STA 260: Statistics and Probability II

Al Nosedal.
University of Toronto.

Winter 2017

"If you can't explain it simply, you don't understand it well enough"

Albert Einstein.

Definition 9.1

Given two unbiased estimators $\hat{\theta}_1$ and $\hat{\theta}_2$ of a parameter θ, with variances $V(\hat{\theta}_1)$ and $V(\hat{\theta}_2)$, respectively, then the **efficiency** of $\hat{\theta}_1$ relative to $\hat{\theta}_2$, denoted eff($\hat{\theta}_1$, $\hat{\theta}_2$), is defined to be the ratio

$$\text{eff}(\hat{\theta}_1, \hat{\theta}_2) = \frac{V(\hat{\theta}_2)}{V(\hat{\theta}_1)}$$

Exercise 9.1

In Exercise 8.8, we considered a random sample of size 3 from an exponential distribution with density function given by

$$f(y) = \begin{cases} (1/\theta)e^{-y/\theta} & y > 0 \\ 0 & elsewhere \end{cases}$$

and determined that $\hat{\theta}_1 = Y_1$, $\hat{\theta}_2 = (Y_1 + Y_2)/2$, $\hat{\theta}_3 = (Y_1 + 2Y_2)/3$, and $\hat{\theta}_5 = \bar{Y}$ are all unbiased estimators for θ. Find the efficiency of $\hat{\theta}_1$ relative to $\hat{\theta}_5$, of $\hat{\theta}_2$ relative to $\hat{\theta}_5$, and of $\hat{\theta}_3$ relative to $\hat{\theta}_5$

Solution

$V(\hat{\theta}_1) = V(Y_1) = \theta^2$ (From Table).

$V(\hat{\theta}_2) = V\left(\frac{Y_1+Y_2}{2}\right) = \frac{2\theta^2}{4} = \frac{\theta^2}{2}$

$V(\hat{\theta}_3) = V\left(\frac{Y_1+2Y_2}{3}\right) = \frac{5\theta^2}{9}$

$V(\hat{\theta}_5) = V\left(\bar{Y}\right) = \frac{\theta^2}{3}$

Solution

$$\text{eff}(\hat{\theta}_1, \hat{\theta}_5) = \frac{V(\hat{\theta}_5)}{V(\hat{\theta}_1)} = \frac{\frac{\theta^2}{3}}{\theta^2} = \frac{1}{3}$$

$$\text{eff}(\hat{\theta}_2, \hat{\theta}_5) = \frac{V(\hat{\theta}_5)}{V(\hat{\theta}_2)} = \frac{\frac{\theta^2}{3}}{\frac{\theta^2}{2}} = \frac{2}{3}$$

$$\text{eff}(\hat{\theta}_3, \hat{\theta}_5) = \frac{V(\hat{\theta}_5)}{V(\hat{\theta}_3)} = \frac{\frac{\theta^2}{3}}{\frac{5\theta^2}{9}} = \frac{3}{5}$$

Exercise 9.3

Let $Y_1, Y_2, ..., Y_n$ denote a random sample from the uniform distribution on the interval $(\theta, \theta + 1)$. Let $\hat{\theta}_1 = \bar{Y} - \frac{1}{2}$ and $\hat{\theta}_2 = Y_{(n)} - \frac{n}{n+1}$.

a. Show that both $\hat{\theta}_1$ and $\hat{\theta}_2$ are unbiased estimators of θ.

b. Find the efficiency of $\hat{\theta}_1$ relative to $\hat{\theta}_2$.

Solution

a. $E(\hat{\theta}_1) = E(\bar{Y} - \frac{1}{2}) = E(\bar{Y}) - E(\frac{1}{2}) = E(\frac{Y_1+Y_2+...+Y_n}{n}) - \frac{1}{2} = \frac{2\theta+1}{2} - \frac{1}{2} = \theta.$

Since Y_i has a Uniform distribution on the interval $(\theta, \theta+1)$, $V(Y_i) = \frac{1}{12}$ (check Table).

