Bios 661: 1 – 5; Bios 673: 2 – 6.

1. C&B 5.21

2. C&B 5.24

3. C&B 5.25

4. Let $X_1, \dots, X_n$ be independent random variables having exponential distribution with respective parameters $\alpha_1, \alpha_2, \dots, \alpha_n$ and probability density functions

$$f_{X_i}(x_i) = \alpha_i e^{-\alpha_i x_i}, \quad x_i > 0, \quad \alpha_i > 0, \quad i = 1, \dots, n.$$

(a) Show that the minimum order statistic $X_{(1)} = \min\{X_1, \dots, X_n\}$ has an exponential distribution with parameter $\sum_{i=1}^n \alpha_i$ and pdf

$$f_{X_{(1)}}(x) = \left(\sum_{i=1}^n \alpha_i\right) e^{-(\sum_{i=1}^n \alpha_i)x}.$$

[Note: The pdf formula for the order statistics does not work since the random variables are not iid. Use CDF method.]

Solution: The CDF of $X_{(1)}$ is

$$\begin{aligned} F_{X_{(1)}}(x) &= P(X_{(1)} \le x) = 1 - P(X_{(1)} \ge x) \\ &= 1 - P(X_1 \ge x) \dots P(X_n \ge x) \\ &= 1 - e^{-\alpha_1 x} \dots e^{-\alpha_n x} \\ &= 1 - e^{-\sum_{i=1}^n \alpha_i x}, \end{aligned}$$

which is CDF of an exponential distribution with parameter $\sum_{i=1}^n \alpha_i$ and pdf

$$f_{X_{(1)}}(x) = \left(\sum_{i=1}^n \alpha_i\right) e^{-(\sum_{i=1}^n \alpha_i)x}.$$

(b) Show that

$$P(X_{(1)} = X_k) = \frac{\alpha_k}{\sum_{i=1}^n \alpha_i}, \quad k \ge 1.$$

Solution: The probability equals

$$
\begin{aligned}
P(X_{(1)} = X_k) &= \int_0^\infty P(X_1 \geq x_k, \ldots, X_n \geq x_k, X_k = x_k) dx_k \\
&= \int_0^\infty \prod_{i=1, i\neq k}^{n} P(X_i \geq x_k) f_{X_k}(x_k) dx_k \\
&= \int_0^\infty \exp\left\{-\left(\sum_{i=1, i\neq k}^{n} \alpha_i\right) x_k\right\} \alpha_k \exp(-\alpha_k x_k) dx_k \\
&= \alpha_k \int_0^\infty \exp\left\{-\left(\sum_{i=1}^{n} \alpha_i\right) x_k\right\} dx_k \\
&= \frac{\alpha_k}{\sum_{i=1}^n \alpha_i}.
\end{aligned}
$$

5. Suppose that iid random variables $X_1, \ldots, X_n$ follow a uniform distribution on the interval $(0, 1)$ with pdf

$$f_X(x) = 1, \quad 0 < x < 1.$$

Let random variables $U = X_{(1)}$ and $V = 1 - X_{(n)}$, where $X_{(1)} = \min_i X_i$ and $X_{(n)} = \max_i X_i$ are minimum and maximum order statistics, respectively.

(a) Find an explicit expression for the joint distribution of the random variables U and V.

Solution: To find the explicit expression, we can first derive the joint pdf of $(X_{(1)}, X_{(n)})$ as

$$f_{X_{(1)},X_{(n)}}(x_1, x_n) = \frac{n!}{1!(n-2)!1!}(x_n - x_1)^{n-2}.$$

Since U and V are transformations of $X_{(1)}$ and $X_{(n)}$, we can have $X_{(1)} = U$, $X_{(n)} = 1 - V$, and $|J| = 1$. The joint pdf of (U, V) is

$$f_{U,V}(u, v) = n(n-1)(1 - v - u)^{n-2}, \quad 0 < u < 1, \quad 0 < v < 1 - u.$$

(b) Let $R = nU$ and $S = nV$. Show that

$$P(R > r, S > s) = \left(1 - \frac{r}{n} - \frac{s}{n}\right)^n.$$

Solution: One can derive the joint survivor function by expressing

$$P(R > r, S > s) = P(U > \frac{r}{n}, V > \frac{s}{n}) = \int_{s/n}^{1-r/n} \int_{r/n}^{1-v} f_{U,V}(u, v) du dv,$$

which leads to the answer by calculating the double integrals.

