Exercises in Introduction to Mathematical Statistics (Ch. 6)

Tomoki Okuno

August 20, 2022

Note

- Not all solutions are provided: exercises that are too simple or not very important to me are skipped.
- Texts in red are just attentions to me. Please ignore them.

6 Maximum Likelihood Method

6.1 Maximum Likelihood Estimation

6.1.1. Let $X_1, X_2, ..., X_n$ be a random sample on X that has a $\Gamma(\alpha = 4, \beta = \theta)$ distribution, $0 < \theta < \infty$.

(a) Determine the mle of θ.

Solution.

$$\begin{aligned}
\ell(\theta) &= \sum_i [-\log\Gamma(4) - 4\log\theta - 3\log x_i - x_i/\theta],\\
\ell'(\theta) &= \sum_i [-4/\theta + x_i/\theta^2] = n(-4\theta + \overline{x})/\theta^2,\\
\ell''(\theta) &= \sum_i [4/\theta^2 - 2x_i/\theta^3].
\end{aligned}$$

Solving $\ell'(\theta) = 0$ obtains $\theta = \overline{x}/4$. Then $\ell''(\overline{x}/4) < 0$. Hence the mle of θ is $\widehat{\theta} = \overline{X}/4$.

(b) Suppose the following data is a realization (rounded) of a random sample on X. Obtain a histogram with the argument pr=T (data are in `ex6111.rda`).

```
9 39 38 23 8 47 21 22 18 10 17 22 14
9 5 26 11 31 15 25 9 29 28 19 8
```

Solution. Skipped.

(c) For this sample, obtain $\widehat{\theta}$ the realized value of the mle and locate $4\widehat{\theta}$ the histogram. Overlay the $\Gamma(\alpha = 4, \beta = \theta)$ pdf on the histogram. Does the data agree with this pdf? Code for overlay:

```
xs=sort(x);y=dgamma(xs,4,1/betahat);hist(x,pr=T);lines(y xs).
```

Solution. Since $\overline{x} = 20.12$, $\widehat{\theta} = 20.12/4 = 5.03$. Graphs are skipped.

6.1.2. Let $X_1, X_2, ..., X_n$ represent a random sample from each of the distributions having the following pdfs:

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

Solution.

$$\ell(\theta) = \sum_i [\log\theta + (\theta-1)\log x_i],$$
$$\ell'(\theta) = \sum_i [1/\theta + \log x_i] = n/\theta + \log\prod x_i,$$
$$\ell''(\theta) = -n/\theta^2 < 0.$$

Solving $\ell'(\theta) = 0$, therefore, we obtain $\widehat{\theta} = -n/\log\prod_i x_i$

(b) $f(x;\theta) = e^{-(x-\theta)}$, $\theta \le x < \infty$, $-\infty < \theta < \infty$, zero elsewhere. Note that this is a nonregular case.

Solution.

$$L(\theta) = \begin{cases} e^{-\sum(x_i-\theta)} & \theta \le x_i,\ i=1,...,n \\ 0 & \text{otherwise} \end{cases} = \begin{cases} e^{-n(\overline{x}-\theta)} & \theta \le x_{(1)} \\ 0 & \text{otherwise} \end{cases}$$

Since $L'(\theta) = ne^{-n(\overline{x}-\theta)} > 0$, $L(\theta)$ is strictly increasing, indicating that θ that maximizes $L(\theta)$ is $x_{(1)}$. Hence, the mle of θ is $\widehat{\theta} = X_{(1)}$.

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

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

is a mle 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, uniqueness is not, in general, a property of mles.

Solution.

$L(\theta;\mathbf{x}) = 1$ (constant) if

$$\theta - \frac{1}{2} \le Y_1 \text{ and } Y_n \le \theta + \frac{1}{2} \ \Rightarrow\ Y_n - \frac{1}{2} \le \theta \le Y_1 + \frac{1}{2},$$

zero elsewhere. Thus, θ that maximizes $L(\theta)$ is inside $[Y_n - 1/2, Y_1 - 1/2]$. That is, let $\widehat{\theta} = u(X_1, ..., X_n)$,

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

For $(4Y_1 + 2Y_n + 1)/6$,

$$\frac{4Y_1 + 2Y_n + 1}{6} - \left(Y_n - \frac{1}{2}\right) = \frac{4(Y_1 - Y_n + 1)}{6} \ge 0,$$
$$\left(Y_1 + \frac{1}{2}\right) - \frac{4Y_1 + 2Y_n + 1}{6} = \frac{2(Y_1 - Y_n + 1)}{6} \ge 0.$$

because $Y_n - Y_1 \le 1$. So do the other two statistics.

6.1.4. Suppose $X_1, ..., X_n$ are iid with pdf $f(x;\theta) = 2x/\theta^2$, $0 < x \le \theta$, zero elsewhere. Note this is a nonregular case. Find:

(a) The mle $\widehat{\theta}$ for θ.

Solution.

$$L(\theta) = \begin{cases} \frac{2^n \sum_i x_i}{\theta^{2n}} & 0 < x_i \le \theta,\ i=1,...,n \\ 0 & \text{otherwise} \end{cases}$$

Since $L'(\theta) < 0$, $L(\theta)$ is strictly decreasing for $\theta \ge x_{(n)} = y_n$. So, θ that maximizes $L(\theta)$ is y_n. Hence, the mle of θ is $\widehat{\theta} = Y_n$.

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

Solution.

By the theorem of pdf of the order statistic,

$$f_{Y_n}(y) = n[F_X(y)]^{n-1} f_X(y) = \frac{2ny^{2n-1}}{\theta^{2n}}.$$

Hence,

$$E(c\widehat{\theta}) = \int_0^\theta cy f_{Y_n}(y)dy = \int_0^\theta \frac{2cny^{2n}}{\theta^{2n}} dy = \frac{2n}{2n+1} c\theta \Rightarrow c = \frac{2n+1}{2n}.$$

(c) The mle for the median of the distribution. Show that it is a consistent estimator.

Solution.

Solving $F_X(x) = 1/2$, we obtain $\theta/\sqrt{2}$. Hence, the mle for the median is $Y_n/\sqrt{2}$. Also,

$$E(Y_n) = \int_0^\theta \frac{2ny^{2n}}{\theta^{2n}} dy = \frac{2n}{2n+1}\theta \to \theta \text{ as } n \to \infty,$$

which implies that $Y_n/\sqrt{2}$ is a consistent estimator.

6.1.5. Consider the pdf in Exercise 6.1.4.

(a) Using Theorem 4.8.1, show how to generate observations from this pdf.

Solution.

Recall $F_X(x) = x^2/\theta^2$. Let $u = F(x)$ then $x = F^{-1}(u) = \theta\sqrt{u}$, $\theta > 0$. Hence, suppose $U \sim U(0,1)$, we would use $X = F^{-1}(u) = \theta U^{1/2}$ to generate observations.

(b) The following data were generated from this pdf. Find the mles of θ and the median.

```
1.2 7.7 4.3 4.1 7.1 6.3 5.3 6.3 5.3 2.8
3.8 7.0 4.5 5.0 6.3 6.7 5.0 7.4 7.5 7.5
```

Solution. $\widehat{\theta} = Y_n = 7.7$, $\widehat{m} = Y_n/\sqrt{2} = 7.7/\sqrt{2} = 5.44$.

6.1.6. Suppose $X_1, X_2, ..., X_n$ are iid with pdf $f(x;\theta) = (1/\theta)e^{-x/\theta}$, $0 < x < \infty$, zero elsewhere. Find the mle of $P(X \le 2)$ and show that it is consistent.

Solution.

