STA 380

Week 7

These problems are not to be handed in, but they are for extra practice for students to be prepared for the quiz.

1. Consider the Pareto distribution:

$$f_X(x) = \frac{\alpha\beta^\alpha}{x^{\alpha+1}}, \quad x > \beta.$$

Let $X_1, X_2, \ldots, X_n$ be a random sample from a Pareto distribution where $\beta = 2$, and we want to estimate α. Note that the **method of moments estimator (MOME)** is:

$$\hat{\alpha}_{MOME} = \frac{\overline{Y}}{\overline{Y} - 2}$$

The solution can be found here. However, you do not need to recall how to find the method of moments estimator for this course.

(a) Install the `extraDistr` package so we can simulate a random sample from the Pareto distribution, where $\beta = 2$ (`install.packages("extraDistr")`).

(b) Pick any value for α to generate a random sample of size n (also of your choice) from the Pareto distribution. Use $m = 10^4$ replicates. (*Hint: read* `??extraDistr::Pareto`).

(c) Compute a Monte Carlo estimate for $\hat{\alpha}_{MOME}$.

(d) Derive the MLE for α. (Yes, you need to remember how to compute the MLE for Unit 6.)

(e) Compute a Monte Carlo estimate for $\hat{\alpha}_{MLE}$.

(f) Determine which estimator is better: $\hat{\alpha}_{MOME}$ or $\hat{\alpha}_{MLE}$?

(d) ① Likelihood.

$$f_X(x) = \frac{\alpha\beta^\alpha}{x^{\alpha+1}},$$

$$L(\alpha) = \prod_{i=1}^{n} \frac{\alpha\beta^\alpha}{x_i^{\alpha+1}} = \alpha^n \cdot \beta^{n\cdot\alpha} \cdot \prod_{i=1}^{n} x_i^{-(\alpha+1)}$$

② Log-likelihood

$$\log\left(\alpha^n \cdot \beta^{n\cdot\alpha} \cdot \prod_{i=1}^{n} x_i^{-(\alpha+1)}\right)$$

$$\ell(\alpha) = n\cdot\log(\alpha) + n\cdot\alpha\log(\beta) - (\alpha+1)\sum_{i=1}^{n}\log x_i$$

③. Differentiate and set to 0.

$$\frac{\partial \ell(\alpha)}{\partial\alpha} = \frac{n}{\alpha} + n\cdot\log(\beta) - \sum_{i=1}^{n}\log x_i = 0$$

④ $$\hat{\alpha}_{MLE} = \frac{n}{\sum_{i=1}^{n}\log x_i - n\log(\beta)} = \frac{n}{\sum_{i=1}^{n}\log x_i - n\log 2}$$

$\beta = 2$

2. Let $X_1, X_2, \ldots, X_{10}$ be a random sample from the *Exponential*(θ) distribution, where θ is the *scale* parameter.

$n = 10$

(a) Use the NP Lemma to derive the **rejection region** (RR) for the uniformly most powerful test (UMP) of:

$$H_0 : \theta = 3 \quad \text{v.s.} \quad H_1 : \theta > 3$$

Let the significance level be $\alpha = 0.10$. *Hint:* for the final step, you are allowed to use `qexp()` or `qgamma()`. Also, for the exponential distribution, $\theta > 0$.

(b) Use the following code to generate samples from the **null** distribution:

```
n <- 10
alpha <- 0.10
theta <- 3
m <- 10000
x <- matrix(rexp(n*m, rate = 1/theta), nrow = m)
```

Using the rejection region from part (a), compute the type-I error rate.

Type II erro

0.6243

(c) Assume we instead had the simple alternative hypothesis $H_1 : \theta = 4$. Using the rejection region from part (a), Compute the type-II error rate. Also, compute the power.

(d) Why do you think the type-II error rate is high, and the power is low? What happens if you increase the value of θ for your alternative hypothesis?

(a). Exponential pdf (θ) $\quad f(x) = \frac{1}{\theta} e^{-\frac{1}{\theta} x}, \quad \theta > 0, \quad x \in [0, \infty)$

Test?

