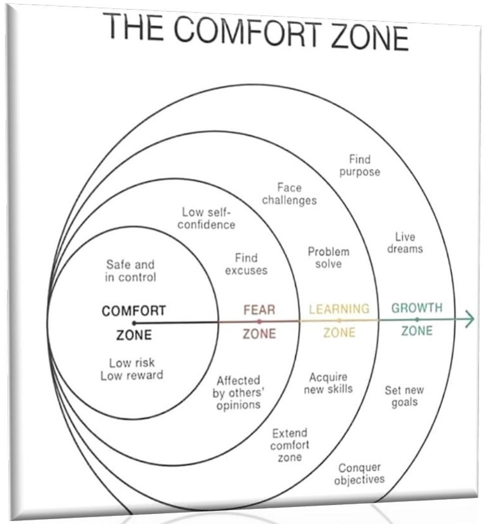
combra.cs.rutgers.edu



206 Discrete Structures II

Jack Klawitter & Stefan Obradovic

Computational Brain Lab

Computer Science | Rutgers University | NJ, USA

Last Lecture | Random Variables, Expectations | Wednesday December 12th 2023

This is the... Last Lecture

Part 1: Counting

- Proofs/induction
- Sum Rule
- Partition Method
- Difference Method
- Bijection Method
- Product Rule
- Generalized product rule
- Pigeonhole Principle
- Inclusion/Exclusion
- Combinatorial proofs, binomial coefficients

Part 2: Probability

- Sample space
- Axioms of probability
- Conditional probability
- Independence
- Bayes rule
- Random Variables
- Expectation and Linearity
- Variance

Random Variables - In a few sentences…

- Random Variable

  - Given experiment and sample space Ω,

  a random variable associates a **real number** to a single, more, or every outcome in Ω.

  - Note that, as a function, a random variable may represent any combination of events in Ω that we can possibly define.

Last Time...

- We understood the notion of expectation in random variables

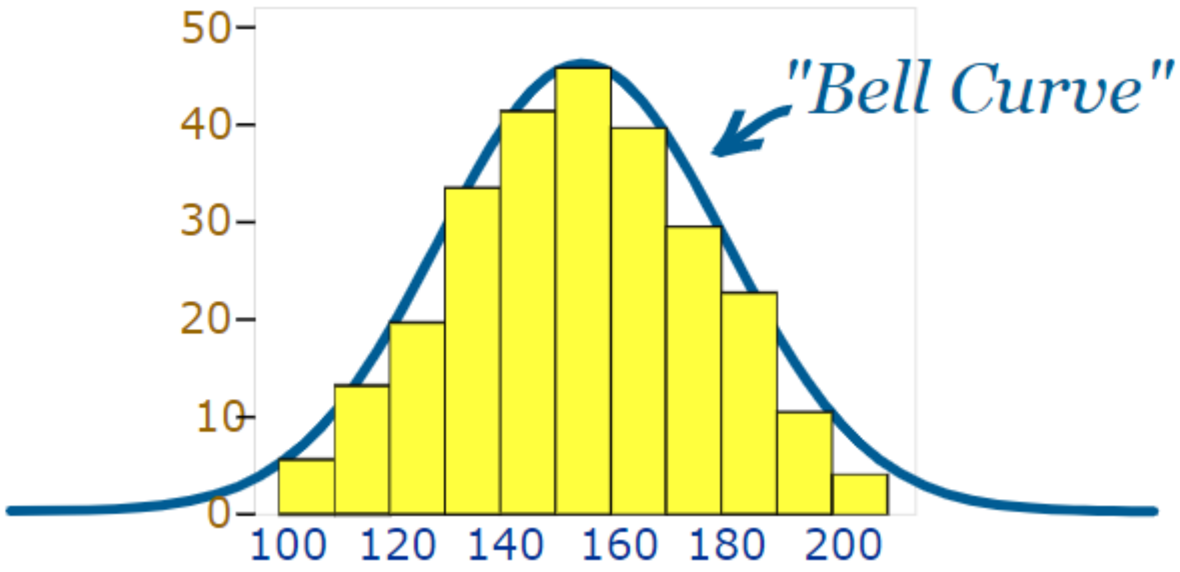
$$E[X] = \sum_{e \in \Omega} X(e) P(e)$$

$$E[X] = \sum_{k} k \cdot P(X=k)$$

k ranges over all possible values that X can take.

