Solutions to Review Problems for Exam 3

1. A random point (X, Y) is distributed uniformly on the square with vertices $(-1,-1)$, $(1,-1)$, $(1,1)$ and $(-1,1)$.

   (a) Give the joint pdf for X and Y.

   (b) Compute the following probabilities:

   (i) $\Pr(X^2 + Y^2 < 1)$,

   (ii) $\Pr(2X - Y > 0)$,

   (iii) $\Pr(|X + Y| < 2)$.

   ***Solution**:* The square is pictured in Figure 1 and has area 4.

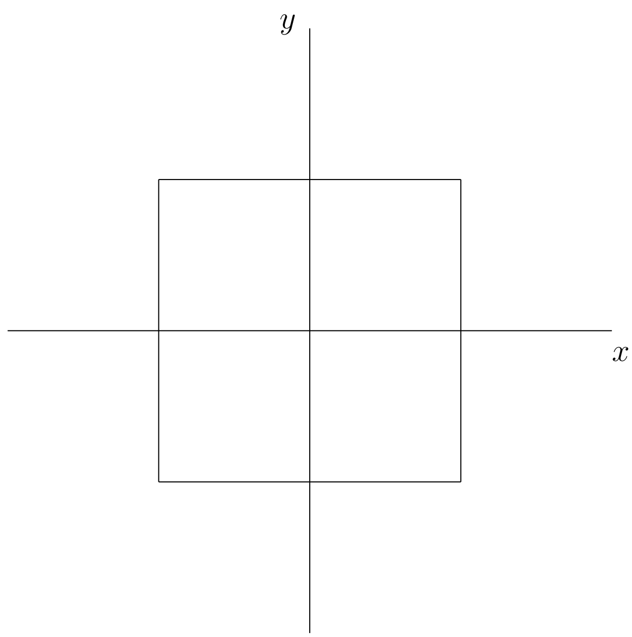

Figure 1: Sketch of square in Problem 1

   (a) Consequently, the joint pdf of (X, Y) is given by

$$f_{(X,Y)}(x,y) = \begin{cases} \dfrac{1}{4}, & \text{for } -1 < x < 1, -1 < y < 1; \\ \\ 0 & \text{elsewhere.} \end{cases} \tag{1}$$

(b) Denoting the square in Figure 1 by R, it follows from (1) that, for any subset A of $\mathbb{R}^2$,

$$\Pr[(x,y)\in A] = \iint_A f_{(X,Y)}(x,y)\ dxdy = \frac{1}{4}\cdot \text{area}(A\cap R); \qquad (2)$$

that is, $\Pr[(x,y)\in A]$ is one–fourth the area of the portion of A in R.

We will use the formula in (2) to compute each of the probabilities in (i), (ii) and (iii).

(i) In this case, A is the circle of radius 1 around the origin in $\mathbb{R}^2$ and pictured in Figure 2.

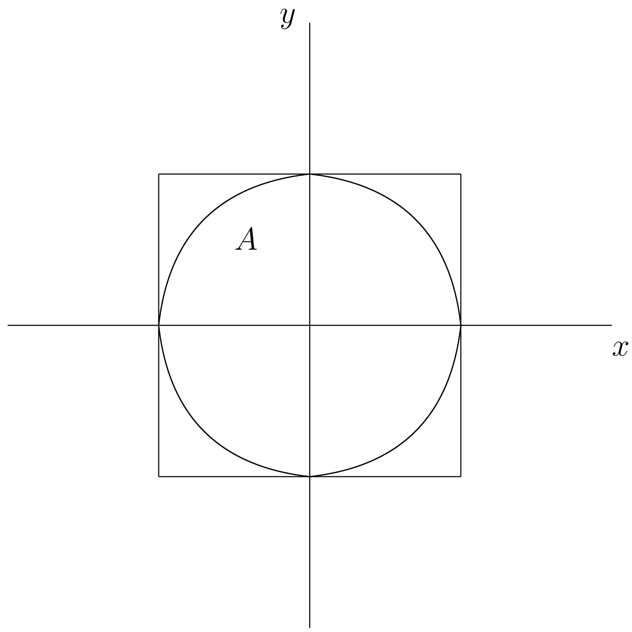

Figure 2: Sketch of A in Problem 1(b)(i)

