STAT 30100 : Mathematical Statistics-1

Winter 2020

•

Homework 5

•

Solutions by

Jinhong Du

12243476

STAT 30100, Homework 5

1. (Casella and Berger Problem 5.48) Using strategies similar to (5.6.5), show how to generate an $F_{m,n}$ random variable, where both m and n are even integers.

If U_i, V_j are iid Uniform$(0,1)$ random variables, then $X_i = -\lambda\log(U_i), Y_j = -\lambda\log(V_j)$ are iid Exp(λ) random variables for $i = 1, \ldots, \frac{m}{2}$ and $j = 1, \ldots, \frac{n}{2}$, and

$$X = -2\sum_{i=1}^{\frac{m}{2}}\log(U_i) \sim \chi^2_m$$

$$Y = -2\sum_{j=1}^{\frac{n}{2}}\log(V_j) \sim \chi^2_n.$$

Since X and Y are independent, we have

$$\frac{\frac{X}{m}}{\frac{Y}{n}} \sim F_{m,n},$$

which can be used to generate $F_{m,n}$ from uniform random variables U_i and V_j.

2. (Casella and Berger Problem 5.25) Let $X_1, \ldots, X_n$ be iid with pdf

$$f_X(x) = \begin{cases} \frac{a}{\theta^a}x^{a-1} & \text{if } 0 < x < \theta \\ 0 & \text{otherwise} \end{cases}$$

Let $X_{(1)} < \ldots < X_{(n)}$ be the order statistics. Show that $\frac{X_{(1)}}{X_{(2)}}, \frac{X_{(2)}}{X_{(3)}}, \ldots, \frac{X_{(n-1)}}{X_{(n)}}$ and $X_{(n)}$ are mutually independent random variables. Find the distribution of each of them.

Proof. The joint distribution of $(X_{(1)}, \ldots, X_{(n)})$ is given by

$$f_{(X_{(1)},\ldots,X_{(n)})}(x_1, \ldots, x_n) = n!\prod_{i=1}^n f_X(x_i)\mathbb{1}_{\{x_1 \le \cdots \le x_n\}} = n!\left(\frac{a}{\theta^a}\right)^n \prod_{i=1}^n x_i^{a-1}\mathbb{1}_{\{0<x_1\le\cdots\le x_n<\theta\}}.$$

The determinant of the Jacobian matrix of the transfromation $(y_1, \ldots, y_n) = g(x_1, \ldots, x_n)$ from $(X_{(1)}, \ldots, X_{(n)})$ to $\left(\frac{X_{(1)}}{X_{(2)}}, \ldots, \frac{X_{(n-1)}}{X_{(n)}}, X_{(n)}\right)$ is given by

$$\left|\begin{bmatrix} \frac{1}{x_2} & & & \\ -\frac{x_1}{x_2^2} & \ddots & & \\ & \ddots & \frac{1}{x_n} & \\ & & -\frac{x_{n-1}}{x_n^2} & 1 \end{bmatrix}\right| = \left|\begin{bmatrix} \frac{1}{x_2} & & & \\ & \ddots & & \\ & & \frac{1}{x_n} & \\ & & & 1 \end{bmatrix}\right| = \prod_{i=2}^n \frac{1}{x_i},$$

when $x_2, \ldots, x_n \neq 0$. The determinant of the Jacobian of the inverse transfromation is $\prod_{i=2}^n x_i = \prod_{i=2}^n y_i^{i-1}$. Therefore, the density function of $\left(\frac{X_{(1)}}{X_{(2)}}, \ldots, \frac{X_{(n-1)}}{X_{(n)}}, X_{(n)}\right)$ is given by

$$\begin{aligned} &f_{\left(\frac{X_{(1)}}{X_{(2)}},\ldots,\frac{X_{(n-1)}}{X_{(n)}},X_{(n)}\right)}(y_1,\ldots,y_n) = f_{(X_{(1)},\ldots,X_{(n)})}(y_1\cdots y_n, \ldots, y_{n-1}y_n, y_n)\left|\prod_{i=2}^n y_i^{i-1}\right| \\ =& n!\left(\frac{a}{\theta^a}\right)^n\prod_{i=1}^n y_i^{i(a-1)}\left|\prod_{i=2}^n y_i^{i-1}\right|\mathbb{1}_{\{0<y_1y_2\le\cdots\le y_{n-1}y_n\le y_n<\theta\}} = \left(\frac{na}{\theta^{na}}y_n^{na-1}\mathbb{1}_{0<y_n<\theta}\right)\prod_{i=1}^{n-1}\left(iay_i^{ia-1}\mathbb{1}_{0<y_i\le 1}\right). \end{aligned}$$

Solution (cont.)