Do we need something else to fully capture how our distribution behaves, other than its **expected value**?

Question 2



Probability mass function

Probability mass functions are used for discrete distributions. We assign a **probability** to each point in the sample space.

Given a random variable X

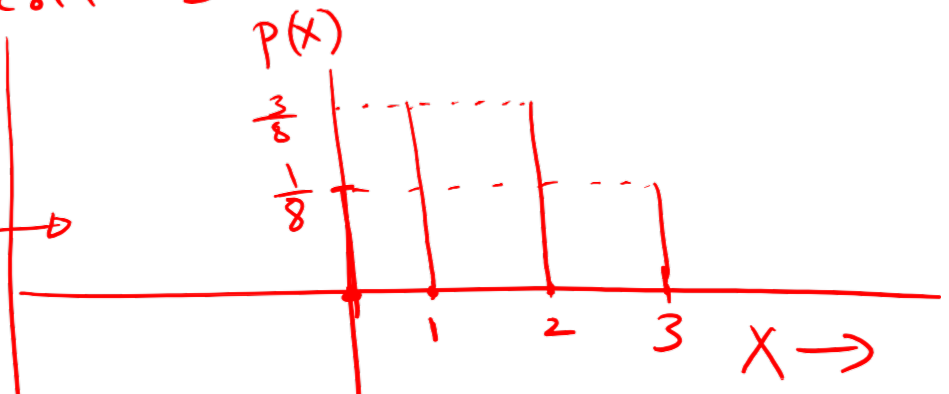
- Table providing values of the form $P(X=a)$ for each possible value a that X can take.

Experiment: Toss a coin 3 times. $X = \#$ Heads

pmf(x) →

$P(X=0)$	$\frac{1}{8}$
$P(X=1)$	$\frac{3}{8}$
$P(X=2)$	$\frac{3}{8}$
$P(X=3)$	$\frac{1}{8}$

OR



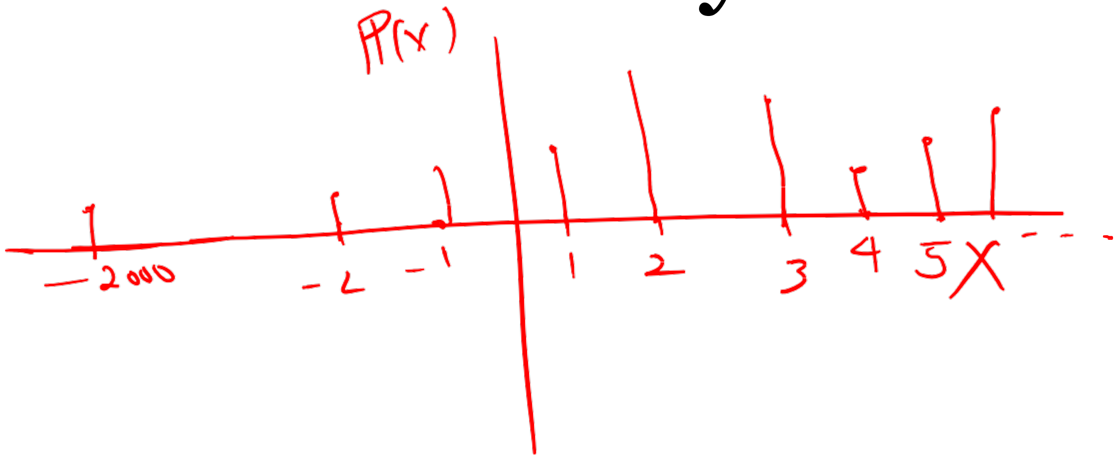
$E[X] \rightarrow$ summary of $P(X)$

Probability mass function – ways to describe

Summaries of Pmf(x)

