Alex Psomas: Lecture 16.

$\boxed{\text{Random Variables}}$

- Regrade requests open.
- Quiz due tomorrow.
- Quiz coming out today.
- Non-technical office hours tomorrow 1-3pm.
- Anonymous questionnaire tonight or tomorrow.

Random Variables

1. Random Variables.
2. Distributions.
3. Combining random variables.
4. Expectation

Questions about outcomes ...

Experiment: roll two dice.
Sample Space: $\{(1,1),(1,2),\dots,(6,6)\} = \{1,\dots,6\}^2$
How many dots?

Experiment: flip 100 coins.
Sample Space: $\{HHH\cdots H, THH\cdots H,\dots, TTT\cdots T\}$
How many heads in 100 coin tosses?

Experiment: choose a random student in cs70.
Sample Space: $\{Peter, Phoebe,\dots,\}$
What midterm score?

Experiment: hand back assignments to 3 students at random.
Sample Space: $\{123, 132, 213, 231, 312, 321\}$
How many students get back their own assignment?

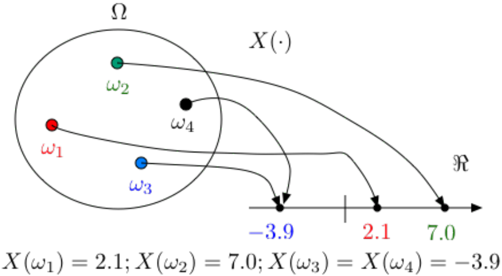
In each scenario, each outcome gives a number.

The number is a (known) function of the outcome.

Random Variables.

A **random variable,** X**,** for an experiment with sample space Ω is a function $X : \Omega \to \Re$.

Thus, $X(\cdot)$ assigns a real number $X(\omega)$ to each $\omega \in \Omega$.



The function $X(\cdot)$ is defined on the outcomes Ω.

A random variable X is not random, not a variable!

What varies at random (from experiment to experiment)? The outcome!

Example 1 of Random Variable

Experiment: roll two dice.
Sample Space: $\{(1,1),(1,2),\ldots,(6,6)\} = \{1,\ldots,6\}^2$
Random Variable X: number of pips.
$X(1,1) = 2$
$X(1,2) = 3,$
$\vdots$
$X(6,6) = 12,$
$X(a,b) = a+b, (a,b) \in \Omega.$

Example 2 of Random Variable

Experiment: flip three coins

Sample Space: $\{HHH, THH, HTH, TTH, HHT, THT, HTT, TTT\}$

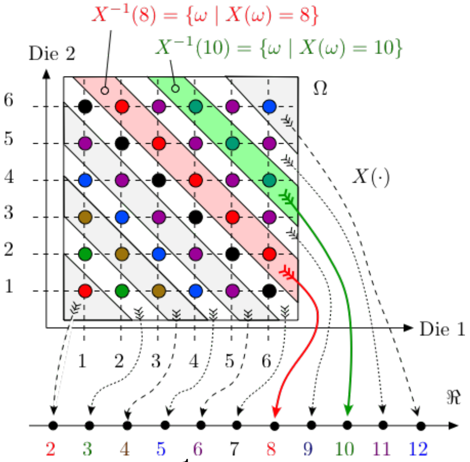
Winnings: if win 1 on heads, lose 1 on tails: X

$X(HHH) = 3$ $X(THH) = 1$ $X(HTH) = 1$ $X(TTH) = -1$

$X(HHT) = 1$ $X(THT) = -1$ $X(HTT) = -1$ $X(TTT) = -3$

Number of dots in two dice.

"What is the likelihood of seeing *n* dots?"



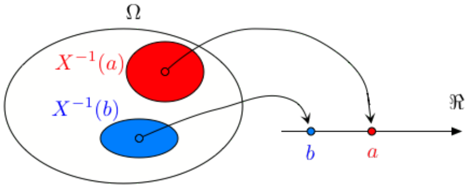
$Pr[X = 10] = 3/36 = Pr[X^{-1}(10)] = \sum_{\omega \in X^{-1}(10)} Pr[\omega]$

$Pr[X = 8] = 5/36 = Pr[X^{-1}(8)]$.

Distribution

The probability of X taking on a value a.

