STAT 410

Examples for 11/06/2020 (2)

Fall 2020
A. Stepanov

Examples for 10/19/2020 (4) & 10/21/2020 (2) & 10/23/2020 (3) (continued)

1. The Pareto probability distribution has many applications in economics, biology, and physics. Let $\beta > 0$ and $\delta > 0$ be the population parameters, and let $X_1, X_2, \ldots, X_n$ be a random sample from the distribution with probability density function

$$f(x; \beta, \delta) = \frac{\delta \cdot \beta^{\delta}}{x^{\delta+1}}, \qquad x > \beta, \qquad \text{zero otherwise.}$$

Suppose β is known.

y) Recall: $Y = \sum_{i=1}^{n} \ln\left(\frac{X_i}{\beta}\right) = \sum_{i=1}^{n} W_i$ has a $\text{Gamma}(\alpha = n, \theta = \frac{1}{\delta})$ distribution.

Suggest a confidence interval for δ with $(1-\alpha)100\%$ confidence level.

z) Suppose $n = 5$, $\beta = 3$, and $x_1 = 3.9$, $x_2 = 4.2$, $x_3 = 6$, $x_4 = 9$, $x_5 = 15$.

Use part (y) to construct a 90% confidence interval for δ.

aa) Find a sufficient statistics for δ.

ab)* Assume $\delta > 2$.

Recall: a method of moments estimator of δ is $\tilde{\delta} = \frac{\overline{X}}{\overline{X} - \beta}$.

Show that $\tilde{\delta}$ is asymptotically normally distributed (as $n \to \infty$). Find the parameters.

① Find $\sigma^2 = \text{Var}(X)$.

② By CLT, $\sqrt{n}\left(\overline{X} - \mu\right) \xrightarrow{D} N\left(0, \sigma^2\right)$.

③ If $g(x)$ is differentiable at μ and $g'(\mu) \neq 0$, …

ac)* Recall: $\hat{\hat{\hat{\delta}}} = \dfrac{n-1}{\sum_{i=1}^{n} \ln\left(\frac{X_i}{\beta}\right)}$ is an unbiased estimator of δ.

Is $\hat{\hat{\hat{\delta}}}$ an efficient estimator of δ? If $\hat{\hat{\hat{\delta}}}$ is not efficient, find its efficiency.

① Find the Fisher information $I(\delta)$.

② Find $\operatorname{Var}(\hat{\hat{\hat{\delta}}})$.

③ Is $\hat{\hat{\hat{\delta}}}$ an efficient estimator of δ? If $\hat{\hat{\hat{\delta}}}$ is not efficient, find its efficiency.

Suppose δ is known.

ad) Find a sufficient statistics for β.

ae)* Recall: the maximum likelihood estimator of β is $\hat{\beta} = \min X_i$.

Let $U_n = n(\hat{\beta} - \beta)$. Find the limiting distribution of U_n.

① Find the cumulative distribution function of U_n, $F_{U_n}(u)$.

② $F_\infty(u) = \lim\limits_{n\to\infty} F_{U_n}(u)$, if the limit exists and if $F_\infty(u)$ is a CDF.

Answers:

1. The Pareto probability distribution has many applications in economics, biology, and physics. Let $\beta > 0$ and $\delta > 0$ be the population parameters, and let $X_1, X_2, \ldots, X_n$ be a random sample from the distribution with probability density function

$$f(x; \beta, \delta) = \frac{\delta \cdot \beta^{\delta}}{x^{\delta+1}}, \qquad x > \beta, \qquad \text{zero otherwise.}$$

Suppose β is known.

y) Recall: $Y = \sum_{i=1}^{n} \ln\left(\frac{X_i}{\beta}\right) = \sum_{i=1}^{n} W_i$ has a Gamma$(\alpha = n, \theta = \frac{1}{\delta})$ distribution.

Suggest a confidence interval for δ with $(1 - \alpha)100\%$ confidence level.

$W = \ln\left(\frac{X}{\beta}\right)$ has Gamma$(\alpha = 1, \theta = \frac{1}{\delta})$ distribution.

$\Rightarrow \quad \sum_{i=1}^{n} \ln\left(\frac{X_i}{\beta}\right)$ has Gamma$(\alpha = n, \theta = \frac{1}{\delta})$ distribution.