$V(\hat{\theta}_1) = V(\bar{Y} - \frac{1}{2}) = V(\bar{Y}) = V(\frac{Y_1+Y_2+...+Y_n}{n}) = \frac{1}{12n}$

Solution

Let $W = Y_{(n)} = max\{Y_1, ..., Y_n\}$.
$F_W(w) = P[W \leq w] = [F_Y(w)]^n$
$f_W(w) = \frac{d}{dw} F_W(w) = n[F_Y(w)]^{n-1} f_Y(w)$
In our case, $f_W(w) = n[w - \theta]^{n-1}$, $\theta < w < \theta + 1$.
Now that we have the pdf of W, we can find its expected value and variance.

Solution

$$E(W) = \int_{\theta}^{\theta+1} nw[w-\theta]^{n-1}dw \text{ (integrating by parts)}$$
$$E(W) = w[w-\theta]^n|_{\theta}^{\theta+1} - \int_{\theta}^{\theta+1}[w-\theta]^n dw$$
$$E(W) = (\theta+1) - \frac{[w-\theta]^{n+1}}{n+1}|_{\theta}^{\theta+1} = (\theta+1) - \frac{1}{n+1}$$
$$E(W) = \theta + \frac{n}{n+1}$$

Solution

$$E(W^2) = \int_{\theta}^{\theta+1} nw^2[w-\theta]^{n-1}dw \text{ (integrating by parts)}$$
$$E(W^2) = w^2[w-\theta]^n\Big|_{\theta}^{\theta+1} - \int_{\theta}^{\theta+1} 2w[w-\theta]^n dw$$
$$E(W^2) = (\theta+1)^2 - \frac{2}{n+1}\int_{\theta}^{\theta+1}(n+1)w[w-\theta]^n dw$$
$$E(W^2) = (\theta+1)^2 - \frac{2}{n+1}\left(\theta + \frac{n+1}{n+2}\right)$$
$$E(W^2) = (\theta+1)^2 - \frac{2\theta}{n+1} - \frac{2}{n+2}$$

Solution

$V(W) = E(W^2) - [E(W)]^2$

$V(W) = \theta^2 + 2\theta + 1 - \frac{2\theta}{n+1} - \frac{2}{n+2} - \left[\theta + \frac{n}{n+1}\right]^2$

(after doing a bit of algebra . . .)

$V(W) = \frac{n}{(n+2)(n+1)^2}$

Solution

$$E(\hat{\theta}_2) = E(W - \tfrac{n}{n+1}) = E(W) - E(\tfrac{n}{n+1}) = \theta + \tfrac{n}{n+1} - \tfrac{n}{n+1} = \theta.$$

$$V(\hat{\theta}_2) = V(W - \tfrac{n}{n+1}) = V(W) = \tfrac{n}{(n+2)(n+1)^2}$$

Solution

$$\text{eff}(\hat{\theta}_1, \hat{\theta}_2) = \frac{V(\hat{\theta}_2)}{V(\hat{\theta}_1)} = \frac{\frac{n}{(n+2)(n+1)^2}}{\frac{1}{12n}} = \frac{12n^2}{(n+2)(n+1)^2}$$

Definition 9.2

The estimator $\hat{\theta}_n$ is said to be **consistent estimator** of θ if, for any positive number ϵ,

$$\lim_{n\to\infty} P(|\hat{\theta}_n - \theta| \leq \epsilon) = 1$$

or, equivalently,

$$\lim_{n\to\infty} P(|\hat{\theta}_n - \theta| > \epsilon) = 0.$$

Theorem 9.1

An unbiased estimator $\hat{\theta}_n$ for θ is a consistent estimator of θ if

$$\lim_{n\to\infty} V(\hat{\theta}_n) = 0.$$

Exercise 9.15

Refer to Exercise 9.3. Show that both $\hat{\theta}_1$ and $\hat{\theta}_2$ are consistent estimators for θ.

Solution

Recall that we have already shown that $\hat{\theta}_1$ and $\hat{\theta}_2$ are unbiased estimator of θ. Thus, if we show that $\lim_{n\to\infty} V(\hat{\theta}_1) = 0$ and $\lim_{n\to\infty} V(\hat{\theta}_2) = 0$, we are done.
Clearly,

$$0 \leq V(\hat{\theta}_1) = \frac{1}{12n}$$

which implies that

$$\lim_{n\to\infty} V(\hat{\theta}_1) = 0$$

Solution

Clearly,

$$0 \leq V(\hat{\theta}_2) = \frac{n}{(n+2)(n+1)^2} = \frac{n}{(n+2)} \frac{1}{(n+1)^2} \leq \frac{(n+2)}{(n+2)} \frac{1}{(n+1)^2}$$

which implies that

$$\lim_{n \to \infty} V(\hat{\theta}_2) \leq \lim_{n \to \infty} \frac{1}{(n+1)^2} = 0$$

Therefore, $\hat{\theta}_1$ and $\hat{\theta}_2$ are consistent estimators for θ.

