-\mu} .$$

(c) Using the Appendix, we obtain

$$k(0.75) = 1 - 0.744 = 0.256, \quad k(1) = 1 - 0.453 = 0.547, \quad k(1.25) = 1 - 0.22 = 0.78 .$$

6.71. Let p denote the probability that, for a particular tennis player, the first serve is good. Since $p = 0.40$, this player decided to take lessons in order to increase p**.** When the lessons are completed, the hypothesis $H_0 : p = 0.40$ will be tested against $H_1 : p > 0.40$ based on $n = 25$ trials. Let y equal the number of first serves that are good, and let the critical region be defined by $C = \{y : y \ge 13\}$.

(a) Determine $\alpha = \Pr(Y \ge 13 : p = 0.40)$.

(b) Find $\beta = \Pr(Y < 13)$ when $p = 0.60;$ that is, $\beta = \Pr(Y \le 12; p = 0.60)$.

Solution

By the sampling theorem, we know that

$$Y \sim b(n, p) .$$

(a) If $n = 25$, the significance level of the test is

$\alpha = \Pr(Y \ge 13 : p = 0.40)$

6.72. The mean birth weight in the United States is $\mu = 3315$ grams with a standard deviation of $\sigma = 575$. Let X equal the birth weight in grams in Jerusalem. Assume that the distribution of X is $N(\mu, \sigma^2)$. We shall test the null hypothesis $H_0 ; \mu = 3315$ against the alternative hypothesis $H_1 ; \mu < 3315$ using a random sample of size $n = 30$.

(a) Define a critical region that has a significance level of $\alpha = 0.05$.

(b) If the random sample of $n = 30$ yields $\bar{x} = 3189$, what is your conclusion?

(c) What is the approximate $p - value$ of your test?

Solution

It follows from the sampling theorem that the sample mean $\bar{X}$ is normally distributed with mean μ and variance σ^2 / n, that is, $\bar{X} \sim N(\mu, \sigma^2 / n)$.

(a) If we shall test the null hypothesis $H_0 ; \mu = 3315$ against the alternative hypothesis $H_1 ; \mu < 3315$, the critical region should be selected as the form

$$C = \{\bar{X} \le c\} ,$$

Moreover, if the significance level is given $\alpha = 0.05$, that is

$$\alpha = \Pr(\bar{X} \le c) = \Pr(\frac{\sqrt{30}(\bar{X} - 3315)}{575} \le \frac{\sqrt{30}(c - 3315)}{575}) = 0.05 ,$$

By the Table III in the Appendix B, we have

$$\frac{\sqrt{30}(c - 3315)}{575} = -1.645, \quad c = 3142 ,$$

Thus the critical of the test is $C = \{\bar{X} \le 3142\}$.

(b) If the random sample of $n=30$ yields $\bar{x}=3189$, since 3189<142, so we should reject the null hypothesis.
(c) If the random sample of $n=30$ yields $\bar{x}=3189$, the approximate $p-value$ of the test is

$$p-value=\Pr(\bar{X}\le 3189)=\Pr(\frac{\sqrt{30}(\bar{X}-3315)}{575}\le\frac{\sqrt{30}(3189-3315)}{575})=\Phi(-1.2)=1-\Phi(1.2)=1-0.885=0.115.$$

6.75. Let $X_1,X_2,\cdots,X_n$ be a random sample from $N(\mu,\sigma^2)$.

(a) If the constant b is defined by the equation $\Pr(X\le b)=0.90$, find the m.l.e. of b.
(b) If c is given constant, find the m.l.e. of $\Pr(X\le c)$.

Solution

If the parameters μ and σ^2 are unknown, then the m.l.e. of the parameters are $\bar{X}$ and S^2, respectively.

(a) Since $\Pr(X\le b)=\Pr(\frac{X-\mu}{\sigma}\le\frac{b-\mu}{\sigma})=\Phi(\frac{b-\mu}{\sigma})=0.90$, so we have

$$\frac{b-\mu}{\sigma}=1.282,$$

Thus we have $b=\mu+1.282\sigma$, from the variance property of the m.l.e., we obtain the m.l.e. of b is

$\hat{b}=\bar{X}=1.282S$.

(b) Since $\Pr(X\le c)=\Pr(\frac{X-\mu}{\sigma}\le\frac{c-\mu}{\sigma})=\Phi(\frac{c-\mu}{\sigma})$, so for the given c, the m.l.e. of c is $\hat{c}=\Phi(\frac{c-\bar{X}}{S})$.

6.76. Let $\bar{X}_1,\bar{X}_2$, and $\bar{X}_3$ and $S^2{}_1,S^2{}_2$ and $S^2{}_3$ denote the means and the variances of three independent random samples, each of size 10, from a normal distribution with mean μ and variance σ^2. Find the constant c so that

$$\Pr(\frac{\bar{X}_1+\bar{X}_2-2\bar{X}_3}{\sqrt{10S^2{}_1+10S^2{}_2+10S^2{}_3}}\le c)=0.95.$$

Solution

It follows from the sampling theorem that

$$\bar{X}_1+\bar{X}_2-2\bar{X}_3\sim N(0,\frac{3}{5}\sigma^2),\ \text{and}\ \frac{10(S_1{}^2+S_2{}^2+S_3{}^2)}{\sigma^2}\sim\chi^2(27),$$

and the statistics are independent with each other. So

$$\frac{\bar{X}_1+\bar{X}_2-2\bar{X}_3}{\sqrt{10S^2{}_1+10S^2{}_2+10S^2{}_3}}\cdot 3\sqrt{5}\sim t(27).$$

Thus

$$\Pr(\frac{\bar{X}_1+\bar{X}_2-2\bar{X}_3}{\sqrt{10S^2{}_1+10S^2{}_2+10S^2{}_3}}\le c)=\Pr(\frac{\bar{X}_1+\bar{X}_2-2\bar{X}_3}{\sqrt{10S^2{}_1+10S^2{}_2+10S^2{}_3}}\cdot 3\sqrt{5}\le 3\sqrt{5}\,c)=0.95,$$

We have

$$3\sqrt{5}\,c=1.703,\ c=0.2538.$$

6.77. Let Y be $b(192, p)$. We reject $H_0: p = 0.75$ and accept $H_1: p > 0.75$ if and only if $Y \geq 152$. Use the normal approximate to determine:

(a) $\alpha = \Pr(Y \geq 152; p = 0.75)$.

(b) $\beta = \Pr(Y < 152)$ when $p = 0.80$.

Solution

It follows the sampling theorem that

$$\frac{Y - np}{\sqrt{np(1-p)}} \to N(0,1) \quad (n \to \infty).$$

Thus

(a)

$$\alpha = \Pr(Y \geq 152; p = 0.75) = \Pr(\frac{Y - np}{\sqrt{np(1-p)}} \geq \frac{152 - np}{\sqrt{np(1-p)}}; p = 0.75)$$
$$= 1 - \Phi(1.333) = 1 - 0.911 = 0.089.$$

(c) When $p = 0.80$,

$$\beta = \Pr(Y < 152; p = 0.80) = \Pr(\frac{Y - np}{\sqrt{np(1-p)}} < \frac{152 - np}{\sqrt{np(1-p)}}; p = 0.80)$$
$$= \Phi(-0.2886) = 1 - \Phi(0.2886) = 1 - 0.618 = 0.382.$$

Chapter 7 Sufficient Statistics

Section 7.1 Measures of Quality Estimation

Exercises

7.1. Show that the mean $\overline{X}$ of a random sample of size n from a distribution having p.d.f. $f(x;\theta)=(1/\theta)e^{-(x/\theta)}$, $0<x,\infty,0<\theta<\infty,$ zero elsewhere, is an unbiased estimator of θ and has variance θ^2/n.

Solution

The expected value of the population is

$$E(X)=\int_0^\infty \frac{x}{\theta}e^{-x/\theta}dx=\theta .$$

Moreover,

$$E\overline{X}=EX=\theta ,$$

Thus the mean $\overline{X}$ is an unbiased estimator of θ.

It is easy to verify $DX=\theta^2$, so the variance of the mean $D\overline{X}=DX/n=\theta^2/n$. This completes the proof.

7.2. Let $X_1,X_2,\cdots,X_n$ denote a random sample from a normal distribution with mean zero and variance $\theta, 0<\theta<\infty$. Show that $\sum_{i=1}^{n}X_i{}^2/n$ is an unbiased estimator of θ and has variance $2\theta^2/n$.

Solution

In fact, the population is $N(0,\theta)$. Since the sample $X_1,X_2,\cdots,X_n$ are independent and identically distributed. So

$$X_i\sim N(0,\theta),\ i=1,2,\cdots,n, \text{ and } \frac{X_i}{\sqrt{\theta}}\sim N(0,1), \frac{X_i{}^2}{\theta}\sim\chi^2(1), i=1,2,\cdots,n .$$

Thus

$$\sum_{i=1}^{n}X_i{}^2/\theta\sim\chi^2(n) .$$

According to the property of the Chi-square distribution, we have

$$E(\sum_{i=1}^{n}X_i{}^2/\theta)=n, D(\sum_{i=1}^{n}X_i{}^2/\theta)=2n,$$

These imply that

$$E(\sum_{i=1}^{n}X_i{}^2/n)=\theta, D(\sum_{i=1}^{n}X_i{}^2/n)=2\theta^2/n .$$

7.3. Let $Y_1<Y_2<Y_3$ be the order statistics of a random sample of size 3 from the uniform distribution having p.d.f. $f(x;\theta)=1/\theta,0<x<\theta,0<\theta<\infty$, zero elsewhere. Show that $4Y_1, 2Y_2$, and $\frac{4}{3}Y_3$ are all unbiased estimators of θ. Find the variance of each of these unbiased estimators.

Solution

The distribution function of the population is

$$F(x)=\begin{cases}0, & x\le 0\\ x/\theta, & 0<x<\theta\\ 1, & x\ge\theta\end{cases},$$

So the p.d.f. of the ith order statistic is

$$f(y)=\frac{3!}{i!(3-i)!}(\frac{y}{\theta})^{i-1}(1-\frac{y}{\theta})^{3-i}\frac{1}{\theta}=\frac{3!}{i!(3-i)!}\frac{1}{\theta^3}y^{i-1}(\theta-y)^{3-i},\ 0<y<\theta,\ i=1,2,3.$$

Then

$$E(4Y_1)=\int_0^\theta 4y\cdot 3/\theta^3 y(\theta-y)^2\,dy=12\theta B(2,3)=\theta,$$

By the similar computation process, we can obtain

$$E(2Y_2)=\theta,\ E(\frac{4}{3}Y_3)=\theta.$$

Thus $4Y_1, 2Y_2$, and $\frac{4}{3}Y_3$ are all unbiased estimators of θ.

