HOMEWORK 02

PSTAT 120B: Mathematical Statistics, I
Summer Session A, 2024 with Instructor: Ethan P. Marzban

1. (**6.23, Modified**) Let Y be a random variable with density
$$f_Y(y) = 2(1-y) \cdot \mathbb{1}_{\{0 \leq y \leq 1\}}$$
Define:
$$\begin{aligned} U_1 &:= 2Y - 1 \\ U_2 &:= 1 - 2Y \\ U_3 &:= Y^2 \end{aligned}$$

(a) Compute $\mathbb{E}[U_1]$, $\mathbb{E}[U_2]$, and $\mathbb{E}[U_3]$ without first finding the densities of $U_1,\ U_2$, and U_3. **Hint:** LOTUS.

Solution:
$$\begin{aligned} \mathbb{E}[U_1] = \mathbb{E}[2Y-1] &= \int_{-\infty}^{\infty} (2y-1) f_Y(y)\ \mathrm{d}y \\ &= \int_0^1 (2y-1) \cdot 2(1-y)\ \mathrm{d}y \\ &= \int_0^1 (6y - 4y^2 - 2)\ \mathrm{d}y = 3 - \frac{4}{3} - 2 = \boxed{-\frac{1}{3}} \\ \mathbb{E}[U_2] = \mathbb{E}[1-2Y] &= -\mathbb{E}[2Y-1] = -\mathbb{E}[U_1] = -\left(-\frac{1}{3}\right) = \boxed{\frac{1}{3}} \\ \mathbb{E}[U_3] = \mathbb{E}[Y^2] &= \int_{-\infty}^{\infty} y^2 f_Y(y)\ \mathrm{d}y \\ &= \int_0^1 y^2 \cdot 2(1-y)\ \mathrm{d}y \\ &= \int_0^1 (2y^2 - 2y^3)\ \mathrm{d}y = \frac{2}{3} - \frac{1}{2} = \boxed{\frac{1}{6}} \end{aligned}$$

(b) Find $f_{U_1}(u),\ f_{U_2}(u)$, and $f_{U_3}(u)$, the densities of U_1, U_2, and U_3, using the **CDF Method**.

Solution: Let's first find the CDF of Y. Since the support of Y is $[0, 1]$, we see that $F_Y(y) = 0$ whenever $y < 0$ and $F_Y(y) = 1$ whenever $y \geq 1$. Hence, fix a $y \in [0, 1)$ and compute
$$F_Y(y) = \int_0^y 2(1-t)\ \mathrm{d}t = 2y - y^2$$
That is to say,
$$F_Y(y) = \begin{cases} 0 & \text{if } y < 0 \\ 2y - y^2 & \text{if } 0 \leq y < 1 \\ 1 & \text{if } y \geq 1 \end{cases}$$

- <u>Considering U_1:</u>

$$F_{U_1}(u) := \mathbb{P}(U_1 \leq u) = \mathbb{P}(2Y - 1 \leq u) = \mathbb{P}\left(Y \leq \frac{u+1}{2}\right) = F_Y\left(\frac{u+1}{2}\right)$$

Now, since the support of Y is $S_Y = [0, 1]$ we see that the support of U_1 will be $S_{U_1} = [-1, 1]$. Hence, for a $u \in [-1, 1]$ we have $y \in [0, 1]$ and

$$F_{U_1}(u) = F_Y\left(\frac{u+1}{2}\right) = 2\left(\frac{u+1}{2}\right) - \left(\frac{u+1}{2}\right)^2$$

Differentiating wrt. u and re-incorporating the support of U_1, we find

$$f_{U_1}(u) = \frac{1-u}{2} \cdot \mathbb{1}_{\{u \in [-1,1]\}}$$

- <u>Considering U_2:</u>

$$F_{U_2}(u) := \mathbb{P}(U_2 \leq u) = \mathbb{P}(1 - 2Y \leq u) = \mathbb{P}\left(Y \geq \frac{1-u}{2}\right) = 1 - F_Y\left(\frac{1-u}{2}\right)$$

