STAT 410

Homework #9

(due Friday, November 6, by 5:00 p.m. CST)

Fall 2020
A. Stepanov

No credit will be given without supporting work.

7. Let $\psi > 0$ be a population parameter, and let $X_1, X_2, \ldots, X_n$ be a random sample from a distribution with probability density function

$$f(x;\psi) = \frac{2}{\sqrt{\pi\psi}}\, e^{-x^2/\psi}, \qquad x > 0, \qquad \text{zero otherwise.}$$

Recall: $W = X^2$ has $\text{Gamma}(\alpha = \frac{1}{2}, \theta = \psi)$ distribution.

i) Construct a consistent estimator of ψ based on $\sum_{i=1}^{n} X_i^4$.

$$E(X^4) = E(W^2) = \text{Var}(W) + [E(W)]^2 = \alpha\theta^2 + (\alpha\theta)^2 = \frac{3}{4}\psi^2.$$

By WLLN, $\overline{X^4} = \frac{1}{n}\cdot\sum_{i=1}^{n} X_i^4 \xrightarrow{P} E(X^4) = \frac{3}{4}\psi^2.$

$$\spadesuit \xrightarrow{P} a, \; g \text{ is continuous at } a \;\Rightarrow\; g(\spadesuit) \xrightarrow{P} g(a)$$

Consider $g(x) = \sqrt{\frac{4}{3}x}$. Since $g(x) = \sqrt{\frac{4}{3}x}$ is continuous at $\frac{3}{4}\psi^2$,

$$\tilde{\tilde{\psi}} = \sqrt{\frac{4}{3n}\sum_{i=1}^{n} X_i^4} = \sqrt{\frac{4}{3}\overline{X^4}} = g(\overline{X^4}) \xrightarrow{P} g(\tfrac{3}{4}\psi^2) = \psi.$$

OR

$$\mathrm{E}(\mathrm{X}^k) = \mathrm{E}(\mathrm{W}^{k/2}) = \ldots$$

W has Gamma$\left(\alpha = \frac{1}{2}, \theta = \psi\right)$ distribution. $\qquad \mathrm{W} = \mathrm{T}_{1/2}$.

If T_α has a Gamma$(\alpha, \theta = 1/\lambda)$ distribution, then

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

$$\ldots = \frac{\psi^{k/2}\,\Gamma\left(\frac{1}{2}+\frac{k}{2}\right)}{\Gamma\left(\frac{1}{2}\right)} = \frac{\psi^{k/2}\,\Gamma\left(\frac{1}{2}+\frac{k}{2}\right)}{\sqrt{\pi}}.$$

$$\mathrm{E}(\mathrm{X}^4) = \frac{\psi^{4/2}\,\Gamma\left(\frac{1}{2}+\frac{4}{2}\right)}{\sqrt{\pi}} = \frac{\psi^2\,\Gamma\left(\frac{5}{2}\right)}{\sqrt{\pi}} = \frac{\psi^2\,\frac{3}{2}\,\Gamma\left(\frac{3}{2}\right)}{\sqrt{\pi}}$$

$$= \frac{\psi^2\,\frac{3}{2}\cdot\frac{1}{2}\,\Gamma\left(\frac{1}{2}\right)}{\sqrt{\pi}} = \frac{3}{4}\,\psi^2.$$

OR

$$\mathrm{E}(\mathrm{X}^4) = \int_0^\infty x^4 \cdot \frac{2}{\sqrt{\pi\psi}}\, e^{-x^2/\psi}\, dx \qquad u = \frac{x^2}{\psi} \qquad x^2 = \psi u$$

$$x = \sqrt{\psi u} \qquad dx = \frac{\sqrt{\psi}}{2\sqrt{u}}\, du$$

$$= \int_0^\infty (\psi u)^2 \cdot \frac{2}{\sqrt{\pi\psi}}\, e^{-u}\, \frac{\sqrt{\psi}}{2\sqrt{u}}\, du$$

$$= \frac{\psi^2}{\sqrt{\pi}} \int_0^\infty u^{3/2}\, e^{-u}\, du = \frac{\psi^2}{\sqrt{\pi}}\,\Gamma\left(\frac{5}{2}\right)$$

$$= \frac{\psi^2}{\sqrt{\pi}} \cdot \frac{3}{2}\,\Gamma\left(\frac{3}{2}\right) = \frac{\psi^2}{\sqrt{\pi}} \cdot \frac{3}{2} \cdot \frac{1}{2}\,\Gamma\left(\frac{1}{2}\right) = \frac{3}{4}\,\psi^2.$$

j) (i) Suggest a confidence interval for ψ with $(1-\alpha)\,100\,\%$ confidence level.