- $E[X]$ is the most basic summary
- Variance → we will stop here
- Skewness
- Kurtosis
- high order moments

⋮



Probability mass function–Example for Variance

Game 1:

Toss a fair coin
- H: win \$1
- T: lose \$1

Game 2:

Toss a fair coin
- H: win \$1 million
- T: lose \$1 million

→ Seems risky

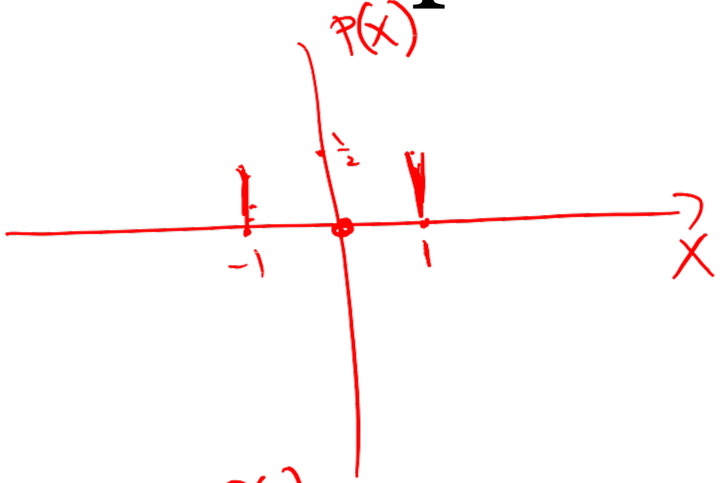
Define X = money earned

$E[X] = \frac{1}{2}(1) + \frac{1}{2}(-1) = 0$ → Game 1

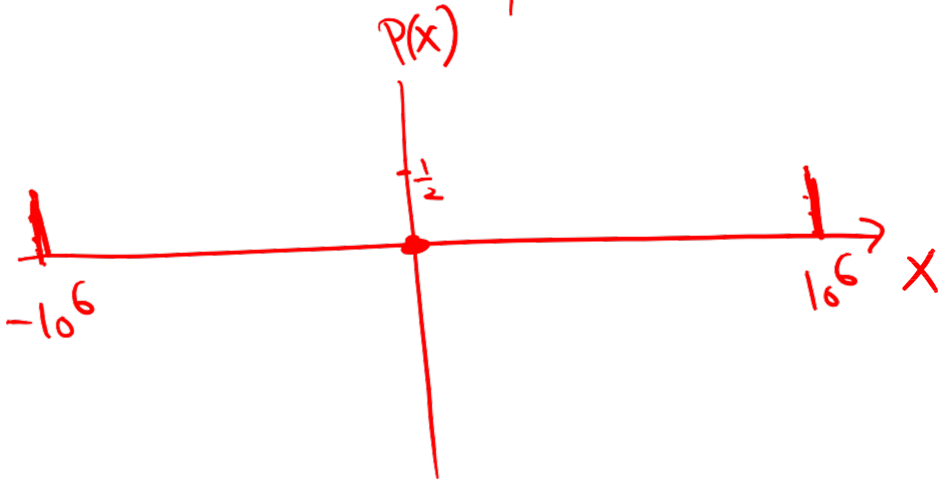
$E[X] = \frac{1}{2}(10^6) + \frac{1}{2}(-10^6) = 0$ → Game 2

Probability mass function–Example for Variance

pmf(X) for Game 1



Pmf(X) for Game2



Variance of X helps
us distinguish Game1 vs Gam2

Probability mass function–Example for Variance

Let $\mu = E[X]$

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

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

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

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

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

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

$$\boxed{\mathrm{Var}(X) = E[X^2] - \mu^2} \rightarrow$$

How to compute $E[X^2]$?

