CS70: Jean Walrand: Lecture 28.

Independence and Variance

1. Expected value of nonnegative integer-valued RV
2. Waiting for a bus
3. A note about expectation
4. Independent random variables.
5. Variance.
6. Variance Calculations.
7. Variance Properties.

Summary of Lecture 27

Expectation; Conditional Expectation; B(n, p); G(p)

Expectation: $E[X] = \sum_a a \times Pr[X = a] = \sum_\omega X(\omega) Pr[\omega]$.

Linearity: $E[a_1 X_1 + \cdots + a_n X_n] = a_1 E[X_1] + \cdots + a_n E[X_n]$.

Binomial: $Pr[X = i] = \binom{n}{k} p^i (1-p)^{(n-i)}; E(X) = pn$.

Geometric: $Pr[X = i] = (1-p)^{(i-1)} p; E(X) = \frac{1}{p}$; memoryless.

Condition Expectation: $E[X|A]$. Linear and

$$E[X] = E[X|A]Pr[A] + E[X|\bar{A}]Pr[\bar{A}].$$

Expected Value of Nonnegative Integer RV

Theorem: For a r.v. X that takes values in $\{0, 1, 2, \ldots\}$, one has

$$E[X] = \sum_{i=1}^{\infty} Pr[X \geq i].$$

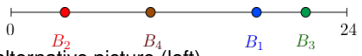
Proof: One has

$$\begin{aligned} E[X] &= \sum_{i=1}^{\infty} i \times Pr[X = i] \\ &= \sum_{i=1}^{\infty} i\{Pr[X \geq i] - Pr[X \geq i+1]\} \\ &= \sum_{i=1}^{\infty} \{i \times Pr[X \geq i] - i \times Pr[X \geq i+1]\} \\ &= \sum_{i=1}^{\infty} \{i \times Pr[X \geq i] - (i-1) \times Pr[X \geq i]\} \\ &= \sum_{i=1}^{\infty} Pr[X \geq i]. \end{aligned}$$

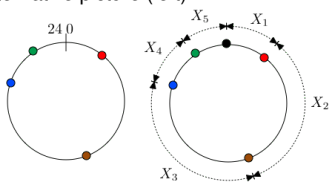
□

Riding the bus.

n buses arrive uniformly at random throughout a 24 hour day. What is the time between buses? How long do you expect to wait for a bus? Here are typical arrival times, independent and uniform in $[0, 24]$.

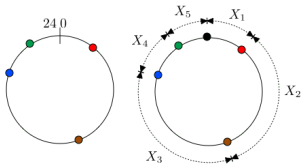


Here is an alternative picture (left)



The n gaps between busses add up to 24. Thus n times the expected gap is 24 and the expected gap between busses is $24/n$.

Riding the bus.



How about the waiting time until the next bus?

Add the black dot uniformly at random and pretend that it represents your arrival time. This is legitimate, because given the black dot, the other dots are uniform at random. Then,

$$24 = E[X_1 + \cdots + X_5] = 5E[X_1], \text{ by linearity and symmetry.}$$

Hence, $E[X_1] = E[X_m] = 24/5 = 24/(n+1)$ for n busses. Note the paradox: the average time between the previous and next bus is $48/(n+1) > 24/n$.

A note about expectation

Recall that

$$E[X] = \sum_x x \; Pr[X = x] = \sum_\omega X(\omega) Pr[\omega]. \tag{1}$$

Theorem
(a) $E[g(X))] = \sum_x g(x) Pr[X = x]$.
(b) $E[g(X, Y, Z)] = \sum_{x,y,z} g(x, y, z) Pr[X = x, Y = y, Z = z]$.

Proof:
(a) One has

$$\begin{aligned} E[g(X)] &= \sum_\omega g(X(\omega)) Pr[\omega], \text{ by (1) applied to } Y(\omega) = g(X(\omega)) \\ &= \sum_x \sum_{\omega \in X^{-1}(x)} g(X(\omega)) Pr[\omega] \\ &= \sum_x \sum_{\omega \in X^{-1}(x)} g(x) Pr[\omega] = \sum_x g(x) Pr[X = x]. \end{aligned}$$

(b) is similar. □

Independent Random Variables.