Definition: The **distribution** of a random variable X, is $\{(a, Pr[X = a]) : a \in \mathscr{A}\}$, where $\mathscr{A}$ is the range of X.



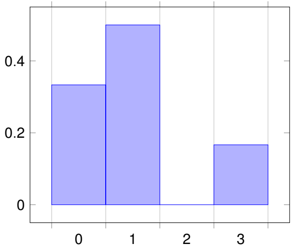
$$Pr[X = a] := Pr[X^{-1}(a)] \text{ where } X^{-1}(a) := \{\omega \mid X(\omega) = a\}.$$

Handing back assignments

Experiment: hand back assignments to 3 students at random.
Sample Space: $\Omega = \{123, 132, 213, 231, 312, 321\}$
How many students get back their own assignment?
Random Variable: values of $X(\omega) : \{3, 1, 1, 0, 0, 1\}$

Distribution:

$$X = \begin{cases} 0, & \text{w.p. } 1/3 \\ 1, & \text{w.p. } 1/2 \\ 3, & \text{w.p. } 1/6 \end{cases}$$

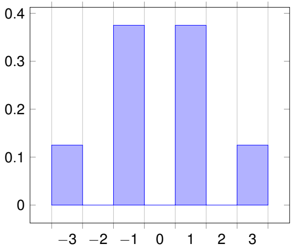


Flip three coins

Experiment: flip three coins
Sample Space: $\{HHH, THH, HTH, TTH, HHT, THT, HTT, TTT\}$
Winnings: if win 1 on heads, lose 1 on tails. X
Random Variable: $\{3, 1, 1, -1, 1, -1, -1, -3\}$

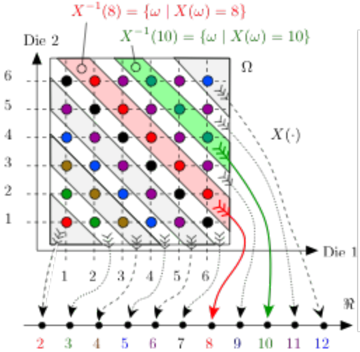
Distribution:

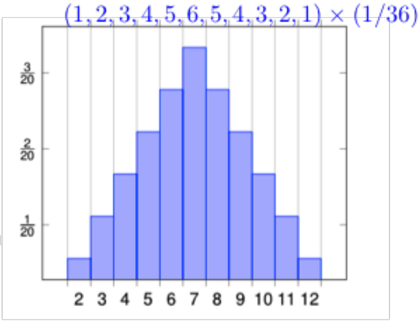
$$X = \begin{cases} -3, & \text{w. p. } 1/8 \\ -1, & \text{w. p. } 3/8 \\ 1, & \text{w. p. } 3/8 \\ 3 & \text{w. p. } 1/8 \end{cases}$$



Number of dots.

Experiment: roll two dice.





The Bernoulli distribution

Flip a coin, with heads probability p.

Random variable X: 1 is heads, 0 if not heads.

X has the Bernoulli distribution.

We will also call this an **indicator random variable**. It indicates whether the event happened.

Distribution:

$$X = \begin{cases} 1 & \text{w.p. } p \\ 0 & \text{w.p. } 1 - p \end{cases}$$

The binomial distribution.

Flip n coins with heads probability p.

Random variable: number of heads.

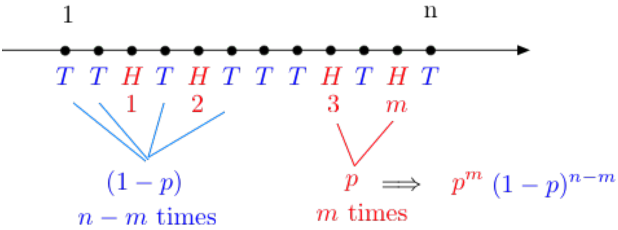
Binomial Distribution: $Pr[X = i]$, for each i.

How many sample points in event "$X = i$"?
i heads out of n coin flips $\implies \binom{n}{i}$

Sample space: $\Omega = \{HHH...HH, HHH...HT, \ldots\}$
What is the probability of ω if ω has i heads?
Probability of heads in any position is p.
Probability of tails in any position is $(1-p)$.
So, we get $Pr[\omega] = p^i(1-p)^{n-i}$.
Probability of "$X = i$" is sum of $Pr[\omega]$, $\omega \in$ "$X = i$".

$$Pr[X = i] = \binom{n}{i} p^i (1-p)^{n-i}, i = 0, 1, \ldots, n : B(n,p) \text{ distribution}$$

The binomial distribution.



