Technical vocab: ① random variable
② expected value
③ Bernoulli Trial
④ Binomial random variable
⑤ Geometric random variable
⑥ Indicator random variable

- X((1,1))= 2,
- X((1,2))= X((2,1))= 3,
- X((1,3))= X((2,2))= X((3,1))= 4,
- X((1,4))= X((2,3))= X((3,2))= X((4,1))= 5,
- X((1,5))= X((2,4))= X((3,3))= X((4,2))= X((5,1))= 6,
- X((1,6))= X((2,5))= X((3,4))= X((4,3))= X((5,2))= X((6,1))= 7,
- X((2,6))= X((3,5))= X((4,4))= X((5,3))= X((6,2))= 8,
- X((3,6))= X((4,5))= X((5,4))= X((6,3))= 9,
- X((4,6))= X((5,5))= X((6,4))= 10,
- X((5,6))= X((6,5))= 11,
- X((6,6))= 12.

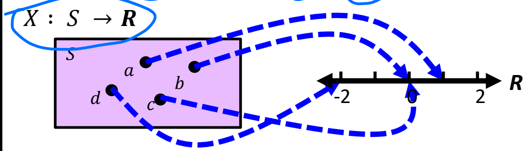
$p: S \to [0,1]$

$X: S \to R$

Lec 25: Random Variables & Expectation - ANSWERS

① A **random variable** is a function from a sample space to the real numbers

$X : S \to \boldsymbol{R}$



$X(a) = 1$
$X(b) = 0$
$X(c) = 0$
$X(d) = -2$
"$X = 0$" $= \{b, c\}$

$\boldsymbol{X} = \boldsymbol{r}$ is the set of all outcomes that map to r

Example: Roll 2 dice. Let X = sum of the dice. S is set of pairs.

- X((2,3)) = 5
- "X = 7" = {(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)}
- "X > 10" = {(5,6), (6,5), (6,6)}

$p(X = 7) = 1/6$ $\qquad p(X > 10) = 3/36 = 1/12$

Random Variable Examples

Experiment (a): roll a (fair) die.

I'll pay you \$5 for a 1, \$10 for a 2, but you pay me \$3 for any other number rolled.

Define the random variables X, Y, Z as follows:

X = the number you rolled, Y = amount you win/lose

Z = 1 if you win money, 0 if not

Experiment (b): roll an unfair die. p(6) = ½ and uniform probability over other rolls

s	$p(s)$ *fair die*	$p(s)$ *unfair die*	$X(s)$	$Y(s)$	$Z(s)$
1	1/6	1/10	1	5	1
2	1/6	1/10	2	10	1
3	1/6	1/10	3	-3	0
4	1/6	1/10	4	-3	0
5	1/6	1/10	5	-3	0
6	1/6	1/2	6	-3	0

Expected Value

The ***expected value*** of $X: S \to \boldsymbol{R}$ is the weighted average of X

Two ways to find E(X):

②

- $E(X) = \sum_{s \in S} p(s) \cdot X(s)$ (weighted sum over outcomes)
- $E(X) = \sum_{r \in range(X)} p(X = r) \cdot r$ (weighted sum over range of X)

Example: Find the expected winnings for the dice game with the *unfair* die

$E(Y) = \sum_{r \in range(Y)} p(Y = r) \cdot r$

$= p(Y = 5)(5) + p(Y = 10)(10) + p(Y = -3)(-3)$

$= (\frac{1}{10})(5) + (\frac{1}{10})(10) + (\frac{1}{10} + \frac{1}{10} + \frac{1}{10} + \frac{1}{2})(-3)$

$= \frac{1}{2} + 1 - \frac{24}{10} \quad = -\frac{9}{10}$

Exercises: Expected Value

Note that you can use either equation for E(X) in both problems, but we just show one.

- Expected value of a roll of a (fair) die: S = {1,2,3,4,5,6}, p(s)=1/6 and X(s)=s, for all s∈S.

$E(X) = \sum_{s \in S} p(s) \cdot X(s)$

$= p(1)X(1) + p(2)X(2) + p(3)X(3) + p(4)X(4) + p(5)X(5) + p(6)X(6)$

$= (1/6) \cdot 1 + (1/6) \cdot 2 + (1/6) \cdot 3 + (1/6) \cdot 4 + (1/6) \cdot 5 + (1/6) \cdot 6$

$= 1/6 + 2/6 + 3/6 + 4/6 + 5/6 + 6/6 \quad = \mathbf{3.5}$

- Expected value of a roll of the unfair die from dice game:

S = {1,2,3,4,5,6},
p(1)=p(2)= p(3)=p(4)=p(5) = 1/10 and p(6) = ½
X(1) = 1, X(2) = 2, X(3) = 3, X(4) = 4, X(5) = 5, X(6) = 6

$E(X) = \sum_{s \in S} p(s) \cdot X(s)$

$= p(1)X(1) + p(2)X(2) + p(3)X(3) + p(4)X(4) + p(5)X(5) + p(6)X(6)$

$= (1/10) \cdot 1 + (1/10) \cdot 2 + (1/10) \cdot 3 + (1/10) \cdot 4 + (1/10) \cdot 5 + (1/2) \cdot 6$

$= \frac{15}{10} + 3 \quad = \mathbf{4.5}$

Expected value of the sum of 2 fair dices: $S = \{(1,1) \ldots (6,6)\}$ $X = \{2, 3, \ldots, 12\}$;

$\Rightarrow |S| = 36, \ |X| = 11$

$E(X) = \sum_{s \in S} p(s) \cdot X(s)$: add up 36 values $(\frac{1}{36} \times (\ldots))$

$E(X) = \sum_{r \in range(X)} p(X=r) \cdot r$: add up 11 values

Thm. if X.Y are random variables, $(X+Y)(s) = X(s) + Y(s)$ is also a random variable

Linearity of Expectation

- The expected value of the sum of random variables is the sum of their expected values

(3)
$$E(X+Y) = E(X) + E(Y)$$
$$E(aX+b) = aE(X) + b \quad \text{for any constants } a, b$$

(4) • Does linearity require that X and Y be *independent*? NO, linearity *always* works

(5) • Does $E(XY) = E(X)E(Y)$? Only when X and Y are independent iff !!

