CHAPTER 4

MATHEMATICAL EXPECTATION

4.1 Mean of a Random Variable

The *expected value*, or *mathematical expectation* $\mathsf{E}(X)$ of a random variable X is the long-run average value of X that would emerge after a very large number of observations. We often denote the expected value as μ_X, or μ if there is no confusion. $\mu_X = \mathsf{E}(X)$ is also referred to the mean of the random variable X, or the mean of the probability distribution of X. In the case of a finite population, the expected value is the population mean.

Consider a university with 15000 students and let X be the number of courses for which a randomly selected student is registered. The probability distribution of X is as follows:

x	1	2	3	4	5
No. of students	300	900	2850	4500	6450
$f(x)$	0.02	0.06	0.19	0.30	0.43

The average number of courses per student, or the average value of X in the population, results from computing the total number of courses taken by all students, and then dividing by the number of students in the population.

The mean, or average value of the random variable X, is therefore

$$\mu = \frac{1(300)+2(900)+3(2850)+4(4500)+5(6450)}{15000} = 4.06$$

Since

$$\frac{300}{15000} = 0.02 = f(1)$$
$$\frac{45900}{15000} = 0.06 = f(2),$$

and so on, an alternative expression for the mean is

$$\begin{aligned}
\mu &= 1\cdot f(1)+2\cdot f(2)+3\cdot f(3)+4\cdot f(4)+5\cdot f(5) \\
&= 1(0.02)+2(0.06)+3(0.19)+4(0.30)+5(0.43) \\
&= 0.02+0.12+0.57+1.20+2.15 \\
&= 4.06
\end{aligned}$$

Mean, or Expected Value of a random variable X

Let X be a random variable with probability distribution $f(x)$. The mean, or expected value, of X is

$$\mu = \mathsf{E}(X) = \begin{cases} \sum_x x\ f(x) & \text{if } X \text{ is discrete} \\ \int_{-\infty}^{\infty} x\ f(x)\ dx & \text{if } X \text{ is continuous} \end{cases}$$

EXAMPLE 4.1 (Discrete). Suppose that a random variable X has the following PMF:

x	-1	0	1	2
$f(x)$	0.3	0.1	0.4	0.2

Find $\mathrm{E}(X)$, the mathematical expectation of X.

EXAMPLE 4.2 (Continuous). Consider a random variable X with PDF

$$f(x) = \begin{cases} 3x^2 & \text{if } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}.$$

Find $\mathrm{E}(X)$.

EXAMPLE 4.3 (Interview). Six men and five women apply for an executive position in a small company. Two of the applicants are selected for interview. Let X denote the number of women in the interview pool. We have found the PMF of X in the previous chapter:

x	0	1	2
$f(x)$	$3/11$	$6/11$	$2/11$

How many women do you expect in the interview pool? That is, what is the expected value of X?

EXAMPLE 4.4 (Train Waiting). A commuter train arrives punctually at a station every half hour. Each morning, a commuter named John leaves his house and casually strolls to the train station. Let X denote the amount of time, in minutes, that John waits for the train from the time he reaches the train station. It is known that the PDF of X is

$$f(x) = \begin{cases} \frac{1}{30}, & \text{for } 0 < x < 30 \\ 0, & \text{otherwise.} \end{cases}$$

Obtain and interpret the expected value of the random variable X.

EXAMPLE 4.5 (DVD Failure). The time to failure in thousands of hours of an important piece of electronic equipment used in a manufactured DVD player has the density function

$$f(x) = \begin{cases} 2e^{-2x}, & x > 0 \\ 0, & \text{otherwise} \end{cases}.$$

Find the expected life of this piece of equipment.

Mean Value of $g(X)$

Let X be a random variable with probability distribution $f(x)$. The expected value of the random variable $g(X)$ is

$$\mu_{g(X)} = \mathsf{E}(g(X)) = \begin{cases} \sum_x g(x)f(x) & \text{if } X \text{ is discrete} \\ \int_{-\infty}^{\infty} g(x)f(x)\ dx & \text{if } X \text{ is continuous} \end{cases}$$

EXAMPLE 4.6. Refer to Example 4.1 (Discrete). Find the expected value of the random variable (X^2+1).

EXAMPLE 4.7. Refer to Example 4.2 (Continuous). Calculate $\mathrm{E}(X^2)$.

