HW9

Sam Olson

Q1

Let $X_1, \dots, X_n$ be iid exponential(θ) and let $\hat{\theta}_n \equiv \bar{X}_n \equiv \sum_{i=1}^n X_i/n$ denote the MLE based on $X_1, \dots, X_n$.

a)

Determine the limiting distribution of $\sqrt{n}(\hat{\theta}_n - \theta)$ as $n \to \infty$.

Answer

As given, $X_1, \dots, X_n$ are iid with $X_i \sim \text{Exponential}(\theta)$.

This is a known distribution, such that:

$\mathbb{E}[X_i] = \theta$

And:

$\text{Var}(X_i) = \theta^2$

By the Central Limit Theorem, we also know:

$$\sqrt{n}(\bar{X}_n - \theta) \xrightarrow{d} N(0, \theta^2)$$

Substituting values, we get our limiting distribution:

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} N(0, \theta^2)$$

b)

Find a variance stabilizing transformation (VST) for $\{\hat{\theta}_n\}$ and use this to determine a large sample confidence interval for θ with approximate confidence coefficient $1 - \alpha$.

Answer

As given, $X_1, \dots, X_n \overset{iid}{\sim} \text{Exp}(\theta)$. Given this distribution, we know it's MLE due to meeting the regularity conditions of the CRLB, such that: $\hat{\theta}_n = \bar{X}_n$.

From part a), we know the limiting distribution is given by:

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} N(0, \theta^2)$$

We arrive at a VST by using the Delta Method.

To that end, define a continuous function $g(\cdot)$:

$$\sqrt{n}(g(\hat{\theta}_n) - g(\theta)) \xrightarrow{d} N\left(0, [g'(\theta)]^2\theta^2\right)$$

Where:

$$[g'(\theta)]^2\theta^2 = 1$$

Isolating the function g', by taking square root, we have:

$$g'(\theta) = \frac{1}{\theta}$$

And integrating to solve for g:

$$g(\theta) = \log\theta + C$$

Where $C = 0$ for our purposes.

Thus, a VST via the Delta Method is:

$$\sqrt{n}(\log\hat{\theta}_n - \log\theta) \xrightarrow{d} N(0, 1)$$

Then, fora large sample confidence interval, we may invert the test to get an approximate $1-\alpha$ confidence interval for $\log(\theta)$:

$$\left(\log(\hat{\theta}_n) \pm \frac{z_{\alpha/2}}{\sqrt{n}}\right)$$

Where $z_{\alpha/2}$ is the $1-\alpha/2$ standard normal quantile.

To then isolate into just an expression of θ then, we have the approximate confidence interval for θ with approximate confidence coefficient $1-\alpha$ is:

$$\left(\hat{\theta}_n \exp\left(-\frac{z_{\alpha/2}}{\sqrt{n}}\right), \hat{\theta}_n \exp\left(\frac{z_{\alpha/2}}{\sqrt{n}}\right)\right)$$

Noting the use of (instead of [given the use of "approximate coverage".

c)

Suppose a random sample $X_1, \dots, X_{100}$ of $n = 100$ observations yields $\bar{x}_n = 1.835464$. Use this information to obtain a large sample confidence interval for θ based on a likelihood ratio statistic, which has approximate confidence coefficient 90%. (Use the chi-squared approximation for this; you should be able to then numerically determine the interval.)

Using this data, compute also a confidence interval with approximate confidence coefficient 90% using the VST approach from part b).

Answer

For testing $H_0 : \theta = \theta_0$, the likelihood ratio statistic satisfies:

$$-2 \log \Lambda(\theta) \xrightarrow{d} \chi^2_1$$

where:

$$\Lambda(\theta) = \frac{L(\theta)}{L(\hat{\theta}_n)}$$

and degrees of freedom 1 (since θ is a single parameter).

