Homework 12

Mengxiang Jiang
Stat 610 Distribution Theory

December 5, 2025

Problem 1. Suppose $X_1, \dots, X_n$ is a simple random sample from the exponential$(1/\lambda)$ distribution with pdf $f(x) = \lambda e^{-\lambda x}$ for $x > 0$.

(a) Find the joint cdf for the smallest and next-to-smallest order statistics ($X_{(1)}$ and $X_{(2)}$, resp.). Hint: start by showing that

$$\mathsf{P}(x_1 < X_{(1)} \leq x_2 < X_{(2)}) = \sum_{i=1}^{n} \left(\mathsf{P}(x_1 < X_i \leq x_2) \prod_{\substack{j=1 \\ j \neq i}}^{n} \mathsf{P}(X_j > x_2) \right)$$

and then simplify. Derive the joint cdf from this (for $x_1 < x_2$, noting the support).

(b) Obtain the joint pdf for $(X_{(1)}, X_{(2)})$. Be careful; it does simplify.

(a) We have

$$\begin{aligned}
\mathsf{P}(x_1 < X_{(1)} \leq x_2 < X_{(2)}) &= \mathsf{P}\left(\bigcup_{i=1}^{n} \{x_1 < X_i \leq x_2, X_j > x_2 \text{ for } j \neq i\} \right) \\
&= \sum_{i=1}^{n} \mathsf{P}(x_1 < X_i \leq x_2, X_j > x_2 \text{ for } j \neq i) \\
&= \sum_{i=1}^{n} \left(\mathsf{P}(x_1 < X_i \leq x_2) \prod_{\substack{j=1 \\ j \neq i}}^{n} \mathsf{P}(X_j > x_2) \right) \\
&= n \left((e^{-\lambda x_1} - e^{-\lambda x_2})(e^{-\lambda x_2})^{n-1} \right) \\
&= n(e^{-\lambda x_1} - e^{-\lambda x_2}) e^{-\lambda(n-1)x_2}
\end{aligned}$$

for $0 < x_1 < x_2$. Therefore, the joint cdf is

$$\begin{aligned}
F_{X_{(1)}, X_{(2)}}(x_1, x_2) &= \mathsf{P}(X_{(1)} \leq x_1, X_{(2)} \leq x_2) \\
&= \mathsf{P}(X_{(1)} \leq x_1) - \mathsf{P}(x_1 < X_{(1)} \leq x_2 < X_{(2)}) \\
&= 1 - e^{-\lambda n x_1} - n(e^{-\lambda x_1} - e^{-\lambda x_2}) e^{-\lambda(n-1)x_2}
\end{aligned}$$

for $0 < x_1 < x_2$.

(b) The joint pdf is

$$
\begin{aligned}
f_{X_{(1)},X_{(2)}}(x_1,x_2) &= \frac{\partial^2}{\partial x_1 \partial x_2} F_{X_{(1)},X_{(2)}}(x_1,x_2) \\
&= \frac{\partial}{\partial x_2}\left(\lambda n e^{-\lambda n x_1} + n\lambda e^{-\lambda x_1} e^{-\lambda(n-1)x_2} - n\lambda n e^{-\lambda n x_2} + n\lambda e^{-\lambda x_2} e^{-\lambda(n-1)x_2}\right) \\
&= n(n-1)\lambda^2 e^{-\lambda x_1} e^{-\lambda(n-1)x_2}
\end{aligned}
$$

for $0 < x_1 < x_2$.

Problem 2. Suppose $T_1, T_2, \ldots$ is an iid sequence from the Lomax distribution with cdf $F(t) = 1 - (1 + t/\beta)^{-\alpha}$ where both α and β are positive. Use Theorem 5.35 in the notes to obtain an asymptotic distribution for $M_n = \max(T_1, \ldots, T_n)$.

Theorem 5.35. If $F(x)$ satisfies $\lim_{x\to\infty} x^\alpha(1 - F(x)) = c > 0$, with $\alpha > 0$, then

$$
\lim_{n\to\infty} \mathsf{P}(M_n \leq n^{1/\alpha} x) = e^{-cx^{-\alpha}},
$$

which is the Fréchet$(\alpha, 1/c)$ cdf.