Probability Distribution – Variance

$$\mathrm{Var}(X) = \Sigma x^2 p - \mu^2$$

To calculate the Variance:
- square each value and multiply by its probability
- sum them up and we get $\mathbf{\Sigma x^2 p}$
- then subtract the square of the Expected Value $\boldsymbol{\mu^2}$

Example continued:

x	1	2	3	4	5	6
p	0.1	0.1	0.1	0.1	0.1	0.5
x^2p	0.1	0.4	0.9	1.6	2.5	18

$\Sigma x^2 p = 0.1+0.4+0.9+1.6+2.5+18 = 23.5$

$$\mathrm{Var}(X) = \Sigma x^2 p - \mu^2 = 23.5 - 4.5^2 = 3.25$$

The variance is 3.25

Probability Distribution – Standard Deviation

$$\sigma = \sqrt{Var(X)}$$

Example continued:

x	1	2	3	4	5	6
p	0.1	0.1	0.1	0.1	0.1	0.5
x^2p	0.1	0.4	0.9	1.6	2.5	18

$$\sigma = \sqrt{Var(X)} = \sqrt{3.25} = 1.803...$$

The Standard Deviation is 1.803...

Linearity of Expectation

- Remember: Expectation is a linear operation

$$E[X_1 + X_2] = E[X_1] + E[X_2]$$

$$E[X_1 + X_2 + \cdots + X_N] = E[X_1] + E[X_2] + \cdots + E[X_N]$$

$$X = X_1 + X_2 + \cdots + X_N$$

$$E[X] = \sum_k k\,P(X=k)$$

$$E[X^2] = \sum_k k^2\,P(X=k)$$

$$E[g(X)] = \sum_k g(k)\,P(X=k)$$

Describing "expectations" from random variable

Consider a random variable X.

Let $\mu = E[X]$

Variance $\equiv$ average deviation of X from its typical value

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

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

$$E[X] = \sum_k k\, P(X=k)$$

$$E[X^2] = \sum_k k^2\, P(X=k)$$

$$E[g(X)] = \sum_k g(k)\, P(X=k)$$

Linearity of Expectation - Example

- Two fair dice are rolled. What is the expected value of the sum?

$X = \text{Sum}$

$X_1 = \text{value of die } 1$

$X_2 = \text{value of die } 2$

Ω

(1,1)
(1,2)
⋮
(6,6)

$$X = X_1 + X_2$$

$$E[X] = E[X_1] + E[X_2]$$

$$E[X_1] = \sum_{k=1}^{6} k P(X_1 = k) = \frac{1}{6} \sum_{k=1}^{6} k = \frac{7}{2}$$

$$E[X_1] = \frac{7}{2}$$

$$E[X_2] = \frac{7}{2}$$

$$\Rightarrow E[X] = 7$$

The expected value of the sum is the sum of the expected values!

Bernoulli Random Variables

- A random variable X is Bernoulli if it takes only two values, 0 and 1.

$$E[X] = 0 \cdot P(X=0) + 1 \cdot P(X=1)$$
$$= P(X=1)$$

Expected Value of a Function

- A fair coin is tossed n times.
  - Expected number of (heads X tails)?

$X = \#$ Heads

$E[X(n-X)]$??

We know: $E[X] = \sum_k k P(X=k)$

Definition $\longrightarrow$ $\boxed{E[X^2] = \sum_k k^2 P(X=k)}$

In general $\longrightarrow$ $E[X^2-3X+4] = \sum_k (k^2-3k+4) P(X=k)$

$\longrightarrow$ $E[g(X)] = \sum_k g(k) P(X=k)$, for any function $g()$

Random Variables

- If I randomly put 100 letters in 100 addressed envelopes, on average how many will go into the right envelope?