EXAMPLE 4.8. Refer to Example 4.3 (Interview). How many men do you expect in the interview pool? That is, find $\mathrm{E}(2-X)$.

EXAMPLE 4.9. Refer to Example 4.4 (Train Waiting). What is the average value of $\mathrm{E}(X/60)$? Can you interpret it?

EXAMPLE 4.10. Refer to Example 4.5 (DVD Failure). Find $\mathrm{E}(e^X)$.

EXAMPLE 4.11 (Insurance Payout). A group health insurance policy of a small business pays 100% of employee medical bills up to a maximum of \$1 million per policy year. The total annual medical bills, X, in millions of dollars, incurred by the employee has PDF given by

$$f(x) = \begin{cases} \dfrac{x(4-x)}{9}, & \text{for } 0<x<3 \\ 0, & \text{otherwise} \end{cases}.$$

Determine the expected annual payout by the insurance company, i.e., the expected value of $g(X)=\min\{X,1\}$.

Mean Value of $g(X,Y)$

Let X and Y be two random variable with joint probability distribution $f(x,y)$. The mean, or expected value of the random variable $g(X,Y)$ is

$$\mu_{g(X,Y)} = \mathrm{E}(g(X,Y)) = \begin{cases} \sum_x\sum_y g(x,y)f(x,y) & \text{discrete} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f(x,y)\,dxdy & \text{continuous} \end{cases}$$

EXAMPLE 4.12 (Joint). If X and Y are two random variables with the joint PMF:

$f(x,y)$	-1	0	1
0	0	0.1	0.2
1	0.3	0	0.1
2	0.2	0.1	0

Find $\mathrm{E}(XY)$ and $\mathrm{E}(XY^2)$.

Calculating $\mathrm{E}(X)$ or $\mathrm{E}(Y)$ based on the joint PMF/PDF

$$\mathrm{E}(X) = \begin{cases} \sum_x\sum_y xf(x,y) = \sum_x xg(x) & \text{discrete} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf(x,y)\,dydx = \int_{-\infty}^{\infty} xg(x)\,dx & \text{continuous} \end{cases}$$

$$\mathrm{E}(Y) = \begin{cases} \sum_y\sum_x yf(x,y) = \sum_y yh(y) & \text{discrete} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yf(x,y)\,dxdy = \int_{-\infty}^{\infty} yh(y)\,dy & \text{continuous} \end{cases}$$

EXAMPLE 4.13. Refer to Example 4.12 (Joint).

(a) Find the marginal PMF of X. Use it to calculate $\mathrm{E}(X)$.

(b) Find the marginal PMF of Y. Use it to calculate $\mathrm{E}(Y)$.

(c) Calculate $\mathrm{E}(X)$ using the joint PMF, i.e., $\mathrm{E}(X)=\sum\sum xf(x,y)$.

(d) Calculate $\mathrm{E}(Y)$ using the joint PMF, i.e., $\mathrm{E}(Y)=\sum\sum yf(x,y)$.

4.2 Variance and Covariance of Random Variables

The *variance* of a random variable X, or the variance of the probability distribution of X, is defined as the expected squared deviation from the expected value.

Variance & Standard Deviation

Let X be a random variable with probability distribution $f(x)$ and mean μ. The *variance* of X is

$$\begin{aligned}\sigma^2 &= \mathsf{Var}(X)\\ &= \mathsf{E}\left[(X-\mu)^2\right]\\ &= \mathsf{E}\left[(X-\mathsf{E}(X))^2\right]\\ &= \begin{cases}\sum_x (x-\mu)^2 f(x) & \text{if } X \text{ is discrete}\\ \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\, dx & \text{if } X \text{ is continuous}\end{cases}\end{aligned}$$

The positive square root of the variance, σ, is called the standard deviation of X.

EXAMPLE 4.14. Refer to Example 4.1 (Discrete). Find $\sigma^2 = \mathrm{Var}(X)$, the variance of X.

Note that

$$\begin{aligned}\mathsf{E}\left[(X-\mu)^2\right] &= \mathsf{E}\left(X^2-2\mu X+\mu^2\right)\\ &= \mathsf{E}\left(X^2\right)-2\mu\mathsf{E}(X)+\mu^2\\ &= \mathsf{E}\left(X^2\right)-\mu^2\end{aligned}$$

We often calculate the variance in the following way:

$$\mathsf{Var}(X) = \mathsf{E}\left(X^2\right) - [\mathsf{E}(X)]^2$$

EXAMPLE 4.15. Refer to Example 4.1 (Discrete). Find $\sigma^2 = \mathrm{Var}(X)$ using the above formula.

