Unit 4 Discrete Random Variables

Probability Mass Functions and Expectations

A random variable is loosely speaking, a numerical quantity whose value is determined by the outcome of a probabilistic experiment e.g. the weight of a randomly selected person

Discrete: takes values in finite or countable set
Random variable examples:
Bernoulli
Uniform
Binomial
Geometric

Expected value of a random variable aka Expectation (mean)
weighted average of the values of the random variable
weighted on their probabilities

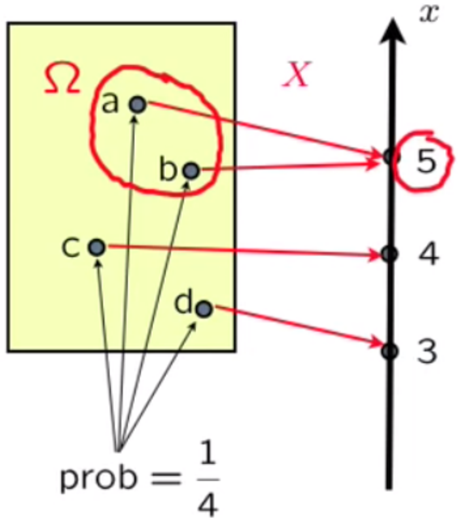
Random variables mathematically: A function from the sample space omega to the real numbers

Notation: random variable X numerical value x

A function of one or more random variables is also a random variable

Probability Mass Function (PMF) aka probability law or probability distribution of X

Probability space of getting different values
P(5) = 1/2 a and b



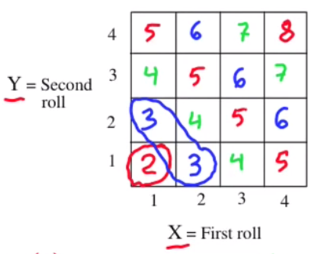
Notation:

$$p_X(x) = \mathbf{P}(X = x) = \mathbf{P}\Big(\{\omega \in \Omega \text{ s.t. } X(\omega) = x\}\Big)$$

- **Properties:** $p_X(x) \geq 0$ $\qquad \sum_x p_X(x) = 1$

PMF calculation

- Two rolls of a tetrahedral die
- Let every possible outcome have probability 1/16



$Z = X + Y$ $\qquad$ Find $p_Z(z)$ $\quad$ for all z

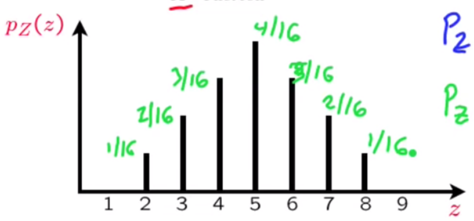
- repeat for all z:
  - collect all possible outcomes for which Z is equal to z
  - add their probabilities

$P_Z(2) = P(Z = 2) = 1/16$

$P_Z(3) = P(Z = 3) = 2/16$

$P_Z(4) = P(Z = 4) = 3/16$

$\vdots$



Exercise: Random variables versus numbers

1/2 points (graded)

Let X be a random variable that takes integer values, with PMF $p_X(x)$. Let Y be another integer-valued random variable and let y be a number.

a) Is $p_X(y)$ a random variable or a number?

Random variable ✖ **Answer:** Number

b) Is $p_X(Y)$ a random variable or a number?

Random variable ✔ **Answer:** Random variable

Solution:

a) Recall that $p_X(\cdot)$ is a function that maps real numbers to real numbers. So, when we give it a numerical argument, y, we obtain a number.

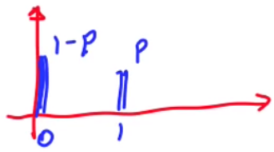
b) In this case, we are dealing with a function, the function being $p_X(\cdot)$, of a random variable Y. And a function of a random variable is a random variable. Intuitively, the "random" value of $p_X(Y)$ is generated as follows: we observe the realized value y of the random variable Y, and then look up the numerical value $p_X(y)$.

The simplest random variable: Bernoulli with parameter $p \in [0,1]$

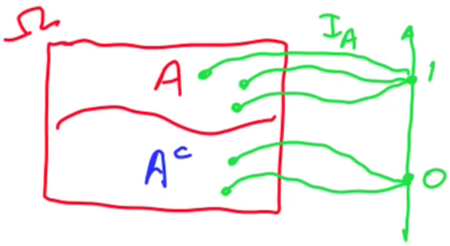
$$X = \begin{cases} 1, & \text{w.p. } p \\ 0, & \text{w.p. } 1-p \end{cases}$$

$p_X(0) = 1 - p$

$p_X(1) = p$



- Models a trial that results in success/failure, Heads/Tails, etc.
- Indicator r.v. of an event A: $I_A = 1$ iff A occurs



$p_{I_A}(1) = P(I_A = 1) = \underline{P(A)}$

p

Exercise: Indicator variables

2/2 points (graded)

Let A and B be two events (subsets of the same sample space Ω), with nonempty intersection. Let I_A and I_B be the associated indicator random variables.