Note that the circle A in Figure 2 is entirely contained in the square R so that, by the formula in (2),

$$\Pr(X^2+Y^2<1) = \frac{\text{area}(A)}{4} = \frac{\pi}{4}.$$

(ii) The set A in this case is pictured in Figure 3 on page 3. Thus, in this case, $A\cap R$ is a trapezoid of area $2\cdot\dfrac{\frac{1}{2}+\frac{3}{2}}{2} = 2$, so that, by the formula in (2),

$$\Pr(2X-Y>0) = \frac{1}{4}\cdot\text{area}(A\cap R) = \frac{1}{2}.$$

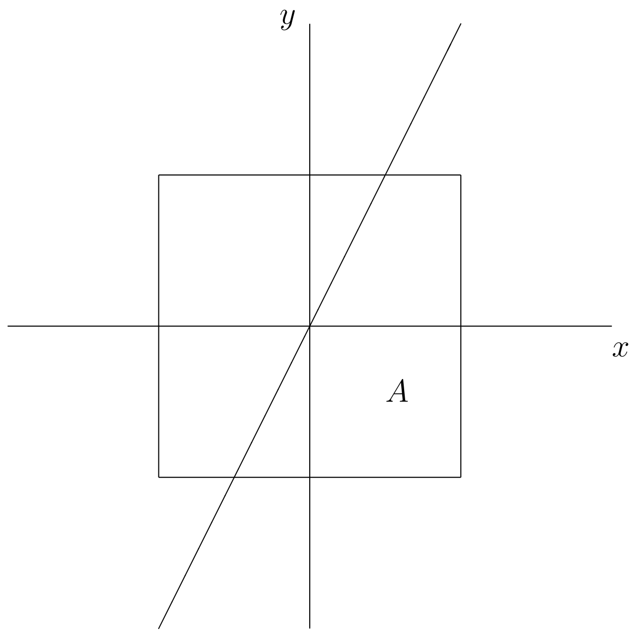

Figure 3: Sketch of A in Problem 1(b)(ii)

(iii) In this case, A is the region in the xy–plane between the lines $x+y=2$ and $x+y=-2$ (see Figure 4 on page 4). Thus, $A \cap R$ is R so that, by the formula in (2),

$$\Pr(|X+Y|<2) = \frac{\text{area}(R)}{4} = 1.$$

□

2. The random pair (X, Y) has the joint distribution shown in Table 1.

$X\backslash Y$	2	3	4
1	$\frac{1}{12}$	$\frac{1}{6}$	0
2	$\frac{1}{6}$	0	$\frac{1}{3}$
3	$\frac{1}{12}$	$\frac{1}{6}$	0

Table 1: Joint Probability Distribution for X and Y, $p_{(X,Y)}$

(a) Show that X and Y are not independent.

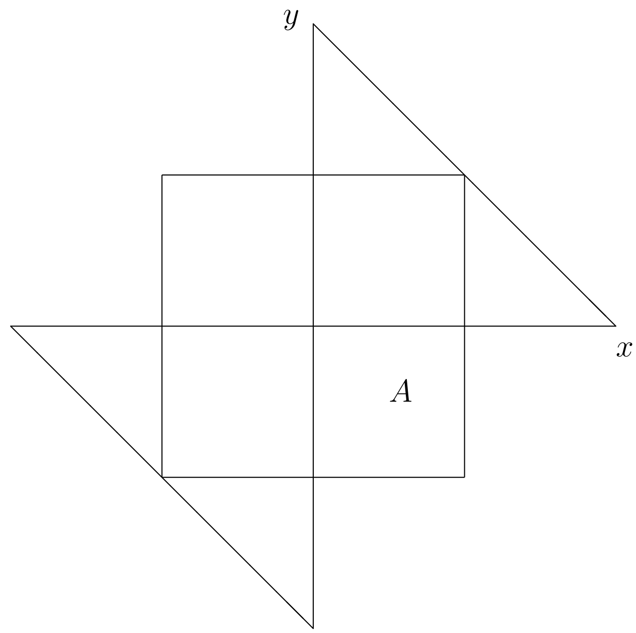

Figure 4: Sketch of A in Problem 1(b)(iii)

