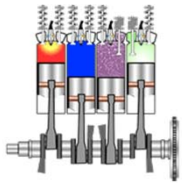
Joint Probability Distribution II

Today's Class

- Expected Values
- Covariance
- Correlation

Example of E(X)

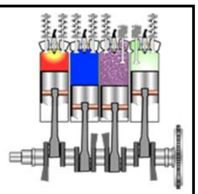


- An internal combustion engine contains several cylinders bored into the engine block. Let X represent the bore diameter of the cylinder (mm). Assume that the pdf of X is

$$f(x) = \begin{cases} 10 & 80.5 < x < 80.6 \\ 0 & otherwise \end{cases}$$

  - Let A = $\pi X^2/4$ represent the area of the bore. Find the expected value of A.

Solution



$$E_A = \int_{-\infty}^{\infty} \frac{\pi x^2}{4} f(x)dx$$

$$= \int_{80.5}^{80.6} \frac{\pi x^2}{4} 10dx$$

$$= \frac{10\pi(80.6^3 - 80.5^3)}{4 \times 3}$$

$$= 5096$$

Expected Value

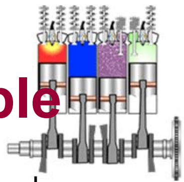
- Expected value of *X* and *Y*,

$$E(X,Y) = \begin{cases} \sum_x \sum_y xyp(x,y) & \text{If X and Y are discrete} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)dxdy & \text{If X and Y are continuous} \end{cases}$$

- Expected value of a function of *X* and *Y*, *h(x,y)*

$$E[h(X,Y)] = \begin{cases} \sum_x \sum_y h(x,y)p(x,y) & \text{If X and Y are discrete} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y)f(x,y)dxdy & \text{If X and Y are continuous} \end{cases}$$

Expected Value Example

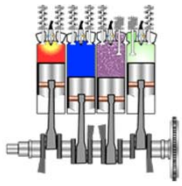


- The displacement of a piston in an internal combustion engine is defined to be the column that the top of the piston moves through from the top to the bottom of its stroke. Let X represent the diameter of the cylinder bore (mm) and Y the length of the piston stroke (mm). The displacement is given by D = $\pi X^2 Y/4$. Assume X and Y are jointly distributed with joint pmf.

$$f(x,y) = \begin{cases} 100 & 80.5 < x < 80.6 \text{ and } 65.1 < y < 65.2 \\ 0 & otherwise \end{cases}$$

  - Find the expected value of D.

Solution



$$E_D = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\pi x^2 y}{4} f(x,y)dxdy$$

$$= \int_{65.1}^{65.2}\int_{80.5}^{80.6} \frac{\pi x^2 y}{4} 100dxdy$$

$$= \int_{80.5}^{80.6} \frac{100\pi x^2(65.2^2-65.1^2)}{4\times 2}$$

$$= \frac{100\pi(80.6^3-80.5^3)(65.2^2-65.1^2)}{4\times 2\times 3}$$

$$= 331{,}998$$

Covariance

- Covariance between two variables is defined as follows:

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

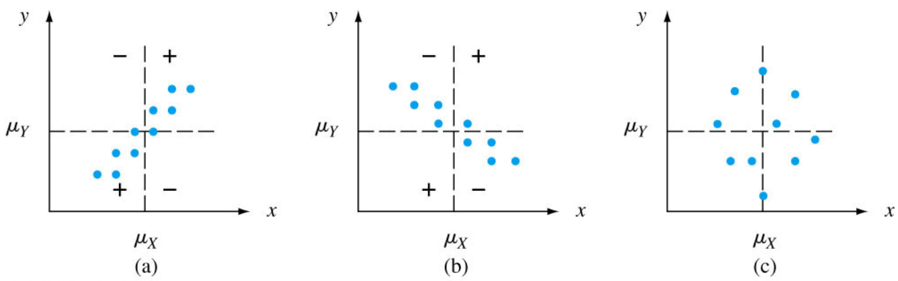
$$= \begin{cases} \sum_x \sum_y (x-\mu_x)(y-\mu_y)p(x,y) & \text{If } X \text{ and } Y \text{ are discrete} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\mu_x)(y-\mu_y)f(x,y)dxdy & \text{If } X \text{ and } Y \text{ are continuous} \end{cases}$$

$$Cov[X,Y] = E[XY] - E[X]E[Y]$$

Proof of Covariance

$$E[(X-\mu_X)\times(Y-\mu_Y)]$$
$$= E(XY - X\mu_Y - Y\mu_Y + \mu_x\mu_Y)$$
$$= E(XY) - E(X\mu_Y) - E(Y\mu_Y) + E(\mu_x\mu_Y)$$
$$= E(XY) - \mu_Y E(X) - \mu_X E(Y) + \mu_x\mu_Y$$
$$= E(XY) - E(X)E(Y) - E(X)E(Y) + E(X)E(Y)$$
$$= E(XY) - E(X)E(Y)$$

