# HW Solution 12 — Due: Not Due

*Lecturer: Prapun Suksompong, Ph.D.*

**Problem 1.** The input $X$ and output $Y$ of a system subject to random perturbations are described probabilistically by the following joint pmf matrix:

$$
\begin{array}{c|ccc}
 _X\!\diagdown^{\,y} & 2 & 4 & 5 \\
\hline
1 & 0.02 & 0.10 & 0.08 \\
3 & 0.08 & 0.32 & 0.40
\end{array}
$$

(a) Evaluate the following quantities:

    (i) The marginal pmf $p_X(x)$

    (ii) The marginal pmf $p_Y(y)$

    (iii) $\mathbb{E}X$

    (iv) $\operatorname{Var} X$

    (v) $\mathbb{E}Y$

    (vi) $\operatorname{Var} Y$

    (vii) $P[XY < 6]$

    (viii) $P[X = Y]$

    (ix) $\mathbb{E}[XY]$

    (x) $\mathbb{E}[(X - 3)(Y - 2)]$

    (xi) $\mathbb{E}[X(Y^3 - 11Y^2 + 38Y)]$

    (xii) $\operatorname{Cov}[X, Y]$

    (xiii) $\rho_{X,Y}$

(b) Find $\rho_{X,X}$

(c) Calculate the following quantities using the values of $\operatorname{Var} X$, $\operatorname{Cov}[X, Y]$, and $\rho_{X,Y}$ that you got earlier.

    (i) $\operatorname{Cov}[3X + 4, 6Y - 7]$

(ii) $\rho_{3X+4,6Y-7}$

(iii) $\text{Cov}\,[X, 6X-7]$

(iv) $\rho_{X,6X-7}$

**Solution**:

(a) The `MATLAB` codes are provided in the file `P_XY_EVarCov.m`.

(i) The marginal pmf $p_X(x)$ is founded by the sums along the rows of the pmf matrix:

$$p_X(x) = \begin{cases} 0.2, & x = 1 \\ 0.8, & x = 3 \\ 0, & \text{otherwise.} \end{cases}$$

(ii) The marginal pmf $p_Y(y)$ is founded by the sums along the columns of the pmf matrix:

$$p_Y(y) = \begin{cases} 0.1, & y = 2 \\ 0.42, & y = 4 \\ 0.48, & y = 5 \\ 0, & \text{otherwise.} \end{cases}$$

(iii) $\mathbb{E}X = \sum_x x p_X(x) = 1 \times 0.2 + 3 \times 0.8 = 0.2 + 2.4 = \boxed{2.6}$.

(iv) $\mathbb{E}[X^2] = \sum_x x^2 p_X(x) = 1^2 \times 0.2 + 3^2 \times 0.8 = 0.2 + 7.2 = 7.4$.

So, $\text{Var}\,X = \mathbb{E}[X^2] - (\mathbb{E}X)^2 = 7.4 - (2.6)^2 = 7.4 - 6.76 = \boxed{0.64}$.

(v) $\mathbb{E}Y = \sum_y y p_Y(y) = 2 \times 0.1 + 4 \times 0.42 + 5 \times 0.48 = 0.2 + 1.68 + 2.4 = \boxed{4.28}$.

(vi) $\mathbb{E}[Y^2] = \sum_y y^2 p_Y(y) = 2^2 \times 0.1 + 4^2 \times 0.42 + 5^2 \times 0.48 = 19.12$.

So, $\text{Var}\,Y = \mathbb{E}[Y^2] - (\mathbb{E}Y)^2 = 19.12 - 4.28^2 = \boxed{0.8016}$.

(vii) Among the 6 possible pairs of $(x, y)$ shown in the joint pmf matrix, only the pairs $(1, 2)$, $(1, 4)$, $(1, 5)$ satisfy $xy < 6$. Therefore, $[XY < 6] = [X = 1]$ which implies $P[XY < 6] = P[X = 1] = \boxed{0.2}$.

(viii) Among the 6 possible pairs of $(x, y)$ shown in the joint pmf matrix, there is no pair which has $x = y$. Therefore, $P[X = Y] = \boxed{0}$.