(ii) Suppose $n = 4$, and $x_1 = 0.2$, $x_2 = 0.6$, $x_3 = 1.1$, $x_4 = 1.7$.

Construct a 95% confidence interval for ψ.

"Hint": Use $\sum_{i=1}^{n} \mathrm{X}_i^2$.

(i) $\mathrm{Y} = \sum_{i=1}^{n} \mathrm{X}_i^2 = \sum_{i=1}^{n} \mathrm{W}_i$ has a $\text{Gamma}\left(\alpha = \frac{n}{2}, \theta = \psi\right)$ distribution.

Let Y be a random variable with a Gamma distribution with parameters α and θ. Then ${}^{2\mathrm{Y}}\!/_{\theta}$ has a chi-square distribution with $r = 2\alpha$ degrees of freedom.

$$\Rightarrow \quad \frac{2\,\mathrm{Y}}{\psi} \text{ has a } \chi^2(2\alpha = n) \text{ distribution.}$$

$$\Rightarrow \quad \mathrm{P}\left(\chi^2_{1-\alpha/2}(n) < \frac{2\,\mathrm{Y}}{\psi} < \chi^2_{\alpha/2}(n)\right) = 1-\alpha.$$

$$\Rightarrow \quad \mathrm{P}\left(\frac{2\,\mathrm{Y}}{\chi^2_{1-\alpha/2}(n)} > \psi > \frac{2\,\mathrm{Y}}{\chi^2_{\alpha/2}(n)}\right) = 1-\alpha.$$

A $(1-\alpha)\,100\,\%$ confidence interval for ψ is

$$\left(\frac{2\,\mathrm{Y}}{\chi^2_{\alpha/2}(n)}, \frac{2\,\mathrm{Y}}{\chi^2_{1-\alpha/2}(n)}\right) = \left(\frac{2\sum_{i=1}^{n}\mathrm{X}_i^2}{\chi^2_{\alpha/2}(n)}, \frac{2\sum_{i=1}^{n}\mathrm{X}_i^2}{\chi^2_{1-\alpha/2}(n)}\right).$$

(ii) $n = 4,$ $\sum_{i=1}^{n} x_i^2 = 0.2^2 + 0.6^2 + 1.1^2 + 1.7^2 = 4.5.$

$\chi^2_{0.975}(4) = 0.484,$ $\chi^2_{0.025}(4) = 11.14.$

$$\left(\frac{2\sum_{i=1}^{n} x_i^2}{\chi^2_{\alpha/2}(n)}, \frac{2\sum_{i=1}^{n} x_i^2}{\chi^2_{1-\alpha/2}(n)} \right) = \left(\frac{9}{11.14}, \frac{9}{0.484} \right) \approx \mathbf{(0.808, 18.595)}.$$

k) Find a sufficient statistic $u(X_1, X_2, \ldots, X_n)$ for ψ.

$$\prod_{i=1}^{n} f(x_i; \psi) = \prod_{i=1}^{n} \frac{2}{\sqrt{\pi\psi}} e^{-x_i^2/\psi} = \left(\frac{2}{\sqrt{\pi\psi}} \right)^n \cdot \exp\left\{ -\frac{1}{\psi} \sum_{i=1}^{n} x_i^2 \right\}.$$

By Factorization Theorem, $Y = \sum_{i=1}^{n} X_i^2$ is a sufficient statistic for ψ.

OR

$$f(x; \psi) = \exp\left\{ -\frac{1}{\psi} \cdot x^2 - \frac{1}{2} \cdot \ln\psi + \ln 2 - \frac{1}{2} \cdot \ln\pi \right\}.$$

$\Rightarrow$ $K(x) = x^2.$

$\Rightarrow$ $Y = \sum_{i=1}^{n} K(X_i) = \sum_{i=1}^{n} X_i^2$ is a sufficient statistic for ψ.

1) Recall that a method of moments estimator for ψ is $\tilde{\psi} = \pi\left(\overline{X}\right)^2$.

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

"Hint": ① By CLT, $\sqrt{n}\left(\overline{X} - \mu\right) \xrightarrow{D} N\left(0, \sigma^2\right)$.