$X =$ # letters that go into right envelope

$$E[X] = 0 \cdot P(X=0) + 1 \cdot P(X=1) + 2 \cdot P(X=2) + \cdots + 100 \cdot P(X=100)$$

$P(X=0) = P_0(100) \leftarrow$ no rabbit goes to correct position

$P(X=1) = P_1(100) \leftarrow$ 1 rabbit goes to correct position

$\vdots$

$P(X=100) = P_{100}(100)$

$$E[X] = 0 \cdot P_0(100) + 1 \cdot P_1(100) + \cdots 100 \cdot P_{100}(100)$$

Random Variables

- If I randomly put 100 letters in 100 addressed envelopes, on average how many will go into the right envelope?

$X =$ # letters that go into right envelope

$X_1 = 1$ if letter 1 goes to right envelope
$= 0$ otherwise

$X_2 = 1$ if letter 2 goes to right envelope
$= 0$ otherwise

$\vdots$

$X_{100} = 1$ if letter 100 goes to right envelope
$= 0$ otherwise

$X = X_1 + X_2 + \cdots + X_{100}$

$\Rightarrow$

Random Variables

- If I randomly put 100 letters in 100 addressed envelopes, on average how many will go into the right envelope?

We have, $E[X_1] = P(X_1 = 1) = \frac{1}{100}$

$E[X_2] = P(X_2 = 1) = \frac{1}{100}$

$\vdots$

$E[X_{100}] = P(X_{100} = 1) = \frac{1}{100}$

$\Rightarrow E[X] = 1$

Random Variables

- What if I flip a coin 3 times? What is the expected number of heads? What is the variance?

$X = \#$ heads

$X_1 = 1$ if flip 1 was H
$\quad = 0$ otherwise

$X_2 = 1$ if flip 2 was H
$\quad = 0$ otherwise

$X_3 = 1$ if flip 3 was H
$\quad = 0$ otherwise

$X = X_1 + X_2 + X_3$

$\Rightarrow E[X] = E[X_1] + E[X_2] + E[X_3]$

$E[X_1] = P(X_1 = 1) = \frac{1}{2}$

$E[X_2] = \frac{1}{2}$

$E[X_3] = \frac{1}{2}$

$\Rightarrow \boxed{E[X] = \frac{3}{2}}$

+ Variance

- What if I flip a coin 3 times? What is the expected number of heads? What is the variance?

Let $X = \#\text{heads}$ $E[X]$, $Var[X]$

$$E[X] = \sum_{k=0}^{3} k\, P(X=k) = 0 \cdot P(X=0) + 1 \cdot P(X=1) + 2 \cdot P(X=2) + 3 \cdot P(X=3)$$

$$= \frac{3}{8} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8} = \frac{3}{2} = \mu$$

+ Variance

- What if I flip a coin 3 times? What is the expected number of heads? What is the variance?

$X = \#$ heads

$E[X]$?

$\mathrm{Var}[X]$?

$E[X] = \frac{3}{2} = \mu$

By formula $\mathrm{Var}[X] = E[X^2] - \mu^2 = E[X^2] - \frac{9}{4}$

$$E[X^2] = \sum_{k=0}^{3} k^2 P(X=k) = 0 \cdot P(X=0) + 1 \cdot P(X=1) + 2^2 \cdot P(X=2) + 3^2 \cdot P(X=3)$$

$$= 1 \cdot \frac{3}{8} + 2^2 \cdot \frac{3}{8} + 3^2 \cdot \frac{1}{8} = \frac{24}{8} = 3$$

$$\mathrm{Var}[X] = 3 - \frac{9}{4} = \frac{3}{4}$$

Variance is never negative

Random Variables

$P(H) = p$

- We flip n coins of bias p. What is the expected number of heads? What is the variance?

$X =$ # heads

$X_1 = 1$ if flip 1 was H
$= 0$ otherwise

$X_2 = 1$ if flip 2 was H
$= 0$ otherwise

$\vdots$

$X_n = 1$ if flip n was H
$= 0$ otherwise

$X = X_1 + X_2 + \cdots + X_n$
$\Rightarrow E[X] = E[X_1] + \quad + E[X_n]$

$E[X_1] = P(X_1 = 1)$
$= p$

$E[X_2] = p$

$\vdots$

$\Rightarrow \boxed{E[X] = np}$

+ Variance

- What if I flip a coin of bias p n times? What is the expected number of heads? What is the variance?

