Probability Theory 2 – Solutions V

1. Let $X \sim U([1,2])$ be a random variable with the continuous uniform distribution. Calculate $\mathbb{E}(X^n)$ for every positive integer n.

   Solution: Since $X \sim U([1,2])$, its probability density function is

   $$f(x) = \begin{cases} 1 & \text{if } 1 \leq x \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

   Hence

   $$\mathbb{E}(X^n) = \int_{-\infty}^{\infty} x^n f(x) dx = \int_1^2 x^n dx = \frac{x^{n+1}}{n+1} \mid_1^2 = \frac{2^{n+1}-1}{n+1}.$$

2. Let $X_1, \ldots, X_n$ be independent exponentially distributed random variables with parameters $\lambda_1, \ldots, \lambda_n$, respectively. Let $X = \min\{X_i : 1 \leq i \leq n\}$ and let $\lambda = \sum_{i=1}^n \lambda_i$. Prove that X is exponentially distributed with parameter λ.

   Solution: In order to prove that X is exponentially distributed with parameter λ, we will prove that $Pr(X > x) = e^{-\lambda x}$ for all $x \geq 0$ and that $Pr(X > x) = 1$ for all $x < 0$. Assume first that $x \geq 0$.

   $$\begin{aligned} Pr(X > x) &= Pr(\min\{X_i : 1 \leq i \leq n\} > x) = Pr(X_1 > x, X_2 > x, \ldots, X_n > x) \\ &= \prod_{i=1}^n Pr(X_i > x) = \prod_{i=1}^n e^{-\lambda_i x} = e^{-x \sum_{i=1}^n \lambda_i} = e^{-\lambda x}, \end{aligned}$$

   where the second equality holds since the minimum of $\{X_i : 1 \leq i \leq n\}$ is larger than x if and only if $X_i > x$ for every $1 \leq i \leq n$, the third equality holds since $X_1, \ldots, X_n$ are independent, and the fourth equality holds since $x \geq 0$ and X_i is exponentially distributed with parameter λ_i for every $1 \leq i \leq n$.

   Similarly, if $x < 0$, then $Pr(X > x) = \prod_{i=1}^n 1 = 1$.

3. A fair 6-sided die is rolled 420 times, all dice rolls being mutually independent. Let S denote the sum of the resulting numbers.

   (a) Use Chebyshev's inequality to find a lower bound on $Pr(1400 \leq S \leq 1540)$.

   (b) Use the central limit theorem (CLT) to estimate $Pr(1400 \leq S \leq 1540)$.

   Solution: For every $1 \leq i \leq 420$, let X_i be the random variable whose value is the result of the ith die roll. Then $S = \sum_{i=1}^{420} X_i$. Fix some $1 \leq i \leq 420$. Then

   $$\mathbb{E}(X_i) = \frac{1}{6} + \frac{2}{6} + \frac{3}{6} + \frac{4}{6} + \frac{5}{6} + \frac{6}{6} = \frac{7}{2}.$$

Similarly,

$$\mathbb{E}(X_i^2) = \frac{1}{6} + \frac{4}{6} + \frac{9}{6} + \frac{16}{6} + \frac{25}{6} + \frac{36}{6} = \frac{91}{6}$$

and thus

$$Var(X_i) = \mathbb{E}(X_i^2) - (\mathbb{E}(X_i))^2 = \frac{91}{6} - \frac{49}{4} = \frac{35}{12}.$$

Therefore

$$\mathbb{E}(S) = \mathbb{E}\left(\sum_{i=1}^{420} X_i\right) = \sum_{i=1}^{420} \mathbb{E}(X_i) = 420 \cdot \frac{7}{2} = 1470$$

and, since the X_i's are mutually independent,

$$Var(S) = Var\left(\sum_{i=1}^{420} X_i\right) = \sum_{i=1}^{420} Var(X_i) = 420 \cdot \frac{35}{12} = 1225.$$

(a) By the above calculations and Chebyshev's inequality we have

$$\begin{aligned} Pr(1400 \leq S \leq 1540) &= 1 - Pr(|S - 1470| > 70) = 1 - Pr(|S - \mathbb{E}(S)| \geq 71) \\ &\geq 1 - \frac{Var(S)}{71^2} = 1 - \frac{1225}{5041} \approx 0.75699. \end{aligned}$$

(b) Applying the CLT we have

$$\begin{aligned} Pr(1400 \leq S \leq 1540) &= Pr\left(\frac{1400 - 420 \cdot 7/2}{\sqrt{35/12} \cdot \sqrt{420}} \leq \frac{\sum_{i=1}^{420} X_i - 420 \cdot 7/2}{\sqrt{35/12} \cdot \sqrt{420}} \leq \frac{1540 - 420 \cdot 7/2}{\sqrt{35/12} \cdot \sqrt{420}}\right) \\ &= Pr\left(-\frac{70}{35} \leq \frac{\sum_{i=1}^{420} X_i - 1470}{35} \leq \frac{70}{35}\right) \\ &= Pr\left(-2 \leq \frac{\sum_{i=1}^{420} X_i - 1470}{35} \leq 2\right) \\ &\approx \Phi(2) - \Phi(-2) = 2\Phi(2) - 1 \approx 0.9544. \end{aligned}$$

4. Let $X_1, \ldots, X_{1200}$ be mutually independent random variables such that, for every $1 \leq i \leq 1200$, X_i is uniformly distributed over the segment $[0, 1)$. For every $1 \leq i \leq 1200$, Let Y_i be a rounding of X_i to the nearest integer, i.e.

$$Y_i = \begin{cases} 0 & \text{if } X_i < 1/2 \\ 1 & \text{if } X_i \geq 1/2 \end{cases}$$

Use the Central Limit Theorem to estimate the following probabilities:

(a) $Pr\left(\left|\sum_{i=1}^{1200} X_i - \sum_{i=1}^{1200} Y_i\right| \leq 10\right)$.