Assume $\theta \neq 0$. Since $X_1, \ldots, X_n$ are iid with pdf $f(x;\theta) = e^{-x/\theta}/\theta$,

$$\ell(\theta) = \log L(\theta) = -\sum_i x_i/\theta - n\log\theta$$
$$\ell'(\theta) = \frac{\sum_i x_i}{\theta^2} - \frac{n}{\theta}.$$

Solving $\ell'(\theta) = 0$, we obtain $\theta = \frac{1}{n}\sum_i x_i = \overline{x}$. Hence, the MLE for θ is $\widehat{\theta X}$. For the second derivative,

$$\frac{d^2\ell(\theta)}{d\theta^2} = \frac{-2\sum_i x_i}{\theta^3} + \frac{n}{\theta^2} = \frac{n(\theta - 2\overline{x})}{\theta^3}$$
$$\Rightarrow \frac{d^2\ell(\widehat{\theta})}{d\theta^2} = \frac{n(\widehat{\theta} - 2\overline{x})}{\theta^3} = -\frac{n\overline{x}}{\theta^3} < 0.$$

Since

$$P(X \le 2) = \int_0^2 e^{-x/\theta}/\theta dx = -e^{-x/\theta}\Big|_0^2 = 1 - e^{-2/\theta},$$

$\widehat{P(X \leq 2)} = 1 - e^{-2/\widehat{\theta}} = 1 - e^{-2/\overline{X}}$ by invariance of MLE. Moreover,

$$E(X) = \int_0^\infty \frac{xe^{-x/\theta}}{\theta} dx = \Gamma(2)\theta = \theta,$$

which provides $\widehat{\theta} = \overline{X} \xrightarrow{P} E(X) = \theta$ by WLLN. Hence, let $g(x) = 1 - e^{-2/x}$ (continuous for $x > 0$),

$$1 - e^{-2/\overline{X}} = g(\overline{X}) \xrightarrow{P} g(\theta) = 1 - e^{-2/\theta}$$

by g function. That is, $\widehat{P(X \leq 2)}$ is consistent for $P(X \leq 2)$.

6.1.7. 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 mle of $P(X \geq 3)$. For the data in the table, using the R function `pbinom` determine the realization of the mle.

Solution.

Let p denote a parameter of the Poisson distribution.

$$f(x; p) = P(X = x) = \binom{5}{x} p^x (1-p)^{50-x}, \quad x = 0, 1, 2, ..., 5.$$

We know that the mle of p is $\widehat{p} = \overline{X}/50$. By invariance of mle, the mle of $P(X \geq 3)$ is

$$\widehat{P(X \geq 3)} = \sum_{i=3}^{5} \binom{5}{x} \widehat{p}^x (1 - \widehat{p})^{50-x}.$$

From the table,

$$\widehat{p} = \frac{\overline{x}}{50} = \frac{0(6) + 1(10) + 2(14) + 3(13) + 4(6) + 5(1)}{5(50)} = \frac{106}{250} = 0.424.$$

Hence, the desired realization is

$$\widehat{P(X \geq 3)} = \texttt{1 - pbinom(2, 5, 0.424)} = 0.3597.$$

6.1.9. Let the table

x	0	1	2	3	4	5
Frequency	7	14	12	13	6	3

represent a summary of a random sample of size 55 from a Poisson distribution. Find the maximum likelihood estimator of $P(X = 2)$. Use the R function `dpois` to find the estimator's realization for the data in the table.

Solution.

Let θ denote a parameter of the Poisson distribution.

$$f(x; \theta) = P(X = x) = \frac{e^{-\theta}\theta^x}{x!}, \quad x = 0, 1, 2, ...$$

The previous exercise shows that the mle of θ is $\widehat{\theta} = \overline{X}$. By invariance of mle, the mle of $P(X = 2)$ is

$$\widehat{P(X = 2)} = \frac{e^{-\overline{X}}\overline{X}^2}{2}.$$

From the table, $\widehat{\theta}$'s realization is

$$\overline{x} = \frac{0(7) + 1(14) + 2(12) + 3(13) + 4(6) + 5(3)}{55} = \frac{116}{55} = 2.11.$$

Hence, the desired realization is

$$\widehat{P(X=2)} = \frac{e^{-2.10}(2.10)^2}{2} = \texttt{dpois(2, 2.11)} = 0.2699$$

6.1.10. Let $X_1, X_2, ..., X_n$ be a random sample from a Bernoulli distribution with parameter p. If p is restricted so that we know that $\frac{1}{2} \le p \le 1$, find the mle of this parameter.

Solution.

$$\begin{aligned}
\ell(p) &= \sum[x_i \log p + (1-x_i)\log(1-p)],\\
\ell'(p) &= \sum[x_i/p - (1-x_i)/(1-p)] = \frac{n(\overline{x}-p)}{p(1-p)},\\
\ell''(p) &= \sum[-x_i/p^2 - (1-x_i)/(1-p)^2] < 0.
\end{aligned}$$

Solving $\ell'(p)$ gets $p = \overline{x} < 1$. But we need to consider the restriction: $\frac{1}{2} \le p \le 1$. If $\overline{x} \ge 1/2$, the mle of p is $\overline{X}$, while the mle of p is $1/2$ if $\overline{x} < 1/2$ since $\ell(p)$ is decreasing for $p \ge 1/2$. That is, $\widehat{p} = \max(1/2, \overline{X})$.

6.1.12. Let $X_1, X_2, ..., X_n$ be a random sample from the Poisson distribution with $0 < \theta \le 2$. Show that the mle of θ is $\widehat{\theta} = \min\{\overline{X}, 2\}$.

Solution.

We know that the mle of θ, parameter for a Poisson distribution, is $\overline{X}$ if $\theta > 0$. In this case, θ is restricted to ≤ 2. Since $L(\theta; \mathbf{x})$ is increasing if $\theta < \overline{X}$, it maximizes at $\theta = 2$ if $\overline{X} > 2$, which gives $\widehat{\theta} = \min\{\overline{X}, 2\}$.

6.1.13. Let $X_1, X_2, ..., X_n$ be a random sample from a distribution with one of two pdfs. If $\theta = 1$, then $f(x; \theta = 1) = \frac{1}{2\pi}e^{-x^2/2}$, $-\infty < x < \infty$. If $\theta = 2$, then $f(x; \theta = 2) = 1/[\pi(1+x^2)]$, $-\infty < x < \infty$. Find the mle of θ.

Solution.

$$\widehat{\theta} = \begin{cases} 1 & L(\theta = 1; \mathbf{x}) > L(\theta = 2; \mathbf{x}) \\ 1, 2 & L(1) = L(2) \\ 2 & L(1) < L(2). \end{cases}$$

6.2. Rao–Cramer Lower Bound and Efficiency

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

Solution.

$$\begin{aligned}
\log f(x;\theta) &= -\log\sqrt{2\pi\sigma^2} - \frac{(x-\theta)^2}{2\sigma^2}\\
\frac{\partial \log f(x;\theta)}{\partial\theta} &= \frac{x-\theta}{\sigma^2}, \quad \frac{\partial^2 \log f(x;\theta)}{\partial\theta^2} = -\frac{1}{\sigma^2}\\
\Rightarrow I(\theta) &= -E\left[\frac{\partial^2 \log f(X;\theta)}{\partial\theta^2}\right] = \frac{1}{\sigma^2}.
\end{aligned}$$

Hence, the CRLB is $1/(nI(\theta)) = \sigma^2/n$. Since $\overline{X}$ is unbiased for θ, $\mathrm{Var}(\overline{X}) = \sigma^2/n$ attains the CRLB, which means that $\overline{X}$ is an efficient estimator of θ.

6.2.2. Given $f(x;\theta) = 1/\theta$, $0 < x < \theta$, zero elsewhere, with $\theta > 0$, formally compute the reciprocal of

$$nE\left\{\left[\frac{\partial \log f(X:\theta)}{\partial\theta}\right]^2\right\}.$$

Compare this with the variance of $(n+1)Y_n/n$, where Y_n is the largest observation of a random sample of size n from this distribution. Comment.

Solution.

Note that this is a non-regular case.

$$nE\left\{\left[\frac{\partial \log f(X:\theta)}{\partial\theta}\right]^2\right\} = \frac{n}{\theta^2}.$$

Thus, the reciprocal is θ^2/n. By the theorem of the order statistic,

$$\begin{aligned} f_{Y_n}(y) &= nF_X(y)fX(y) = \frac{ny^{n-1}}{\theta^n} \\ \Rightarrow E(Y_n) &= \cdots = \frac{n}{n+1}\theta, \quad E(Y_n^2) = \cdots = \frac{n}{n+2}\theta^2, \\ \Rightarrow \mathrm{Var}(Y_n) &= E(Y_n^2) - E(Y_n)^2 = \frac{n}{(n+1)^2(n+2)}\theta^2. \end{aligned}$$

Hence,

$$\mathrm{Var}\left(\frac{n+1}{n}Y_n\right) = \frac{(n+1)^2}{n^2}\mathrm{Var}\,(Y_n) = \frac{\theta^2}{n(n+2)} < \frac{\theta^2}{n},$$

which indicates that the variance violates CRLB because of the non-regular case.

6.2.7. Recall Exercise 6.1.1 where $X_1, X_2, ..., X_n$ is a random sample on X that has a $\Gamma(\alpha = 4, \beta = \theta)$ distribution, $0 < \theta < \infty$.

