Solutions to Review Problems for Exam 3

1. Let X denote a positive random variable such that $\ln(X)$ has a Normal$(0,1)$ distribution.

   (a) Give the pdf of X and compute its expectation.
   Solution: Set $Z = \ln(X)$, so that $Z \sim$ Normal$(0,1)$; thus,

   $$f_Z(y) = \frac{1}{\sqrt{2\pi}}\, e^{-z^2/2}, \qquad \text{for } z \in \mathbb{R}. \tag{1}$$

   Next, compute the cdf for X,

   $$F_X(x) = \Pr(X \leqslant x), \qquad \text{for } x > 0,$$

   to get

   $$\begin{aligned} F_X(x) &= \Pr[\ln(X) \leqslant \ln(x)] \\ &= \Pr[Z \leqslant \ln(x)] \\ &= F_Z(\ln(x)), \end{aligned}$$

   so that

   $$F_X(x) = \begin{cases} F_Z(\ln(x)), & \text{for } x > 0; \\ 0 & \text{for } x \leqslant 0. \end{cases} \tag{2}$$

   Differentiating (2) with respect to x, for $x > 0$, we obtain

   $$f_X(x) = F_Z'(\ln(x)) \cdot \frac{1}{x},$$

   so that

   $$f_X(x) = f_Z(\ln(x)) \cdot \frac{1}{x}, \tag{3}$$

   where we have used the Chain Rule. Combining (1) and (3) yields

   $$f_X(x) = \begin{cases} \dfrac{1}{\sqrt{2\pi}\; x}\, e^{-(\ln x)^2/2}, & \text{for } x > 0; \\ 0 & \text{for } x \leqslant 0. \end{cases} \tag{4}$$

   In order to compute the expected value of X, use the pdf in (4) to get

   $$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} x f_X(x)\; dx \\ &= \int_0^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{-(\ln x)^2/2}\; dx. \end{aligned} \tag{5}$$

Make the change of variables $u = \ln x$ in the last integral in (5) to get $du = \dfrac{1}{x}dx$, so that $dx = e^u du$ and

$$E(X) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{u-u^2/2}\, du. \tag{6}$$

Complete the square in the exponent of the integrand in (6) to obtain

$$E(X) = e^{1/2}\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{-(u-1)^2/2}\, du. \tag{7}$$

Next, make the change of variables $w = u - 1$ for the integral in (7) to get

$$E(X) = e^{1/2}\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{-w^2/2}\, dw = \sqrt{e}.$$

□

(b) Estimate $\Pr(X \leq 6.5)$.

Solution: Use the result in (2) to compute

$$\Pr(X \leq 6.5) = F_Z(\ln(6.5)), \quad \text{where } Z \sim \text{Normal}(0, 1).$$

Thus,

$$\Pr(X \leq 6.5) \doteq F_Z(1.8718) \doteq 0.969383,$$

or about 97%. □

2. Forty seven digits are chosen at random and with replacement from $\{0, 1, 2, \ldots, 9\}$. Estimate the probability that their average lies between 4 and 6.

Solution: Let $X_1, X_2, \ldots, X_n$, where $n = 47$, denote the 47 digits. Since the sampling is done without replacement, the random variables $X_1, X_2, \ldots, X_n$ are identically uniformly distributed over the digits $\{0, 1, 2, \ldots, 9\}$ with pmf given by

$$p_X(k) = \begin{cases} \dfrac{1}{10}, & \text{for } k = 0, 1, 2, \ldots, 9; \\ \\ 0, & \text{elsewhere.} \end{cases} \tag{8}$$

Consequently, the mean of the distribution is

$$\mu = \sum_{k=0}^{9} k p_X(k) = \frac{1}{10}\sum_{k=1}^{9} k = \frac{1}{10}\cdot\frac{9\cdot 10}{2} = \frac{9}{2}. \tag{9}$$