$\binom{n}{m}$ outcomes with m Hs and $n - m$ Ts

$$\Longrightarrow \boxed{Pr[X = m] = \binom{n}{m} p^m (1-p)^{n-m}}$$

Combining Random Variables.

Let X and Y be two RV on the same probability space.

That is, $X : \Omega \to \Re$ assigns the value $X(\omega)$ to ω. Also, $Y : \Omega \to \Re$ assigns the value $Y(\omega)$ to ω.

Then $Z = X + Y$ is a random variable: It assigns the value

$$Z(\omega) = X(\omega) + Y(\omega)$$

to outcome ω.

Experiment: Roll two dice. X = outcome of first die, Y = outcome of second die.

$$X(a,b) = a \text{ and } Y(a,b) = b \text{ for } (a,b) \in \Omega = \{1,\ldots,6\}^2.$$

Then $Z = X + Y$ = sum of two dice is defined by

$$Z(a,b) = X(a,b) + Y(a,b) = a + b.$$

Combining Random Variables

Other random variables:

- $X^k : \Omega \to \Re$ is defined by $X^k(\omega) = [X(\omega)]^k$.
  In the dice example, $X^3(a,b) = a^3$.
- $(X-2)^2 + 4XY$ assigns the value $(X(\omega)-2)^2 + 4X(\omega)Y(\omega)$ to ω.
- $g(X,Y,Z)$ assigned the value $g(X(\omega),Y(\omega),Z(\omega))$ to ω.

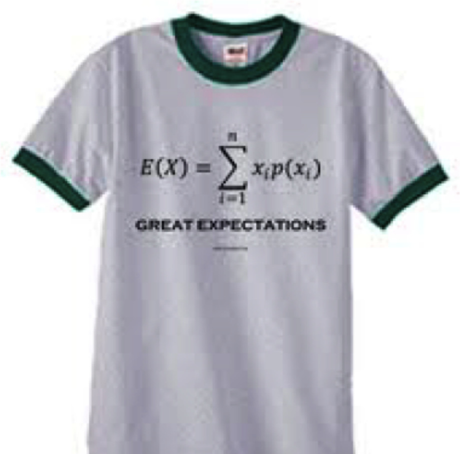
Expectation.

How did people do on the midterm?

Distribution.

Summary of distribution?

Average!



Expectation - Intuition

Flip a loaded coin with $Pr[H] = p$ a large number N of times.

We expect heads to come up a fraction p of the times and tails a fraction $1-p$.

Say that you get 5 for every H and 3 for every T.

If there are N_H outcomes equal to H and N_T outcomes equal to T, you collect

$$5 \times N_H + 3 \times N_T.$$

Your average gain per experiment is

$$\frac{5N_H + 3N_T}{N}.$$

Since $\frac{N_H}{N} \approx p = Pr[X = 5]$ and $\frac{N_T}{N} \approx 1 - p = Pr[X = 3]$, we find that the average gain per outcome is approximately equal to

$$5Pr[X = 5] + 3Pr[X = 3].$$

We use this frequentist interpretation as a definition.

Expectation - Definition

Definition: The **expected value** of a random variable X is

$$E[X] = \sum_a a \times Pr[X = a].$$

a in the range of X.
The expected value is also called the mean.

According to our intuition, we expect that if we repeat an experiment a large number N of times and if $X_1, \ldots, X_N$ are the successive values of the random variable, then

$$\frac{X_1 + \cdots + X_N}{N} \approx E[X].$$

That is indeed the case, in the same way that the fraction of times that $X = x$ approaches $Pr[X = x]$.

This (nontrivial) result is called the Law of Large Numbers.