Since the above joint density is separable for each variable, $\frac{X_{(1)}}{X_{(2)}}, \ldots, \frac{X_{(n-1)}}{X_{(n)}}, X_{(n)}$ are independent and their density functions are given by

$$f_{\frac{X_{(i)}}{X_{(i+1)}}}(y) = iay^{ia-1}\mathbb{1}_{0<y\leq 1}, \quad i = 1, \ldots n-1, \qquad f_{X_{(n)}}(y) = \frac{na}{\theta^{na}} y^{na-1}\mathbb{1}_{0<y<\theta}.$$

$\square$

3. Let $Y_1, \ldots, Y_{n+1}$ be i.i.d. exponential random variables with mean 1, and let $S_j = \sum_{i=1}^{j} Y_i$ for $i = 1, \ldots, n+1$. Let $W = \left(\frac{S_1}{S_{n+1}}, \ldots, \frac{S_n}{S_{n+1}}\right)$. Find the joint distribution of (W, S_{n+1}). In particular, show that W and S_{n+1} are independent, and W has the same distribution as the vector of order statistics of an i.i.d. sample of size n from the $U(0,1)$ distribution.

Proof. Let $g(y_1, \ldots, y_{n+1}) = \left(\frac{s_1}{s_{n+1}}, \ldots, \frac{s_n}{s_{n+1}}, s_{n+1}\right)$ where $s_j = \sum_{i=1}^{j} y_j$ which maps $(Y_1, \ldots, Y_{n+1})$ to $\left(\frac{S_1}{S_{n+1}}, \ldots, \frac{S_n}{S_{n+1}}, S_{n+1}\right)$.
The determinant of the Jacobian of g is given by

$$\begin{aligned}
J &= \left|\begin{bmatrix}
\frac{s_{n+1}-s_1}{s_{n+1}^2} & -\frac{s_1}{s_{n+1}^2} & -\frac{s_1}{s_{n+1}^2} & \cdots & -\frac{s_1}{s_{n+1}^2} \\
\frac{s_{n+1}-s_2}{s_{n+1}^2} & \frac{s_{n+1}-s_2}{s_{n+1}^2} & -\frac{s_2}{s_{n+1}} & \cdots & -\frac{s_2}{s_{n+1}^2} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
\frac{s_{n+1}-s_n}{s_{n+1}^2} & \frac{s_{n+1}-s_n}{s_{n+1}^2} & \cdots & \frac{s_{n+1}-s_n}{s_{n+1}^2} & -\frac{s_n}{s_{n+1}^2} \\
1 & 1 & 1 & \cdots & 1
\end{bmatrix}\right| \\
&= \left|\begin{bmatrix}
0 & -\frac{1}{s_{n+1}} & -\frac{1}{s_{n+1}} & \cdots & -\frac{1}{s_{n+1}} \\
0 & 0 & -\frac{1}{s_{n+1}} & \cdots & -\frac{1}{s_{n+1}} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & 0 & -\frac{1}{s_{n+1}} \\
1 & 1 & 1 & \cdots & 1
\end{bmatrix}\right| = -\left|\begin{bmatrix}
1 & 1 & 1 & \cdots & 1 \\
0 & -\frac{1}{s_{n+1}} & -\frac{1}{s_{n+1}} & \cdots & -\frac{1}{s_{n+1}} \\
0 & 0 & -\frac{1}{s_{n+1}} & \cdots & -\frac{1}{s_{n+1}} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & 0 & -\frac{1}{s_{n+1}}
\end{bmatrix}\right| \\
&= -\frac{1}{s_{n+1}^n}.
\end{aligned}$$