Definition 9.3

Let $Y_1,\ Y_2, ...,\ Y_n$ denote a random sample from a probability distribution with unknown parameter θ. Then the statistic $U = g(Y_1, Y_2, ..., Y_n)$ is said to be **sufficient** for θ if the conditional distribution of $Y_1, Y_2, ..., Y_n$, given U, does not depend on θ.

Example

Let X_1, X_2, X_3 be a sample of size 3 from the Bernoulli distribution. Consider $U = g(X_1, X_2, X_3) = X_1 + X_2 + X_3$. We will show that $g(X_1, X_2, X_3)$ is sufficient.

Solution

	Values of U	$f_{X_1,X_2,X_3\|U}$
(0,0,0)	0	1
(0,0,1)	1	1/3
(0,1,0)	1	1/3
(1,0,0)	1	1/3
(0,1,1)	2	1/3
(1,0,1)	2	1/3
(1,1,0)	2	1/3
(1,1,1)	3	1

Example

The conditional densities given in the last column are routinely calculated. For instance,

$$f_{X_1,X_2,X_3|U=1}(0,1,0|1) = P[X_1 = 0, X_2 = 1, X_3 = 0|U = 1]$$
$$= \frac{P[X_1{=}0 \ \ and \ \ X_2{=}1 \ \ and \ \ X_3{=}0 \ \ and \ \ U{=}1]}{P[U{=}1]}$$
$$= \frac{(1-p)(p)(1-p)}{\binom{3}{1}p(1-p)^2} = \frac{1}{3}$$

Definition 9.4

Let $y_1, y_2, ..., y_n$ be sample observations taken on corresponding random variables Y_1, Y_2, ..., Y_n whose distribution depends on a parameter θ. Then, if Y_1, Y_2, ..., Y_n are discrete random variables, the **likelihood of the sample**, $L(y_1, y_2, ..., y_n|\theta)$, is defined to be the joint probability of $y_1, y_2, ..., y_n$. If Y_1, Y_2, ..., Y_n are continuous random variables, the likelihood $L(y_1, y_2, ..., y_n|\theta)$, is defined to be the joint density of $y_1, y_2, ..., y_n$.

Theorem 9.4

Let U be a statistic based on the random sample $Y_1, Y_2, ..., Y_n$. Then U is a **sufficient statistic** for the estimation of a parameter θ if and only if the likelihood $L(\theta) = L(y_1, y_2, ..., y_n|\theta)$ can be factored into two nonnegative functions,

$$L(y_1, y_2, ..., y_n|\theta) = g(u, \theta)h(y_1, y_2, ..., y_n)$$

where $g(u, \theta)$ is a function **only** of u and θ and $h(y_1, y_2, ..., y_n)$ is **not** a function of θ.

Exercise 9.37

Let $X_1, X_2, ..., X_n$ denote n independent and identically distributed Bernoulli random variables such that

$$P(X_i = 1) = \theta \quad \text{and} \quad P(X_i = 0) = 1 - \theta,$$

for each $i = 1, 2, ..., n$. Show that $\sum_{i=1}^{n} X_i$ is sufficient for θ by using the factorization criterion given in Theorem 9.4.

Solution

$$L(x_1, x_2, ..., x_n|\theta) = P(x_1|\theta)P(x_2|\theta)...P(x_n|\theta)$$
$$= \theta^{x_1}(1-\theta)^{1-x_1}\theta^{x_2}(1-\theta)^{1-x_2}...\theta^{x_n}(1-\theta)^{1-x_n}$$
$$= \theta^{\sum_{i=1}^n x_i}(1-\theta)^{n-\sum_{i=1}^n x_i}$$

By Theorem 9.4, $\sum_{i=1}^n x_i$ is sufficient for θ with

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

and

$$h(x_1, x_2, ..., x_n) = 1$$

.