It is not difficult to verify that

$$D(Y_1)=\frac{3}{4\cdot 5}\theta^2=\frac{3}{20}\theta^2,\ D(Y_2)=\frac{2\cdot 2}{4\cdot 5}\theta^2=\frac{1}{5}\theta^2,\ D(Y_3)=\frac{3}{4\cdot 5}\theta^2=\frac{3}{20}\theta^2,$$

So the variance of the statistics is

$$D(4Y_1)=\frac{3\cdot 16}{4\cdot 5}\theta^2=\frac{12}{5}\theta^2,\ D(2Y_2)=\frac{2\cdot 2\cdot 4}{4\cdot 5}\theta^2=\frac{4}{5}\theta^2,\ D(\frac{4}{3}Y_3)=\frac{3}{4\cdot 5}\theta^2\cdot\frac{16}{9}=\frac{4}{15}\theta^2.$$

7.4. Let Y_1 and Y_2 be two independent unbiased estimators of θ. Say the variance of Y_1 is twice the variance of Y_2. Find the constants k_1 and k_2 so that $k_1Y_1+k_2Y_2$ is an unbiased estimator with smallest possible variance for such a linear combination.

Solution

Since $D(Y_1)=2D(Y_2)$ and , so

$$D(k_1Y_1+k_2Y_2)=k_1{}^2D(Y_1)+2k_2{}^2D(Y_1)=(k_1{}^2+2k_2{}^2)D(Y_1)=[k_1{}^2+2(1-k_1)^2]D(Y_1)$$
$$=(3k_1{}^2-4k_1+2)D(Y_1).$$

The function $3k_1{}^2-4k_1+2$ can be minimized at the point $k_1=2/3$, thus $k_2=1/3$.

7.5. In Example 1of this section, take $L[\theta,\delta(Y)]=|\theta-\delta(Y)|$. Show that $R(\theta,\delta_1)=\frac{1}{5}\sqrt{2/\pi}$ and $R(\theta,\delta_2)=|\theta|$ of these two decision functions δ_1 and δ_2, which yields the smallest maximum risk?

Solution

Since the sample mean $\overline{X}\sim N(\theta,1/25)$, so $\overline{X}-\theta\sim N(0,1/25)$.

$$R(\theta,\delta_1)=E(|\theta-\overline{X}|)=2\int_0^\infty x\frac{5}{\sqrt{2\pi}}e^{-\frac{25x^2}{2}}dx=\frac{1}{5}\sqrt{2/\pi},\ R(\theta,\delta_2)=E(|\theta|)=|\theta|,$$

This completes the proof.

The maximum risk of the function $R(\theta,\delta_1)=\frac{1}{5}\sqrt{2/\pi}$ is $\frac{1}{5}\sqrt{2/\pi}$, whereas the maximum risk of the function

$R(\theta,\delta_2)=|\theta|$ is $+\infty$, so in these two decision functions δ_1 and δ_2, δ_1 yields the smallest maximum risk.

7.6. Let $X_1, X_2, \cdots, X_n$ denote a random sample from a Poisson distribution with parameter θ, $0<\theta<\infty$. Let $Y=\sum_{i=1}^{n} X_i$ and let $L[\theta,\delta(Y)]=(\theta-\delta(Y))^2$. If we restrict our considerations to decision functions of the form $\delta(y)=b+y/n$, where b does not depend upon y, show that $R(\theta,\delta)=b^2+\theta/n$. What decision function of this form yields a uniformly smaller risk than every other decision function of this from? With this solution, say δ, and $0<\theta<\infty$, determine $\max_{\theta} R(\theta,\delta)$ if it exists.

Solution

Since $Y=\sum_{i=1}^{n} X_i$ is Poisson distribution with parameter $n\theta$, $0<\theta<\infty$, so if we restrict our considerations to decision functions of the form $\delta(y)=b+y/n$, where b does not depend upon y, we have

$$R(\theta,\delta)=E(\theta-b-Y/n)^2=b^2+E(\theta-Y/n)^2=b^2+D(Y/n)=b^2+\theta/n.$$

It is easy to understand that when $b=0$, that is $\delta=Y/n$, the decision function has uniformly smaller risk. Under conditions, $\max_{\theta} R(\theta,\delta)$ does not exist.

7.7. Let $X_1, X_2, \cdots, X_n$ denote a random sample from a distribution that is $N(\mu,\theta)$, $0<\theta<\infty$, where μ is unknown. Let $Y=\sum_{i=1}^{n}(X_i-\overline{X})/n=S^2$ and let $L[\theta,\delta(Y)]=(\theta-\delta(Y))^2$. If we consider decision functions of the from $\delta(y)=by$, where b does not depend upon y, show that $R(\theta,\delta)=(\theta^2/n^2)[(n^2-1)b^2-2n(n-1)b+n^2]$. Show that $b=n/(n+1)$ yields a minimum risk for decision functions of this form. Note that $nY/(n+1)$ is not an unbiased estimator of θ. With $\delta(y)=ny/(n+1)$ and $0<\theta<\infty$, determine $\max_{\theta} R(\theta,\delta)$ if it exists.

Solution

It follows from the sampling distribution theorem that $\frac{nS^2}{\theta}=\frac{nY}{\theta}\sim\chi^2(n-1)$, we have

$$E(bY)=\frac{(n-1)b\theta}{n},\ \ \mathrm{D}(bY)=\frac{2(n-1)b^2\theta^2}{n^2}.$$

Thus

$$R(\theta,\delta)=E[\theta-bY]^2=E[\theta-\frac{(n-1)b\theta}{n}+\frac{(n-1)b\theta}{n}-bY]^2=(\theta^2/n^2)[(n^2-1)b^2-2n(n-1)b+n^2].$$

Since the derivative of the function $(n^2-1)b^2-2n(n-1)b=2(n^2-1)-2n(n-1)$ with respective to b. Let

$2(n^2-1)-2n(n-1)=0$, we obtain $b=n/(n+1)$ which yields a minimum risk for decision functions of this form.

Whereas $R(\theta,\delta)=\theta^2$, $\max_{\theta} R(\theta,\delta)$ does not exist.

7.8. Let $X_1, X_2, \cdots, X_n$ denote a random sample from a distribution that is $b(1,\theta),\ 0 \le \theta \le 1$. Let $Y = \sum_{i=1}^{n} X_i$ and let $L[\theta, \delta(Y)] = (\theta - \delta(Y))^2$. Consider decision functions of the form $\delta(y) = by$, where b does not depend upon y. Prove that $R(\theta,\delta) = b^2 n\theta(1-\theta) + (bn-1)^2\theta^2$. Show that

$$\max_{\theta} R(\theta,\delta) = \frac{b^4 n^2}{4[b^2 n - (bn-1)^2]},$$

Provided that the value b is such that $b^2 n \ge 2(bn-1)^2$. Prove that $b = 1/n$ does not minimize $\max_{\theta} R(\theta,\delta)$.

Solution

In fact, $Y = \sum_{i=1}^{n} X_i$ is binomial distribution with parameters n and θ, thus $E(bY) = nb\theta,\ D(bY) = b^2 n\theta(1-\theta)$.

So

$$R(\theta,\delta) = E(\theta - bY)^2 = E(\theta - nb\theta + nb\theta - bY)^2 = (bn-1)^2\theta^2 + D(bY) = b^2 n\theta(1-\theta) + (bn-1)^2\theta^2.$$

Since the derivative of the function $b^2 n\theta(1-\theta) + (bn-1)^2\theta^2 = [2(bn-1)^2 - 2b^2 n]\theta + b^2 n$ with respect to θ, let

$[2(bn-1)^2 - 2b^2 n]\theta + b^2 n = 0$ whose solution is $\theta = -\dfrac{b^2 n}{2(bn-1)^2 - 2b^2 n}$.

Thus we obtain

$$\max_{\theta} R(\theta,\delta) = \frac{b^4 n^2}{4[b^2 n - (bn-1)^2]}.$$

It is easy to verify $b = 1/n$ does not minimize $\max_{\theta} R(\theta,\delta)$.

Section 7.2. A Sufficient Statistic for a Parameter

Exercises

7.10. Let $X_1, X_2, \cdots, X_n$ denote a random sample from the normal distribution $N(0,\theta), 0<\theta<\infty$. Show that $\sum_{i=1}^{n} X_i^{\,2}$ is a sufficient statistic for θ.

Solution

The joint p.d.f. of the sample $X_1, X_2, \cdots, X_n$ is

$$(\frac{1}{\sqrt{2\pi\theta}})^n e^{-\frac{\sum x_i^2}{2\theta}},$$

It follows from the factorization theorem, $\sum_{i=1}^{n} X_i^{\,2}$ is a sufficient statistic for θ.

7.11. Prove that the sum of the observation of a random sample of size n from a Poisson distribution having parameter $\theta, 0<\theta, \infty$, is a sufficient statistic for θ.

Solution

The joint p.d.f. of the sample $X_1, X_2, \cdots, X_n$ is

$$\frac{\theta^{x_1}\theta^{x_2}\cdots\theta^{x_n}}{x_1!x_2!\cdots x_n!}e^{-n\theta} = \theta^{\sum x_i} e^{-n\theta}\cdot\frac{1}{x_1!x_2!\cdots x_n!},$$

It follows from the factorization theorem, $\sum_{i=1}^{n} X_i$ is a sufficient statistic for θ.

7.12. Show that the nth order statistic of a random sample of size n from the uniform distribution having p.d.f. $f(x;\theta)=1/\theta,\ 0<x<\theta,\ 0<\theta<\infty$, zero elsewhere, is a sufficient statistic for θ. Generalize this result by considering the p.d.f. $f(x;\theta)=Q(\theta)M(x),\ 0<x<\theta, 0<\theta<\infty$, zero elsewhere. Here, of course,

$$\int_0^\theta M(x)\,dx = \frac{1}{Q(\theta)}.$$

Solution

The joint p.d.f. of the sample $X_1, X_2, \cdots, X_n$ is

$$\prod_{i=1}^{n} f(x_i;\theta) = \begin{cases} \frac{1}{\theta^n},\ 0\le x_{(1)}\le x_{(n)}\le\theta \\ 0, \quad \text{otherwise}\end{cases} = \frac{1}{\theta^n} I_{\{x_{(n)}\le\theta\}} I_{\{x_{(1)}\ge 0\}} = k_1(T,\theta)k_2(x_1,x_2,\cdots,x_n),$$

where $k_1(T,\theta)=\frac{1}{\theta^n} I_{\{x_{(n)}\le\theta\}}, k_2(x_1,x_2,\cdots,x_n)=I_{\{x_{(10}\ge 0\}}$.

