Linear Combinations of Random Variables

Prerequisites

You should already have met the concept of the linear scaling and translation of the probability distribution of a random variable X in a number of contexts.

Expectation and variance under a scaling and transformation

Let X be a random variable with expectation $E(X)$ and variance $\text{var}(X)$. Let a and b be constants. Then

$$E(aX+b) = aE(X)+b$$
$$\text{var}(aX+b) = a^2\,\text{var}(X)$$

The expected value of X – that is, the expected mean is translated by the addition of b and multiplied by the scale factor a. The variance is multiplied by the square of the of the scale factor a and is unaffected by the addition of b.

Example (1)

The random variable X has mean 12 and variance 4. Find the mean and variance of the random variable $Y = 3X+5$.

Solution

$$\begin{aligned} E(Y) &= E(3X+5) \\ &= 3\times E(X)+5 \\ &= 3\times 12+5 \\ &= 41 \end{aligned}$$

$$\begin{aligned} \text{var}(Y) &= \text{var}(3X+5) \\ &= (3^2)\,\text{var}(X) \\ &= 9\times 4 \\ &= 36 \end{aligned}$$

The above results apply also to continuous probability distributions and so we can extend these formulae to the normal distribution.

Example (2)

X is a random variable that follows the normal distribution with mean μ and standard deviation σ. Find the mean and variance of the random variable $Y = aX + b$.

Solution

$X \sim N(\mu, \sigma^2)$

$$\begin{aligned} E(Y) &= E(aX + b) \\ &= a \times E(X) + b \\ &= a\mu + b \end{aligned}$$

$$\begin{aligned} \operatorname{var}(Y) &= \operatorname{var}(aX + b) \\ &= a^2 \operatorname{var}(X) \\ &= a^2\sigma^2 \end{aligned}$$

Example (3)

In an observatory the luminosity X of stars from a cluster is found to be normally distributed with mean 20 arbitrary units and variance 25. In order to standardise the results each luminosity is first multiplied by 4 and then 30 is subtracted. Find the probability that the resultant standard Y is greater than or equal to 100. No star is in fact given a standard value greater than 100. If 300 stars are observed in the cluster find to the nearest whole number the expected number of stars that will given the maximum standard value of 100.

Solution

$X \sim N(20, 16) = N(\mu, \sigma^2)$

$Y = 4X - 30$

$E(Y) = 4\mu - 30 = 4 \times 20 - 30 = 50$

$\operatorname{var}(Y) = (4)^2 \sigma^2 = 16 \times 25 = 400$

Thus $Y \sim N(50, 400)$

We require $P(Y > 100)$. The corresponding z-value is

$$z = \frac{y - \mu}{\sigma} = \frac{100 - 50}{\sqrt{400}} = 2.5$$

$P(Y > 100) = P(Z > 2.5) = 1 - \Phi(2.5) = 1 - 0.9938 = 0.0062$

The number of stars in the cluster that are expected to score 100 is $0.0062 \times 300 = 1.86$, which is 2 stars to the nearest whole number.

In this last example we are scaling and transforming a **single** probability distribution. A different kind of problem arises when two independent variables are combined. For example, suppose the weight of a printed sheet of card is normally distributed with mean 13.7 g and variance 0.2 g. and that the weight of an envelope is normally distributed with mean 5.7 g and variance 0.4 g. To find the probability that the combined weight of one sheet of card and one envelop will be greater than 20 g would be an instance where **two** independent random variables are linearly combined.

Linear combinations of independent random variables

Linear combination

Let X_1 and X_2 be two independent random variables. Let *a* and *b* be scalars. Then a *linear combination* of the variables X_1 and X_2 is defined to be any other random variable of the form $Y = aX_1 + bX_2$.

Independence

Two events, *A* & *B*, are independent if the probability of both *A* and *B* occurring together is equal to the product of the probability of *A* occurring and the probability of *B* occurring. $P(A+B) = P(A)\times P(B)$.