Indicator Function

For $a < b$,

$$I_{(a,b)}(y) = \begin{cases} 1 & \text{if } a < y < b \\ 0 & \text{otherwise.} \end{cases}$$

Exercise 9.49

Let $Y_1, Y_2, ..., Y_n$ denote a random sample from the Uniform distribution over the interval $(0, \theta)$. Show that $Y_{(n)} = max(Y_1, Y_2, ..., Y_n)$ is sufficient for θ.

Solution

$$\begin{aligned} L(y_1, y_2, ..., y_n|\theta) &= f(y_1|\theta)f(y_2|\theta)...f(y_n|\theta) \\ &= \tfrac{1}{\theta} I_{(0,\theta)}(y_1)\tfrac{1}{\theta} I_{(0,\theta)}(y_2)...\tfrac{1}{\theta} I_{(0,\theta)}(y_n) \\ &= \tfrac{1}{\theta^n} I_{(0,\theta)}(y_1) I_{(0,\theta)}(y_2)... I_{(0,\theta)}(y_n) \\ &= \tfrac{1}{\theta^n} I_{(0,\theta)}(y_{(n)}) \end{aligned}$$

Therefore, Theorem 9.4 is satisfied with

$$g(y_{(n)}, \theta) = \frac{1}{\theta^n} I_{(0,\theta)}(y_{(n)})$$

and

$$h(y_1, y_2, ..., y_n) = 1$$

.

Exercise 9.51

Let $Y_1, Y_2, ..., Y_n$ denote a random sample from the probability density function

$$f(y|\theta) = \begin{cases} e^{-(y-\theta)} & y \geq \theta \\ 0 & elsewhere \end{cases}$$

Show that $Y_{(1)} = min(Y_1, Y_2, ..., Y_n)$ is sufficient for θ.

Solution

$$\begin{aligned}
L(y_1, y_2, ..., y_n|\theta) &= f(y_1|\theta)f(y_2|\theta)...f(y_n|\theta) \\
&= e^{-(y_1-\theta)} I_{[\theta,\infty)}(y_1) e^{-(y_2-\theta)} I_{[\theta,\infty)}(y_2)...e^{-(y_n-\theta)} I_{[\theta,\infty)}(y_n) \\
&= e^{n\theta} e^{-\sum_{i=1}^{n} y_i} I_{[\theta,\infty)}(y_1) I_{[\theta,\infty)}(y_2)...I_{[\theta,\infty)}(y_n) \\
&= e^{n\theta} e^{-\sum_{i=1}^{n} y_i} I_{[\theta,\infty)}(y_{(1)}) \\
&= e^{n\theta} I_{[\theta,\infty)}(y_{(1)}) e^{-\sum_{i=1}^{n} y_i}
\end{aligned}$$

Solution

Therefore, Theorem 9.4 is satisfied with

$$g(y_{(1)}, \theta) = e^{n\theta} I_{[\theta,\infty)}(y_{(1)})$$

and

$$h(y_1, y_2, ..., y_n) = e^{-\sum_{i=1}^{n} y_i}$$

and $Y_{(1)}$ is sufficient for θ.

Theorem 9.5

The Rao-Blackwell Theorem. Let $\hat{\theta}$ be an unbiased estimator for θ such that $V(\hat{\theta}) < \infty$. If U is a sufficient statistic for θ, define $\hat{\theta}^* = E(\hat{\theta}|U)$. Then, for all θ,

$$E(\hat{\theta}^*) = \theta \quad \text{and} \quad V(\hat{\theta}^*) \leq V(\hat{\theta}).$$

Proof

Check page 465.
(It is almost identical to what we did in class, Remember?)

Exercise 9.61

Refer to Exercise 9.49. Use $Y_{(n)}$ to find an MVUE of θ.

Exercise 9.62

Refer to Exercise 9.51. Find a function of $Y_{(1)}$ that is an MVUE for θ.

Solution

Please, see review 2.

The Method of Moments

The method of moments is a very simple procedure for finding an estimator for one or more population parameters. Recall that the kth moment of random variable, taken about the origin, is

$$\mu_k^{'} = E(Y^k).$$

The corresponding kth sample moment is the average

$$m_k^{'} = \frac{1}{n}\sum_{i=1}^{n} Y_i^k.$$

The Method of Moments

Choose as estimates those values of the parameters that are solutions of the equations $\mu_k^{'}$, for $k = 1, 2, ..., t$, where t is the number of parameters to be estimated.

