Exam Statistics 2^{nd} February

Maura Mezzetti

1 Exercise 1

Let $(X_1, \ldots, X_n)$ be a random sample of i.i.d. random variables distributed as a Pareto distribution with parameters α and x_m

$$f(x; \alpha, x_\mathrm{m}) = \frac{\alpha \, x_\mathrm{m}^\alpha}{x^{\alpha+1}} \qquad \text{for } x \geq x_m$$

.

1. Find $\hat{\alpha}_{MLE}$ and $\hat{x_\mathrm{m}}_{MLE}$maximum likelihood estimate (MLE) for α and x_m and discuss properties of this estimator,
2. Find a sufficient statistics for α,
3. Show that the LRT of $\alpha = 1$ and x_m unknown and $\alpha \neq 1$ and x_m unknown, has critical region of the form $\{x : \Lambda(x) \leq c_1, \Lambda(x) \geq 1\}$, where $0 < c_1 < c_2$ and
$$\Lambda(x) = log \left[\frac{\prod_{i=1}^n x_i}{(minx_i)^n} \right] > 0$$

Assume x_m known:

4. Compute the score function and the Fisher information for α in X.
5. Find the asymptotic variance of the $\hat{\alpha}_{MLE}$.
6. Complete ONE of the following questions:
   (a) Find the Wald test statistic for testing $H_0 : \alpha = 2$ versus $H_1 : \alpha \neq 2$ and specify the distribution.
   (b) Find the Score test statistic for testing $H_0 : \alpha = 2$ versus $H_1 : \alpha \neq 2$ and specify the distribution.

2 Exercise 2

Let $(X_1, \ldots, X_n)$ be independent identically distributed random variables with p.d.f.

$$f(x) = \theta^2 x \exp(-\theta x) \qquad x > 0$$

Is $T(X_1, \ldots, X_n) = 1/X_1$ an unbiased estimator of θ?

3 Exercise 3

Choose one of the following questions:

1. Provide correct statement for Neyman Pearson Lemma
2. Provide correct statement for Factorization Theorem

4 Solution Exercise 1

1. The log-likelihood function is

$$l(\alpha, x_{\mathrm{m}}) = nlog\alpha + n\alpha log(x_{\mathrm{m}})) - (\alpha+1)\sum_{i=1}^{n} logx_i$$

   Thus

$$\begin{aligned}\frac{\delta l(\alpha, x_{\mathrm{m}})}{\delta\alpha} &= \frac{n}{\alpha} + nlog(x_{\mathrm{m}})) - \sum_{i=1}^{n} logx_i \\ \frac{\delta l(\alpha, x_{\mathrm{m}})}{\delta x_{\mathrm{m}}} &= \frac{n\alpha}{x_{\mathrm{m}}} \qquad x \geq x_{\mathrm{m}}\end{aligned}$$

   Solving for $\frac{\delta l(\alpha, x_{\mathrm{m}})}{\delta\alpha} = 0$, the mle of α is given by

$$\hat{\alpha} = \frac{n}{\sum_{i=1}^{n} logx_i - nlog(x_{\mathrm{m}}))}$$

   First derivative always positive, $\hat{x_{\mathrm{m}}}_{MLE} = min(x_i)$

   Check second derivative....

2. Observe that the joint pdf of $X = (X_1, \ldots, X_n)$

$$\begin{aligned}f(x; \alpha, x_{\mathrm{m}}) &= \prod_{i=1}^{n} \frac{\alpha\, x_{\mathrm{m}}^{\alpha}}{x_i^{\alpha+1}} \\ &= \frac{\alpha^n\, x_{\mathrm{m}}^{n\alpha}}{\prod_{i=1}^{n} x_i^{\alpha+1}} \\ &= g(t, \alpha)h(x)\end{aligned}$$

   where $t = \prod_{i=1}^{n} x_i$ $g(t, \alpha) = c\alpha^n\, x_{\mathrm{m}}^{n\alpha} t^{-(\alpha+1)}$ and $h(x) = 1$. By the factorization theorem, $T(X) = \prod_{i=1}^{n} X_i$ is sufficient for α.

3. Under H_0, the MLE of α is $\hat{\alpha} = 1$ and and the MLE of x_{m} is still $\hat{x_{\mathrm{m}}} = min_i x_i$. So the likelihood ratio statistic is

$$\begin{aligned}LRT(x) &= \frac{\frac{(1min_i x_i)^n}{\left(\prod_{i=1}^{n} x_i\right)^{1+1}}}{\left(\frac{n}{\Lambda(x)}\right)^n \frac{((min_i x_i)^n)^{\frac{n}{\Lambda(x)}}}{\left(\prod_{i=1}^{n} x_i\right)^{\frac{n}{\Lambda(x)}}} \prod_{i=1}^{n} x_i} \\ &= \frac{\frac{\exp(-\Lambda(x))}{\prod_{i=1}^{n} x_i}}{\left(\frac{n}{\Lambda(x)}\right)^n \exp -\Lambda(x) \frac{n}{\Lambda(x)} \prod_{i=1}^{n} x_i}\end{aligned}$$

Note now that

$$\frac{\delta log LRT(x)}{\delta \Lambda(x)} = \delta\delta\Lambda(x) \left\{ nlog\left(\frac{\Lambda(x)}{n}\right) - \Lambda(x) + n \right\} = \frac{n}{\Lambda(x)} - 1$$

so that $LRT(x)$ is increasing for $\Lambda(x) \leq n$ and decreasing if $\Lambda(x) > n$. Thus, $LRT(x) < c$ is equivalent to $\{x : \Lambda(x) \leq c_1, \Lambda(x) \geq 1\}$, for appropriately chosen constants c_1 and c_2.

4.

$$I(\alpha) = -E\left(\frac{\delta^2 l(\alpha)}{\delta\alpha^2} l(\alpha)\right) = \frac{1}{\alpha^2}$$

Hence the asymptotic variance of $\hat{\alpha}_{MLE}$ is α^2/n

5 Solution Exercise 2

The CLT tells us that

$$Z = \frac{\sum_{i=1}^{n} x_i - np}{\sqrt{np(1-p)}}$$

is approximately $N(0,1)$. For a test that rejects H_0 when $\sum_{i=1}^{n} x_i \geq c$, we need to find c and n satisfies:

$$P\left(Z > \frac{c - n(0.49)}{\sqrt{n0.490.51}}\right) = 0.01$$

and

$$P\left(Z > \frac{c - n(0.51)}{\sqrt{n0.490.51}}\right) = 0.99$$

Solving these equations, gives n=13567 and c=6783.5