Since the data are iid Exponential(θ), the joint likelihood function is:

$$L(\theta) = \prod_{i=1}^{n} \left(\frac{1}{\theta} e^{-x_i/\theta} \right) = \left(\frac{1}{\theta} \right)^n \exp\left(-\frac{1}{\theta} \sum_{i=1}^{n} x_i \right)$$

Taking the log (noting that the logarithm is a monotonic transformation and preserves maxima):

$$\ell(\theta) = \log L(\theta) = -n \log \theta - \frac{1}{\theta} \sum_{i=1}^{n} x_i$$

Since:

$$\sum_{i=1}^{n} x_i = n\bar{x}_n$$

the log-likelihood simplifies to:

$$\ell(\theta) = -n \log \theta - \frac{n\bar{x}_n}{\theta}$$

Thus, the likelihood ratio test statistic simplifies to:

$$-2 \log \Lambda(\theta) = 2n \left[\log\left(\frac{\theta}{\hat{\theta}_n} \right) + \frac{\hat{\theta}_n}{\theta} - 1 \right],$$

Where $\hat{\theta}_n = \bar{X}_n$ is the MLE for θ.

We then have:

$$2n \left[\log\left(\frac{\theta}{\hat{\theta}_n} \right) + \frac{\hat{\theta}_n}{\theta} - 1 \right] \leq \chi^2_{1,0.90}$$

Where $\chi^2_{1,0.90} \approx 2.7055$.

We have everything we then need to construct a confidence interval, using $n = 100$ observations with $\bar{x}_n = 1.835464$.

We have the inequality:

$$2n\left[\log\left(\frac{\theta}{\bar{x}_n}\right)+\frac{\bar{x}_n}{\theta}-1\right]\leq 2.7055.$$

Our goal then is to isolate into θ terms, when possible. To that end:

$$\log\left(\frac{\theta}{1.835464}\right)+\frac{1.835464}{\theta}-1\leq 0.0135275\rightarrow \log\left(\frac{\theta}{1.835464}\right)+\frac{1.835464}{\theta}=1.0135275$$

This is rather tricky to solve for analytically! I will turn to the computer for help.

```
n <- 100
x_bar <- 1.835464
chi_sq_crit <- qchisq(0.90, df = 1)

rhs_value <- chi_sq_crit / (2 * n)

lrt_equation <- function(theta) log(theta/x_bar) + (x_bar/theta) - (1 + rhs_value)

lower_theta <- uniroot(lrt_equation, lower = 0.5, upper = x_bar)$root
upper_theta <- uniroot(lrt_equation, lower = x_bar, upper = 5)$root

c(lower_theta, upper_theta)
```

```
## [1] 1.563927 2.173671
```

Interestingly, this is similar (but different) to the interval via the VST method.

Aside (VST Method) Using the VST approach from part b):

```
x_bar <- 1.835464
n <- 100
z_90 <- qnorm(0.95)

lower_vst <- x_bar * exp(-z_90 / sqrt(n))
upper_vst <- x_bar * exp(z_90 / sqrt(n))

c(lower_vst, upper_vst)
```

```
## [1] 1.557079 2.163620
```

Q3

Suppose $X_1, \dots, X_n$ are a random sample with common cdf given by

$$P(X_1 \leq x|\theta) = \begin{cases} 1 - e^{-(x/\theta)^2} & \text{if } x > 0 \\ 0 & \text{otherwise,} \end{cases} \qquad \theta > 0$$

a)

Use the Mood-Graybill-Boes Method to derive a CI for θ with C.C. $1 - \alpha$ based on the statistic $X_{(1)} = \min_{1 \leq i \leq n} X_i$.

Answer

Since $X_1, \dots, X_n$ are a random sample with common cdf, they are iid, such that we may write:

$$P(X_{(1)} \leq x) = 1 - P(X_1 > x, \dots, X_n > x) = 1 - (P(X_1 > x))^n$$

With the cdf as given this simplifies:

$$P(X_1 > x) = 1 - P(X_1 \leq x|\theta) = 1 - (1 - e^{-(x/\theta)^2}) = e^{-(x/\theta)^2}$$

By the definition of $X_{(1)}$ then:

$$P(X_{(1)} \leq x) = 1 - e^{-n(x/\theta)^2}$$

Let:

$$V = n\left(\frac{X_{(1)}}{\theta}\right)^2 \to P(V \leq v) = 1 - e^{-v}$$

The above cdf is from an Exponential distribution!

So, $V \sim$ Exponential(1), and V is a pivotal quantity.

