STAT 712 hw 9

Order statistics, convergence in probability

Do problems 5.23, 5.24, 5.27(a) from CB. In addition:

1. For a random sample $X_1, \dots, X_n \overset{\text{ind}}{\sim} f_X$, the joint density of all the order statistics $X_{(1)} < \cdots < X_{(n)}$ is given by

$$f_{X_{(1)},\dots,X_{(n)}}(x_1, \dots, x_n) = n! \cdot \prod_{i=1}^{n} f_X(x_i) \cdot \mathbf{1}(-\infty < x_1 < \cdots < x_n < \infty).$$

Suppose $X_1, \dots, X_n \overset{\text{ind}}{\sim} \text{Exponential}(1)$, with order statistics $X_{(1)} < \cdots < X_{(n)}$, and let

$$\begin{aligned} Z_1 &= nX_{(1)} \\ Z_2 &= (n-1)(X_{(2)} - X_{(1)}) \\ Z_3 &= (n-2)(X_{(3)} - X_{(2)}) \\ &\vdots \\ Z_n &= X_{(n)} - X_{(n-1)}. \end{aligned}$$

(a) Find the joint pdf of $Z_1, \dots, Z_n$.

We first note that the joint support of $(Z_1, \dots, Z_n)$ is the set $(0, \infty)^n$. The inverse transformation is

$$\begin{aligned} X_{(1)} &= n^{-1}Z_1 \\ X_{(2)} &= (n-1)^{-1}Z_2 + n^{-1}Z_1 \\ X_{(3)} &= (n-2)^{-1}Z_3 + (n-1)^{-1}Z_2 + n^{-1}Z_1 \\ &\vdots \\ X_{(n)} &= Z_n + 2^{-1}Z_{n-1} + \cdots + (n-2)^{-1}Z_3 + (n-1)^{-1}Z_2 + n^{-1}Z_1, \end{aligned}$$

which has Jacobian $J(z_1, \dots, z_n) = (n!)^{-1}$. Noting that

$$\sum_{i=1}^{n} X_{(i)} = Z_1 + \cdots + Z_n,$$

the joint pdf of $(Z_1, \dots, Z_n)$ is given by

$$f_{Z_1,\dots,Z_n}(z_1, \dots, z_n) = n! \cdot e^{-(z_1+\cdots+z_n)} \cdot (n!)^{-1} = e^{-(z_1+\cdots+z_n)}$$

for $(z_1, \dots, z_n) \in (0, \infty)^n$.

(b) State whether $Z_1, \dots, Z_n$ are mutually independent.

Since the joint pdf can be factored into a product of functions such that each function depends on only one of the variables z_j, the random variables $Z_1, \dots, Z_n$ are mutually independent.

(c) Give the marginal distribution of each of $Z_1, \dots, Z_n$.

Each Z_j has the Exponential(1) distribution, for $j = 1, \dots, n$.

2. For a random sample $X_1, \dots, X_n$ from a continuous distribution with median M, give

$$P(X_{(1+l)} \leq M \leq X_{(n-l)})$$

for each integer $0 \leq l < \frac{n-1}{2}$. Evaluate the probability for $n = 15$ and $l = 3$. The idea is that we can choose l so the interval $(X_{(1+l)}, X_{(n-l)})$ will contain M with some desired probability (this is a confidence interval for the median). *Hint: When $n = 3$ and $l = 0$ the probability is 0.75.*

We have

$$\begin{aligned} P(X_{(1+l)} \leq M \leq X_{(n-l)}) &= P(\{X_{(1+l)} \leq M\} \cap \{M \leq X_{(n-l)}\}) \\ &= 1 - P(\{X_{(1+l)} > M\} \cup \{X_{(n-l)} < M\}) \\ &= 1 - P(\{X_{(1+l)} > M\} - P(\{X_{(n-l)} < M\}), \end{aligned}$$

since $\{X_{(1+l)} > M\}$ and $\{X_{(n-l)} < M\}$ are disjoint events. Now consider the probabilities

$$\begin{aligned} P(X_{(1+l)} > M) &= P(\text{at least } n - l \text{ are greater than } M) \\ P(X_{(n-l)} < M) &= P(\text{at least } n - l \text{ are less than } M), \end{aligned}$$

noting that both are equal to $P(Y \geq n - l)$, where $Y \sim \text{Binomial}(n, 1/2)$. So we have

$$\begin{aligned} P(X_{(1+l)} \leq M \leq X_{(n-l)}) &= 1 - 2P(Y \geq n - l) \\ &= 1 - 2(2^{-n} \textstyle\sum_{y=n-l}^{n} \binom{n}{y}) \\ &= \texttt{1 - 2*(1 - pbinom(n-l-1,n,0.5))} \end{aligned}$$

