Session-15

Variance of sum of random variables, Expected value of product of independent random variable
And
Covariance & corelation

Outline

- Variance
- Variance of sum of random variables
- Expected value of product of independent random variable
- Covariance
- Corelation

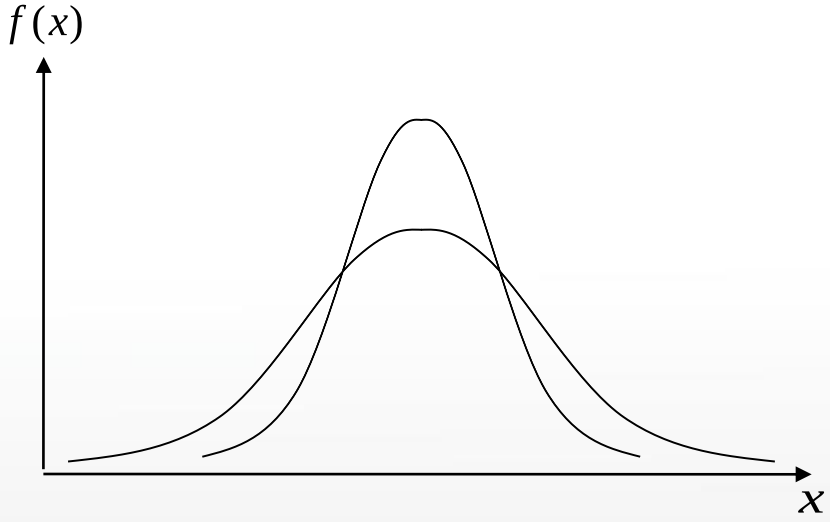
Variance

- The variance of a random variable X is a measure of how ***spread out*** it is. Are the values of X clustered tightly around their mean, or can we commonly observe values of X a long way from the mean value?
- The ***variance*** measures how far the values of X are from their mean, on average.
- If X has ***high variance***, we can observe values of X a long way from the mean.
- If X has ***low variance***, the values of X tend to be clustered tightly around the mean value.

Definition and Interpretation of Variance

$$\begin{aligned}\mathrm{Var}(X) &= E((X - E(X))^2) \\ &= E(X^2 - 2XE(X) + (E(X))^2) \\ &= E(X^2) - 2E(X)E(X) + (E(X))^2 \\ &= E(X^2) - (E(X))^2\end{aligned}$$

Two distribution with identical mean values but different variances



Expection of random variable

- **Discrete case:**

$$E(X) = \sum_{\text{all } x} x_i p(x_i)$$

- **Continuous case:**

$$E(X) = \int_{\text{all } x} x_i p(x_i)\,\text{dx}$$

Example - 1

- Let *X* takes the values 2, 3, and 4 with probabilities 0.1, 0.7, and 0.2. We can compute that *E[X]*= 3.1.

$$\begin{aligned}
\mathrm{Var}(X) = \mathrm{E}\left[(X-3.1)^2\right] &= 0.1 \cdot (2-3.1)^2 + 0.7 \cdot (3-3.1)^2 + 0.2 \cdot (4-3.1)^2 \\
&= 0.1 \cdot (-1.1)^2 + 0.7 \cdot (-0.1)^2 + 0.2 \cdot (0.9)^2 \\
&= 0.1 \cdot 1.21 + 0.7 \cdot 0.01 + 0.2 \cdot 0.81 \\
&= 0.121 + 0.007 + 0.162 \\
&= 0.29.
\end{aligned}$$

Using the rule is neater and somewhat faster: $\mathrm{Var}(X) = E[X^2] - E$

$$\begin{aligned}
\mathrm{Var}(X) = \mathrm{E}\left[X^2\right] - (3.1)^2 &= 0.1 \cdot 2^2 + 0.7 \cdot 3^2 + 0.2 \cdot 4^2 - 9.61 \\
&= 0.1 \cdot 4 + 0.7 \cdot 9 + 0.2 \cdot 16 - 9.61 \\
&= 0.4 + 6.3 + 3.2 - 9.61 \\
&= 0.29.
\end{aligned}$$