Before we compute the variance, we first compute the second moment of X:

$$E(X^2) = \sum_{k=0}^{9} k^2 p_{_X}(k) = \sum_{k=1}^{9} k^2 p_{_X}(k);$$

thus, using the pmf of X in (8),

$$\begin{aligned} E(X^2) &= \frac{1}{10}\sum_{k=1}^{9} k^2 \\ &= \frac{1}{10}\cdot\frac{9\cdot(9+1)(2\cdot 9+1)}{6} \\ &= \frac{3\cdot(19)}{2} \\ &= \frac{57}{2}. \end{aligned}$$

Thus, the variance of X is

$$\begin{aligned} \sigma^2 &= E(X^2) - \mu^2 \\ &= \frac{57}{2} - \frac{81}{4} \\ &= \frac{33}{4}; \end{aligned}$$

so that

$$\sigma^2 = 8.25. \tag{10}$$

We would like to estimate

$$\Pr(4 \leqslant \overline{X}_n \leqslant 6),$$

or

$$\Pr(4-\mu \leqslant \overline{X}_n - \mu \leqslant 6-\mu),$$

where μ is given in (9), so that

$$\Pr(4 \leqslant \overline{X}_n \leqslant 6) = \Pr(-0.5 \leqslant \overline{X}_n - \mu \leqslant 1.5) \tag{11}$$

Next, divide the last inequality in (11) by $\sigma/\sqrt{n}$, where σ is as given in (10), to get

$$\Pr(4 \leqslant \overline{X}_n \leqslant 6) \doteq \Pr\left(-1.19 \leqslant \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \leqslant 3.58\right) \tag{12}$$

Since $n = 47$ can be considered a large sample size, we can apply the Central Limit Theorem to obtain from (12) that

$$\Pr(4 \leqslant \overline{X}_n \leqslant 6) \approx \Pr\left(-1.19 \leqslant Z \leqslant 3.58\right), \quad \text{where } Z \sim \text{Normal}(0,1). \tag{13}$$

It follows from (13) and the definition of the cdf that

$$\Pr(4 \leqslant \overline{X}_n \leqslant 6) \approx F_{_Z}(3.58) - F_{_Z}(-1.19), \tag{14}$$

where $F_{_Z}$ is the cdf of $Z \sim \text{Normal}(0,1)$. Using the symmetry of the pdf of $Z \sim \text{Normal}(0,1)$, we can rewrite (14) as

$$\Pr(4 \leqslant \overline{X}_n \leqslant 6) \approx F_{_Z}(1.19) + F_{_Z}(3.58) - 1. \tag{15}$$

Finally, using a table of standard normal probabilities, we obtain from (15) that

$$\Pr(4 \leqslant \overline{X}_n \leqslant 6) \approx 0.8830 + 0.9998 - 1 = 0.8828.$$

Thus, the probability that the average of the 47 digits is between 4 and 6 is about 88.3%. □

3. Let $X_1, X_2, \ldots, X_{30}$ be independent random variables each having a discrete distribution with pmf:

$$p(x) = \begin{cases} 1/4, & \text{if } x = 0 \text{ or } x = 2; \\ 1/2, & \text{if } x = 1; \\ 0, & \text{otherwise.} \end{cases}$$

Estimate the probability that $X_1 + X_2 + \cdots + X_{30}$ is at most 33.

***Solution**:* First, compute the mean, $\mu = E(X)$, and variance, $\sigma^2 = \text{Var}(X)$, of the distribution:

$$\mu = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 2\frac{1}{4} = 1. \tag{16}$$

$$\sigma^2 = E(X^2) - [E(X)]^2, \tag{17}$$

where

$$E(X^2) = 0^2 \cdot \frac{1}{4} + 1^2 \cdot \frac{1}{2} + 2^2\frac{1}{4} = 1.5; \tag{18}$$

so that, combining (16), (17) and (18),

$$\sigma^2 = 1.5 - 1 = 0.5. \tag{19}$$

Next, let $Y = \sum_{k=1}^{n} X_k$, where $n = 30$. We would like to estimate

$$\Pr[Y \leqslant 33],$$

using the continuity correction,

$$\Pr[Y \leqslant 33.5], \tag{20}$$

By the Central Limit Theorem

$$\Pr\left(\frac{Y - n\mu}{\sqrt{n}\,\sigma} \leqslant \right) \approx \Pr(Z \leqslant z), \quad \text{for } z \in \mathbb{R}, \tag{21}$$