Covariance



Positive Covariance   Negative Covariance   Covariance ≈ 0

Covariance Example

- Quality-control checks on wood paneling involve counting the number of surface flaws. On a given 2×8 ft panel, let X and Y be the numbers of surface flaws due to uneven application of the final coat of finishing material, and due to inclusions of foreign particles in the finish, respectively. The joint pmf p(x,y) of X and Y is given as follows.

	y		
x	0	1	2
0	0.05	0.10	0.20
1	0.05	0.15	0.05
2	0.25	0.10	0.05

  - Find the E[XY]

Solution

	y		
x	0	1	2
0	0.05	0.10	0.20
1	0.05	0.15	0.05
2	0.25	0.10	0.05

$$E[XY] = \sum_{x=0}^{2} \sum_{y=0}^{2} xy\, p(x,y)$$

$$=1\times1\times0.15 + 1\times2\times0.05 + 2\times1\times0.1 + 2\times2\times0.05$$

$$=0.65$$

Covariance Example

- Quality-control checks on wood paneling involve counting the number of surface flaws. On a given 2×8 ft panel, let X and Y be the numbers of surface flaws due to uneven application of the final coat of finishing material, and due to inclusions of foreign particles in the finish, respectively. The joint pmf p(x,y) of X and Y is given as follows.

	y		
x	0	1	2
0	0.05	0.10	0.20
1	0.05	0.15	0.05
2	0.25	0.10	0.05

  - Find the covariance of X and Y

Solution

	y			
x	0	1	2	$p_x(x)$
0	0.05	0.10	0.20	**0.35**
1	0.05	0.15	0.05	**0.25**
2	0.25	0.10	0.05	**0.40**
$p_y(y)$	**0.35**	**0.35**	**0.30**	

$E[X]$
$= 0\times0.35 + 1\times0.25 + 2\times0.40 = 1.05$
$E[Y]$
$= 0\times0.35 + 1\times0.35 + 2\times0.30 = 0.95$

$Cov[X,Y] = E[XY] - E[X]E[Y]$
$= 0.65 - 1.05\times0.95 = -0.3475$

Correlation

- Correlation Coefficient

$$\rho_{X,Y} = Corr(X,Y) = \frac{Cov[X,Y]}{\sigma_x \sigma_y}$$

- For any two rv's X and Y

$$-1 \le \rho_{X,Y} \le 1$$

Correlation Example

- Quality-control checks on wood paneling involve counting the number of surface flaws. On a given 2×8 ft panel, let X and Y be the numbers of surface flaws due to uneven application of the final coat of finishing material, and due to inclusions of foreign particles in the finish, respectively. The joint pmf p(x,y) of X and Y is given as follows.

		y	
x	0	1	2
0	0.05	0.10	0.20
1	0.05	0.15	0.05
2	0.25	0.10	0.05

  - Find the correlation of X and Y.

Solution

	y			
x	0	1	2	$p_x(x)$
0	0.05	0.10	0.20	**0.35**
1	0.05	0.15	0.05	**0.25**
2	0.25	0.10	0.05	**0.40**
$p_y(y)$	**0.35**	**0.35**	**0.30**	

$$V[X] = E[X^2] - E[X]^2$$

$$= 0^2 \times 0.35 + 1^2 \times 0.25 + 2^2 \times 0.40 - 1.05^2$$

$$= 0.7475$$

$$V[Y] = 0^2 \times 0.35 + 1^2 \times 0.35 + 2^2 \times 0.30 - 0.95^2$$

$$= 0.6475$$

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sqrt{V[X]V[Y]}} = \frac{-0.3475}{\sqrt{0.7475 \times 0.6475}} = = -0.4995$$

Independence

- If $Cov(X,Y) = \rho_{X,Y} = 0$,
  then X and Y are said to be uncorrelated
- If X and Y are independent, then ρ=0,
  but ρ=0 does not imply independence

Linear Combination of 2 rvs

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

$V(X+Y) = V(X) + V(Y) + 2Cov(X,Y)$

$V(X-Y) = V(X) + V(Y) - 2Cov(X,Y)$

- If X and Y are independent then:

  $V(X+Y) = V(X)+V(Y)$

  $V(X-Y) = V(X)+V(Y)$

Example

- Assume that the mobile computer moves from a random position (X,Y) vertically to the point (X,0) and then along the x axis to the origin

  $E(X)=0.80$, $E(Y)=0.53$

  $V(X)=0.027$, $V(Y)=0.049$

  $Cov(X,Y)= 0.018$

  - Find the mean of the distance traveled

Solution

E(X)=0.80

E(Y)=0.53

E(X+Y) = E(X) + E(Y)

= 0.80 + 0.53

= 1.33

Example

- Assume that the mobile computer moves from a random position (X,Y) vertically to the point (X,0) and then along the x axis to the origin

  E(X)=0.80, E(Y)=0.53

  V(X)=0.027, V(Y)=0.049

  Cov(X,Y)= 0.018

  - Find the variance of the distance traveled

Solution

V(X)=0.027

V(Y)=0.049

Cov(X,Y)= 0.018

$$V(X+Y) = V(X) + V(Y) + 2Cov(X,Y)$$

$$= 0.027 + 0.049 + 2\times 0.018$$

$$= 0.112$$