Variance and Covariance

July, 23, 2020

1) Applications of Linearity of Expectation

2) Variance

3) Covariance and Correlation

Last time:

Distributions:

$$X \sim \text{Bernoulli}(p), \quad Pr[X] = \begin{cases} p & \text{if } X=1 \\ 1-p & \text{if } X=0 \end{cases}$$

$$X \sim \text{Bin}(n,p), \quad Pr[X=i] = \binom{n}{i} p^i (1-p)^{n-i}$$

Linearity of Expectation:

$$E[c_1 X_1 + \cdots + c_n X_n] = c_1 E[X_1] + \cdots + c_n E[X_n]$$

1) Applications of Linearity of Expectation:

Example: Hand back assignments to n students at random.

Expected number of students who get back their hw?

X_n: the number of students who get back their hw.

Use indicator r.v. $\boxed{I_i} = \begin{cases} 1 & i^{th} \text{ student get back their own assignment} \quad 1/n \\ 0 & \text{others} \end{cases}$

$Pr[I_i = 1] = \frac{1}{n}$

$X_n = I_1 + \cdots + I_n$

$E[X_n] = E[I_1 + \cdots + I_n] = E[I_1] + \cdots + E[I_n]$

linearity

$= \sum_{i=1}^{n} E[I_i] = \sum_{i=1}^{n} \frac{1}{n} = n \times \frac{1}{n}$

$= 1$

$E[I_i] = \sum_{a \in \{0,1\}} a \, Pr[I_i = a] = 1 \times Pr[I_i = 1] + \cancel{0 \times Pr[I_i = 0]}$

$= \frac{1}{n}$

The expected number of students who get their own homeworks in a class size n is 1.

Example: Throw m balls in n bins.

The expected number of empty bins.

X: number of empty bins

Help Indicators, help!

$$I_i = \begin{cases} 1 & i^{th} \text{ bin is empty} \quad \underbrace{\left(\frac{n-1}{n}\right)^m}_{Pr} \Rightarrow E[I_i] = Pr[I_i=1] \\ 0 & \text{otherwis} \end{cases}$$

$$X = \sum_{i=1}^{n} I_i$$

linearity

$$E[X] = E\left[\sum_{i=1}^{n} I_i\right] = \sum_{i=1}^{n} \underline{E[I_i]} = \sum_{i=1}^{n} \left(\frac{n-1}{n}\right)^m = \left(\frac{n-1}{n}\right)^m \sum_{i=1}^{n} 1$$

$$E[I_i] = \sum_{a\in\{0,1\}} a \times Pr[I_i=a] = 1\times Pr[I_i=1] + 0\times Pr[I_i=0]$$

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

$$E[X] = n\left(\frac{n-1}{n}\right)^m$$

For $n=m$ and $n\to\infty$?

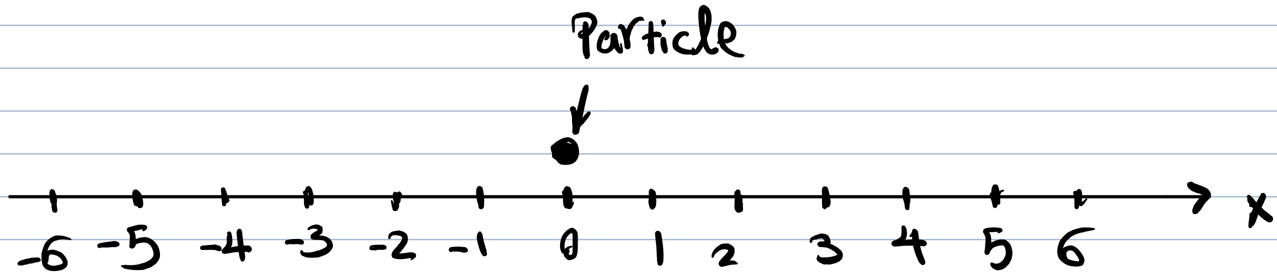
$$E[X] = n\left(\frac{n-1}{n}\right)^n = n\underbrace{\left(1-\frac{1}{n}\right)^n} = \frac{n}{e} \simeq 0.37n$$

$$\lim_{n\to\infty}\left(1-\frac{1}{n}\right)^n = \frac{1}{e}$$

~37% of bins will be empty.