(ix) First, we calculate the values of $x \times y$:

$$\begin{array}{c|ccc} x \backslash y & 2 & 4 & 5 \\ \hline 1 & \left[\begin{array}{ccc} 2 & 4 & 5 \\ 3 & 6 & 12 & 15 \end{array}\right. \end{array}$$

Then, each $x \times y$ is weighted (multiplied) by the corresponding probability $p_{X,Y}(x, y)$:

$$
\begin{array}{c|ccc}
x \setminus y & 2 & 4 & 5 \\
1 & \begin{bmatrix} 0.04 & 0.40 & 0.40 \\ 0.48 & 3.84 & 6.00 \end{bmatrix}
\end{array}
$$

Finally, $\mathbb{E}[XY]$ is sum of these numbers. Therefore, $\mathbb{E}[XY] = \boxed{11.16}$.

(x) First, we calculate the values of $(x - 3) \times (y - 2)$:

$$
\begin{array}{c|ccc}
x \setminus y & 2 & 4 & 5 \\
1 & \begin{bmatrix} 0 & -4 & -6 \\ 0 & 0 & 0 \end{bmatrix}
\end{array}
$$

Then, each $(x - 3) \times (y - 2)$ is weighted (multiplied) by the corresponding probability $p_{X,Y}(x, y)$:

$$
\begin{array}{cc|ccc}
 & y - 2 & 0 & 2 & 3 \\
x - 3 & x \setminus y & 2 & 4 & 5 \\
-2 & 1 & \begin{bmatrix} 0 & -0.40 & -0.48 \\ 0 & 0 & 0 \end{bmatrix} \\
0 & 3 & 
\end{array}
$$

Finally, $\mathbb{E}[(X - 3)(Y - 2)]$ is sum of these numbers. Therefore,

$$\mathbb{E}[(X - 3)(Y - 2)] = \boxed{-0.88}.$$

(xi) First, we calculate the values of $x(y^3 - 11y^2 + 38y)$:

$$
\begin{array}{c|ccc}
y^3 - 11y^2 + 38y & 40 & 40 & 40 \\
x \setminus y & 2 & 4 & 5 \\
1 & \begin{bmatrix} 40 & 40 & 40 \\ 120 & 120 & 120 \end{bmatrix} \\
3 & 
\end{array}
$$

Then, each $x(y^3 - 11y^2 + 38y)$ is weighted (multiplied) by the corresponding probability $p_{X,Y}(x, y)$:

$$
\begin{array}{c|ccc}
x \setminus y & 2 & 4 & 5 \\
1 & \begin{bmatrix} 0.8 & 4.0 & 3.2 \\ 9.6 & 38.4 & 48.0 \end{bmatrix}
\end{array}
$$

Finally, $\mathbb{E}[X(Y^3 - 11Y^2 + 38Y)]$ is sum of these numbers. Therefore,

$$\mathbb{E}[X(Y^3 - 11Y^2 + 38Y)] = \boxed{104}.$$

(xii) $\text{Cov}[X, Y] = \mathbb{E}[XY] - \mathbb{E}X\mathbb{E}Y = 11.16 - (2.6)(4.28) = \boxed{0.032}$.

(xiii) $\rho_{X,Y} = \frac{\text{Cov}[X,Y]}{\sigma_X \sigma_Y} = \frac{0.032}{\sqrt{0.64}\sqrt{0.8016}} = \boxed{0.044677}$

(b) $\rho_{X,X} = \frac{\text{Cov}[X,X]}{\sigma_X \sigma_X} = \frac{\text{Var}[X]}{\sigma_X^2} = \boxed{1}$.

(c)

(i) $\text{Cov}[3X + 4, 6Y - 7] = 3 \times 6 \times \text{Cov}[X, Y] \approx 3 \times 6 \times 0.032 \approx \boxed{0.576}$.

(ii) Note that

$$
\begin{aligned}
\rho_{aX+b,cY+d} &= \frac{\text{Cov}[aX + b, cY + d]}{\sigma_{aX+b}\sigma_{cY+d}} \\
&= \frac{ac\text{Cov}[X, Y]}{|a|\sigma_X|c|\sigma_Y} = \frac{ac}{|ac|}\rho_{X,Y} = \text{sign}(ac) \times \rho_{X,Y}.
\end{aligned}
$$

Hence, $\rho_{3X+4,6Y-7} = \text{sign}(3 \times 4)\rho_{X,Y} = \rho_{X,Y} = \boxed{0.0447}$.

(iii) $\text{Cov}[X, 6X - 7] = 1 \times 6 \times \text{Cov}[X, X] = 6 \times \text{Var}[X] \approx \boxed{3.84}$.

(iv) $\rho_{X,6X-7} = \text{sign}(1 \times 6) \times \rho_{X,X} = \boxed{1}$.

**Problem 2.** Suppose $X \sim \text{binomial}(5, 1/3)$, $Y \sim \text{binomial}(7, 4/5)$, and $X \perp\!\!\!\perp Y$. Evaluate the following quantities.