Likelihood Ratio Test.

Review. NP Lemma. the RR is determined by $\frac{L(\theta_0)}{L(\theta_\alpha)} < k$

① $L(\theta) = \frac{1}{\theta} e^{-\frac{1}{\theta} x_1} * \frac{1}{\theta} e^{-\frac{1}{\theta} x_2} \cdots\cdots * \frac{1}{\theta} e^{-\frac{1}{\theta} x_{10}}$

$= \frac{1}{\theta^{10}} * e^{-\frac{1}{\theta} \sum_{i=1}^{10} x_i}.$

$e^a \cdot e^b = e^{a+b}$

$$\frac{L(\theta_0)}{L(\theta_\alpha)} = \frac{\frac{1}{\theta_0^{10}} \cdot e^{-\frac{1}{\theta_0} \sum_{i=1}^{10} x_i}}{\frac{1}{\theta_\alpha^{10}} \cdot e^{-\frac{1}{\theta_\alpha} \cdot \sum_{i=1}^{10} x_i}}$$

$$= \left(\frac{\theta_\alpha}{\theta_0}\right)^{10} e^{-\frac{1}{\theta_0} - (-\frac{1}{\theta_\alpha}) \sum_{i=1}^{10} x_i} = \left(\frac{\theta_\alpha}{\theta_0}\right)^{10} e^{(\frac{1}{\theta_\alpha} - \frac{1}{\theta_0}) \sum_{i=1}^{10} x_i} < k.$$

$k_1 = k / (\frac{\theta_\alpha}{\theta_0})^{10}$

$$\Rightarrow e^{\left(\frac{1}{\theta_a}-\frac{1}{\theta_0}\right)\sum_{i=1}^{n} x_i} < k_1$$

$$\log\left(e^{\left(\frac{1}{\theta_a}-\frac{1}{\theta_0}\right)\sum_{i=1}^{n}}\right) < \log(k_1)$$

$$\Rightarrow \left(\frac{1}{\theta_a}-\frac{1}{\theta_0}\right)\sum_{i=1}^{n} x_i < k_2$$

$$k^* = k_2 \Big/ \underbrace{\left(\frac{1}{\theta_a}-\frac{1}{\theta_0}\right)}_{<0}.$$

$\theta_a > \theta_0$

$\frac{1}{\theta_a} < \frac{1}{\theta_0}$

$\frac{1}{\theta_a} - \frac{1}{\theta_0} < 0$

$$\Rightarrow \sum_{i=1}^{n} x_i \boxed{>} k^*$$

$(RR \quad | \quad H_0)$

$$0.1 = \alpha = P\left(\sum_{i=1}^{10} x_i > k^* \quad \middle| \quad \theta_0 = 3\right)$$

Review. if the $x_1 \ldots x_n$ iid Exponential (θ).

$$\Rightarrow T \sim \sum_{i=1}^{n} x_i \sim \text{Gamma}(n, \theta)$$

$$\Rightarrow \sum_{i=1}^{10} x_i \sim \text{Gamma}(10, \theta) \quad \theta_0 = 3.$$

k -> qgamma (0.1 , shap = 10, scale = 3, lower.tail = FALSE).

k -> 42.61797.

$k^* \approx 42.61797.$

$$RR = \left\{\sum_{i=1}^{n} x_i > 42.61797\right\}.$$

(c). Type II error $= P($ fail to reject $| \theta_a = 4)$

H_a is true

RR $\Rightarrow \sum_{i=1}^{10} x_i > k^*$

$\sum_{i=1}^{10} x_i \leq k^*$

Type II error $= P(\sum_{i=1}^{10} x_i \leq k^* | \theta_a = 4)$

(d) $T \sim Gamma(10, \theta) \Rightarrow E[T] = 10 \cdot \theta$

H_0 $\quad E_{H_0}[T] = 30$

H_a $\quad E_{H_a}[T] = 40.$ $\qquad H_a = 6 \Rightarrow E_{H_a}[T] = 60.$

$k^* \approx 42 \ldots$