Definition: Independence

The random variables X and Y are **independent** if and only if

$$Pr[Y=b|X=a]=Pr[Y=b], \text{ for all } a \text{ and } b.$$

Fact:

X, Y are independent if and only if

$$Pr[X=a, Y=b]=Pr[X=a]Pr[Y=b], \text{ for all } a \text{ and } b.$$

Indeed:

$$Pr[Y=b|X=a]=\frac{Pr[X=a, Y=b]}{Pr[X=a]}=Pr[Y=b]$$
$$\Leftrightarrow Pr[X=a, Y=b]=Pr[X=a]Pr[Y=b].$$

□

Independence: Examples

Example 1
Roll two dices. X, Y = number of pips on the two dice. X, Y are independent.

Indeed: $Pr[X=a, Y=b]=\frac{1}{36}, Pr[X=a]=Pr[Y=b]=\frac{1}{6}$.

Example 2
Roll two dices. X = total number of pips, Y = number of pips on die 1 minus number on die 2. X and Y are not independent.

Indeed: $Pr[X=12, Y=1]=0 \neq Pr[X=12]Pr[Y=1]>0$.

Example 3
Flip a fair coin five times, X = number of Hs in first three flips, Y = number of Hs in last two flips. X and Y are independent.

Indeed:

$$Pr[X=a, Y=b]=\binom{3}{a}\binom{2}{b}2^{-5}=\binom{3}{a}2^{-3}\times\binom{2}{b}2^{-2}=Pr[X=a]Pr[Y=b].$$

A useful observation about independence

Theorem

X and Y are independent if and only if

$$Pr[X\in A, Y\in B]=Pr[X\in A]Pr[Y\in B] \text{ for all } A, B\subset\Re.$$

Proof:
If ($\Leftarrow$): Choose $A=\{a\}$ and $B=\{b\}$.

This shows that $Pr[X=a, Y=b]=Pr[X=a]Pr[Y=b]$.

Only if ($\Rightarrow$):

$$\begin{aligned}
&Pr[X\in A, Y\in B]\\
&=\sum_{a\in A}\sum_{b\in B}Pr[X=a, Y=b]=\sum_{a\in A}\sum_{b\in B}Pr[X=a]Pr[Y=b]\\
&=\sum_{a\in A}[\sum_{b\in B}Pr[X=a]Pr[Y=b]]=\sum_{a\in A}Pr[X=a][\sum_{b\in B}Pr[Y=b]]\\
&=\sum_{a\in A}Pr[X=a]Pr[Y\in B]=Pr[X\in A]Pr[Y\in B].
\end{aligned}$$

□

Functions of Independent random Variables

Theorem Functions of independent RVs are independent
Let X, Y be independent RV. Then

$$f(X) \text{ and } g(Y) \text{ are independent, for all } f(\cdot), g(\cdot).$$

Proof:
Recall the definition of inverse image:

$$h(z)\in C \Leftrightarrow z\in h^{-1}(C):=\{z \mid h(z)\in C\}. \qquad (2)$$

Now,

$$\begin{aligned}
&Pr[f(X)\in A, g(Y)\in B]\\
&=Pr[X\in f^{-1}(A), Y\in g^{-1}(B)], \text{ by (2)}\\
&=Pr[X\in f^{-1}(A)]Pr[Y\in g^{-1}(B)], \text{ since } X, Y \text{ ind.}\\
&=Pr[f(X)\in A]Pr[g(Y)\in B], \text{ by (2).}
\end{aligned}$$

□

Mean of product of independent RV

Theorem
Let X, Y be independent RVs. Then

$$E[XY]=E[X]E[Y].$$

Proof:
Recall that $E[g(X,Y)]=\sum_{x,y}g(x,y)Pr[X=x, Y=y]$. Hence,

$$\begin{aligned}
E[XY] &= \sum_{x,y}xyPr[X=x, Y=y]=\sum_{x,y}xyPr[X=x]Pr[Y=y], \text{ by ind.}\\
&= \sum_{x}[\sum_{y}xyPr[X=x]Pr[Y=y]]=\sum_{x}[xPr[X=x](\sum_{y}yPr[Y=y])]\\
&= \sum_{x}[xPr[X=x]E[Y]]=E[X]E[Y].
\end{aligned}$$

□

Examples

(1) Assume that X, Y, Z are mutually independent, with $E[X]=E[Y]=E[Z]=0$ and $E[X^2]=E[Y^2]=E[Z^2]=1$.