(a) $\mathbb{E}[(X - 3)(Y - 2)]$

(b) $\text{Cov}[X, Y]$

(c) $\rho_{X,Y}$

    ***Solution***:

(a) First, because $X$ and $Y$ are independent, we have $\mathbb{E}[(X - 3)(Y - 2)] = \mathbb{E}[X - 3]\mathbb{E}[Y - 2]$. Recall that $\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$. Therefore, $\mathbb{E}[X - 3]\mathbb{E}[Y - 2] = (\mathbb{E}[X] - 3)(\mathbb{E}[Y] - 2)$ Now, for Binomial$(n, p)$, the expected value is $np$. So,

$$
(\mathbb{E}[X] - 3)(\mathbb{E}[Y] - 2) = \left(5 \times \frac{1}{3} - 3\right)\left(7 \times \frac{4}{5} - 2\right) = -\frac{4}{3} \times \frac{18}{5} = \boxed{-\frac{24}{5}} = -4.8.
$$

(b) $\text{Cov}[X, Y] = \boxed{0}$ because $X \perp\!\!\!\perp Y$.

(c) $\rho_{X,Y} = \boxed{0}$ because $\text{Cov}[X, Y] = 0$

# Extra Questions

Here are some optional questions for those who want more practice.

**Problem 3.** Let a continuous random variable $X$ denote the current measured in a thin copper wire in milliamperes. Assume that the probability density function of $X$ is

$$f_X(x) = \begin{cases} 5, & 4.9 \leq x \leq 5.1, \\ 0, & \text{otherwise.} \end{cases}$$

(a) Find the probability that a current measurement is less than 5 milliamperes.

(b) Find and plot the cumulative distribution function of the random variable $X$.

(c) Find the expected value of $X$.

(d) Find the variance and the standard deviation of $X$.

(e) Find the expected value of power when the resistance is 100 ohms?

***Solution***:

(a) $P[X < 5] = \int\limits_{-\infty}^{5} f_X(x)dx = \int\limits_{-\infty}^{0} \underbrace{f_X(x)}_{0}dx + \int\limits_{0}^{5} \underbrace{f_X(x)}_{5}dx = 0 + 5x|_{x=4.9}^{5} = \boxed{0.5}$.

(b) By definition, $F_X(x) = P[X \leq x] = \int\limits_{-\infty}^{x} f_X(t)dt$.

- For $x < 4.9$, $f_X(t) = 0$ for all $t$ inside $(-\infty, 4.9)$. Therefore,
$F_X(x) = \int_{-\infty}^{x} 0\,dt = 0$.

- For $4.9 \leq x \leq 5.1$,
$F_X(x) = \int\limits_{-\infty}^{x} f_X(t)dt = \int\limits_{-\infty}^{4.9} \underbrace{f_X(t)}_{0}dt + \int\limits_{4.9}^{x} \underbrace{f_X(t)}_{5}dt = 0 + 5t|_{t=4.9}^{x} = 5x - 24.5$.

- For $x > 5.1$,
$F_X(x) = \int\limits_{-\infty}^{x} f_X(t)dt = \int\limits_{-\infty}^{4.9} \underbrace{f_X(t)}_{0}dt + \int\limits_{4.9}^{5.1} \underbrace{f_X(t)}_{5}dt + \int\limits_{5.1}^{x} \underbrace{f_X(t)}_{0}dt = 0 + 5t|_{t=4.9}^{5.1} + 0 = 1$.

Combining the three cases above, we have the complete description of the cdf:

$$F_X(x) = \begin{cases} 0, & x < 4.9, \\ 5x - 24.5, & 4.9 \leq x \leq 5.1, \\ 1, & x > 5.1. \end{cases}$$

The corresponding plot is shown in Figure 12.1. Note that $F_X(x)$ is a continuous function; this is expected because $X$ is a continuous RV.
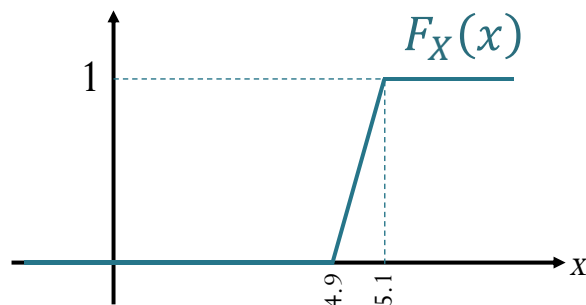
Figure 12.1: Plot of cdf for Problem 3.