$X\backslash Y$	2	3	4	p_X
1	$\frac{1}{12}$	$\frac{1}{6}$	0	$\frac{1}{4}$
2	$\frac{1}{6}$	0	$\frac{1}{3}$	$\frac{1}{2}$
3	$\frac{1}{12}$	$\frac{1}{6}$	0	$\frac{1}{4}$
p_Y	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	1

Table 2: Joint pdf for X and Y and marginal distributions p_X and p_Y

***Solution**:* Table 2 shows the marginal distributions of X and Y on the margins.
Observe from Table 2 that

$$p_{(X,Y)}(1,4) = 0,$$

while

$$p_X(1) = \frac{1}{4} \quad \text{and} \quad p_Y(4) = \frac{1}{3}.$$

Thus,

$$p_X(1) \cdot p_Y(4) = \frac{1}{12};$$

so that

$$p_{(X,Y)}(1,4) \neq p_X(1) \cdot p_Y(4),$$

and, therefore, X and Y are not independent. □

(b) Give a probability table for random variables U and V that have the same marginal distributions as X and Y, respectively, but are independent.

Solution: Table 3 on page 5 shows the joint pmf of (U, V) and the marginal distributions, p_U and p_V. □

$U\backslash V$	2	3	4	p_U
1	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$
2	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{2}$
3	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$
p_V	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	1

Table 3: Joint pdf for U and V and their marginal distributions.

3. An experiment consists of independent tosses of a fair coin. Let X denote the number of trials needed to obtain the first head, and let Y be the number of trials needed to get two heads in repeated tosses. Are X and Y independent random variables?

Solution: X has a geometric distribution with parameter $p = \dfrac{1}{2}$, so that

$$p_X(k) = \frac{1}{2^k}, \quad \text{for } k = 1, 2, 3, \ldots \tag{3}$$

On the other hand,

$$\Pr[Y = 2] = \frac{1}{4}, \tag{4}$$

since, in two repeated tosses of a coin, the events are HH, HT, TH and TT, and these events are equally likely.

Next, consider the joint event $(X = 2, Y = 2)$. Note that

$$(X = 2, Y = 2) = [X = 2] \cap [Y = 2] = \emptyset,$$

since $[X = 2]$ corresponds to the event TH, while $[Y = 2]$ to the event HH. Thus,

$$\Pr(X = 2, Y = 2) = 0,$$

while

$$p_X(2) \cdot p_Y(2) = \frac{1}{4} \cdot \frac{1}{4} = \frac{1}{16},$$

by (3) and (4). Thus,

$$p_{(X,Y)}(2,2) \neq p_X(2) \cdot p_X(2).$$

Hence, X and Y are not independent. □

4. Let $g(t)$ denote a non–negative, integrable function of a single variable with the property that
$$\int_0^\infty g(t) \ \mathrm{d}t = 1.$$
Define
$$f(x,y) = \begin{cases} \dfrac{2g(\sqrt{x^2+y^2})}{\pi\sqrt{x^2+y^2}}, & \text{for } 0 < x < \infty,\ 0 < y < \infty, \\ \\ 0, & \text{otherwise.} \end{cases}$$
Show that $f(x,y)$ is a joint pdf for two random variables X and Y.