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

Solution.

$$\begin{aligned} \log f(x;\theta) &= K - 4\log\theta + 3\log x - x/\theta \\ \frac{\partial \log f(x;\theta)}{\partial\theta} &= -4/\theta + x/\theta^2, \quad \frac{\partial^2 \log f(x;\theta)}{\partial\theta^2} = 4/\theta^2 - 2x/\theta^3 \\ \Rightarrow I(\theta) &= -E\left[\frac{\partial^2 \log f(x;\theta,\sigma^2)}{\partial\theta^2}\right] = \frac{2E(X)}{\theta^3} - \frac{4}{\theta^2} = \frac{4}{\theta^2}. \end{aligned}$$

(b) Show that the mle of θ, which was derived in Exercise 6.1.1, is an efficient estimator of θ.

Solution.

The mle of θ is $\widehat{\theta} = \overline{X}/4$. Since $E(\widehat{\theta}) = E(\overline{X}/4) = \theta$ and

$$\mathrm{Var}(\widehat{\theta}) = \mathrm{Var}(\overline{X}/4) = \mathrm{Var}(\overline{X})/16 = \theta^2/4n = 1/nI(\theta),$$

$\widehat{\theta}$ is an efficient estimator of θ.

(c) Using Theorem 6.2.2, obtain the asymptotic distribution of $\sqrt{n}(\widehat{\theta} - \theta)$.

Solution. By the asymptotic distribution of MLE, $\sqrt{n}(\widehat{\theta} - \theta) \xrightarrow{D} N(0, \theta^2/4)$.

(d) For the data of Exercise 6.1.1, find the asymptotic 95% confidence interval for θ.

Solution.

$$\frac{\sqrt{n}(\widehat{\theta}-\theta)}{\theta/2} \xrightarrow{D} N(0,1) \;\Rightarrow\; \frac{\sqrt{n}(\widehat{\theta}-\theta)}{\widehat{\theta}/2} = \frac{\sqrt{n}(\widehat{\theta}-\theta)}{\theta/2}\frac{\theta}{\widehat{\theta}} \xrightarrow{D} N(0,1) \quad \text{by WLLN and Slutsky.}$$

Hence,

$$0.95 = P\left(-1.96 < \frac{\sqrt{n}(\widehat{\theta}-\theta)}{\widehat{\theta}/2} < 1.96\right) = P\left(\widehat{\theta} - \frac{0.98\widehat{\theta}}{\sqrt{n}} < \theta < \widehat{\theta} + \frac{0.98\widehat{\theta}}{\sqrt{n}}\right),$$

which gives us the asymptotic 95% confidence interval for θ:

$$\widehat{\theta} \pm \frac{0.98\widehat{\theta}}{\sqrt{n}} = \widehat{\theta}\left(1 \pm \frac{0.98}{\sqrt{n}}\right) = 5.03\left(1 \pm \frac{0.98}{\sqrt{25}}\right) = (4.04, 6.02).$$

because We obtained $\widehat{\theta} = 5.03$ in Exercise 6.1.1.

6.2.8. Let X be $\mathrm{N}(0,\theta)$, $0 < \theta < \infty$.

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

Solution.

$$\begin{aligned}
\log f(x;\theta) &= -\frac{1}{2}\log 2\pi\theta - \frac{x^2}{2\theta} \\
\frac{\partial \log f(x;\theta)}{\partial\theta} &= -\frac{1}{2\theta} + \frac{x^2}{2\theta^2}, \quad \frac{\partial^2 \log f(x;\theta)}{\partial\theta^2} = \frac{1}{2\theta^2} - \frac{x^2}{\theta^3} \\
\Rightarrow I(\theta) &= -E\left[\frac{\partial^2 \log f(x;\theta)}{\partial\theta^2}\right] = -\frac{1}{2\theta^2} + \frac{E(X^2)}{\theta^3} = \frac{1}{2\theta^2}
\end{aligned}$$

because $E(X^2) = \mathrm{Var}(X) = \theta$.

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

Solution.

Solving $\ell'(\theta) = 0$, we obtain the mle of θ: $\widehat{\theta} = \frac{1}{n}\sum_i X_i^2$. Since $X_i/\sqrt{\theta} \sim N(0,1) \Rightarrow \sum X_i^2/\theta \sim \chi^2(n)$, $\mathrm{Var}(\sum X_i^2/\theta) = 2n$, or $\mathrm{Var}(\sum X_i^2) = 2n\theta^2$. Hence

$$\mathrm{Var}(\widehat{\theta}) = Var\left(\frac{1}{n}\sum_i X_i^2\right) = \frac{\mathrm{Var}(\sum_i X_i^2)}{n^2} = \frac{2\theta^2}{n} = \frac{1}{nI(\theta)},$$

meaning that $\widehat{\theta}$ is an efficient estimator of θ.

(c) What is the asymptotic distribution of $\sqrt{n}(\widehat{\theta}-\theta)$?

Solution. By the asymptotic distribution of MLE, $\sqrt{n}(\widehat{\theta}-\theta) \xrightarrow{D} N(0, 2\theta^2)$.

6.2.11. Let $\overline{X}$ be the mean of a random sample of size n from a $N(\theta,\sigma^2)$ distribution, $-\infty < \theta < \infty$, $\sigma^2 > 0$. Assume that σ^2 is known. Show that $\overline{X}^2 - \frac{\sigma^2}{n}$ is an unbiased estimator of θ^2 and find its efficiency.

Solution.

$$E(\overline{X}^2) = \mathrm{Var}(\overline{X}) + [E(\overline{X})]^2 = \frac{\sigma^2}{n} + \theta^2 \;\Rightarrow\; E\left(\overline{X}^2 - \frac{\sigma^2}{n}\right) = \theta^2.$$

For the Fisher information, let $\theta^2 = \mu$,

$$\frac{\partial^2 \log f(x,\mu)}{\partial \mu^2} = \cdots = -\frac{x}{2\sigma^2 \mu}.$$

Hence,

$$I(\mu) = -E\left[\frac{\partial^2 \log f(X,\mu)}{\partial \mu^2}\right] = \frac{E(X)}{2\sigma^2\mu} = \frac{1}{2\sigma^2\sqrt{\mu}} \Rightarrow I(\theta^2) = \frac{1}{2\sigma^2\theta}.$$

Since $E\left(\overline{X}^2 - \frac{\sigma^2}{n}\right) = \theta^2$, the CRLB of the variance of $\overline{X}^2 - \frac{\sigma^2}{n}$ is

$$\mathrm{Var}\left(\overline{X}^2 - \frac{\sigma^2}{n}\right) = \mathrm{Var}(\overline{X}^2) \geq \frac{2\theta}{nI(\theta^2)} = \frac{4\sigma^2\theta^2}{n}.$$

Finally, compute $\mathrm{Var}(\overline{X}^2)$.

$$\begin{aligned}
&\left[\frac{\sqrt{n}(\overline{X}-\theta)}{\sigma}\right]^2 = \frac{n(\overline{X}-\theta)^2}{\sigma^2} \sim \chi^2(1) \\
&\Rightarrow \mathrm{Var}\left(\frac{n(\overline{X}-\theta)^2}{\sigma^2}\right) = \frac{n^2}{\sigma^4}\mathrm{Var}[(\overline{X}-\theta)^2] = 2 \\
&\Rightarrow \mathrm{Var}[(\overline{X}-\theta)^2] = \mathrm{Var}(\overline{X}^2) + 4\theta^2\mathrm{Var}(\overline{X}) = \frac{2\sigma^4}{n^2} \\
&\Rightarrow \mathrm{Var}(\overline{X}^2) = \frac{2\sigma^4}{n^2} - 4\theta^2\mathrm{Var}(\overline{X}) = \frac{2\sigma^4}{n^2} - \frac{4\sigma^2\theta^2}{n}.
\end{aligned}$$

Thus, the efficacy is

$$\frac{1/(nI(\theta^2))}{\mathrm{Var}(\overline{X})} = \frac{\frac{4\sigma^2\theta^2}{n}}{\frac{2\sigma^4}{n^2} - \frac{4\sigma^2\theta^2}{n}},$$

which converges to -1 as $n \to \infty$. Note that it should be incorrect.

6.2.12. Recall that $\widehat{\theta} = -n/\sum_{i=1}^n \log X_i$ is the mle of θ for a beta$(\theta, 1)$ distribution. Also, $W = -\sum_{i=1}^n \log X_i$ has the gamma distribution $\Gamma(n, 1/\theta)$.