② If $g(x)$ is differentiable at μ and $g'(\mu) \neq 0$, then

$$\sqrt{n}\left(g\left(\overline{X}\right) - g(\mu)\right) \xrightarrow{D} N\left(0, \left[g'(\mu)\right]^2 \sigma^2\right).$$

That is, for large n,

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

$$\mu = E(X) = \frac{\sqrt{\psi}}{\sqrt{\pi}}. \qquad E(X^2) = \frac{\psi}{2}.$$

$$\sigma^2 = \operatorname{Var}(X) = E(X^2) - [E(X)]^2 = \frac{\psi}{2} - \frac{\psi}{\pi} = \psi\left(\frac{\pi - 2}{2\pi}\right).$$

Consider $g(x) = \pi x^2$. Then $g(\overline{X}) = \tilde{\psi}$, $g(\mu) = \psi$.

$$g'(x) = 2\pi x. \qquad g'(\mu) = 2\sqrt{\pi\psi}.$$

$$\left[g'(\mu)\right]^2 \sigma^2 = 4\pi\psi \cdot \psi\left(\frac{\pi - 2}{2\pi}\right) = 2(\pi - 2)\psi^2.$$

$$\Rightarrow \quad \sqrt{n}\left(\tilde{\psi} - \psi\right) \xrightarrow{D} N\left(0, 2(\pi - 2)\psi^2\right).$$

For large n, $\tilde{\psi}$ is approximately $N\left(\psi, \dfrac{2(\pi - 2)\psi^2}{n}\right)$.

For fun:

m) Recall that the maximum likelihood estimator for ψ is $\hat{\psi} = \frac{2}{n} \sum_{i=1}^{n} X_i^2 = 2\,\overline{X^2}$.

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

$$\hat{\psi} = 2\,\overline{X^2} = 2\,\overline{W}.$$

$$\mu_W = E(W) = \alpha\theta = \frac{1}{2}\psi. \qquad \sigma_W^2 = Var(W) = \alpha\theta^2 = \frac{1}{2}\psi^2.$$

By CLT, $$\sqrt{n}\left(\overline{W} - \mu_W\right) \xrightarrow{D} N\left(0, \sigma_W^2\right).$$

$$\Rightarrow \qquad \sqrt{n}\left(2\overline{W} - 2\mu_W\right) \xrightarrow{D} N\left(0, 4\sigma_W^2\right)$$

$$\Rightarrow \qquad \sqrt{n}\left(\hat{\psi} - \psi\right) \xrightarrow{D} N\left(0, 2\psi^2\right).$$

For large n, $\hat{\psi}$ is approximately $N(\psi, \frac{2\psi^2}{n})$.

Since $\pi - 2 > 1$, $\hat{\psi}$ is "better" than $\tilde{\psi}$.

8. Let $\beta > 0$ be a population parameter, and let $X_1, X_2, \dots, X_n$ be a random sample from a distribution with probability density function

$$f(x;\beta) = \beta\,(1-x)^{\beta-1}, \qquad 0 < x < 1, \qquad \text{zero otherwise.}$$

Recall: $W = -\ln(1-X)$ has an $\text{Exponential}(\theta = \frac{1}{\beta})$

$= \text{Gamma}(\alpha = 1, \theta = \frac{1}{\beta})$ distribution.

j) (i) Suggest a confidence interval for β with $(1-\alpha)\,100\,\%$ confidence level.

(ii) Suppose $n = 3$, and $x_1 = 0.31$, $x_2 = 0.77$, $x_3 = 0.93$.

Construct a 90% confidence interval for β.

"Hint": Use $\sum_{i=1}^{n} \big(-\ln(1-X_i)\big)$.

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

Let Y be a random variable with a Gamma distribution with parameters α and $\theta = {}^{1}\!/_{\lambda}$.
Then ${}^{2Y}\!/_{\theta} = 2\lambda Y$ has a chi-square distribution with $r = 2\alpha$ degrees of freedom.

$\Rightarrow$ $2\beta Y$ has a $\chi^2(2\alpha = 2n)$ distribution.

