STAT 410 **Examples for 10/19/2020 (4)** Fall 2020
A. Stepanov

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.

a) Find the probability distribution of $W = \ln\left(\frac{X}{\beta}\right) = \ln X - \ln \beta$.

b) Find the probability distribution of $Y = \sum_{i=1}^{n} \ln\left(\frac{X_i}{\beta}\right) = \sum_{i=1}^{n} \ln X_i - n \ln \beta$.

c) Suppose $n = 5$, $\beta = 3$, and $\delta = 1.5$. Find $P\left(\sum_{i=1}^{5} \ln\left(\frac{X_i}{3}\right) \le 4\right)$.

d) Suppose $n = 5$, $\beta = 3$, and $\delta = 1.5$. Find c such that $P\left(\sum_{i=1}^{5} \ln\left(\frac{X_i}{3}\right) \le c\right) = 0.90$.

e) Obtain the maximum likelihood estimator of δ, $\hat{\delta}$.

f) Suppose $n = 5$, $\beta = 3$, and $x_1 = 3.9$, $x_2 = 4.2$, $x_3 = 6$, $x_4 = 9$, $x_5 = 15$.
Find the maximum likelihood estimate of δ.

g) Is the maximum likelihood estimator $\hat{\delta}$ an unbiased estimator of δ?
If $\hat{\delta}$ is not an unbiased estimator of δ, construct an unbiased estimator of δ based on $\hat{\delta}$.
"Hint": Recall part (b).

h) Find $\text{MSE}(\hat{\delta}) = (\text{bias}(\hat{\delta}))^2 + \text{Var}(\hat{\delta})$.

i) Assume $\delta > 1$. Obtain a method of moments estimator of δ, $\tilde{\delta}$.

j) Suppose $n = 5$, $\beta = 3$, and $x_1 = 3.9$, $x_2 = 4.2$, $x_3 = 6$, $x_4 = 9$, $x_5 = 15$.
Find a method of moments estimate of δ.

k) Is the method of moments estimator of δ, $\tilde{\delta}$, an unbiased estimator of δ?
If $\tilde{\delta}$ is not an unbiased estimator of δ, does $\tilde{\delta}$ underestimate or overestimate δ (on average)?

"Hint": $\text{E}(\overline{\text{V}}) = \mu_{\text{V}} = \text{E}(\text{V})$. $\qquad \text{Var}(\overline{\text{V}}) = \dfrac{\sigma_{\text{V}}^2}{n} = \dfrac{\text{Var}(\text{V})}{n}$.

$\text{Var}(\text{V}) = \text{E}(\text{V}^2) - [\text{E}(\text{V})]^2$.

$\text{E}(a ☺) = a\,\text{E}(☺)$. $\qquad \text{Var}(a ☺) = a^2\,\text{Var}(☺)$.

If T_α has a $\text{Gamma}(\alpha,\ \theta = \frac{1}{\lambda})$ distribution, then

$$\text{E}(\text{T}_\alpha^k) = \frac{\theta^k\,\Gamma(\alpha+k)}{\Gamma(\alpha)} = \frac{\Gamma(\alpha+k)}{\lambda^k\,\Gamma(\alpha)}, \qquad k > -\alpha.$$

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 $X_1, X_2, \dots, X_n$ 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.

a) Find the probability distribution of $W = \ln\left(\frac{X}{\beta}\right) = \ln X - \ln \beta$.

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

$$F_W(w) = P(W \le w) = P\left(\ln\left(\frac{X}{\beta}\right) \le w\right) = P(X \le \beta e^{w}) = F_X(\beta e^{w})$$

$$= 1 - \frac{\beta^{\delta}}{\beta^{\delta} e^{\delta w}} = 1 - e^{-\delta w}, \qquad w > 0.$$

$\Rightarrow$ W has an $\text{Exponential}(\theta = \frac{1}{\delta}) = \text{Gamma}(\alpha = 1, \theta = \frac{1}{\delta})$ distribution.