(b) $Pr\left(\sum_{i=1}^{1200} |X_i - Y_i| > 310\right)$.

Solution:

(a) For every $1 \leq i \leq 1200$, let $Z_i = X_i - Y_i$. The following claim should be intuitively clear; nevertheless we include a formal proof.

Claim 1. *For every $1 \leq i \leq 1200$, Z_i is uniformly distributed over the segment $[-1/2, 1/2)$.*

Proof. Let F_i denote the cumulative probability function of Z_i; we wish to prove that

$$F_i(x) = \begin{cases} 0 & \text{if } \ x < -1/2 \\ x + 1/2 & \text{if } \ -1/2 \leq x < 1/2 \\ 1 & \text{if } \ x \geq 1/2 \end{cases}$$

It readily follows from the definitions of X_i and Y_i that $Pr(-1/2 \leq Z_i < 1/2) = 1$. Hence, $F_i(x) = Pr(Z_i \leq x) = 0$ holds for every $x < -1/2$ and $F_i(x) = Pr(Z_i \leq x) = 1$ holds for every $x \geq 1/2$ as required.

Next, fix some $-1/2 \leq x < 0$. Observe that $Z_i \leq x$ if and only if $1/2 \leq X_i \leq 1 + x$. Hence, in this case we have

$$F_i(x) = Pr(Z_i \leq x) = Pr(1/2 \leq X_i \leq 1 + x) = 1/2 + x,$$

where the last equality holds since $X_i \sim U([0,1))$ and $-1/2 \leq x < 0$.

Finally, fix some $0 \leq x < 1/2$. Observe that $Z_i \leq x$ if and only if $0 \leq X_i \leq x$ or $1/2 \leq X_i < 1$. Hence, in this case we have

$$F_i(x) = Pr(Z_i \leq x) = Pr(1/2 \leq X_i < 1) + Pr(0 \leq X_i \leq x) = 1/2 + x,$$

where the second equality holds since $x < 1/2$ and the last equality holds since $X_i \sim U([0,1))$ and $x \geq 0$. $\square$

By Claim 1 we have $\mathbb{E}(Z_i) = 0$ and $Var(Z_i) = 1/12$. Moreover, $Z_1, \ldots, Z_{1200}$ are mutually independent. Applying the CLT then yields

$$\begin{aligned}
& Pr\left(\left|\sum_{i=1}^{1200} X_i - \sum_{i=1}^{1200} Y_i\right| \leq 10\right) = Pr\left(-10 \leq \sum_{i=1}^{1200} X_i - \sum_{i=1}^{1200} Y_i \leq 10\right) \\
= \ & Pr\left(-10 \leq \sum_{i=1}^{1200} (X_i - Y_i) \leq 10\right) = Pr\left(-10 \leq \sum_{i=1}^{1200} Z_i \leq 10\right) \\
= \ & Pr\left(\frac{-10 - 1200 \cdot 0}{\sqrt{1200} \cdot \sqrt{1/12}} \leq \frac{\sum_{i=1}^{1200} Z_i - 1200 \cdot 0}{\sqrt{1200} \cdot \sqrt{1/12}} \leq \frac{10 - 1200 \cdot 0}{\sqrt{1200} \cdot \sqrt{1/12}}\right) \\
= \ & Pr\left(-1 \leq \frac{\sum_{i=1}^{1200} Z_i}{10} \leq 1\right) \approx \Phi(1) - \Phi(-1) = 2\Phi(1) - 1 \approx 0.6826.
\end{aligned}$$

(b) For every $1 \leq i \leq 1200$, let $W_i = |X_i - Y_i|$. Similarly to Part (a) of this question, it readily follows from the definitions of X_i and Y_i that W_i is uniformly distributed over the segment $[0, 1/2]$ (this is not entirely accurate as for $w \in \{0, 1/2\}$ only one value of X_i leads to $W_i = w$, whereas, for any $w \in (0, 1/2)$, there are two values of X_i which lead to $W_i = w$. However, since $Pr(W_i \in \{0, 1/2\}) = 0$, this inaccuracy is negligible). Hence, $\mathbb{E}(W_i) = 1/4$ and $Var(W_i) = 1/48$. Moreover, $W_1, \ldots, W_{1200}$ are mutually independent. Applying the CLT then yields

$$
\begin{aligned}
& Pr\left(\sum_{i=1}^{1200} |X_i - Y_i| > 310\right) = Pr\left(\sum_{i=1}^{1200} W_i > 310\right) = 1 - Pr\left(\sum_{i=1}^{1200} W_i \leq 310\right) \\
= \quad & 1 - Pr\left(\frac{\sum_{i=1}^{1200} W_i - 1200 \cdot 1/4}{\sqrt{1200} \cdot \sqrt{1/48}} \leq \frac{310 - 1200 \cdot 1/4}{\sqrt{1200} \cdot \sqrt{1/48}}\right) \\
= \quad & 1 - Pr\left(\frac{\sum_{i=1}^{1200} W_i - 300}{5} \leq 2\right) \approx 1 - \Phi(2) \approx 0.0228.
\end{aligned}
$$