***Solution**:* First observe that f is non–negative since g is non–negative. Next, compute
$$\iint_{\mathbb{R}^2} f(x,y) \ \mathrm{d}x \ \mathrm{d}y \ = \ \int_0^\infty \int_0^\infty \frac{2g(\sqrt{x^2+y^2})}{\pi\sqrt{x^2+y^2}} \ \mathrm{d}x \ \mathrm{d}y.$$
Switching to polar coordinates we then get that
$$\begin{aligned} \iint_{\mathbb{R}^2} f(x,y) \ \mathrm{d}x \ \mathrm{d}y &= \int_0^{\pi/2} \int_0^\infty \frac{2g(r)}{\pi r} \ r \ \mathrm{d}r \ \mathrm{d}\theta \\ &= \frac{\pi}{2} \int_0^\infty \frac{2}{\pi} g(r) \ \mathrm{d}r \\ &= \int_0^\infty g(r) \ \mathrm{d}r \\ &= 1; \end{aligned}$$
therefore, $f(x,y)$ is indeed a joint pdf for two random variables X and Y. □

5. Suppose that two persons make an appointment to meet between 5 PM and 6 PM at a certain location and they agree that neither person will wait more than 10 minutes for each person. If they arrive independently at random times between 5 PM and 6 PM, what is the probability that they will meet?

Solution: Let X denote the arrival time of the first person and Y that of the second person. Then X and Y are independent and uniformly distributed on the interval $(5 \text{ PM}, 6 \text{ PM})$, in hours. It then follows that the joint pdf of X and Y is

$$f_{(X,Y)}(x,y) = \begin{cases} 1, & \text{if } 5 \text{ PM} < x < 6 \text{ PM}, 5 \text{ PM} < x < 6 \text{ PM}, \\ 0, & \text{elsewhere.} \end{cases}$$

Define $W = |X - Y|$; this is the time that one person would have to wait for the other one. Then, W takes on values, w, between 0 and 1 (in hours). The probability that that a person would have to wait more than 10 minutes is

$$\Pr(W > 1/6),$$

since the time is being measured in hours. It then follows that the probability that the two persons will meet is

$$1 - \Pr(W > 1/6) = \Pr(W \leqslant 1/6) = F_W(1/6).$$

We will therefore need to find the cdf of W. To do this, we compute

$$\begin{aligned} \Pr(W \leqslant w) &= \Pr(|X - Y| \leqslant w), \quad \text{for } 0 < w < 1, \\ &= \iint_A f_{(X,Y)}(x,y) \, \mathrm{d}x \, \mathrm{d}y, \end{aligned}$$

where A is the event

$$A = \{(x,y) \in \mathbb{R}^2 \mid 5 \text{ PM} < x < 6 \text{ PM}, 5 \text{ PM} < y < 6 \text{ PM}, |x - y| \leqslant w\}.$$

This event is pictured in Figure 5.

We then have that

$$\begin{aligned} \Pr(W \leqslant w) &= \iint_A \mathrm{d}x \, \mathrm{d}y \\ &= \text{area}(A), \end{aligned}$$