OR

$$w = \ln\left(\frac{x}{\beta}\right) \qquad x = \beta e^{w} = g^{-1}(w) \qquad \frac{dx}{dw} = \beta e^{w}$$

$$f_W(w) = f_X(g^{-1}(w)) \cdot \left| \frac{dx}{dw} \right| = \frac{\delta \cdot \beta^{\delta}}{\left(\beta e^{w}\right)^{\delta+1}} \cdot \beta e^{w} = \delta e^{-\delta w}, \qquad w > 0.$$

$\Rightarrow$ W has an Exponential$\left(\theta = \frac{1}{\delta}\right)$ = Gamma$\left(\alpha = 1, \theta = \frac{1}{\delta}\right)$ distribution.

b) Find the probability distribution of $Y = \sum_{i=1}^{n} \ln\left(\frac{X_i}{\beta}\right) = \sum_{i=1}^{n} \ln X_i - n \ln\beta$.

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

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

c) Suppose $n = 5$, $\beta = 3$, and $\delta = 1.5$. Find $P\left(\sum_{i=1}^{5} \ln\left(\frac{X_i}{3}\right) \le 4\right)$.

$$P\left(\sum_{i=1}^{5} \ln\left(\frac{X_i}{3}\right) \le 4\right) = P\left(\text{Gamma}\left(\alpha = 5, \theta = \frac{1}{1.5}\right) \le 4\right) = P(T_5 \le 4)$$

$$= P(\text{Poisson}(1.5 \cdot 4) \ge 5) = 1 - P(\text{Poisson}(6) \le 4)$$

$$= 1 - 0.285 = \mathbf{0.715}.$$

OR

$$\mathrm{P}(\mathrm{T}_5 \le 4) = \int_0^4 \frac{3^5}{\Gamma(5)}\, t^{5-1} e^{-3t}\, dt = \int_0^4 \frac{3^5}{4!}\, t^4 e^{-3t}\, dt = \ldots$$

OR

$\dfrac{2\,\mathrm{T}_5}{\theta} = 2\,\delta\,\mathrm{T}_5 = 3\,\mathrm{T}_5$ has a $\chi^2(2\alpha = 10)$ distribution.

$\mathrm{P}(\mathrm{T}_5 \le 4) = \mathrm{P}(\chi^2(10) \le 12) \approx$ **0.714943**.

d) Suppose $n = 5$, $\beta = 3$, and $\delta = 1.5$. Find c such that $\mathrm{P}\left(\sum_{i=1}^{5} \ln\left(\dfrac{\mathrm{X}_i}{3}\right) \le c \right) = 0.90$.

$\dfrac{2\,\mathrm{T}_5}{\theta} = 2\,\delta\,\mathrm{T}_5 = 3\,\mathrm{T}_5$ has a $\chi^2(2\alpha = 10)$ distribution.

$\chi^2_{0.10}(10) = 15.99.$ $\quad\quad\quad\quad$ $\mathrm{P}(\chi^2(10) \le 15.99) = 0.90.$

$$0.10 = \mathrm{P}\left(\sum_{i=1}^{5} \ln\left(\frac{\mathrm{X}_i}{3}\right) \le c \right) = \mathrm{P}(\mathrm{T}_5 \le c) = \mathrm{P}(\chi^2(10) \le 3\,c).$$

$3\,c = 15.99.$ $\quad\Rightarrow\quad$ $c =$ **5.33**.

With a bit “gentler” rounding:

$\chi^2_{0.10}(10) = 15.987.$ $\quad\quad\quad\quad$ $\chi^2_{0.10}(10) = 15.98718.$

$c = 5.329.$ $\quad\quad\quad\quad$ $c = 5.32906.$

OR

$$P\left(\sum_{i=1}^{5} \ln\left(\frac{X_i}{3}\right) \le c\right) = P(\text{Gamma}(\alpha = 5, \theta = \frac{1}{1.5}) \le c) = P(T_5 \le c)$$

$$= P(\text{Poisson}(1.5 \cdot c) \ge 5) = 1 - P(\text{Poisson}(1.5 \cdot c) \le 4) = 0.90.$$

$$P(\text{Poisson}(8.0) \le 4) = 0.10.$$

$$1.5 \cdot c = 8.0. \qquad \Rightarrow \qquad c = \frac{16}{3} \approx 5.333333.$$

