1. (WMS, Problem 6.15.) Let Y have a distribution function given by

$$F(y) = \begin{cases} 0 & y < 0, \\ 1 - e^{-y^2}, & y \geq 0. \end{cases}$$

Find a transformation $G(U)$ such that, if U has a uniform distribution on the interval (0,1), $G(U)$ has the same distribution as Y.

Solution. By probability integral transformation, if Y has CDF F and $U \sim \text{Unif}(0,1)$, then $X = F^{-1}(U)$ has the same distribution as Y. In this problem, for $y \geq 0$

$$u = F(y) = 1 - e^{-y^2} \implies e^{-y^2} = 1 - u \implies y^2 = -\log(1-u) \implies y = \sqrt{-\log(1-u)}.$$

Therefore, required transformation is $G(U) = F^{-1}(U) = \sqrt{-\log(1-U)}$. □

2. (WMS, Problem 6.20) Let the random variable Y possess a uniform distribution on the interval (0,1). Derive the

(a) distribution of the random variable $W_1 = Y^2$.

Solution. Because the distribution of a RV is specified by both PDF and CDF, here we'll find the CDF. Note that the support of W_1 is also $[0,1]$. Hence, for $0 \leq w \leq 1$, the CDF of W_1 is given by

$$F_{W_1}(w) = P(W_1 \leq w) = P(Y^2 \leq w) = P(-\sqrt{w} \leq Y \leq \sqrt{w}) = P(Y \leq \sqrt{w}) = F_Y(\sqrt{w}) = \sqrt{w}$$

Hence, the CDF of W_1 is given by

$$F_{W_1}(w) = \begin{cases} 0, & w < 0, \\ \sqrt{w}, & 0 \leq w \leq 1 \\ 1, & w > 1. \end{cases}$$

□

(b) distribution of the random variable $W_2 = \sqrt{Y}$.

Solution. The support of W_2 is also $[0,1]$. Hence, for $0 \leq w \leq 1$, the CDF of W_2 is given by

$$F_{W_1}(w) = P(W_1 \leq w) = P(\sqrt{Y} \leq w) = P(Y \leq w^2) = F_Y(w^2) = w^2$$

Hence, the CDF of W_1 is given by

$$F_{W_2}(w) = \begin{cases} 0, & w < 0, \\ w^2, & 0 \leq w \leq 1 \\ 1, & w > 1. \end{cases}$$

□

3. (WMS, Problem 6.33.) The proportion of impurities in certain ore samples is a random variable Y with a density function given by

$$f(y) = \begin{cases} (3/2)y^2 + y, & 0 \le y \le 1, \\ 0, & \text{elsewhere.} \end{cases}$$

The dollar value of such samples is $U = 5 - (Y/2)$. Find the probability density function for U.

Solution. Note that $u = g(y) = 5 - \frac{y}{2} \implies y = 10 - 2u = h(u)$, which implies, $\frac{d}{du}h(u) = -2$. Also,

$$0 \le y \le 1 \implies 0 \le \frac{y}{2} \le \frac{1}{2} \implies 5 - \frac{1}{2} = \frac{9}{2} \le 5 - \frac{y}{2} = u \le 5$$

which means the support of U is $\mathscr{U} = [9/2, 5]$. By the method of transformation, the PDF of U is therefore given by,

$$f_U(u) = f_Y(h(u)) \left| \frac{d}{du} h(u) \right| = 2 \left[\frac{3}{2}(10 - 2u)^2 + (10 - 2u) \right], \quad \frac{9}{2} \le u \le 5$$

□

4. Suppose $X \sim \text{Bin}(n, p)$. Find the distribution of $Y = n - X$. For which value of p will Y and X have the same distribution?

Solution. The PMF of X is given by

$$P(X = x) = \binom{n}{x} p^x q^{n-x}, \quad x = 0, 1, \cdots, n.$$

where $q = 1 - p$. Note that the support of $Y = n - X$ is also $\{0, 1, \cdots, n\}$. Therefore, for $y = 0, 1, \cdots, n$,

$$P(Y = y) = P(n - X = y) = P(X = n - y) = \binom{n}{n-y} p^{n-y} q^{n-(n-y)} = \binom{n}{y} q^y p^{n-y}.$$

Hence $Y \sim \text{Bin}(n, q)$. Clearly, X and Y will have the same distribution if $p = q = 1 - p \implies p = 1/2$. □

5. Let $Z \sim N(0, 1)$. Use the method of transformation to show that $X = \mu + \sigma Z$, for $\sigma > 0$ has $N(\mu, \sigma^2)$ distribution.

Solution. The PDF of Z is given by

$$f_Z(z) = \frac{1}{\sqrt{2\pi)}} e^{-\frac{1}{2}z^2}, \ -\infty < z < \infty$$

Here $X = \mu + \sigma Z$. So the support of X is also $(-\infty, \infty)$. Also, $x = g(z) = \mu + \sigma z$ is an increasing function of z (since $\sigma > 0$) and $x = \mu + \sigma z \implies z = (x - \mu)/\sigma = h(x)$, with $\frac{d}{dx}h(x) = 1/\sigma$. Therefore, for $x \in (-\infty, \infty)$, the PDF of X is given by

$$f_X(x) = f_Z(h(x)) \left| \frac{d}{dx} h(x) \right| = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}.$$

which is the PDF of $N(\mu, \sigma^2)$. □

6. (WMS, Problem 6.88.) Suppose that the length of time Y it takes a worker to complete a certain task has the probability density function given by

$$f(y) = \begin{cases} e^{-(y-\theta)}, & y > \theta, \\ 0, & \text{elsewhere,} \end{cases}$$

where θ is a positive constant that represents the minimum time until task completion. Let $Y_1, \cdots, Y_n$ denote a random sample of completion times from this distribution. Find

(a) the density function for $Y_{(1)} = \min(Y_1, Y_2, ..., Y_n)$.