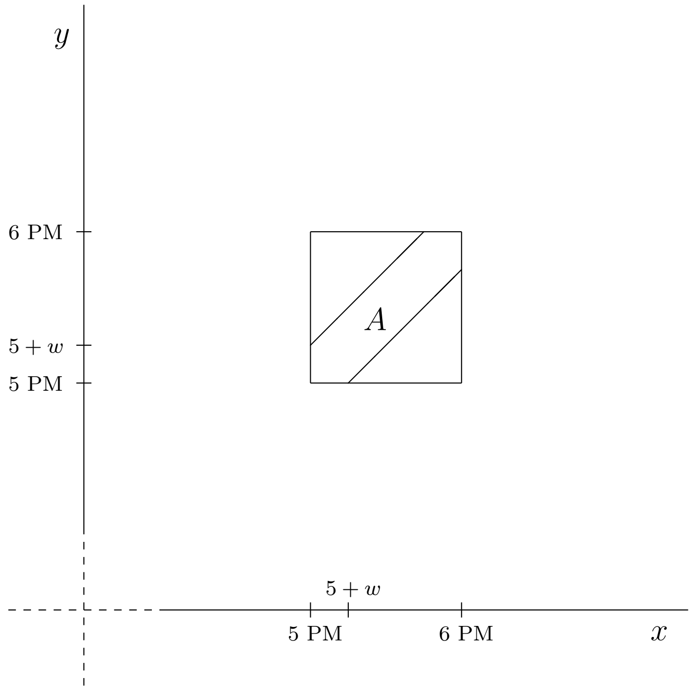

Figure 5: Event A in the xy–plane

where the area of A can be computed by subtracting from 1 the area of the two corner triangles shown in Figure 5:

$$\begin{aligned} \Pr(W \leqslant w) &= 1 - (1-w)^2 \\ &= 2w - w^2. \end{aligned}$$

Consequently, $F_W(w) = 2w - w^2$ for $0 < w < 1$. Thus the probability that the two persons will meet is

$$F_W(1/6) = 2 \cdot \frac{1}{6} - \left(\frac{1}{6}\right)^2 = \frac{11}{36},$$

or about 30.56%. □

6. Assume that the number of calls coming per minute into a hotel's reservation center follows a Poisson distribution with mean 3.

   (a) Find the probability that no calls come in a given 1 minute period.

Solution: Let Y denote the number of calls that come to the hotel's reservation center in one minute. Then, $Y \sim \text{Poisson}(3)$; so that,

$$p_{_Y}(k) = \frac{3^k}{k!}\, e^{-3}, \quad \text{for } k = 0, 1, 2, \ldots$$

Then, the probability that no calls will come in the given minute is

$$\Pr(Y = 0) = p_{_Y}(0) = e^{-3} \approx 0.05,$$

or about 5%. □

(b) Assume that the number of calls arriving in two different minutes are independent. Find the probability that at least two calls will arrive in a given two minute period.

Solution: Let Y_1 denote the number of calls that arrive in one minute and Y_2 denote the number of calls that arrive in another minute. We then have that

$$Y_i \sim \text{Poisson}(3), \quad \text{for } i = 1, 2,$$

and Y_i and Y_2 are independent. We want to compute

$$\Pr(Y_1 + Y_2 \geqslant 2).$$

To do this, we determine the distribution of $W = Y_1 + Y_2$.
Since Y_1 and Y_2 are independent,

$$\psi_{_W}(t) = \psi_{_{Y_1+Y_2}}(t) = \psi_{_{Y_1}}(t) \cdot \psi_{_{Y_2}}(t);$$

so that,

$$\psi_{_W}(t) = e^{3(e^t-1)} \cdot e^{3(e^t-1)} = e^{6(e^t-1)},$$

which is the mgf of a Poisson(6) distribution. Thus, by the mgf Uniqueness Theorem, $W \sim \text{Poisson}(6)$. We then have that

$$p_{_W}(k) = \frac{6^k}{k!}\, e^{-6}, \quad \text{for } k = 0, 1, 2, \ldots$$

Therefore,

$$\begin{aligned}
\Pr(Y_1 + Y_2 \geqslant 2) &= \Pr(W \geqslant 2) \\
&= 1 - \Pr(W < 2) \\
&= 1 - \Pr(W = 0) - \Pr(W = 1) \\
&= 1 - e^{-6} - 6e^{-6} \\
&= 1 - \frac{7}{e^6} \\
&\approx 0.9826.
\end{aligned}$$

Hence, the probability that at least two calls will arrive in a given two minute period is about 98.3%. □

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

   Suggestion: Observe that $\Pr(Y = 50) = \Pr(49.5 < Y \leq 50.5)$, since Y is discrete.