(c) $\mathbb{E}X = \int\limits_{-\infty}^{\infty} x f_X(x)\, dx = \int\limits_{-\infty}^{4.9} x \underbrace{f_X(x)}_{0} dx + \int\limits_{4.9}^{5.1} x \underbrace{f_X(x)}_{5} dx + \int\limits_{5.1}^{x} x \underbrace{f_X(x)}_{0} dx = 0 + 5\frac{x^2}{2}\Big|_{x=4.9}^{5.1} +$

$0 = \boxed{5}$ mA.

Alternatively, for $X \sim \mathcal{U}(a,b)$, we have $\mathbb{E}X = \frac{a+b}{2} = \frac{4.9+5.1}{2} = 5$.

(d) $\operatorname{Var} X = \mathbb{E}[X^2] - (\mathbb{E}X)^2$. From the previous part, we know that $\mathbb{E}X = 5$. SO, to find $\operatorname{Var} X$, we need to find $\mathbb{E}[X^2]$: $\mathbb{E}[X^2] = \int\limits_{-\infty}^{\infty} x^2 f_X(x)\, dx = \int\limits_{-\infty}^{4.9} x^2 \underbrace{f_X(x)}_{0} dx +$

$\int\limits_{4.9}^{5.1} x^2 \underbrace{f_X(x)}_{5} dx + \int\limits_{5.1}^{x} x^2 \underbrace{f_X(x)}_{0} dx = 0 + 5\frac{x^3}{3}\Big|_{x=4.9}^{5.1} + 0 = 25 + \frac{1}{300}$.

Therefore, $\operatorname{Var} X = \mathbb{E}[X^2] - (\mathbb{E}X)^2 = \left(25 + \frac{1}{300}\right) = \boxed{\dfrac{1}{300}} \approx 0.0033$ (mA)$^2$

and

$\sigma_X = \sqrt{\operatorname{Var} X} = \boxed{\dfrac{1}{10\sqrt{3}}} \approx 0.0577$ mA.

Alternatively, for $X \sim \mathcal{U}(a,b)$, we have $\operatorname{Var} X = \frac{(b-a)^2}{12} = \frac{(5.1-4.9)^2}{12} = \frac{1}{300}$.

(e) Recall that $P = I \times V = I^2 r$. Here, $I = X$. Therefore, $P = X^2 r$ and $\mathbb{E}P = \mathbb{E}[X^2 r] = r\mathbb{E}[X^2] = 100 \times \left(25 + \frac{1}{300}\right) = 2500 + \frac{1}{3} \approx 2.50033 \times 10^3$ [(mA)$^2 \Omega$]. Factoring out $m^2$, we have $\mathbb{E}P \approx 2.50033$ mW. ([A$^2\Omega$] = [W].)

**Problem 4.** Let $X$ be a uniform random variable on the interval $[0,1]$. Set

$$A = \left[0, \frac{1}{2}\right), \quad B = \left[0, \frac{1}{4}\right) \cup \left[\frac{1}{2}, \frac{3}{4}\right), \quad \text{and } C = \left[0, \frac{1}{8}\right) \cup \left[\frac{1}{4}, \frac{3}{8}\right) \cup \left[\frac{1}{2}, \frac{5}{8}\right) \cup \left[\frac{3}{4}, \frac{7}{8}\right).$$

Are the events $[X \in A], [X \in B]$, and $[X \in C]$ independent?

*Solution*: Note that

$$P[X \in A] = \int_0^{\frac{1}{2}} dx = \frac{1}{2},$$

$$P[X \in B] = \int_0^{\frac{1}{4}} dx + \int_{\frac{1}{2}}^{\frac{3}{4}} dx = \frac{1}{2}, \text{ and}$$

$$P[X \in C] = \int_0^{\frac{1}{8}} dx + \int_{\frac{1}{4}}^{\frac{3}{8}} dx + \int_{\frac{1}{2}}^{\frac{5}{8}} dx + \int_{\frac{3}{4}}^{\frac{7}{8}} dx = \frac{1}{2}.$$

Now, for pairs of events, we have

$$P([X \in A] \cap [X \in B]) = \int_0^{\frac{1}{4}} dx = \frac{1}{4} = P[X \in A] \times P[X \in B], \tag{12.1}$$

$$P([X \in A] \cap [X \in C]) = \int_0^{\frac{1}{8}} dx + \int_{\frac{1}{4}}^{\frac{3}{8}} dx = \frac{1}{4} = P[X \in A] \times P[X \in C], \text{ and} \tag{12.2}$$

$$P([X \in B] \cap [X \in C]) = \int_0^{\frac{1}{8}} dx + \int_{\frac{1}{2}}^{\frac{5}{8}} dx = \frac{1}{4} = P[X \in B] \times P[X \in C]. \tag{12.3}$$

