Random Variables Lecture 5

Properties of Variance

We have seen that $E[X_1+\dots+X_n] = E[X_1]+\dots+E[X_n]$ in every circumstance. It is important to know whether $Var(X_1+\dots+X_n) = Var(X_1)+\dots+Var(X_n)$? We need the following idea

Covariance

Def: $Cov(X,Y) = E[(X-\mu_x)(Y-\mu_y)]$

Observe that $E[(X-\mu_x)(Y-\mu_y)] = E[XY-\mu_x Y - X\mu_y + \mu_x\mu_y]$

$= E[XY] - \mu_x E[Y] - \mu_y E[X] + \mu_x \mu_y =$

$= E[XY] - \mu_x\mu_y - \mu_x\mu_y + \mu_x\mu_y = E[XY] - E[X]E[Y]$.

That is $\boxed{Cov(X,Y) = E[XY] - E[X]E[Y]}$

Here are some other simple properties:

(a) $Cov(X,Y) = Cov(Y,X)$

(b) $Cov(X,X) = Var(X)$

(c) $Cov(aX,Y) = a\,Cov(X,Y)$

(d) $Cov(X_1+X_2,Y) = Cov(X_1,Y) + Cov(X_2,Y)$

<u>Proof:</u>

(a) $Cov(X,Y) = E[X,Y] - E[X]E[Y] =$

$= E[YX] - E[Y]E[X] = Cov(Y,X)$.

(b) $Cov(X,X) = E[XX] - E[X]E[X] = E[X^2] - (E[X])^2$

$= Var(X)$.

(c) $Cov(aX,Y) = E[aX \cdot Y] - E[aX]E[Y] =$

$= aE[XY] - aE[X]E[Y] = a\,Cov(X,Y)$.

(d) $Cov(X_1+X_2,Y) = E[(X_1+X_2)Y] - E[X_1+X_2]E[Y]$

$= E[X_1Y] + E[X_2Y] - (E[X_1]+E[X_2])E[Y] =$

$= \underline{E[X_1Y]} + \underline{E[X_2Y]} - \underline{E[X_1]E[Y]} - \underline{E[X_2]E[Y]}$

$= Cov(X_1,Y) + Cov(X_2,Y)$.

Thus $Cov\left(\sum_{k=1}^{n} X_k, Y\right) = \sum_{k=1}^{n} Cov(X_k,Y)$

and

$$Cov\left(\sum_{k=1}^{n} X_k, \sum_{j=1}^{m} Y_j\right) = \sum_{k=1}^{n}\sum_{j=1}^{m} Cov(X_k,Y_j)$$

<u>Corollary:</u> $Var\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} Var(X_k) +$

$$+ \sum_{k=1}^{n}\sum_{\substack{j=1 \\ j\neq k}}^{n} Cov(X_k,X_j)$$

Ex. $Var(X+Y) = Var(X) + Var(Y) + 2\,Cov(X,Y)$.

$$Var(X+Y+Z) = Var(X) + Var(Y) + Var(Z) + 2\,Cov(X,Y) + 2\,Cov(X,Z) + 2\,Cov(Y,Z).$$

Note: The above calculations indicate that, in general, $Var(X_1 + \dots + X_n) \neq Var(X_1) + \dots + Var(X_n)$.

Thm: If X and Y are independent random variables then $Cov(X,Y) = 0$

Proof: $Cov(X,Y) = E[XY] - E[X]E[Y]$

where $E[XY] = \sum_i \sum_j x_i y_j P(X = x_i, Y = y_j)$

$$= \sum_i \sum_j x_i y_j P(X = x_i) P(Y = y_j)$$

$$= \left(\sum_i x_i P(X = x_i) \right) \left(\sum_j y_j P(\cancel{X} = y_j) \right) = E[X]E[Y].$$

Corollary: If $X_1, \dots, X_n$ are independent random variables,

$$Var(X_1 + \dots + X_n) = Var(X_1) + \dots + Var(X_n).$$

Ex. There are 5 fair coins. One coin is picked at random and tossed. Let X - # of fair coin (coin 1, ..., coin 5) and $Y = \begin{cases} 0 & \text{if tails} \\ 1 & \text{if heads} \end{cases}$.