$X = \#$ Heads, $E[X], Var[X]$?

Define

$X_1 = 1$ if flip 1 was H

$= 0$ otherwise

$X_2 = 1$ if flip 2 was H

$= 0$ otherwise

$\vdots$

$X_n = 1$ if flip n was H

$= 0$ otherwise

Then $X = X_1 + X_2 + \cdots + X_n$

$\Rightarrow E[X] = E[X_1] + E[X_2] + \cdots + E[X_n]$

+ Variance

- What if I flip a coin of bias p n times? What is the expected number of heads? What is the variance?

We have for $i = 1$ to n, $E[X_i] = P(X_i = 1) = p$

Hence $E[X] = np$

Random Variables

- We flip a coins of bias p until a H turns up. What is the expected number of flips?

$X =$ # flips till H

$X_1 = 1$ if flip 1 = T
$= 0$ otherwise

$X_2 = 1$ if flip 1 and flip 2 = T
$= 0$ otherwise

$X_3 = 1$ if flips 1,2,3 = T
$= 0$ otherwise

$\vdots$

$X = 1 + X_1 + X_2 + \cdots$

Random Variables

- We flip a coins of bias p until a H turns up. What is the expected number of flips?

$$E[X] = E[1] + E[X_1] + E[X_2] + \cdots$$
$$= 1 + E[X_1] + E[X_2] + \cdots$$

$$E[X_1] = 1-p, \quad E[X_2] = (1-p)^2, \quad E[X_3] = (1-p)^3$$

$$E[X] = 1 + (1-p) + (1-p)^2 + (1-p)^3 + \cdots$$

$$1 + x + x^2 + \cdots = \frac{1}{1-x} \quad [\text{If } x < 1]$$

$$E[X] = \frac{1}{1-(1-p)} = \frac{1}{p}$$

Random Variables

- A building has n floors numbered 1,2,⋯,n plus a ground floor G. At the ground floor, m people get on the elevator together, and each gets off at a uniformly random one of the n floors (independently of everybody else).
- What is the expected number of floors the elevator stops at (not counting the ground floor)?

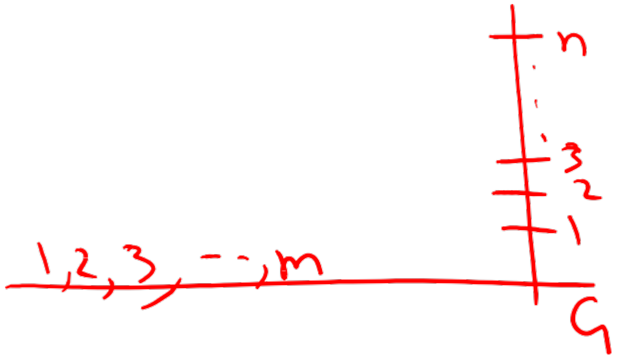
$X =$ # floors elevator stops at

$X_1 = 1$ if someone chose floor 1
$\quad = 0$ otherwise

$X_2 = 1$ if someone chose floor 2
$\quad = 0$ otherwise

⋮

$X_n = 1$ if someone chose floor n



Random Variables

$$E[X] = E[X_1] + E[X_2] + \text{---} + E[X_n]$$

$$E[X_1] = P(X_1 = 1) = P(\text{Someone chose floor } 1)$$

$$= P(\text{at least one picked floor } 1)$$

$$= 1 - P(\text{no one picked floor } 1)$$

$$= 1 - P\begin{pmatrix} \text{Person 1 did not pick 1} \\ \text{and} \\ \text{person 2 did not pick 1} \\ \text{and} \\ \vdots \\ \text{Person m did not pick 1} \end{pmatrix}$$