Then

$$\begin{aligned}
E[(X+2Y+3Z)^2]&=E[X^2+4Y^2+9Z^2+4XY+12YZ+6XZ]\\
&=1+4+9+4\times 0+12\times 0+6\times 0\\
&=14.
\end{aligned}$$

(2) Let X, Y be independent and $U[1,2,\ldots n]$. Then

$$\begin{aligned}
E[(X-Y)^2] &= E[X^2+Y^2-2XY]=2E[X^2]-2E[X]^2\\
&= \frac{1+3n+2n^2}{3}-\frac{(n+1)^2}{2}.
\end{aligned}$$

Mutually Independent Random Variables

Definition

(a) X, Y, Z are mutually independent if

$Pr[X=x, Y=y, Z=z] = Pr[X=x]Pr[Y=y]Pr[Z=z]$, for all x, y, z.

(b) $\{X_j, j \in J\}$ are mutually dependent if

$Pr[X_k = x_k, k \in K] = \Pi_{k \in K} Pr[X_k = x_k]$, for all finite $K \subset J$ and all x_k.

Note: If A, B, C are pairwise independent but not mutually independent, then $X = 1_A, Y = 1_B, Z = 1_C$ are pairwise independent but not mutually independent.

Functions of mutually independent RVs

If X, Y, Z are pairwise independent, but not mutually independent, it may be that

$$f(X) \text{ and } g(Y,Z) \text{ are not independent.}$$

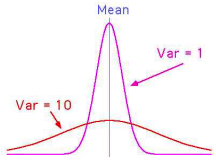
Example: Flip two fair coins,
$X = 1\{\text{coin 1 is } H\}, Y = 1\{\text{coin 2 is } H\}, Z = X \oplus Y$. Then, X, Y, Z are pairwise independent. Let $g(Y,Z) = Y \oplus Z$. Then $g(Y,Z) = X$ is not independent of X.

Theorem
Let $\{X_j, j \in J\}$ be mutually independent RVs. Then

$$f_1(X_j, j \in K_1), f_2(X_j, j \in K_2), \ldots \text{ are mutually independent}$$

for all pairwise disjoint K_m and functions $f_m(\cdot)$.
Proof:
Try it, you will like it. □

Variance



The variance measures the deviation from the mean value.

Definition: The variance of X is

$$\sigma^2(X) := var[X] = E[(X - E[X])^2].$$

$\sigma(X)$ is called the standard deviation of X.

Variance and Standard Deviation

Fact:

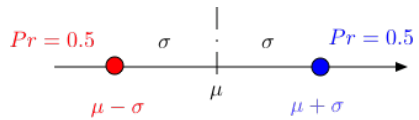
$$var[X] = E[X^2] - E[X]^2.$$

Indeed:

$$\begin{aligned} var(X) &= E[(X - E[X])^2] \\ &= E[X^2 - 2XE[X] + E[X]^2) \\ &= E[X^2] - 2E[X]E[X] + E[X]^2, \text{ by linearity} \\ &= E[X^2] - E[X]^2. \end{aligned}$$

A simple example

This example illustrates the term 'standard deviation.'



Consider the random variable X such that

$$X = \begin{cases} \mu - \sigma, & \text{w.p. } 1/2 \\ \mu + \sigma, & \text{w.p. } 1/2. \end{cases}$$

Then, $E[X] = \mu$ and $(X - E[X])^2 = \sigma^2$. Hence,

$$var(X) = \sigma^2 \text{ and } \sigma(X) = \sigma.$$

Example

Consider X with

$$X = \begin{cases} -1, & \text{w. p. } 0.99 \\ 99, & \text{w. p. } 0.01. \end{cases}$$

Then

$$\begin{aligned} E[X] &= -1 \times 0.99 + 99 \times 0.01 = 0. \\ E[X^2] &= 1 \times 0.99 + (99)^2 \times 0.01 \approx 100. \\ Var(X) &\approx 100 \implies \sigma(X) \approx 10. \end{aligned}$$

Also,

$$E(|X|) = 1 \times 0.99 + 99 \times 0.01 = 1.98.$$

Thus, $\sigma(X) \neq E[|X - E[X]|]$!

Exercise: How big can you make $\frac{\sigma(X)}{E[|X - E[X]|]}$?