Example - 2

$$\mathrm{Var}(X) = E((X - E(X))^2) = \sum_i p_i (x_i - E(X))^2$$

$$= 0.3(50 - 230)^2 + 0.2(200 - 230)^2 + 0.5(350 - 230)^2$$

$$= 17,100 = \sigma^2$$

Standard deviation = $\sigma = \sqrt{17,100} = 130.77$

Sum of random variables

- Let X and Y be two independent integer-valued random variables, with distribution functions $m_1(x)$ and $m_2(x)$ respectively. Then the convolution of $m_1(x)$ and $m_2(x)$ is the distribution function $m_3 = m_1 * m_2$ given by

$$m_3(j) = \sum_k m_1(k) \cdot m_2(j-k) \ ,$$

for $j = \ldots, \ -2, \ -1, \ 0, \ 1, \ 2, \ \ldots$. The function $m_3(x)$ is the distribution function of the random variable $Z = X + Y$. $\square$

Now let $S_n = X_1 + X_2 + \cdots + X_n$ be the sum of n independent random variables of an independent trials process with common distribution function m defined on the integers. Then the distribution function of S_1 is m. We can write $S_n = S_{n-1} + X_n$.
since we know the distribution function of X_n is m, we can find the distribution function of S_n by induction.

Example

- A die is rolled twice. Let X 1 and X 2 be the outcomes, and let S2 = X1 + X2 be the sum of these outcomes. Then X1 and X2 have the common distribution function:

$$m = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{pmatrix}.$$

The distribution function of S_2 is then the convolution of this distribution with itself. Thus,

$$\begin{aligned}
P(S_2 = 2) &= m(1)m(1) \\
&= \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}\ , \\
P(S_2 = 3) &= m(1)m(2) + m(2)m(1) \\
&= \frac{1}{6} \cdot \frac{1}{6} + \frac{1}{6} \cdot \frac{1}{6} = \frac{2}{36}\ , \\
P(S_2 = 4) &= m(1)m(3) + m(2)m(2) + m(3)m(1) \\
&= \frac{1}{6} \cdot \frac{1}{6} + \frac{1}{6} \cdot \frac{1}{6} + \frac{1}{6} \cdot \frac{1}{6} = \frac{3}{36}\ .
\end{aligned}$$

Cont...

Continuing in this way we would find $P(S_2 = 5) = 4/36$, $P(S_2 = 6) = 5/36$, $P(S_2 = 7) = 6/36$, $P(S_2 = 8) = 5/36$, $P(S_2 = 9) = 4/36$, $P(S_2 = 10) = 3/36$, $P(S_2 = 11) = 2/36$, and $P(S_2 = 12) = 1/36$.

The distribution for S_3 would then be the convolution of the distribution for S_2 with the distribution for X_3. Thus

$$\begin{aligned}
P(S_3 = 3) &= P(S_2 = 2)P(X_3 = 1) \\
&= \frac{1}{36} \cdot \frac{1}{6} = \frac{1}{216}\ , \\
P(S_3 = 4) &= P(S_2 = 3)P(X_3 = 1) + P(S_2 = 2)P(X_3 = 2) \\
&= \frac{2}{36} \cdot \frac{1}{6} + \frac{1}{36} \cdot \frac{1}{6} = \frac{3}{216}\ ,
\end{aligned}$$

Variance of sum of random variables

If 'X' and 'Y' are two random variables
Then, variance of $X+Y$ is

$$Var(X+Y) = E\left((X+Y) - E(X+Y)\right)^2 \quad \text{—— (1)}$$

we know that $E(X+Y) = E(X) + E(Y)$ —(2)

So, we can rewrite (1) as below using (2)

$$Var(X+Y) = E\left((X - E(X)) + (Y - E(Y))\right)^2$$

$$= E\left((X - EX) + (Y - EY)\right)^2$$

Just written $E(X) = EX$
$E(Y) = EY$

Cont...

$$= E\left((X-EX)^2\right) + 2E(X-EX)(Y-EY) + E\left((Y-EY)^2\right)$$

$$\underbrace{E\left((X-EX)^2\right)}_{\text{variance of } X} + \underbrace{2E(X-EX)(Y-EY)}_{\text{Covariance } X \text{ and } Y} + \underbrace{E\left((Y-EY)^2\right)}_{\text{variance of } Y}$$

If X and Y are independent then

Covariance $(X, Y) = 0$

Now

$$Var(X+Y) = E\left((X-EX)^2\right) + E\left((Y-EY)^2\right)$$

$$= Var(X) + Var(Y)$$

Expected value of product of independent random variable

The Expectation/expected value of a Random Variable

- Mathmatical expection is a point around which all random variables are concentrated.
- Eample: When you measure the temprature of let it be 100 people then their temperature is different but they are close to normal temp which is mathmatical expection. Expection can be
  - Expectations of Discrete Random Variables
  - Expectations of Continuous Random Variables