Let X_1 and X_2 be two random variables. Then to say that X_1 and X_2 are independent is to say that any observation (or trial) of X_1 does not affect the probability of an observation of X_2 and vice-versa. Let $X = x$ and $Y = y$ be the events X takes the value x and Y takes the value y respectively. Then *X* and *Y* are independent if

$P((X = x) \text{ and } (Y = y)) = P(X = x)\times P(Y = y)$.

Example (4)

An aeroplane has 100 seats and on every flight all the seats are taken. The passengers are either male or female. The number of male passengers *X* on any one flight is known to follow a normal distribution with mean 45 and variance 16.

(*a*) Let *Y* denote the number of female passengers on any one flight. Explain why *Y* is **not** independent of *X*.

(*b*) Find the mean and variance of *Y*.

Solution.

(*a*) Since the total number of passengers is always fixed, then once the number of male passengers X is known, then the number of female passengers Y is immediately determined by $Y = 100 - X$. If there are 50 male passengers then there must be 50 female passengers. Then, for instance, $P(Y = 40) = 0$ and

$P((X = 50) \text{ and } (Y = 40)) = 0 \neq P(X = x) \times P(Y = y)$.

(*b*) $Y = 100 - X$ and $X \sim N(45,16) = N(\mu, \sigma^2)$

Y arises from X by a scale factor of -1 and a translation of $+100$. Hence

$E(Y) = 100 - \mu = 100 - 45 = 55$

$\text{var}(Y) = \sigma^2 = 16$

We begin by stating the result for the expectation and variance when a random variable is formed from the linear combination of two independent random variables.

Expectation and variance of the linear combination of independent random variables

Let X and Y be independent random variables with expectations $E(X)$ and $E(Y)$ and variances $\text{var}(X)$ and $\text{var}(Y)$ respectively. Then

$$E(aX + bY) = aE(X) + bE(Y)$$

$$E(aX - bY) = aE(X) - bE(Y)$$

$$\text{var}(aX + bY) = a^2 \text{var}(X) + b^2 \text{var}(Y)$$

$$\text{var}(aX - bY) = a^2 \text{var}(X) + b^2 \text{var}(Y)$$

Pay particular attention to the last of these formulae as the variances are still added even when Y is subtracted from X.

Example (5)

The random variables X and Y are independent random variables with

$E(X) = 3$ $\quad$ $\text{var}(X) = 0.5$

$E(Y) = 2$ $\quad$ $\text{var}(Y) = 3$

Find the expectation and variance of

(a) $\quad 2X + 3Y$

(b) $\quad \frac{1}{2}X - 2Y$

Solution

(a) $E(2X+3Y)=2E(X)+3E(Y)=2\times 3+3\times 2=12$

$$\operatorname{var}(2X+3Y)=2^2\operatorname{var}(X)+3^2\operatorname{var}(Y)=4\times 0.5+9\times 3=29$$

(b) $E\left(\frac{1}{2}X-2Y\right)=\frac{1}{2}E(X)-2E(Y)=\frac{1}{2}\times 3-2\times 2=-\frac{5}{2}$

$$\operatorname{var}\left(\frac{1}{2}X-2Y\right)=\left(\frac{1}{2}\right)^2\operatorname{var}(X)+2^2\operatorname{var}(Y)=\frac{1}{4}\times\frac{1}{2}+4\times 3=12\frac{1}{8}$$

The above results apply to all linear combinations of random variables regardless of the form of their distribution. However, we will also be interested in the form the linear combination of two random variables takes – that is, in the nature of the distribution that arises when two distributions are combined in some way. The normal distribution will be of particular importance.

The difference between linear combinations and scalar multiples

Throughout this chapter we are looking at the following.