Now, since the support of Y is $S_Y = [0, 1]$ we see that the support of U_2 will be $S_{U_2} = [-1, 1]$. Hence, for a $u \in [-1, 1]$ we have $y \in [0, 1]$ and

$$F_{U_2}(u) = 1 - F_Y\left(\frac{1-u}{2}\right) = 1 - 2\left(\frac{1-u}{2}\right) + \left(\frac{1-u}{2}\right)^2 = \left(\frac{u+1}{2}\right)^2$$

Differentiating wrt. u and re-incorporating the support of U_2, we find

$$f_{U_2}(u) = \frac{u+1}{2} \cdot \mathbb{1}_{\{u \in [-1,1]\}}$$

- <u>Considering U_3:</u>

$$F_{U_3}(u) := \mathbb{P}(U_3 \leq u) = \mathbb{P}(Y^2 \leq u) = \mathbb{P}(-\sqrt{u} \leq Y \leq \sqrt{u}) = F_Y(\sqrt{u}) - F_Y(-\sqrt{u})$$

Now, since the support of Y is $S_Y = [0, 1]$ we see that the support of U_3 will be $S_{U_3} = [0, 1]$. So, let's fix a $u \in [0, 1]$; note then that $F_Y(-\sqrt{u}) = 0$, and so

$$F_{U_3}(u) = F_Y(\sqrt{u}) = 2\sqrt{u} - u$$

Differentiating wrt. u and re-incorporating the support of U_2, we find

$$f_{U_3}(u) = \left(\frac{1}{\sqrt{u}} - 1\right) \cdot \mathbb{1}_{\{u \in [0,1]\}}$$

(c) Find $f_{U_1}(u),\ f_{U_2}(u)$, and $f_{U_3}(u)$, the densities of U_1, U_2, and U_3, using the **Change of Variable formula**.

Solution:

- <u>Considering U_1</u>: the transformation is $g(y) = 2y - 1$, which has inverse

$$g^{-1}(u) = \frac{u+1}{2} \implies \left| \frac{\mathsf{d}}{\mathsf{d}u} g^{-1}(u) \right| = |1/2| = 1/2$$

  Therefore, by the change of variable formula,

$$\begin{aligned} f_{U_1}(u) &= f_Y[g^{-1}(u)] \cdot \left| \frac{\mathsf{d}}{\mathsf{d}u} g^{-1}(u) \right| \\ &= 2\left(1 - \frac{u+1}{2}\right) \cdot \frac{1}{2} \cdot \mathbb{1}_{\{\frac{u+1}{2} \in [0,1]\}} \\ &= \boxed{\frac{1-u}{2} \cdot \mathbb{1}_{\{u \in [-1,1]\}}} \end{aligned}$$

- <u>Considering U_2</u>: the transformation is $g(y) = 1 - 2y$, which has inverse

$$g^{-1}(u) = \frac{1-u}{2} \implies \left| \frac{\mathsf{d}}{\mathsf{d}u} g^{-1}(u) \right| = |-1/2| = 1/2$$

  Therefore, by the change of variable formula,

$$\begin{aligned} f_{U_2}(u) &= f_Y[g^{-1}(u)] \cdot \left| \frac{\mathsf{d}}{\mathsf{d}u} g^{-1}(u) \right| \\ &= 2\left(1 - \frac{1-u}{2}\right) \cdot \frac{1}{2} \cdot \mathbb{1}_{\{\frac{1-u}{2} \in [0,1]\}} \\ &= \boxed{\frac{u+1}{2} \cdot \mathbb{1}_{\{u \in [-1,1]\}}} \end{aligned}$$

- <u>Considering U_2</u>: the transformation is $g(y) = y^2$. Again, this transformation is *not* strictly monotonic over $\mathbb{R}$, but it *is* strictly monotonic over $S_Y = [0,1]$, with inverse given by

$$g^{-1}(u) = \sqrt{u} \implies \left| \frac{\mathsf{d}}{\mathsf{d}u} g^{-1}(u) \right| = \left| \frac{1}{2\sqrt{u}} \right| = \frac{1}{2} \cdot \left| \frac{1}{\sqrt{u}} \right|$$

  Therefore, by the change of variable formula,

$$\begin{aligned} f_{U_3}(u) &= f_Y[g^{-1}(u)] \cdot \left| \frac{\mathsf{d}}{\mathsf{d}u} g^{-1}(u) \right| \\ &= 2\left(1 - \sqrt{u}\right) \cdot \cdot \left| \frac{1}{\sqrt{u}} \right| \cdot \mathbb{1}_{\{\sqrt{u} \in [0,1]\}} \\ &= \boxed{\left(\frac{1}{\sqrt{u}} - 1\right) \cdot \mathbb{1}_{\{u \in [0,1]\}}} \end{aligned}$$

(d) Recompute $\mathbb{E}[U_1]$, $\mathbb{E}[U_2]$, and $\mathbb{E}[U_3]$, now using the densities you derived in parts (b) and (c) above.