We have

$$
\begin{aligned}
\lim_{t\to\infty} t^\alpha(1 - F(t)) &= \lim_{t\to\infty} t^\alpha (1 + t/\beta)^{-\alpha} \\
&= \lim_{t\to\infty} \left(\frac{t}{1 + t/\beta}\right)^\alpha \\
&= \lim_{t\to\infty} \left(\frac{1}{1/t + 1/\beta}\right)^\alpha \\
&= \beta^\alpha
\end{aligned}
$$

Therefore, by Theorem 5.35, we have

$$
\lim_{n\to\infty} \mathsf{P}(M_n \leq n^{1/\alpha} x) = e^{-\beta^\alpha x^{-\alpha}},
$$

which is the Fréchet$(\alpha, 1/\beta^\alpha)$ cdf.

Problem 3. Suppose $V_1, \dots, V_n \overset{iid}{\sim}$ beta(a, 1), with cdf $F_V(v) = v^a$ for $0 \le v \le 1$.

(a) Find the marginal pdf for each i-th order statistic, $V_{(i)}$.

(b) Let $M_n = V_{(n)} = \max(_1, \dots, V_n)$. Find $\mathsf{E}(M_n)$.

(c) Determine a sequence a_n so that $\frac{1-M_n}{a_n}$ converges in distribution and identify the limit. Hint: Theorem 5.36 in the notes.

(d) Let $W_i = 1 - V_i$, which has beta(1, a) distribution. (What is the cdf?) Determine a sequence a'_n so that $\frac{1-W_{(n)}}{a'_n}$ converges in distribution and identify the limit. Note that this also determines the asymptotic behavior for $V_{(1)} = \min(_1, \dots, V_n) = 1 - W_{(n)}$.

(a) The marginal pdf for $V_{(i)}$ is

$$\begin{aligned} f_{V_{(i)}}(v) &= \frac{n!}{(i-1)!(n-i)!}(F_V(v))^{i-1}(1-F_V(v))^{n-i}f_V(v) \\ &= \frac{n!}{(i-1)!(n-i)!}(v^a)^{i-1}(1-v^a)^{n-i}av^{a-1} \\ &= \frac{n!a}{(i-1)!(n-i)!}v^{ai-1}(1-v^a)^{n-i} \end{aligned}$$

(b) We have

$$\begin{aligned} \mathsf{E}(M_n) &= \int_0^1 v f_{V_{(n)}}(v)dv \\ &= \int_0^1 v\frac{n!a}{(n-1)!}v^{an-1}(1-v^a)^0 dv \\ &= na\int_0^1 v^{an}dv \\ &= na\frac{1}{an+1} \\ &= \frac{na}{an+1} \end{aligned}$$

(c) **Theorem 5.36.** If $F(x)$ satisfies $\lim_{x\to 0} x^{-\gamma}(1-F(b-x)) = c > 0$, with $\gamma > 0$, then

$$\lim_{n\to\infty} \mathsf{P}(b - M_n \le n^{-1/\gamma}x) = 1 - e^{-cx^\gamma},$$

which is the Weibull(γ, $1/c$) distribution.

We want a_n such that $\frac{1-M_n}{a_n}$ converges in distribution. We have

$$\begin{aligned} \mathsf{P}(1 - M_n \le a_n x) &= \mathsf{P}\left(M_n \ge 1 - a_n x\right) \\ &= 1 - \mathsf{P}\left(M_n < 1 - a_n x\right) \\ &= 1 - (F_V(1-a_n x))^n \\ &= 1 - (1-(a_n x)^a)^n \end{aligned}$$

We want

$$\lim_{n\to\infty} \mathsf{P}(1 - M_n \le a_n x) = 1 - e^{-cx^\gamma}$$

for some $c > 0$ and $\gamma > 0$. This requires

$$\lim_{n\to\infty} n(a_n x)^a = cx^\gamma$$

which implies $\gamma = a$ and $a_n = (c/n)^{1/a}$. Therefore,

$$\frac{1 - M_n}{(c/n)^{1/a}} \xrightarrow{d} \text{Weibull}(a, 1/c).$$

(d) The cdf for W_i is

$$F_W(w) = 1 - (1 - w)^a$$

for $0 \le w \le 1$. We want a_n' such that $\frac{1-W_{(n)}}{a_n'}$ converges in distribution. We have

$$\begin{aligned}
\mathsf{P}(1 - W_{(n)} \le a_n' x) &= \mathsf{P}\left(W_{(n)} \ge 1 - a_n' x\right) \\
&= 1 - \mathsf{P}\left(W_{(n)} < 1 - a_n' x\right) \\
&= 1 - (F_W(1 - a_n' x))^n \\
&= 1 - (1 - (a_n' x)^a)^n
\end{aligned}$$