For each of the two cases below, select one statement that is true.

a) $I_A + I_B$:

is not the indicator random variable of any event ✔

Answer: is not the indicator random variable of any event

b) $I_A \cdot I_B$:

is the indicator variable of the event $A \cap B$ ✔

Answer: is the indicator variable of the event $A \cap B$

(*Bug warning:* In some browsers, the mathematical content in each choice in the dropdown menu may appear duplicated, e.g. $A \cup B$ may show up twice as $A \cup BA \cup B$.)

Solution:

a) If the outcome of the experiment lies in the intersection of the events A and B, then $I_A + I_B$ takes the value of 2. But indicator random variables can take only the values 0 or 1. Therefore, $I_A + I_B$ is not an indicator random variable.

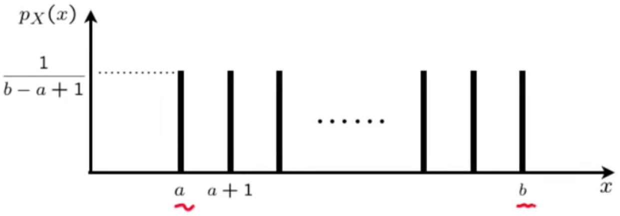
b) Note that $I_A \cdot I_B$ can take only the values 0 or 1. It is equal to 1 if and only if $I_A = 1$ (i.e., event A occurs) and $I_B = 1$ (i.e., event B occurs). Thus, $I_A \cdot I_B$ takes the value of 1 if and only if both A and B occur, and so it is the indicator random variable of the event $A \cap B$.

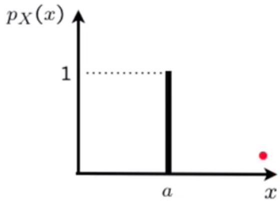
Discrete uniform random variable; parameters a, b

- **Parameters:** integers a, b; $a \leq b$
- **Experiment:** Pick one of $a, a+1, \ldots, b$ at random; all equally likely
- **Sample space:** $\{a, a+1, \ldots, b\}$ $b - a + 1$ possible values
- **Random variable** X: $X(\omega) = \omega$ 11:52:26 $\{0, 1, \ldots, 59\}$
- **Model of:** complete ignorance

Special case: $a = b$

constant/deterministic r.v.



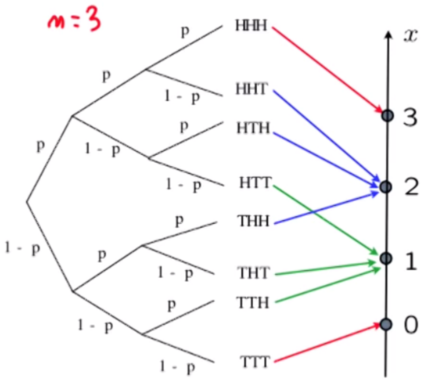


e.g. imagine looking at a clock at time 11:52:26 but only looking at the seconds the probability of getting a number 0 to 59 is equally likely

Special case where a = b is still a random variable in the mathematical sense, it just so happens that it can only be one value in this case

Binomial random variable; **parameters:** positive integer n; $p \in [0,1]$

- **Experiment:** n independent tosses of a coin with $\mathbf{P}(\text{Heads}) = p$
- **Sample space:** Set of sequences of H and T, of length n
- **Random variable** X**:** number of Heads observed
- **Model of:** number of successes in a given number of independent trials



$$P_X(2) = P(X=2)$$
$$= P(HHT) + P(HTH) + P(THH)$$
$$= 3p^2(1-p) = \binom{3}{2} p^2 (1-p)$$

$$\boxed{p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad \text{for } k = 0, 1, \ldots, n}$$

Same formula as earlier but with slightly different notation

Exercise: The binomial PMF

2/2 points (graded)

You roll a fair six-sided die (all 6 of the possible results of a die roll are equally likely) 5 times, independently. Let X be the number of times that the roll results in 2 or 3. Find the numerical values of the following.

a) $p_X(2.5) =$ 0 ✔ **Answer:** 0

b) $p_X(1) =$ 0.3292 ✔ **Answer:** 0.32922

Solution:

a) A value of 2.5 is not possible for X since the number of rolls must be an integer, and therefore $p_X(2.5) = 0$.

b) For each die roll, there is a probability $2/6 = 1/3$ of obtaining a 2 or a 3. Hence, the random variable X is binomial with parameters $n = 5$ and $p = 1/3$, so that $p_X(1) = \binom{5}{1} \cdot (1/3) \cdot (2/3)^4 \approx 0.32922$.

