class09

February 1, 2021 10:26 AM

Outline

§7.2 Methods of point estimation
Maximum likelihood estimation

§8.1 Basic properties of confidence intervals

Let $X_1, \ldots, X_n$ have joint pmf or pdf

$$f(x_1, \ldots, x_n; \theta_1, \ldots, \theta_m) \quad (7.6)$$

where the parameters $\theta_1, \ldots, \theta_m$ have unknown values. When $x_1, \ldots, x_n$ are the observed sample values and (7.6) is regarded as a function of $\theta_1, \ldots, \theta_m$, it is called the **likelihood function**.

The **maximum likelihood estimates** $\hat{\theta}_1, \ldots, \hat{\theta}_m$ are those values of the θ_i's that maximize the likelihood function, so that

$$f(x_1, \ldots, x_n; \hat{\theta}_1, \ldots, \hat{\theta}_m) \geq f(x_1, \ldots, x_n; \theta_1, \ldots, \theta_m)$$

for all $\theta_1, \ldots, \theta_m$.

When the X_i's are substituted in place of the x_i's, the **maximum likelihood estimators** (mle's) result.

$f(x;\theta)$

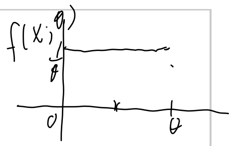
$L(\theta) = f(x_1;\theta) f(x_2;\theta) \cdots f(x_n;\theta)$

$\hat{\theta}$

$\ell(\theta) = \ln(L(\theta))$

$= \ln f(x_1;\theta) + \ln f(x_2;\theta) + \cdots + \ln f(x_n;\theta)$

Example (Uniform).
X is uniform on $(0, \theta)$.



$$f(x) = \begin{cases} \frac{1}{\theta}, 0 < x < \theta. \\ 0, \text{ elsewhere} \end{cases}$$

Method of moments: $E(X) = \frac{\theta}{2} = \bar{X} \Rightarrow \hat{\theta} = 2\bar{X}$.

$\mu = \frac{\theta}{2} = \bar{X}$

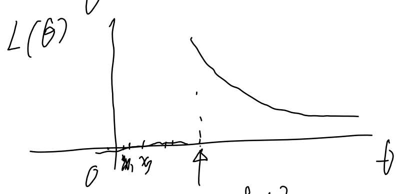
Method of maximum likelihood:
The likelihood function

$$L(\theta) = \frac{1}{\theta^n}, 0 < x_1, \ldots, x_n < \theta$$

$$\hat{\theta} = \max(X_i)$$

$\ell(\theta) = -n \ln(\theta)$, $0 < x_1, \cdots, x_n < \theta$

$\frac{d\ell(\theta)}{d\theta} = -\frac{n}{\theta} = 0$?



$\hat{\theta} = \max\{x_i\}$

θ $\hat{\theta} = \max\{x_i\}$

The invariance principle

Let $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_m$ be the mle's of the parameters $\theta_1, \theta_2, \ldots, \theta_m$.

Then the mle of any function $h(\theta_1, \theta_2, \ldots, \theta_m)$ is $h(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_m)$.

Example 7.21

In the normal case, the mle's of μ and σ^2 are $\hat{\mu} = \bar{X}$ and $\hat{\sigma}^2 = \frac{1}{n}\sum(X_i - \bar{X})^2$. biased

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

To obtain the mle of the function $h(\mu, \sigma^2) = \sqrt{\sigma^2} = \sigma$, substitute the mle's into the function:

$$\hat{\sigma} = \sqrt{\hat{\sigma}^2} = \left[\frac{1}{n}\sum(X_i - \bar{X})^2\right]^{1/2}$$

$X \sim N(\mu, \sigma^2)$

of size $n=10$

Based on a sample, $\bar{x} = 2.5$, $s = 1.5$. Using the invariance principle, find the maximum likelihood estimate of

$\hat{\mu} = \bar{x} = 2.5$, $\hat{\sigma} = \sqrt{\hat{\sigma}^2} = \sqrt{\frac{1}{n}\sum(x_i - \bar{x})^2} = 1.42$