We want

$$\lim_{n\to\infty} \mathsf{P}(1 - W_{(n)} \le a_n' x) = 1 - e^{-cx^\gamma}$$

for some $c > 0$ and $\gamma > 0$. This requires

$$\lim_{n\to\infty} n(a_n' x)^a = cx^\gamma$$

which implies $\gamma = a$ and $a_n' = (c/n)^{1/a}$. Therefore,

$$\frac{1 - W_{(n)}}{(c/n)^{1/a}} \xrightarrow{d} \text{Weibull}(a, 1/c).$$

This also implies the asymptotic behavior for $V_{(1)} = \min({}_1, \ldots, V_n) = 1 - W_{(n)}$ is

$$\frac{V_{(1)}}{(c/n)^{1/a}} \xrightarrow{d} \text{Weibull}(a, 1/c).$$

Problem 4. Assume $X_1, X_2, \ldots \overset{iid}{\sim}$ exponential(β) and $M_n = \max(_1, \ldots, X_n)$.

(a) Write down the cdf for M_n and use it to obtain the cdf for $Y_n = M_n - \beta \log(n)$.

(b) Show that $-\log(\Pr(Y_n > y)) \to h(y)$ for some increasing function $h(y)$ and use this fact to deduce the limit distribution for $Y_n = M_n - \beta \log(n)$.

(a) The cdf for M_n is

$$\begin{aligned} F_{M_n}(x) &= \mathsf{P}(M_n \le x) \\ &= (F_X(x))^n \\ &= (1 - e^{-x/\beta})^n \end{aligned}$$

for $x > 0$. Therefore, the cdf for Y_n is

$$\begin{aligned} F_{Y_n}(y) &= \mathsf{P}(Y_n \le y) \\ &= \mathsf{P}(M_n - \beta \log(n) \le y) \\ &= \mathsf{P}(M_n \le y + \beta \log(n)) \\ &= (1 - e^{-(y+\beta \log(n))/\beta})^n \\ &= (1 - e^{-y/\beta} e^{-\log(n)})^n \\ &= (1 - e^{-y/\beta}/n)^n \end{aligned}$$

for $y > -\beta \log(n)$.

(b) We have

$$\begin{aligned} \mathsf{P}(Y_n > y) &= 1 - F_{Y_n}(y) \\ &= 1 - (1 - e^{-y/\beta}/n)^n. \end{aligned}$$

The limit cdf is

$$\begin{aligned} \lim_{n\to\infty} F_{Y_n}(y) &= \lim_{n\to\infty} (1 - e^{-y/\beta}/n)^n \\ &= e^{-e^{-y/\beta}} \end{aligned}$$

for $y \in \mathbb{R}$. Therefore,

$$-\log(\Pr(Y_n > y)) \to e^{-y/\beta} = h(y).$$

This implies that the limit distribution for $Y_n = M_n - \beta \log(n)$ is the Gumbel distribution with cdf

$$F_Y(y) = e^{-e^{-y/\beta}}.$$

Problem 5. Recall the definition of a compound Poisson distribution (Slide 398 in the notes) where $N \sim \text{Poisson}(\lambda)$, independent of $Y_1, Y_2, \ldots$, and $T = \sum_{i \le N} Y_i$ (with $T = 0$ if $N = 0$).

(a) Use iterated expectations or variance partition, conditioning on N, to find $\mathsf{E}(T)$ and $\mathsf{Var}(T)$ (e.g., $\mathsf{E}(T) = \mathsf{E}(\mathsf{E}(T|N))$), in terms of λ and the moments of T.

(b) Use the mgf from Corollary 5.38 to do the same.

(c) Apply the above to the case that each $Y_i \sim \text{exponential}(\beta)$, and compare with Example 4.18 in the notes.

(a) We have

$$\begin{aligned}
\mathsf{E}(T) &= \mathsf{E}(\mathsf{E}(T|N)) \\
&= \mathsf{E}\left(\mathsf{E}\left(\sum_{i=1}^{N} Y_i \middle| N\right)\right) \\
&= \mathsf{E}\left(N\mathsf{E}(Y_1)\right) \\
&= \mathsf{E}(N)\mathsf{E}(Y_1) \\
&= \lambda\mathsf{E}(Y_1)
\end{aligned}$$

