C. Confidence intervals for other distributions.

Use exponential distribution for illustration.

Suppose $X_1, X_2, \ldots, X_n \sim \exp(\lambda) = \text{Gamma}(1, \lambda)$

Let $\mu = E[X_1] = \frac{1}{\lambda}$.

Construct CI. for $\mu$.

(a) Exact CI.

Recall: $\mu \overset{est}{=} \bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$

$n\bar{X} = \sum_{i=1}^{n} X_i \sim \text{Gamma}(n, \lambda)$

So, $\frac{n\bar{X}}{\mu} \sim \text{Gamma}(n, 1)$ -- parameter free

From the probability table or statistical software

to get the 2.5% quantile of Gamma$(n, 1)$, say $C_1$

the 97.5% quantile, say $C_2$
Then, \( P \left\{ c_1 \leq \frac{n \bar{x}}{\mu} \leq c_2 \right\} \approx 0.95 \)

So, \( P \left\{ \frac{n \bar{x}}{c_2} \leq \mu \leq \frac{n \bar{x}}{c_1} \right\} = 0.95 \)

The 95% C.I. of \( \mu \) is \( \left( \frac{n \bar{x}}{c_2}, \frac{n \bar{x}}{c_1} \right) \).

(b) Approximate C.I. (when \( n \) is large)

Recall: \( \mu \) est. \( \hat{\mu} = \bar{x} \),

\( \text{var}(X_i) = \frac{1}{\theta} = \mu^2 \).

By the CLT, we have

\[
\frac{\sqrt{n} (\bar{X} - \mu)}{\mu} = \sqrt{\frac{\bar{X}}{\mu^2}} \rightarrow N(0,1)
\]

Let \( d_1 = 1.96 \) be the 2.5% quantile of \( N(0,1) \)

\( d_2 = 1.96 \) be the 97.5% quantile

So, \( P \left\{ -1.96 \leq \frac{\sqrt{n} (\bar{X} - \mu)}{\mu} \leq 1.96 \right\} = 0.95 \)
So, \( P \left\{ \frac{\bar{x}}{\text{sn}^{1/2} \sqrt{n}} + 1 \leq \mu \leq \frac{\bar{x}}{\text{sn}^{1/2} \sqrt{n}} + 1 \right\} \approx 0.95 \).

The approximate 95\% C.I. is

\[
\left( \frac{\bar{x}}{1.96 / \text{sn}^{1/2} \sqrt{n}} + 1, \frac{\bar{x}}{1.96 / \text{sn}^{1/2} \sqrt{n}} + 1 \right).
\]

**Numerical Example:** \( X_1, \ldots, X_{20} \sim \text{Exp}(\lambda) \).

\( n = 20 \) and \( \bar{x} = 5.70 \).

Find a 95\% C.I. for \( \mu = \frac{1}{\lambda} = E X_1 \).

**Solution:**

(a) For \( n = 20 \),

\[
C_1 = 12.2165,
\]

\[
C_2 = 29.67.
\]

So,

\[
\frac{n \bar{x}}{C_1} = \frac{20 \times 5.70}{12.2165} = 9.33
\]

\[
\frac{n \bar{x}}{C_2} = \frac{20 \times 5.70}{29.67} = 3.84.
\]

\( \Rightarrow \) **Exact 95\% C.I. is** \((3.84, 9.33)\).

(b) \( n = 20 \),

\[
\frac{\bar{x}}{1.96 / \text{sn}^{1/2} \sqrt{n}} + 1 = \frac{5.70}{1.96 / 20 + 1} = 10.47
\]

\[
\frac{\bar{x}}{1.96 / \text{sn}^{1/2} \sqrt{n}} - 1.96 / \text{sn}^{1/2} \sqrt{n} + 1 = \frac{5.70}{1.96 / 20 + 1} = 3.96
\]

\( \Rightarrow \) **Approximate 95\% C.I. is** \((3.96, 10.47)\).
4. Unbiased Estimators

\[ X_1, \ldots, X_n \] random sample.

\[ \hat{\sigma}(x) = \frac{\sum (x_i - \bar{x})^2}{n-1} \] is an estimator of \( \sigma^2 \).

Then, \( \hat{\sigma}(x) \) is said to be an unbiased estimator of \( \sigma^2 \), if

\[
E(\hat{\sigma}(x)) = \frac{\sum (x_i - \bar{x})^2}{n-1} = \frac{3}{2} \sigma^2
\]

Example: \( X_1, \ldots, X_n \sim N(\theta, \sigma^2) \), interested to estimate \( \theta \) and \( \sigma^2 \).

\[ X \] is an unbiased estimator of \( \theta \)

\[ \frac{X_1 + X_2}{2} \] is NOT an unbiased estimator of \( \theta \)

\[
E(\frac{X_1 + X_2}{2}) = \frac{\theta + 2\theta}{2} = \frac{3}{2} \theta = \theta + \theta
\]

\[ S^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2 \] is an unbiased estimator of \( \sigma^2 \).

\[ \hat{\sigma}_{MLE}^2 = \frac{1}{n} \sum (x_i - \bar{x})^2 \] is NOT an unbiased estimator of \( \sigma^2 \).

\[
E(\hat{\sigma}_{MLE}^2) = \frac{n-1}{n} \sigma^2 = \frac{n-1}{n} \sigma^2 + \sigma^2
\]
Unbiasedness V.S. Consistency

(Recall consistency: \( \delta(x) \to 0 \) when \( n \to +\infty \))

\( x_i \) is an unbiased estimator of \( \theta \), but not a consistent estimator.