(1) The effect of taking a scalar multiple of a **single** random variable *X*. Let *X* be a random variable and let a be a scalar. Then $Y = aX$ is another random variable. We can find the distribution of *Y* in terms of the distribution of *X* and find the expectation and variance of *Y* in terms of the expectation and variance of *X*.

$$E(aX)=aE(X)$$
$$\operatorname{var}(aX)=a^2\operatorname{var}(X)$$

(2) The effect of taking a linear combination of **two** random variables *X* and *Y*. Let *X* and *Y* be independent random variables. Then $X+Y$ is another independent random variable. Sometimes, given the distributions of X and Y we can determine the distribution of $X+Y$. In particular we can find the expectation and variance of $X+Y$ in terms of the expectation and variance of *X* and *Y* respectively.

$$E(X+Y)=E(X)+E(Y)$$
$$E(X-Y)=E(X)-E(Y)$$
$$\operatorname{var}(X+Y)=\operatorname{var}(X)+\operatorname{var}(Y)$$
$$\operatorname{var}(X-Y)=\operatorname{var}(X)+\operatorname{var}(Y)$$

It should be quite clear that cases (1) and (2) are distinct. The formulae for the expectation and variance of the derived distributions are also different. It is certainly a mistake to confuse the two cases.

(3) We can also combine the two cases into a single formula dealing with the simultaneous effect of taking a linear combination of two independent random variables also subject to scalings. If X and Y are independent random variables and a and b are scalars, then

$$E(aX+bY)=aE(X)+bE(Y)$$

$$E(aX-bY)=aE(X)-bE(Y)$$

$$\operatorname{var}(aX+bY)=a^2\operatorname{var}(X)+b^2\operatorname{var}(Y)$$

$$\operatorname{var}(aX-bY)=a^2\operatorname{var}(X)+b^2\operatorname{var}(Y)$$

As indicated above it is important to maintain a clear distinction between the two cases, not least because the results for both are also different. However, there is one situation where confusion can naturally arise. In case (2) we deal with the linear combination of **two** independent random variables X and Y; specifically to form the random variable $X+Y$. However, suppose for instance that X and Y **both** represent the role of one fair cubical die, so that the probability distribution for X and Y is the same. Then X and Y are independent observations of the **same** random variable whose distribution is the uniform distribution

$$P(X=r)=P(Y=r)=\frac{1}{6} \qquad 0 \le r \le 6$$

In situations of this type we sometimes use suffixes, so that we denote the background distribution by, say, X and let X_1 and X_2 stand for two linearly independent observations of X. Their linear combination is X_1+X_2. However, since these are both observations of the same random variable X, we might also write this as $X+X$. Finally, a potential ambiguity can arise when instead of writing $X+X$ we also write $2X$. The expression $2X$ can consequently stand ambiguously for **two** totally different concepts

(1) Where $2X$ stands for the scalar multiple of the **single** random variable X. In this case

$$E(2X)=2E(X)$$
$$\operatorname{var}(2X)=2^2\operatorname{var}(X)=4\operatorname{var}(X)$$

(2) Where $2X$ stands for the linear combination $2X=X_1+X_2$ formed by adding the results of **two** separate independent observations (X_1, X_2) of a single random variable X. In this case

$$E(2X)=E(X_1+X_2)=E(X)+E(X)=2E(X)$$
$$\operatorname{var}(2X)=\operatorname{var}(X_1+X_2)=\operatorname{var}(X)+\operatorname{var}(X)=2\operatorname{var}(X)$$

In these two cases formulae for the variance differ. It would be nice to be able to avoid discussing this ambiguity altogether, but unfortunately, when you study sums of Poisson variables it will be necessary to keep this distinction in mind. This is because it has become customary practice to use the symbol $2X$ in the context of a Poisson variable to stand for the linear combination of Poisson variables and not for the scalar multiple. We will discuss this in more depth when we deal with sums of independent Poisson variables in a subsequent chapter.

Example (6)

A cubical die is thrown twice. Let X denote the random variable representing the score of one throw of this die. Assume both throws of the die are independent.

(a) Find the probability distribution of X and determine its expectation and variance.

(b) Let D represent the random variable that stands for double the score of one throw of the die. Determine the probability distribution of D and find its expectation and variance.

(c) Let $Y = X + X$ represent the random variable that stands for the score of the first throw of the die added to the score of the second throw. Determine the probability distribution of Y and find its expectation and variance.