$\Rightarrow \quad 2\delta \sum_{i=1}^{n} \ln\left(\frac{X_i}{\beta}\right)$ has a $\chi^2(2\alpha = 2n)$ distribution.

$$\Rightarrow \quad P\left(\chi^2_{1-\alpha/2}(2n) < 2\delta \sum_{i=1}^{n} \ln\left(\frac{X_i}{\beta}\right) < \chi^2_{\alpha/2}(2n)\right) = 1 - \alpha.$$

$$\Rightarrow \quad P\left(\frac{\chi^2_{1-\alpha/2}(2n)}{2\sum_{i=1}^{n} \ln\left(\frac{X_i}{\beta}\right)} < \delta < \frac{\chi^2_{\alpha/2}(2n)}{2\sum_{i=1}^{n} \ln\left(\frac{X_i}{\beta}\right)}\right) = 1 - \alpha.$$

A $(1-\alpha)100\%$ confidence interval for δ:

$$\left(\frac{\chi^2_{1-\alpha/2}(2n)}{2\sum_{i=1}^{n} \ln\left(\frac{x_i}{\beta}\right)}, \frac{\chi^2_{\alpha/2}(2n)}{2\sum_{i=1}^{n} \ln\left(\frac{x_i}{\beta}\right)} \right).$$

OR

$$P\left(0 < 2\delta\sum_{i=1}^{n} \ln\left(\frac{X_i}{\beta}\right) < \chi^2_{\alpha}(2n)\right) = 1-\alpha.$$

A $(1-\alpha)100\%$ confidence interval for δ:
$$\left(0, \frac{\chi^2_{\alpha}(2n)}{2\sum_{i=1}^{n} \ln\left(\frac{x_i}{\beta}\right)} \right).$$

OR

$$P\left(\chi^2_{1-\alpha}(2n) < 2\delta\sum_{i=1}^{n} \ln\left(\frac{X_i}{\beta}\right) < \infty\right) = 1-\alpha.$$

A $(1-\alpha)100\%$ confidence interval for δ:
$$\left(\frac{\chi^2_{1-\alpha}(2n)}{2\sum_{i=1}^{n} \ln\left(\frac{x_i}{\beta}\right)}, \infty \right).$$

z) Suppose $n = 5$, $\beta = 3$, and $x_1 = 3.9$, $x_2 = 4.2$, $x_3 = 6$, $x_4 = 9$, $x_5 = 15$.

Use part (y) to construct a 90% confidence interval for δ.

$$\sum_{i=1}^{n} \ln\left(\frac{x_i}{3}\right) \approx 4. \qquad \chi^2_{0.95}(10) = 3.940, \qquad \chi^2_{0.05}(10) = 18.31.$$

$$\left(\frac{3.940}{2 \cdot 4}, \frac{18.31}{2 \cdot 4} \right) = (0.4925, 2.28875).$$

OR

$$\chi^2_{0.10}(10) = 15.99. \qquad \chi^2_{0.90}(10) = 4.865.$$

$$\left(0, \frac{15.99}{2 \cdot 4} \right) \approx (0, 1.999). \qquad \left(\frac{4.865}{2 \cdot 4}, \infty \right) \approx (0.608, \infty).$$

aa) Find a sufficient statistics for δ.

Define
$$\mathrm{I}_{\{A\}} = \begin{cases} 1 & \text{if A is true} \\ 0 & \text{if A is false} \end{cases}.$$

$$f(x; \beta, \delta) = \frac{\delta \cdot \beta^{\delta}}{x^{\delta+1}} \cdot \mathrm{I}_{\{x > \beta\}}.$$

$$\prod_{i=1}^{n} f(x_i; \beta, \delta) = \delta^n \cdot \beta^{n\delta} \cdot \left(\prod_{i=1}^{n} x_i \right)^{-(\delta+1)} \cdot \left(\prod_{i=1}^{n} \mathrm{I}_{\{x_i > \beta\}} \right).$$

By Factorization Theorem, $\prod_{i=1}^{n} X_i$ is a sufficient statistic for δ.

$$\Rightarrow \qquad \sum_{i=1}^{n} \ln X_i \quad \text{is also a sufficient statistic for } \delta.$$

$$\Rightarrow \qquad \text{IF } \beta \text{ is known, } \quad \sum_{i=1}^{n} \ln\left(\frac{X_i}{\beta}\right) \quad \text{is also a sufficient statistic for } \delta.$$

ab)* Assume $\delta > 2$.

Recall: a method of moments estimator of δ is $\tilde{\delta} = \dfrac{\overline{X}}{\overline{X} - \beta}$.

Show that $\tilde{\delta}$ is asymptotically normally distributed (as $n \to \infty$). Find the parameters.

① Find $\sigma^2 = \operatorname{Var}(X)$.

② By CLT, $\sqrt{n}\left(\overline{X} - \mu\right) \xrightarrow{D} N\left(0, \sigma^2\right)$.