Geometric random variable; parameter p: $0 < p \leq 1$

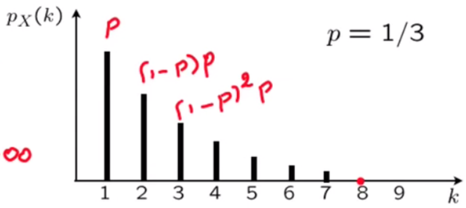
- **Experiment:** infinitely many independent tosses of a coin; $\mathbf{P}(\text{Heads}) = p$
- **Sample space:** Set of infinite sequences of H and T $\qquad \underbrace{TTTTH}HT\cdots$
- **Random variable** X**:** number of tosses until the first Heads $\qquad X=5$

- **Model of:** waiting times; number of trials until a success

$$p_X(k) = \mathbf{P}(X=k) = \mathbf{P}(\underbrace{T\cdots T}_{k-1}H) = (1-p)^{k-1}p \qquad k=1,2,3,\ldots$$

$\mathbf{P}(\text{no Heads ever})$ $\leq \mathbf{P}(\underbrace{T\cdots T}_{k}) = \underbrace{(1-p)^k}$ $\downarrow k\to\infty$ 0

$(TTT\cdots)$ "$X=\infty$" $\Big\} = 0$



P(Tails forever) becomes 0 as k goes to infinity

Exercise: Geometric random variables

0/2 points (graded)

Let X be a geometric random variable with parameter p. Find the probability that $X \geq 10$. Express your answer in terms of p using standard notation (click on the "STANDARD NOTATION" button below.)

$\mathbf{P}(X \geq 10) =$ p/(1-((1-p)^9)) ✖ **Answer:** (1-p)^9

$$\frac{p}{1-\left((1-p)^9\right)}$$

STANDARD NOTATION

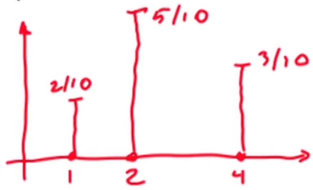
Solution:

We can calculate the desired probability by adding the probabilities of the events $\{X=10\}$, $\{X=11\}$, $\{X=12\}$, etc., and using the formula for the sum of a geometric series. However, we can get the answer in an easier way, using the interpretation of geometric random variables as the number of trials until the first success. The event $\{X \geq 10\}$ is the event that the the first 9 trials resulted in failure, and therefore its probability is $(1-p)^9$.

Expectation/mean of a random variable

- **Motivation:** Play a game 1000 times.
  Random gain at each play described by:

$$X = \begin{cases} 1, & \text{w.p. } 2/10 \\ 2, & \text{w.p. } 5/10 \\ 4, & \text{w.p. } 3/10 \end{cases} \qquad \begin{matrix} \sim 200 \\ \sim 500 \\ \sim 300 \end{matrix}$$



- "Average" gain:

$$\frac{1\cdot 200 + 2\cdot 500 + 4\cdot 300}{1000}$$

$$= 1\cdot\frac{2}{10} + 2\cdot\frac{5}{10} + 4\cdot\frac{3}{10}$$

- **Definition:** $\mathrm{E}[X] = \sum_x x p_X(x)$

- **Interpretation:** Average in large number of independent repetitions of the experiment

- Caution: If we have an infinite sum, it needs to be well-defined.
  We assume $\sum_x |x|\, p_X(x) < \infty$

Expectation of a Bernoulli r.v.

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

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

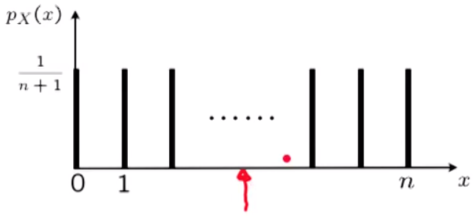
If X is the indicator of an event A, $X = I_A$:

$X = 1$ iff A occurs $\qquad p = P(A)$

$$E[I_A] = P(A)$$

Expectation of a uniform r.v.