Therefore, for $0 < x_1 < \ldots < x_n < 1, x_{n+1} > 0$, we have

$$\begin{aligned}
f_{\left(\frac{S_1}{S_{n+1}}, \ldots, \frac{S_n}{S_{n+1}}, S_{n+1}\right)}(x_1, \ldots, x_{n+1}) &= f_{Y_1}(x_1 x_{n+1}) \left[\prod_{i=2}^{n} f_{Y_i}(x_i x_{n+1} - x_{i-1}x_{n+1})\right] \\
&\quad \cdot f_{Y_{n+1}}(x_{n+1} - x_n x_{n+1}) \cdot x_{n+1}^n \\
&= x_{n+1}^n e^{-x_1 x_{n+1} - \sum_{i=2}^{n}(x_i x_{n+1} - x_{i-1}x_{n+1}) - (x_{n+1} - x_n x_{n+1})} \\
&= n! \cdot \frac{1}{n!} x_{n+1}^n e^{-x_{n+1}}.
\end{aligned}$$

Since $\int_0^{+\infty} \frac{1}{n!} x_{n+1}^n e^{-x_{n+1}} \mathrm{d}x_{n+1} = 1$, we know that $S_{n+1} \sim \Gamma(n+1, 1)$. So $f_W(\boldsymbol{w}) = n!$, i.e., W has the same distribution as the vector of order statistics of an i.i.d. sample of size n from the $U(0,1)$ distribution. As $f_{\left(\frac{S_1}{S_{n+1}}, \ldots, \frac{S_n}{S_{n+1}}, S_{n+1}\right)}(x_1, \ldots, x_{n+1}) = f_W(\boldsymbol{w}) f_{S_{n+1}}(x_{n+1})$, we have that W and S_{n+1} are independent. $\square$

4. Let $0 < p_1 < \ldots < p_k < 1$, and let $X_{(\lceil npi \rceil)}$ be the corresponding sample quantiles (as defined in Ferguson p. 87) for a sample of size n from a distribution with location parameter θ having distribution function $F(x-\theta)$ and density $f(x-\theta)$. Let u_i denote the p_ith quantile of F (i.e. $F(u_i) = p_i$).

   (a) Let $Z_i = X_{(\lceil npi \rceil)} - u_i$. Let $\boldsymbol{Z}$ represent the vector $(Z_1, \ldots, Z_k)^\top$ and $\mathbf{1}$ represent the k-vector of all 1's. Show that $\sqrt{n}(\boldsymbol{Z} - \theta\mathbf{1})$ converges in distribution to $\mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$, where Σ is the symmetric matrix with components $\sigma_{ij} = \frac{p_i(1-p_j)}{f(u_i)f(u_j)}$ for $i \le j$.

Proof. Let $U_{(i)} = F(X_{(i)} - \theta)$ for $i = 1, \ldots, n$. Then $U_{(1)}, \ldots, U_{(n)}$ are the order statistics of $U_1, \ldots, U_n \overset{iid}{\sim} U(0,1)$. From Problem 3, we have $U_{(i)} = \frac{S_i}{S_{n+1}}$ where $S_i = \sum_{j=1}^{i} Y_j$ and $Y_1, \ldots, Y_{n+1} \overset{iid}{\sim} \text{Exp}(1)$.

Let $n_i = \lceil np_i \rceil$ for $i = 1, \ldots, k$ and let $n_0 = 0$, $p_0 = 0$, $S_0 = 0$, $n_{k+1} = n$, and $p_{k+1} = 1$. For $i = 1, \ldots, k+1$, since $S_{n_i} - S_{n_{i-1}} \sim \Gamma(n_i - n_{i-1}, 1)$, by Central Limit Theorem, we have $\sqrt{n_i - n_{i-1}} \frac{(S_{n_i} - S_{n_{i-1}}) - (n_i - n_{i-1})}{n_i - n_{i-1}} \xrightarrow{D} \mathcal{N}(0,1)$ as $n_i - n_{i-1} \to \infty$. Since $\lim_{n\to\infty} \sqrt{\frac{n_i - n_{i-1}}{n+1}} = \sqrt{p_i - p_{i-1}}$ and $\lim_{n\to\infty} \sqrt{\frac{n+1}{n}} = 1$, by Slutsky Theorem, we have $\sqrt{n} \frac{(S_{n_i} - S_{n_{i-1}}) - (n_i - n_{i-1})}{n+1} = \sqrt{n+1} \sqrt{\frac{n}{n+1}} \sqrt{\frac{n_i - n_{i-1}}{n+1}} \sqrt{n_i - n_{i-1}} \frac{(S_{n_i} - S_{n_{i-1}}) - (n_i - n_{i-1})}{n_i - n_{i-1}} \xrightarrow{D} \mathcal{N}(0, p_i - p_{i-1})$. As $\lim_{n\to\infty} \sqrt{n} \frac{(n_i - n_{i-1}) - (np_i - np_{i-1})}{n+1} = 0$, by Slutsky Theorem, we have $\sqrt{n+1}\left(\frac{(S_{n_i} - S_{n_{i-1}})}{n+1} - (p_i - p_{i-1})\right) \xrightarrow{D} \mathcal{N}(0, p_i - p_{i-1})$ for $i = 1, \ldots, k+1$. Since $S_1 - S_0, \ldots, S_{k+1} - S_k$ are independent, we have