It follows from the factorization theorem, Y_n is a sufficient statistic for θ.

By similar reasoning, Y_n is a sufficient statistic for θ in the p.d.f. $f(x;\theta)=Q(\theta)M(x),\ 0<x<\theta, 0<\theta<\infty$.

7.13. Let $X_1, X_2, \cdots, X_n$ be a random sample of size n from a geometric distribution that has p.d.f. $f(x;\theta)=(1-\theta)^x\theta, x=0,1,2,\cdots,\ 0<\theta<1$, zero elsewhere. Show that $\sum_{i=1}^{n} X_i$ is a sufficient statistic for θ

Solution

The joint p.d.f. of the sample $X_1, X_2, \cdots, X_n$ is

$$\prod_{i=1}^{n} f(x_i;\theta) = (1-\theta)^{\sum x_i}\theta^n ,$$

It follows from the factorization theorem, $\sum_{i=1}^{n} X_i$ is a sufficient statistic for θ.

7.14. Show that the sum of the observation of a random sample of size n from a gamma distribution that has p.d.f. $f(x;\theta) = (1/\theta)e^{-x/\theta}$, $0 < x < \infty, 0 < \theta < \infty$,zero elsewhere, is a sufficient statistic for θ.

Solution

The joint p.d.f. of the sample $X_1, X_2, \cdots, X_n$ is

$$\prod_{i=1}^{n} f(x_i;\theta) = \frac{1}{\theta^n} e^{-\sum x_i/\theta} , ,$$

It follows from the factorization theorem, $\sum_{i=1}^{n} X_i$ is a sufficient statistic for θ.

7.15. Let $X_1, X_2, \cdots, X_n$ be a random sample of size n from a beta distribution with parameters $\alpha = \theta > 0$ and $\beta = 2$. Show that the product $X_1 X_2 \cdots X_n$ is a sufficient statistic for θ.

Solution

The joint p.d.f. of the sample $X_1, X_2, \cdots, X_n$ is

$$\prod_{i=1}^{n} f(x_i;\theta) = (\frac{\Gamma(2+\alpha)}{\Gamma(\alpha)})^n (x_1 x_2 \cdots x_n)^{\theta-1}(1-x_1)(1-x_2)\cdots(1-x_n), ,$$

It follows from the factorization theorem, $X_1 X_2 \cdots X_n$ is a sufficient statistic for θ.

7.16. Show that the product of the sample observation is a sufficient statistic for $\theta > 0$ if the random sample is taken from a gamma distribution with parameters $\alpha = \theta$ and $\beta = 6$.

Solution

The joint p.d.f. of the sample $X_1, X_2, \cdots, X_n$ is

$$\prod_{i=1}^{n} f(x_i;\theta) = (\frac{1}{\Gamma(\theta)6^\theta})^n (x_1 x_2 \cdots x_n)^{\theta-1} e^{-\sum x_i/6},$$

It follows from the factorization theorem, $X_1 X_2 \cdots X_n$ is a sufficient statistic for θ.

7.17. What is the sufficient statistic for θ if the sample arises from a beta distribution in which $\alpha = \beta = \theta > 0$?

The joint p.d.f. of the sample $X_1, X_2, \cdots, X_n$ is

$$\prod_{i=1}^{n} f(x_i;\theta) = (\frac{\Gamma(2\theta)}{\Gamma(\theta)\Gamma(\theta)})^n (x_1 x_2 \cdots x_n)^{\theta-1}((1-x_1)(1-x_2)\cdots(1-x_n))^{\theta-1},$$

It follows from the factorization theorem, $X_1 X_2 \cdots X_n(1-X_1)(1-X_2)\cdots(1-X_n)$ is a sufficient statistic for θ.

Section 7.3. Properties of a Sufficient Statistics

Solutions to Exercises

7.18. In each of the Exercises 7.10., 7.11, 7.13, and 7.14, show that the m.l.e. of θ is a sufficient statistic for θ.

Solution

In the exercises 7.10, the m.l.e. of the unknown θ is $\frac{1}{n}\sum_{i=1}^{n} X_i^2$ which is also a sufficient for θ.

In the exercises 7.11, the m.l.e. of the unknown θ is $\frac{1}{n}\sum_{i=1}^{n} X_i$ which is also a sufficient for θ.

In the exercises 7.13, the m.l.e. of the unknown θ is $\dfrac{n}{n+\sum_{i=1}^{n} X_i}$ which is also a sufficient for θ.

In the exercises 7.14, the m.l.e. of the unknown θ is $\frac{1}{n}\sum_{i=1}^{n} X_i$ which is also a sufficient for θ.

7.19. Let $Y_1 < Y_2 < Y_3 < Y_4 < Y_5$ be the order statistics of a random sample of size 5 from the uniform distribution having $f(x;\theta) = 1/\theta, 0 < x < \theta, 0 < \theta < \infty$, zero elsewhere. Show that $2Y_3$ is an unbiased estimator of θ. Determine the joint p.d.f. of Y_3 and the sufficient statistic Y_5 for θ. Find the conditional expectation $E(2Y_3 \mid y_5) = \varphi(y_5)$. Compare the variance of $2Y_3$ and $\varphi(Y_5)$.

Solution

The distribution function of the population is

$$F(x) = \begin{cases} 0, & x < 0 \\ \dfrac{x}{\theta}, & 0 \le x < \theta \\ 1, & x \ge \theta \end{cases}$$

So the p.d.f. of the order statistic Y_3 is

$$f(y) = 30y^2(\theta - y)^2/\theta^5,\ 0 < y < \theta,$$

Thus the expected value of $2Y_3$ is

$$E(2Y_3) = 2\int_0^\theta y \cdot 30y^2(\theta-y)^2/\theta^5 dy = 60\theta\int_0^1 t^3(1-t)^2\,dt = 60\theta B(4,3) = \theta.$$

This implies that $2Y_3$ is an unbiased estimator of θ.

The joint p.d.f. of Y_3 and the sufficient statistic Y_5 for θ is

$$f(x,y) = 60x^2(y-x)/\theta^5,\ 0 < x \le y < \theta,$$

The conditional p.d.f. of Y_3 given $Y_5 = y$ is

$$f(x \mid y) = \frac{12x^2(y-x)}{y^4},\ \ 0 < x < y < \theta,$$

So the conditional expectation $E(2Y_3 \mid y_5) = \frac{6}{5} y_5 = \varphi(y_5)$.

The variance of $2Y_3$ is

$$D(2Y_3) = 4 \cdot \frac{3 \cdot 3}{6 \cdot 6 \cdot 7} \theta^2 = \frac{1}{7} \theta^2.$$

Whereas the variance of $\varphi(Y_5)$ is

$$D(\frac{6}{5} Y_5) = \frac{36}{25} \cdot \frac{5}{6 \cdot 6 \cdot 7} \theta^2 = \frac{1}{35} \theta^2.$$

It is obvious that $D(\varphi(Y_5)) < D(2Y_3)$.

7.20. If X_1, X_2 is a random sample of size 2 from a distribution having p.d.f.

$$f(x;\theta) = (1/\theta)e^{-x/\theta},\ 0 < x < \infty,\ 0 < \theta < \infty, \text{ zero elsewhere,}$$

find the joint p.d.f. of the sufficient statistic $Y_1 = X_1 + X_2$ for θ and $Y_2 = X_2$. Show that Y_2 is an unbiased estimator of θ with variance θ^2. Find $E(Y_2 \mid y_1) = \varphi(y_1)$ and the variance of $\varphi(Y_1)$.

Solution

The joint p.d.f. of the sample X_1, X_2 is

$$f(x_1, x_2) = \frac{1}{\theta^2} e^{-(x_1+x_2)/\theta},\ \ 0 < x_1, x_2 < \infty.$$

Let $Y_1 = X_1 + X_2$, $Y_2 = X_2$, then the Jacobian of the transformation is

$$J = \begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} = 1,$$

So the joint of the the sufficient statistic $Y_1 = X_1 + X_2$ for θ and $Y_2 = X_2$ is

$$f(y_1, y_2) = \frac{1}{\theta^2} e^{-y_1/\theta},\ \ 0 < y_2 < y_1 < \infty.$$

$E(Y_2) = E(X_2) = \theta,\ D(Y_2) = D(X_2) = \theta^2.$

The marginal p.d.f. of Y_1 is

$$f(y_1) = \int_0^{y_1} \frac{1}{\theta^2} e^{-y_1/\theta} dy_2 = \frac{1}{\theta^2} y_1 e^{-y_1/\theta},\ \ 0 < y_1 < \infty.$$

So the conditional p.d.f. of Y_2 given $Y_1 = y_1$ is

$$f(y_2 \mid y_1) = \frac{1}{y_1},\ 0 < y_2 < y_1 < \infty.$$

Thus

$$E(Y_2 \mid y_1) = \int_0^{y_1} \frac{y_2}{y_1} dy_2 = y_1 / 2 = \varphi(y_1).$$

The variance of $\varphi(Y_1)$ is $D(\varphi(Y_1)) = \theta^2 / 2$.

7.21. Let the random variables X and Y have the joint p.d.f.

$f(x, y) = (2/\theta^2)e^{-(x+y)/\theta},\ 0 < x < y < \infty$, zero elsewhere.

(a) Show that the mean and the variance of Y are, respectively, $3\theta/2$ and $5\theta^2/4$.

(b) Show that $E(Y \mid x) = x + \theta$. In accordance with the theory, the expected value of $X + \theta$ is that of Y, namely, $3\theta/2$, and the variance of $X + \theta$ is less than that of Y. Show that the variance of $X + \theta$ is in fact $\theta^2/4$.

Solution

(a)

$$E(Y)=\iint_{R^2} yf(x,y)dxdy=\int_0^{\infty}dx\int_x^{+\infty}(2/\theta^2)ye^{-(x+y)/\theta}dy=3\theta/2,$$

$$E(Y^2)=\iint_{R^2} y^2 f(x,y)dxdy=\int_0^{\infty}dx\int_x^{+\infty}(2/\theta^2)y^2e^{-(x+y)/\theta}dy=13\theta^2/4,$$

So the variance of Y is

$$D(Y)=E(Y^2)-(E(Y))^2=5\theta^2/4.$$

(b) The marginal p.d.f. of X is

$$f(x)=\int_{-\infty}^{\infty}f(x,y)dy=\int_x^{+\infty}(2/\theta^2)e^{-(x+y)/\theta}dy=\frac{2}{\theta}e^{-2x/\theta},\ \ 0<x<\infty.$$

The conditional p.d.f.of Y given $X=x$ is

$$f(y\,|\,x)=\frac{1}{\theta}e^{-(y-x)/\theta},\ \ 0<x<y<\infty,$$

Thus the conditional expected value

$$E(Y\,|\,x)=\int_x^{+\infty}y\cdot\frac{1}{\theta}e^{-(y-x)/\theta}dy=x+\theta.$$

In fact, X is exponential distribution with parameter $\theta/2$, so the expected value of $X+\theta$ is

$$E(X+\theta)=\theta/2+\theta=3\theta/2,$$

And the variance of $X+\theta$ is

$$D(X+\theta)=D(X)=\theta^2/4,$$

This completes the proof.