(c) Following the result in (b), show that R and S are asymptotically independent. You may need the fact that

$$\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n = e^x.$$

Solution: Taking the limit of the joint survivor function, one has

$$\lim_{n\to\infty} P(R>r, S>s) = \lim_{n\to\infty}\left(1-\frac{r}{n}-\frac{s}{n}\right)^n = e^{-(r+s)} = e^{-r}e^{-s}.$$

Since the limiting joint survivor function is the product of two separate functions of r and s, one can claim that R and S are asymptotically independent.

(d) What is the asymptotic distribution of R and S?

Solution: The asymptotic survivor functions of R and S are e^{-r} and e^{-s}, respectively. Both follow exponential distribution with mean 1 when the sample size is large.

6. Let $X_1, \ldots, X_n$ be a random sample from the exponential distribution with pdf

$$\beta^{-1}e^{(\alpha-x)/\beta}, \quad \alpha < x < \infty,$$

where $\alpha \in \mathcal{R}$ and $\beta > 0$ are parameters. Let $X_{(1)} \le \cdots \le X_{(n)}$ be order statistics, and let $Z_1 = X_{(1)}$ and $Z_i = X_{(i)} - X_{(i-1)}$ for $i = 2, \ldots, n$. Show that

(a) $Z_1, \ldots, Z_n$ are independent and $2(n-i+1)Z_i/\beta$ has the χ^2_2 distribution.

Solution: First, we have

$$f_{X_{(1)},\ldots,X_{(n)}}(x_1,\ldots,x_n) = n!\beta^{-n}e^{\sum_{i=1}^n(\alpha-x_i)/\beta}I(\alpha < x_1 < x_2 < \ldots < x_n < \infty).$$

Then, the inverse transformation equals $X_{(1)} = Z_1$, $X_{(2)} = Z_1 + Z_2$, …, $X_{(n)} = \sum_{i=1}^n Z_i$. The Jacobian $J = 1$. Therefore,

$$\begin{aligned} f_{Z_1,\ldots,Z_n}(z_1,\ldots,z_n) &= n!\beta^{-n}e^{(\alpha-z_1)/\beta}e^{(\alpha-z_1-z_2)/\beta}\cdots e^{(\alpha-\sum_{i=1}^n z_i)/\beta} \\ &= f_{Z_1}(z_1)f_{Z_2}(z_2)\cdots f_{Z_n}(z_n), \end{aligned}$$

where $f_{Z_1}(z_1) = n\beta^{-1}e^{n(\alpha-z_1)/\beta}I(\alpha < z_1 < \infty)$ and

$$f_{Z_i}(z_i) = (n-i+1)\beta^{-1}e^{-(n-i+1)z_i/\beta}I(0 < z_i < \infty), \quad i = 2, \ldots, n.$$

We hence can claim $Z_1, \ldots, Z_n$ are independent. Let $W_i = 2(n-i+1)Z_i/\beta$ for $i = 2, \ldots, n$. The pdf of W_i is

$$f_{W_i}(w_i) = \frac{1}{2} e^{-w_i/2} I(0 < w_i < \infty),$$

which is a pdf of $\chi^2_2 \equiv \text{Gamma}(1, 2)$.

(b) $X_{(1)}$ and Y are independent, where $Y = (n-1)^{-1} \sum_{i=1}^n (X_i - X_{(1)})$.

Solution: One can see $Y = (n-1)^{-1} \sum_{i=1}^n (X_i - X_{(1)}) = (n-1)^{-1} \sum_{i=1}^n (X_{(i)} - X_{(1)}) = (n-1)^{-1} \sum_{i=2}^n (X_{(i)} - X_{(1)}) = (n-1)^{-1} \sum_{i=2}^n (n-i+1) Z_i$. One can easily see that Y is a function of $Z_2, \ldots, Z_n$, which are mutually independent with $Z_i = X_{(1)}$. We can claim $X_{(1)}$ and Y are independent.

(c) (Bios 673 class material, no need to return) $T = (X_{(1)} - \alpha)/Y$ has a pdf

$$f_Y(t) = n \left(1 + \frac{nt}{n-1}\right)^{-n},$$

for $0 < t < \infty$ and 0 otherwise.