Expectation: A Useful Fact

Theorem:

$$E[X] = \sum_{\omega \in \Omega} X(\omega) \times Pr[\omega].$$

Proof:

$$\begin{aligned}
E[X] &= \sum_a a \times Pr[X = a] \\
&= \sum_a a \times \sum_{\omega : X(\omega) = a} Pr[\omega] \\
&= \sum_a \sum_{\omega : X(\omega) = a} a \times Pr[\omega] \\
&= \sum_a \sum_{\omega : X(\omega) = a} X(\omega) Pr[\omega] \\
&= \sum_\omega X(\omega) Pr[\omega]
\end{aligned}$$

An Example

Flip a fair coin three times.

$\Omega = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$.

$X =$ number of H's: $\{3,2,2,2,1,1,1,0\}$.

Thus,

$$\sum_{\omega} X(\omega) Pr[\omega] = \{3+2+2+2+1+1+1+0\} \times \frac{1}{8}.$$

Also,

$$\sum_{a} a \times Pr[X=a] = 3 \times \frac{1}{8} + 2 \times \frac{3}{8} + 1 \times \frac{3}{8} + 0 \times \frac{1}{8}.$$

Expectation and Average.

There are n students in the class;

$X(m)$ = score of student m, for $m = 1, 2, \ldots, n$.

"Average score" of the n students: add scores and divide by n:

$$\text{Average} = \frac{X(1) + X(1) + \cdots + X(n)}{n}.$$

Experiment: choose a student uniformly at random.
Uniform sample space: $\Omega = \{1, 2, \cdots, n\}, Pr[\omega] = 1/n$, for all ω.
Random Variable: midterm score: $X(\omega)$.
Expectation:

$$E(X) = \sum_{\omega} X(\omega) Pr[\omega] = \sum_{\omega} X(\omega) \frac{1}{n}.$$

Hence,

$$\text{Average } = E(X).$$

Our intuition matches the math.

Handing back assignments

We give back assignments randomly to three students.
What is the expected number of students that get their own assignment back?

The expected number of **fixed points** in a random permutation.

Expected value of a random variable:

$$E[X] = \sum_a a \times Pr[X = a].$$

For 3 students (permutations of 3 elements):

$$Pr[X = 3] = 1/6, Pr[X = 1] = 3/6, Pr[X = 0] = 2/6.$$

$$E[X] = 3 \times \frac{1}{6} + 1 \times \frac{3}{6} + 0 \times \frac{2}{6} = 1.$$

Win or Lose.

Expected winnings for heads/tails games, with 3 flips?
Every time it's H ,I get 1,. Every time it's T, I lose 1.

$$E[X] = 3 \times \frac{1}{8} + 1 \times \frac{3}{8} - 1 \times \frac{3}{8} - 3 \times \frac{1}{8} = 0.$$

Can you ever win 0?

Apparently: expected value is not a common value, by any means.

Expectation

Recall: $X : \Omega \to \Re; Pr[X = a]; = Pr[X^{-1}(a)];$

Definition: The **expectation** of a random variable X is

$$E[X] = \sum_a a \times Pr[X = a].$$

Indicator:

Let A be an event. The random variable X defined by

$$X(\omega) = \begin{cases} 1, & \text{if } \omega \in A \\ 0, & \text{if } \omega \notin A \end{cases}$$

is called the indicator of the event A.

Note that $Pr[X = 1] = Pr[A]$ and $Pr[X = 0] = 1 - Pr[A]$.

Hence,

$$E[X] = 1 \times Pr[X = 1] + 0 \times Pr[X = 0] = Pr[A].$$

The random variable X is sometimes written as

$$1\{\omega \in A\} \text{ or } 1_A(\omega).$$

Linearity of Expectation

Theorem:

$$E[X] = \sum_{\omega} X(\omega) \times Pr[\omega].$$

Theorem: Expectation is linear

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

Proof:

$$\begin{aligned}
&E[a_1 X_1 + \cdots + a_n X_n] \\
&\quad = \sum_{\omega} (a_1 X_1 + \cdots + a_n X_n)(\omega) Pr[\omega] \\
&\quad = \sum_{\omega} (a_1 X_1(\omega) + \cdots + a_n X_n(\omega)) Pr[\omega] \\
&\quad = a_1 \sum_{\omega} X_1(\omega) Pr[\omega] + \cdots + a_n \sum_{\omega} X_n(\omega) Pr[\omega] \\
&\quad = a_1 E[X_1] + \cdots + a_n E[X_n].
\end{aligned}$$

$\square$

Using Linearity - 1: Dots on dice

Roll a die n times.