For $n = 15$ and $l = 3$ we have `1 - 2*(1 - pbinom(11,15,0.5))` $= 0.9648438$, so the interval $(X_{(4)}, X_{(12)})$ will contain the median with probability 0.9648438.

3. (Optional) Let $X_1, \dots, X_n \overset{\text{ind}}{\sim} \text{Uniform}(\mu - \theta, \mu + \theta)$, $n \geq 2$, and consider the sequences of random variables $\{R_n\}_{n \geq 2}$ and $\{M_n\}_{n \geq 2}$ given by

$$R_n = \frac{X_{(n)} - X_{(1)}}{2} \quad \text{and} \quad M_n = \frac{X_{(1)} + X_{(n)}}{2}$$

for $n \geq 2$.

(a) Find the joint pdf of (R_n, M_n).

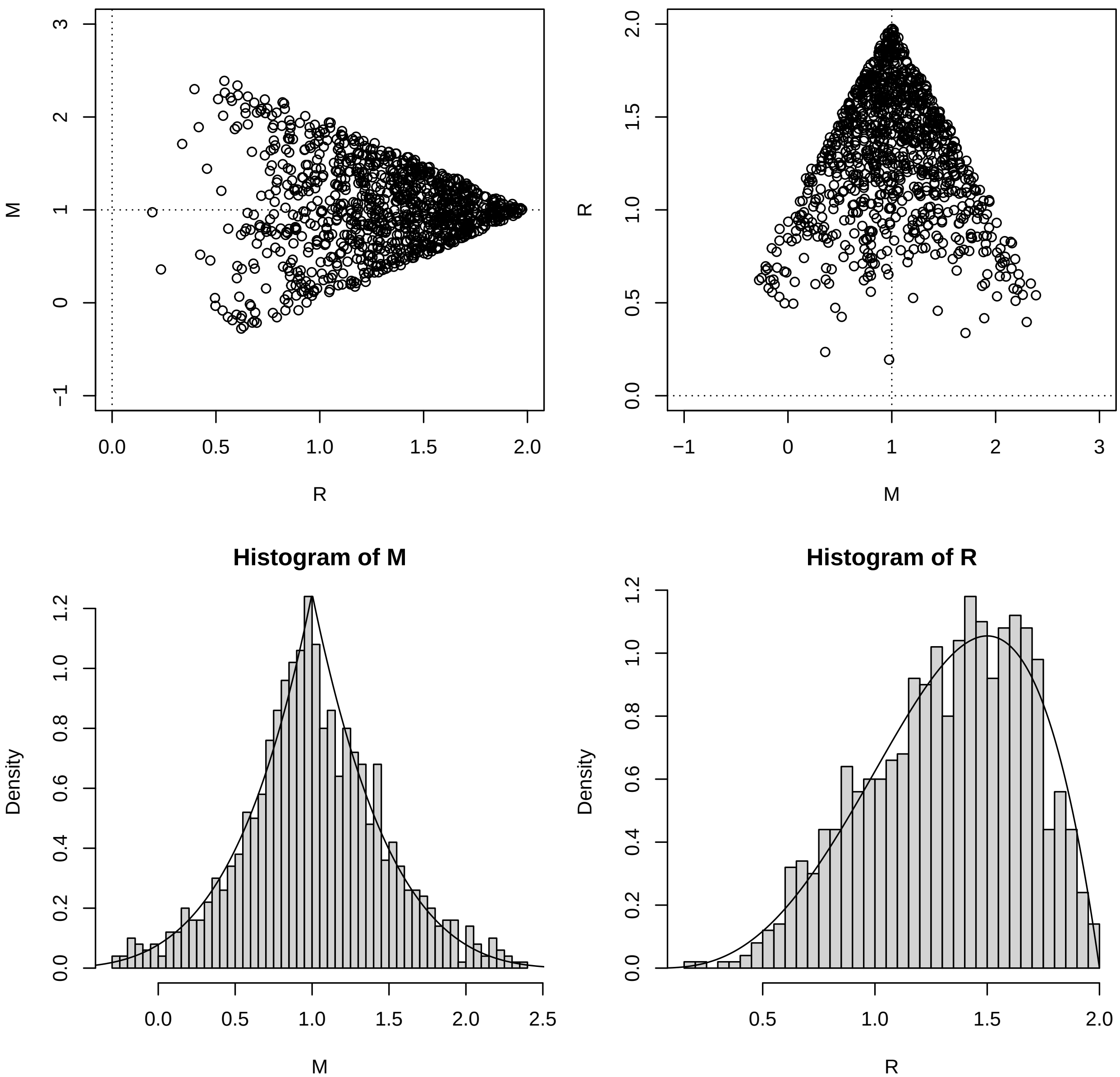
We first write down the joint pdf of $(X_{(1)}, X_{(n)})$ as