$$\sqrt{n} \begin{bmatrix} \frac{S_{n_1}}{n+1} - p_1 \\ \frac{S_{n_2} - S_{n_1}}{n+1} - (p_2 - p_1) \\ \vdots \\ \frac{S_{n+1} - S_{n_k}}{n+1} - (1 - p_k) \end{bmatrix} \xrightarrow{D} \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} p_1 & & & \\ & p_2 - p_1 & & \\ & & \ddots & \\ & & & 1 - p_k \end{bmatrix}\right).$$

Let $g(x_1, \ldots, x_{k+1}) = \frac{1}{\sum_{i=1}^{k+1} x_i} \begin{bmatrix} x_1 & x_1 + x_2 & \cdots & \sum_{i=1}^{k} x_i \end{bmatrix}^\top$, then

$$\frac{\partial g(\boldsymbol{x})}{\partial \boldsymbol{x}} = \frac{1}{\left(\sum_{i=1}^{k+1} x_i\right)^2} \begin{bmatrix} \sum_{i=2}^{k+1} x_i & -x_1 & -x_1 & \cdots & -x_1 & -x_1 \\ \sum_{i=3}^{k+1} x_i & \sum_{i=3}^{k+1} x_i & -\sum_{i=1}^{2} x_i & \cdots & -\sum_{i=1}^{2} x_i & -\sum_{i=1}^{2} x_i \\ \vdots & \vdots & \vdots & \ddots & \cdots & \vdots \\ \sum_{i=k}^{k+1} x_i & \sum_{i=k}^{k+1} x_i & \sum_{i=k}^{k+1} x_i & \cdots & \sum_{i=k}^{k+1} x_i & -\sum_{i=1}^{k-1} x_i \\ x_{k+1} & x_{k+1} & x_{k+1} & \cdots & x_{k+1} & -\sum_{i=1}^{k} x_i \end{bmatrix},$$

which is continuous near $\boldsymbol{\mu} = [p_1, p_2 - p_1, \ldots, 1 - p_k]^\top$. Then,

$$\left.\frac{g(\boldsymbol{x})}{\partial \boldsymbol{x}}^\top\right|_{\boldsymbol{x}=\boldsymbol{\mu}} = \begin{bmatrix} 1 - p_1 & -p_1 & -p_1 & \cdots & -p_1 & -p_1 \\ 1 - p_2 & 1 - p_2 & -p_2 & \cdots & -p_2 & -p_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 - p_k & 1 - p_k & 1 - p_k & \cdots & 1 - p_k & -p_k \end{bmatrix}$$

$$\left.\frac{\partial g(\boldsymbol{x})}{\partial \boldsymbol{x}} \operatorname{diag}(p_1, p_2 - p_1, \ldots, 1 - p_k) \frac{\partial g(\boldsymbol{x})}{\partial \boldsymbol{x}}^\top\right|_{\boldsymbol{x}=\boldsymbol{\mu}} = \begin{bmatrix} p_1(1-p_1) & p_1(1-p_2) & \cdots & p_1(1-p_k) \\ p_1(1-p_2) & p_2(1-p_2) & \cdots & p_2(1-p_k) \\ \vdots & \vdots & \ddots & \vdots \\ p_1(1-p_k) & p_2(1-p_k) & \cdots & p_k(1-p_k) \end{bmatrix}.$$

Solution (cont.)