- Uniform on $0, 1, \ldots, n$



- **Definition:** $\mathbf{E}[X] = \sum_x x p_X(x)$

$$\mathbf{E}[X] = 0 \cdot \frac{1}{n+1} + 1 \cdot \frac{1}{n+1} + \cdots + n \frac{1}{n+1}$$

$$= \frac{1}{n+1}(0+1+\cdots+n) = \frac{1}{n+1} \cdot \frac{n(n+1)}{2} = \frac{n}{2}$$

Elementary properties of expectations

- **Definition:** $\mathbf{E}[X] = \sum_x \underbrace{x}_{\geq 0} \underbrace{p_X(x)}_{\geq 0}$

≥ 0

- If $X \geq 0$, then $\mathbf{E}[X] \geq 0$

for all ω: $X(\omega) \geq 0$

- If $a \leq X \leq b$, then $a \leq \mathbf{E}[X] \leq b$

for all ω: $a \leq X(\omega) \leq b$

$$E[X] = \sum_x x p_X(x) \geq \sum_x a p_X(x)$$

$$= a \sum_x p_X(x) = a \cdot 1 = a$$

- If c is a constant, $\mathbf{E}[c] = c$

$$E[c] = c \cdot p(c) = c$$

Exercise: Random variables with bounded range

3/3 points (graded)

Suppose a random variable X can take any value in the interval $[-1, 2]$ and a random variable Y can take any value in the interval $[-2, 3]$.

a) The random variable $X - Y$ can take any value in an interval $[a, b]$. Find the values of a and b:

$a =$ -4 ✔

$b =$ 4 ✔

b) Can the expected value of $X + Y$ be equal to 6?

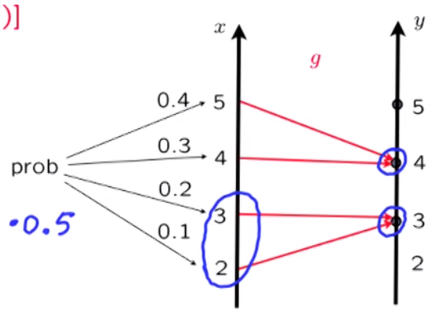
No ✔

The expected value rule, for calculating $\mathbf{E}[g(X)]$

- Let X be a r.v. and let $Y = g(X)$
- Averaging over y: $\mathbf{E}[Y] = \sum_y y p_Y(y)$

  $3 \cdot (0.1 + 0.2) + 4 \cdot (0.3 + 0.4)$
- Averaging over x: $3 \cdot 0.1 + 3 \cdot 0.2 + 4 \cdot 0.3 + 4 \cdot 0.5$

$$\boxed{\mathbf{E}[Y] = \mathbf{E}[g(X)] = \sum_x g(x) p_X(x)}$$



Proof: $\sum_y \sum_{x: g(x)=y} g(x) p_X(x)$

$= \sum_y \sum_{x: g(x)=y} y \, p_X(x) = \sum_y y \sum_{x: g(x)=y} p_X(x)$

$= \sum_y y \, p_Y(y) = E[Y]$

- $\mathbf{E}[X^2] = \sum_x x^2 p_X(x)$

  $g(x) = x^2$
- Caution: In general, $\mathbf{E}[g(X)] \neq g(\mathbf{E}[X])$

  $E[X^2] \neq (E[X])^2$

Linearity of expectation: $\mathbf{E}[aX+b] = a\mathbf{E}[X]+b$

X = salary $\quad E[X]$ = average salary

Y = new salary $= 2X + 100 \quad E[Y] = E[2X+100] = 2E[X] + 100$

- Intuitive

- **Derivation**, based on the expected value rule: $\quad g(x) = ax+b$

$Y = g(X)$

$$E[Y] = \sum_x g(x) p_X(x)$$

$$= \sum_x (ax+b) p_X(x) = a \sum_x x\, p_X(x) + b \underbrace{\sum_x p_X(x)}_{1}$$

$$= aE[X] + b$$

$E[g(X)] = g(E[X])$ exceptional g

The blue equality is only true for linear functions

Exercise: Linearity of expectations

3/3 points (graded)

The random variable X is known to satisfy $\mathbf{E}[X] = 2$ and $\mathbf{E}[X^2] = 7$. Find the expected value of $8 - X$ and of $(X-3)(X+3)$.

a) $\mathbf{E}[8 - X] =$ 6 ✔

b) $\mathbf{E}[(X-3)(X+3)] =$ -2 ✔