$$f_{X_{(1)},X_{(n)}}(u,v) = \frac{n(n-1)}{(2\theta)^n} \cdot (v-u)^{n-2} \cdot \mathbf{1}(\mu - \theta < u < v < \mu + \theta).$$

Now, the trickiest part of finding the joint pdf of (R_n, M_n) is writing down the joint support. We have

$$\begin{aligned}(R_n, M_n) &\in \{(r,m) : 0 < r < \theta, \mu - \theta + r < m < \mu + \theta - r\} \\ &= \{(r,m) : \mu - \theta < m < \mu + \theta, r < \theta - |m - \mu|\}\end{aligned}$$

To see this, it might help to run a simulation and plot many realizations of (R_n, M_n), as below (these plots were made with $\theta = 2$, $\mu = 1$, and $n = 5$):



Now we have

$$\begin{array}{l} r = (v-u)/2 =: g_1(u,v) \\ m = (v+u)/2 =: g_2(u,v) \end{array} \iff \begin{array}{l} u = m - r =: g_1^{-1}(r,m) \\ v = m + r =: g_2^{-1}(r,m) \end{array}$$

with Jacobian

$$J(x,y) = \begin{vmatrix} \frac{d}{dr} m - r & \frac{d}{dm} m - r \\ \frac{d}{dr} m + r & \frac{d}{dm} m + r \end{vmatrix} = \begin{vmatrix} -1 & 1 \\ 1 & 1 \end{vmatrix} = -2.$$

Now the joint pdf of (R_n, M_n) is given by

$$f_{R_n,M_m}(r,m) = \frac{n(n-1)}{2\theta^n} \cdot r^{n-2} \cdot \mathbf{1}(\mu - \theta < m < \mu + \theta, r < \theta - |m - \mu|).$$

(b) Find the marginal pdf of R_n.

For $r \in (0, \theta)$, the marginal pdf of R_n is given by

$$\begin{aligned} f_{R_n}(r) &= \int_{\mu-\theta+r}^{\mu+\theta-r} \frac{n(n-1)}{2\theta^n} \cdot r^{n-2} \cdot dm \\ &= \frac{n(n-1)}{2\theta^n} \cdot r^{n-2}[(\mu+\theta-r) - (\mu-\theta+r)] \\ &= \frac{n(n-1)}{2\theta^n} \cdot r^{n-2}[2\theta - 2r] \end{aligned}$$

We can write this as

$$f_{R_n}(r) = \frac{1}{\theta} \cdot \frac{\Gamma(n-1+2)}{\Gamma(n-1)\Gamma(2)} \left(\frac{r}{\theta}\right)^{(n-1)-1} \left(1 - \frac{r}{\theta}\right)^{2-1} \cdot \mathbf{1}(0 < r < \theta).$$

(c) Find the marginal pdf of M_n.