Random Variables

$$E[X] = E[X_1] + E[X_2] + \text{---} + E[X_n]$$

$$E[X_1] = P(X_1 = 1) = P(\text{Someone chose flower } 1)$$

$$= P(\text{at least one picked flower } 1)$$

$$= 1 - P(\text{no one picked flower } 1)$$

$$= 1 - P\begin{pmatrix} \text{Person 1 did not pick 1} \\ \text{and} \\ \text{person 2 did not pick 1} \\ \text{and} \\ \vdots \\ \text{Person m did not pick 1} \end{pmatrix}$$

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

Random Variables

$$E[X] = E[X_1] + E[X_2] + \text{---} + E[X_n]$$

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

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

Properties of Mean and Variance

Expectation

① $E[a] = a$

② $E[aX] = aE[X]$

③ $E[a_1X_1 + a_2X_2 + \cdots + a_nX_n]$
$= a_1E[X_1] + a_2E[X_2] + \cdots + a_nE[X_n]$

Variance

$Var[a] = 0$

$Var[aX] = a^2 Var[X]$

$Var[a_1X_1 + a_2X_2 + \cdots + a_nX_n]$
$= a_1^2 Var[X_1] + a_2^2 Var[X_2] + \cdots + a_n^2 Var[X_n]$

↳ only if $X_1, X_2, \ldots, X_n$ are independent.

Properties of Mean and Variance

If a is a constant then

— $E[a] = a = \mu$

— $Var[a] = E[a^2] - \mu^2$

$= a^2 - a^2 = 0$

If a is a constant, X is a random variable

$$E[aX] = aE[X]$$

$$Var[aX] = a^2 Var[X]$$

Proof: Let $\mu = E[X]$, $E[ax] = \mu a$

$$Var[ax] = E[a^2x^2] - \mu^2 a^2 = a^2 E[X^2] - \mu^2 a^2$$

$$= a^2\left(E[X^2] - \mu^2\right)$$

$$= a^2 \, Var[X]$$

Properties of Mean and Variance

$E[X_1 + X_2] = E[X_1] + E[X_2]$

$Var[X_1 + X_2] = Var[X_1] + Var[X_2]$

↳ only if X_1 and X_2 are independent

$E[a_1X_1 + a_2X_2 + a_3X_3 + \cdots + a_nX_n] = a_1E[X_1] + a_2E[X_2] + \cdots + a_nE[X_n]$

↳ always true

$Var[a_1X_1 + a_2X_2 + \cdots + a_nX_n] = a_1^2Var[X_1] + a_2^2Var[X_2] + \cdots + a_n^2Var[X_n]$

↳ only true if $X_1, X_2, \ldots, X_n$ are independent

Properties of Variance - Example

- X is a random variable such that $E[X] = 100, Var[X] = 15$
  - $E[X^2]$?
  - $E[3X + 10]$? $\rightarrow 3E[X] + 10 = 310$
  - $E[-X]$? $\rightarrow (-1)E[X] = -100$
  - $Var[-X]$? $\rightarrow (-1)^2 Var[X] = 15$

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

$$15 = E[X^2] - (100)^2$$

$$\Rightarrow E[X^2] = 15 + (100)^2$$

Variance

- What if I flip a coin of bias p n times? What is the expected number of heads? What is the variance?

Since $X_1, X_2, \ldots, X_n$ are independent

we have

$$Var[X] = Var[X_1] + Var[X_2] + \cdots + Var[X_n]$$

Let's compute $Var[X_i]$ for $i = 1$ to n.

we have, $\mu = E[X_i] = p(X_i = 1) = p$

Hence $Var[X_i] = E[X_i^2] - \mu^2 = E[X_i^2] - p^2$

$$= 0^2 \cdot p(X_i = 0) + 1^2 \cdot p(X_i = 1) - p^2$$

$$= p - p^2 = p(1-p)$$

$$\Rightarrow \boxed{Var[X] = np(1-p)}$$

Variance

- We flip a coins of bias p until a H turns up. What is the expected number of flips? What is the variance?

$X = \#$ flips until H, $E[X]$, $Var[X]$?