Section 7.4. Completeness and Uniqueness

7.23. If $az^2 + bz + c = 0$ for more than two values of z, then $a = b = c = 0$. Use this result to show that the family $\{b(2,\theta): 0 < \theta < 1\}$ is complete.

Solution

Let X be a binomial distribution with parameters 2 and θ, that is, $X \sim b(2,\theta)$. Suppose that there is a statistic $u(X)$ such that $E(u(X)) = 0$, for every $0 < \theta < 1$. Then

$$E(u(X)) = u(0)\binom{2}{0}\theta^0(1-\theta)^2 + u(1)\binom{2}{1}\theta^1(1-\theta)^1 + u(2)\binom{2}{2}\theta^2(1-\theta)^0 = 0,$$

Equivalently, we have

$$(u(0) - 2u(1) + u(2))\theta^2 + (-2u(0) + 2u(1))\theta + u(0) = 0,$$

The equation holds for every $0 < \theta < 1$, from the preceding discussion we know

$$u(0) = u(1) = u(2) = 0,$$

this implies that the family $\{b(2,\theta): 0 < \theta < 1\}$ is complete.

7.24. Show that each of the following families is not complete by finding at least one nonzero function $u(x)$ such that $E(u(X)) = 0,$ for all $\theta > 0$.

(a)

$$f(x;\theta) = \frac{1}{2\theta}, \quad -\theta < x < \theta, \text{ where } 0 < \theta < \infty$$
$$= 0, \quad \text{elsewhere.}$$

(b) $N(0,\theta),$ where $0 < \theta < \infty$.

Solution

(a) Here let $u(X) = -X$, it is obvious that $E(u(X)) = 0,$ for all $\theta > 0$, but $u(x) \not\equiv 0,$ $-\theta < x < \theta$, for any $\theta > 0$.

(b) Here let $u(X) = -X$, it is obvious that $E(u(X)) = 0,$ for all $\theta > 0$, but $u(x) \not\equiv 0,$ $-\theta < x < \theta$, for any $\theta > 0$.

7.25. Let $X_1, X_2, \cdots, X_n$ represent a random sample from the discrete distribution having the probability density function

$$f(x;\theta) = \theta^x(1-\theta)^{1-x}, \quad x = 0, 1, \ 0 < \theta < 1,$$
$$= 0, \quad \text{elsewhere.}$$

Show that $Y_1 = \sum_{i=1}^{n} X_i$ is a complete sufficient statistic for θ. Find the unique function of Y_1 that is the unbiased minimum variance estimator of θ.

Solution

It follows from the factorization theorem that the statistic $Y_1 = \sum_{i=1}^{n} X_i$ is a sufficient statistic for θ, and the statistic

$Y_1 = \sum_{i=1}^{n} X_i$ is binomial distribution with parameters n and θ.

Suppose that there exists a statistic $u(Y_1)$ such that $E(u(Y_1)) = 0$, so we have

$$E(u(Y_1)) = \sum_{i=0}^{n} u(i)\Pr(Y_1 = i) = \sum_{i=0}^{n} u(i)\binom{n}{i}\theta^i(1-\theta)^{n-i} = u(0)(1-\theta)^n + u(1)\theta(1-\theta)^{n-1} + \cdots + u(n)\theta^n = 0.$$

Obviously, the constant in the preceding equation is $u(0)$ and every term in the equation is nonnegative, so we have $u(0)=0$, thus the equation becomes the following

$$u(1)\theta(1-\theta)^{n-1}+\cdots+u(n)\theta^{n}=0,$$

Divide both members of the equation by $\theta \neq 0$, we obtain

$$u(1)(1-\theta)^{n-1}+\cdots+u(n)\theta^{n-1}=0,$$

which implies that the coefficient $u(1)=0$. Repeat the process n times, we can get

$$u(0)=u(1)=\cdots=u(n)=0,$$

This proves that the sufficient statistic $Y_1=\sum_{i=1}^{n} X_i$ is a complete sufficient statistic for θ.

7.26. Consider the family of probability density functions $\{h(z;\theta):\theta \in \Omega\}$, where $h(z;\theta)=1/\theta, 0<x<\theta$, zero elsewhere.

(a) Show that the family is complete provided that $\Omega=\{\theta: 0<\theta<\infty\}$.

(b) Show that this family is not complete if $\Omega=\{\theta: 1<\theta<\infty\}$.

Solution

(a) Suppose that there exists a statistic $u(Z)$ such that $E(u(Z))=0$ for all $\theta>0$, that is,

$$E[u(Z)]=\int_0^\theta u(z)\frac{1}{\theta}dz=\frac{1}{\theta}\int_0^\theta u(z)dz=0, \text{ for all } \theta>0,$$

From the formulation, we have

$$\int_0^\theta u(z)dz=0, \text{ for all } \theta>0.$$

Note that the derivative of $E(u(Z))$ with respect to θ is equal to zero also, thus

$$u(\theta)=0, \text{ for all } \theta>0.$$

This implies that the family of probability density function is complete provided that $\Omega=\{\theta: 0<\theta<\infty\}$.

(b) By the similar reasoning, we can obtain

$$\int_0^\theta u(z)dz=0, \text{ for all } \theta>1,$$

Since

$$\int_0^\theta u(z)dz=\int_0^1 u(z)dz+\int_1^\theta u(z)dz=0, \text{ for all } \theta>1,$$

We can construct the function

$$u(z)=\begin{cases}1-2z, & 0<z<1\\ 0, & 1<z<\theta\end{cases}.$$

It is obvious that the function $u(Z)$ satisfies $E(u(Z))=0$ for all $\theta>0$, whereas $u(Z)\neq 0$ for all $\theta>1$, this prove that the family is not complete if $\Omega=\{\theta: 1<\theta<\infty\}$.

Section 7.5. The Exponential Class of Probability Density Functions

7.29. Write the p.d.f

$$f(x;\theta)=\frac{1}{6\theta^4}x^3e^{-x/\theta},\ 0<x<\infty,\ 0<\theta<\infty,$$

Zero elsewhere, in the exponential form. If $X_1,X_2,\cdots,X_n$ is a random sample from this distribution, find a complete sufficient statistic Y_1 for θ and the unique function $\varphi(Y_1)$ of this statistic that is the unbiased minimum variance estimator of θ. Is $\varphi(Y_1)$ itself a complete sufficient statistic?

Solution

The p.d.f. of the distribution can be converted into the following form

$$f(x;\theta)=\exp(-x/\theta+3\ln x-\ln 6\theta^4),\ 0<x<\infty,\ 0<\theta<\infty,$$

So the statistic $Y_1=\sum_{i=1}^{n}X_i$ is a complete sufficient statistic for θ.

In fact, the population is gamma distribution with parameters 4 and θ, so the expected value of the distribution is $E(X)=4\theta$, thus

$$E(Y_1)=E(\sum_{i=1}^{n}X_i)=4n\theta,\quad E(Y_1/4n)=\theta,$$

So we obtain the unique unbiased minimum variance estimator of θ $\varphi(Y_1)=Y_1/4n$.

Here, the statistic $\varphi(Y_1)$ itself is also a complete sufficient statistic for θ.

7.30. Let $X_1,X_2,\cdots,X_n$ denote a random sample of size $n>1$ from a distribution with p.d.f. $f(x;\theta)=\theta e^{-\theta x}, 0<x<\infty$, zero elsewhere, and $\theta>0$. Then $Y=\sum_{1}^{n}X_i$ is a sufficient statistic for θ. Prove that $(n-1)/Y$ is the unbiased minimum variance estimator of θ.

Solution

Since

$$f(x,\theta)=\exp(-\theta x+\ln\theta),$$

Thus,

$$Y=\sum_{1}^{n}X_i$$

is a sufficient complete statistics of θ.

In fact, Y is Gamma distribution with parameters n and $\frac{1}{\theta}$, so

$$E(n-1/Y)=(n-1)E(1/Y)=(n-1)\int_0^\infty\frac{1}{y}\frac{\theta^n}{\Gamma(n)}y^{n-1}e^{-\theta y}=\theta,$$

In accordance with the Rao-Blackwell theorem, we have $(n-1)/Y$ is the unbiased minimum variance estimator of θ.

7.31. Let $X_1,X_2,\cdots,X_n$ denote a random sample of size n from a distribution with p.d.f.

$f(x;\theta)=\theta x^{\theta-1}$, $0<x<1$, zero elsewhere, and $\theta>0$.

(a) Show that the *geometric mean* $(X_1X_2\cdots X_n)^{1/n}$ of the sample is a complete sufficient statistic for θ.

(b) Find the maximum likelihood estimator of θ, and observe that it is a function of this geometric mean.

Solution

(a) Since the p.d.f. can be changed into the following formula

$$f(x;\theta)=\exp((\theta-1)\ln x+\ln\theta),\quad 0<x<1,$$

So the statistic $\sum_{i=1}^{n}\ln X_i=\ln(X_1X_2\cdots X_n)$ is a complete sufficient statistic for θ, furthermore, the geometric mean $(X_1X_2\cdots X_n)^{1/n}$ of the sample which is a monotonic increasing function of $\ln(X_1X_2\cdots X_n)$ is also a complete sufficient statistic for θ.

(b) The likelihood function of the distribution is

$$L(\theta)=\theta^n(x_1x_2\cdots x_n)^{\theta-1},\ 0<x_i<1,\ i=1,2,\cdots n,$$

The logarithm likelihood function is

$$\ln L(\theta)=n\ln\theta+(\theta-1)\sum_{i=1}^{n}\ln x_i,\ 0<x_i<1,\ i=1,2,\cdots n,$$

The derivative of the logarithm likelihood function with respect to θ is

$$\frac{d\ln L(\theta)}{d\theta}=\frac{n}{\theta}+\sum_{i=1}^{n}\ln x_i\ ,$$

Let $\frac{d\ln L(\theta)}{d\theta}=0$ whose solution is

$$\hat{\theta}=-\frac{n}{\sum_{i=1}^{n}\ln x_i}=-\frac{1}{\ln(x_1x_2\cdots x_n)^{1/n}},$$

which is a function of this geometric mean.

