Theorem: Let X be a discrete R.V. with PMF $f(x)$, then $\mathbb{V}[X] = \sigma^2 = \mathbb{E}[X-\mu]^2 = \mathbb{E}[X^2] - \mu^2$

$$\Rightarrow \boxed{\mathbb{V}[X] = \mathbb{E}[X^2] - \{\mathbb{E}[X]\}^2}$$

Proof: By definition, $\mathbb{V}[X] = \sigma^2 = \mathbb{E}\{[X-\mu]^2\}$ where $\mu = \mathbb{E}[X]$

$$\Rightarrow \sigma^2 = \mathbb{E}[X^2 - 2X\mu + \mu^2]$$

$$\Rightarrow \sigma^2 = \mathbb{E}[X^2] - 2\mu\mathbb{E}[X] + \mathbb{E}[\mu^2]$$

$$\Rightarrow \sigma^2 = \mathbb{E}[X^2] - 2\mu^2 + \mu^2$$

$$\Rightarrow \sigma^2 = \mathbb{E}[X^2] - \mu^2$$

$$\Rightarrow \sigma^2 = \mathbb{E}[X^2] - \{\mathbb{E}[X]\}^2$$

$\mathbb{E}[\mu^2] = \mu^2 \mathbb{E}[1]$
$= \mu^2 \cdot 1$
$= \mu^2$

$\mathbb{E}[X^2] = \sum x^2 f(x)$

$\mathbb{E}[2X+1]$
$= \sum_x (2x+1) f(x)$

Th^m: The expected value of the sum of 2 R.Vs X and Y is the sum of the expected values of the R.V.s.

$$\mathbb{E}[X+Y] = \mathbb{E}[X] \pm \mathbb{E}[Y]$$

Theorem: ~~Let~~ Let X be a R.V. with a finite mean. Then for any numerical constants a and b, $\mathbb{E}[aX+b] = a\,\mathbb{E}[X] + b$.

Th^m: The expected value of the product of 2 R.V.s X and Y is equal to the product of their respective ~~perspected~~ expected values only if the R.Vs are independent.

$\boxed{\mathbb{E}[XY] = \mathbb{E}[X] \cdot \mathbb{E}[Y]}$ iff X & Y are independent.

$E[X] = \sum_{x} x f(x) = (-3)\cdot 6 \cdot 1/8 + (5)\cdot 2 \cdot 1/8 = -18/8 + 10/8 = -1$

Thus, the man is expected to lose TK.1 in the long-run.

<u>Problem:</u> A life insure company in BD offers to sell a TK.25,000 one-year term life insurance policy to a 25-year old man for a premium of TK.2500. According to BD life table, the probability of surviving 1-year for a 25-year old main is 0·97 & of his dying is 0·03. What is the company's expected gain in the long-run?

<u>Solution:</u> The gain X is a R.V that many take on the values TK.2500 if the man survives or 2500−25000 = TK.−22,500 if he dies. Consequently the probability distribution of X is shown in the following table:

$X:x$	2500	~~25000~~ −22,500
$P(X=x) = f(x)$	0·97	0·03

$E[X] = \sum x f(x) = (2500 * 0{\cdot}97) + (-22500 * 0{\cdot}03) = 1750$

<u>H.W</u>

P: Let X denotes the no. of spots showing on the face of a well-balanced dice after it is rolled once. If $Y = X^2 + 2X$, find (a) $E[X]$ · (b) $E[Y]$ (c) $E[X^2]$ (d) $E[Y^2]$ (e) $V[X]$ (f) $V[Y]$

[X 6 values
Y 6 values.
Draw Table.]

Covariance and Correlation:

18/7/25

When we consider the joint distribution of two R.Vs, it is useful to have a numerical summary that enables us to measure the association between the 2 variables. The covariance and correlation are the attempts to measure that association, or dependence.