   ***Solution**:* We use the Central Limit Theorem to estimate

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

   By the Central Limit Theorem,

$$\Pr(49.5 < Y \leqslant 50.5) \approx \Pr\left(\frac{49.5 - n\mu}{\sqrt{n}\sigma} < Z \leqslant \frac{50.5 - n\mu}{\sqrt{n}\sigma}\right), \tag{5}$$

   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 (5) that

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

Thus,

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

or about 8%. □

8. 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

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

   and the variance is

   $$\sigma^2 = \frac{(6+1)(6-1)}{12} = \frac{35}{12}. \tag{7}$$

   We also have that

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

   where $n = 36$.

By the Central Limit Theorem,

$$\Pr(107.5 < Y \leqslant 144.5) \approx \Pr\left(\frac{107.5 - n\mu}{\sqrt{n}\sigma} < Z \leqslant \frac{144.5 - n\mu}{\sqrt{n}\sigma}\right), \tag{8}$$

where $Z \sim \text{Normal}(0,1)$, $n = 36$, and μ and σ are given in (6) and (7), respectively. We then have from (8) 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%. □

9. Forty nine 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 $n = 49$, denote the 49 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

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

Consequently, the mean of the distribution is

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

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 (9),

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

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

$$\sigma^2 = 8.25. \tag{11}$$

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 (10), so that

$$\Pr(4 \leqslant \overline{X}_n \leqslant 6) = \Pr(-0.5 \leqslant \overline{X}_n - \mu \leqslant 1.5) \tag{12}$$

Next, divide the last inequality in (12) by $\sigma/\sqrt{n}$, where σ is as given in (11), to get

$$\Pr(4 \leqslant \overline{X}_n \leqslant 6) \doteq \Pr\left(-1.22 \leqslant \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \leqslant 3.66\right) \tag{13}$$

Since $n = 49$ can be considered a large sample size, we can apply the Central Limit Theorem to obtain from (13) that

$$\Pr(4 \leqslant \overline{X}_n \leqslant 6) \approx \Pr\left(-1.22 \leqslant Z \leqslant 3.66\right), \quad \text{where } Z \sim \text{Normal}(0,1). \tag{14}$$

It follows from (14) and the definition of the cdf that

$$\Pr(4 \leqslant \overline{X}_n \leqslant 6) \approx F_Z(3.66) - F_Z(-1.22), \tag{15}$$

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 (15) as

$$\Pr(4 \leqslant \overline{X}_n \leqslant 6) \approx F_Z(1.22) + F_Z(3.66) - 1. \tag{16}$$

Finally, using a table of standard normal probabilities, we obtain from (16) that

$$\Pr(4 \leqslant \overline{X}_n \leqslant 6) \approx 0.8888 + 1 - 1 = 0.8888.$$

Thus, the probability that the average of the 49 digits is between 4 and 6 is about 88.9%. □

10. 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:

$$\mu = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 2\frac{1}{4} = 1. \tag{17}$$

$$\sigma^2 = E(X^2) - [E(X)]^2, \tag{18}$$

where

$$E(X^2) = 0^2 \cdot \frac{1}{4} + 1^2 \cdot \frac{1}{2} + 2^2\frac{1}{4} = 1.5; \tag{19}$$

so that, combining (17), (18) and (19),

$$\sigma^2 = 1.5 - 1 = 0.5. \tag{20}$$

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,

$$\Pr[Y \leqslant 33.5], \tag{21}$$

By the Central Limit Theorem

$$\Pr\left(\frac{Y - n\mu}{\sqrt{n}\ \sigma} \leqslant \right) \approx \Pr(Z \leqslant z), \quad \text{for } z \in \mathbb{R}, \tag{22}$$

where $Z \sim \text{Normal}(0,1)$, $\mu = 1$, $\sigma^2 = 1.5$ and $n = 30$. It follows from (22) that we can estimate the probability in (21) by

$$\Pr[Y \leqslant 33.5] \approx \Pr(Z \leqslant 0.52) \doteq 0.6985. \tag{23}$$

Thus, according to (23), the probability that $X_1 + X_2 + \cdots + X_{30}$ is at most 33 is about 70%. □