Exercise 9.69

Let $Y_1, Y_2, ..., Y_n$ denote a random sample from the probability density function

$$f(y|\theta) = \begin{cases} (\theta + 1)y^{\theta} & 0 < y < 1; \quad \theta > -1 \\ 0 & \textit{elsewhere} \end{cases}$$

Find an estimator for θ by the method of moments.

Solution

Note that Y is a random variable with a Beta distribution where $\alpha = \theta + 1$ and $\beta = 1$. Therefore,

$$\mu_1^{'} = E(Y) = \frac{\alpha}{\alpha + \beta} = \frac{\theta + 1}{\theta + 2}$$

(we can find this formula on our Table).

Solution

The corresponding first sample moment is

$$m_1^{'} = \frac{1}{n}\sum_{i=1}^{n} Y_i = \bar{Y}.$$

Solution

Equating the corresponding population and sample moment, we obtain

$$\frac{\theta+1}{\theta+2} = \bar{Y}$$

(solving for θ)

$$\hat{\theta}_{MOM} = \frac{2\bar{Y}-1}{1-\bar{Y}} = \frac{1-2\bar{Y}}{\bar{Y}-1}$$

Exercise 9.75

Let $Y_1, Y_2, ..., Y_n$ be a random sample from the probability density function given by

$$f(y|\theta) = \begin{cases} \frac{\Gamma(2\theta)}{[\Gamma(2\theta)]^2} y^{\theta-1}(1-y)^{\theta-1} & 0 < y < 1; \quad \theta > -1 \\ 0 & \textit{elsewhere} \end{cases}$$

Find the method of moments estimator for θ.

Solution

Note that Y is a random variable with a Beta distribution where $\alpha = \theta$ and $\beta = \theta$. Therefore,

$$\mu_1^{'} = E(Y) = \frac{\alpha}{\alpha + \beta} = \frac{\theta}{2\theta} = \frac{1}{2}$$

Solution

The corresponding first sample moment is

$$m_1^{'} = \frac{1}{n}\sum_{i=1}^{n} Y_i = \bar{Y}.$$

Solution

Equating the corresponding population and sample moment, we obtain

$$\frac{1}{2} = \bar{Y}$$

(since we **can't** solve for θ, we have to repeat the process using second moments).

Solution

Recalling that $V(Y) = E(Y^2) - [E(Y)]^2$ and solving for $E(Y^2)$, we have that
(we can easily get $V(Y)$ from our table, Right?)

$$\mu_2' = E(Y^2) = \frac{\theta^2}{(2\theta)^2(2\theta+1)} + \frac{1}{4} = \frac{1}{4(2\theta+1)} + \frac{1}{4}$$

(after a little bit of algebra...)

$$E(Y^2) = \frac{\theta+1}{4\theta+2}.$$

Solution

The corresponding second sample moment is

$$m_2^{'} = \frac{1}{n}\sum_{i=1}^{n} Y_i^2.$$

Solution

Solving for θ

$$\hat{\theta}_{MOM} = \frac{1 - 2m_2^{'}}{4m_2^{'} - 1}$$

Method of Maximum Likelihood

Suppose that the likelihood function depends on k parameters θ_1, θ_2, . . . , θ_k. Choose as estimates those values of the parameters that maximize the likelihood $L(y_1, y_2, ..., y_n|\theta_1, \theta_2, ..., \theta_k)$.

Problem

Given a random sample $Y_1, Y_2, ..., Y_n$ from a population with pdf $f(x|\theta)$, show that maximizing the likelihood function, $L(y_1, y_2, ..., y_n|\theta)$, as a function of θ is equivalent to maximizing $lnL(y_1, y_2, ..., y_n|\theta)$.

Proof

Let $\hat{\theta}_{MLE}$, which implies that

$$L(y_1, y_2, ..., y_n|\theta) \leq L(y_1, y_2, ..., y_n|\hat{\theta}_{MLE}) \text{ for all } \theta.$$

We know that $g(\theta) = ln(\theta)$ is monotonically increasing function of θ, thus for $\theta_1 \leq \theta_2$ we have that $ln(\theta_1) \leq ln(\theta_2)$.
Therefore

$$lnL(y_1, y_2, ..., y_n|\theta) \leq lnL(y_1, y_2, ..., y_n|\hat{\theta}_{MLE}) \text{ for all } \theta.$$

We have shown that $lnL(y_1, y_2, ..., y_n|\theta)$ attains its maximum at $\hat{\theta}_{MLE}$.