X_m = number of dots on roll m.

$X = X_1 + \cdots + X_n$ = total number of dots in n rolls.

$$\begin{aligned} E[X] &= E[X_1 + \cdots + X_n] \\ &= E[X_1] + \cdots + E[X_n], \text{ by linearity} \\ &= nE[X_1], \text{ because the } X_m \text{ have the same distribution} \end{aligned}$$

Now,

$$E[X_1] = 1 \times \frac{1}{6} + \cdots + 6 \times \frac{1}{6} = \frac{6 \times 7}{2} \times \frac{1}{6} = \frac{7}{2}.$$

Hence,

$$E[X] = \frac{7n}{2}.$$

Using Linearity - 2: Fixed point.

Hand out assignments at random to n students.

$X =$ number of students that get their own assignment back.

$X = X_1 + \cdots + X_n$ where
$X_m = 1\{$student m gets his/her own assignment back$\}$.

One has

$$
\begin{aligned}
E[X] &= E[X_1 + \cdots + X_n] \\
&= E[X_1] + \cdots + E[X_n], \text{ by linearity} \\
&= nE[X_1], \text{ because all the } X_m \text{ have the same distribution} \\
&= nPr[X_1 = 1], \text{ because } X_1 \text{ is an indicator} \\
&= n(1/n), \text{ because student 1 is equally likely} \\
&\qquad\qquad \text{to get any one of the } n \text{ assignments} \\
&= 1.
\end{aligned}
$$

Note that linearity holds even though the X_m are not independent (whatever that means).

Using Linearity - 3: Binomial Distribution.

Flip n coins with heads probability p. X - number of heads

Binomial Distibution: $Pr[X = i]$, for each i.

$$Pr[X = i] = \binom{n}{i} p^i (1-p)^{n-i}.$$

$$E[X] = \sum_i i \times Pr[X = i] = \sum_i i \times \binom{n}{i} p^i (1-p)^{n-i}.$$

No no no no no. NO ... Or... a better approach: Let

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

$E[X_i] = 1 \times Pr[\text{``}heads''] + 0 \times Pr[\text{``}tails''] = p.$

Moreover $X = X_1 + \cdots X_n$ and

$E[X] = E[X_1] + E[X_2] + \cdots E[X_n] = n \times E[X_i] = np.$

Today's gig: St. Petersburg paradox

I offer the following game:

We start with a pot of 2 dollars.

Flip a fair coin. If it's tails, you take the pot. If it's heads, I double the pot.

So, if the sequence is HHT, you make 8 dollars.

How much would you we willing to pay?

Today's gig: St. Petersburg paradox

Well, how much money should you expect to make?

Let X be the random variable indicating how much money you make for each outcome:

$X = 2$ with probability $\frac{1}{2}$

$X = 4$ with probability $\frac{1}{4}$

$X = 8$ with probability $\frac{1}{8}$

$$E[X] = 2\frac{1}{2} + 4\frac{1}{4} + 8\frac{1}{8} + \ldots$$

$$= 1 + 1 + 1 + \ldots = \infty$$

So, if you were rational you would be willing to pay anything!

Is there a trick here?

Today's gig: St. Petersburg paradox

What if I didn't have infinite money?

Banker	Bankroll	Expected value of lottery
Friendly game	$100	$7.56
Millionaire	$1,000,000	$20.91
Billionaire	$1,000,000,000	$30.86
Bill Gates (2015)	$79,200,000,000[5]	$37.15
U.S. GDP (2007)	$13.8 trillion[6]	$44.57
World GDP (2007)	$54.3 trillion[6]	$46.54
Googolaire	$$10^{100}$	$333.14

Summary

Random Variables

- A random variable X is a function $X : \Omega \to \Re$.
- $Pr[X = a] := Pr[X^{-1}(a)] = Pr[\{\omega \mid X(\omega) = a\}]$.
- $Pr[X \in A] := Pr[X^{-1}(A)]$.
- The distribution of X is the list of possible values and their probability: $\{(a, Pr[X = a]), a \in \mathscr{A}\}$.
- $g(X, Y, Z)$ assigns the value .... .
- $E[X] := \sum_a a Pr[X = a]$.
- Expectation is Linear.
- $B(n, p)$.