(a) Show that $2\theta W$ has a $\chi^2(2n)$ distribution.

Solution.

Since $M_W(t) = (1-t/\theta)^{-n}$, $M_{2\theta W}(t) = M_W(2\theta t) = (1-2t)^{-n}$, indicating $2\theta W \sim \chi^2(2n)$.

(b) Using part **(a)**, find c_1 and c_2 so that

$$P\left(c_1 < \frac{2\theta n}{\widehat{\theta}} < c_2\right) = 1 - \alpha,$$

for $0 < \alpha < 1$. Next, obtain a $(1-\alpha)100\%$ confidence interval for θ.

Solution.

Since $\widehat{\theta} = -n/\sum_{i=1}^n \log X_i = n/W$,

$$\begin{aligned}
1-\alpha &= P\left(\chi^2_{2n,\alpha/2} < 2\theta W < \chi^2_{2n,1-\alpha/2}\right) \\
&= P\left(\chi^2_{2n,\alpha/2} < \frac{2\theta n}{\widehat{\theta}} < \chi^2_{2n,1-\alpha/2}\right) \\
&= P\left(\frac{\widehat{\theta}\chi^2_{2n,\alpha/2}}{2n} < \theta < \frac{\widehat{\theta}\chi^2_{2n,1-\alpha/2}}{2n}\right).
\end{aligned}$$

Hence, $c_1 = \chi^2_{2n,\alpha/2}$ and $c_2 = \chi^2_{2n,1-\alpha/2}$. Also, a $(1-\alpha)100\%$ confidence interval for θ is

$$\left[\frac{\widehat{\theta}\chi^2_{2n,\alpha/2}}{2n}, \frac{\widehat{\theta}\chi^2_{2n,1-\alpha/2}}{2n}\right].$$

(c) For $\alpha = 0.05$ and $n = 10$, compare the length of this interval with the length of the interval found in Example 6.2.6.

Solution.

The length of this interval is

$$\frac{\widehat{\theta}\chi^2_{20,0.975}}{20} - \frac{\widehat{\theta}\chi^2_{20,0.025}}{20} = \frac{\widehat{\theta}(34.17)}{20} - \frac{\widehat{\theta}(9.59)}{20} = 1.22\widehat{\theta}.$$

On the other hand, the length found in Example 6.2.6 is

$$2\frac{z_{0.025}\widehat{\theta}}{\sqrt{10}} = 1.24\widehat{\theta},$$

which means that the length of the approximate CI is very close to that of the exact CI.

6.2.16. Let S^2 be the sample variance of a random sample of size $n > 1$ from $N(\mu, \theta)$, $0 < \theta < \infty$, where μ is known. We know $E(S^2) = \theta$.

(a) What is the efficiency of S^2?

Solution.

First compute the Fisher information for θ.

$$\begin{aligned}
\log f(x;\theta) &= -\frac{1}{2}\log 2\pi\theta - \frac{(x-\mu)^2}{2\theta}, \\
\frac{\partial \log f(x;\theta)}{\partial\theta} &= -\frac{1}{2\theta} + \frac{(x-\mu)^2}{2\theta^2}, \\
\frac{\partial^2 \log f(x;\theta)}{\partial\theta^2} &= \frac{1}{2\theta^2} - \frac{(x-\mu)^2}{\theta^3}.
\end{aligned}$$

Since $E[(X-\mu)^2] = \mathrm{Var}(X) = \theta$,

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

Next, consider $\mathrm{Var}(S^2)$. We have

$$\frac{(n-1)S^2}{\theta} \sim \chi^2(n-1) \;\Rightarrow\; \mathrm{Var}\left(\frac{(n-1)S^2}{\theta}\right) = 2(n-1) \;\Rightarrow\; \mathrm{Var}(S^2) = \frac{2\theta^2}{n-1}.$$

Hence, the efficiency is

$$\frac{1/(nI(\theta))}{\mathrm{Var}(S^2)} = \frac{n-1}{n}.$$

(b) Under these conditions, what is the mle $\widehat{\theta}$ of θ?

Solution.

Part **(a)** implies that

$$\widehat{\theta} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2.$$

(c) What is the asymptotic distribution of $\sqrt{n}(\widehat{\theta}-\theta)$?

Solution. By the asymptotic distribution of MLE, $\sqrt{n}(\widehat{\theta}-\theta) \xrightarrow{D} N(0, 2\theta^2)$.

6.3. Maximum Likelihood Methods

Note that I use the reverise definition of Λ:

$$\Lambda = \frac{L(\widehat{\theta})}{L(\theta_0)}$$

because I learned this in a class. Accordingly, I use $2\log\Lambda$, not $-2\log\Lambda$.

6.3.1. . The following data were generated from an exponential distribution with pdf $f(x;\theta) = (1/\theta)e^{-x/\theta}$, for $x > 0$, where $\theta = 40$.

(a) Histogram the data and locate $\theta_0 = 50$ on the plot.

Solution. Skipped.

(b) Use the test described in Example 6.3.1 to test $H_0 : \theta = 50$ versus $H_1 : \theta \neq 50$. Determine the decision at level $\alpha = 0.10$.

```
19 15 76 23 24 66 27 12 25 7 6 16 51 26 39
```

Solution.

$$\frac{2}{\theta_0}\sum_{1}^{15} X_i = \frac{2}{50}(432) = 17.28.$$

Since $\chi^2_{0.05,30} = 18.49$ and $\chi^2_{0.95,30} = 43.77$, we reject $H_0 : \theta = 50$.

6.3.3. Show that the test with decision rule (6.3.6) is like that of Example 4.6.1 except that here σ^2 is known.

Solution.

$$\left(\frac{\overline{X}-\theta_0}{\sigma/\sqrt{n}}\right)^2 \geq \chi^2_\alpha(1) \;\Leftrightarrow\; \left|\frac{\overline{X}-\theta_0}{\sigma/\sqrt{n}}\right| > z_{\alpha/2}.$$

The decision rule in Example 4.6.1 is an approximate one, but if σ^2 is known, this is the exact decision rule.

6.3.6. Let $X_1, X_2, ..., X_n$ be a random sample from a $N(\mu_0, \sigma^2 = \theta)$ distribution, $where 0 < \theta < \infty$ and μ_0 is known. Show that the likelihood ratio test of $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$ can be based upon the statistic $W = \sum_{i=1}^n (X_i - \mu_0)^2/\theta_0$. Determine the null distribution of W and give, explicitly, the rejection rule for a level α test.

Solution.

We have

$$L(\theta) = (2\pi\theta)^{-n/2}\exp\left[-\sum_{i=1}^n (x_i-\mu_0)^2/(2\theta)\right], \quad \widehat{\theta} = \frac{1}{n}\sum_{i=1}^n (X_i-\mu_0)^2.$$

Hence,

$$\begin{aligned}\Lambda = \frac{L(\widehat{\theta})}{L(\theta_0)} &= \left(\frac{\theta_0}{\widehat{\theta}}\right)^{n/2} \exp\left[-\sum_{i=1}^n (x_i-\mu_0)^2/(2\widehat{\theta}) + \sum_{i=1}^n (x_i-\mu_0)^2/(2\theta_0)\right] \\ &= \left(\frac{n\theta_0}{\sum_{i=1}^n (x_i-\mu_0)^2}\right)^{n/2} \exp\left[-\frac{n}{2} + \frac{1}{2\theta_0}\sum_{i=1}^n (x_i-\mu_0)^2\right] \\ &= (n^{n/2}e^{-n/2})w^{-n/2}e^{w/2} \geq k \quad \Rightarrow w^{-n/2}e^{w/2} \geq k'.\end{aligned}$$

Let $g(w) = \log(w^{-n/2}e^{w/2}) = -(n/2)\log w + w/2$. Then

$$g'(w) = -\frac{n}{2w} + \frac{1}{2}, \quad g''(w) = \frac{n}{2w^2} > 0$$

Hence, $g(w)$ is a convex function with a minimum at $w = n$, which implies that

$$\Lambda \geq k \Rightarrow W \leq c_1,\ W \geq c_2.$$

Moreover, since $W \sim \chi^2(n)$ under H_0, we obtain the rejection rule for level α test as

$$W \leq \chi^2_{\alpha/2,n},\ W \geq \chi^2_{1-\alpha/2,n},$$

where $\chi^2_{\alpha/2,n}$ and $\chi^2_{1-\alpha/2,n}$ are lower and upper critical regions of the chi-square distribution, respectively.