where $Z \sim \text{Normal}(0, 1)$, $\mu = 1$, $\sigma^2 = 1.5$ and $n = 30$. It follows from (21) that we can estimate the probability in (20) by

$$\Pr[Y \leqslant 33.5] \approx \Pr(Z \leqslant 0.52) \doteq 0.6985. \tag{22}$$

Thus, according to (22), the probability that $X_1 + X_2 + \cdots + X_{30}$ is at most 33 is about 70%. □

4. Roll a balanced die 36 times. Let Y denote the sum of the outcomes in each of the 36 rolls. Estimate the probability that $108 \leq Y \leq 144$.

   Suggestion: Since the event of interest is ($Y \in \{108, 109, \ldots, 144\}$), rewrite $\Pr(108 \leq Y \leq 144)$ as

   $$\Pr(107.5 < Y \leqslant 144.5).$$

   ***Solution**:* Let $X_1, X_2, \ldots, X_n$, where $n = 36$, denote the outcomes of the 36 rolls. Since we are assuming that the die is balanced, the random variables $X_1, X_2, \ldots, X_n$ are identically uniformly distributed over the digits $\{1, 2, \ldots, 6\}$; in other words, $X_1, X_2, \ldots, X_n$ is a random sample from the discrete Uniform(6) distribution. Consequently, the mean of the distribution is

   $$\mu = \frac{6+1}{2} = 3.5, \tag{23}$$

   and the variance is

   $$\sigma^2 = \frac{(6+1)(6-1)}{12} = \frac{35}{12}. \tag{24}$$

   We also have that

   $$Y = \sum_{k=1}^{n} X_k,$$

where $n = 36$.

By the Central Limit Theorem,

$$\Pr(107.5 < Y \leqslant 144.5) \approx \Pr\left(\frac{107.5 - n\mu}{\sqrt{n}\sigma} < Z \leqslant \frac{144.5 - n\mu}{\sqrt{n}\sigma}\right), \qquad (25)$$

where $Z \sim \text{Normal}(0, 1)$, $n = 36$, and μ and σ are given in (23) and (24), respectively. We then have from (25) that

$$\begin{aligned}
\Pr(107.5 < Y \leqslant 144.5) &\approx \Pr(-1.81 < Z \leqslant 1.81) \\
&\approx F_Z(1.81) - F_Z(-1.81) \\
&\approx 2F_Z(1.81) - 1 \\
&\approx 2(0.9649) - 1 \\
&\approx 0.9298;
\end{aligned}$$

so that the probability that $108 \leqslant Y \leqslant 144$ is about 93%. □

5. Let $Y \sim \text{Binomial}(100, 1/2)$. Use the Central Limit Theorem to estimate the value of $\Pr(Y = 50)$.

***Solution**:* We use the continuity correction and estimate

$$\Pr(49.5 < Y \leqslant 50.5).$$

By the Central Limit Theorem,

$$\Pr(49.5 < Y \leqslant 50.5) \approx \Pr\left(\frac{49.5 - n\mu}{\sqrt{n}\sigma} < Z \leqslant \frac{50.5 - n\mu}{\sqrt{n}\sigma}\right), \qquad (26)$$

where $Z \sim \text{Normal}(0, 1)$, $n = 100$, and $n\mu = 50$ and

$$\sigma = \sqrt{\frac{1}{2}\left(1 - \frac{1}{2}\right)} = \frac{1}{2}.$$

We then obtain from (26) that

$$\begin{aligned}
\Pr(49.5 < Y \leqslant 50.5) &\approx \Pr\left(-0.1 < Z \leqslant 0.1\right) \\
&\approx F_{Z}(0.1) - F_{Z}(-0.1) \\
&\approx 2F_{Z}(0.1) - 1 \\
&\approx 2(0.5398) - 1 \\
&\approx 0.0796.
\end{aligned}$$