Therefore, by Cramer's Theorem, we have

$$\sqrt{n}\left(\begin{bmatrix} U_{(\lceil np_1 \rceil)} \\ \vdots \\ U_{(\lceil np_k \rceil)} \end{bmatrix} - \begin{bmatrix} p_1 \\ \vdots \\ p_k \end{bmatrix}\right) \xrightarrow{D} \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} p_1(1-p_1) & p_1(1-p_2) & \cdots & p_1(1-p_k) \\ p_1(1-p_2) & p_2(1-p_2) & \cdots & p_2(1-p_k) \\ \vdots & \vdots & \ddots & \vdots \\ p_1(1-p_k) & p_2(1-p_k) & \cdots & p_k(1-p_k) \end{bmatrix}\right).$$

Notice that $X_{\lceil np_i \rceil} - \theta = F^{-1}(U_{\lceil np_i \rceil})$ and $u_i = F^{-1}(p_i)$. Let $h(x_1, \dots, x_k) = (F^{-1}(x_1), \dots, F^{-1}(x_k))^\top$, then $\frac{\partial h(\boldsymbol{x})}{\partial \boldsymbol{x}} = \operatorname{diag}\left(\frac{1}{f[F^{-1}(x_1)]}, \cdots \frac{1}{f[F^{-1}(x_k)]}\right)$. If f is continuous and positive near $p_1, \dots, p_k$, then $\frac{\partial h}{\partial \boldsymbol{x}}$ is continuous near $\begin{bmatrix} p_1, \cdots, p_k \end{bmatrix}^\top$. Then by Cramer's Theorem, we have

$$\sqrt{n}(\boldsymbol{Z} - \theta\mathbf{1}) \xrightarrow{D} \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}),$$

where $\boldsymbol{\Sigma}$ is the symmetric matrix with components $\sigma_{ij} = \frac{p_i(1-p_j)}{f(u_i)f(u_j)}$ for $i \leq j$. □

(b) Find the asymptotic best linear unbiased estimate of θ based on $\boldsymbol{Z}$. That is, for $\theta = \boldsymbol{a}^\top \boldsymbol{Z}$, find $\boldsymbol{a}$ to minimize $\boldsymbol{a}^\top \boldsymbol{\Sigma} \boldsymbol{a}$ subject to $\mathbf{1}^\top \boldsymbol{a} = 1$ (in terms of $\boldsymbol{\Sigma}^{-1}$).

Proof. Let $L(\boldsymbol{\alpha}) = \boldsymbol{a}^\top \boldsymbol{\Sigma} \boldsymbol{a} + \lambda(1 - \mathbf{1}^\top \boldsymbol{a})$. Let

$$\frac{\partial L}{\partial \boldsymbol{a}} = 2\Sigma \boldsymbol{a} - \lambda \mathbf{1} = \mathbf{0} \tag{1}$$

$$\frac{\partial L}{\partial \lambda} = 1 - \mathbf{1}^\top \boldsymbol{a} = 0, . \tag{2}$$

From (1) we have $\boldsymbol{a} = \frac{\lambda}{2} \boldsymbol{\Sigma}^{-1} \mathbf{1}$. Substituting it in to (2), we get $\lambda = \frac{2}{\mathbf{1}^\top \boldsymbol{\Sigma}^{-1} \mathbf{1}}$. Therefore,

$$\boldsymbol{a} = \frac{\boldsymbol{\Sigma}^{-1} \mathbf{1}}{\mathbf{1}^\top \boldsymbol{\Sigma}^{-1} \mathbf{1}}.$$

□

(c) In view of (b), it is comforting to know that the inverse of $\boldsymbol{\Sigma}$ has a simple form. It is a tridiagonal matrix. Find it.