Finally,

$$P([X \in A] \cap [X \in B] \cap [X \in C]) = \int_0^{\frac{1}{8}} dx = \frac{1}{8} = P[X \in A] P[X \in B] P[X \in C]. \tag{12.4}$$

From (12.1), (12.2), (12.3) and (12.4), we can conclude that the events $[X \in A], [X \in B]$, and $[X \in C]$ are $\boxed{\text{independent}}$.

**Problem 5.** Cholesterol is a fatty substance that is an important part of the outer lining (membrane) of cells in the body of animals. Its normal range for an adult is 120–240 mg/dl. The Food and Nutrition Institute of the Philippines found that the total cholesterol level for Filipino adults has a mean of 159.2 mg/dl and 84.1% of adults have a cholesterol level below 200 mg/dl. Suppose that the cholesterol level in the population is normally distributed.

(a) Determine the standard deviation of this distribution.

(b) What is the value of the cholesterol level that exceeds 90% of the population?

(c) An adult is at moderate risk if cholesterol level is more than one but less than two standard deviations above the mean. What percentage of the population is at moderate risk according to this criterion?

(d) An adult is thought to be at high risk if his cholesterol level is more than two standard deviations above the mean. What percentage of the population is at high risk?

**Solution**: Let $X$ be the cholesterol level of a randomly chosen adult. It is given that $X \sim \mathcal{N}(m, \sigma^2)$ where $m = 159.2$ mg/dl. We also know that $P[X < 200] = 0.841$.

(a) For any Gaussian random variable, $P[X < 200] = \Phi\left(\frac{200-m}{\sigma}\right)$. It is given that this probability should be 0.841. Our plan is then to first find the number $z$ whose $\Phi(z) = 0.841$. Then, solve for $\sigma$ from $z = \frac{200-m}{\sigma}$.

From the $\Phi$ table, $\Phi(0.99) \approx 0.8389$ and $\Phi(1) \approx 0.8413$. Because 0.841 is closer to 0.8413 than 0.8389, we conclude that the value of $z$ that makes $\Phi(z) = 0.841$ is $z \approx 1$.

From $z = \frac{200-m}{\sigma}$, plugging-in $z \approx 1$ gives $\sigma \approx 200 - m = 200 - 159.2 = \boxed{40.8 \text{ mg/dl.}}$

(b) Here, we want to find the value of $x$ such that $P[X \leq x] = 0.9$. (90% of the population has cholesterol level lower than this $x$.)

For any Gaussian random variable, $P[X \leq x] = \Phi\left(\frac{x-m}{\sigma}\right)$. Our plan is then to first find the number $z$ whose $\Phi(z) = 0.9$. Then, solve for $x$ from $z = \frac{x-m}{\sigma}$.

From the $\Phi$ table, $\Phi(1.28) \approx 0.8997$ and $\Phi(1.29) \approx 0.9015$. Because 0.9 is closer to 0.8997 than 0.9015, we conclude that the value of $z$ that makes $\Phi(z) = 0.9$ is $z \approx 1.28$.

From $z = \frac{x-m}{\sigma}$, plugging-in $z \approx 1.28$ gives $x \approx 1.28\sigma + m \approx \boxed{211.424 \text{ mg/dl.}}$

(c) Here we want to find the probability that $m + \sigma < X < m + 2\sigma$:

$$P[m + \sigma < X < m + 2\sigma] = F_X(m + 2\sigma) - F_X(m + \sigma).$$

For any Gaussian random variable, $F_X(x) = \Phi\left(\frac{x-m}{\sigma}\right)$. Therefore,

$$P[m + \sigma < X < m + 2\sigma] = \Phi\left(\frac{(m + 2\sigma) - m}{\sigma}\right) - \Phi\left(\frac{(m + \sigma) - m}{\sigma}\right)$$

$$= \Phi(2) - \Phi(1) \approx 0.97725 - 0.8413 \approx \boxed{0.1359 = 13.59\%}$$

(d) $P[X > m + 2\sigma] = 1 - P[X \leq m + 2\sigma] = 1 - F_X(m + 2\sigma) = 1 - \Phi\left(\frac{(m+2\sigma)-m}{\sigma}\right) =$ $1 - \Phi(2) \approx 1 - 0.97725 \approx \boxed{0.0228 = 2.28\%}$.