7.32. Let $\overline{X}$ denote the mean of the random sample $X_1,X_2,\cdots,X_n$ from a gamma-type distribution with parameters $\alpha>0$ and $\beta=\theta>0$. Compute $E[X_1\,|\,\overline{x}]$.

Solution

The likelihood function of the population is

$$L(\theta)=\left(\frac{1}{\Gamma(\alpha)\theta^{\alpha}}\right)^n(x_1x_2\cdots x_n)^{\alpha-1}e^{-\sum x_i/\theta},$$

By the factorization theorem, we know that the statistic $\overline{X}$ is a sufficient statistic for θ, and it is not difficult to prove that

$$E(\overline{X}/\alpha)=\theta\,.$$

In fact, the population is gamma distribution with parameters $\alpha>0$ and $\beta=\theta>0$, that is, $X\sim\Gamma(\alpha,\theta)$.

So we have $E(X_1)=E(X)=\alpha\theta$, thus by the Rao and Blackwell theorem, we have

$$E[X_1 \mid \bar{x}]=\alpha\cdot\overline{X}/\alpha=\overline{X}\ .$$

7.33. Let X be a random variable with a p.d.f. of a regular case of the exponential class. Show that $E[K(X)]=-q'(\theta)/p'(\theta)$, provided these derivatives exist, by differentiating both member of the equality

$$\int_a^b \exp[p(\theta)K(x)+S(x)+q(\theta)]dx=1 \qquad (1)$$

with respect to θ. By a second differentiation, find the variance of $K(X)$.

Solution

By differentiating both member of the equality

$$\int_a^b \exp[p(\theta)K(x)+S(x)+q(\theta)]dx=1$$

with respect to θ, we have

$$\int_a^b [K(x)p'(\theta)+q'(\theta)]\exp[p(\theta)K(x)+S(x)+q(\theta)]dx=0, \qquad (2)$$

this indicates that

$$E[K(X)p'(\theta)+q'(\theta)]=0,$$

equivalently,

$$E[K(X)]=-q'(\theta)/p'(\theta)\,.$$

By a second differentiation of the equality (2), we obtain

$$\int_a^b\{[K(x)p''(\theta)+q''(\theta)]\exp[p(\theta)K(x)+S(x)+q(\theta)]+[K(x)p'(\theta)+q'(\theta)]^2\exp[p(\theta)K(x)+S(x)+q(\theta)\}dx=0 \qquad (3)$$

from which we can deduce the variance of $K(X)$,

$$D(K(X))=\frac{p''(\theta)q'(\theta)-p'(\theta)q''(\theta)}{[p'(\theta)]^3}\,.$$

7.34. Given that $f(x;\theta)=\exp[\theta K(x)+S(x)+q(\theta)],\ a<x<b,\ \gamma<\theta<\delta$, represents a regular case of the exponential class, show that the moment-generating function $M(t)$ of $Y=K(X)$ is

$$M(t)=\exp[q(\theta)-q(\theta+t)],\ \gamma<\theta+t<\delta\,.$$

Solution

By the definition of the moment-generating function, we have

$$\begin{aligned}M(t)=E(e^{tK(X)})&=\int_a^b e^{tK(x)}\exp[\theta K(x)+S(x)+q(\theta)]dx\\&=\int_a^b \exp[(\theta+t)K(x)+S(x)+q(\theta+t)+q(\theta)-q(\theta+t)]dx\\&=\exp[q(\theta)-q(\theta+t)],\ \gamma<\theta+t<\delta.\end{aligned}$$

Thus we complete the proof.

7.35. Given, in the preceding exercise, that $E(Y)=E[K(X)]=\theta$. Prove that Y is $N(\theta,1)$.

Solution

In the preceding exercise,

$$M'(t)=-q'(\theta+t)\exp[q(\theta)-q(\theta+t)],$$

Since

$$M'(0)=E(Y)\,.$$

Thus we obtain the differential equation

$$-q'(\theta) = \theta,$$

So we obtain

$$q(\theta) = -\theta^2 / 2,$$

Thus the moment-generating of $Y = K(X)$ is

$$M(t) = \exp(\theta t + t^2 / 2),$$

which is exact the moment- generating of the normal distribution with mean θ and variance 1, that is $Y = K(X) \sim N(\theta, 1)$.

7.36. If $X_1, X_2, \cdots, X_n$ is a random sample from a distribution that has a p.d.f. which is a regular case of the exponential class, show that the p.d.f. of $Y_1 = \sum_{i=1}^{n} K(X_i)$ is of the form $g_1(y_1;\theta) = R(y_1)\exp[p(\theta)y_1 + nq(\theta)]$.

Solution

The joint p.d.f. of the sample $X_1, X_2, \cdots, X_n$ is

$$f(x_1, x_2, \cdots, x_n) = \exp[p(\theta)\sum_{i=1}^{n} K(x_i) + \sum_{i=1}^{n} S(x_i) + nq(\theta)].$$

Let $Y_1 = \sum_{i=1}^{n} K(X_i)$ and $Y_2 = X_2, Y_3 = X_3, \cdots, Y_n = X_n$, then the Jacobian of the transformation is

$|J| = J(y_1, y_2, \cdots, y_n)$,

then the joint p.d.f. of the sample $Y_1, Y_2, \cdots, Y_n$ is

$$f(y_1, y_2, \cdots, y_n) = \exp[p(\theta)y_1 + \sum_{i=1}^{n} S(u^{-1}(y_i)) + nq(\theta)]J(y_1, y_2 \cdots y_n),$$

The marginal p.d.f. of the statistic $Y_1 = \sum_{i=1}^{n} K(X_i)$ is

$$g_1(y_1;\theta) = \underset{R^{n-1}}{\int\int\cdots\int} f(y_1, y_2, \cdots, y_n)\, dy_2 \cdots dy_n = R(y_1)\exp[p(\theta)y_1 + nq(\theta)].$$

7.37. Let Y denote the median and let $\overline{X}$ denote the mean of a random sample of size $n = 2k+1$ from a distribution that is $N(\mu, \sigma^2)$. Compute $E[Y \mid \overline{x}]$.

Solution

7.38. Let $X_1, X_2, \cdots, X_n$ be a random sample from a distribution with p. d. f.

$$f(x;\theta) = \theta^2 x e^{-\theta x},\ 0 < x < \infty, \text{ where } \theta > 0.$$

(a) Argue that $Y = \sum_{1}^{n} X_i$ is a complete sufficient statistic for θ.

(b) Compute $\mathrm{E}(1/Y)$ and find the function of Y which is the unique unbiased minimum variance estimator of θ.

Solution

(a) since

$$f(x;\theta)=\exp\{-\theta x+\ln x+2\ln\theta\}$$

So $Y=\sum_1^n X_i$ is a complete sufficient statistic for θ.

(b) In fact, X_i is gamma distribution with parameters 2 and $1/\theta$, thus

$Y=\sum_1^n X_i$ is also gamma distribution with parameters $2n$ and $1/\theta$.

$$E(1/Y)=\int_0^\infty \frac{1}{y}\frac{\theta^{2n}}{\Gamma(2n)}y^{2n-1}e^{-\theta y}dy=\frac{\theta}{2n-1}$$

Accordingly, $E((2n-1)/Y)=\theta$, $\frac{2n-1}{Y}$ is the unique unbiased minimum variance estimator of θ.

7.39. Let $X_1,X_2,\cdots,X_n$, $n>2$, be a random sample from the binomial distribution $b(1,\theta)$.

(a) Show that $Y_1=X_1+X_2+\cdots+X_n$ is a complete sufficient statistic for θ.
(b) Find the function $\varphi(Y_1)$ which is the unbiased minimum variance estimator of θ.
(c) Let $Y_2=(X_1+X_2)/2$ and compute $E(Y_2)$.
(d) Determine $E(Y_2\,|\,Y_1=y_1)$.

Solution

(a) The p.d.f. can be changed into the form

$$\theta^x(1-\theta)^{1-x}=\exp[x\ln\frac{\theta}{1-\theta}+\ln(1-\theta)],$$

from which we can know that the statistic $Y_1=X_1+X_2+\cdots+X_n$ is a complete sufficient statistic for θ.
(b) In fact, $Y_1=X_1+X_2+\cdots+X_n\sim b(n,\theta)$, so we have

$$E(Y_1)=n\theta,\ E(Y_1/n)=\theta.$$

It follows the Lehmann and Scheffe theorem that

$$\varphi(Y_1)=Y_1/n=\overline{X}$$

is the unbiased minimum variance estimator of θ.
(c) It is very easy to verify that $E(Y_2)=\theta$.
(d) By the Rao and Blackwell theorem , we have

$$E(Y_2\,|\,Y_1=y_1)=\varphi(y_1)=y_1/n=\bar{x}.$$

ADDITIONAL EXERCISES

7.68. Let $X_1, X_2, \cdots, X_n$ denote a random sample of size $n > 1$ from a distribution with p.d.f. $f(x;\theta) = \theta e^{-\theta x}$, $0 < x < \infty$, zero elsewhere, and $\theta > 0$.

(a) What is the complete sufficient statistic, say Y, for θ?
(b) What function of Y is an unbiased estimator of θ?

Solution

(a) Since

$$f(x;\theta) = \theta e^{-\theta x} = \exp\{-\theta x + \ln\theta\}$$

So $Y = \sum_1^n X_i$ is a complete sufficient estimator for θ.

(b) In fact, according to the property of gamma distribution, we have

$Y = \sum_1^n X_i$ is gamma distribution with parameters n and $1/\theta$.

then

$$E((n-1)/Y) = \int_0^\infty \frac{n-1}{y} \frac{\theta^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\theta y} dy = \theta$$

This indicates that $(n-1)/Y$ is the unbiased minimum variance estimator of θ.

7.69. Let $Y_1 < Y_2 < \cdots < Y_n$ be the order statistics of a random sample of size n from a distribution with p.d.f. $f(x;\theta) = 1/\theta, 0 < x < \theta$, zero elsewhere. The statistics Y_n is a complete sufficient statistics for θ and it has p.d.f.

$$g(y_n;\theta) = \frac{n{y_n}^{n-1}}{\theta^n},\ 0 < y_n < \theta,$$

and zero elsewhere.

(a) Find the distribution function $H_n(z;\theta)$ of $Z = n(\theta - Y_n)$.

(b) Find the $\lim_{n\to\infty} H_n(z;\theta)$ and thus the limiting distribution of Z.

Solution

(a) Since $Z = n(\theta - Y_n)$, we have the p.d.f. of Z is

$$h(z;\theta) = \frac{n(n\theta - z)^{n-1}}{(n\theta)^n},\ 0 < z < n\theta.$$

Thus the distribution function of Z is

$$H_n(z;\theta) = \int_0^z \frac{n(n\theta - t)^{n-1}}{(n\theta)^n} dt = 1 - (1 - \frac{z}{n\theta})^n,\ 0 < z < n\theta.$$

(b) $\lim_{n\to\infty} H_n(z;\theta) = \lim_{n\to\infty} 1 - (1 - \frac{z}{n\theta})^n = 1 - e^{-z/\theta}$, thus the limiting distribution of Z is exponential distribution with mean θ.