Solution:

$$\begin{aligned}
\mathbb{E}[U_1] &:= \int_{\mathbb{R}} u \cdot f_{U_1}(u) \, \mathrm{d}u \\
&= \int_{-1}^{1} u \cdot \frac{1-u}{2} \, \mathrm{d}u = \frac{1}{2} \int_{-1}^{1} (u - u^2) \, \mathrm{d}u = \frac{1}{2}\left[0 - \frac{2}{3}\right] = \boxed{-\frac{1}{3}} \\
\mathbb{E}[U_2] &:= \int_{\mathbb{R}} u \cdot f_{U_2}(u) \, \mathrm{d}u \\
&= \int_{-1}^{1} u \cdot \frac{u+1}{2} \, \mathrm{d}u = \frac{1}{2} \int_{-1}^{1} (u^2 + u) \, \mathrm{d}u = \frac{1}{2}\left[\frac{2}{3} - 0\right] = \boxed{\frac{1}{3}} \\
\mathbb{E}[U_3] &:= \int_{\mathbb{R}} u \cdot f_{U_3}(u) \, \mathrm{d}u \\
&= \int_{0}^{1} u \left(\frac{1}{\sqrt{u}} - 1\right) \, \mathrm{d}u \int_{0}^{1} (\sqrt{u} - u) \, \mathrm{d}u = \frac{2}{3} - \frac{1}{2} = \boxed{\frac{1}{6}}
\end{aligned}$$

2. Let $Y \sim \text{Exp}(\theta)$, and set $U := \alpha Y + \delta$ for positive constants α, β. Find the density $f_U(u)$ of u, using whichever method you like. As an aside: the distribution of U is called the **two-parameter exponential distribution**.

Solution: Either the CDF method or the Change of Variable formula would work here. I'll demonstrate using the Change of Variable formula: take $g(y) = \alpha y + \delta$, so that

$$g^{-1}(u) = \frac{u - \delta}{\alpha} \implies \left| \frac{\mathrm{d}}{\mathrm{d}u} g^{-1}(u) \right| = \left| \frac{1}{\alpha} \right| = \frac{1}{\alpha}$$

where we have dropped the absolute values in the final step, since α is assumed to be positive. Therefore, by the change of variable formula,

$$\begin{aligned}
f_U(u) &= f_Y[g^{-1}(u)] \cdot \left| \frac{\mathrm{d}}{\mathrm{d}u} g^{-1}(u) \right| \\
&= \frac{1}{\theta} \exp\left\{ -\frac{\left(\frac{u-\delta}{\alpha}\right)}{\theta} \right\} \cdot \mathbb{1}_{\left\{\left\{\frac{u-\delta}{\alpha} \geq 0\right\}\right\}} = \boxed{\frac{1}{\alpha\theta} \exp\left\{ -\frac{u - \delta}{\alpha\theta} \right\} \cdot \mathbb{1}_{\{u \geq \delta\}}}
\end{aligned}$$

Note that this is essentially just the density of the Exp$(\alpha\theta)$ distribution, shifted δ units to the right.

3. The **Rayleigh Distribution**, which admits a single parameter $\beta > 0$, is widely used throughout statistics and engineering. If $X \sim \text{Ray}(\beta)$, then X has density

$$f_X(x) = \frac{2x}{\beta} \cdot e^{-x^2/\beta} \cdot \mathbb{1}_{\{x \geq 0\}}$$

   (a) Let $Y \sim \text{Exp}(\theta)$, and set $U := \sqrt{Y}$. Show that U follows the Rayleigh distribution, and identify its parameter.