Def^n (Covariance): The covariance of 2 R.Vs X and Y having finite expectations, $\mathbb{E}[X] = \mu_x$ and $\mathbb{E}[Y] = \mu_y$, denoted by $Cov(X,Y)$ is defined by.

$$Cov(X,Y) = \mathbb{E}[(X-\mu_x)(Y-\mu_y)] \quad \text{--- (1)}$$

It is provided that the expectations in equation (1) exist.

The value of $Cov(X,Y)$ may be positive, 0 or negative.

It follows from equ^n (1) that

$$Cov(X,Y) = \mathbb{E}[XY - \mu_x Y - \mu_y X + \mu_x \mu_Y]$$

$$= \mathbb{E}[XY] - \mu_x \mathbb{E}[Y] - \mu_Y \mathbb{E}[X] + \mathbb{E}[\mu_x \mu_Y]$$

$$= \mathbb{E}[XY] - \mu_x \mu_Y - \mu_Y \mu_x + \mu_x \mu_Y$$

$$\therefore \boxed{Cov(X,Y) = \mathbb{E}[XY] - \mu_x \mu_Y} \rightarrow$$ Formula used in mathematical problems.

Note that if X and Y are independent then $\mathbb{E}[XY] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$

$$Cov(X,Y) = \mathbb{E}[X]\mathbb{E}[Y] - \mu_x \mu_y = \mu_x \mu_y - \mu_x \mu_Y$$

$$\boxed{\therefore Cov(X,Y) = 0}$$

Problem: A discrete R.V X has a probability function shown in the following table:

Value of ~~X=x~~ $X:x$	-3	-2	0	1	2
$P(X=x)=f(x)$	0·10	0·30	0·15	0·40	0·05

Find $E[X]$ & $V[X]$.

Soln: By defn, $E[X]=\sum_x x f(x) = \sum_{x=-3}^{2} x f(x) = -0{\cdot}4$

$V[X]=E[(X-\mu)^2]$
$=\sum_x (x-\mu)^2 f(x)$
$= ~~(-3+0{\cdot}4)^2~~$

$V[X] = (-3+0{\cdot}4)^2 * 0{\cdot}10 + (-2+0{\cdot}4)^2 * 0{\cdot}30 + (0+0{\cdot}4)^2 * 0{\cdot}15 + (1+0{\cdot}4)^2 * 0{\cdot}40 + (2+0{\cdot}4)^2 * 0{\cdot}05$

$V[X] = E[X^2] - \{E[X]\}^2$
$= \sum_x x^2 f(x) - (-0{\cdot}4)^2$
$= \{(9 * 0{\cdot}10) + (4 * 0{\cdot}30) + (0 * 0{\cdot}15) + (1 * 0{\cdot}40) + (4 * 0{\cdot}05)\} - 0{\cdot}16$

Problem: In a coin-tossing game, a man is promised to receive TK.5 if he gets all heads or all tails when 3 coins are tossed. He loses TK.3 if either one or 2 head appear. How much is he expected to gain in the long run?

Soln: The R.V. X here is the amount of money (in TK) the man can win. Here the R.V. will take on a value 5 when the coins shown all heads or all tails. ~~On~~ Otherwise, -3.

Outcome	HHH	HHT	HTH	HTT	THH	THT	TTH	TTT
$X:x$	5	-3	-3	-3	-3	-3	-3	5
$P(X=x)=f(x)$	1/8	1/8	1/8	1/8	1/8	1/8	1/8	1/8

← for unbaised coins

Mathematical Expectation / Expected Values of a RV: $\mathbb{E}$ 17/7/25

Defⁿ: If X is a discrete random variable having a PMF $f_X(x)$ or $f(x)$, then the expected value of X or the mathematical expectation of X is denoted with $E[X]$ and is defined by

$$\mathbb{E}[X] = \sum_x x f(x)$$

In other words, the expected value of X is a weighted average of the possible values that X can take on; each value being weighted by the probability that X assumes.