Thus,

$$\Pr(Y = 50) \approx 0.08,$$

or about 8%. □

6. Let $Y \sim \text{Binomial}(n, 0.55)$. Find the smallest value of n such that, approximately,

$$\Pr(Y/n > 1/2) \geqslant 0.95. \tag{27}$$

***Solution**:* By the Central Limit Theorem,

$$\frac{\dfrac{Y}{n} - 0.55}{\sqrt{(0.55)(1-0.55)}/\sqrt{n}} \xrightarrow{D} Z \sim \text{Normal}(0,1) \text{ as } n \to \infty. \tag{28}$$

Thus, according to (27) and (28), we need to find the smallest value of n such that

$$\Pr\left(Z > \frac{0.5 - 0.55}{(0.4975)/\sqrt{n}}\right) \geqslant 0.95,$$

or

$$\Pr\left(Z > -\frac{\sqrt{n}}{10}\right) \geqslant 0.95. \tag{29}$$

The expression in (29) is equivalent to

$$1 - \Pr\left(Z \leqslant -\frac{\sqrt{n}}{10}\right) \geqslant 0.95,$$

which can be re-written as

$$1 - F_{Z}\left(-\frac{\sqrt{n}}{10}\right) \geqslant 0.95, \tag{30}$$

where F_{z} is the cdf of $Z \sim \text{Normal}(0,1)$.

By the symmetry of the pdf for $Z \sim \text{Normal}(0,1)$, (30) is equivalent to

$$F_{z}\left(\frac{\sqrt{n}}{10}\right) \geqslant 0.95. \tag{31}$$

The smallest value of n for which (31) holds true occurs when

$$\frac{\sqrt{n}}{10} \geqslant z^*, \tag{32}$$

where z^* is a positive real number with the property

$$F_{z}(z^*) = 0.95. \tag{33}$$

The equality in (33) occurs approximately when

$$z^* = 1.645. \tag{34}$$

It follows from (32) and (34) that (27) holds approximately when

$$\frac{\sqrt{n}}{10} \geqslant 1.645,$$

or $n \geqslant 270.6025$. Thus, $n = 271$ is the smallest value of n such that, approximately,

$$\Pr(Y/n > 1/2) \geqslant 0.95.$$

□

7. Let $X_1, X_2, \ldots, X_n$ be a random sample from a Poisson distribution with mean λ. Thus, $Y = \displaystyle\sum_{i=1}^{n} X_i$ has a Poisson distribution with mean $n\lambda$. Moreover, by the Central Limit Theorem, $\overline{X} = Y/n$ has, approximately, a $\text{Normal}(\lambda, \lambda/n)$ distribution for large n. Show that $u(Y/n) = \sqrt{Y/n}$ is a function of Y/n which is essentially free of λ.

   ***Solution**:* We will show that, for large values of n, the distribution of

$$2\sqrt{n}\left(\sqrt{\frac{Y}{n}} - \sqrt{\lambda}\right) \tag{35}$$

is independent of λ. In fact, we will show that, for large values of n, the distribution of the random variables in (35) can be approximated by a Normal$(0,1)$ distribution.

First, note that by the Weak Law of Large Numbers,

$$\frac{Y}{n} \xrightarrow{\text{Pr}} \lambda, \qquad \text{as } n \to \infty.$$

Thus, for large values of n, we can approximate $u(Y/n)$ by its linear approximation around λ

$$u(Y/n) \approx u(\lambda) + u'(\lambda)\left(\frac{Y}{n} - \lambda\right), \tag{36}$$

where

$$u'(\lambda) = \frac{1}{2\sqrt{\lambda}}, \tag{37}$$

since $u(\lambda) = \sqrt{\lambda}$, for $\lambda > 0$. Combining (36) and (37) we see that, for large values of n,

$$\sqrt{\frac{Y}{n}} - \sqrt{\lambda} \approx \frac{1}{2\sqrt{n}} \cdot \frac{\dfrac{Y}{n} - \lambda}{\sqrt{\dfrac{\lambda}{n}}}. \tag{38}$$