For $m \in (\mu - \theta, \mu + \theta)$ the marginal pdf of M_n is given by

$$f_{M_n}(m) = \int_0^{\theta - |m-\mu|} \frac{n(n-1)}{2\theta^n} \cdot r^{n-2} \cdot dr = \frac{n}{2\theta^n} \cdot (\theta - |m-\mu|)^{n-1}.$$

So we write

$$f_{M_n}(m) = \frac{n}{2\theta^n} \cdot (\theta - |m-\mu|)^{n-1} \cdot \mathbf{1}(\mu - \theta < m < \mu + \theta).$$

(d) Show that R_n converges in probability to θ as $n \to \infty$.

If $B_n \sim \text{Beta}(n-1, 2)$, then R_n has the same distribution as θB_n. Therefore

$$\begin{aligned} \mathbb{E} R_n &= \mathbb{E}(\theta B_n) = \theta \cdot \frac{n-1}{n-1+2} \to \theta \quad \text{as } n \to \infty, \quad \text{and} \\ \text{Var}\, R_n &= \text{Var}(\theta B_n) = \theta^2 \cdot \frac{(n-1)2}{(n-1+2)^2(n-1+2+1)} \to 0 \quad \text{as } n \to \infty, \end{aligned}$$

from the mean and variance formulas of the Beta distribution. Therefore $R_n \xrightarrow{\text{p}} \theta$.

(e) Show that M_n converges in probability to μ as $n \to \infty$.

Fixing $\varepsilon > 0$, we have $P(|M_n - \mu| < \varepsilon) = 1$ if $\varepsilon > \theta$; otherwise, we have

$$\begin{aligned} P(|M_n - \mu| < \varepsilon) &= \int_{\mu-\varepsilon}^{\mu+\varepsilon} \frac{n}{2\theta^n} \cdot (\theta - |m - \mu|)^{n-1} dm \\ &= 2\int_{\mu}^{\mu+\varepsilon} \frac{n}{2\theta^n} \cdot (\theta - (m - \mu))^{n-1} dm \\ &= \frac{1}{\theta^n} (\theta - (m - \mu))^n \Big|_{\mu}^{\mu+\varepsilon} \\ &= 1 - \left(\frac{\theta - \varepsilon}{\theta}\right)^n. \end{aligned}$$

Since the right hand side goes to 1 as $n \to \infty$, $M_n \xrightarrow{\text{p}} \mu$ as $n \to \infty$.

Problems 5.23, 5.24, 5.27 (a) from CB

5.23 Let $U_1, \dots, U_X \sim U(0,1)$, where X is a rv with pmf

$$p_X(x) = \frac{1}{e-1} \frac{1}{x!}, \qquad x = 1, 2, 3, \dots$$

We want to find the marginal pdf of $Z = U_{(1)}$.

We have

$$Z \mid X \sim f(z \mid x) = x(1-z)^{x-1} \qquad \left(= x\left[1 - F_U(z)\right]^{x-1} f_U(z) \right)$$

$$X \sim p_X$$

Note that $Z \in (0,1)$.

For $z \in (0,1)$, we have

$$\begin{aligned}
f_Z(z) &= \sum_{x=1}^{\infty} x(1-z)^{x-1} \frac{1}{e-1} \frac{1}{x!} \\
&= \frac{1}{e-1} \sum_{x=1}^{\infty} \frac{(1-z)^{x-1}}{(x-1)!} \\
&= \frac{1}{e-1} \sum_{k=0}^{\infty} \frac{(1-z)^k}{k!} \\
&= \frac{e^{(1-z)}}{e-1} \\
&= \frac{e^{-z}}{1-e^{-1}},
\end{aligned}$$

which is the pdf of an exponential dist. truncated to $(0,1)$:

$$f_Z(z) = \frac{e^{-z}}{1-e^{-1}} \mathbb{1}(0 < z < 1).$$

5.24 Let $X_1, \dots, X_n \overset{ind}{\sim} f_X(x) = \frac{1}{\theta} \mathbb{1}(0 < x < \theta)$.

Let

$$R = X_{(1)} / X_{(n)}$$
$$M = X_{(n)}$$

The joint density of $(X_{(1)}, X_{(n)})$ is

$$f(u,v) = n(n-1) \left[F_X(v) - F_X(u) \right]^{n-2} f_X(u) f_X(v) \; \mathbb{1}(u < v)$$

$$= n(n-1) \left[\frac{v}{\theta} - \frac{u}{\theta} \right]^{n-2} \frac{1}{\theta} \frac{1}{\theta} \; \mathbb{1}(u < v)$$