R:

c)

```
> pgamma(4,5,1.5)
[1] 0.7149435
```

```
> 1-ppois(5-1,1.5*4)
```

```
> pchisq(2*1.5*4,2*5)
```

d)

```
qgamma(0.90,5,1.5)
5.32906
```

```
> qchisq(0.90,2*5)/(2*1.5)
```

Excel:

c)

=GAMMA.DIST(4,5,1/1.5,1)

0.714943

=1-POISSON.DIST(5-1,1.5*4,1)

=CHISQ.DIST(2*1.5*4,2*5,1)
=1-CHISQ.DIST.RT(2*1.5*4,2*5)

d)

=GAMMA.INV(0.90,5,1/1.5)

5.32906

=CHISQ.INV(0.9,2*5)/(2*1.5)
=CHISQ.INV.RT(1-0.90,2*5)/(2*1.5)

e) Obtain the maximum likelihood estimator of δ, $\hat{\delta}$.

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

$$\ln L(\delta) = n\cdot\ln\delta + n\,\delta\cdot\ln\beta - (\delta+1)\cdot\sum_{i=1}^{n}\ln x_i\,.$$

$$\frac{d}{d\delta}\ln L(\delta) = \frac{n}{\delta} + n\cdot\ln\beta - \sum_{i=1}^{n}\ln x_i = 0.$$

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

f) Suppose $n = 5$, $\beta = 3$, and $x_1 = 3.9$, $x_2 = 4.2$, $x_3 = 6$, $x_4 = 9$, $x_5 = 15$.
Find the maximum likelihood estimate of δ.

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

$$\sum_{i=1}^{n}\ln x_i \approx 9.4931. \qquad \hat{\delta} \approx \frac{5}{9.4931 - 5\cdot\ln 3} \approx \mathbf{1.25}.$$

OR $$\sum_{i=1}^{n}\ln\left(\frac{x_i}{\beta}\right) = \sum_{i=1}^{n}\ln\left(\frac{x_i}{3}\right) \approx 4. \qquad \hat{\delta} \approx \frac{5}{4} = \mathbf{1.25}.$$

g) Is the maximum likelihood estimator $\hat{\delta}$ an unbiased estimator of δ?
If $\hat{\delta}$ is not an unbiased estimator of δ, construct an unbiased estimator of δ based on $\hat{\delta}$.

"Hint": Recall part (b).

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

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

If T_α has a Gamma$\left(\alpha,\ \theta = \frac{1}{\lambda}\right)$ distribution, then

$$E(T_\alpha^k) = \frac{\theta^k\,\Gamma(\alpha+k)}{\Gamma(\alpha)} = \frac{\Gamma(\alpha+k)}{\lambda^k\,\Gamma(\alpha)}, \qquad k > -\alpha.$$

$$E(\hat{\delta}) = E(n\,Y^{-1}) = n\,E(Y^{-1}) = n\,\frac{\delta\,\Gamma(n-1)}{\Gamma(n)} = \frac{n}{n-1}\,\delta \neq \delta.$$

$\hat{\delta}$ is NOT an unbiased estimator of δ.

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

$$E(\hat{\hat{\delta}}) = \frac{n-1}{n}\,E(\hat{\delta}) = \delta.$$

$\hat{\hat{\delta}}$ is an unbiased estimator of δ.

h) Find $\operatorname{MSE}(\hat{\delta}) = (\operatorname{bias}(\hat{\delta}))^2 + \operatorname{Var}(\hat{\delta})$.

$$\operatorname{Var}(\hat{\delta}) = \operatorname{Var}(n\,\mathrm{Y}^{-1}) = n^2 \operatorname{Var}(\mathrm{Y}^{-1}) = n^2 \left\{ \mathrm{E}(\mathrm{Y}^{-2}) - [\mathrm{E}(\mathrm{Y}^{-1})]^2 \right\}.$$