Proof. Since $\boldsymbol{\Sigma}$ is symmetric, $\boldsymbol{\Sigma}^{-1}$ is also symmetric. Assume that $\boldsymbol{\Sigma}^{-1} = \begin{bmatrix} c_1 & d_1 & & & \\ d_1 & c_2 & d_2 & & \\ & \ddots & \ddots & \ddots & \\ & & d_{k-2} & c_{k-1} & d_{k-1} \\ & & & d_{k-1} & c_k \end{bmatrix}$. Since $\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-1} = \boldsymbol{I}_k$, we have $\forall\ i, j$,

$$\frac{p_i(1-p_{j-1})}{f(u_i)f(u_{j-1})} d_{j-1} + \frac{p_i(1-p_j)}{f(u_i)f(u_j)} c_j + \frac{p_i(1-p_{j+1})}{f(u_i)f(u_{j+1})} d_j = 0, \qquad i < j \tag{3}$$

$$\frac{p_{j-1}(1-p_j)}{f(u_j)f(u_{j-1})} d_{j-1} + \frac{p_j(1-p_j)}{f(u_j)f(u_j)} c_j + \frac{p_j(1-p_{j+1})}{f(u_j)f(u_{j+1})} d_j = 1, \qquad i = j \tag{4}$$

Solution (cont.)

where $d_0 = d_k = 0$, $p_0 = 0$, $p_{k+1} = 1$, u_0 and u_{k+1} are any constants such that $f(u_0), f(u_{k+1}) > 0$. From (3), we have for $j = 2, \dots, k$,

$$\frac{1-p_j}{f(u_j)}c_j + \frac{1-p_{j+1}}{f(u_{j+1})}d_j = -\frac{1-p_{j-1}}{f(u_{j-1})}d_{j-1}. \tag{5}$$

Substituting it in (4), we have

$$\frac{p_{j-1}-p_j}{f(u_j)f(u_{j-1})}d_{j-1} = 1,\ j = 2,\dots,k \qquad \Longrightarrow \qquad d_j = \frac{f(u_{j+1})f(u_j)}{p_j - p_{j+1}},\ \forall\ j \le k-1.$$

Substituing it in to (5), we have

$$c_j = \frac{f^2(u_j)(p_{i+1}-p_{i-1})}{(p_{j-1}-p_j)(p_j-p_{j+1})} = f^2(u_j)\left(\frac{1}{p_j - p_{j-1}} + \frac{1}{p_{j+1}-p_j}\right),\ \forall\ j.$$

Therefore,

$$\boldsymbol{\Sigma}^{-1} = \begin{bmatrix} \frac{f^2(u_1)p_2}{p_1(p_2-p_1)} & \frac{f(u_1)f(u_2)}{p_1-p_2} & & \\ \frac{f(u_1)f(u_2)}{p_1-p_2} & \frac{f^2(u_2)(p_3-p_1)}{(p_3-p_2)(p_2-p_1)} & & \\ & \ddots & \ddots & \ddots \\ & \frac{f(u_{k-2})f(u_{k-1})}{p_{k-2}-p_{k-1}} & \frac{f^2(u_{k-1})(p_k-p_{k-2})}{(p_k-p_{k-1})(p_{k-1}-p_{k-2})} & \frac{f(u_{k-1})f(u_k)}{p_{k-1}-p_k} \\ & & \frac{f(u_{k-1})f(u_k)}{p_{k-1}-p_k} & \frac{f^2(u_k)(1-p_{k-1})}{(1-p_k)(p_k-p_{k-1})} \end{bmatrix}.$$

□

(d) Find $\hat{\theta}$ of (b) explicitly, for the uniform distribution, $F(x) = x$ for $0 \le x \le 1$.

Proof. For uniform distribution, we have $f(x) = \mathbb{1}_{[0,1]}$. Therefore,

$$\boldsymbol{\Sigma}^{-1} = \begin{bmatrix} \frac{p_2}{p_1(p_2-p_1)} & \frac{1}{p_1-p_2} & & \\ \frac{1}{p_1-p_2} & \frac{p_3-p_1}{(p_3-p_2)(p_2-p_1)} & & \\ & \ddots & \ddots & \ddots \\ & \frac{1}{p_{k-2}-p_{k-1}} & \frac{p_k-p_{k-2}}{(p_k-p_{k-1})(p_{k-1}-p_{k-2})} & \frac{1}{p_{k-1}-p_k} \\ & & \frac{1}{p_{k-1}-p_k} & \frac{1-p_{k-1}}{(1-p_k)(p_k-p_{k-1})} \end{bmatrix}.$$