Uniform

Assume that $Pr[X=i]=1/n$ for $i \in \{1,\ldots,n\}$. Then

$$\begin{aligned} E[X] &= \sum_{i=1}^{n} i \times Pr[X=i] = \frac{1}{n}\sum_{i=1}^{n} i \\ &= \frac{1}{n}\frac{n(n+1)}{2} = \frac{n+1}{2}. \end{aligned}$$

Also,

$$\begin{aligned} E[X^2] &= \sum_{i=1}^{n} i^2 Pr[X=i] = \frac{1}{n}\sum_{i=1}^{n} i^2 \\ &= \frac{1+3n+2n^2}{6}, \text{ as you can verify.} \end{aligned}$$

This gives

$$var(X) = \frac{1+3n+2n^2}{6} - \frac{(n+1)^2}{4} = \frac{n^2-1}{12}.$$

Fixed points.

Number of fixed points in a random permutation of n items.
"Number of student that get homework back."

$X = X_1 + X_2 \cdots + X_n$

where X_i is indicator variable for ith student getting hw back.

$$\begin{aligned} E(X^2) &= \sum_i E(X_i^2) + \sum_{i \neq j} E(X_i X_j). \\ &= n \times \frac{1}{n} + (n)(n-1) \times \frac{1}{n(n-1)} \\ &= 1+1=2. \end{aligned}$$

$E(X_i^2) = 1 \times Pr[X_i = 1] + 0 \times Pr[X_i = 0]$
$= \frac{1}{n}$
$E(X_i X_j) = 1 \times Pr[X_i = 1 \cap X_j = 1] + 0 \times Pr[\text{"anything else"}]$
$= \frac{1 \times 1 \times (n-2)!}{n!} = \frac{1}{n(n-1)}$

$Var(X) = E(X^2) - (E(X))^2 = 2 - 1 = 1.$

Variance: binomial.

$$\begin{aligned} E[X^2] &= \sum_{i=0}^{n} i^2 \binom{n}{i} p^i (1-p)^{n-i}. \\ &= \text{Really???!!\#\#...} \end{aligned}$$

Too hard!

Ok.. fine.
Let's do something else.
Maybe not much easier...but there is a payoff.

Properties of variance.

1. $Var(cX) = c^2 Var(X)$, where c is a constant.
   Scales by c^2.
2. $Var(X+c) = Var(X)$, where c is a constant.
   Shifts center.

Proof:

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

□

Variance of sum of independent random variables

Theorem:
If X and Y are independent, then

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

Proof:
Since shifting the random variables does not change their variance, let us subtract their means.

That is, we assume that $E(X)=0$ and $E(Y)=0$.

Then, by independence,

$$E(XY) = E(X)E(Y) = 0.$$

Hence,

$$\begin{aligned} var(X+Y) &= E((X+Y)^2) = E(X^2 + 2XY + Y^2) \\ &= E(X^2) + 2E(XY) + E(Y^2) = E(X^2) + E(Y^2) \\ &= var(X) + var(Y). \end{aligned}$$

Variance of Binomial Distribution.

Flip coin with heads probability p.
X- how many heads?

$$X_i = \begin{cases} 1 & \text{if } i\text{th flip is heads} \\ 0 & \text{otherwise} \end{cases}$$

$E(X_i^2) = 1^2 \times p + 0^2 \times (1-p) = p.$
$Var(X_i) = p - (E(X))^2 = p - p^2 = p(1-p).$

$p = 0 \implies Var(X_i) = 0$
$p = 1 \implies Var(X_i) = 0$

$X = X_1 + X_2 + \ldots X_n.$

X_i and X_j are independent: $Pr[X_i = 1 | X_j = 1] = Pr[X_i = 1].$

$$Var(X) = Var(X_1 + \cdots X_n) = np(1-p).$$

Summary

Independence and Variance

- $E[g(X,Y,Z)] = \sum_{x,y,z} g(x,y,z) Pr[X = x, Y = y, Z = z]$
- X, Y independent
  $\Leftrightarrow Pr[X \in A, Y \in B] = Pr[X \in A] Pr[Y \in B]$
- Then, $f(X), g(Y)$ are independent
- Also, $E[XY] = E[X]E[Y]$ and $var[X+Y] = var[X] + var[Y]$.

Distributions:

- $U[1,\ldots,n] : E[X] = (n+1)/2; var[X] = (n^2-1)/12$
- $B(n,p) : E[X] = np; var[X] = np(1-p)$.