On the other hand, if X is a continuous random variable having a PDF, $f(x)$ then the expected value of X is ~~is~~ defined by $\mathbb{E}[X] = \int_{-\infty}^{\infty} x f(x)\,dx$

Expected value = Mean

$\Rightarrow \mathbb{E}[X] = \mu$

Theorem: Let X be a discrete R.V with probability function $f(x)$ and c be any constant. Then $\mathbb{E}[c] = c$.

Proof: By definition of expectation, $\mathbb{E}[c] = \sum_x c f(x) = c \sum_x f(x) = c \cdot 1 = c$

$[\because \sum_x f(x) = 1$ for PMF$]$

$\therefore \mathbb{E}[c] = c$

Variance: Let X be a random variable with finite mean $\mu = \mathbb{E}[X]$, then the variance of X is defined with $\mathbb{V}[X]$ or $\mathbb{V}ar[X]$ and is defined by $\mathbb{V}[X] = \mathbb{E}[X-\mu]^2 = \mathbb{E}[\{X - E[X]\}^2] = \sigma^2$

Standard Deviation: The positive square root of the variance is known as standard deviation (SD), i.e, $\sigma = \sqrt{\mathbb{V}[X]} = \sqrt{\mathbb{E}[X-\mu]^2}$

$\therefore \sigma^2 = \mathbb{V}[X]$; ~~SD =~~ $\sigma = SD[X]$

Properties of Covariance:

For any random variables X, Y, Z and constants a, b

~~① $Cov(X,Y) = \mathbb{V}[X]$~~

① $Cov(X,X) = \mathbb{V}[X]$ ② $Cov(X,Y) = Cov(Y,X)$

③ $Cov(aX, bY) = ab\,Cov(X,Y)$ ④ $Cov(X+Y, Z) = Cov(X,Z) + Cov(Y,Z)$

Proof of ①: By defn

$$Cov(X,Y) = \mathbb{E}[XY] - \mu_X \mu_Y$$

$$\therefore Cov(X,X) = \mathbb{E}[X^2] - \mu_X^2 = \mathbb{V}[X]$$

[Proved]

Proof of ②: $Cov(X,Y) = \mathbb{E}[(X-\mu_X)(Y-\mu_Y)]$

$$= \mathbb{E}[(Y-\mu_Y)(X-\mu_X)] = Cov(Y,X)$$

~~$= \mathbb{E}[Y,$~~

Proof of ③: $Cov(aX, bY) = \mathbb{E}[\{aX - \mathbb{E}(aX)\}\{bY - \mathbb{E}(bY)\}]$

$$= \mathbb{E}[(aX - a\mathbb{E}[X])(bY - b\mathbb{E}[Y])]$$

$$= ab\,\mathbb{E}[(X-\mu_X)(Y-\mu_Y)]$$

$$= ab\,Cov(X,Y)$$

[Proved]

Proof of ④: $Cov(X+Y, Z) = \mathbb{E}[(X+Y).Z] - \mathbb{E}[(X+Y)]\,\mathbb{E}[Z]$

$$= \mathbb{E}[XZ + YZ] - \{\mathbb{E}[X] + \mathbb{E}[Y]\}\cdot\mathbb{E}[Z]$$

$$= \mathbb{E}[XZ] + \mathbb{E}[YZ] - \mathbb{E}[X]\cdot\mathbb{E}[Z] - \mathbb{E}[Y]\mathbb{E}[Z]$$

$$= Cov(X,Z) + Cov(Y,Z)$$

Also marginal.

Problem: Consider the following joint PDF of X and Y:

$$f(x,y) = \begin{cases} 6/5\,(x^2+2xy) \;;\; 0 \le x \le 1,\; 0 \le y \le 1 \\ 0 \;;\; \text{otherwise} \end{cases}$$

Find the covariance between X and Y, i.e, $Cov(X,Y) = ?$

Solution: The marginal density function of X is

$$g(x) = \begin{cases} 6/5\{x(x+1) \;;\; 0 \le x \le 1 \\ 0 \;;\; \text{Otherwise} \end{cases}$$

$g(x) = \int_{y=0} f(x,y)\,dy$ for all $0 \le x \le 1$.