If T_α has a $\operatorname{Gamma}(\alpha,\ \theta = \frac{1}{\lambda})$ distribution, then

$$\mathrm{E}(\mathrm{T}_\alpha^k) = \frac{\theta^k\,\Gamma(\alpha+k)}{\Gamma(\alpha)} = \frac{\Gamma(\alpha+k)}{\lambda^k\,\Gamma(\alpha)}, \qquad k > -\alpha.$$

$$\mathrm{E}(\mathrm{Y}^{-2}) = \frac{\delta^2\,\Gamma(n-2)}{\Gamma(n)} = \frac{\delta^2}{(n-2)(n-1)}.$$

$$\operatorname{Var}(\mathrm{Y}^{-1}) = \frac{\delta^2}{(n-2)(n-1)} - \frac{\delta^2}{(n-1)^2} = \frac{\delta^2}{(n-2)(n-1)^2}$$

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

$$\operatorname{bias}(\hat{\delta}) = \mathrm{E}(\hat{\delta}) - \delta = \frac{n}{n-1}\,\delta - \delta = \frac{\delta}{n-1}.$$

$$\operatorname{MSE}(\hat{\delta}) = (\operatorname{bias}(\hat{\delta}))^2 + \operatorname{Var}(\hat{\delta}) = \frac{\delta^2}{(n-1)^2} + \frac{n^2\,\delta^2}{(n-2)(n-1)^2}$$

$$= \frac{\left(n^2+n-2\right)\delta^2}{(n-2)(n-1)^2} = \frac{(n+2)\,\delta^2}{(n-2)(n-1)}.$$

i) Assume $\delta > 1$. Obtain a method of moments estimator of δ, $\tilde{\delta}$.

$$\mathrm{E}(\mathrm{X}) = \int_{-\infty}^{\infty} x \cdot f\left(x;\beta,\delta\right) dx = \int_{\beta}^{\infty} x \cdot \frac{\delta \cdot \beta^{\delta}}{x^{\delta+1}} dx = \delta\,\beta^{\delta} \cdot \int_{\beta}^{\infty} x^{-\delta} dx = \frac{\beta\,\delta}{\delta-1}.$$

$$\overline{\mathrm{X}} = \frac{\beta\,\delta}{\delta-1}. \qquad \Rightarrow \qquad \tilde{\delta} = \frac{\overline{\mathrm{X}}}{\overline{\mathrm{X}}-\beta}.$$

j) Suppose $n = 5$, $\beta = 3$, and $x_1 = 3.9$, $x_2 = 4.2$, $x_3 = 6$, $x_4 = 9$, $x_5 = 15$.
Find a method of moments estimate of δ.

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

$$\sum_{i=1}^{n} x_i = 38.1. \qquad \bar{x} = 7.62. \qquad \tilde{\delta} = \frac{7.62}{7.62-3} \approx \mathbf{1.65}.$$

k) Is the method of moments estimator of δ, $\tilde{\delta}$, an unbiased estimator of δ?
If $\tilde{\delta}$ is not an unbiased estimator of δ, does $\tilde{\delta}$ underestimate or overestimate δ (on average)?

Consider $g(x) = \dfrac{x}{x-\beta}$. Then $g(\overline{\mathrm{X}}) = \tilde{\delta}$, $g(\mu) = \delta$.

Also $g''(x) = \dfrac{2\beta}{(x-\beta)^3} > 0$ for $x > \beta$, i.e., $g(x)$ is strictly convex.

By Jensen's Inequality,

$$\mathrm{E}(\tilde{\delta}) = \mathrm{E}[g(\overline{\mathrm{X}})] > g(\mathrm{E}(\overline{\mathrm{X}})) = g(\mu) = \delta.$$

Therefore, $\tilde{\delta}$ is NOT an unbiased estimator of δ.
On average, $\tilde{\delta}$ **overestimates** δ.