We have

$$\boldsymbol{\Sigma}^{-1}\mathbf{1} = \begin{bmatrix} \frac{1}{p_1} \\ 0 \\ \vdots \\ 0 \\ \frac{1}{1-p_k} \end{bmatrix}, \qquad \mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1} = \frac{1}{p_1} + \frac{1}{1-p_k} = \frac{1+p_1-p_k}{p_1(1-p_k)}, \qquad \boldsymbol{a} = \begin{bmatrix} \frac{1-p_k}{1+p_1-p_k} \\ 0 \\ \vdots \\ 0 \\ \frac{p_1}{1+p_1-p_k} \end{bmatrix}$$

and therefore $\hat{\boldsymbol{\theta}} = \boldsymbol{a}^\top \boldsymbol{Z} = \frac{1-p_k}{1+p_1-p_k}Z_1 + \frac{p_1}{1+p_1-p_k}Z_k$. □

5. Suppose X has the $G_{1,\gamma}(x)$ distribution, and let $Y = \gamma(X-1)$. Show that as $\gamma \to \infty$, Y converges in distribution to a random variable having the G_3 distribution. Here,

$$G_{1,\gamma}(x) = \begin{cases} e^{-x^{-\gamma}} & , \text{ for } x > 0 \\ 0 & , \text{ for } x \le 0 \end{cases}, \qquad G_3(x) = e^{-e^{-x}}.$$

Proof.

$$F_Y(y) = \mathbb{P}(Y \le y) = \mathbb{P}\left(X \le \frac{y}{\gamma} + 1\right) = G_{1,\gamma}\left(\frac{y}{\gamma}+1\right) = \begin{cases} e^{-\left(\frac{y}{\gamma}+1\right)^{-\gamma}} & , \text{ for } y > -\gamma \\ 0 & , \text{ for } y \le -\gamma \end{cases}$$

Since $\lim\limits_{\gamma\to+\infty} e^{-(\frac{y}{\gamma}+1)^{-\gamma}} = e^{-\lim\limits_{\gamma\to+\infty}(\frac{y}{\gamma}+1)^{-\gamma}} = e^{-\lim\limits_{\gamma\to+\infty}(\frac{y}{\gamma}+1)^{\frac{\gamma}{y}\cdot(-y)}} = e^{-e^{-y}}$ and the support goes to $\mathbb{R}$ as $\gamma \to +\infty$, we have $\lim\limits_{\gamma\to+\infty} F_Y(y) = G_3(y)$ for all $y \in \mathbb{R}$, i.e., Y converges in distribution to a random variable having the G_3 distribution. □

6. Find the asymptotic joint distribution of the range, $R_n = X_{(n:n)} - X_{(n:1)}$, and midrange, $M_n = \frac{1}{2}(X_{(n:n)} + X_{(n:1)})$, when sampling from a Pareto distribution with density $f(x) = \frac{1}{x^2}$ for $x > 1$.

Proof. The distribution of Pareto random variable is given by $F(x) = 1 - \frac{1}{x}$ for $x > 1$ and 0 otherwise. Since

$$\mathbb{P}(X_{(n:1)} > x) = [1-F(x)]^n = \begin{cases} x^{-n} & , \ x > 1 \\ 0, & \ x \le 1 \end{cases}$$

we have for all $y > 0$,

$$\mathbb{P}(n(X_{(n:1)} - 1) > y) = \mathbb{P}\left(X_{(1)} > 1 + \frac{y}{n}\right) = \left(1 + \frac{y}{n}\right)^{-n} \to e^{-y}$$

as $n \to \infty$. Therefore, $n(X_{(n:1)} - 1) \xrightarrow{D} \mathrm{Exp}(1)$. Furthermore, we have $X_{(n:1)} \xrightarrow{P} 1$.
From Theorem 15 in Ferguson, we have $nF(X_{(n:1)}) \xrightarrow{D} Y_1$ and $n[1 - F(X_{(n:n)})] \xrightarrow{D} Y_2$, where $Y_1, Y_2 \overset{iid}{\sim} \Gamma(1,1)$. Then $n\frac{X_{(n:1)}-1}{X_{(n:1)}} = n\left(1 - \frac{1}{X_{(n:1)}}\right) \xrightarrow{D} Y_1$ and $n\left(\frac{1}{X_{(n:n)}}\right)^{\top} \xrightarrow{D} Y_2$ are asymptotically independent. So by Slutsky Theorem, we have $\frac{X_{(n:n)}}{n} \xrightarrow{D} \frac{1}{Y_2}$ and

$$\begin{pmatrix} \frac{X_{(n:1)}}{n} \\ \frac{X_{(n:n)}}{n} \end{pmatrix} \xrightarrow{D} \begin{pmatrix} 0 \\ \frac{1}{Y_2} \end{pmatrix}.$$