2) Variance:

Examle: Random Walk



- The Particle moves in one dimension.
- At each time step, Particle moves to the right or left with equal Probability.

S_n = The Position of the Particle after n moves

X_i = the i^{th} movement $\Rightarrow X = \begin{cases} +1 & Pr[1] = \frac{1}{2} \\ -1 & Pr[-1] = \frac{1}{2} \end{cases}$

$$\underline{S_n} = 0 + X_1 + X_2 + \cdots + X_n = \underline{\sum_{i=1}^{n} X_i}$$

$$E[S_n] = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i] = 0 \rightarrow \text{not useful}$$

$$E[X_i] = \sum_{a \in \{1,-1\}} a \times Pr[X_i] = 1 \times Pr[X_i = 1] + (-1) \times Pr[X_i = -1]$$

$$= 1 \times \frac{1}{2} + (-1) \times \frac{1}{2} = 0$$

what is the expected distance from 0? $\underline{|S_n|}$

we work $\boxed{\underline{S_n^2}}$.

Hint:

$$(X_1 + \cdots + X_n)^2 = \sum_{i=1}^{n} X_i^2 + 2\sum_{i<j} X_i X_j$$

$$E[S_n^2] = E[(X_1 + \cdots + X_n)^2]$$

$$= E\left[\sum_{i=1}^{n} X_i^2 + 2\sum_{i<j} X_i X_j\right]$$

$$= \sum_{i=1}^{n} E[X_i^2] + 2\sum_{i<j} \cancel{E[X_i X_j]} = \sum_{i=1}^{n} 1 = n$$

$$E[X_i^2] = \sum_{a\in\{\pm 1\}} a^2 \times Pr[X_i = a] = (1)^2 Pr[X_i = 1] + (-1)^2 Pr[X_i = -1]$$

$$= 1 \times \frac{1}{2} + 1 \times \frac{1}{2} = 1$$

$$E[X_i X_j] = \sum_{a,b\in\{\pm 1\}} ab \times Pr[X_i = a, X_j = b]$$

$\to Pr[X_i = a] \times Pr[X_j = b]$

$$= (1)(1)\, Pr[X_i = 1]\, Pr[X_j = 1]$$
$$+ (1)(-1)\, Pr[X_i = 1]\, Pr[X_j = -1]$$
$$+ (-1)(1)\, Pr[X_i = -1]\, Pr[X_j = 1]$$
$$+ (-1)(-1)\, Pr[X_i = -1]\, Pr[X_j = -1]$$
$$= 1 \times \frac{1}{4} - 1 \times \frac{1}{4} - 1 \times \frac{1}{4} + 1 \times \frac{1}{4} = 0$$

So $E[S_n^2] = n$ ↓ Variance of S_n

Definition: For a r.v. X with expectation $E[X] = \mu$ the variance of X is defined to be

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

what does variance measure? The deviation from mean value.

The square root $\sigma(X) := \sqrt{Var(X)}$ is the standard deviation of variable X.

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

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

linearity

$$= E[X^2 - 2X\mu + \mu^2]$$

$$= E[X^2] - 2\mu E[X] + E[\mu^2]$$

$$= E[X^2] - 2\mu\mu + \mu^2$$

$$= E[X^2] - \mu^2.$$

Some Properties of Variance:

$Var(X+c) = Var(X)$ → exercise

$Var(cX) = c^2 Var(X)$ → exercise.

