Math 502AB - Lecture 26

Dr. Jamshidian

November 29, 2017

1 Lecture - Part 1

1.1 Chapter 9: Interval Estimation

Example

1. Suppose that you have $X_1, X_2, X_3, X_4 \sim \textit{exponential}(\lambda)$. You know that the **MLE** of $\hat{\lambda} = \overline{X}$. Now consider the interval estimate:

$$\left[\frac{\overline{X}}{4}, 4\overline{X}\right]$$

Let's talk about some properties of this interval. For example, consider:

$$\begin{aligned}
P\left(\lambda \in \left[\frac{\overline{X}}{4}, 4\overline{X}\right]\right) &= P\left(\frac{\overline{X}}{4} \leq \lambda \leq 4\overline{X}\right) \\
&= P\left(\frac{1}{4\overline{X}} \leq \frac{1}{\lambda} \leq \frac{4}{\overline{X}}\right) \\
&= P\left(2 \leq \frac{8\overline{X}}{\lambda} \leq 32\right) \\
&= P(2 \leq \chi^2_{(8)} \leq 32) = 0.98
\end{aligned}$$

Thus, the interval $\left[\frac{\overline{X}}{4}, 4\overline{X}\right]$ is called a 98% Confidence interval for λ.

1.1.1 Confidence Interval

Definition:

Let $X = (X_1, ..., X_n)$ be a random vector from a distribution with parameter θ. Then, the interval $[L(X), U(X)]$ is an interval estimate of θ where $L(X)$ and $U(X)$ are functions such that $L(X) \leq U(X)$.If $P\left[\theta \in (L(X), U(X))\right] = 1 - \alpha$, then the interval is called a $100(1-\alpha)\%$ confidence interval for θ.

In general:

$$P(\theta \in [L(X), U(X)])$$

is called the **coverage probability** of the interval.

1.1.2 Pivotal Quantity

Definition:

If $Q = q(X_1, ..., X_n, \theta)$ is a random variable that is a function only of $X_1, ..., X_n$ and θ, then q is called a **pivotal quantity** if its distribution does not depend on θ. Such pivotal quantities can be used to obtain confidence intervals for θ.

1.1.3 Example:

1. Let $X_1, ..., X_n \sim exponential(\theta)$. Then the quantity:

$$\frac{2n\overline{X}}{\theta} \sim \chi^2_{(2n)}$$

To obtain a $100(1-\alpha)\%$ confidence interval, we compute:

$$\begin{aligned}
1-\alpha &= P\left[c_1 \leq \frac{2n\overline{X}}{\theta} \leq c_2\right] \\
&= P\left[\frac{1}{c_2} \leq \frac{\theta}{2n\overline{X}} \leq \frac{1}{c_1}\right] \\
&= P\left[\frac{2n\overline{X}}{c_2} \leq \theta \leq \frac{2n\overline{X}}{c_1}\right]
\end{aligned}$$

Theorem:

Let $X_1, ..., X_n$ be a random sample from a distribution with location-scale parameters:

$$f(x|\theta_1, \theta_2) = \frac{1}{\theta_2} f\left(\frac{x-\theta_1}{\theta_2}\right)$$

If **MLEs** $\hat{\theta}_1$ and $\hat{\theta}_2$ exist, then:

$$\frac{\hat{\theta}_1 - \theta_1}{\hat{\theta}_2} \text{ and } \frac{\hat{\theta}_2}{\theta_2} \text{ are pivotal quantities for } \theta_1 \text{ and } \theta_2.$$

1.1.4 Example

1. Let $X_1, ..., X_n \sim N(\mu, \sigma^2)$ with both μ and σ^2 unknown. We have:

$$\hat{\mu} = \overline{X} \quad \text{and} \quad \hat{\sigma} = \sqrt{\frac{n-1}{n}S^2}$$

According to the theorem,

$$Q = \frac{\overline{X} - \mu}{\sqrt{\frac{n-1}{n}S^2}} \quad \text{is a pivotal quantity}$$

We then notice that,

$$\sqrt{n-1} \cdot Q = \frac{\sqrt{n}(\overline{X} - \mu)}{S} \sim t^{(n-1)}$$

Now, let's consider $\frac{\hat{\theta}_2}{\theta_2}$. We have:

$$\begin{aligned}
Q &= \left(\frac{\sqrt{\frac{n-1}{n} S^2}}{\sigma} \right) \\
Q^2 &= \frac{(n-1)S^2}{n\sigma^2} \\
nQ^2 &= \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}
\end{aligned}$$