6.3.9. Let $X_1, X_2, ..., X_n$ be a random sample from a Poisson distribution with mean $\theta > 0$.

(a) Show that the likelihood ratio test of $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$ is based upon the statistic $Y = \sum_{i=1}^n X_i$. Obtain the null distribution of Y.

Solution.

Since we have $\widehat{\theta} = \overline{X}$ (omitted the proof),

$$\Lambda = \frac{L(\widehat{\theta})}{L(\theta_0)} = \frac{e^{-\sum x_i}(\sum x_i/n)^{\sum x_i}}{e^{-n\theta_0}\theta_0^{\sum x_i}} = e^{n\theta_0}e^{-\sum x_i}\left(\frac{\sum x_i}{n\theta_0}\right)^{\sum x_i} = e^{n\theta_0}e^{-y}\left(\frac{y}{n\theta_0}\right)^y \equiv e^{n\theta_0}g(y).$$

Since $g(y)$ is a convex function (omitted the proof), for $k > 0$,

$$\Lambda > k \Rightarrow Y < c_1,\ Y > c_2\ (c_1 < c_2).$$

(b) For $\theta_0 = 2$ and $n = 5$, find the significance level of the test that rejects H_0 if $Y \leq 4$ or $Y \geq 17$.

Solution.

Since $Y \sim \text{Poisson}(n\theta_0 = 10)$ under H_0,

$$\alpha = P_{\theta_0=2}(Y \leq 4) + P_{\theta_0=2}(Y \geq 17) = 0.0293 + 0.0270 = 0.0563.$$

6.3.10. Let $X_1, X_2, ..., X_n$ be a random sample from a Bernoulli $b(1, \theta)$ distribution, where $0 < \theta < 1$.

(a) Show that the likelihood ratio test of $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$ is based upon the statistic $Y = \sum_{i=1}^n X_i$. Obtain the null distribution of Y.

Solution.

Since we have $\widehat{\theta} = \overline{X}$ (omitted the proof),

$$\Lambda = \frac{L(\widehat{\theta})}{L(\theta_0)} = \frac{L(\sum x_i/n)}{L(\theta_0)} = \left(\frac{y}{n\theta_0}\right)^y\left(\frac{n-y}{n(1-\theta_0)}\right)^{n-y} = K_1\left(K_2\frac{y}{n-y}\right)^y \equiv K_1 g(y).$$

Since $g(y)$ is a convex function ($g''(y) > 0$),

$$\Lambda > k \Rightarrow Y < c_1,\ Y > c_2\ (c_1 < c_2).$$

(b) For $n = 100$ and $\theta_0 = 1/2$, find c_1 so that the test rejects H_0 when $Y \leq c_1$ or $Y \geq c_2 = 100 - c_1$ has the approximate significance level of $\alpha = 0.05$. Hint: Use the Central Limit Theorem.

Solution.

Since $n\theta_0(1-\theta_0) = 25$, CLT can be applied, thus, $Y \overset{D}{\sim} N(n\theta_0, n\theta_0(1-\theta_0)) = N(50, 25)$ under H_0. Thus,

$$Y < c_1 \Rightarrow \frac{Y-50}{5} < \frac{c-50}{5} = -1.96 \Rightarrow c_1 = 40.2\ (c_2 = 59.8).$$

6.3.11. Let $X_1, X_2, ..., X_n$ be a random sample from a $\Gamma(\alpha = 4, \beta = \theta)$ distribution, where $0 < \theta < \infty$.

(a) Show that the likelihood ratio test of $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$ is based upon the statistic $W = \sum_{i=1}^n X_i$. Obtain the null distribution of $2W/\theta_0$.

Solution.

Since $\widehat{\theta} = \overline{X}/4 = \sum_i X_i/(4n)$ (omitted the proof) and $L(\theta) = (\Gamma(4)\theta^4)^{-n} \prod_i x_i^3 e^{-\sum_i x_i/\theta}$, the LRT statistic is

$$\Lambda = \frac{L(\widehat{\theta})}{L(\theta_0)} = \left(\frac{4n\theta_0}{\sum_i x_i}\right)^{4n} e^{-4n} e^{-\sum_i x_i/\theta_0} = Kw^{-4n}e^{-w/\theta} > k,$$

where $K = (4n\theta_0/e)^{4n}$ and $w = \sum_i x_i$. Let $g(w) = w^{-4n}e^{-w/\theta}$. Consider $\log g(w)$, then we have $(\log g(w))'' > 0 \Rightarrow g''(w) > 0$, meaning that $g(w)$ is a convex function with a minimum. Hence, the likelihood ratio test rejects H_0 if

$$\Lambda > k \;\Rightarrow\; W < c_1,\; W > c_2.$$

Also, we have $W \sim \Gamma(4n, \theta)$ using the mgf of X. Then

$$M_W(t) = (1 - \theta t)^{-4n} \;\Rightarrow\; M_{2W/\theta_0}(t) = M_W(2t/\theta_0) = (1 - 2t)^{-4n},$$

which indicates that $2W/\theta_0 \sim \chi^2(8n)$ under H_0.

(b) For $\theta_0 = 3$ and $n = 5$, find c_1 and c_2 so that the test that rejects H_0 when $W \leq c_1$ or $W \geq c_2$ has significance level 0.05.

Solution.

By part **(a)**,

$$W < c_1,\; W > c_2 \;\Rightarrow\; \frac{2W}{\theta_0} < \frac{2c_1}{\theta_0} = \chi^2_{0.025,8n},\; \frac{2W}{\theta_0} > \frac{2c_2}{\theta_0} = \chi^2_{0.975,8n}.$$

Substituting $\theta_0 = 3$ and $n = 5$, we obtain

$$\begin{aligned} c_1 &= \frac{3}{2}\chi^2_{0.025,40} = 1.5(24.43) = 36.7, \\ c_2 &= \frac{3}{2}\chi^2_{0.975,40} = 1.5(59.34) = 89.0. \end{aligned}$$

6.3.12. Let $X_1, X_2, ..., X_n$ be a random sample from a distribution with pdf $f(x; \theta) = \theta \exp\{-|x|^\theta\}/2\Gamma(1/\theta)$, $-\infty < x < \infty$, where $\theta > 0$. Suppose $\Omega = \{\theta : \theta = 1, 2\}$. Consider the hypotheses $H_0 : \theta = 2$ (a normal distribution) versus $H_1 : \theta = 1$ (a double exponential distribution). Show that the likelihood ratio test can be based on the statistic $W = \sum_{i=1}^n (X_i^2 - |X_i|)$.

Solution.

Since $\Omega = \{\theta : \theta = 1, 2\}$ and $H_0 : \theta = 2$, the LRT statistic is

$$\Lambda = \frac{L(1)}{L(2)} = \frac{e^{-\sum_i |x_i|}/2^n}{e^{-\sum_i x_i^2}/(\sqrt{\pi})^n} = Ke^{\sum_i (x_i^2 - |x_i|)} = Ke^w,$$

where $K > 0$. Then $\Lambda > k \;\Rightarrow\; W > c$, which is the desired result.

6.3.17. Let $X_1, X_2, ..., X_n$ be a random sample from a Poisson distribution with mean $\theta > 0$. Consider testing $H_0 : \theta = \theta_0$ against $H_1 : \theta = \theta_0$.

(a) Obtain the Wald type test of expression (6.3.13).

Solution.

Since $\widehat{\theta} = \overline{X}$ and $I(\theta) = 1/\theta$,

$$\chi^2_W = \left\{\sqrt{nI(\overline{X})}(\overline{X} - \theta_0)\right\}^2 = \left\{\sqrt{\frac{n}{\overline{X}}}(\overline{X} - \theta_0)\right\}^2.$$

(b) Write an R function to compute this test statistic.

Solution. Skipped.

(c) For $\theta_0 = 23$, compute the test statistic and determine the p-value for the following data.

```
27 13 21 24 22 14 17 26 14 22
21 24 19 25 15 25 23 16 20 19
```

Solution.

Since $n = 20$ and $\overline{X} = 20.35$,

$$\chi^2_W = \left\{\sqrt{\frac{20}{20.35}}(20.35 - 23)\right\}^2 = 6.90$$

$$\Rightarrow \; p = P(\chi^2_1 > 6.90) = \texttt{1 - pchisq(6.9, 1)} = 0.0086.$$

Note that for some reason, the textbook answer doubles it (0.0172), which does not make sense for me.