$$= \frac{n(n-1)}{\theta^n} (v-u)^{n-2} \mathbb{1}(0 < u < v < \theta)$$

Now write

$$\begin{aligned} r &= u/v = g_1(u,v) \\ m &= v \quad = g_2(u,v) \end{aligned} \quad \Longleftrightarrow \quad \begin{aligned} u &= mr = g_1^{-1}(r,m) \\ v &= m = g_2^{-1}(r,m) \end{aligned}$$

$$J(r,m) = \begin{vmatrix} \frac{\partial}{\partial r} mr & \frac{\partial}{\partial r} m \\ \frac{\partial}{\partial m} mr & \frac{\partial}{\partial m} m \end{vmatrix} = \begin{vmatrix} m & 0 \\ r & 1 \end{vmatrix} = m.$$

So the joint density of (R, M) is given by

$$f(r,m) = \frac{n(n-1)}{\theta^n} (m - mr)^{n-2} \cdot |m| \; \mathbb{1}(0 < mr < m < \theta)$$

$$= n(n-1) \frac{m^{n-1}}{\theta^n} (1-r)^{n-2} \; \mathbb{1}(0 < r < 1) \; \mathbb{1}(0 < m < \theta)$$

$$= n(n-1)(1-r)^{n-2} \mathbb{1}(0 < r < 1) \cdot \frac{m^{n-1}}{\theta^n} \mathbb{1}(0 < m < \theta).$$

We see that R and M are independent by the factorization theorem.

5.27 Let $X_1, \ldots, X_n \overset{ind}{\sim} f_X$ (a density) with cdf F_X.

Find the conditional pdf of $X_{(j)} | X_{(k)}$.

For $j < k$ the joint pdf of $(X_{(j)}, X_{(k)})$ is given by

$$f(u,v) = \frac{n!}{(j-1)!\,(k-j-1)!\,(n-k)!} f_X(u)\, f_X(v) \left[F_X(u)\right]^{j-1} \left[F_X(v) - F_X(u)\right]^{k-j-1} \left[1 - F_X(v)\right]^{n-k}$$

for $u < v$.

The marginal pdf of $X_{(k)}$ is given by

$$f_V(v) = \frac{n!}{(k-1)!\,(n-k)!} \left[F_X(v)\right]^{k-1} \left[1 - F_X(v)\right]^{n-k} f_X(v)$$

So the conditional pdf of $X_{(j)} | X_{(k)}$, $j < k$ is given by

$$f(u|v) = \frac{\frac{n!}{(j-1)!\,(k-j-1)!\,(n-k)!} f_X(u)\, f_X(v) \left[F_X(u)\right]^{j-1} \left[F_X(v) - F_X(u)\right]^{k-j-1} \left[1 - F_X(v)\right]^{n-k}}{\frac{n!}{(k-1)!\,(n-k)!} \left[F_X(v)\right]^{k-1} \left[1 - F_X(v)\right]^{n-k} f_X(v)}$$

$$= \frac{(k-1)!}{(j-1)!\,(k-j-1)!} \frac{f_X(u)}{F_X(v)} \left[\frac{F_X(u)}{F_X(v)}\right]^{j-1} \left[1 - \frac{F_X(u)}{F_X(v)}\right]^{k-j-1}$$

for $u < v$.

The conditional pdf of $X_{(k)} \mid X_{(j)}$ is

$$f(v|u) = \frac{\frac{n!}{(j-1)!\,(k-j-1)!\,(n-k)!} f_X(u)\, f_X(v) \left[F_X(u)\right]^{j-1} \left[F_X(v) - F_X(u)\right]^{k-j-1} \left[1 - F_X(v)\right]^{n-k}}{\frac{n!}{(j-1)!\,(n-j)!} \left[F_X(u)\right]^{j-1} \left[1 - F_X(u)\right]^{n-j} f_X(u)}$$

$$= \frac{(n-j)!}{(k-j-1)!\,(n-k)!} \frac{f_X(v)}{1 - F_X(u)} \left[\frac{F_X(v) - F_X(u)}{1 - F_X(u)}\right]^{k-j-1} \left[\frac{1 - F_X(v)}{1 - F_X(u)}\right]^{n-k}$$

for $u < v$.