EXAMPLE 4.16. Refer to Example 4.2 (Continuous). Find $\mathrm{Var}(X)$.

EXAMPLE 4.17. Refer to Example 4.4 (Train Waiting). Calculate σ, the standard deviation of X.

EXAMPLE 4.18. Refer to Example 4.5 (DVD Failure). Calculate the variance of X.

Similar to the mathematical expectation, we can extend the concept of the variance of a random variable X to the variance of a function of X, say, $g(X)$.

Variance of $g(X)$

Let X be a random variable with probability distribution $f(x)$. The variance of the random variable $g(X)$ is

$$\begin{aligned}\sigma^2_{g(X)} &= \mathsf{E}\left\{\left[g(X)-\mu_{g(X)}\right]^2\right\}\\ &= \begin{cases}\sum_x \left[g(x)-\mu_{g(X)}\right]^2 f(x) & \text{discrete}\\ \int_{-\infty}^{\infty} \left[g(x)-\mu_{g(X)}\right]^2 f(x)\, dx & \text{continuous}\end{cases}\end{aligned}$$

It can also be calculated as follows:

$$\mathsf{Var}[g(X)] = \mathsf{E}\left\{[g(X)]^2\right\} - \{\mathsf{E}[g(X)]\}^2$$

EXAMPLE 4.19. Refer to Example 4.1 (Discrete). Find the variance of $\left(X^2+1\right)$.

EXAMPLE 4.20. Refer to Example 4.2 (Continuous). Find Var $\left(X^2\right)$.

Covariance of X and Y

Let X and Y be random variables with joint probability distribution $f(x,y)$. The *covariance* of X and Y is

$$\begin{aligned}\sigma_{XY} &= \mathsf{Cov}(X,Y)\\ &= \mathsf{E}\left[(X-\mu_X)(Y-\mu_Y)\right]\\ &= \begin{cases}\sum\limits_x\sum\limits_y (x-\mu_X)(y-\mu_Y)f(x,y) & \text{discrete}\\ \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}(x-\mu_X)(y-\mu_Y)f(x,y)\,dx\,dy & \text{continuous.}\end{cases}\end{aligned}$$

We often calculate $\mathsf{Cov}(X,Y)$ in the following way:

$$\mathsf{Cov}(X,Y)=\mathsf{E}(XY)-\mathsf{E}(X)\,\mathsf{E}(Y)$$

NOTE.
- The covariance is a measure of the association between the two random variables. The **sign** of the covariance indicates whether the relationship between two dependent random variables is positive or negative.
- If X and Y are statistically independent, then the covariance is zero. The converse, however, is not generally true.
- The association that the covariance measures between X and Y is the linear relationship.

EXAMPLE 4.21. Refer to Example 4.12 (Joint). Calculate the covariance of X and Y.

Correlation Coefficient of X and Y

Let X and Y be random variables with covariance σ_{XY} and standard deviations σ_X and σ_Y , respectively. The *correlation coefficient* of X and Y is

$$\rho_{XY}=\frac{\sigma_{XY}}{\sigma_X\sigma_Y}=\frac{\mathsf{Cov}(X,Y)}{\sqrt{\mathsf{Var}(X)}\sqrt{\mathsf{Var}(Y)}}$$

NOTE.
- Unlike the variance, the correlation coefficient ρ_{XY} is a **scale-free** measure. The magnitude of ρ_{XY} does not depend on the units used to measure both X and Y.
- The correlation coefficient ρ_{XY} indicates the strength of the relationship. $-1\le\rho_{XY}\le 1$.

EXAMPLE 4.22. Refer to Example 4.12 (Joint). Calculate the correlation coefficient of X and Y.