Solution. Note that for all $i = 1, \cdots, n$, Y_i has support (θ, ∞) and CDF

$$F(y) = \int_\theta^y e^{-(x-\theta)}\, dx = \left[-e^{-(x-\theta)}\right]_\theta^y = 1 - e^{-(y-\theta)}, \quad y \geq \theta.$$

Therefore, for $0 \leq y \leq \theta$, the CDF of $Y_{(1)}$ is given by

$$\begin{aligned} F_{Y_{(1)}}(y) &= P(Y_{(1)} \leq y) \\ &= 1 - P(Y_{(1)} > y) \\ &= 1 - P(Y_1 > y, \cdots, Y_n > y) \\ &= 1 - P(Y_1 \leq y) \cdots P(Y_n \leq y) & \text{(independence)} \\ &= 1 - \{1 - F(y)\}^n = 1 - \left(e^{-(y-\theta)}\right)^n = 1 - e^{-n(y-\theta)}. \end{aligned}$$

Hence, the PDF of $Y_{(1)}$ is given by:

$$f_{Y_{(1)}}(y) = \begin{cases} \frac{d}{dy}\left(1 - e^{-n(y-\theta)}\right) = ne^{-n(y-\theta)}, & y \geq \theta, \\ 0, & \text{otherwise.} \end{cases}$$

□

(b) $E(Y_{(1)})$.

Solution. We have

$$\begin{aligned} E(Y_{(1)}) &= \int_\theta^\infty y\, ne^{-n(y-\theta)}\, dy \\ &= \int_0^\infty (x+\theta)\, ne^{-nx}\, dx & \text{(substitute } x = y - \theta) \\ &= \underbrace{\int_0^\infty x\, ne^{-nx}\, dx}_{\text{mean of Exp}(\beta=1/n)} + \theta \underbrace{\int_0^\infty ne^{-nx}\, dx}_{\text{integral of Exp}(\beta=1/n)\text{ PDF}} \\ &= \frac{1}{n} + \theta. \end{aligned}$$

□

7. Let $Y_1, Y_2, \cdots, Y_n$ be independent, uniformly distributed random variables on the interval $[0, \theta]$. Find the PDF of $Y_{(n)} = \max\{Y_1, Y_2, ..., Y_n\}$.

Solution. Recall that the CDF of Y_i, for $i = 1, \cdots, n$ is given by:

$$F(y) = \begin{cases} 0, & y < 0, \\ \frac{y}{\theta}, & 0 \le y \le \theta, \\ 1, & y > 1. \end{cases}$$

Note that the support of (the marginal distribution of) $Y_{(n)}$ is the same as that of any Y_i, $i = 1, \cdots, n$, which is $[0, \theta]$. Therefore, for $0 \le y \le \theta$, the CDF of $Y_{(n)}$ is given by

$$F_{Y_{(n)}}(y) = P(Y_{(n)} \le y) = P(Y_1 \le y, \cdots, Y_n \le y) \overset{\text{indep}}{=} P(Y_1 \le y) \cdots P(Y_n \le y) = \{F(y)\}^n = \left(\frac{y}{\theta}\right)^n.$$

Hence, the PDF of $Y_{(n)}$ is given by:

$$f_{Y_{(n)}}(y) = \begin{cases} \frac{d}{dy}\left(\frac{y}{\theta}\right)^n = n\frac{y^{n-1}}{\theta^n}, & 0 \le y \le \theta, \\ 0, & \text{otherwise.} \end{cases}$$

□

8. Suppose $X_1, \cdots, X_n$ denote a random sample from Exp(β) distribution. Find the distribution of the sample mean $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$.

Solution. Recall that the MGF of $X_i \sim$ Exp(β) is $M_{X_i}(t) = (1 - \beta t)^{-1}$. Therefore, the MGF of $\overline{X}$ is given by

$$\begin{aligned} M_{\overline{X}}(t) &= E\left(e^{t\overline{X}}\right) \\ &= E\left(e^{\frac{t}{n}\sum_{i=1}^n X_i}\right) \\ &= E\left(e^{\frac{t}{n}X_1} \cdots e^{\frac{t}{n}X_1}\right) \\ &= E\left(e^{\frac{t}{n}X_1}\right) \cdots E\left(e^{\frac{t}{n}X_1}\right) && \text{(independence)} \\ &= M_{X_1}(t/n) \cdots M_{X_n}(t/n) = (1 - (\beta/n)\, t)^{-n} \end{aligned}$$

which is the MGF of Gamma($\alpha = n, \beta = \beta/n$) distribution. Therefore, $\overline{X} \sim$ Gamma($n, \beta/n$).

□