We can construct a **confidence interval for** μ as:

$$\begin{aligned}
1 - \alpha &= P\left(c_1 < \frac{\sqrt{n}(\overline{X} - \mu)}{S} \leq c_2 \right) \\
&= P\left(\overline{X} + c_1 \frac{S}{\sqrt{n}} < \mu < \overline{X} + c_2 \frac{S}{\sqrt{n}} \right)
\end{aligned}$$

And a **confidence interval for** σ^2 as:

$$\begin{aligned}
1 - \alpha &= P\left(c_1 \leq \frac{(n-1)S^2}{\sigma^2} \leq c_2 \right) \\
&= P\left(\frac{1}{c_2} \leq \frac{\sigma^2}{(n-1)S^2} \leq \frac{1}{c_1} \right) \\
&= P\left(\frac{(n-1)S^2}{c_2} \leq \sigma^2 \leq \frac{(n-1)S^2}{c_1} \right)
\end{aligned}$$

1.1.5 Approximate Confidence Intervals

Example

Let $X_1, .., X_n \sim Bernouli(\theta)$ with $\hat{\theta} = \overline{X}$. By the **CLT**, for large n, we have:

$$\begin{aligned}
\overline{X} &\sim N\left(\theta, \frac{\theta(1-\theta)}{n} \right) \\
\frac{\overline{X} - \theta}{\sqrt{\frac{\theta(1-\theta)}{n}}} &\sim N(0,1)
\end{aligned}$$

Suppose we want to obtain θ such that:

$$P\left(-Z_{\alpha/2} \leq \frac{\overline{X}-\theta}{\sqrt{\frac{\theta(1-\theta)}{n}}} \leq Z_{\alpha/2}\right) = 1-\alpha$$

1.1.6 Inverting a Test Statistic (Duality between CI and Tests)

Confidence Intervals are intervals of the form $I = [L(X), U(X)]$ such that $P(\theta \in I) = 1-\alpha$. The acceptance region $\mathcal{A}$ for a test of hypothesis is:

$$X : P(X \in A | H_0 \text{ is true}) = 1-\alpha$$

Example:

1. Let $X_1, ..., X_n \sim N(\mu, \sigma^2)$. Suppose that we are testing:

$$\begin{cases} H_0 : & \mu = \mu_0 \\ H_a : & \mu \neq \mu_0 \end{cases}$$

The **LRT** uses the statistic,

$$T = \frac{\sqrt{n}(\overline{X} - \mu_0}{S}$$

and accepts H_0 at the α-level of significance for all X in

$$\mathcal{A}(\mu) = \left\{X : -t^*_{\alpha/2} \leq \frac{\sqrt{n}(\overline{X}-\mu_0)}{S} \leq t^*_{\alpha/2}\right\}$$

Now, if we find all μ_0's for which the H_0 is accepted, we obtain:

$$c(X) = \left\{\mu_0 : \mu_0 \in \left[\overline{X} - t^*_{\alpha/2}\frac{S}{\sqrt{n}}, \overline{X} + t^*_{\alpha/2}\frac{S}{\sqrt{n}}\right]\right\}$$

2 Lecture - Part 2

2.0.1 Examples, cont'd

1. **Inverting an LRT: Exponential**

Suppose $X_1, ..., X_n \sim exp(\theta)$ and consider the test:

$$\begin{cases} H_0 : & \theta = \theta_0 \\ H_a : & \theta \neq \theta_0 \end{cases}$$

The LRT acceptance region for this test is:

$$\mathcal{A}(\theta_0) = \left\{X : \left(\frac{T}{\theta_0}\right) e^{-T/\theta_0} \geq c\right\}$$

Where $T = \sum x_i$.

