DISCUSSION WORKSHEET 03

PSTAT 120B: Mathematical Statistics, I
Summer Session A, 2024 with Instructor: Ethan P. Marzban

Conceptual Review

(a) What is a **bivariate transformation**? How can we find densities of bivariate transformations of random variables?

(b) What is an **order statistic**?

(c) How do we find the densities of **sample minima** and **sample maxima**?

Problem 1: Minimalism

Let $Y_1, Y_2, \cdots \overset{\text{i.i.d.}}{\sim}$ Pareto(θ, α). Recall from Midterm 01 that this means each of the Y_i's have the following density:

$$f_Y(y) = \frac{\alpha \theta^\alpha}{y^{\alpha+1}} \cdot \mathbb{1}_{\{y \geq \theta\}}$$

(a) Derive an expression for $\overline{F_Y}(y)$, the survival function of the Pareto(θ, α) distribution.

Solution: By definition, $\overline{F_Y}(y) := 1 - F_Y(y) = \mathbb{P}(Y > y)$. Hence, since the support of Y is $S_Y = [\theta, \infty)$, we see that $\overline{F_Y}(y) = 1$ for any $y < \theta$. As such, fix a $y \geq \theta$ so that

$$\begin{aligned} \overline{F_Y}(y) = \mathbb{P}(Y > y) &= \int_y^\infty f_Y(t) \, \mathrm{d}t \\ &= \alpha\theta^\alpha \cdot \int_y^\infty t^{-\alpha-1} \, \mathrm{d}t \\ &= \cancel{\alpha}\theta^\alpha \cdot \frac{1}{\cancel{\alpha}} \left[t^{-\alpha}\right]_{t=\infty}^{t=y} = \frac{\theta^\alpha}{y^\alpha} = \left(\frac{\theta}{y}\right)^\alpha \end{aligned}$$

Therefore, putting everything together,

$$\overline{F_Y}(y) = \begin{cases} 1 & \text{if } y < \theta \\ \left(\frac{\theta}{y}\right)^\alpha & \text{if } y \geq \theta \end{cases}$$

(b) Find the density of $Y_{(1)} := \min_{1 \leq i \leq n}\{Y_i\}$, the first order statistic. Use this to show that $Y_{(1)}$ follows a Pareto distribution, and identify the parameters.

Solution: We have a formula for the density of $Y_{(1)}$, that was derived in lecture:

$$\begin{aligned} f_{Y_{(1)}}(y) &= n[\overline{F_Y}(y)]^{n-1} \cdot f_Y(y) \\ &= n \left[\left(\frac{\theta}{y} \right)^{\alpha} \right]^{n-1} \cdot \frac{\alpha \theta^{\alpha}}{y^{\alpha+1}} \cdot \mathbb{1}_{\{y \geq \theta\}} \\ &= n \cdot \frac{\theta^{n\alpha - \alpha}}{y^{n\alpha - \alpha}} \cdot \alpha \theta^{\alpha} \cdot \frac{1}{y^{\alpha+1}} \cdot \mathbb{1}_{\{y \geq \theta\}} \\ &= \frac{\alpha n \theta^{\alpha n}}{y^{n\alpha+1}} \cdot \mathbb{1}_{\{y \geq \theta\}} = \boxed{\frac{(n\alpha)\theta^{(n\alpha)}}{y^{(n\alpha)+1}} \cdot \mathbb{1}_{\{y \geq \theta\}}} \end{aligned}$$

which we recognize as the density of the Pareto distribution with parameters θ and $n\alpha$; that is,

$$\boxed{Y_{(1)} \sim \text{Pareto}(\theta, n\alpha)}$$

Problem 2: Drink Up!

GauchoPop, the hit new soda brand, has implemented a new state-of-the-art bottling machine in its factory. The amount of soda (in oz) dispensed by the machine into a bottle is normally distributed about some mean μ and with some variance σ^2. As a notational aid, you may let $F_{\chi^2_\nu}(x)$ denote the CDF of the χ^2_ν distribution evaluated at the point x.

Hint: Use previously-derived results wherever possible!

(a) Assume $\sigma^2 = 8.2$. Compute the probability that the average amount of fill in a random sample of 25 *GauchoPop* bottles lies within 1 oz of μ, the true average amount of soda dispensed by the machine.

Solution: Let Y_i denote the amount of fill (in oz) in a randomly-selected *GauchoPop* bottle of soda; then, by the problem statement, $Y_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, 8.2)$. Define the sample mean amount of fill to be

$$\overline{Y}_{25} := \frac{1}{25}\sum_{i=1}^{n} Y_i$$

From results discussed in lecture, we know that

$$\overline{Y}_{25} \sim \mathcal{N}\left(\mu\,,\ \frac{8.2}{25}\right) \sim \mathcal{N}(\mu, 0.328)$$

We seek $\mathbb{P}(|\overline{Y}_{25} - \mu| < 1)$, which we can compute using standardization:

$$\begin{aligned} \mathbb{P}(|\overline{Y}_{25} - \mu| < 1) &= \mathbb{P}(-1 < \overline{Y}_{25} - \mu < 1) \\ &= \mathbb{P}\left(-\frac{1}{\sqrt{0.328}} < \frac{\overline{Y}_{25} - \mu}{\sqrt{0.328}} < \frac{1}{\sqrt{0.328}}\right) \\ &= 2\Phi\left(\frac{1}{\sqrt{0.328}}\right) - 1 \approx 0.9192 = 91.92\% \end{aligned}$$

(b) Again assume $\sigma^2 = 8.2$. Suppose a random sample of 10 *GauchoPop* bottles is taken, and the sample variance of the amount of fill in these 10 bottles is recorded. What is the probability that this sample variance lies between 8oz and 8.5oz?

Solution: Again let Y_i denote the amount of fill (in oz) in a randomly-selected *GauchoPop* bottle of soda. Define the sample mean of $\{Y_i\}_{i=1}^{10}$ as

$$\overline{Y}_{10} := \frac{1}{10}\sum_{i=1}^{n} Y_i$$

and define the sample variance to be

$$S^2_{10} := \frac{1}{10-1}\sum_{i=1}^{10}(Y_i - \overline{Y}_{10})^2 = \frac{1}{9}\sum_{i=1}^{10}(Y_i - \overline{Y}_{10})^2$$

We know that, in general,

$$\frac{n-1}{\sigma^2} S_n^2 \sim \chi_{n-1}^2$$

meaning, plugging in $n = 10$ and $\sigma^2 = 8.2$,

$$\frac{9}{8.2} S_{10}^2 \sim \chi_9^2$$

Hence, since we seek $\mathbb{P}(8 \leq S_{10}^2 \leq 8.5)$, we compute

$$\begin{aligned} \mathbb{P}(8 \leq S_{10}^2 \leq 8.5) &= \mathbb{P}\left(\frac{9}{8.2} \cdot 8 \leq \frac{9}{8.2} \cdot S_{10}^2 \leq \frac{9}{8.2} \cdot 8.5 \right) \\ &= \mathbb{P}\left(\frac{360}{41} \leq \frac{9}{8.2} \cdot S_{10}^2 \leq \frac{765}{82} \right) \\ &= F_{\chi_9^2}\left(\frac{765}{82} \right) - F_{\chi_9^2}\left(\frac{360}{41} \right) \end{aligned}$$

If you're curious, we can use a computer software to compute this numerically; the final answer ends up being around $0.0503 \approx 5.03\%$.