**Problem 6** (Q3.5.6). Solve this question using the $\Phi/Q$ table.

A professor pays 25 cents for each blackboard error made in lecture to the student who points out the error. In a career of $n$ years filled with blackboard errors, the total amount in dollars paid can be approximated by a Gaussian random variable $Y_n$ with expected value $40n$ and variance $100n$.

(a) What is the probability that $Y_{20}$ exceeds 1000?

(b) How many years $n$ must the professor teach in order that $P[Y_n > 1000] > 0.99$?

    **Solution**: We are given[1] that $Y_n \sim \mathcal{N}(40n, 100n)$. Recall that when $X \sim \mathcal{N}(m, \sigma^2)$,

$$F_X(x) = \Phi\left(\frac{x - m}{\sigma}\right). \tag{12.5}$$

(a) Here $n = 20$. So, we have $Y_n \sim \mathcal{N}(40 \times 20, 100 \times 20) = \mathcal{N}(800, 2000)$. For this random variable $m = 800$ and $\sigma = \sqrt{2000}$.

We want to find $P[Y_{20} > 1000]$ which is the same as $1 - P[Y_20 \leq 1000]$. Expressing this quantity using cdf, we have

$$P[Y_{20} > 1000] = 1 - F_{Y_{20}}(1000).$$

Apply (12.5) to get

$$P[Y_{20} > 1000] = 1 - \Phi\left(\frac{1000 - 800}{\sqrt{2000}}\right) = 1 - \Phi(4.472) \approx Q(4.47) \approx \boxed{3.91 \times 10^{-6}}.$$

(b) Here, the value of $n$ is what we want. So, we will need to keep the formula in the general form. Again, from (12.5), for $Y_n \sim \mathcal{N}(40n, 100n)$, we have

$$P[Y_n > 1000] = 1 - F_{Y_n}(1000) = 1 - \Phi\left(\frac{1000 - 40n}{10\sqrt{n}}\right) = 1 - \Phi\left(\frac{100 - 4n}{\sqrt{n}}\right).$$

To find the value of $n$ such that $P[Y_n > 1000] > 0.99$, we will first find the value of $z$ which make

$$1 - \Phi(z) > 0.99. \tag{12.6}$$

At this point, we may try to solve for the value of $Z$ by noting that (12.6) is the same as

$$\Phi(z) < 0.01. \tag{12.7}$$

---

[1]Note that the expected value and the variance in this question are proportional to $n$. This naturally occurs when we consider the sum of i.i.d. random variables. The approximation by Gaussian random variable is a result of the central limit theorem (CLT).

Unfortunately, the tables that we have start with $\Phi(0) = 0.5$ and increase to something close to 1 when the argument of the $\Phi$ function is large. This means we can't directly find 0.01 in the table. Of course, 0.99 is in there and therefore we will need to solve (12.6) via another approach.

To do this, we use another property of the $\Phi$ function. Recall that $1 - \Phi(z) = \Phi(-z)$. Therefore, (12.6) is the same as

$$\Phi(-z) > 0.99. \tag{12.8}$$

From our table, we can then conclude that (12.7) (which is the same as (12.8)) will happen when $-z > 2.33$. (If you have `MATLAB`, then you can get a more accurate answer of 2.3263.)

Now, plugging in $z = \frac{100-4n}{\sqrt{n}}$, we have $\frac{4n-100}{\sqrt{n}} > 2.33$. To solve for $n$, we first let $x = \sqrt{n}$. In which case, we have $\frac{4x^2-100}{x} > 2.33$ or, equivalently, $4x^2 - 2.33x - 100 > 0$. The two roots are $x = -4.717$ and $x > 5.3$. So, We need $x < -4.717$ or $x > 5.3$. Note that $x = \sqrt{n}$ and therefore can not be negative. So, we only have one case; that is, we need $x > 5.3$. Because $n = x^2$, we then conclude that we need $\boxed{n > 28.1 \text{ years}}$.

**Problem 7.** The time until a chemical reaction is complete (in milliseconds) is approximated by the cumulative distribution function

$$F_X(x) = \begin{cases} 1 - e^{-0.01x}, & x \geq 0, \\ 0, & \text{otherwise.} \end{cases}$$

(a) Determine the probability density function of $X$.

(b) What proportion of reactions is complete within 200 milliseconds?