$\Rightarrow$ $P\big(\chi^2_{1-\alpha/2}(2n) < 2\beta Y < \chi^2_{\alpha/2}(2n)\big) = 1-\alpha.$

$\Rightarrow$ $P\Big(\dfrac{\chi^2_{1-\alpha/2}(2n)}{2Y} < \beta < \dfrac{\chi^2_{\alpha/2}(2n)}{2Y}\Big) = 1-\alpha.$

A $(1-\alpha)\,100\,\%$ confidence interval for β is

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

(ii) $n = 3$, $\displaystyle\sum_{i=1}^{n}\left(-\ln(1-x_i)\right) = -\ln 0.69 - \ln 0.23 - \ln 0.07 \approx 4.5.$

$\chi^2_{0.95}(6) = 1.635$, $\chi^2_{0.05}(6) = 12.59$.

$$\left(\frac{\chi^2_{1-\alpha/2}(2n)}{2\sum_{i=1}^{n}\left(-\ln(1-x_i)\right)}, \frac{\chi^2_{\alpha/2}(2n)}{2\sum_{i=1}^{n}\left(-\ln(1-x_i)\right)} \right) \approx \left(\frac{1.635}{9}, \frac{12.59}{9} \right)$$

$$\approx \mathbf{(0.1817, 1.3989)}.$$

k) Find a sufficient statistic $u(\mathrm{X}_1, \mathrm{X}_2, \ldots, \mathrm{X}_n)$ for β.

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

By Factorization Theorem, $\mathrm{Y}_1 = \prod_{i=1}^{n}(1-\mathrm{X}_i)$ is a sufficient statistic for β.

$\Rightarrow$ $\mathrm{Y}_2 = \ln \mathrm{Y}_1 = \sum_{i=1}^{n} \ln(1-\mathrm{X}_i)$ is also a sufficient statistic for β.

$\Rightarrow$ $\mathrm{Y} = -\mathrm{Y}_2 = \sum_{i=1}^{n}\left(-\ln(1-\mathrm{X}_i)\right)$ is also a sufficient statistic for β.

OR

$$f(x;\beta) = \exp\left\{ (\beta-1)\cdot\ln(1-x) + \ln\beta \right\}.$$

$\Rightarrow \quad \mathrm{K}(x) = \ln(1-x).$

$\Rightarrow \quad \mathrm{Y}_2 = \sum_{i=1}^{n} \mathrm{K}\left(\mathrm{X}_i\right) = \sum_{i=1}^{n} \ln\left(1-\mathrm{X}_i\right)$ is a sufficient statistic for β.

$\Rightarrow \quad \mathrm{Y}_1 = e^{\mathrm{Y}_2} = \prod_{i=1}^{n} \left(1-\mathrm{X}_i\right)$ is also a sufficient statistic for β,

$\quad \mathrm{Y} = -\mathrm{Y}_2 = \sum_{i=1}^{n} \left(-\ln\left(1-\mathrm{X}_i\right)\right)$ is also a sufficient statistic for β.

l) Let $\mathrm{Y}_1 < \mathrm{Y}_2 < \ldots < \mathrm{Y}_n$ denote the corresponding order statistics.

$\Big[$ Proving that $\mathrm{Y}_1 \xrightarrow{P} 0$ is super easy, barely an inconvenience. $\Big]$

Find δ so that $\mathrm{V}_n = n^{\delta}\,\mathrm{Y}_1$ converges in distribution.

Find the limiting distribution of V_n.

"Hint":

① Use $\mathrm{F}_{\mathrm{X}}(x)$ to find the c.d.f. of Y_1, $\mathrm{F}_{\mathrm{Y}_1}(x) = \mathrm{F}_{\min \mathrm{X}_i}(x)$.

② Use $\mathrm{F}_{\mathrm{Y}_1}(x)$ to find the c.d.f. of V_n, $\mathrm{F}_{\mathrm{V}_n}(v) = \mathrm{P}(\mathrm{V}_n \le v)$.

③ $\mathrm{F}_{\infty}(v) = \lim_{n\to\infty} \mathrm{F}_{\mathrm{V}_n}(v)$. IF the limit exists and IF $\mathrm{F}_{\infty}(v)$ is a c.d.f. of a probability distribution, then that is the limiting distribution of V_n.

④ $\lim_{n\to\infty}\left(1+\dfrac{a}{n^{\mathbf{1}}}\right)^{n} = e^{a}$. Only "interesting" case is interesting.

$$\mathrm{F}_{\mathrm{X}}(x) = \int_0^x \beta\,(1-u)^{\beta-1}\,du = -(1-u)^{\beta}\,\Big|_0^x = 1-(1-x)^{\beta}, \qquad 0 < x < 1.$$

$$F_{Y_1}(x) = F_{\min X_i}(x) = 1 - \big(1 - F_X(x)\big)^n = 1 - (1-x)^{\beta n}, \qquad 0 < x < 1.$$

$\Big[$ Proving that $Y_1 \xrightarrow{P} 0$ (super easy, barely an inconvenience):

Let $\varepsilon > 0$.