③ If $g(x)$ is differentiable at μ and $g'(\mu) \neq 0$, …

① $$\mu = E(X) = \int_{\beta}^{\infty} x \cdot \frac{\delta \cdot \beta^{\delta}}{x^{\delta+1}}\, dx = \frac{\beta\,\delta}{\delta-1}.$$

$$E(X^2) = \int_{\beta}^{\infty} x^2 \cdot \frac{\delta \cdot \beta^{\delta}}{x^{\delta+1}}\, dx = \frac{\beta^2\,\delta}{\delta-2}.$$

$$\sigma^2 = \operatorname{Var}(X) = \frac{\beta^2\,\delta}{\delta-2} - \left(\frac{\beta\,\delta}{\delta-1}\right)^2 = \frac{\beta^2\,\delta}{(\delta-2)\,(\delta-1)^2}.$$

② By CLT, $$\sqrt{n}\left(\overline{X} - \mu\right) \xrightarrow{D} N\left(0, \sigma^2\right).$$

$$\overline{X} \text{ is approximately } N\left(\frac{\beta\,\delta}{\delta-1}, \frac{\beta^2\,\delta}{(\delta-2)\,(\delta-1)^2\, n}\right) \quad \text{for large } n.$$

③ Consider $g(x) = \dfrac{x}{x-\beta}$.

Then $$g(\overline{X}) = \frac{\overline{X}}{\overline{X}-\beta} = \tilde{\delta}, \qquad g(\mu) = \frac{\dfrac{\beta\,\delta}{\delta-1}}{\dfrac{\beta\,\delta}{\delta-1}-\beta} = \delta.$$

$$g'(x) = \frac{-\beta}{(x-\beta)^2}. \qquad g'(\mu) = \frac{-\beta}{\left(\dfrac{\beta\,\delta}{\delta-1}-\beta\right)^2} = \frac{-(\delta-1)^2}{\beta} \neq 0.$$

$g(\overline{X})$ is approximately $N\left(g(\mu), [g'(\mu)]^2\,\dfrac{\sigma^2}{n}\right)$ for large n.

Therefore, $\tilde{\delta}$ is approximately $N\left(\delta, \dfrac{\delta\,(\delta-1)^2}{(\delta-2)\,n}\right)$ for large n.

OR $$\sqrt{n}\,\left(\tilde{\delta} - \delta\right) \xrightarrow{D} N\left(0, \frac{\delta\,(\delta-1)^2}{(\delta-2)}\right).$$

ac)* Recall: $\hat{\hat{\delta}} = \dfrac{n-1}{\sum\limits_{i=1}^{n} \ln\left(\dfrac{X_i}{\beta}\right)}$ is an unbiased estimator of δ.

Is $\hat{\hat{\delta}}$ an efficient estimator of δ? If $\hat{\hat{\delta}}$ is not efficient, find its efficiency.

① Find the Fisher information $I(\delta)$.

② Find $\operatorname{Var}(\hat{\hat{\delta}})$.

③ Is $\hat{\hat{\delta}}$ an efficient estimator of δ? If $\hat{\hat{\delta}}$ is not efficient, find its efficiency.

① $\ln f(x;\beta,\delta) = \ln\delta + \delta\cdot\ln\beta - (\delta+1)\cdot\ln x.$

$$\frac{\partial}{\partial\delta}\ln f(x;\beta,\delta) = \frac{1}{\delta} + \ln\beta - \ln x = \frac{1}{\delta} - \ln\left(\frac{x}{\beta}\right).$$

$$\frac{\partial^2}{\partial\delta^2}\ln f(x;\beta,\delta) = -\frac{1}{\delta^2}.$$

$$I(\delta) = -E\left[\frac{\partial^2}{\partial\delta^2}\ln f(X;\beta,\delta)\right] = -E\left[-\frac{1}{\delta^2}\right] = \frac{1}{\delta^2}.$$

OR

$$I(\delta) = \operatorname{Var}\left[\frac{\partial}{\partial\delta}\ln f(X;\beta,\delta)\right] = \operatorname{Var}\left[\frac{1}{\delta} - \ln\left(\frac{X}{\beta}\right)\right] = \operatorname{Var}\left[\ln\left(\frac{X}{\beta}\right)\right] = \frac{1}{\delta^2},$$

since $W = \ln\left(\dfrac{X}{\beta}\right)$ has an Exponential$\left(\theta = \dfrac{1}{\delta}\right)$ distribution.