Example : X = sum of two dice, i.e., X((a,b)) = a+b

- X_1 = outcome of 1st die, X_2 = outcome of 2nd die
- $X = X_1 + X_2$.
- $E(X) = E(X_1+X_2)$ LoE
- $= E(X_1) + E(X_2)$
- $= 7/2 + 7/2$
- $= 7$

(6)
- *Events* E, F are *independent* if $p(E \cap F) = p(E) \cdot p(F)$.
- *Random variables* X, Y are *independent* if any events derived from X, Y are independent.
  - E.g., E = event that $X \leq 5$; F = event that Y is an odd integer.
  - If X, Y are independent then $\Pr(E \cap F) = \Pr(E) \cdot \Pr(F)$

Indicator RVs

(8)
- An ***indicator variable*** I_F indicates whether an event F happened or not:

$$I_F = \begin{cases} 1 & \text{, if } F \text{ happened} \\ 0 & \text{, if not} \end{cases}$$

- Let $p(F) = p$. Then $E(I_F) = 1 \cdot p + 0 \cdot (1-p) = p$ (7)
- Together with Linearity of Expectation, Indicator variables can help us determine the expected number of "successes" in repeated trials of the same experiment.

Example: You play the dice game 10 times, with the fair die.
What is the expected number of games you win?

Let Z_1 indicate whether we won game 1 (Z_1=1 means won game 1),
Let Z_2 indicate whether we won game 2, etc.
And let Z be the total number of wins over 10 games.

$Z = Z_1 + Z_2 + \ldots + Z_{10}$
$E(Z) = E(Z_1 + Z_2 + \ldots + Z_{10})$
$= E(Z_1) + E(Z_2) + \ldots + E(Z_{10})$
$= 1/3 + 1/3 + \ldots + 1/3$
$= 10/3$

$E(Z_i) = 1 \cdot p(win) + 0 \cdot p(lose)$
$= 1 \cdot 1/3$
$= 1/3$

Bernoulli Trials & Binomial Distribution

Bernoulli Trial

Experiment that has exactly 2 outcomes,

- S = {success, failure}
- p(success) = p
- p(failure) = $q = (1-p)$

Binomial Experiment

Repeat the Bernoulli trial n times, where

- Each trial has the same probability of success
- All trials are mutually independent
- Let X = total # of successes

(9)
Binomial Distribution:
The probability of exactly $\boldsymbol{k}$ successes in $\boldsymbol{n}$ independent (and identically distributed) Bernoulli trials is

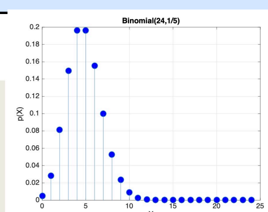
$$p(X=k) = \binom{n}{k} p^k q^{n-k}$$

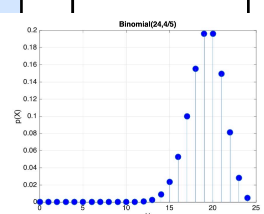
And $E(X) = np$

Expected # of successes in a **single** Bernoulli trial:
$E(X_i) = 1 \cdot p + 0 \cdot (1-p) = p$

Expected # of successes in $\boldsymbol{n}$ Bernoulli trials:

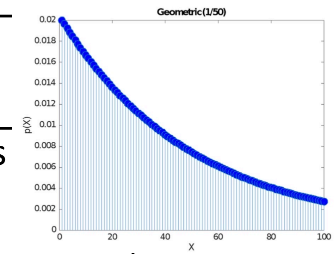
$E(X) = E(X_1 + X_2 + \cdots + X_n)$
$= E(X_1) + E(X_2) + \cdots + E(X_n)$ by linearity of expectation
$= p + p + \cdots + p$
$= np$





Geometric Random Variables

aka "Waiting Time" Experiment



Repeat a Bernoulli trial until you get your first "success".

- The number of trials it took can be described with a Geometric RV

Example:

A coin has probability p of Heads. Toss it until you get the first Heads.
Let X be the number of flips including the last one.

- What is $p(X=k)$?
- What is $E(X)$?

(10)
Geometric Distribution:
The probability of requiring exactly $\boldsymbol{k}$ trials to achieve the first success in a sequence of Bernoulli trials is

$$p(X=k) = (1-p)^{k-1} p$$

And $E(X) = 1/p$

$$E(X) = p(T)E(X|T) + p(H)E(X|H)$$
$$E(X) = p \cdot 1 \;+\; (1-p)(E(X)+1)$$

Solve for E(X) and you get…

$$E(X) = 1/p$$

Seinfeld Reruns:

$X_{i \to j}$ = the number of days you have to watch to go from having watched i distinct shows to having watched j distinct shows.

$\Rightarrow X_{0\to180} = X_{0\to1} + X_{1\to2} + \ldots + X_{179\to180} = \sum_{k=0}^{179} X_{k\to k+1}$

$\because X_{k\to k+1}$ is a geometric r.v. with success prob. $\frac{180-k}{180}$

$\therefore E(X_{k\to k+1}) = \frac{180}{180-k}$

By linearity of Expectation $\therefore E(X_{0\to180}) = \frac{180}{180} + \frac{180}{179} + \ldots + \frac{180}{1} \approx 180.6$