Solution

(a) We indicated above that X is uniformly distributed

$$P(X=r)=P(Y=r)=\frac{1}{6} \qquad 0 \le r \le 6$$

Explicitly its probability distribution table is as follows.

r	0	1	2	3	4	5	6
$P(X=r)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

$$E(X)=\sum r\,P(X=r)=\frac{1+2+3+4+5+6}{6}=\frac{21}{6}=\frac{7}{2}$$

$$E(X^2)=\sum r^2\,P(X=r)=\frac{1+2^2+3^2+4^2+5^2+6^2}{6}=\frac{91}{6}$$

$$\operatorname{var}(X)=E(X^2)-\left[E(X)\right]^2=\left(\frac{91}{6}\right)-\left(\frac{7}{2}\right)^2=\frac{35}{12}$$

(b) The random variable D arises from just doubling the score for each throw of one die.

r	0	2	4	6	8	10	12
$P(X=r)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

$$E(D)=E(2X)=2E(X)=7$$

$$\mathrm{var}(D)=\mathrm{var}(2X)=4\,\mathrm{var}(X)=4\times\frac{35}{12}=\frac{35}{3}$$

(*c*) The following diagram represents the sample space of *Y*.

		first die					
		1	2	3	4	5	6
second die	1	2	3	4	5	6	7
	2	3	4	5	6	7	8
	3	4	5	6	7	8	9
	4	5	6	7	8	9	10
	5	6	7	8	9	10	11
	6	7	8	9	10	11	12

Its probability distribution is

r	2	3	4	5	6	7
$P(X=r)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$
r	8	9	10	11	12	
$P(X=r)$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	

$$E(Y)=E(X+X)=E(X)+E(X)=3.5+3.5=7$$

$$\mathrm{var}(Y)=\mathrm{var}(X)+\mathrm{var}(X)=2\,\mathrm{var}(X)=2\times\frac{35}{12}=\frac{35}{6}$$

Linear combinations of independent normal distributions

(1) Scaling and translation of a normal distribution

If X is normally distributed then so is $aX + b$, where a and b are constants.

$$X \sim N\left(\mu, \sigma^2\right) \quad \Rightarrow \quad aX + b \sim N\left(a\mu + b, a^2\sigma^2\right)$$

Example (7)

A company produces plastic wallets. Let X represent the width of the wallets. It is given that X is normally distributed with mean 32 cm and variance 0.4 cm. The machine producing these wallets is recalibrated to give a new width $Y = 2X - 10$ cm.

(*a*) State the distribution of *Y*.

(*b*) Find the probability that the new wallets will be within 1.0 cm of their expected mean. Give your answer to 2 significant figures.

Solution

(a) $X \sim N(32, 0.4)$

$$E(X) = 32 \qquad \text{var}(X) = 0.4$$

$$Y = 2X - 10$$

$$E(Y) = 2E(X) - 10 = 2 \times 32 - 10 = 54$$

$$\text{var}(Y) = 2^2 \text{var}(X) = 4 \times 0.4 = 1.6$$

$$Y \sim N(54, 1.6)$$

(b) We require $P(53 \le Y \le 55)$

$$z = \frac{x - \mu}{\sigma} = \frac{1}{\sqrt{1.6}} = 0.791$$

$$P(Z \le 0.791) = \Phi(0.791) = 0.7855 \left[\text{from tables}\right]$$

$$P(54 \le Y \le 55) = 0.7855 - 0.5 = 0.2855$$

$$P(53 \le Y \le 55) = 2 \times P(54 \le Z \le 55) = 2 \times 0.2855 = 0.571 = 0.57 \ \text{(2 s.f.)}$$

(2) Linear Combinations of independent normal distributions

If $X \sim N\left(\mu, \sigma^2\right)$ and $Y \sim N\left(\mu', (\sigma')^2\right)$ have independent normal distributions, then $aX + bY$ is also normally distributed and

$$E(aX + bY) = aE(X) + bE(Y)$$

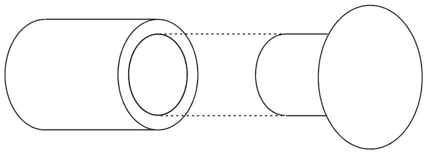
$$\text{var}(aX + bX) = a^2\text{var}(X) + b^2\text{var}(Y)$$

Hence

$$aX + bY \sim N\left(a\mu + b\mu',\ a^2\sigma^2 + b^2(\sigma')^2\right)$$

Example (8)

A certain manufacturing process depends on the insertion of a metal rivet into a metal gasket.