② $Y = \sum_{i=1}^{n} \ln\left(\frac{X_i}{\beta}\right) = \sum_{i=1}^{n} W_i$ has a $\text{Gamma}\left(\alpha = n, \theta = \frac{1}{\delta}\right)$ distribution.

$$\operatorname{Var}\left(Y^{-1}\right) = \frac{\delta^2}{(n-2)\,(n-1)^2}$$ (Examples for 10/19/2020 (4) **1** (h)).

$$\hat{\hat{\delta}} = \frac{n-1}{\sum_{i=1}^{n} \ln\left(\frac{X_i}{\beta}\right)} = \frac{n-1}{Y}.$$

$$\operatorname{Var}(\hat{\hat{\delta}}) = (n-1)^2 \operatorname{Var}\left(Y^{-1}\right) = \frac{\delta^2}{n-2}.$$

③ Rao-Cramer Lower Bound:

$$\frac{\left(k'(\delta)\right)^2}{n \cdot I(\delta)} = \frac{1}{n \cdot I(\delta)} = \frac{\delta^2}{n}.$$

$$\operatorname{Var}(\hat{\hat{\delta}}) = \frac{\delta^2}{n-2} > \frac{\delta^2}{n} = \text{R.C.L.B.}$$

$\Rightarrow$ $\hat{\hat{\delta}}$ is NOT an efficient estimator of δ,

its efficiency $= \frac{n-2}{n} \to 1$ as $n \to \infty$.

Suppose δ is known.

ad) Find a sufficient statistics for β.

$$\prod_{i=1}^{n} f\left(x_i\,;\,\beta,\delta\right) = \delta^{n}\cdot\beta^{n\delta}\cdot\left(\prod_{i=1}^{n} x_i\right)^{-(\delta+1)}\cdot\left(\prod_{i=1}^{n} \mathrm{I}_{\{x_i>\beta\}}\right)$$

$$= \delta^{n}\cdot\beta^{n\delta}\cdot\left(\prod_{i=1}^{n} x_i\right)^{-(\delta+1)}\cdot \mathrm{I}_{\{\min x_i>\beta\}}.$$

By Factorization Theorem, $\min \mathrm{X}_i$ is a sufficient statistic for β.

ae)* Recall: the maximum likelihood estimator of β is $\hat{\beta} = \min \mathrm{X}_i$.

Let $\mathrm{U}_n = n(\hat{\beta}-\beta)$. Find the limiting distribution of U_n.

① Find the cumulative distribution function of U_n, $\mathrm{F}_{\mathrm{U}_n}(u)$.

② $\mathrm{F}_{\infty}(u) = \lim\limits_{n\to\infty} \mathrm{F}_{\mathrm{U}_n}(u)$, if the limit exists and if $\mathrm{F}_{\infty}(u)$ is a CDF.

① $$\mathrm{F}_{\mathrm{X}}(x) = \mathrm{P}(\mathrm{X}\le x) = \int_{\beta}^{x} \frac{\delta\cdot\beta^{\delta}}{u^{\delta+1}}\,du = 1-\frac{\beta^{\delta}}{x^{\delta}}, \qquad x>\beta.$$

$$\mathrm{F}_{\min \mathrm{X}_i}(x) = 1-\left(1-\mathrm{F}_{\mathrm{X}}(x)\right)^{n} = 1-\left(\frac{\beta^{\delta}}{x^{\delta}}\right)^{n} = 1-\frac{\beta^{\delta n}}{x^{\delta n}}, \qquad x>\beta.$$

$$\mathrm{F}_{\mathrm{U}_n}(u) = \mathrm{P}(\mathrm{U}_n\le u) = \mathrm{P}\left(\hat{\beta}\le\beta+\frac{u}{n}\right) = \mathrm{P}\left(\min \mathrm{X}_i\le\beta+\frac{u}{n}\right)$$

$$= F_{\min X_i}\left(\beta + \frac{u}{n}\right) = 1 - \left(\frac{\beta}{\beta + \frac{u}{n}}\right)^{n\delta} = 1 - \left(1 + \frac{u}{\beta n}\right)^{-n\delta}, \qquad u > 0.$$

② $$F_{\infty}(u) = \lim_{n \to \infty} F_{U_n}(u) = 1 - e^{-u\delta/\beta}, \qquad u > 0.$$

The limiting distribution is an Exponential distribution with mean $\theta = \frac{\beta}{\delta}$.

For fun:

Recall: $\hat{\beta} \xrightarrow{P} \beta$.

Consider $U_n = n^{\gamma}(\hat{\beta} - \beta)$.

$$F_{U_n}(u) = 1 - \left(1 + \frac{u}{\beta n^{\gamma}}\right)^{-n\delta}, \qquad u > 0.$$

If $\gamma = 1$, $$F_{\infty}(u) = \lim_{n \to \infty} F_{U_n}(u) = 1 - e^{-u\delta/\beta}, \qquad u > 0.$$

The limiting distribution is an Exponential distribution with mean $\theta = \frac{\beta}{\delta}$.

If $\gamma < 1$, $$\lim_{n \to \infty} F_{U_n}(u) = 1, \qquad u > 0.$$

Then $U_n \xrightarrow{D} 0$, and thus $U_n \xrightarrow{P} 0$.

If $\gamma > 1$, $$\lim_{n \to \infty} F_{U_n}(u) = 0, \qquad u > 0.$$

Then U_n does not have a limiting distribution.