$S = 1.5$

(a) the population coefficient of variation $100\sigma/\mu\%$

$\widehat{h(\mu,\sigma)} = \frac{100\hat{\sigma}}{\hat{\mu}} = 56.9(\%)$ $\hat{\sigma}^2 = \frac{(n-1)S^2}{n} = \frac{1}{n}\sum(x_i - \bar{x})^2$

$\hat{\sigma} = \sqrt{\frac{(n-1)S^2}{n}} = \sqrt{\frac{(10-1)}{10}(1.5)^2} = 1.42$

(b) $P(X < 3.0) = \int_{-\infty}^{3.0} f(x;\mu,\sigma)\,dx = F(3;\mu,\sigma)$

$F(3, \hat{\mu}, \hat{\sigma})$

(c) the 95th percentile for the distribution of X.

$F^{-1}(.95, \hat{\mu}, \hat{\sigma})$

qnorm(.95, 2.5, 1.42)

$= 4.84$.

95th percentile of $Z = 1.645$

$\hat{\mu} + (1.645)(\hat{\sigma}) = 4.84$.

$\widehat{P(X<3.0)} = P\left(\frac{X-\mu}{\sigma} < \frac{3.0-\mu}{\sigma}\right)$

$= P\left(Z < \frac{3.0-\hat{\mu}}{\hat{\sigma}}\right) = P\left(Z < \frac{3.0-2.5}{1.42}\right)$

= pnorm(3, 2.5, 1.42) = .64

Large-sample behavior of the mle

Under very general conditions on the joint distribution of the sample, when the **sample size is large**,

the mle $\hat{\theta}$ is close to θ (consistency),

is approximately unbiased $[E(\hat{\theta}) \approx \theta]$, and

has variance that is nearly as small as can be achieved by any unbiased estimator, i.e., $\hat{\theta}$ is approximately the MVUE of θ.

$\hat{\mu} + (1.645)(\hat{\sigma}) = 4.84.$

Large-sample behavior of the mle

Under very general conditions on the joint distribution of the sample, when the **sample size is large**,

the mle $\hat{\theta}$ is close to θ (consistency),

is approximately unbiased $[E(\hat{\theta}) \approx \theta]$, and

has variance that is nearly as small as can be achieved by any unbiased estimator, i.e., $\hat{\theta}$ is approximately the MVUE of θ.

Maximum likelihood estimation is the most widely used estimation technique among statisticians.

Note that there is no similar result for method of moments estimators. In general, if there is a choice between maximum likelihood and moment estimators, the mle is preferable.

$\hat{\theta} = 2\overline{X}$

$\hat{\theta} = \max\{X_i\}$

Chapter 8. Statistical Intervals Based on a Single Sample

- A point estimate by itself provides no info about the precision and reliability of estimation.
- Standard error is a measure of the precision and reliability.
- Alternatively, we may calculate and report an entire **interval of plausible estimates** – an **interval estimate or confidence interval (CI)**.

μ, σ

$\overline{x} = 7.5$

$SE = \frac{s}{\sqrt{n}}$

Outline

§7.2 Methods of point estimation

Maximum likelihood estimation

§8.1 Basic properties of confidence intervals

Normal population with known σ

Consider a simple situation –

Suppose that the parameter of interest is a population mean μ and that

1. The population distribution is normal
2. the value of population standard deviation σ is known.

$= 1.5$

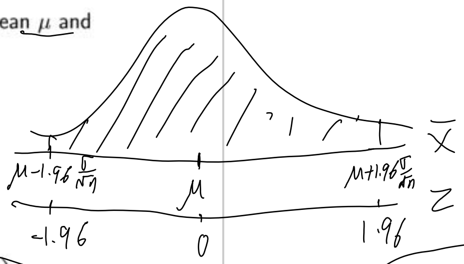
Random interval

- The sample mean $\bar{X}$ is normal with mean μ and standard deviation $\sigma/\sqrt{n}$.
- Standardizing $\bar{X}$:

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

  is standard normal.
- It then follows that

$$P\left(-1.96 < \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < 1.96\right) = .95$$



$\mu > \bar{X} - 1.96\frac{\sigma}{\sqrt{n}}$

$P\left(-1.96\frac{\sigma}{\sqrt{n}} < \bar{X} - \mu < 1.96\frac{\sigma}{\sqrt{n}}\right) = .95$

$\Updownarrow$

$\mu < \bar{X} + 1.96\frac{\sigma}{\sqrt{n}}$

Random interval

- Rearranging terms of the inequalities,

$$P(\bar{X} - 1.96\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + 1.96\frac{\sigma_0}{\sqrt{n}}) = .95$$

- The **random interval**

$$(\bar{X} - 1.96\frac{\sigma}{\sqrt{n}}, \bar{X} + 1.96\frac{\sigma}{\sqrt{n}}) \qquad (8.4)$$

  includes or covers the true value of μ with a probability of 0.95.