More Examples:

- Uniform distribution

X: a uniform random variable on the set $\{1, \dots, n\}$

$X \sim \text{Uniform}\{1, \dots, n\}$, $\Pr[X=i] = \frac{1}{n}$, $\forall i \in \{1, \dots, n\}$

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

$$\begin{cases} E[X] = \sum_{a \in \{1,\dots,n\}} a \times \Pr[X=a] = \sum_{a=1}^{n} \frac{a}{n} = \frac{1}{n} \sum_{a=1}^{n} a = \frac{1}{n} \frac{n(n+1)}{2} = \frac{n+1}{2}. \\ E[X^2] = \sum_{a \in \{1,\dots,n\}} a^2 \times \Pr[X=a] = \sum_{a=1}^{n} \frac{a^2}{n} \end{cases}$$

$$= \frac{1}{n} \underbrace{\sum_{a=1}^{n} a^2} = \frac{1}{\cancel{n}} \frac{\cancel{n}(n+1)(2n+1)}{6}$$

$$= \frac{(n+1)(2n+1)}{6}$$

$$Var(X) = \frac{n^2-1}{12}.$$

Example: Hand back assignments to n students at random.

X_n: the number of students who get back their hw.

use indicators! $I_i = \begin{cases} 1 & \text{student } i \text{ gets their hw} \\ 0 & \text{otherwise} \end{cases}$ Pr $\frac{1}{n}$

- $X_n = I_1 + I_2 + \dots + I_n = \sum_{i=1}^{n} I_i$, exercise

- $E[X_n] = 1$

- $Var(X) = ?$

Hint: $(I_1 + \dots + I_n)^2 = \sum_{i=1}^{n} I_i^2 + 2\sum_{i<j} I_i I_j$

$Var(X) = E[X^2] - E[X]^2 = E[X^2] - 1$

$E[X^2] = E[(\sum_{i=1}^{n} I_i)^2] = \sum_{i=1}^{n} E[I_i^2] + 2\sum_{i<j} E[I_i I_j]$

1) $E[I_i^2] = 1 \times Pr[I_i = 1] + 0 \times Pr[I_i = 0] = \frac{1}{n}$

2) $E[\underbrace{I_i I_j}_{=1}] = 1 \times Pr[I_i = 1] \times Pr[I_j = 1] = \frac{1}{n(n-1)}$

$E[X^2] = \sum_{i=1}^{n} \frac{1}{n} + 2\sum_{i<j} \frac{1}{n(n-1)}$ Choose a Pair of indecis

$= n \times \frac{1}{n} + 2\binom{n}{2}\frac{1}{n(n-1)} = 2$

$\Rightarrow \boxed{Var(X) = 2 - 1 = 1}$

Independent Random Variables:

Theorem: For independent random variables X, Y we have

$$E[XY] = E[X]\,E[Y]$$

Proof:

independent X and Y

$$E[XY] = \sum_{a \in A} \sum_{b \in B} a\, b \times \Pr[X=a, Y=b]$$

$$= \sum_a \sum_b ab \Pr[X=a] \Pr[Y=b]$$

$$= \Big(\sum_a a \Pr[X=a]\Big) \times \Big(\sum_b b \Pr[X=b]\Big)$$

$$= E[X] \times E[Y]. \qquad \square$$

Theorem: For independent random variables X, Y we have $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$

Proof:

$$\mathrm{Var}(X+Y) = E[(X+Y)^2] - (\overset{E[X]+E[Y]}{E[X+Y]})^2$$

$$= E[X^2 + 2XY + Y^2] - \left(E[X]^2 + 2E[X]E[Y] + E[Y]^2\right)$$

linearity

$$= E[X^2] + 2E[XY] + E[Y^2] - E[X]^2 - 2E[X]E[Y] - E[Y]^2$$

$$= (E[X^2] - E[X]^2) + (E[Y^2] - E[Y]^2) + 2(E[XY] - E[X]E[Y])$$

$= Var(X) + Var(Y) + 2(\underbrace{E[XY] - E[X]E[Y]}_{=0})$

If x and y are independent $\longrightarrow$

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

If x and y are independent, then

$$Var(X-Y) = ?$$

remember $Var(cX) = c^2 Var(X)$

$$Var(X + (-Y)) = Var(X) + Var(-Y)$$
$$= Var(X) + (-1)^2 Var(Y)$$
$$= Var(X) + Var(Y)$$

Example: X_n = is the number of heads in n tosses of a biased coin with heads probability P.