7.71. Let $X_1, X_2, \cdots, X_n$ be a random sample of size n from a Poisson distribution with mean θ. Find the conditional expectation $E[X_1 + 2X_2 + 3X_3 \mid \sum_{i=1}^{n} X_i]$.

Solution

It follows from the factorization theorem that the sum of the sample $\sum_{i=1}^{n} X_i$ is a sufficient statistic for θ, and

$\sum_{i=1}^{n} X_i \sim P(n\theta)$, so $E\left(\sum_{i=1}^{n} X_i\right) = n\theta,\ E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \theta$.

Moreover, $E[X_1 + 2X_2 + 3X_3] = 6\theta$, thus we have

$$E[X_1 + 2X_2 + 3X_3 \mid \sum_{i=1}^{n} X_i] = \frac{6}{n}\sum_{i=1}^{n} X_i .$$

7.72. Let $X_1, X_2, \cdots, X_n$ be a random sample of size n from the normal distribution $N(\theta,1)$. Find the unbiased minimum variance estimator of θ^2.

Solution

The p.d.f. of the distribution can be changed into the following formula

$$\frac{1}{\sqrt{2\pi}} e^{-(x-\theta)^2/2} = \exp[\theta x - x^2/2 - \theta^2/2 - \ln 2\pi/2],$$

It is easy seen that the statistic $\frac{1}{n}\sum_{i=1}^{n} X_i$ is a complete sufficient statistic for θ^2.

In fact, $\frac{1}{n}\sum_{i=1}^{n} X_i \sim N(\theta, \frac{1}{n})$, and $E(\overline{X}) = \theta, D(\overline{X}) = \frac{1}{n}$, so $E(\overline{X})^2 = D(\overline{X}) + [E(\overline{X})]^2 = \theta^2 + \frac{1}{n}$, thus we have

$$E[(\overline{X})^2 - \frac{1}{n}] = \theta^2,$$

So the statistic $(\overline{X})^2 - \frac{1}{n}$ is the unbiased minimum variance estimator of θ^2.

7.73. Let $X_1, X_2, \cdots, X_n$ be a random sample of size n from a Poissonl distribution with mean θ. Find the unbiased minimum variance estimator of θ^2.

Solution

The p.d.f. of the population can be changed into

$$\frac{\theta^x}{x!} e^{-\theta} = \exp[x \ln\theta - \ln x! - \theta],$$

So the statistic $\sum_{i=1}^{n} X_i$ is a complete sufficient statistic for θ^2, and $\sum_{i=1}^{n} X_i$ is Poisson distribution with mean $n\theta$,

$$E(\sum_{i=1}^{n} X_i) = D(\sum_{i=1}^{n} X_i) = n\theta,\ E(\overline{X})^2 = D(\overline{X}) + [E(\overline{X})]^2 = \theta^2 + \frac{\theta}{n}.$$

So we can construct the statistic $(\overline{X})^2 - \overline{X}/n$ such that $E[(\overline{X})^2 - \overline{X}/n] = \theta^2$, so the unbiased minimum variance estimator of θ^2 is $(\overline{X})^2 - \overline{X}/n$.

7.76. Let $X_1, X_2, \cdots, X_n$ be a random sample from a distribution with p.d.f. $f(x;\theta) = \theta^x(1-\theta), x = 0, 1, 2, \cdots$, zero elsewhere, where $0 \le \theta \le 1$.

(a) Find the m.l.e. $\hat{\theta}$ of θ.

(b) Show that $\sum_{i=1}^{n} X_i$ is a complete sufficient statistic for θ.

(c) Determine the unbiased minimum variance estimator of θ.

Solution

The likelihood function of the population is

$$L(\theta) = \theta^{\sum X_i}(1-\theta)^n,$$

The logarithm likelihood function is

$$\ln L(\theta) = \sum_{i=1}^{n} X_i \cdot \ln\theta + n\ln(1-\theta),$$

Let the derivative of the function $\ln L(\theta)$ with respect to θ be equate to zero, we have

$$\sum_{i=1}^{n} X_i / n - \frac{n}{1-\theta} = 0,$$

whose solution is

$$\hat{\theta} = \sum_{i=1}^{n} X_i / (n + \sum_{i=1}^{n} X_i),$$

the statistic is exact the m.l.e. of θ.

(b) The p.d.f. of the population can be changed into

$$\theta^x(1-\theta) = \exp[x\ln\theta + \ln(1-\theta)],$$

So the statistic $\sum_{i=1}^{n} X_i$ is a complete sufficient statistic for θ .

(b) The expected value of the population is

$$E(X) = \frac{\theta}{1-\theta},$$

Chapter 8 More About Estimation

Section 8.1 Bayesian Estimation

8.1. Let $X_1, X_2, \cdots, X_n$ be a random sample from a distribution that is $b(1,\theta)$. Let the prior p.d.f. of Θ be a beta one with parameters α and β. Show that the posterior p.d.f. $k(\theta \mid x_1, x_2, \cdots, x_n)$ is exactly the same as $k(\theta \mid y)$ given in Example 2.

Solution

The conditional p.d.f. of Θ, given $X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n$, is

$$k(\theta \mid x_1, x_2, \cdots, x_n) \propto \theta^{\sum x_I}(1-\theta)^{n-\sum x_i}\theta^{\alpha-1}(1-\theta)^{\beta-1} \propto \theta^{\sum x_i+\alpha-1}(1-\theta)^{n-\sum x_i+\beta-1}$$

Provided that $0<\theta<\infty$, and is equal to zero elsewhere. This conditional p.d.f. is one of the beta type with parameters $\sum x_i + \alpha$ and $n - \sum x_i + \beta$.
We can see that the the posterior p.d.f. $k(\theta \mid x_1, x_2, \cdots, x_n)$ is exactly the same as $k(\theta \mid y)$ given in Example 2.

8.2. Let $X_1, X_2, \cdots, X_n$ denote a random sample from a distribution that is $N(\theta, \sigma^2), -\infty<\theta<\infty$, where σ^2 is a given positive number. Let $Y = \overline{X}$, the mean of the random sample. Take the loss function to be $L[\theta, \delta(y)] = |\theta - \delta(y)|$. If θ is an observed value of the random variable Θ that is $N(\mu, \tau^2)$, where $\tau^2 > 0$ and μ are known numbers, find the Bayes' solution $\delta(y)$ for a point estimate of θ.

Solution

The sufficient statistic of θ is $Y = \overline{X}$, and $\overline{X} \sim N(\theta, \sigma^2/n)$. If the prior p.d.f. is taken as $N(\mu, \tau^2)$, then

$$k(\theta \mid y) \propto \frac{1}{\sqrt{2\pi}\sigma/\sqrt{n}}\frac{1}{\sqrt{2\pi}\mu}\exp\left[-\frac{(y-\theta)^2}{2(\sigma^2/n)} - \frac{(\theta-\mu)}{2\tau^2}\right] \propto \exp\left[-\frac{\left(\theta - \dfrac{y\tau^2+\mu\sigma^2/n}{\tau^2+\sigma^2/n}\right)^2}{\dfrac{2(\sigma^2/n)\tau^2}{(\tau^2+\sigma^2/n)}}\right].$$

That is, the posterior p.d.f. of the parameter is obviluosly normal with mean $\dfrac{y\tau^2+\mu\sigma^2/n}{\tau^2+\sigma^2/n}$ and variance $\dfrac{2(\sigma^2/n)\tau^2}{(\tau^2+\sigma^2/n)}$. If we take the loss function to be $L[\theta, \delta(y)] = |\theta - \delta(y)|$, then the Bayes' solution $\delta(y)$ for a point estimate of θ is

$$\frac{y\tau^2+\mu\sigma^2/n}{\tau^2+\sigma^2/n}.$$

8.3. Let $X_1, X_2, \cdots, X_n$ denote a random sample from a Poisson distribution with mean $\theta, 0<\theta<\infty$. Let and take the loss function to be $L(\theta, \delta(Y)) = [\theta - \delta(Y)]^2$. Let θ be an observed value of the random variable Θ. If Θ has the p.d.f. $h(\theta) = \theta^{\alpha-1}e^{-\theta/\beta}/\Gamma(\alpha)\beta^\alpha, 0<\theta<\infty$, zero elsewhere, where $\alpha>0, \beta>0$. Find the Bayes' solution $\delta(y)$ for a point estimate of θ.

Solution

In accordance with the factorization theorem of Neyman, it is easily proved that

$$Y = \sum_{1}^{n} X_i$$

is a sufficient of θ.

Here $Y = \sum_{1}^{n} X_i$ is Poisson distribution with mean $n\theta$, this means that the p.d.f. of Y is

$$f(y;\theta) = \frac{(n\theta)^y}{y!} e^{-n\theta}, y = 0, 1, 2, \cdots$$

Since the prior distribution of Θ is $h(\theta) = \theta^{\alpha-1}e^{-\theta/\beta} / \Gamma(\alpha)\beta^{\alpha}$, $0 < \theta < \infty$,

So the posterior distribution of Θ is

$$k(\theta \mid y) \propto (n\theta)^y \theta^{\alpha-1} e^{-n\theta} e^{-\theta/\beta} \propto \theta^{y+\alpha-1} e^{-\theta/(\frac{\beta}{n\beta+1})}$$

This explains that the posterior distribution of Θ is gamma distribution with parameters $y+\alpha$ and $\frac{\beta}{n\beta+1}$.

If we take the loss function to be $L(\theta, \delta(Y)) = [\theta - \delta(Y)]^2$, then the Bayes' solution $\delta(y)$ is

$$\hat{\theta} = \frac{\beta(Y+\alpha)}{n\beta+1} \quad .$$

8.4. Let Y_n be the nth order statistic of a random sample of size n from a distribution with p.d.f. $f(x;\theta) = 1/\theta,\ 0 < x < \theta,$ zero elsewhere. Take the loss function to be $L(\theta, \delta(Y_n)) = [\theta - \delta(Y_n)]^2$. let Θ have p.d.f. $h(\theta) = \beta\alpha^{\beta} / \theta^{\beta+1},\ \alpha < \theta < \infty,$ with $\alpha > 0, \beta > 0$. Find the Bayes' solution $\delta(y_n)$ for a point estimate of θ.

Solution

In accordance with the factorization theorem of Neyman, it is easily proved that Y_n is a sufficient of θ.

