2.6 Variance

How far are values of X away from $E[X]$?

Definition 43 X RV, $\mu := E[X]$

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

is the variance of X.

Remark $Var(X)$ has better mathematical properties than $E[|X-\mu|]$

Note:
$$\begin{aligned} Var(X) &= E[(X-\mu)^2] \\ &= E[X^2 - 2\mu X + \mu^2] \\ &= E[X^2] - 2\mu E[X] + E[\mu^2] \\ &= E[X^2] - 2\mu^2 + \mu^2 \\ &= E[X^2] - E[X]^2 \end{aligned}$$

Example 44 X = number shown when rolling a die. We know $E[X] = \frac{7}{2}$.

$E[X^2] = \frac{91}{6}$ (Example 37)

$$\Rightarrow Var(X) = E[X^2] - E[X]^2 = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{91}{6} - \frac{49}{4} = \frac{182-147}{12} = \frac{35}{12}$$

Properties of $Var(\cdot)$. Let $E[X] = \mu$

$$Var(aX+b) = E[(aX+b-E[aX+b])^2]$$
$$= E[(aX+b-a\mu-b)^2]$$
$$= E[(aX-a\mu)^2]$$
$$= E[a^2(X-\mu)^2] = a^2 E[(X-\mu)^2]$$
$$= a^2 Var(X)$$

We call $\sigma := \sqrt{Var(X)}$ the <u>standard deviation</u> of X.

Note that then

$$Var(X) = \sigma^2$$

2.7 Covariance

We note

$$Var(X + X) = Var(2X) = 4\,Var(X)$$
$$\neq Var(X) + Var(X)$$

What about $Var(X+Y)$ if X, Y are indep.?

<u>Definition 45</u> X, Y RVs, $\mu_X = E[X]$, $\mu_Y = E[Y]$

Then

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

Note: This assumes a joint distribution of X and Y.

Property of Cov

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

Proposition 46 $Cov(X_1 + X_2, Y) = Cov(X_1, Y) + Cov(X_2, Y)$

$$
\begin{aligned}
Cov(X_1 + X_2, Y) &= E[(X_1 + X_2) \cdot Y] - E[X_1 + X_2]E[Y] \\
&= E[X_1 Y] + E[X_2 Y] - E[X_1]E[Y] - E[X_2]E[Y] \\
&= E[X_1 Y] - E[X_1]E[Y] + E[XY] - E[X_2]E[Y] \\
&= Cov(X_1 Y) + Cov(X_2 Y)
\end{aligned}
$$

Theorem 47

$$
Cov\left(\sum_i X_i, \sum_j Y_j\right) = \sum_i \sum_j Cov(X_i, Y_j)
$$

Consequence:

$$
\begin{aligned}
Var\left(\sum_i X_i\right) &= Cov\left(\sum_i X_i, \sum_i X_i\right) \\
&= \sum_i \sum_j Cov(X_i, Y_j) \\
&= \sum_i \left(\sum_{j \neq i} Cov(X_i, X_j) + Cov(X_i, X_i) \right) \\
&= \sum_i \sum_{j \neq i} Cov(X_i, X_j) + \sum_i Cov(X_i, X_i) \\
&= \sum_i Var(X_i, X_i) + \sum_i \sum_{j \neq i} Cov(X_i, X_j)
\end{aligned}
$$

For $n = 2$:

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

<u>Theorem</u> <u>48</u>

$$X, Y \text{ independent} \Rightarrow \mathrm{Cov}(X, Y) = 0$$

Proof: Show $E[XY] = E[X]E[Y]$. For continuous case let f, g be the densities of X, Y, resp.

$$\begin{aligned} E[XY] &= \int_{\mathbb{R}} \int_{\mathbb{R}} x \cdot y \, f(x) g(y) \, dx \, dy \\ &= \int_{\mathbb{R}} y \, g(y) \left(\int_{\mathbb{R}} x \, f(x) \, dx \right) dy \\ &= \int_{\mathbb{R}} y \, g(y) \, E[X] \, dy \\ &= E[X] \, E[Y] \end{aligned}$$

For discr. case by distributive law.

<u>Example</u> 49 Variance of sum of 10 independent rolls of a die.

Let X_i be i-th roll.

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) \underbrace{=}_{\text{independence}} \sum_{i=1}^{n} \mathrm{Var}(X_i) =$$

$$= \sum_{i=1}^{n} \frac{35}{12} = 10 \cdot \frac{35}{12} = \frac{175}{6}$$

Meaning of covariance $Cov(X, Y)$

> 0: X, Y grow together

< 0: X, Y grow in opposite direction

Normalize: correlation!

$$Corr(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X)\ Var(Y)}}$$

This value is always in $[-1, 1]$