Marginal density func of Y is

$$h(y) = \begin{cases} 2/5\,(1+3y) \;;\; 0 \le y \le 1 \\ 0 \;;\; \text{otherwise} \end{cases}$$

$$\mathbb{E}[X] = \int x\, f(x)\,dx$$

$$\mathbb{E}[X] = \sum_x x f(x)$$

$$\cancel{\mu_x}\ \mu_X = \mathbb{E}[X] = \int_0^1 x\, \frac{6}{5} x(x+1)\,dx = 6/5 \int_0^1 x^2(x+1)\,dx$$

$$= 7/10$$

$$\mu_Y = \mathbb{E}[Y] = \int_0^1 y\{2/5(1+3y)\,dy = 2/5 \int_0^1 (y+3y^2)\,dy = 3/5$$

$$Cov(X,Y) = \mathbb{E}[XY] - \mu_X \mu_Y$$

$$\mathbb{E}[XY] = \int_0^1 \int_0^1 xy\; 6/5\,(x^2+2xy)\,dy\,dx$$

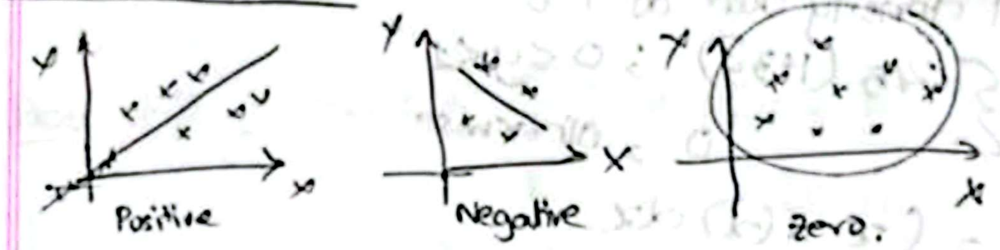
$$= 6/5 \int_0^1 \int_0^1 (x^3y + 2x^2y^2)\,dy\,dx = 5/12$$

Thus, $\mathrm{Cov}(X,Y) = \mathbb{E}[XY] - \mu_X \mu_Y = 5/12 - (7/10)(3/5)$
$= -1/300$.

Correlation: Correlation means association more precisely it is a measure of the extent to which 2 variables are related. There are 3 variables/possible results of a correlation study:

① Positive Correlation ② Negative Correlation ③ Zero correlation/ No correlation.

Scatter Diagram:



Number, no unit.

Correlation Coefficient: (Determine Correlation Strength.)

Defn: Let X & Y be 2 R.Vs with finite variances $\mathbb{V}[X]$ & $\mathbb{V}[Y]$ respectively. The correlation coefficient $\rho(X,Y)$ is defined to be zero if $\mathbb{V}[X]=0$ or $\mathbb{V}[Y]=0$, and otherwise

$$\boxed{\rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathbb{V}[X]\cdot\mathbb{V}[Y]}}}$$

Previous Problem: Also compute the correlation-coefficient betw X and Y.

$\mathbb{V}[X] = \mathbb{E}[X^2] - \{\mathbb{E}[X]\}^2$

$\mathbb{E}[X^2] = \int_0^1 x^2 g(x)\,dx = \frac{27}{50}$.

$$\mathbb{E}[Y^2] = \int_0^1 y^2 \, h(y)\, dy = 13/30$$

$$\mathbb{V}[X] = \mathbb{E}[X^2] - \{\mathbb{E}[X]\}^2 = \frac{27}{50} - (7/10)^2 = 1/20$$

$$\mathbb{V}[Y] = \mathbb{E}[Y^2] - \{\mathbb{E}[Y]\}^2 = 13/30 - (3/5)^2 = 11/150$$

$$\rho(X,Y) = \frac{-1/300}{\sqrt{1/20 \cdot 11/150}} = -0.055$$

low strength.

∴ Therefore, X and Y are negatively correlated.