Solution: Either the CDF method or the Change of Variable formula would work here. For variety's sake, I'll demonstrate using the CDF method:

$$F_U(u) := \mathbb{P}(U \le u) = \mathbb{P}(\sqrt{Y} \le u) = \mathbb{P}(Y \le u^2) = F_Y(u^2)$$
$$= \begin{cases} 0 & \text{if } u^2 < 0 \\ 1 - e^{-u^2/\theta} & \text{if } u^2 \ge 0 \end{cases} = \begin{cases} 0 & \text{if } u < 0 \\ 1 - e^{-u^2/\theta} & \text{if } u \ge 0 \end{cases}$$

Differentiating wrt. u yields

$$f_U(u) = \frac{2u}{\theta} \cdot e^{-u^2/\theta} \cdot \mathbb{1}_{\{u \ge 0\}}$$

which is indeed the density of the Ray(θ) distribution.

(b) Let $Y_1, Y_2 \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$. Show that $R := \sqrt{Y_1^2 + Y_2^2}$ follows the Rayleigh distribution, and identify its parameter. As an aside: note how this implies the Rayleigh distribution is well-suited for modeling distances! **Hint:** consider using previously-derived results from lecture, along with the result of part (a) above.

Solution: There are a couple of ways to go about this problem. Here's how I think through it:

- We've previously seen that $Y_1^2, Y_2^2 \overset{\text{i.i.d.}}{\sim} \chi_2^2 \overset{\text{d}}{=} \chi_1^2 \overset{\text{d}}{=}$ Gamma$(1/2\ ,\ 2)$
- Hence, $(Y_1^2 + Y_2^2) \sim$ Gamma$(1/2 + 1/2\ ,\ 2) \overset{\text{d}}{=}$ Gamma$(1, 2) \overset{\text{d}}{=}$ Exp(2)
- By the result of part (a), the square root of a Exp(θ) distribution follows the Ray(θ) distribution. Therefore, $R := \sqrt{Y_1^2 + Y_2^2}$ must follow a Ray(2) distribution.

4. Consider a collection $\{X_i\}_{i=1}^n$ of random variables, and define the sample mean in the usual manner:

$$\overline{X}_n := \frac{1}{n} \sum_{i=1}^{n} X_i$$

(a) Suppose that $X_i \sim \chi^2_{\nu_i}$ for positive integers $\{\nu_i\}_{i=1}^n$. Use the MGF method to derive the distribution of $\overline{X}_n$. (Yes, there is a way to do this using the closure property of the Gamma distribution, but I'd like you to use the MGF method for this part.) **You may assume the X_i are independent.**

Solution: Since $X_i \sim \chi^2_{\nu_i}$, we have

$$M_{X_i} = \begin{cases} (1-2t)^{-\nu_i/2} & \text{if } t < 1/2 \\ \infty & \text{otherwise} \end{cases}$$

Hence, applying the theorem titled "Important MGF Formula" with $a_i = 1/n$,

$$M_{\overline{X}_n}(t) \prod_{i=1}^{n} M_{Y_i}\left(\frac{1}{n} t\right)$$

$$= \prod_{i=1}^{n} \left(\begin{cases} (1 - 2t/n)^{-\nu_i/2} & \text{if } t < 1/2 \\ \infty & \text{otherwise} \end{cases} \right)$$

$$= \begin{cases} (1 - (2/n)t)^{-\sum_{i=1}^{n} \nu_i/2} & \text{if } t < 1/2 \\ \infty & \text{otherwise} \end{cases}$$

which we recognize as the MGF of the Gamma$(\sum_{i=1}^{n} \nu_i/2 \, , \, 2/n)$ distribution. In other words:

$$\overline{X}_n \sim \text{Gamma}\left(\sum_{i=1}^{n} \nu_i/2 \, , \, \frac{2}{n} \right)$$

(b) Suppose that $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$ for constants $\{\mu_i\}_{i=1}^{n}$ and positive constants $\{\sigma_i^2\}_{i=1}^{n}$. Derive the distribution of $\overline{X}_n$ (for this part you can use previously-derived results from lecture). **You may assume the X_i are independent.**

Solution: From the result titled "Closure of Normal Distribution under Linear Combinations)" with $a_i = 1/n$ we have

$$\frac{1}{n} \sum_{i=1}^{n} X_i \sim \mathcal{N}\left(\frac{1}{n} \sum_{i=1}^{n} \mu_i \, , \, \frac{1}{n^2} \sum_{i=1}^{n} \sigma_i^2 \right)$$