We require quantiles of the Exponential(1) distribution in order to form confidence intervals.

Let q_p denote the p-th quantile of the Exponential(1) distribution.

We want coverage coefficient:

$$P_\theta\left(q_{\alpha/2} \leq V \leq q_{1-\alpha/2}\right) = 1 - \alpha$$

In terms of θ, solving:

$$q_{\alpha/2} \leq n\left(\frac{X_{(1)}}{\theta}\right)^2 \leq q_{1-\alpha/2} \to \sqrt{\frac{q_{\alpha/2}}{n}} \leq \frac{X_{(1)}}{\theta} \leq \sqrt{\frac{q_{1-\alpha/2}}{n}}$$

After some more algebra, we have:

$$\theta \in \left(\frac{X_{(1)}}{\sqrt{q_{1-\alpha/2}/n}}, \frac{X_{(1)}}{\sqrt{q_{\alpha/2}/n}} \right)$$

I believe the proof may end here, so leaving some space before continuing...

We may simplify further with a note:

By definition, the p-th quantile q_p satisfies the expression:

$$P(X \leq q_p) = p$$

Substituting the given cdf:

$$F(q_p) = p \quad \Rightarrow \quad 1 - e^{-q_p} = p$$

Rearranging:

$$e^{-q_p} = 1 - p \rightarrow -q_p = \log(1-p) \rightarrow q_p = -\log(1-p)$$

So, we may write the p-th quantile q_p of an Exponential(1) random variable as:

$$q_p = -\log(1-p)$$

So, taking our CI for θ given above, we may then write:

For a confidence coefficient of $1-\alpha$, we have:

$$q_{1-\alpha/2} = -\log(\alpha/2)$$

And:

$$q_{\alpha/2} = -\log(1-\alpha/2)$$

Substituting into the above interval:

$$\theta \in \left(\frac{X_{(1)}}{\sqrt{\frac{-\log(\alpha/2)}{n}}}, \frac{X_{(1)}}{\sqrt{\frac{-\log(1-\alpha/2)}{n}}} \right)$$

Which is equivalent to:

$$\theta \in \left(X_{(1)}\sqrt{\frac{n}{-\log(\alpha/2)}}, X_{(1)}\sqrt{\frac{n}{-\log(1-\alpha/2)}} \right)$$

b)

Use the Mood-Graybill-Boes Method to derive a CI for θ with C.C. $1-\alpha$ based on the statistic $T = \sum_{i=1}^{n} X_i^2$. Express your confidence interval using chi-squared quantiles.

Note: One can show X_i^2 is Exponential(θ^2) distributed so that $2T/\theta^2$ is χ^2_{2n} distributed with $2n$ degrees of freedom.

Answer

As given, we know:

$$X_i^2 \sim \text{Exponential}(\theta^2)$$

Note:

$$T = \sum_{i=1}^{n} X_i^2 \to T \sim \text{Gamma}(n, \theta^2)$$

Using the above note, let:

$$T = \sum_{i=1}^{n} X_i^2 \to \frac{2T}{\theta^2} \sim \chi^2_{2n}$$

Where T is a pivotal quantity.

Regarding the coverage coefficient, we define:

$$P_\theta\left(\chi^2_{2n,\alpha/2} \leq \frac{2T}{\theta^2} \leq \chi^2_{2n,1-\alpha/2}\right) = 1 - \alpha$$

Solving for θ^2:

$$\frac{2T}{\chi^2_{2n,1-\alpha/2}} \leq \theta^2 \leq \frac{2T}{\chi^2_{2n,\alpha/2}} \to \sqrt{\frac{2T}{\chi^2_{2n,1-\alpha/2}}} \leq \theta \leq \sqrt{\frac{2T}{\chi^2_{2n,\alpha/2}}}$$

Thus, the confidence interval for θ is:

$$\left(\sqrt{\frac{2\sum_{i=1}^{n} X_i^2}{\chi^2_{2n,1-\alpha/2}}}, \sqrt{\frac{2\sum_{i=1}^{n} X_i^2}{\chi^2_{2n,\alpha/2}}}\right)$$

With the desired coverage $1 - \alpha$.