The p.d.f. of Y_n is

$$g(y;\theta) = \frac{ny^{n-1}}{\theta^n}, \qquad 0 < y < \theta$$

Since the prior distribution of Θ is $h(\theta) = \beta\alpha^{\beta} / \theta^{\beta+1},\ \alpha < \theta < \infty,$ So the posterior distribution of Θ is

$$k(\theta \mid y) \propto \frac{1}{\theta^{n+\beta+1}}, \ \ y < \theta$$

This explains that the posterior distribution of Θ is $k(\theta \mid y) = (n+\beta)y^{n+\beta} / \theta^{n+\beta+1},\ \ y < \theta$

If we take the loss function to be $L(\theta, \delta(Y)) = [\theta - \delta(Y)]^2$, then the Bayes' solution is

$$\hat{\theta} = E(\Theta \mid y) = \int_y^{\infty} (n+\beta)y^{n+\beta} / \theta^{n+\beta} d\theta = \frac{n+\beta}{n+\beta-1} Y_n .$$

8.5. Let Y_1 and Y_2 be statistics that have a trinomial distribution with parameters n, θ_1, and θ_2. Here θ_1 and

θ_2 are observed values of the random variables Θ_1 and Θ_2, which have a Dirichlet distribution with known parameters α_1, α_2, and α_3. Show that the conditional distribution of Θ_1 and Θ_2 is Dirichlet distribution and determine the conditional means $E(\Theta_1 \mid y_1, y_2)$ and $E(\Theta_2 \mid y_1, y_2)$.

Solution

8.9. Let Y_4 be the largest order statistic of a sample of size $n = 4$ from a distribution with uniform p.d.f. $f(x;\theta) = 1/\theta, 0 < x < \theta$, zero elsewhere. If the prior p.d.f. of the parameter is $g(\theta) = 2/\theta^3$, $1 < \theta < \infty$, zero elsewhere, find the Bayesian estimator $\delta(Y_4)$ of θ, based upon the sufficient statistic Y_4, using the loss function $L[\theta, \delta(y)] = |\theta - \delta(y)|$.

Solution

If $y \le 1$, the posterior distribution of Θ is

$$k(\theta \mid y) \propto \frac{2}{\theta^3}, \quad 0 < y < \theta, 1 < \theta < \infty,$$

Then the distribution function of the posterior distribution is

If $y \le 1$, $F(\theta) =$

If we take the loss function to be $L[\theta, \delta(y)] = |\theta - \delta(y)|$, then the Bayes' solution $\delta(y)$ is the median of the posterior distribution

.

Section 8.2 Fisher Information and the Rao-Cramer Inequality

8.11. Prove that $\overline{X}$, the mean of a random sample of size n from a distribution that is $N(\theta,\sigma^2)$, is , for every known $\sigma^2 > 0$, an efficient estimator of θ.

Solution

The p.d.f. of the population

$$f(x\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\theta)^2}{2\sigma^2}\right], \qquad -\infty < x < \infty,$$

where $-\infty < \theta < \infty$, and

$$\ln f(x\theta) = -\frac{1}{2}\ln(2\pi\sigma^2) - \frac{(x-\theta)^2}{2\sigma^2},$$

Thus

$$\frac{\partial \ln f(x;\theta)}{\partial\theta} = \frac{x-\theta}{\sigma^2}$$

and

$$\frac{\partial^2 \ln f(x;\theta)}{\partial\theta^2} = -\frac{1}{\sigma^2}.$$

So

$$I(\theta) = -E\left[\frac{\partial^2 \ln f(x;\theta)}{\partial\theta^2}\right] = \frac{1}{\sigma^2}.$$

On one hand,

$$E(\overline{X}) = \theta$$

which indicates that $\overline{X}$ is an unbiased estimator of θ.

On the other hand,

$$D(\overline{X}) = \frac{\sigma^2}{n} = \frac{1}{nI(\theta)}$$

This implies that the variance of $\overline{X}$ achieves the lower bound of the Rao-Cramer inequality, so $\overline{X}$ is an efficient estimator of θ.

8.2. Show that the mean $\overline{X}$ of a random sample of size n from a distribution which is $b(1,\theta)$, $0 < \theta < 1$, is an efficient estimator of θ.

Solution

Firstly, we have

$$E\overline{X} = EX = \theta.$$

Then

$$I(\theta) = -E\left(\frac{\partial^2 \ln f(X;\theta)}{\partial\theta^2}\right) = \frac{1}{\theta(1-\theta)}$$

So the Rao-Cramer lower bound is

$$\frac{1}{nI(\theta)} = \frac{\theta(1-\theta)}{n}$$

Furthermore,

$$D(\overline{X}) = \frac{D(X)}{n} = \frac{\theta(1-\theta)}{n}$$

This indicates that the variance of $\overline{X}$ achieves the Rao-Cramer lower bound, so the mean $\overline{X}$ of a random sample of size n is an efficient estimator of θ.

8.14. Given the p.d.f.

$$f(x;\theta) = \frac{1}{\pi[1+(x-\theta)^2]}, -\infty < x < \infty, \ -\infty < \theta < \infty.$$

Show that the Rao-Cramer lower bound is $2/n$, where n is the size of a random sample from Cauchy distribution.

Solution

Since

$$f(x;\theta) = \frac{1}{\pi[1+(x-\theta)^2]}, -\infty < x < \infty, \ -\infty < \theta < \infty.$$

and

$$\frac{\partial \ln f(x;\theta)}{\partial \theta} = \frac{2(x-\theta)}{1+(x-\theta)^2},$$

So the Fisher information is

$$I(\theta) = E\left(\frac{\partial \ln f(x;\theta)}{\partial \theta}\right)^2 = \frac{1}{2},$$

Thus the Rao-Cramer lower bound is

$$\frac{1}{nI(\theta)} = \frac{2}{n}.$$

8.15. Let X have a gamma distribution with $\alpha = 4$ and $\beta = \theta > 0$.

(a) Find the Fisher information $I(\theta)$.

(b) If $X_1, X_2, \cdots, X_n$ is a random sample from this distribution, show that the m.l.e. of θ is an efficient estimator of θ.

Solution

(a)

$$f(x;\theta) = \frac{1}{6\theta^4} x^3 e^{-x/\theta}, \ x > 0,$$

$$\frac{\partial^2 \ln f(X;\theta)}{\partial \theta^2} = \frac{4}{\theta^2} - \frac{2X}{\theta^3}$$

$$I(\theta) = -E(\frac{\partial^2 \ln f(X;\theta)}{\partial \theta^2}) = \frac{4}{\theta^2}$$

(b) It is easily proved that the m.l.e. of θ is $\hat{\theta} = \dfrac{\sum_{i=1}^{n} X_i}{4n}$

and that $$E(\hat{\theta}) = \theta$$

$$D(\hat{\theta}) = \frac{1}{16n} 4\theta^2 = \frac{\theta^2}{4n} = \frac{1}{nI(\theta)}$$

This indicates that the m.l.e. of θ is an efficient estimator for θ.

8.16. Let X be $N(0,\theta), 0<\theta<\infty.$

(a) Find the Fisher information $I(\theta)$.

(b) If $X_1, X_2, \cdots, X_n$ is a random sample from this distribution ,show that the m.l.e. of θ is an efficient estimator of θ.

Solution

(a)The p.d.f. of X is

$$f(x;\theta)=\frac{1}{\sqrt{2\pi\theta}}e^{-\frac{x^2}{2\theta}}, -\infty<x<\infty, 0<\theta<\infty$$

$$\ln f(x;\theta)=-\frac{x^2}{2\theta}-\frac{1}{2}\ln(2\pi\theta)$$

and

$$\frac{\partial \ln f(x;\theta)}{\partial\theta}=\frac{x^2}{2\theta^2}-\frac{1}{2\theta},$$
$$\frac{\partial^2 \ln f(x;\theta)}{\partial\theta^2}=\frac{-x^2}{\theta^3}+\frac{1}{2\theta^2}$$

Thus

$$I(\theta)=-E[\frac{\partial^2 \ln f(X;\theta)}{\partial\theta^2}]=E[\frac{X^2}{\theta^3}-\frac{1}{2\theta^2}]$$
$$=\frac{1}{\theta^3}E(X^2)-\frac{1}{2\theta^2}=\frac{\theta}{\theta^3}-\frac{1}{2\theta^2}=\frac{1}{2\theta^2}$$

Hence $I(\theta)$ is $\frac{1}{2\theta^2}$.

(b) since $I(\theta)$ is $\frac{1}{2\theta^2}$,then the Rao-Cramer lower bound $\frac{1}{nI(\theta)}$ is $\frac{2\theta^2}{n}$.

and $X_1, X_2, \cdots, X_n$ is a random sample from this distribution, so the likelihood function is

$$L(\theta;x_1,x_2,\cdots,x_n)=f(x_1;\theta)f(x_2;\theta)\cdots f(x_n;\theta)$$
$$=(2\pi\theta)^{-\frac{n}{2}}e^{-\sum_1^n\frac{x_i^2}{2\theta}}$$

$$\ln L(\theta;x_1,x_2,\cdots,x_n)=-\frac{n}{2}\ln(2\pi\theta)-\sum_1^n\frac{x_i^2}{2\theta}$$

We note that $L(\theta;x_1,x_2,\cdots,x_n)$ and $\ln L(\theta;x_1,x_2,\cdots,x_n)$ is maximized at the value of θ.so it may be easier to solve

$$\frac{d\ln L(\theta;x_1,x_2,\cdots,x_n)}{d\theta}=0$$

and we have $\frac{d\ln L(\theta;x_1,x_2,\cdots,x_n)}{d\theta}=-\frac{n}{2\theta}+\frac{\sum_1^n x_i^2}{2\theta^2}=\frac{-n\theta+\sum_1^n x_i^2}{2\theta^2}=0$, the solution of the parameter is

$\frac{\sum_{i=1}^{n} x_i^2}{n}$ .Thus the statistic $\hat{\theta} = \frac{\sum_{i=1}^{n} X_i^2}{n}$ is the m.l.e. of θ.

It is easily shown that $E\hat{\theta} = \frac{\sum_{1}^{n} E(X_i^2)}{n} = E(X_i^2) = \theta$, so $\hat{\theta} = \frac{\sum_{i=1}^{n} X_i^2}{n}$ is an unbiased estimator of θ and the variance of $\hat{\theta}$ is

$$\operatorname{var}(\hat{\theta}) = \operatorname{var}(\frac{\sum_{1}^{n} X_i^2}{n}) = \frac{\operatorname{var}(X_i^2)}{n} = \frac{E(X_i^4) - [E(X_i^2)]^2}{n} = \frac{3\theta^2 - \theta^2}{n} = \frac{2\theta^2}{n}$$

Thus

$$\operatorname{var}(\hat{\theta}) = \frac{1}{nI(\theta)} \cdot$$

So the m.l.e. of θ is an efficient estimator of θ.