6.3.18. Let $X_1, X_2, ..., X_n$ be a random sample from a $\Gamma(\alpha, \beta)$ distribution where α is known and $\beta > 0$. Determine the likelihood ratio test for $H_0 : \beta = \beta_0$ against $H_1 : \beta = \beta_0$.

Solution.

We have $\widehat{\beta} = \overline{X}/\alpha = \sum_i X_i/(n\alpha)$ (omitted the proof). Hence, the LRT statistic is

$$\Lambda = \frac{L(\widehat{\beta})}{L(\beta_0)} = \cdots = \left(\frac{n\alpha}{e}\right)^{n\alpha}\left(\frac{\beta_0}{\sum_i x_i}\right)^{n\alpha} e^{\sum_i x_i/\beta_0} = Kw^{-n\alpha}e^w,$$

where $K > 0$ and $W = \sum_i X_i/\beta_0 \sim \Gamma(n\alpha, 1)$. Let $g(w) = w^{-n\alpha}e^w$, then $g'(n\alpha) = 0$ and $g''(w) > 0$. Thus, $g(w)$ is a convex function with minimum. Hence, the likelihood ratio test rejects H_0 if $W < c_1$ or $W > c_2$.

6.3.19. Let $Y_1 < Y_2 < \cdots < Y_n$ be the order statistics of a random sample from a uniform distribution on $(0, \theta)$, where $\theta > 0$.

(a) Show that Λ for testing $H_0 : \theta = \theta_0$ against $H_1 : \theta = \theta_0$ is $\Lambda = (Y_n/\theta_0)^n$, $Y_n \le \theta_0$, and $\Lambda = 0$ if $Y_n > \theta_0$.

Solution.

$$L(\theta, \mathbf{x}) = \begin{cases} \theta^{-n} & \theta \ge y_n \\ 0 & \theta < y_n. \end{cases}$$

Since $L'(\theta) < 0$, i.e., $L(\theta)$ is strictly decreasing for $\theta > y_n$, $\widehat{\theta} = Y_n$. Hence,

$$\Lambda = \frac{L(\widehat{\theta})}{L(\theta_0)} = \begin{cases} (\theta_0/Y_n)^n & \theta_0 \ge Y_n \\ 0 & \theta_0 < Y_n \end{cases} \quad \text{under } H_0.$$

(b) When H_0 is true, show that $-2\log\Lambda$ has an exact $\chi^2(2)$ distribution, not $\chi^2(1)$. Note that the regularity conditions are not satisfied.

Solution.

We have the pdf of Y_n:

$$f_{Y_n}(y) = \frac{n!}{(n-1)!}[F_X(y)]^{n-1} f_X(y) = \frac{ny^{n-1}}{\theta_0^n}.$$

Let $W = 2\log\Lambda = 2n(\log\theta_0 - \log Y_n)$. The inverse one-to-one transformation is

$$\log y_n = \log\theta_0 - \frac{w}{2n} \;\Rightarrow\; y_n = \theta_0 e^{-w/2n} \;\Rightarrow\; \frac{dy}{dw} = -\frac{\theta_0}{2n}e^{-w/2n}.$$

Hence, the pdf of W is

$$f_W(w) = f_{Y_n}(\theta_0 e^{-w/2n})\left|\frac{dy}{dw}\right| = \frac{n\theta_0^{n-1}e^{-w(n-1)/2n}}{\theta_0^n}\frac{\theta_0}{2n}e^{-w/2n} = \frac{1}{2}e^{-w/2},$$

which means $W \sim \Gamma(1,2) = \chi^2(2)$.

6.4. Multiparameter Case: Estimation

6.4.2. Let $X_1, X_2, ..., X_n$ and $Y_1, Y_2, ..., Y_m$ be independent random samples from $N(\theta_1, \theta_3)$ and $N(\theta_2, \theta_4)$ distributions, respectively.

(a) If $\Omega \subset R^3$ is defined by $\Omega = \{(\theta_1, \theta_2, \theta_3) : -\infty < \theta_i < \infty,\ i = 1, 2; 0 < \theta_3 = \theta_4 < \infty\}$, find the mles of θ_1, θ_2, and θ_3.

Solution.

Let $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)'$.

$$L(\boldsymbol{\theta}) = \left(\frac{1}{2\pi\theta_3}\right)^{(n+m)/2} \exp\left[-\frac{\sum_{i=1}^n (x_i-\theta_1)^2 + \sum_{j=1}^m (y_i-\theta_2)^2}{2\theta_3}\right],$$

$$\ell(\boldsymbol{\theta}) = -\frac{n+m}{2}\log 2\pi\theta_3 - \frac{\sum_{i=1}^n (x_i-\theta_1)^2 + \sum_{j=1}^m (y_i-\theta_2)^2}{2\theta_3}.$$

Hence,

$$\frac{\partial\ell(\boldsymbol{\theta})}{\partial\theta_1} = 0 \;\Rightarrow\; \widehat{\theta}_1 = \overline{X} \quad \frac{\partial\ell(\boldsymbol{\theta})}{\partial\theta_2} = 0 \;\Rightarrow\; \widehat{\theta}_2 = \overline{Y},$$

$$\frac{\partial\ell(\boldsymbol{\theta})}{\partial\theta_3} = 0 \;\Rightarrow\; \widehat{\theta}_3 = \frac{1}{n+m}\left[\sum_{i=1}^n (X_i - \overline{X})^2 + \sum_{j=1}^m (Y_i - \overline{Y})^2\right].$$

We also need to check the second derivatives of $\ell(\boldsymbol{\theta})$ w.r.t θ_1, θ_2, and θ_3 are all negative.

(b) If $\Omega \subset R^2$ is defined by $\Omega = \{(\theta_1, \theta_3) : -\infty < \theta_1 = \theta_2 < \infty; 0 < \theta_3 = \theta_4 < \infty\}$, find the mles of θ_1 and θ_3.

Solution.

$$\ell(\boldsymbol{\theta}) = -\frac{n+m}{2}\log 2\pi\theta_3 - \frac{\sum_{i=1}^n (x_i-\theta_1)^2 + \sum_{j=1}^m (y_i-\theta_1)^2}{2\theta_3}.$$

Hence,

$$\frac{\partial\ell(\boldsymbol{\theta})}{\partial\theta_1} = 0 \;\Rightarrow\; \widehat{\theta}_1 = \frac{n\overline{X} + m\overline{Y}}{n+m},$$

$$\frac{\partial\ell(\boldsymbol{\theta})}{\partial\theta_3} = 0 \;\Rightarrow\; \widehat{\theta}_3 = \frac{1}{n+m}\left[\sum_{i=1}^n (X_i - \widehat{\theta}_1)^2 + \sum_{j=1}^m (Y_i - \widehat{\theta}_1)^2\right].$$

We also need to check the second derivatives of $\ell(\boldsymbol{\theta})$ with respect to θ_1 and θ_3 are all negative.

6.4.3. Let $X_1, X_2, ..., X_n$ be iid, each with the distribution having pdf $f(x;\theta_1,\theta_2) = (1/\theta_2)e^{-(x-\theta_1)/\theta_2}$, $\theta_1 \le x < \infty$, $-\infty < \theta_2 < \infty$, zero elsewhere. Find the maximum likelihood estimators of θ_1 and θ_2.

Solution.

This is a nonregular case because of the support of θ_1.

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

for $\forall i$. Since $\partial L/\partial\theta_1 > 0$, L is strictly increasing for θ_1. Hence the minimum of $X_1, X_2, ..., X_n$ maximizes $\partial L(\theta_1,\theta_2;\mathbf{x})$ in terms of θ_1: $\widehat{\theta}_1 = Y_1$. Also,

$$\begin{aligned}
\ell(\theta_1,\theta_2) &= -n\log\theta_2 - \frac{\sum_i x_i - n\theta_1}{\theta_2} \\
\frac{\partial\ell(\theta_1,\theta_2)}{\partial\theta_2} &= -\frac{n}{\theta_2} + \frac{\sum_i x_i - n\theta_1}{\theta_2^2}.
\end{aligned}$$

Hence, solving $\frac{\partial\ell(\theta_1,\theta_2)}{\partial\theta_2} = 0$, we obtain

$$\widehat{\theta}_2 = \frac{\sum_i X_i - n\widehat{\theta}_1}{n} = \frac{\sum_i X_i - nY_1}{n} = \overline{X} - Y_1.$$

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

$$F(x;\theta_1,\theta_2) = \begin{cases} 1 - (\theta_1/x)^{\theta_2} & \theta_1 \le x \\ 0 & \text{elsewhere,} \end{cases}$$

where $\theta_1 > 0$ and $\theta_2 > 0$. If $X_1, X_2, ..., X_n$ is a random sample from this distribution, find the maximum likelihood estimators of θ_1 and θ_2. (Hint: This exercise deals with a nonregular case.)