Define $X_1 = 1$ if flip 1 is T
$\quad = 0$ otherwise

$X_2 = 1$ if flip 1, flip 2 are T
$\quad = 0$ otherwise

$\vdots$

$X_i = 1$ if flips 1 to i are all Ts
$\quad = 0$ otherwise

Then $X = 1 + X_1 + X_2 + \cdots$

Variance

- We flip a coins of bias p until a H turns up. What is the expected number of flips? What is the variance?

Hence,
$$E[X] = 1 + E[X_1] + E[X_2] + \cdots$$

Next,
$$E[X_i] = P(X_i = 1) = (1-p)^i$$

Hence,
$$E[X] = 1 + (1-p) + (1-p)^2 + \cdots$$
$$= \frac{1}{p}$$

Variance

- We flip a coins of bias p until a H turns up. What is the expected number of flips? What is the variance?

$$Var[X] = Var[1 + X_1 + X_2 + \cdots]$$
$$= Var[1] + Var[X_1] + Var[X_2] + \cdots$$
$$= 0 + Var[X_1] + Var[X_2] + \cdots$$

— This is wrong. Can't use linearity since $X_1, X_2, \ldots$ are not independent.

— Hence, go back to original formula

Variance

- We flip a coins of bias p until a H turns up. What is the expected number of flips? What is the variance?

$$Var[X] = E[X^2] - \mu^2 = E[X^2] - \frac{1}{p^2}$$

$$E[X^2] = \sum_{k=1}^{\infty} k^2 P(X=k)$$

$$P(X=k) = P(\text{H turns up after } k \text{ tries})$$
$$= P(\text{Try } 1 = T, \text{Try } 2 = T, \ldots \text{Try } k-1 = T, \text{Try}_k = H)$$
$$= (1-p)^{k-1} \cdot p$$

Variance

- We flip a coins of bias p until a H turns up. What is the expected number of flips? What is the variance?

$$E[X^2] = \sum_{k=1}^{\infty} k^2 p (1-p)^{k-1}$$

$$= p\Big[\underbrace{1 + 2^2(1-p) + 3^2(1-p)^2 + 4^2(1-p)^3 + \cdots}_{S}\Big]$$

$$= \frac{2-p}{p^2}$$

$$\Rightarrow Var[X] = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}$$

Variance

- We flip a coins of bias p until a H turns up. What is the expected number of flips? What is the variance?

Proof that

$$p\left[1+2^2(1-p)+3^2(1-p)^2+4^2(1-p)^3+\cdots\right]$$

$$=\frac{2-p}{p^2}$$

Let

$$S=1+2^2(1-p)+3^2(1-p)^2+4^2(1-p)^3+\cdots$$

We want to prove that $pS=\frac{2-p}{p^2}$

In other words, $S=\frac{2-p}{p^3}$

Variance

- We flip a coins of bias p until a H turns up. What is the expected number of flips? What is the variance?

(i) — $S = 1 + 2^2(1-p) + 3^2(1-p)^2 + 4^2(1-p)^3 + \cdots$

(ii) — $(1-p)S = (1-p) + 2^2(1-p)^2 + 3^2(1-p)^3 + \cdots$

Subtracting (ii) from (i) we get

(iii) — $pS = 1 + 3(1-p) + 5(1-p)^2 + 7(1-p)^3 + \cdots$

(iv) $p(1-p)S = (1-p) + 3(1-p)^2 + 5(1-p)^3 + \cdots$

subtracting (iv) from (iii) we get

$p^2 S = 1 + 2(1-p) + 2(1-p)^2 + 2(1-p)^3 + \cdots$

Variance

- We flip a coins of bias p until a H turns up. What is the expected number of flips? What is the variance?

$$p^2 S = 1 + 2(1-p) + 2(1-p)^2 + 2(1-p)^3 + \cdots$$

$$= 1 + 2\left[(1-p) + (1-p)^2 + (1-p)^3 + \cdots\right]$$

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

$$\Rightarrow S = \frac{2-p}{p^3}$$

$\square$