In order to create a high seal the gasket is heated. The initial diameter of the gasket, in mm, is $X \sim N(20.000,\ 0.001)$. The external diameter of the rivet, in mm, is

$Y \sim N(20.080,\ 0.001)$

(*a*) Find the probability that a randomly selected rivet would fit inside a randomly selected gasket without heating.

(*b*) The gasket is heated so that it expands by 1%. Find to 4 significant figures the proportion of rivets that fit the gasket.

(*c*) What percentage of rivets and gaskets randomly chosen will produce a high seal first time? Give your answer to 3 significant figures.

Solution

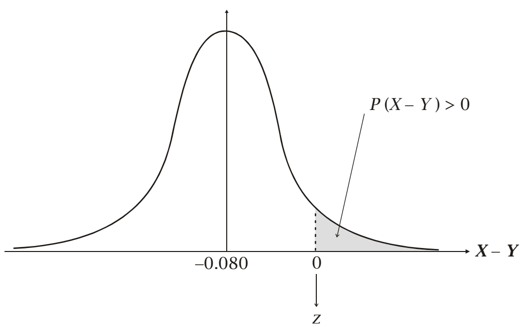
(*a*) We consider the distribution $X - Y$

$$E(X-Y) = E(X) - E(Y) = 20.000 - 20.080 = -0.080$$

$$\mathrm{var}(X-Y) = \mathrm{var}(X) + \mathrm{var}(Y) = 0.001 + 0.001 = 0.002$$

Then $X - Y \sim N(-0.080,\ 0.002)$

We require $P(Y < X) = P(X-Y) > 0$



The corresponding Z value is given by

$$z = \frac{x-\mu}{\sigma} = \frac{0-(-0.080)}{\sqrt{0.002}} = 1.788$$

$$P(Y < X) = P(Z > 1.788) = 0.0368 (3 \text{ s.f.})$$

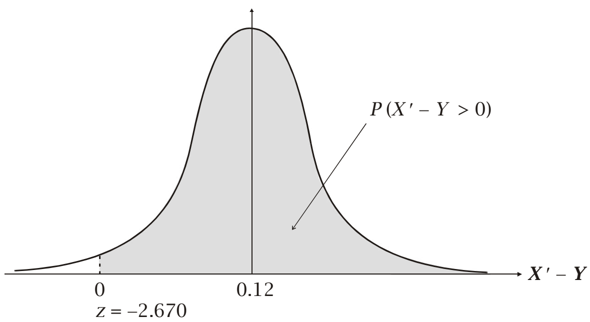
(*b*) The heating corresponds to scale factor 1.01. Let X' be the diameter of the gasket after heating then $X' = 1.01X$.

$$\begin{aligned} E(X'-Y) &= E(X') - E(Y) \\ &= 1.01E(X) - E(Y) \\ &= 1.01 \times 20.000 - 20.080 \\ &= 20.2 - 20.080 \\ &= 0.12 \end{aligned}$$

$$\begin{aligned} \operatorname{var}(X'-Y) &= \operatorname{var}(X') + \operatorname{var}(Y) \\ &= (1.01)^2 \operatorname{var}(X) + \operatorname{var}(Y) \\ &= (1.01)^2 \times 0.001 + 0.001 \\ &= 0.0020201 \end{aligned}$$

$$P(Y < X') = P(X'-Y > 0)$$

$$z = \frac{x-\mu}{\sigma} = \frac{0-0.12}{\sqrt{0.0020201}} = -2.6698... = -2.670 \ (3 \text{ d.p.})$$



$$P(Z > -2.670) = 0.9962 \ (4 \text{ s.f.})$$

(c) Proportion $= 1 -$ proportion not sealing when cold $-$ proportion not sealing when hot

$$\begin{aligned} &= 1 - 0.0368 - 0.0038 \\ &= 0.9594 \\ &= 95.9\% \ \ (3 \text{ s.f.}) \end{aligned}$$

Proofs

This section is optional. We will prove here the formulae for expectation and variance under a scaling and translation. We will then prove the formulae for the linear combination of independent random variables.