$X_n \sim ?$ $Bin(n,p)$

i^{th} toss

$$I_i = \begin{cases} 1 & \text{heads} \quad p \\ 0 & \text{tails} \quad 1-p \end{cases}$$

$$X_n = \sum_{i=1}^{n} I_i$$

$$E[X_n] = E\left[\sum_{i=1}^{n} I_i\right] = \sum_{i=1}^{n} E[I_i] = \sum_{i=1}^{n} p = np$$

$$Var(X_n) = Var\left(\sum_{i=1}^{n} I_i\right) = \sum_{i=1}^{n} Var(I_i) = \sum_{i=1}^{n} P(1-P) = nP(1-P)$$

$$Var(I_i) = E[I_i^2] - \underbrace{E[I_i]}^2$$

$$= 1^2 \times Pr[I_i = 1] + 0^2 \times Pr[I_i = 0] - P^2 = P - P^2 = \underline{P(1-P)}$$

3) Covariance and Correlation:

Definition: The covariance of random variables X and Y is

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

Remarks:

1. If X, Y are independent, $Cov(X,Y) = 0$

- Can we say

if $Cov(X,Y) = 0 \Rightarrow X, Y$ are independent? No

The converse of 1 is not true.

2. $Cov(X,X) = E[X \cdot X] - E[X]E[X]$

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

3. Covariance is bilinear

For any collection of random variables $\{X_1, \dots, X_n\}$, $\{Y_1, \dots, Y_m\}$ and fixed constants $\{a_1, \dots, a_n\}$, $\{b_1, \dots, b_m\}$ then

$$\mathrm{Cov}\left(\sum_{i=1}^{n} a_i X_i, \underbrace{\sum_{j=1}^{m} b_j Y_j}_{Y}\right) = \sum_{i=1}^{n} \sum_{j=1}^{m} a_i b_j \,\mathrm{Cov}(X_i, Y_j)$$

About $\mathrm{Cov}(X, Y)$

1. The sign of X, Y determines how X and Y are related
2. The magnitude is hard to interpret.

Definition: Correlation.

Suppose X and Y are random variables with $\sigma(X) > 0$ and $\sigma(Y) > 0$. The correlation of X and Y is

$$\mathrm{Corr}(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sigma(X)\,\sigma(Y)}.$$

Theorem:

$$-1 \leq corr(X,Y) \leq 1$$

$X = aY + b$ if $a>0 \Rightarrow corr(X,Y) = 1$
if $a<0 \Rightarrow corr(X,Y) = -1$

- $corr(X, X) = 1$

- $corr(X, -X) = -1$

- $corr(X,Y) = \dfrac{cov(X,Y)}{\sigma(X)\sigma(Y)} = \dfrac{E[XY] - E[X]E[Y]}{\sigma(X)\sigma(Y)}$
  $\boxed{Y = aX+b}$

$$\sigma(Y) = \sigma(aX+b) = \sqrt{Var(aX+b)} = \sqrt{a^2 Var(X)} = |a|\sqrt{Var(X)} = |a|\sigma(X)$$

$$= \frac{E[X(aX+b)] - E[X]E[aX+b]}{\sigma(X)\,|a|\sigma(X)}$$

$$= \frac{aE[X^2] + \cancel{bE[X]} - aE[X]^2 - \cancel{bE[X]}}{|a|\,Var(X)}$$

$$= \frac{a(E[X^2] - E[X]^2)}{|a|\,Var(X)} = \frac{a}{|a|}\frac{Var(X)}{Var(X)}$$

$$= \frac{a}{|a|} = \begin{cases} 1 & a>0 \\ -1 & a<0. \end{cases}$$