***Solution***: Note that the cdf $F_X(x)$ is a continuous function. Therefore, $X$ is a continuous RV.

(a) $f_X(x) = \frac{d}{dx} F_X(x) = \begin{cases} 0.01e^{-0.01x}, & x > 0, \\ 0, & x < 0. \end{cases}$

At $x = 0$, the derivative does not exist. However, $X$ is a continuous RV. Therefore, we can assign $f_X(0)$ to be any arbitrary value. Here, we set $f_X(0) = 0$:

$$f_X(x) = \boxed{\begin{cases} 0.01e^{-0.01x}, & x > 0, \\ 0, & \text{otherwise.} \end{cases}}$$

(b) $P[X < 200] = P[X \leq 200] = F_X(200) = 1 - e^{-0.01 \times 200} = \boxed{1 - e^{-2} \approx 0.8647.}$

**Problem 8.** Suppose $X \sim \text{binomial}(5, 1/3)$, $Y \sim \text{binomial}(7, 4/5)$, and $X \perp\!\!\!\perp Y$.

(a) A vector describing the pmf of $X$ can be created by the MATLAB expression:

$$\texttt{x = 0:5; pX = binopdf(x,5,1/3)}.$$

What is the expression that would give $\texttt{pY}$, a corresponding vector describing the pmf of $Y$?

(b) Use $\texttt{pX}$ and $\texttt{pY}$ from part (a), how can you create the joint pmf matrix in MATLAB? Do not use "for-loop", "while-loop", "if statement". Hint: Multiply them in an appropriate orientation.

(c) Use $\texttt{MATLAB}$ to evaluate the following quantities. Again, do not use "for-loop", "while-loop", "if statement".

   (i) $\mathbb{E}X$

   (ii) $P[X = Y]$

   (iii) $P[XY < 6]$

   *Solution*: The $\texttt{MATLAB}$ codes are provided in the file $\texttt{P\_XY\_jointfromMarginal\_indp.m}$.

(a) $\boxed{\texttt{y = 0:7; pY = binopdf(y,7,4/5);}}$

(b) $\boxed{\texttt{P = pX.'*pY;}}$

(c)

   (i) $\mathbb{E}X = \boxed{1.667}$

   (ii) $P[X = Y] = \boxed{0.0121}$

   (iii) $P[XY < 6] = \boxed{0.2727}$

**Problem 9.** Suppose $\text{Var}\, X = 5$. Find $\text{Cov}\,[X, X]$ and $\rho_{X,X}$.
   *Solution*:

(a) $\text{Cov}\,[X, X] = \mathbb{E}\,[(X - \mathbb{E}X)(X - \mathbb{E}X)] = \mathbb{E}\,[(X - \mathbb{E}X)^2] = \text{Var}\, X = \boxed{5}$.

(b) $\rho_{X,X} = \frac{\text{Cov}[X,X]}{\sigma_X \sigma_X} = \frac{\text{Var}\, X}{\sigma_X^2} = \frac{\text{Var}\, X}{\text{Var}\, X} = \boxed{1}$.

**Problem 10.** Suppose we know that $\sigma_X = \frac{\sqrt{21}}{10}$, $\sigma_Y = \frac{4\sqrt{6}}{5}$, $\rho_{X,Y} = -\frac{1}{\sqrt{126}}$.

(a) Find $\text{Var}[X + Y]$.

(b) Find $\mathbb{E}\left[(Y - 3X + 5)^2\right]$. Assume $\mathbb{E}\left[Y - 3X + 5\right] = 1$.

    ***Solution***:

(a) First, we know that $\operatorname{Var} X = \sigma_X^2 = \frac{21}{100}$, $\operatorname{Var} Y = \sigma_Y^2 = \frac{96}{25}$, and $\operatorname{Cov}\left[X, Y\right] = \rho_{X,Y} \times \sigma_X \times \sigma_Y = -\frac{2}{25}$. Now,

$$\operatorname{Var}\left[X + Y\right] = \mathbb{E}\left[((X + Y) - \mathbb{E}\left[X + Y\right])^2\right] = \mathbb{E}\left[((X - \mathbb{E}X) + (Y - \mathbb{E}Y))^2\right]$$
$$= \mathbb{E}\left[(X - \mathbb{E}X)^2\right] + 2\mathbb{E}\left[(X - \mathbb{E}X)(Y - \mathbb{E}Y)\right] + \mathbb{E}\left[(Y - \mathbb{E}Y)^2\right]$$
$$= \operatorname{Var} X + 2\operatorname{Cov}\left[X, Y\right] + \operatorname{Var} Y$$
$$= \boxed{\frac{389}{100}} = 3.89.$$

    Remark: It is useful to remember that

$$\operatorname{Var}\left[X + Y\right] = \operatorname{Var} X + 2\operatorname{Cov}\left[X, Y\right] + \operatorname{Var} Y.$$

    Note that when $X$ and $Y$ are uncorrelated, $\operatorname{Var}\left[X + Y\right] = \operatorname{Var} X + \operatorname{Var} Y$. This simpler formula also holds when $X$ and $Y$ are independence because independence is a stronger condition.