Expectation and variance under a scaling and transformation

Let X be a random variable with expectation $E(X)$. Let a and b be constants. Then

$$E(aX+b)=aE(X)+b$$
$$\operatorname{var}(aX+b)=a^2\operatorname{var}(X)$$

We will prove these formulae separately for a discrete probability distribution, and then for a continuous one.

<u>Discrete probability distribution</u>

Let X be a random variable with expectation $E(X)$ and variance $Var(X)$. Let a and b be constants. The definition of the expected mean is $E(X)=\sum xp$ where the expression xp means the product of the value and its probability and the expression $\sum$ indicates that we should sum all of these.

$$\begin{aligned}E(aX+b)&=\sum(ax+b)p\\&=\sum a(xp)+bp\\&=\sum a(xp)+\sum bp\\&=a\sum xp+b\sum p\\&=a\,E(X)+b\qquad\qquad \textstyle\sum p=1 \text{ is the law of total probability, and } E(X)=\sum xp\end{aligned}$$

To prove the formula for variance, note that

$$E\left[(aX)^2\right]=E\left(a^2X^2\right)=a^2E\left(X^2\right)$$

Then

$$\begin{aligned}\operatorname{var}(aX)&=E\left[(aX)^2\right]-\left[E(aX)\right]^2\\&=a^2E\left(X^2\right)-\left(aE(X)\right)^2\\&=a^2\left(E\left(X^2\right)-E(X)\right)\\&=a^2\operatorname{var}(X)\end{aligned}$$

Continuous probability distribution

Let X be a continuous random variable with probability density function $p(x)$. Then the expectation and variance of X are defined to be

$$E(X) = \int_{-\infty}^{\infty} xp(x)dx$$

$$E(X^2) = \int_{-\infty}^{\infty} x^2 p(x)dx$$

$$Var(X) = E(X^2) - [E(X)]^2$$

For the expectation of $aX + b$ we have

$$\begin{aligned}
E(aX+b) &= \int_{-\infty}^{\infty} (ax+b)p(ax+b)dx \\
&= \int_{-\infty}^{\infty} (ax+b)p(x)dx \\
&= \int_{-\infty}^{\infty} (axp(x)+bp(x))dx \\
&= \int_{-\infty}^{\infty} axp(x)dx + \int_{-\infty}^{\infty} bp(x)dx \\
&= a\int_{-\infty}^{\infty} xp(x)dx + b\int_{-\infty}^{\infty} p(x)dx \\
&= aE(X)+b
\end{aligned}$$

$$\begin{aligned}
\mathrm{var}(aX+b) &= \int_{-\infty}^{\infty} (ax+b)^2 p(ax+b) - [E(aX+b)]^2 \\
&= \int_{-\infty}^{\infty} (a^2x^2+2abx+b^2)p(x)dx - [aE(X)+b]^2 \\
&= a^2\int_{-\infty}^{\infty} x^2p(x)dx + 2ab\int_{-\infty}^{\infty} xp(x)dx + b\int_{-\infty}^{\infty} p(x)dx \\
&\quad -a^2[E(X)]^2 - 2abE(X) - b^2 \\
&= a^2\left(E(X^2) - [E(X)]^2\right) \\
&= a^2\,\mathrm{var}(X)
\end{aligned}$$