and

$$\begin{aligned}
\mathsf{Var}(T) &= \mathsf{E}(\mathsf{Var}(T|N)) + \mathsf{Var}(\mathsf{E}(T|N)) \\
&= \mathsf{E}\left(\mathsf{Var}\left(\sum_{i=1}^{N} Y_i \middle| N\right)\right) + \mathsf{Var}\left(\mathsf{E}\left(\sum_{i=1}^{N} Y_i \middle| N\right)\right) \\
&= \mathsf{E}\left(N\mathsf{Var}(Y_1)\right) + \mathsf{Var}\left(N\mathsf{E}(Y_1)\right) \\
&= \mathsf{E}(N)\mathsf{Var}(Y_1) + (\mathsf{E}(Y_1))^2\mathsf{Var}(N) \\
&= \lambda\mathsf{Var}(Y_1) + (\mathsf{E}(Y_1))^2\lambda \\
&= \lambda(\mathsf{Var}(Y_1) + (\mathsf{E}(Y_1))^2)
\end{aligned}$$

(b) From corollary 5.38, the mgf is

$$M_T(t) = e^{\lambda(M_Y(t)-1)}.$$

where $M_Y(t)$ is the mgf of Y_i. Therefore,

$$\begin{aligned}
\mathsf{E}(T) &= M_T'(0) \\
&= \lambda M_Y'(0) \\
&= \lambda\mathsf{E}(Y_1)
\end{aligned}$$

and

$$\begin{aligned}
\mathsf{Var}(T) &= M_T''(0) - (M_T'(0))^2 \\
&= \lambda M_Y''(0) + \lambda^2(M_Y'(0))^2 - (\lambda M_Y'(0))^2 \\
&= \lambda M_Y''(0) \\
&= \lambda(\mathsf{Var}(Y_1) + (\mathsf{E}(Y_1))^2)
\end{aligned}$$

(c) For $Y_i \sim$ exponential(β), we have

$$\mathsf{E}(Y_1) = \beta, \quad \mathsf{Var}(Y_1) = \beta^2.$$

Therefore,

$$\mathsf{E}(T) = \lambda\beta, \quad \mathsf{Var}(T) = \lambda(\beta^2 + \beta^2) = 2\lambda\beta^2,$$

which is consistent with Example 4.18 in the notes.

Problem 6. Consider the interval [0, 1] divided into n intervals of length $1/n$ each:

$$[0, \frac{1}{n}), [\frac{1}{n}, \frac{2}{n}), \dots, [1 - \frac{1}{n}, 1].$$

Let $X_{n,i}$ be the indicator of the event a point occurs in the i-th interval, and suppose

$$p_{n,i} = \mathsf{P}(X_{n,i} = 1) = \frac{g(i/n)}{n}$$

for some continuous positive function $g(x)$. Assume also that $X_{n,1}, \dots, X_{n,n}$ are independent. Define $S_n = \sum_{i=1}^{n} X_{n,i}$ to be the number of intervals with points, for a fixed n, and use the Law of Rare Events (Theorem 5.37 in the notes) to show that $S_n \xrightarrow{D}$ Poisson(λ), where

$$\lambda = \int_0^1 g(x)\,dx.$$

Hint: you also need the definition of Riemann integral. Aside from that, this is just a couple lines of argument.

Theorem 5.37. (Law of Rare Events) Assume $X_{n,1}, \dots, X_{n,n}$ are independent Bernoulli random variables with $\mathsf{P}(X_{n,i} = 1) = p_{n,i}$, $S_n = \sum_{i=1}^{n} X_{n,i}$, and $\lambda_n = \mathsf{E}(S_n) = \sum_{i=1}^{n} p_{n,i}$. If $\lambda_n \to \lambda \in (0, \infty)$, and $\delta_n = \max_{i \le n} p_{n,i} \to 0$ as $n \to \infty$, then

$$S_n \xrightarrow{D} \text{Poisson}(\lambda).$$

We have

$$\lambda_n = \mathsf{E}(S_n) = \sum_{i=1}^{n} p_{n,i} = \sum_{i=1}^{n} \frac{g(i/n)}{n}.$$

By the definition of Riemann integral, we have

$$\lim_{n\to\infty} \lambda_n = \lim_{n\to\infty} \sum_{i=1}^{n} \frac{g(i/n)}{n} = \int_0^1 g(x)\,dx = \lambda.$$

Also, we have

$$\delta_n = \max_{i \le n} p_{n,i} = \max_{i \le n} \frac{g(i/n)}{n} \le \frac{\max_{x\in[0,1]} g(x)}{n} \to 0$$

as $n \to \infty$. Therefore, by Theorem 5.37, we have

$$S_n \xrightarrow{D} \text{Poisson}(\lambda).$$