Example 9.16

Let $Y_1, Y_2, ..., Y_n$ be a random sample of observations from a uniform distribution with probability density function $f(y_i|\theta) = \frac{1}{\theta}$, for $0 \leq y_i \leq \theta$ and $i = 1, 2, ..., n$. Find the MLE of θ.

MLEs have some additional properties that make this method of estimation particularly attractive. Generally, if θ is the parameter associated with a distribution, we are sometimes interested in estimating some function of θ - say $t(\theta)$ - rather than θ itself. In exercise, 9.94, you will prove that if $t(\theta)$ is a one-to-one function of θ and if $\hat{\theta}$ is the MLE for θ, then the MLE of $t(\theta)$ is given by

$$\hat{t(\theta)} = t(\hat{\theta}).$$

This result, sometimes referred to as the **invariance property** of MLEs, also holds for any function of a parameter of interest (**not just one-to-one functions**).

Exercise 9.81

Suppose that $Y_1, Y_2, ..., Y_n$ denote a random sample from an exponentially distributed population with mean θ. Find the MLE of the population variance θ^2.

Exercise 9.85

Let $Y_1, Y_2, ..., Y_n$ be a random sample from the probability density function given by

$$f(y|\alpha,\theta) = \begin{cases} \frac{1}{\Gamma(\alpha)\theta^{\alpha}} y^{\alpha-1} e^{-y/\theta} & y > 0, \\ 0 & \text{elsewhere}, \end{cases}$$

where $\alpha > 0$ is known.

Exercise 9.85

a. Find the MLE $\hat{\theta}$ of θ.
b. Find the expected value and variance of $\hat{\theta}_{MLE}$.
c. Show that $\hat{\theta}_{MLE}$ is consistent for θ.
d. What is the best sufficient statistic for θ in this problem?

Solution a)

$$L(\theta) = \frac{1}{[\Gamma(\alpha)\theta^{\alpha}]^n} \Pi_{i=1}^n y_i^{\alpha-1} e^{-\sum_{i=1}^n y_i/\theta}$$

Solution a)

$$lnL(\theta) = (\alpha - 1)\sum_{i=1}^{n} ln(y_i) - \frac{\sum_{i=1}^{n} y_i}{\theta} - nln\Gamma(\alpha) - n\alpha ln(\theta)$$

Solution a)

$$\frac{d lnL(\theta)}{d\theta} = \frac{\sum_{i=1}^{n} y_i - n\alpha\theta}{\theta^2}$$

$$\hat{\theta}_{MLE} = \frac{\bar{y}}{\alpha}$$

Solution a)

(Let us check that we actually have a maximum...)

$$\frac{d^2 lnL(\theta)}{d\theta^2} = \frac{-2\sum_{i=1}^{n} y_i + n\alpha\theta}{\theta^3}$$

$$\frac{d^2 lnL(\hat{\theta}_{MLE})}{d\theta^2} = \frac{-\alpha^3 n}{\bar{y}^2} < 0$$

Solution b)

$$E(\hat{\theta}_{MLE}) = E\left(\frac{\bar{y}}{\alpha}\right) = \theta.$$

$$V(\hat{\theta}_{MLE}) = V\left(\frac{\bar{y}}{\alpha}\right) = \frac{\theta^2}{\alpha n}.$$

Solution c)

Since $\hat{\theta}$ is unbiased, we only need to show that

$$lim_{n\to\infty} V(\hat{\theta}_{MLE}) = lim_{n\to\infty} \frac{\theta^2}{\alpha n} = 0.$$

Solution d)

By the Factorization Theorem,

$$g(u,\theta) = g(\sum_{i=1}^{n} yi, \theta) = \frac{e^{-\sum_{i=1}^{n} y_i/\theta}}{\theta^{\alpha n}}$$

and

$$h(y_1, y_2, ..., y_n) = \frac{\prod_{i=1}^{n} y_i^{\alpha-1}}{[\Gamma(\alpha)]^n}$$

Example 9.15

Let $Y_1, Y_2, ..., Y_n$ be a random sample from a normal distribution with mean μ and variance σ^2. Find the MLEs of μ and σ^2.