Statistical Inference

B. Statistical Data Science 2nd Year Indian Statistical Institute

Teacher: Soham Sarkar

Exercise Series 5

Exercise 1. In this exercise, we will derive the density of the range and mid-range of a random sample from the $\mathsf{Uniform}$ distribution. For a random sample $(X_1, \ldots, X_n)$, we call $R = X_{(n)} - X_{(1)}$ to be its range. The mid-range is defined as $V = \dfrac{X_{(1)} + X_{(n)}}{2}$.

(a) Let $X_1, \ldots, X_n \overset{\text{i.i.d.}}{\sim} \mathsf{Uniform}(0, \theta)$, $\theta > 0$. Find the joint pdf of $X_{(1)}$ and $X_{(n)}$.

(b) Show that the joint pdf of R and V is given by

$$f_{R,V}(r, v) = n(n-1)\,\frac{r^{n-2}}{\theta^n}, \qquad 0 < r < \theta, \frac{r}{2} < v < \theta - \frac{r}{2}.$$

Hint: *Use the Jacobian formula with* $r = z - y, v = \dfrac{y+z}{2}$. *With this transformation,* $y = v - \dfrac{r}{2}$ *and* $z = v + \dfrac{r}{2}$. *Also, for* $\mathsf{Uniform}(0, \theta)$, $0 < y < \theta$ *and* $0 < z < \theta$. *Use these to deduce the joint support of* (R, V).

(c) Show that the marginal pdf of R is given by

$$f_R(r) = n(n-1)\,\frac{r^{n-2}(\theta - r)}{\theta^n}, \qquad 0 < r < \theta.$$

Hence, deduce that $R/\theta \sim \mathsf{Beta}(n-1, 2)$.

Hint: *The marginal of* R *is obtained by integrating the joint pdf w.r.t.* v *over the appropriate range. Here, we have* $\dfrac{r}{2} < v < \theta - \dfrac{r}{2}$.

(d) Show that the marginal pdf of V is given by

$$f_V(v) = \begin{cases} n2^{n-1}\,\dfrac{v^{n-1}}{\theta^n} & \text{if } 0 < v < \dfrac{\theta}{2}, \\ n2^{n-1}\,\dfrac{(\theta - v)^{n-1}}{\theta^n} & \text{if } \dfrac{\theta}{2} < v < \theta. \end{cases}$$

Verify that the pdf of V is symmetric about $\dfrac{\theta}{2}$ and the mode of the pdf is at $\dfrac{\theta}{2}$.

Hint: *Notice that* $\dfrac{r}{2} < v < \theta - \dfrac{r}{2}$ *implies* $r < 2v$ *and* $r < 2(\theta - v)$. *Also,* $2v \gtreqless 2(\theta - v) \Leftrightarrow v \gtreqless \dfrac{\theta}{2}$. *You can also find out the range of integration by drawing the support of* (R, V).

(e) Calculate expectations and variances of R and V.

(f) Find the joint pdf of (R, V) for a random sample from $\mathsf{Uniform}(a, b)$, $-\infty < a < b < \infty$. Find the marginal pdfs, expectations and variances of R and V.

Hint: *You can modify and repeat all the steps. Alternatively, notice that for* $X_1, \dots, X_n \overset{\text{i.i.d.}}{\sim} \mathsf{Uniform}(a, b)$, $X_1 - a, \dots, X_n - a \overset{\text{i.i.d.}}{\sim} \mathsf{Uniform}(0, b-a)$. *If we define* $Y_i = X_i - a$, *then* $Y_{(1)} = X_{(1)} - a$ *and* $Y_{(n)} = X_{(n)} - a$, *so that* $R = X_{(n)} - X_{(1)} = Y_{(n)} - Y_{(1)}$ *and* $V = \dfrac{X_{(1)} + X_{(n)}}{2} = \dfrac{Y_{(1)} + Y_{(n)}}{2} + a$.

Exercise 2. Let $X_1, \dots, X_n \overset{\text{i.i.d.}}{\sim} \mathsf{Uniform}(0, \theta)$, $\theta > 0$.

(a) Find the joint density of $\dfrac{X_{(1)}}{X_{(n)}}$ and $X_{(n)}$.

Hint: *Start with the joint pdf of* $X_{(1)}$ *and* $X_{(n)}$, *and use the Jacobian formula.*

(b) Find the marginal distributions of $\dfrac{X_{(1)}}{X_{(n)}}$ and $X_{(n)}$. Verify that $\dfrac{X_{(1)}}{X_{(n)}}$ and $X_{(n)}$ are independent.

(c) Verify that the distribution of $\dfrac{X_{(1)}}{X_{(n)}}$ is free of θ. Hence, deduce that $\dfrac{X_{(1)}}{X_{(n)}}$ is an ancillary statistic.

(d) Repeat (a)–(c) with $X_{(i)}$ and $X_{(j)}$ for $1 \le i < j \le n$. Show that $\dfrac{X_{(i)}}{X_{(j)}}$ and $X_{(j)}$ are independent, and $\dfrac{X_{(i)}}{X_{(j)}}$ is ancillary.

Exercise 3. Let $X_1, \dots, X_n$ be i.i.d. random variables with pdf f_X and cdf F_X. Find the conditional pdf of $X_{(i)}$ given $X_{(j)}$.

Hint: *Recall that the conditional pdf of* X *given* Y *is* $f_{X,Y}(x, y)/f_Y(y)$.

Exercise 4. Let $X_1, \dots, X_n$ be i.i.d. random variables with pdf f_X and cdf F_X. Find the conditional pdf of $R = X_{(n)} - X_{(1)}$ given $V = \dfrac{X_{(1)} + X_{(n)}}{2}$.

Exercise 5. Let $X_1, \dots, X_n \overset{\text{i.i.d.}}{\sim} f_\theta$ with $f_\theta(x) = \dfrac{1}{\theta} f\left(\dfrac{x}{\theta}\right)$. Here, $f(\cdot)$ is a completely known pdf (free of unknown parameters).

(a) Show that the cdf corresponding to f_θ is of the form $F_\theta(x) = F\left(\dfrac{x}{\theta}\right)$, where $F(\cdot)$ is the cdf corresponding to the pdf $f(\cdot)$.

(b) Show that the distribution of $\dfrac{X_j}{X_i}$ is free of θ for every i, j. Hence, $\dfrac{X_j}{X_i}$ is ancillary.

(c) Show that the distribution of $\dfrac{X_{(j)}}{X_{(i)}}$ is free of θ for every i, j. Hence, $\dfrac{X_{(j)}}{X_{(i)}}$ is ancillary.

(d) Show that the joint distribution of $\dfrac{X_1}{X_n}, \dfrac{X_2}{X_n}, \dots, \dfrac{X_{n-1}}{X_n}$ is free of θ. Therefore, $\left(\dfrac{X_1}{X_n}, \dfrac{X_2}{X_n}, \dots, \dfrac{X_{n-1}}{X_n}\right)$ is ancillary.