Solution.

$$\begin{aligned}
f(x;\theta_1,\theta_2) &= -\theta_2\left(\frac{\theta_1}{x}\right)^{\theta_2-1}\left(-\frac{\theta_1}{x^2}\right) = \frac{\theta_2\theta_1^{\theta_2}}{x^{\theta_2+1}}, \quad \theta_1 \le x \\
\Rightarrow L(\theta_1,\theta_2;\mathbf{x}) &= \frac{(\theta_2\theta_1^{\theta_2})^n}{\prod_i x_i^{\theta_2+1}}, \quad \theta_1 \le x_1,
\end{aligned}$$

zero elsewhere. Since $\partial L/\partial\theta_1 > 0$, or L is strictly increasing for θ_1, $\widehat{\theta}_1 = X_{(1)} = Y_1$.

$$\begin{aligned}
\ell(\theta_1,\theta_2) &= \sum\left[\log\theta_2 + \theta_2\log\theta_1 - (\theta_2+1)\log x_i\right], \\
\frac{\partial\ell(\theta_1,\theta_2)}{\partial\theta_2} &= \sum\left[1/\theta_2 + \log\theta_1 - \log x_i\right] = n/\theta_2 + n\log\theta_1 - \log\prod x_i.
\end{aligned}$$

Hence, solving $\frac{\partial\ell(\theta_1,\theta_2)}{\partial\theta_2} = 0$, we obtain

$$\widehat{\theta}_2 = \frac{n}{\log\prod_i x_i - n\log\widehat{\theta}_1} = \frac{n}{\log[\prod_i x_i/Y_1^n]}.$$

6.4.5. 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 $[\theta-\rho, \theta+\rho]$. Find the maximum likelihood estimators for θ and ρ. Are these two unbiased estimators?

Solution.

$L(\theta, \rho) = (2\rho)^{-n}, \ \theta - \rho < x_i < \theta + \rho$, zero elsewhere. Hence,

$$\widehat{\theta} - \widehat{\rho} = Y_1, \quad \widehat{\theta} + \widehat{\rho} = Y_n \quad \Rightarrow \quad \widehat{\theta} = \frac{Y_1 + Y_n}{2}, \quad \widehat{\rho} = \frac{Y_n - Y_1}{2}.$$

(Omitted the check of unbiasness, but they both should be biased).

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

(a) If the constant b is defined by the equation $P(X \le b) = 0.90$, find the mle of b.

Solution.

$$0.90 = P(X \le b) = P\left(\frac{X - \mu}{\sigma} \le \frac{b - \mu}{\sigma}\right) \ \Rightarrow \ \frac{b - \mu}{\sigma} = 1.28 \ \Rightarrow \ b = \mu + 1.28\sigma.$$

We know

$$\widehat{\mu} = \overline{X}, \quad \widehat{\sigma} = \sqrt{\frac{1}{n}\sum_i (X_i - \overline{X})^2} = \sqrt{\frac{n-1}{n}}S.$$

Thus, the mle of b is

$$\widehat{b} = \overline{X} + 1.28\sqrt{\frac{n-1}{n}}S.$$

(b) If c is given constant, find the mle of $P(X \le c)$.

Solution.

$$P(X \le c) = P\left(\frac{X - \mu}{\sigma} \le \frac{c - \mu}{\sigma}\right) = \Phi\left(\frac{c - \mu}{\sigma}\right)$$
$$\Rightarrow P(\widehat{X \le c}) = \Phi\left(\frac{c - \widehat{\mu}}{\widehat{\sigma}}\right) = \Phi\left(\frac{c - \overline{X}}{\sqrt{(n-1)/n}S}\right).$$

6.4.10. Show that if X_i follows the model (6.4.14), then its pdf is $b^{-1}f((x - a)/b)$.

Solution.

Since $X = a + be$ can be transformed to $e = (X - a)/b$,

$$f_X(x) = f((X - a)/b)\left|\frac{de}{dx}\right| = b^{-1}f((x - a)/b).$$

6.5. Multiparameter Case: Testing

Note that I use the reverise definition of Λ:

$$\Lambda = \frac{L(\widehat{\theta})}{L(\theta_0)}$$

because I learned this in a class. Accordingly, I use $2\log\Lambda$, not $-2\log\Lambda$.

6.5.1. On page 80 of their test, Hollander and Wolfe (1999) present measurements of the ratio of the earth's mass to that of its moon that were made by 7 different spacecraft (5 of the Mariner type and 2 of the Pioneer type). These measurements are presented below (also in the file `earthmoon.rda`). Based on earlier Ranger voyages, scientists had set this ratio at 81.3035. Assuming a normal distribution, test the hypotheses $H_0 : \mu = 81.3035$ versus $H_1 : \mu = 81.3035$, where μ is the true mean ratio of these later voyages. Using the p-value, conclude in terms of the problem at the nominal α-level of 0.05.

Earth to Moon Mass Ratios						
81.3001	81.3015	81.3006	81.3011	81.2997	81.3005	81.3021

Solution.

From the LRT statistic:

$$\Lambda = \frac{L(\widehat{\mu}, \widehat{\sigma}^2)}{L(\mu_0, \widehat{\sigma}_0^2)} = \frac{L(\overline{X}, (n-1/n)S^2)}{L(\mu_0, (n-1/n)S^2)} > k \quad (k > 0),$$

we obtain the rejection criteria under H_0:

$$\left| \frac{\sqrt{n}(\overline{X} - \mu_0)}{S} \right| > t_{0.025, n-1}.$$

Since $t_{0.025,n-1} = t_{0.025,6} = 2.45$ and

$$\frac{\sqrt{n}(\overline{X} - \mu_0)}{S} = \frac{\sqrt{7}(81.3008 - 81.3035)}{0.000827} = -8.64,$$

we reject H_0.

6.5.2. Obtain the boxplot of the data in Exercise 6.5.1. Mark the value 81.3035 on the plot. Compute the 95% confidence interval for μ, (4.2.3), and mark its endpoints on the plot. Comment.

Solution.

Omitted the boxplot, the mark, and the plot of the endpoints. 95% confidence interval for μ is

$$\overline{X} \pm t_{\alpha/2,n-1} \frac{S}{\sqrt{n}} = 81.3008 \pm 2.45 \frac{0.000827}{\sqrt{7}} = (81.30004, 81.30156).$$

6.5.4. Let $X_1, X_2, ..., X_n$ be a random sample from the distribution $N(\theta_1, \theta_2)$. Show that the likelihood ratio principle for testing $H_0 : \theta_2 = \theta_2'$ specified, and θ_1 unspecified against $H_1 : \theta_2 \neq \theta_2'$, θ_1 unspecified, leads to a test that rejects when $\sum_1^n (x_i - \overline{x})^2 \leq c_1$ or $\sum_1^n (x_i - \overline{x})^2 \geq c_2$, where $c_1 < c_2$ are selected appropriately.

Solution.

By the previous exercises, we have

$$\widehat{\theta}_1 = \overline{X}, \quad \widehat{\theta}_2 = n^{-1} \sum_{i=1}^{n} (X_i - \overline{X})^2 \quad \text{under } \Omega,$$

$$\widehat{\theta}_{10} = \overline{X} \quad \text{under } H_0.$$

Hence, the LRT statistic is

$$\Lambda = \frac{L(\widehat{\theta}_1, \widehat{\theta}_2)}{L(\widehat{\theta}_{10}, \theta_2')} = \cdots = \left(\frac{n}{e}\right)^{n/2} w^{-n/2} e^{w/2} = K g(w),$$

where $K > 0$, $w = \sum_{i=1}^n (x_i - \overline{x})^2 / \theta_2'$, and $g(w) = w^{-n/2} e^{w/2}$. Since $g(w)$ is a convex function with a minimum at $w = n$ (omitted the proof),

$$\Lambda > k \;\Rightarrow\; w \leq k_1 \text{ or } w \geq k_2 \;\Rightarrow\; \sum_{i=1}^{n} (x_i - \overline{x})^2 \leq c_1 \text{ or } \sum_{i=1}^{n} (x_i - \overline{x})^2 \geq c_2,$$

where $c_1 = \theta_2' k_1$ and $c_2 = \theta_2' k_2$.