Expectations of Discrete Random Variable

- The mean, expected value, or expectationof a random variable X is written as E(X) or μX.
- Let X be a discrete random variable with probability function $P(x_i)$ then expection is

$$E(X) = \sum_{\text{all } x} x_i p(x_i)$$

Independence of random variables

Random variables X and Y are **independent** if

$$P_{(X,Y)}(x,y) = P_X(x)P_Y(y)$$

for *all* values of x and y. This means, events $\{X = x\}$ and $\{Y = y\}$ are independent for all x and y; in other words, variables X and Y take their values independently of each other.

To show independence of X and Y , we have to check whether the joint pmf factors into the product of marginal pmfs for all pairs x and y.
To prove dependence, we only need to present one counterexample, a pair (x, y) with P (x, y) != $P_X(x)\ P_Y(y)$

Expection of product of two random variables

If X and Y are two random variables then

$$E(X+Y) = E(X) + E(Y) \quad \text{is True}$$

But

$$E(XY) = E(X) \cdot E(Y) \quad \text{is true only when X and Y are independent.}$$

<u>Proof.</u>

Let

$$X = \begin{Bmatrix} x_1, & x_2 & \cdots & x_M \\ p_1 & p_2 & \cdots & p_M \end{Bmatrix} \qquad Y = \begin{Bmatrix} y_1, & y_2 & \cdots & y_N \\ q_1, & q_2 & \cdots & q_N \end{Bmatrix}$$

$$E(XY) = E(X) \cdot E(Y) \quad (\text{is true only when X and Y are independent})$$

Two random variables are independent if

$$\boxed{P\{X = x_i \text{ AND } Y = y_j\} = r_{ij} = P(X = x_i) \cdot P(Y = y_j) = p_i q_j}$$

Cont...

L.H.S

$$E(XY) = (x_1 * y_1) r_{11} + (x_1 * y_2) r_{12} + \cdots + (x_1 * y_N) r_{1N}$$
$$+ (x_2 * y_1) r_{21} + (x_2 * y_2) r_{22} + \cdots + (x_2 * y_N) r_{2N}$$
$$\cdots$$
$$+ (x_M * y_1) r_{M1} + (x_M * y_2) r_{M2} + \cdots + (x_M * y_N) r_{MN}$$

R.H.S

$$E(X) \cdot E(Y) = (x_1 p_1 + x_2 p_2 + \cdots + x_M p_M)(y_1 q_1 + y_2 q_2 + \cdots + y_N q_N)$$
$$= (x_1 y_1) p_1 q_1 + (x_1 y_2) p_1 q_2 + \cdots + (x_1 y_N) p_1 q_N$$
$$+ \cdots$$
$$+ (x_M y_1) p_M q_1 + (x_M y_2) p_M q_2 + \cdots + (x_M y_N) p_M q_N$$

L.H.S is equal to RHS only when X and Y are independent. because r_{ij} is equal to $p_i q_j$ only when X an Y are independent.

Expected value of product of independent random variable(Proof other way)

Consider two independent random variables
X, Y.

Let's assume X and Y are independent

Then, $E(XY) = E(X) \cdot E(Y)$

Proof.

Suppose $X = \{x_1, \dots x_m\}$
$Y = \{y_1, \dots y_n\}$.

Distribution of X and Y

$$P(X = x_i) = P_i$$
$$P(Y = y_i) = q_j$$

If X and Y are independent then

$$(P(X = x_i) \text{ and } P(Y = y_j)) = P(X = x_i) \cdot P(Y = y_i)$$
$$= P_i \cdot q_j \quad (\text{Joint Probability})$$

Expected value of product of independent random variable

$$E(X \cdot Y) = \sum_{i=1}^{m} \sum_{j=1}^{n} P(X = x_i \cap Y = y_j) \cdot x_i y_j$$

$$= \sum_{i=1}^{m} \sum_{j=1}^{n} p_i p_j x_i y_j$$

$$= \sum_{i=1}^{m} \left(p_i x_i \sum_{j=1}^{n} q_j y_j \right)$$

$$= \left(\sum_{i=1}^{n} p_i x_i \right) \left(\sum_{i=1}^{n} q_j y_j \right)$$

$$= E(X) \cdot E(Y)$$