(b) First, we write

$$Y - aX - b = (Y - \mathbb{E}Y) - a(X - \mathbb{E}X) - \underbrace{(a\mathbb{E}X + b - \mathbb{E}Y)}_{c}.$$

    Now, using the expansion

$$(u + v + t)^2 = u^2 + v^2 + t^2 + 2uv + 2ut + 2vt,$$

    we have

$$(Y - aX - b)^2 = (Y - \mathbb{E}Y)^2 + a^2(X - \mathbb{E}X)^2 + c^2$$
$$- 2a(X - \mathbb{E}X)(Y - \mathbb{E}Y) - 2c(Y - \mathbb{E}Y) + 2a(X - \mathbb{E}X)c.$$

    Recall that $\mathbb{E}\left[X - \mathbb{E}X\right] = \mathbb{E}\left[Y - \mathbb{E}Y\right] = 0$. Therefore,

$$\mathbb{E}\left[(Y - aX - b)^2\right] = \operatorname{Var} Y + a^2 \operatorname{Var} X + c^2 - 2a\operatorname{Cov}\left[X, Y\right]$$

    Plugging back the value of $c$, we have

$$\boxed{\mathbb{E}\left[(Y - aX - b)^2\right] = \operatorname{Var} Y + a^2 \operatorname{Var} X + (\mathbb{E}\left[(Y - aX - b)\right])^2 - 2a\operatorname{Cov}\left[X, Y\right]}.$$

    Here, $a = 3$ and $b = -5$. Plugging these values along with the given quantities into the formula gives

$$\mathbb{E}\left[(Y - aX - b)^2\right] = \boxed{\frac{721}{100}} = 7.21.$$

**Problem 11.** The input $X$ and output $Y$ of a system subject to random perturbations are described probabilistically by the joint pmf $p_{X,Y}(x, y)$, where $x = 1, 2, 3$ and $y = 1, 2, 3, 4, 5$. Let $P$ denote the joint pmf matrix whose $i,j$ entry is $p_{X,Y}(i, j)$, and suppose that

$$P = \frac{1}{71} \begin{bmatrix} 7 & 2 & 8 & 5 & 4 \\ 4 & 2 & 5 & 5 & 9 \\ 2 & 4 & 8 & 5 & 1 \end{bmatrix}$$

(a) Find the marginal pmfs $p_X(x)$ and $p_Y(y)$.

(b) Find $\mathbb{E}X$

(c) Find $\mathbb{E}Y$

(d) Find $\text{Var}\, X$

(e) Find $\text{Var}\, Y$

   **Solution**: All of the calculations in this question are simply plugging numbers into appropriate formula. The `MATLAB` codes are provided in the file `P_XY_marginal_2.m`.

(a) The marginal pmf $p_X(x)$ is founded by the sums along the rows of the pmf matrix:

$$p_X(x) = \begin{cases} 26/71, & x = 1 \\ 25/71, & x = 2 \\ 20/71, & x = 3 \\ 0, & \text{otherwise} \end{cases} \approx \begin{cases} 0.3662, & x = 1 \\ 0.3521, & x = 2 \\ 0.2817, & x = 3 \\ 0, & \text{otherwise.} \end{cases}$$

The marginal pmf $p_Y(y)$ is founded by the sums along the columns of the pmf matrix:

$$p_Y(y) = \begin{cases} 13/71, & y = 1 \\ 8/71, & y = 2 \\ 21/71, & y = 3 \\ 15/71, & y = 4 \\ 14/71, & y = 5 \\ 0, & \text{otherwise} \end{cases} \approx \begin{cases} 0.1831, & y = 1 \\ 0.1127, & y = 2 \\ 0.2958, & y = 3 \\ 0.2113, & y = 4 \\ 0.1972, & y = 5 \\ 0, & \text{otherwise.} \end{cases}$$

(b) $\mathbb{E}X = \frac{136}{71} \approx 1.9155$

(c) $\mathbb{E}Y = \frac{222}{71} \approx 3.1268$

(d) $\text{Var}\, X = \frac{3230}{5041} \approx 0.6407$

(e) $\text{Var}\, Y = \frac{9220}{5041} \approx 1.8290$