. Let $X_1, ..., X_n$ and $Y_1, ..., Y_m$ be independent random samples from the distributions $N(\theta_1, \theta_3)$ and $N(\theta_2, \theta_4)$, respectively.

(a) Show that the likelihood ratio for testing $H_0 : \theta_1 = \theta_2,\ \theta_3 = \theta_4$ against all alternatives is given by

$$\frac{\left[\sum_1^n (x_i - \overline{x})^2/n\right]^{n/2} \left[\sum_1^m (y_i - \overline{y})^2/m\right]^{m/2}}{\left\{\left[\sum_1^n (x_i - u)^2 + \sum_1^m (y_i - u)^2\right]/(n+m)\right\}^{(n+m)/2}}$$

where $u = (n\overline{x} + m\overline{y})/(n+m)$.

Solution.

On the whole space Ω, by the previous exercises,

$$\widehat{\theta}_1 = \overline{X},\ \widehat{\theta}_2 = \overline{Y},$$
$$\widehat{\theta}_3 = \frac{1}{n}\sum_1^n (X_i - \overline{X})^2,\ \widehat{\theta}_4 = \frac{1}{m}\sum_1^n (Y_i - \overline{Y})^2.$$

Under H_0, on the other hand,

$$\widehat{\theta}_1 = \overline{X},$$
$$\widehat{\theta}_{30} = \widehat{\theta}_{40} = \frac{1}{n+m}\left[\sum_1^n (X_i - U)^2 + \sum_1^m (Y_i - U)^2\right].$$

Hence, $\Lambda = L(\widehat{\theta}_1, \widehat{\theta}_2, \widehat{\theta}_3, \widehat{\theta}_4)/L(\widehat{\theta}_{10}, \widehat{\theta}_{30})$ gives the desired result.

(b) Show that the likelihood ratio test for testing $H_0 : \theta_3 = \theta_4$, θ_1 and θ_2 unspecified, against $H_1 : \theta_3 \neq \theta_4$, θ_1 and θ_2 unspecified, can be based on the random variable

$$F = \frac{\sum_1^n (X_i - \overline{X})^2/(n-1)}{\sum_1^m (Y_i - \overline{Y})^2/(m-1)}$$

Solution.

Note that H_0 is different from that in part (a). Under Ω, the mles are the same as in part (a), while under H_0,

$$\widehat{\theta}_{10} = \overline{X},\ \widehat{\theta}_{20} = \overline{Y},$$
$$\widehat{\theta}_{30} = \widehat{\theta}_{40} = \frac{1}{n+m}\left[\sum_1^n (X_i - \overline{X})^2 + \sum_1^m (Y_i - \overline{Y})^2\right].$$

Hence, the LRT statistic is given by

$$\Lambda = \frac{\left[\sum_1^n (x_i - \overline{x})^2/n\right]^{n/2} \left[\sum_1^m (y_i - \overline{y})^2/m\right]^{m/2}}{\left\{\left[\sum_1^n (x_i - \overline{x})^2 + \sum_1^m (y_i - \overline{y})^2\right]/(n+m)\right\}^{(n+m)/2}}$$

Here, let S_x^2 and S_y^2 denote the sample variances. Then the F statistic is $F = S_x^2/S_y^2$ and thus

$$\begin{aligned}
\Lambda &= K\frac{(S_x^2)^{n/2}(S_y^2)^{m/2}}{[(n-1)S_x^2 + (m-1)S_y^2]^{(n+m)/2}} \\
&= K\frac{(S_x^2)^{n/2}(S_y^2)^{m/2}/(S_y^2)^{(n+m)/2}}{[(n-1)S_x^2 + (m-1)S_y^2]^{(n+m)/2}/(S_y^2)^{(n+m)/2}} \\
&= K\frac{(S_x^2/S_y^2)^{n/2}}{[(n-1)S_x^2/S_y^2 + (m-1)]^{(n+m)/2}} \\
&= K\frac{F^{n/2}}{[(n-1)F + (m-1)]^{(n+m)/2}},
\end{aligned}$$

which is a function of random variable $F \sim F_{n-1,m-1}$,

6.5.6. Let $X_1, X_2, ..., X_n$ and $Y_1, Y_2, ..., Y_m$ be independent random samples from the two normal distributions $N(0, \theta_1)$ and $N(0, \theta_2)$.

(a) Find the likelihood ratio Λ for testing the composite hypothesis $H_0 : \theta_1 = \theta_2$ against the composite alternative $H_1 : \theta_1 \neq \theta_2$.

Solution.

On the whole space Ω, by the previous exercises,

$$\widehat{\theta}_1 = \frac{1}{n}\sum_1^n X_i^2, \ \widehat{\theta}_2 = \frac{1}{m}\sum_1^n Y_i^2.$$

Under H_0, on the other hand, solving $\ell'(\theta_1) = 0$ gets

$$\widehat{\theta}_1 = \widehat{\theta}_2 = \frac{1}{n+m}\left[\sum_1^n X_i^2 + \sum_1^m Y_i^2\right].$$

Hence, the LRT statistic is

$$\Lambda = \frac{L(\widehat{\theta}_1, \widehat{\theta}_1)}{L(\widehat{\theta}_1)} = \frac{\left\{\left[\sum_1^n x_i^2 + \sum_1^m y_i^2\right]/(n+m)\right\}^{(n+m)/2}}{\left[\sum_1^n x_i^2/n\right]^{n/2}\left[\sum_1^m y_i^2/m\right]^{m/2}}$$

(b) This Λ is a function of what F-statistic that would actually be used in this test?

Solution.

Similarly to part (b) in Exercise 6.5.5, under $H_0 : \theta_1 = \theta_2$,

$$F = \frac{(\sum_1^n X_i^2/\theta_1)/n}{(\sum_1^m Y_i^2/\theta_1)/m} = \frac{\sum_1^n X_i^2/n}{\sum_1^m Y_i^2/m} \sim F_{n,m}$$

can be used in Λ as a random variable.

6.5.7. Let X and Y be two independent random variables with respective pdfs

$$f(x;\theta_i) = \begin{cases} \left(\frac{1}{\theta_i}\right) e^{-x/\theta_i} & 0 < x < \infty, \ 0 < \theta_i < \infty \\ 0 & \text{elsewhere.} \end{cases}$$

for $i = 1, 2$. To test $H_0 : \theta_1 = \theta_2$ against $H_1 : \theta_1 = \theta_2$, two independent samples of sizes n_1 and n_2, respectively, were taken from these distributions. Find the likelihood ratio Λ and show that Λ can be written as a function of a statistic having an F-distribution, under H_0.

Solution.

Given that

$$f(x, \theta_1) = \left(\frac{1}{\theta_1}\right) e^{-x/\theta_1}, \ 0 < x < \infty,$$
$$f(y, \theta_2) = \left(\frac{1}{\theta_2}\right) e^{-y/\theta_2}, \ 0 < y < \infty.$$

Under Ω, we obtain the mles (omitted the proof)

$$\widehat{\theta}_1 = \overline{X}, \quad \widehat{\theta}_2 = \overline{Y}.$$

While, under H_0, solving $\ell'(\theta_1) = 0$ obtains

$$\widehat{\theta}_{10} = \widehat{\theta}_{20} = \frac{n_1\overline{X} + n_2\overline{Y}}{n_1 + n_2}.$$

Hence, the LRT statistic is

$$\Lambda = \frac{L(\widehat{\theta}_1, \widehat{\theta}_2)}{L(\widehat{\theta}_{10})} = \cdots = K\frac{(n_1\overline{x} + n_2\overline{y})^{n_1+n_2}}{\overline{x}^{n_1}\overline{y}^{n_2}} = K\frac{(n_1(\overline{x}/\overline{y}) + n_2)^{n_1+n_2}}{(\overline{x}/\overline{y})^{n_1}},$$

which is a function of a random variable $\overline{X}/\overline{Y}$.

Under H_0, $X, Y \sim \Gamma(1, \theta_1)$,

$$\frac{2\sum_1^{n_1} X_k}{\theta_1} \sim \chi^2(2n_1) \quad \frac{2\sum_1^{n_2} Y_k}{\theta_1} \sim \chi^2(2n_2).$$

Therefore,

$$\frac{\overline{X}}{\overline{Y}} = \frac{(2\sum_1^{n_1} X_k/\theta_1)/(2n_1)}{(2\sum_1^{n_1} Y_k/\theta_1)/(2n_2)} \sim F_{2n_1,2n_2},$$

which is the desired result.