Random interval

- Rearranging terms of the inequalities,

$$P(\bar{X} - 1.96\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + 1.96\frac{\sigma_0}{\sqrt{n}}) = .95$$

- The **random interval**

$$(\bar{X} - 1.96\frac{\sigma}{\sqrt{n}}, \bar{X} + 1.96\frac{\sigma}{\sqrt{n}}) \qquad (8.4)$$

includes or covers the true value of μ with a probability of 0.95.

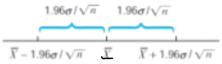


Figure 8.2 The random interval (8.4) centered at $\bar{X}$

95% Confidence interval

If, after observing $X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n$, we compute the observed sample mean $\bar{x}$ and then substitute $\bar{x}$ into the random interval in place of $\bar{X}$, the resulting **fixed** interval is called a **95% confidence interval for** μ. This CI can be expressed as

$$\left(\bar{x} - 1.96\frac{\sigma}{\sqrt{n}}, \bar{x} + 1.96\frac{\sigma}{\sqrt{n}}\right) \text{ is a 95\% CI for } \mu$$

or as

$$\bar{x} - 1.96\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + 1.96\frac{\sigma}{\sqrt{n}} \text{ with 95\% confidence}$$

A concise expression for the interval is $\bar{x} \pm 1.96\sigma/\sqrt{n}$, where $-$ gives the left endpoint (lower limit) and $+$ gives the right endpoint (upper limit).

$\bar{x} = 2.6$ $n = 36$

Example. The **average zinc concentration** recovered from a sample of zinc measurements in 36 different locations is found to be 2.6 grams per milliliter. Find a 95% confidence interval for the **mean zinc concentration in the river**. (μ) Assume that the population distribution is normal and that the population standard deviation is 0.3. $= \sigma$

95% confidence

$\bar{x} = 2.6$, $n = 36$, $\sigma = 0.3$. The resulting interval is

$$\bar{x} \pm 1.96\frac{\sigma}{\sqrt{n}} = 2.6 \pm (1.96)\frac{0.3}{\sqrt{36}} = 2.6 \pm 0.10 = (2.50, 2.70)$$

We are 95% confident that $2.50 < \mu < 2.70$.

Interpreting a confidence interval

The statement

$$P(2.50 < \mu < 2.70) = 0.95$$

is not quite right: (2.50, 2.70) is a fixed interval and μ is a constant; either μ is within the interval or it does not!

Interpretation of the 95% confidence level:
If we could take samples (of size n) over and over again, 95% of these samples would give an interval that includes μ, and 5% of all samples would yield an erroneous interval.

Interpreting a confidence interval

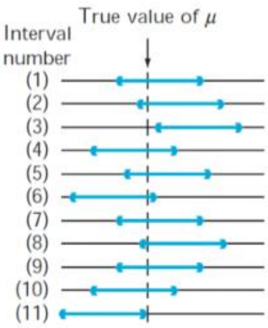


Figure 8.3 Repeated construction of 95% CIs

An R simulation

```
# M: number of samples.
# n: sample size.
M<-20
n<-10
mu<-0; sigma=1

for (m in 1:M){
# Generate a sample of size n from standard normal.
x<-rnorm(n, mean=mu, sd=sigma)
xbar<-mean(x); ME<-1.96*sigma/sqrt(n)
LowerLimit<-xbar-ME; UpperLimit<-xbar+ME
Outside<-ifelse(mu<LowerLimit | mu>UpperLimit, 1, 0)
# 95% CI.
cat("Sample ", m, ": ",
round(c(xbar-ME, xbar+ME), 2), "\t ", Outside, "\n")
}
```