\( \overline{x} + \frac{1}{n} \) is a consistent estimator of \( \theta \), but a biased estimator of \( \theta \). \( \mathbb{E}(\overline{x} + \frac{1}{n}) = \theta + \frac{1}{n} \neq \theta \).

\( \overline{x} \) is both an unbiased and consistent estimator.
5. Sufficient Statistic (Sections 6.7 & 6.8)

Theoretical value

Loosely speaking:

Sufficient statistic can summarize all relevant information about unknown $\theta$ contained in a sample set (data set).

Formal definition:

Let $X_1, X_2, \ldots, X_n \sim f_0(x)$ and $T = r(X_1, \ldots, X_n)$ is a statistic.

The statistic $T$ is a sufficient statistic if

$$f_0(x_1, x_2, \ldots, x_n \mid T)$$

is independent of $\theta$.

(or $P_\theta(x_1 = x_1, \ldots, x_n = x_n \mid T)$ for discrete distribution)

Example: Poisson $X_1, \ldots, X_n \sim$ Poisson ($\theta$)

Let $T = \sum_{i=1}^{n} X_i$

$$P_\theta(x_1 = x_1, \ldots, x_n = x_n) = \frac{\theta^x}{x_1!} \cdots \frac{\theta^x}{x_n!} = \frac{\theta^x}{x_1! \cdots x_n!} = \frac{\theta^x}{x_1! \cdots x_n!} \frac{x_1! \cdots x_n!}{\lambda^n} (x_1! \cdots x_n!)$$
Note, $T \sim \text{Poisson} (\theta \phi)$

$$P_\theta (T = t) = \frac{e^{-\theta \phi} (\theta \phi)^t}{t!}$$

So, $P_\theta (x_1 = x_i, \ldots, x_u = x_i \mid T = t) = \frac{P_\theta (x_1 = x_i, \ldots, x_u = x_i, T = t)}{P_\theta (T = t)}$

$$= \frac{\frac{e^{-\theta \phi} \theta \phi^{\frac{u}{i}} x_i^1}{i=1} x_i!}{\frac{e^{u \theta \phi} (\theta \phi)^t}{t!}} = \frac{t!}{(\prod_{i=1}^{u} x_i)!} \frac{t!}{\text{if } t \geq \frac{u}{i} x_i}
$$

$$0 \quad \text{if } t < \frac{u}{i} x_i$$

free of $\theta$!!

$\Rightarrow T = \frac{u}{i} x_i$ is a sufficient statistic of $\theta$.

Q: how about $3T + 4$? Yes, it is a sufficient statistic too.

""" $T' = x_1 + 2x_2$? No, since for example,$$P (x_1 = \theta, x_2 = 1 \mid x_1 + 2x_2 = 2) = \ldots$$

$= \frac{1}{1 + \theta}$, depend on $\theta$"""
Theorem (The factorization criterion)

Let \( x_1, \ldots, x_n \sim f_\theta(x) \).

\( T = r(x_1, \ldots, x_n) \) is a sufficient statistic for \( \theta \)

if and only if \( f_\theta(x_1, \ldots, x_n) \) can be factorized as follows.

\[
f_\theta(x_1, \ldots, x_n) = u(x_1, \ldots, x_n) \cdot V(r(x_1, \ldots, x_n), \theta) \\
\text{free of } \theta \in \text{ involves only } T \text{ and } \theta
\]

for any \( \theta \) and \( x_1, \ldots, x_n \).

(Note, the function \( u(\cdot) \) and \( V(\cdot, \cdot) \) is not uniquely defined.)

Proof: See textbook page 373

(key: \( f_\theta(T=t) \sim V(t, \theta) \) )
Example (Poisson) \( X_1, \ldots, X_n \sim \text{Poisson}(\theta) \), \( T = \sum_{i=1}^{n} X_i \).

\[
P_0(X_1=x_1, \ldots, X_n=x_n) = \frac{e^{-\theta} \theta^{\sum_{i=1}^{n} x_i}}{(\prod_{i=1}^{n} x_i!)} = \frac{1}{\prod_{i=1}^{n} (x_i!)} \cdot \frac{e^{-\theta} \theta^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i} = U(x_1, \ldots, x_n) \cdot V(T, \theta).
\]

So, \( T \) is sufficient.

This concept extends to no-scalar parameters.

Example: \( N(\mu, \sigma^2) \) both \( \mu \) and \( \sigma^2 \) unknown.

\[
\sum_{i=1}^{n} X_i \text{ and } \frac{\sum_{i=1}^{n} X_i^2}{n} \text{ are jointly sufficient for } \mu \text{ and } \sigma.
\]

\( \bar{X} \text{ and } \sigma^2 \) are sufficient.

Minimal Sufficient Statistic

\( (X_1, \ldots, X_n) \) is sufficient

\( T \) is sufficient \( \Rightarrow \) \( (T, T') \) is sufficient.

for any statistic \( T' \).

many, many choices.
We prefer to deal with $T$ rather than $(T, T')$ or $(X_1, \ldots, X_n)$.

If there is no further reduction from $T$, then $T$ is minimally sufficient.

Formal definition:

A statistic $T$ is minimal sufficient statistic if $T$ is a sufficient statistic and is a function of every other sufficient statistic.

Poisson example: $T = \bar{X}$, or $\frac{1}{\lambda}X_i$ or $2\bar{X}$ etc.