MH2500 AY18/19

Solution 1.

(a) $P(X^2 + Y^2 \leq 3) = \dfrac{1}{2}$.

(b) $P(Y \leq X^2 | Y \geq X^2) = \dfrac{P(Y = X^2)}{P(Y \geq X^2)} = \dfrac{2}{3}$.

(c) $P(|X| = 0) = \dfrac{1}{4}, \quad P(|X| = 1) = \dfrac{1}{2}, \quad P(|X| = 2) = \dfrac{1}{4}$.

(d) Note that $Var(X^7|Y = k) \geq 0$. When $Y = \pm 2$, X only takes one possible value, thus X^7 also takes one possible value given $Y = \pm 2$. Therefore, $Var(X^7|Y = k) = 0$, when $k = \pm 2$. This minimizes the variance.

(e) No, because $P(X = 0, Y = 0) \neq P(X = 0)P(Y = 0)$. Here, $P(X = 0, Y = 0) = 0$ and $P(X = 0) = P(Y = 0) = \frac{1}{4}$.

Solution 2.

(a) The joint density function of X, Y is $f(x, y) = 1$, if $x, y \in (0, 1]$. When $0 < z \leq 1$,

$$F_Z(z) = P\left(\frac{X}{Y} \leq z\right) = P(\left(Y \geq \frac{X}{z}\right) = \int_0^z \int_{\frac{x}{z}}^1 1 dy dx = \frac{z}{2} \implies f_Z(z) = \frac{1}{2}.$$

When $z > 1$,

$$F_Z(z) = P\left(\frac{X}{Y} \leq z\right) = 1 - P(\left(Y \leq \frac{X}{z}\right) = 1 - \int_0^1 \int_0^{\frac{z}{x}} 1 dy dx = 1 - \frac{1}{2z} \implies f_Z(z) = \frac{1}{2z^2}.$$

To get the double integrals, you should sketch the domain that is being integrated.
Let m be the median of Z, then $P(Z \leq m) = F_Z(m) = \frac{1}{2}$. We have $m = 1$.

(b)

$$\mathbb{E}(\sqrt{Z}) = \int_0^1 \sqrt{z} \cdot \frac{1}{2} dz + \int_1^\infty \sqrt{z} \cdot \frac{1}{2z^2} dz = \frac{1}{3} + 1 = \frac{4}{3}.$$

$$E(Z) = \int_0^1 \frac{z}{2} dz + \int_1^\infty \frac{1}{2z} dz.$$

The improper integral in $\mathbb{E}(Z)$ diverges. Hence, $Var(Z) = \mathbb{E}(Z) - \mathbb{E}(\sqrt{Z})^2$ does not exist.

Solution 3.

(a) $\mathbb{E}(S_n) = \sum_{i=1}^n \mathbb{E}(X_i) = n\mathbb{E}(X_1) = 0, \qquad Var(S_n) = \sum_{i=1}^n Var(X_i) = nVar(X_1) = 2n.$
By Chebyshev's inequality,

$$P(S_n \geq \sqrt{n \ln n}) \leq P(|S_n| \geq \sqrt{n \ln n}) \leq \frac{Var(S_n)}{n \ln n} = \frac{2}{\ln n} \to 0, \qquad \text{as } n \to \infty.$$

(b) By central limit theorem, $\frac{S_n}{n}$ converges in distribution to $N(0, \frac{2}{n})$. Therefore,

$$\lim_{n\to\infty} P(S_n \geq 0) = P(Z \geq 0) = \frac{1}{2}, \qquad Z \sim N(0,1).$$

(c) For $t \geq 0$,

$$P(S_n \leq \sqrt{n \ln n}) = P(tS_n \leq t\sqrt{n \ln n}) = P(e^{tS_n} \leq e^{t\sqrt{n \ln n}}) \leq \frac{\mathbb{E}(e^{tS_n})}{e^{t\sqrt{n \ln n}}}$$