8.3. Limiting Distribution of Maximum Likelihood Estimation

8.17. Let $X_1, X_2, \cdots, X_n$ be a random sample from each of the following distributions. In each case, find the m.l.e. $\hat{\theta}$, $\text{var}(\hat{\theta})$, $1/nI(\theta)$, where $I(\theta)$ is the Fisher information of a single observation X, and compare $\text{var}(\hat{\theta})$ and $1/nI(\theta)$.

(a) $b(1,\theta),\ 0 \le \theta \le 1$.

(b) $N(\theta,1), -\infty < \theta < \infty$.

(c) $N(0,\theta),\ 0 < \theta < \infty$.

(d) Gamma ($\alpha = 5, \beta = \theta$), $0 < \theta < \infty$.

Solution

(a),

The m.l.e. of θ is $\hat{\theta} = \overline{X}$, and $\text{var}(\hat{\theta}) = \dfrac{\theta(1-\theta)}{n}$, $\dfrac{1}{nI(\theta)} = \dfrac{\theta(1-\theta)}{n}$. So $\text{var}(\hat{\theta}) = \dfrac{1}{nI(\theta)} = \dfrac{\theta(1-\theta)}{n}$.

(b)

The m.l.e. of θ is

$$\hat{\theta} = \overline{X},$$

and $\overline{X} \sim N(\theta, 1/n)$, thus $\text{var}(\hat{\theta}) = \dfrac{1}{n}$, and $I(\theta) = 1$, so $\text{var}(\hat{\theta}) = \dfrac{1}{nI(\theta)} = \dfrac{1}{n}$.

(c) The m.l.e.of θ is

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} X_i^{\,2},$$

$$\text{var}(\hat{\theta}) = \text{var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i^{\,2}\right) = \text{var}\left(\frac{\theta}{n}\cdot\frac{1}{\theta}\sum_{i=1}^{n} X_i^{\,2}\right) = \frac{2\theta^2}{n}.$$

The Fisher information of a single observation X is

$I(\theta) = \dfrac{1}{2\theta^2}$. So the lower bound of Rao-Cramer inequality is

$$1/nI(\theta) = \frac{2\theta^2}{n} = \text{var}(\hat{\theta}).$$

(d)

The m.l.e. of θ is

$$\hat{\theta} = \frac{1}{5n}\sum_{i=1}^{n} X_i.$$

And $\text{var}(\hat{\theta}) = \dfrac{\theta^2}{5n}$.

The Fisher information of X is

$$I(\theta) = \frac{5}{\theta^2},$$

Thus the lower bound of the Rao-Cramer inequality is

$1/nI(\theta)=\dfrac{\theta^2}{5n}$. So $\operatorname{var}(\hat{\theta})=\dfrac{1}{nI(\theta)}=\dfrac{\theta^2}{5n}$.

8.18. Referring to Exercise 8.17 and using the fact that $\hat{\theta}$ has an approximate $N[\theta, 1/nI(\theta)]$, in each case construct an approximate 95 percent confidence interval for θ.

Solution

(a)

It follows from the central limiting theorem that the m.l.e. $\overline{X}$ has a limiting distribution as following

$$\overline{X} \to N(\theta, \frac{\theta(1-\theta)}{n}),$$

8.19. Let $(X_1,Y_1),(X_2,Y_2),\cdots,(X_n,Y_n)$ be a random sample from a bivariate normal distribution with unknown means θ_1 and θ_2 and with known variance and correlation coefficient, $\sigma_1{}^2, \sigma_2{}^2$, and ρ, respectively. Find the maximum likelihood estimator $\hat{\theta}_1$ and θ_2 of θ_1 and θ_2 and their approximate variance-covariance matrix. In this case, does the latter provide the exact variance and covariance?

Solution

8.20. Let $(X_1,Y_1),(X_2,Y_2),\cdots,(X_n,Y_n)$ be a random sample from a bivariate normal distribution with means equal to zero and variances θ_1 and θ_2, respectively and known correlation coefficient ρ .Find the maximum likelihood estimator $\hat{\theta}_1$ and θ_2 of θ_1 and θ_2 and their approximate variance-covariance matrix.

Solution

Chapter 9 Theory of statistical Tests

Section 9.1 Certain Best Tests

9.1. In Example2 of this section, let the simple hypotheses read $H_0:\theta=\theta'=0$ and $H_1:\theta=\theta'=-1$. Show that the best test of H_0 against H_1 may be carried out by use of the statistic $\overline{X}$, and that if $n=25$ and $\alpha=0.05$, the power of the test is $0.999+$ when H_1 is true.

Solution

Now

$$\frac{L(\theta';x_1,x_2,\cdots,x_n)}{L(\theta'';x_1,x_2,\cdots,x_n)}=\frac{(1/\sqrt{2\pi})\exp\left[-(\sum_{i=1}^{n}x_i^{\,2})/2\right]}{(1/\sqrt{2\pi})\exp\left[-(\sum_{i=1}^{n}(x_i+1)^2)/2\right]}=\exp\left(\sum_{i=1}^{n}x_i+\frac{n}{2}\right).$$

If $k>0$, the set of all points $(x_1,x_2,\cdots,x_n)$ such that

$$\exp\left(\sum_{i=1}^{n}x_i+\frac{n}{2}\right)\le k$$

is a best critical region. This inequality holds if and only if

$$\sum_{i=1}^{n}x_i+\frac{n}{2}\le\ln k$$

or, equivalently,

$$\overline{X}\le(\ln k-\frac{n}{2})/n=c.$$

If $n=25$ and $\alpha=0.05$, the number $c=-0.329$ can be found from Table III in Appendix B, so that

$$\Pr(\overline{X}\le c;H_0)=0.05\,.$$

Hence the power function of the test when H_1 is true is

$$\Pr(\overline{X}\le c;H_1)=\Pr(5(\overline{X}+1)\le 3.355)=0.999+\,.$$

9.2. Let the random variable X have the p.d.f. $f(x;\theta)=(1/\theta)e^{-x/\theta}, 0<x<\infty$, zero elsewhere. Consider the simple hypothesis $H_0:\theta=\theta'=2$ and the alternative hypothesis $H_1:\theta=\theta'=4$. Let X_1,X_2 denote a random sample of size 2 from this distribution. Show that the best test of of H_0 against H_1 may be carried out by use of the statistic X_1+X_2 and that the assertion in Example 2 of Section 6.4 is correct.

Solution

Now

$$\frac{L(\theta';x_1,x_2,\cdots,x_n)}{L(\theta'';x_1,x_2,\cdots,x_n)}=\frac{\frac{1}{4}e^{-(x_1+x_2)/2}}{\frac{1}{16}e^{-(x_1+x_2)/4}}=4e^{-(x_1+x_2)/4}.$$

If $k>0$, the set of all points (x_1,x_2) such that

$$4e^{-(x_1+x_2)/4}\le k$$

is a best critical region. This inequality holds if and only if

$$-(x_1+x_2)/4\le \ln\frac{k}{4}$$

or, equivalently,

$$x_1+x_2\ge 4(\ln\frac{k}{4})=c.$$

Hence, the critical region of the test is of the form

$$C=\{(x_1,x_2):x_1+x_2\ge c\}.$$

9.3. Repeat Exercise 9.2 when $H_1:\theta=\theta'=6$. Generalize this for every $\theta''>2$.

Solution

Now

$$\frac{L(\theta';x_1,x_2,\cdots,x_n)}{L(\theta'';x_1,x_2,\cdots,x_n)}=\frac{\frac{1}{4}e^{-(x_1+x_2)/2}}{\frac{1}{36}e^{-(x_1+x_2)/6}}=9e^{-(x_1+x_2)/3}.$$

If $k>0$, the set of all points (x_1,x_2) such that

$$9e^{-(x_1+x_2)/3}\le k$$

is a best critical region. This inequality holds if and only if

$$-(x_1+x_2)/3\le \ln\frac{k}{9}$$

or, equivalently,

$$x_1+x_2\ge 3(\ln\frac{k}{9})=c.$$

Hence, the critical region of the test is of the form

$$C=\{(x_1,x_2):x_1+x_2\ge c\}.$$

Similarly, for any $\theta''>2$, we can obtain the critical region as following
$C=\{(x_1,x_2):x_1+x_2\ge c\}$.

9.4. Let $X_{1j},X_{2j},\cdots,X_{a_j j}$ represent independent random samples of sizes a_j from normal distributions with means μ_j and variances $\sigma^2,\ j=1,\ 2,\cdots,b$. Show that

$$\sum_{j=1}^{b}\sum_{i=1}^{a_j}(X_{ij}-\overline{X}_{..})^2=\sum_{j=1}^{b}\sum_{i=1}^{a_j}(X_{ij}-\overline{X}_{.j})^2+\sum_{j=1}^{b}a_j(\overline{X}_{.j}-\overline{X}_{..})^2,\quad \text{or}\quad Q'=Q'_3+Q'_4.$$

Here. If $\mu_1=\mu_2=\cdots=\mu_b$, show that Q'/σ^2 and Q_3'/σ^2 have chi-square distributions.

Solution

$$\sum_{j=1}^{b}\sum_{i=1}^{a_j}(X_{ij}-\overline{X}_{..})^2=\sum_{j=1}^{b}\sum_{i=1}^{a_j}(X_{ij}-\overline{X}_{.j}+\overline{X}_{.j}-\overline{X}_{..})^2$$

$$=\sum_{j=1}^{b}\sum_{i=1}^{a_j}(X_{ij}-\overline{X}_{.j})^2+2\sum_{j=1}^{b}\sum_{i=1}^{a_j}(X_{ij}-\overline{X}_{.j})(\overline{X}_{.j}-\overline{X}_{..})+\sum_{j=1}^{b}a_j(\overline{X}_{.j}-\overline{X}_{..})^2$$

$$=\sum_{j=1}^{b}\sum_{i=1}^{a_j}(X_{ij}-\overline{X}_{.j})^2+\sum_{j=1}^{b}a_j(\overline{X}_{.j}-\overline{X}_{..})^2$$

If $\mu_1=\mu_2=\cdots=\mu_b$, then X_{ij} is (μ,σ^2) regardless of i and j, so

$$Q'/\sigma^2 \text{ is } \chi^2(\sum_{j=1}^{b}a_j-1)$$

$$Q_3'/\sigma^2=\sum a_jS_j{}^2/\sigma^2 \sim \chi^2(\sum(a_j-1))=\chi^2(\sum_{j=1}^{b}a_j-b).$$