4.3 Means and Variances of Linear Combinations of Random Variables

Theorem. The expected value of the sum or difference of two or more functions of a random variable X is the sum or difference of the expected values of the functions. That is,

$$\mathrm{E}\left[g(X)\pm h(X)\right]=\mathrm{E}\left[g(X)\right]\pm\mathrm{E}\left[h(X)\right]$$

Proof. For continuous case,

$$\begin{aligned}\mathrm{E}\left[g(X)\pm h(X)\right] &= \int\left[g(x)\pm h(x)\right]f(x)\,dx\\ &= \int g(x)f(x)\,dx\pm\int h(x)f(x)\,dx\\ &= \mathrm{E}\left[g(X)\right]\pm\mathrm{E}\left[h(X)\right]\end{aligned}$$

□

Theorem. The expected value of the sum or difference of two or more functions of the random variables X and Y is the sum or difference of the expected values of the functions. That is,

$$\mathrm{E}[g(X,Y) \pm h(X,Y)] = \mathrm{E}[g(X,Y)] \pm \mathrm{E}[h(X,Y)]$$

Proof. For continuous case,

$$\begin{aligned}\mathrm{E}[g(X,Y) \pm h(X,Y)] &= \int\int [g(x,y) \pm h(x,y)]\, f(x,y)\; dxdy \\ &= \int\int g(x,y) f(x,y)\; dxdy \pm \int\int h(x,y) f(x,y)\; dxdy \\ &= \mathrm{E}[g(X,Y)] \pm \mathrm{E}[h(X,Y)]\end{aligned}$$

□

Corollary. $\mathrm{E}[g(X) \pm h(Y)] = \mathrm{E}[g(X)] \pm \mathrm{E}[h(Y)]$

Corollary. $\mathrm{E}[X \pm Y] = \mathrm{E}[X] \pm \mathrm{E}[Y]$

Theorem. If a, b and c are constants, then

$$\mathrm{E}(aX + bY + c) = a\mathrm{E}(X) + b\mathrm{E}(Y) + c$$

and

$$\mathrm{Var}(aX + bY + c) = a^2\mathrm{Var}(X) + b^2\mathrm{Var}(Y) + 2ab\mathrm{Cov}(X,Y)$$

Proof. $\mathrm{E}(aX + bY + c) = a\mathrm{E}(X) + b\mathrm{E}(Y) + c$.

$$\begin{aligned}\mathrm{Var}(aX + bY + c) &= \mathrm{E}\left\{[(aX + bY + c) - \mathrm{E}(aX + bY + c)]^2\right\} \\ &= \mathrm{E}\left\{[(aX + bY + c) - (a\mathrm{E}(X) + b\mathrm{E}(Y) + c)]^2\right\} \\ &= \mathrm{E}\left\{[a(X - \mathrm{E}(X)) + b(Y - \mathrm{E}(Y))]^2\right\} \\ &= a^2\mathrm{E}\left\{[X - \mathrm{E}(X)]^2\right\} + b^2\mathrm{E}\left\{[Y - \mathrm{E}(Y)]^2\right\} \\ &\quad + 2ab\mathrm{E}\{[X - \mathrm{E}(X)][Y - \mathrm{E}(Y)]\} \\ &= a^2\mathrm{Var}(X) + b^2\mathrm{Var}(Y) + 2ab\mathrm{Cov}(XY).\end{aligned}$$

□

Corollary. It can be easily verified that

$$\begin{array}{ll} \mathrm{E}(c) = c & \mathrm{Var}(c) = 0 \\ \mathrm{E}(X + c) = \mathrm{E}(X) + c & \mathrm{Var}(X + c) = \mathrm{Var}(X) \\ \mathrm{E}(aX) = a\mathrm{E}(X) & \mathrm{Var}(aX) = a^2\mathrm{Var}(X) \\ \mathrm{E}(aX + c) = a\mathrm{E}(X) + c & \mathrm{Var}(aX + c) = a^2\mathrm{Var}(X) \end{array}$$

Example 4.23. Suppose that X and Y are random variables with $\mathrm{E}(X) = 2$ and $\mathrm{E}(Y) = 3$, $\mathrm{Var}(X) = 4$, $\mathrm{Var}(Y) = 5$, and correlation coefficient $\rho = 0.6$. Let $Z = -2X + 4Y - 3$. Find

(a) $\mathrm{E}(Z)$

(b) $\mathrm{Cov}(X,Y)$

(c) $\mathrm{Var}(Z)$

Example 4.24. Refer to Example 4.1 (Discrete). Find $\mathrm{E}[(X-2)(X+1)]$.

Example 4.25. Refer to Example 4.2 (Continuous). Find $\mathrm{E}\left(3X^2 + 5X - 8\right)$.

Example 4.26. Refer to Example 4.12 (Joint). Find $\mathrm{Var}(X - 2Y + 3)$.