Calculate

(a) $E[X]$ and $E[Y]$

(b) $Var(X)$ and $Var(Y)$

(c) $Cov(X,Y)$

(d) $Var(X+Y)$

Solution:

(a) $E[X] = \sum_{k=1}^{5} k \frac{1}{5} = \frac{5\cdot 6}{2} \cdot \frac{1}{5} = 3.$

$E[Y] = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = \frac{1}{2}.$

(b) $E[X^2] = \sum_{k=1}^{5} k^2 \cdot \frac{1}{5} = \frac{5\cdot 6 \cdot 11}{6} \cdot \frac{1}{5} = 11$

Hence $Var(X) = E[X^2] - (E[X])^2 = 11 - 9 = 2$

$E[Y^2] = 0^2 \cdot \frac{1}{2} + 1^2 \cdot \frac{1}{2} = \frac{1}{2}.$

Hence $Var(Y) = E[Y^2] - (E[Y])^2 = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}.$

(c) $Cov(X,Y) = 1\cdot 0\, P(X=1, Y=0) + 1\cdot 1\, P(X=1, Y=1) +$
$+ 2\cdot 0\, P(X=2, Y=0) + 2\cdot 1\, P(X=2, Y=1) + 3\cdot 0\, P(X=3, Y=0) +$
$+ 3\cdot 1\, P(X=3, Y=1) + 4\cdot 0\, P(X=4, Y=0) + 4\cdot 1\, P(X=4, Y=1) +$
$+ 5\cdot 0\, P(X=5, Y=0) + 5\cdot 1\, P(X=5, Y=1) - 3 \cdot \frac{1}{2}.$

$= 1\cdot P(X=1, Y=1) + 2P(X=2, Y=1) + 3P(X=3, Y=1)$

$+ 4P(X=4, Y=1) + 5P(X=5, Y=1) - \frac{3}{2}$

$= 1\cdot P(X=1)P(Y=1) + 2P(X=2)P(Y=1) + 3P(X=3)P(Y=1)$

$+ 4P(X=4)P(Y=1) + 5P(X=5)P(Y=1)$

$= \Big(1\cdot P(X=5) + 2P(X=2) + 3P(X=3) + 4P(X=4) +$

$+ 5P(X=5)\Big)\Big(0\cdot P(X=0) + 1\cdot P(Y=1)\Big) - \frac{3}{2} =$

$= \left(\frac{5\cdot 6}{2}\cdot\frac{1}{5}\right)\frac{1}{2} - \frac{3}{2} = \frac{3}{2} - \frac{3}{2} = 0.$

(d) $Var(X+Y) = Var(X) + Var(Y) - 2Cov(X,Y)$

$= Var(X) + Var(Y) = 2 + \frac{1}{4} = \frac{9}{4}.$

Ex. Two fair dice are rolled. Compute the expected value and variance of the sum on the dice.

Solution: Let $X = \#$ on die 1 $Y = \#$ on die 2.

$E[X] = E[Y] = \sum_{k=1}^{6} k\frac{1}{6} = \frac{6\cdot 7}{2}\cdot\frac{1}{6} = \frac{7}{2}$

$E[X^2] = E[Y^2] = \sum_{k=1}^{6} k^2\frac{1}{6} = \frac{6\cdot 7\cdot 13}{6}\cdot\frac{1}{6} = \frac{91}{6}$

Hence $Var(X) = Var(Y) = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}.$

$E[X+Y] = \underbrace{E[X] + E[Y]}_{\text{Always.}} = \frac{7}{2} + \frac{7}{2} = 7$

$$\underbrace{Var(X+Y) = Var(X) + Var(Y)}_{\text{By independence of variables.}} = \frac{35}{12} + \frac{35}{12} = \frac{35}{6}$$

Ex. Each of n gentlemen pick one of n hats at random. What is the expected number of gentlemen that pick their own hat? What's the variance?

Solution: For $1 \leq k \leq n$ let $X_k = \begin{cases} 1 & \text{if } k^{th} \text{ man picks own hat} \\ 0 & \text{otherwise} \end{cases}$.

$E[X_k] = E[X_1] = \frac{1}{n}$.

$E[X_k^2] = E[X_1^2] = \frac{1}{n}$

$Var(X_k) = Var(X_1) = E[X_1^2] - (E[X_1])^2 = \frac{1}{n} - \left(\frac{1}{n}\right)^2 =$

$= \frac{1}{n}\left(1 - \frac{1}{n}\right)$

Expected number of men that pick own hat =

$= E[X_1 + \dots + X_n] = n\,E[X_1] = n \cdot \frac{1}{n} = 1$.

Notice that $E[X_1 X_2] = \frac{1}{n(n-1)}$.

Thus $Cov(X_1, X_2) = E[X_1X_2] - E[X_1]E[X_2] =$

$= \frac{1}{n(n-1)} - \left(\frac{1}{n}\right)^2 \neq 0$.

$$Var(X_1 + \dots + X_n) = \sum_{k=1}^{n} Var(X_k) + \sum_{k=1}^{n} \sum_{\substack{j=1 \\ j \neq k}}^{n} Cov(X_k, X_j)$$

$$= n\,Var(X_1) + 2\binom{n}{2} Cov(X_1, X_2)$$

$$= \left(1 - \frac{1}{n}\right) + n(n-1)\left(\frac{1}{n(n-1)} - \left(\frac{1}{n}\right)^2\right) = 1 - \frac{1}{n} + 1 - \frac{n-1}{n} = 1.$$