Evaluating $\mathbb{E}(e^{tS_n})$, we have

$$P(S_n \leq \sqrt{n \ln n}) \leq e^{-t\sqrt{n \ln n}} \mathbb{E}(e^{t(X_1 + \ldots X_n)}) = e^{-t\sqrt{n \ln n}} \mathbb{E}(e^{t(X_1)})^n = e^{-t\sqrt{n \ln n}} \left(\frac{1}{3}e^{-2t} + \frac{2}{3}e^t\right)^n$$

Using Taylor series,

$$\frac{1}{3}e^{-2t} + \frac{2}{3}e^t \leq \frac{1}{3}(1 - 2t + 2t^2) + \frac{2}{3}(1 + t + t^2) \leq 1 + 2t^2$$

Choose $t = \frac{3}{2}\frac{\ln \ln n}{\sqrt{n \ln n}}$, then we have

$$P(S_n \leq \sqrt{n \ln n}) \leq e^{-\frac{3}{2}\ln \ln n} \left(1 + \frac{9}{2}\frac{(\ln \ln n)^2}{n \ln n}\right)^n \leq \frac{1}{(\ln n)^{3/2}} \left(1 + \frac{4.5}{n}\right)^n \leq \frac{e^{4.5}}{(\ln n)^{3/2}}.$$

We used the fact that for $n \geq 2$, $\frac{(\ln \ln n)^2}{\ln n} \leq 1$ and $\left(1 + \frac{x}{n}\right)^n \leq e^x$.

Solution 4. Here $X_{(1)} = \min\{X_1, \ldots, X_n\}$ and $X_{(n)} = \max\{X_1, \ldots, X_n\}$. Let $F(\cdot)$ and $f(\cdot)$ be the CDF and PDF of X_i respectively, for all i. For $i = 1, 2, \ldots, n$, we have $f(x) = 1$, for $x \in (0, 1)$.

(a)

$$F_{X_{(1)}}(x) = P(X_{(1)} \le x) = 1 - P(X_{(1)} \ge x) = 1 - P\left(\bigcap_{i=1}^{n}\{X_i \ge x\}\right) = 1 - P(X_1 \ge x)^n = 1 - (1 - F(x))^n.$$

$$f_{X_{(1)}}(x) = \frac{d}{dx}F_{X_{(1)}}(x) = nf(x)(1 - F(x))^{n-1}.$$

$$\mathbb{E}(X_{(1)}) = \int_0^1 x \cdot nf(x)(1 - F(x))^{n-1}dx = n\int_0^1 x(1 - x)^{n-1}dx = n\int_0^1 x^{n-1}(1 - x)dx = \frac{1}{n+1}.$$

$$F_{X_{(n)}}(x) = P(X_{(n)} \le x) = P\left(\bigcap_{i=1}^{n}\{X_i \le x\}\right) = P(X_1 \le x)^n = F(x)^n.$$

$$f_{X_{(n)}}(x) = \frac{d}{dx}F_{X_{(n)}}(x) = nf(x)(F(x))^{n-1}.$$

$$\mathbb{E}(X_{(n)}) = \int_0^1 x \cdot nf(x)(F(x))^{n-1}dx = n\int_0^1 x^n dx = \frac{n}{n+1}.$$

(b) The joint PDF of $X_{(1)}, X_{(n)}$ is

$$f_{X_{(1)},X_{(n)}}(u, v) = \frac{n!}{(n-2)!}(F(v) - F(u))^{n-2}f(u)f(v) = n(n-1)(v - u)^{n-2}, \quad \text{for } 0 \le u \le v \le 1.$$

The idea is that besides the minimum sample $X_{(1)}$ and the maximum sample $X_{(n)}$, the other $n-2$ samples must be between $X_{(1)}$ and $X_{(n)}$ and the total number of ways to arrange these n samples is $\frac{n!}{(n-2)!}$, the $n - 2$ samples are considered indistinguishable as their positions do not matter.

$$\mathbb{E}(X_{(1)}X_{(n)}) = \int_0^1\int_0^v uv \cdot n(n-1)(u - v)^{n-2}dudv = \frac{1}{n+2}.$$

$$Cov(X_{(1)}, X_{(n)}) = E(X_{(1)}X_{(n)}) - E(X_{(1)})E(X_{(n)}) = \frac{1}{n+2} - \frac{n}{(n+1)^2}.$$