Now, by the Central Limit Theorem,

$$\frac{\dfrac{Y}{n} - \lambda}{\sqrt{\dfrac{\lambda}{n}}} \xrightarrow{D} Z \sim \text{Normal}(0,1) \text{ as } n \to \infty; \tag{39}$$

it then follows from (38) and (39) that

$$2\sqrt{n}\left(\sqrt{\frac{Y}{n}} - \sqrt{\lambda}\right) \xrightarrow{D} Z \sim \text{Normal}(0,1) \text{ as } n \to \infty,$$

which was to be shown. □

8. Suppose that factory produces a number X of items in a week, where X can be modeled by a random variable with mean 50. Suppose also that the variance for a week's production is known to be 25. What can be said about the probability that this week's production will be between 40 and 60?

   ***Solution**:* Write

$$\Pr(40 < X < 60) = \Pr(-10 < X - \mu < 10),$$

where $\mu = 50$, so that

$$\Pr(40 < X < 60) = \Pr(|X - \mu| < 2\sigma),$$

where $\sigma = 5$. We can then write

$$\Pr(40 < X < 60) = 1 - \Pr(|X - \mu| \geqslant 2\sigma);$$

where, using Chebyshev's inequality,

$$\Pr(|X - \mu| \geqslant 2\sigma) \leqslant \frac{\sigma^2}{4\sigma^2} = \frac{1}{4}.$$

Consequently,

$$\Pr(40 < X < 60) \geqslant 1 - \frac{1}{4} = \frac{3}{4};$$

so that the he probability that this week's production will be between 40 and 60 is at least 75%. □

9. Let (X_n) denote a sequence of nonnegative random variables with means $\mu_n = E(X_n)$, for each $n = 1, 2, 3, \ldots$. Assume that $\lim\limits_{n\to\infty} \mu_n = 0$. Show that X_n converges in probability to 0 as $n \to \infty$.

   ***Solution**:* We need to show that, for any $\varepsilon > 0$,

$$\lim_{n\to\infty} \Pr(|X_n| < \varepsilon) = 1. \tag{40}$$

   We will establish (40) by first applying the Markov inequality: Given a random variable X and any $\varepsilon > 0$,

$$\Pr(|X| \geqslant \varepsilon) \leqslant \frac{E(|X|)}{\varepsilon}, \tag{41}$$

   to $X = X_n$, for each n.

   Observe that

$$E(|X_n|) = E(X_n) = \mu_n, \quad \text{for all } n,$$

   since we are assuming that X_n is nonnegative. It then follows from (41) that

$$\Pr(|X_n| \geqslant \varepsilon) \leqslant \frac{\mu_n}{\varepsilon}, \quad \text{for } n = 1, 2, 3, \ldots \tag{42}$$

   It follows from (42) that

$$\Pr(|X_n| < \varepsilon) \geqslant 1 - \frac{\mu_n}{\varepsilon}, \quad \text{for } n = 1, 2, 3, \ldots$$

so that

$$1 - \frac{\mu_n}{\varepsilon} \leqslant \Pr(|X_n| < \varepsilon) \leqslant 1, \qquad \text{for } n = 1, 2, 3, \ldots \tag{43}$$

Next, use the assumption that $\lim\limits_{n\to\infty} \mu_n = 0$ and the Squeeze Lemma to obtain from (43) that

$$\lim_{n\to\infty} \Pr(|X_n| < \varepsilon) = 1,$$

which is the statement in (40). Hence, X_n converges in probability to 0 as $n \to \infty$. □

10. Let $X_1, X_2, \ldots, X_n$ be a random sample from a Poisson distribution with mean λ. Thus, $Y_n = \sum\limits_{i=1}^{n} X_i$ has a Poisson distribution with mean $n\lambda$. Moreover, by the Central Limit Theorem, $\overline{X}_n = Y_n/n$ has, approximately, a Normal$(\lambda, \lambda/n)$ distribution for large n. Show that, for large values of n, the distribution of $2\sqrt{n}\left(\sqrt{\frac{Y_n}{n}} - \sqrt{\lambda}\right)$ is independent of λ.

    Solution: See solution to Problem 7. □