Theorem. Let X and Y be two independent random variables. Then

$$\mathrm{E}(XY) = \mathrm{E}(X)\ \mathrm{E}(Y).$$

Proof. For continuous case,

$$\begin{aligned}
\mathrm{E}(XY) &= \int\int xyf(x,y)\ dxdy \\
&= \int\int xyg(x)h(y)\ dxdy \\
&= \left[\int xg(x)\ dx\right]\left[\int yh(y)\ dy\right] \\
&= \mathrm{E}(X)\ \mathrm{E}(Y)
\end{aligned}$$

$\square$

COROLLARY. Let X and Y be two independent random variables. Then

$$\sigma_{XY} = \mathrm{Cov}(X,Y) = 0.$$

EXAMPLE 4.27. If X and Y are random variables with the joint density function

$$f(x,y) = \begin{cases} 6e^{-(2x+3y)} & \text{if } x>0, y>0 \\ 0 & \text{otherwise} \end{cases}.$$

Find ρ_{XY}.

COROLLARY. If X and Y are independent random variables, then

$$\mathrm{Var}(aX \pm bY) = a^2\mathrm{Var}(X) + b^2\mathrm{Var}(Y).$$

COROLLARY. If $X_1, X_2, \ldots, X_n$ are independent random variables, then

$$\mathrm{Var}\left(\sum_{i=1}^n a_iX_i\right) = \sum_{i=1}^n a_i^2\mathrm{Var}(X_i)$$

That is,

$$\mathrm{Var}(a_1X_1 + a_2X_2 + \cdots + a_nX_n) = a_1^2\mathrm{Var}(X_1) + a_2^2\mathrm{Var}(X_2) + \cdots + a_n^2\mathrm{Var}(X_n).$$

4.4 Other properties

In general,

$$\begin{aligned}
\mathsf{E}\left(\sum_{i=1}^n a_iX_i\right) &= \sum_{i=1}^n a_i\mathsf{E}(X_i) \\
\mathsf{Var}\left(\sum_{i=1}^n a_iX_i\right) &= \sum_{i=1}^n\sum_{j=1}^n a_ia_j\mathsf{Cov}(X_i,X_j) \\
&= \sum_{i=1}^n a_i^2\mathsf{Var}(X_i) + \sum_{i\neq j} a_ia_j\mathsf{Cov}(X_i,X_j) \\
&= \sum_{i=1}^n a_i^2\mathsf{Var}(X_i) + 2\sum_{1\leq i<j\leq n} a_ia_j\mathsf{Cov}(X_i,X_j) \\
\mathsf{Cov}\left(\sum_{i=1}^n a_iX_i, \sum_{j=1}^m b_jY_j\right) &= \sum_{i=1}^n\sum_{j=1}^m a_ib_j\mathsf{Cov}(X_i,Y_j)
\end{aligned}$$

As follows are more interesting properties of covariance.

$$\begin{aligned}
\mathsf{Cov}(X,a) &= 0 \\
\mathsf{Cov}(X,X) &= \mathsf{Var}(X) \\
\mathsf{Cov}(X,Y) &= \mathsf{Cov}(Y,X) \\
\mathsf{Cov}(aX,bY) &= ab\mathsf{Cov}(X,Y) \\
\mathsf{Cov}(X+a,Y+b) &= \mathsf{Cov}(X,Y) \\
\mathsf{Cov}(aX+bY,cZ+dW) &= ac\mathsf{Cov}(X,Z)+ad\mathsf{Cov}(X,W)+bc\mathsf{Cov}(Y,Z)+bd\mathsf{Cov}(Y,Z)
\end{aligned}$$

EXAMPLE 4.28. Prove that $\mathrm{Cov}(aX,bY) = ab\mathrm{Cov}(X,Y)$ where a and b are constants.

EXAMPLE 4.29. Suppose that X and Y are random variables with $\mathrm{Var}(X) = 4$, $\mathrm{Var}(Y) = 5$, and $\mathrm{Cov}(X,Y) = -3$. Calculate

(a) $\mathrm{Cov}(12X-2013,\ 2014)$

(b) $\mathrm{Cov}(5X,\ 6Y)$

(c) $\mathrm{Cov}(X+2013,\ -Y+2014)$

(d) $\mathrm{Cov}(X+2Y,\ -3X+4Y)$