Let $g(x_1, x_2) = \left(x_2 - x_1, \frac{1}{2}(x_2 + x_1)\right)$, then $\mathcal{C}_g$ the set of continuity points of g satisfies $\mathbb{P}\left((x_1, x_2) \notin \mathcal{C}_g\right) = 0$. By Slutsky Theorem, we have

$$\frac{1}{n}\begin{pmatrix} R_n \\ M_n \end{pmatrix} = g\left(\frac{X_{(n:1)}}{n}, \frac{X_{(n:n)}}{n}\right) \xrightarrow{D} \begin{pmatrix} \frac{1}{Y_2} \\ \frac{1}{2Y_2} \end{pmatrix}.$$

□

7. (Casella and Berger Problem 6.3) Let $X_1, \dots, X_n$ be a random sample from the pdf

$$f(x|\mu, \sigma) = \frac{1}{\sigma} e^{-(x-\mu)/\sigma}, \mu < x < \infty, 0 < \sigma < \infty.$$

Find a two-dimensional sufficient statistic for (μ, σ).

Proof. The joint density of $X_1, \dots, X_n$ is given by

$$\begin{aligned} f_{(X_1,\dots,X_n)}(x_1, \dots, x_n|\mu, \sigma) &= \frac{1}{\sigma^n} e^{-\frac{1}{\sigma}\sum_{i=1}^n (x_i-\mu)} \prod_{i=1}^n \mathbb{1}_{(\mu,+\infty)}(x_i) \\ &= \frac{1}{\sigma^n} e^{-\frac{n}{\sigma}(\overline{x}_n-\mu)} \mathbb{1}_{(\mu,+\infty)}(x_{(1)}) \end{aligned}$$

where $\overline{x}_n = \frac{1}{n}\sum_{i=1}^n x_i$ and $x_{(1)} = \min\{x_1, \dots, x_n\}$. Since this is a function of $(\overline{x}_n, x_{(1)})^\top$ and $(\mu, \sigma)^\top$, by Fisher-Neyman Factorization Theorem, a two-dimensional sufficient statistic for (μ, σ) is $(\overline{X}_n, X_{(1)})^\top$.

□

8. Consider the experiment $(\mathcal{X}, \mathcal{A}, \{f_\theta(x) : \theta \in \Theta\})$, where $\{f_\theta(x) : \theta \in \Theta\}$ is a family of pdfs or pmfs all defined with respect to a common measure. For each $x \in \mathcal{X}$, define $\Theta_x = \{\theta : f_\theta(x) > 0\}$. Assume $\Theta_x \neq \varnothing$ for each $x \in \mathcal{X}$. Assume T is a sufficient statistic. Prove the following lemma. (This lemma was used in class as part of the proof of the version of the Lehmann-Scheffe Theorem that allows the support of the distribution to depend on the parameter.)

Lemma: If $T(x) = T(y)$ for $x, y \in \mathcal{X}$, then $\Theta_x = \Theta_y$.

Proof. Since T is a sufficient statistic, by Fisher-Neymann Factorization Theorem, there exists functions $g(t, \theta)$ and $h(x)$ such that $f_\theta(x) = g(T(x), \theta)h(x)$ for all $x \in \mathcal{X}$ and $\theta \in \Theta$.
If for $x, y \in \mathcal{X}$, $T(x) = T(y)$, then we have $T(x) = T(y)$ and $g(T(x), \theta) = g(T(y), \theta)$. Notice that $f_\theta(x) = g(T(x), \theta)h(x)$, if $h(x) = 0$, then $f_\theta(x) \equiv 0$, $\Theta_x = \varnothing$, which is a contradiction. Thus, $h(x) > 0$. Analogously, $h(y) > 0$. Then

$$\Theta_x = \{\theta : f_\theta(x) > 0\} = \{\theta : g(T(x, \theta) > 0\} = \{\theta : g(T(y, \theta) > 0\} = \{\theta : f_\theta(y) > 0\} = \Theta_y.$$

□