If $0 < \varepsilon < 1$,

$$P(|Y_1 - 0| \ge \varepsilon) = P(Y_1 \le -\varepsilon) + P(Y_1 \ge \varepsilon) = 0 + 1 - F_{Y_1}(\varepsilon)$$

$$= (1-\varepsilon)^{\beta n} \to 0 \quad \text{as } n \to \infty.$$

If $\varepsilon \ge 1$, $\quad P(|Y_1 - 0| \ge \varepsilon) = 0. \Big]$

$$F_{V_n}(v) = P(V_n \le v) = P(Y_1 \le \frac{v}{n^\delta}) = 1 - \left(1 - \frac{v}{n^\delta}\right)^{\beta n}, \qquad 0 < v < n^\delta.$$

If $\delta = 1$,

$$F_{V_n}(v) = 1 - \left(1 - \frac{v}{n}\right)^{\beta n}, \qquad 0 < v < n.$$

$$F_\infty(v) = \lim_{n\to\infty} F_{V_n}(v) = 1 - e^{-\beta v}, \qquad 0 < v < \infty.$$

The limiting distribution of V_n is Exponential with mean $\frac{1}{\beta}$.

For fun: Only "interesting" case is interesting. However, …

If $\delta < 1$,

$$F_\infty(v) = \lim_{n\to\infty} F_{V_n}(v) = 1, \qquad 0 < v < \infty.$$

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

If $\delta > 1$,

$$F_\infty(v) = \lim_{n\to\infty} F_{V_n}(v) = 0, \qquad 0 < v < \infty.$$

Then V_n does not have a limiting distribution.

For fun:

m) $\left[\text{Proving that } Y_n \overset{P}{\to} 1 \text{ is super easy, barely an inconvenience.}\right]$

Find γ so that $W_n = n^{\gamma}(1 - Y_n)$ converges in distribution.

Find the limiting distribution of W_n.

$$F_{Y_n}(x) = F_{\max X_i}(x) = \left(F_X(x)\right)^n = \left(1 - (1-x)^{\beta}\right)^n, \qquad 0 < x < 1.$$

Let $\varepsilon > 0$.

If $0 < \varepsilon < 1$,

$$P(|Y_n - 1| \ge \varepsilon) = P(Y_n \le 1 - \varepsilon) + P(Y_n \ge 1 + \varepsilon) = F_{Y_n}(1-\varepsilon) + 0$$

$$= \left(1 - \varepsilon^{\beta}\right)^n \to 0 \quad \text{as } n \to \infty.$$

If $\varepsilon \ge 1$, $\qquad P(|Y_n - 1| \ge \varepsilon) = 0.$

$$\Rightarrow \qquad Y_n \overset{P}{\to} 1.$$

$$F_{W_n}(w) = P(W_n \le w) = P(Y_n \ge 1 - \frac{w}{n^{\beta}}) = 1 - \left(1 - \left(\frac{w}{n^{\gamma}}\right)^{\beta}\right)^n$$

$$= 1 - \left(1 - \frac{w^{\beta}}{n^{\gamma\beta}}\right)^n, \qquad 0 < w < n^{\gamma}.$$

If $\gamma = \frac{1}{\beta}$, $\qquad F_{W_n}(w) = 1 - \left(1 - \frac{w^{\beta}}{n}\right)^n, \qquad 0 < w < n^{\gamma}.$

$$F_{\infty}(w) = \lim_{n\to\infty} F_{W_n}(w) = 1 - e^{-w^{\beta}}, \qquad 0 < w < \infty.$$

The limiting distribution is a Weibull distribution.

Only "interesting" case is interesting. However, …

If $\gamma < \frac{1}{\beta}$, $\quad F_{\infty}(w) = \lim_{n\to\infty} F_{W_n}(w) = 1, \qquad 0 < w < \infty.$

Then $W_n \overset{D}{\to} 0$, and thus $W_n \overset{P}{\to} 0$.

If $\gamma > \frac{1}{\beta}$, $\quad F_{\infty}(w) = \lim_{n\to\infty} F_{W_n}(w) = 0, \qquad 0 < w < \infty.$

Then W_n does not have a limiting distribution.