To invert the test to get a $100(1-\alpha)\%$ confidence interval, we seek, for a given data X and summary statistic $T(X) = \sum x_i$, a set:

$$C(T(X)) = \left\{\theta : \left(\frac{T(X)}{\theta}\right)^n e^{-T(X)/\theta} \geq c\right\}$$

Consider $g(T) = \left(\frac{T}{\theta}\right)^n e^{-T/\theta}$. Our confidence interval must take the form:

$$C(T) = \{\theta : L(T) \leq \theta \leq U(T)\}$$

Now, $L(T)$ and $U(T)$ must satisfy:

$$g(L(T)) = g(U(T)) \text{ or } \left(\frac{T}{L(T)}\right)^n e^{-T/L(T)} = \left(\frac{T}{U(T)}\right)^n e^{-T/U(T)}$$

Let $a = \frac{T}{L(T)}$ and $b = \frac{T}{U(T)}$ where $a > b$. Then, the interval must satisfy:

$$a^n e^{-a} = b^n e^{-b}$$

Then the interval is:

$$\{\theta : \frac{T}{a} \leq \theta \leq \frac{T}{b}\}$$

Where a and b must satisfy:

$$\begin{aligned} 1-\alpha &= P\left[\frac{T}{a} < \theta < \frac{T}{b}\right] = P\left[\frac{1}{a} < \frac{\theta}{T} < \frac{1}{b}\right] \\ &= P\left[b < \frac{T}{\theta} < a\right] \end{aligned}$$

Remember:

$$\begin{aligned} T = \sum_{i=1}^n X_i &\sim gamma(n, \theta) \\ \frac{T}{\theta} &\sim gamma(n, 1) \end{aligned}$$

Coming back, we must find a and b such that:

$$a^n e^{-a} = b^n e^{-b} \text{ and } \int_b^a \frac{1}{\Gamma(n)} x^{n-1} e^{-x} dx = 1-\alpha$$

We can do this with numerical methods as seen in class...

2.1 Bayesian Interval Estimation

Let $f_{\Theta|X}$ be the posterior distribution. An interval estimate for a parameter θ is the interval $[u(x), v(x)]$ such that

$$P[u(x) \leq \theta \leq v(x)|X = x] = \int_{u(x)}^{v(x)} f_{\Theta|X}(\theta|x)d\theta = 1 - \alpha$$

In Bayesian estimation, these intervals are called **credible intervals**.

2.1.1 Examples

1. Suppose you have $X_1, ..., X_n \sim N(\theta, \sigma^2)$ (where σ^2 is known). Consider a prior with $\theta \sim N(\theta_0, \sigma_0^2)$ (where both θ_0 and σ_0^2 are known).

$$\Theta|X \sim N\left(\frac{\overline{X}\sigma_0^2 + \theta_0\sigma^2/n}{\sigma_0^2 + \frac{\sigma^2}{n}}, \frac{\sigma^2\sigma_0^2}{n\left(\sigma_0^2 + \frac{\sigma^2}{n}\right)}\right)$$

   To get a 95% confidence interval we then have:

$$\frac{\overline{X}\sigma_0^2 + \theta_0\sigma^2/n}{\sigma_0^2 + \frac{\sigma^2}{n}} \pm 1.96\sqrt{\frac{\sigma^2\sigma_0^2}{n\left(\sigma_0^2 + \frac{\sigma^2}{n}\right)}}$$

2. Suppose you have $X_1, ..., X_n \sim Poisson(\theta)$. Consider the prior $\theta \sim gamma(\alpha, \beta)$.

$$\Theta|X \sim gamma\left(y + \alpha, \frac{\beta}{n\beta + 1}\right) \text{ , where } y = \sum x_i$$

   We then have:

$$2\left(\frac{n\beta + 1}{\beta}\right)\Theta|X \sim \chi^2_{(2\sum x_i + \alpha)}$$

   To find a $100(1 - \alpha)\%$ credible interval for θ, we have:

$$\frac{\beta}{2(n\beta + 1)}\left(2\left(\sum x_i + \alpha\right)\right)\chi^2_{1-\alpha/2}, \frac{\beta}{2(n\beta + 1)}\left(2\left(\sum x_i + \alpha\right)\right)\chi^2_{\alpha/2}$$

2.1.2 "Best" (narrowest) Confidence interval

Theorem

Let $f(x)$ be a unimodal *pdf*. If the interval $[a, b]$ satisfies:

1. $$\int_a^b f(x)dx = 1 - \alpha$$

2. $f(a) = f(b)$

3. $a \leq X^* \leq b$, where X^* is **the** mode of $f(x)$

then $[a, b]$ is the shortest amongst all intervals which satisfy

$$\int_a^b f(x)dx = 1 - \alpha$$

2.1.3 Highest posterior Density

Definition:

The **Highest Posterior Density** (HPD) region corresponds to the shortest interval $c(X)$ such that $P(\theta \in C(X)) = 1 - \alpha$