Thus, for any random variable, whether discrete or continuous

$$E(aX+b) = aE(X)+b$$
$$\mathrm{var}(aX+b) = a^2\,\mathrm{var}(X)$$

Linear combinations of independent random variables

To prove: Let X and Y be independent random variables with expectations $E(X)$ and $E(Y)$ and variances $\operatorname{var}(X)$ and $\operatorname{var}(Y)$ respectively. Then

$E(X+Y)=E(X)+E(Y)$ $\quad\quad$ $E(X-Y)=E(X)-E(Y)$

$\operatorname{var}(X+Y)=\operatorname{var}(X)+\operatorname{var}(Y)$ $\quad\quad$ $\operatorname{var}(X-Y)=\operatorname{var}(X)+\operatorname{var}(Y)$

For the expectation

$$E(X+Y)=\sum_{\text{all } x}\left[\sum_{\text{all } y}(x+y)P(X=x \textbf{ and } Y=y)\right]$$

$$=\sum_{\text{all } x}\left[\sum_{\text{all } y}(x+y)P(X=x)P(Y=y)\right] \qquad [\text{Independent events}]$$

$$=\sum_{\text{all } x}\left[\sum_{\text{all } y}x\,P(X=x)P(Y=y)\right]+\sum_{\text{all } x}\left[\sum_{\text{all } y}y\,P(X=x)P(Y=y)\right] \qquad [\text{Linear in } x \text{ and } y]$$

$$=\sum_{\text{all } y}\left[\sum_{\text{all } x}x\,P(X=x)P(Y=y)\right]+\sum_{\text{all } x}\left[\sum_{\text{all } y}y\,P(X=x)P(Y=y)\right] \qquad \left[\begin{array}{l}\text{We can change the}\\ \text{sums around}\end{array}\right]$$

$$=\sum_{\text{all } y}E(X)P(Y=y)+\sum_{\text{all } x}E(Y)P(X=x) \qquad \left[E(X)=\sum_{\text{all } x}x\,P(X=x)\right]$$

$$=E(X)\sum_{\text{all } y}P(Y=y)+E(Y)\sum_{\text{all } x}P(X=x) \qquad [E(X) \text{ is a scalar}]$$

$$=E(X)+E(Y) \qquad \left[\begin{array}{l}\sum_{\text{all } x}P(X=x)=1\\ \text{law of total probability}\end{array}\right]$$

From this

$$E\left[(X+Y)^2\right]=E\left[X^2+2XY+Y^2\right]$$
$$=E\left(X^2\right)+E(2XY)+E\left(Y^2\right)$$
$$=E\left(X^2\right)+2E(X)E(Y)+E\left(Y^2\right)$$

For the variance

$$\operatorname{var}(X+Y)=E\left[(X+Y)^2\right]-\left[E(X+Y)\right]^2$$
$$=E\left(X^2\right)+2E(X)E(Y)+E\left(Y^2\right)-\left(E(X)+E(Y)\right)^2$$
$$=E\left(X^2\right)+2E(X)E(Y)+E\left(Y^2\right)-\left[E(X)\right]^2-2E(X)E(Y)-\left[E(Y)\right]^2$$
$$=E\left(X^2\right)-\left[E(X)\right]^2+E\left(Y^2\right)-\left[E(Y)\right]^2$$
$$=\operatorname{var}(X)+\operatorname{var}(Y)$$

Example (8)

By adapting the proofs above, show that

$$E(X-Y)=E(X)-E(Y)$$

$$\mathrm{var}(X-Y)=\mathrm{var}(X)+\mathrm{var}(Y)$$

Solution

$$\begin{aligned}
E(X-Y) &= \sum_{\text{all } x}\left[\sum_{\text{all } y}(x-y)P(X=x \textbf{ and } Y=y)\right] \\
&= \sum_{\text{all } x}\left[\sum_{\text{all } y}(x-y)P(X=x)P(Y=y)\right] \\
&= \sum_{\text{all } x}\left[\sum_{\text{all } y}x\,P(X=x)P(Y=y)\right]-\sum_{\text{all } x}\left[\sum_{\text{all } y}y\,P(X=x)P(Y=y)\right] \\
&= E(X)-E(Y)
\end{aligned}$$

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

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