Solution: The pdf of $T_1 = (X_{(1)} - \alpha)/\beta$ is $f_{T_1}(t_1) = ne^{-nt_1} I(0 < t_1 < \infty)$. Following (b), $T_2 = 2(n-1)Y/\beta = \sum_{i=2}^n 2(n-i+1)Z_i/\beta = \sum_{i=2}^n W_i$ follows $\chi^2_{2(n-1)}$ distribution. According to (a), T_1 and T_2 are independent. Let $U = T_1/T_2$ and $V = T_2$ with inverse function $T_1 = UV$ and $T_2 = V$. Jacobian is v. The joint pdf of (U, V) is

$$f_{U,V}(u, v) = f_{T_1,T_2}(uv, v) = ne^{-nuv} \frac{1}{\Gamma(n-1) 2^{n-1}} v^{n-1} e^{-v/2}.$$

The marginal pdf of U is

$$f_U(u) = \int_0^\infty f_{U,V}(u, v) dv = \frac{n(n-1)(nu + 1/2)^{-n}}{2^{n-1}}.$$

Now, we need to derive the pdf of $T = 2(n-1)U$. That is,

$$f_T(t) = \frac{n(n-1)\{\frac{n}{2(n-1)} t + 1/2\}^{-n}}{2^{n-1} 2(n-1)} = n \left(1 + \frac{nt}{n-1}\right)^{-n}, \quad 0 < t < \infty.$$

7. (Bios 673 class material) Let $X_1, \ldots, X_n$ be a random sample from a distribution with unknown mean $\mu \in \mathcal{R}$ and unknown variance $\sigma^2 > 0$. Let $\bar{X}$ and S^2 be the sample mean and sample variance, respectively. One is interested in comparing three estimators, $T_1 = \bar{X}^2$, $T_2 = \bar{X}^2 - S^2/n$, and $T_3 = \max\{0, T_2\}$ for μ^2.

(a) When $\mu \neq 0$, show that three estimators have the same limiting distribution, i.e.,

$$\sqrt{n}(T_i - \mu^2) \to_d N(0, \tau^2),$$

for $i = 1, 2, 3$. Express τ^2 as a function of μ and σ^2.

Solution: By Central Limit Theorem, we have

$$\sqrt{n}(\bar{X} - \mu) \to_d N(0, \sigma^2).$$

By Delta Method, we have

$$\sqrt{n}(T_1 - \mu^2) \to_d N(0, 4\mu^2\sigma^2).$$

One can show that

$$\sqrt{n}(T_2 - \mu^2) = \sqrt{n}(T_1 - \mu^2) - \sqrt{n}\frac{S^2}{n} \to_d N(0, 4\mu^2\sigma^2),$$

by Slutsky Theorem, since $S^2 \to_p \sigma^2$ and $S^2/\sqrt{n} \to_p 0$. One can also show that T_3 has the same limiting distribution as T_2 since

$$\lim_{n\to\infty} P(T_2 < 0) = \lim_{n\to\infty} P\left(\frac{\sqrt{n}(T_2 - \mu^2)}{\sqrt{4\mu^2\sigma^2}} < -\frac{\sqrt{n}\mu^2}{\sqrt{4\mu^2\sigma^2}}\right) = \Phi(-\infty) = 0.$$

(b) When $\mu = 0$, show that

$$nT_1 \to_d \sigma^2 W,$$

and

$$nT_2 \to_d \sigma^2(W - 1),$$

where W follows a χ^2_1 distribution.

Solution: When $\mu = 0$, one can see $\sqrt{n}\bar{X}/\sigma \to_d N(0, 1)$. Hence $nT_1 \to_d \sigma^2 W$, where W follows a χ^2_1 distribution. By the definition of T_2, one can have

$$nT_2 = n(T_1 - \frac{S^2}{n}) = nT_1 - S^2.$$

Since $T_1 \to_d \sigma^2 W$ and $S^2 \to_p \sigma^2$, we can claim that $T_2 \to_d \sigma^2(W - 1)$ by Slutsky's Theorem.

(c) When $\mu = 0$, show that T_2 has a smaller asymptotic mean square error (AMSE) than T_1, where AMSE is defined by EX^2/a_n^2 when $a_n X_n \to_d X$ with $EX^2 < \infty$.

Solution: By the definition of AMSE,

$$AMSE(T_1) = E(\sigma^4 W^2)/n^2 = 3\sigma^4/n^2,$$

and

$$AMSE(T_2) = E(\sigma^4 (W-1^) 2)/n^2 = 2\sigma^4/n^2.$$

One can see the AMSE of T_2 is smaller than the AMSE of T_1. The asymptotic relative efficiency (ARE) is $AMSE(T_1)/AMSE(T_2) = 1.5$.

(d) When $\mu = 0$, T_3 is in fact a better estimator of μ^2 with the smallest AMSE. Without any theoretical proof, comment on why this is true.

Solution: When $\mu = 0$, the probability of T_3 being 0 is $P(T_2 < 0) > 0$. Hence, T_3 has a better chance of hitting the true value $\mu = 0$. A smaller variation is anticipated for T_3 since whenever a value is repeated, the variation is less. Therefore, one can anticipate T_3 has a smaller AMSE.