Unit 4 Discrete Random Variables

- Random Variables (r.v.) → Discrete / → Continuous
- their "distribution"
- expectation, variance

- Four main threads:
- definitions, notation
- properties of expectation and variance
- conditioning and independence
- ~~total~~ probability / expectation theorem

* Random variable → A numerical quantity whose value is determined by the outcome of a probabilistic experiment.
* Discrete → take values in finite or countable set

* Probability Mass Function (PMF) → Tells likelihood of the each possible value of random variable

* Random Variable Examples
→ Bernoulli, Uniform, Binomial, Geometric

* Expectation (mean) and its properties
- the expected value rule - Linearity

* Random Variables: the idea

a . b . c . d → randomly pick 4 students and record their weight : $W = \{62, 75, 52, 48\}$; $W \to$ Discrete r.v.

We can do Height: $H = \{\_, \_, \_, \_\}$

We can define new random variable: $B = \frac{W}{H^2}$

→ a r.v. associate a value (a number) to every possible outcome.

→ Mathematically: A function from the sample space Ω to the real numbers.

→ It can take discrete or continuous values

Notation: $\boxed{\text{random variable } X}$ $\boxed{\text{numerical value } x}$

X: (function) / abstract

x: (real value

→ We can have several r.vs defined on the same sample space

→ A function of one or several random variables is also a random variable.

— meaning of $X+Y$: r.v. takes value $x+y$, when X takes value x, Y takes value y.

Exercise: Random Variables

Background: a number, e.g. 2, can be thought of as a trivial r.v. that always takes value 2. Let x be a number. Let X be a r.v. associated with some probabilistic experiment.

a) It is always true that $X+x$ is a r.v.

→ Think of an example: Let X be the height of a randomly selected student and $x=10$. We're dealing with the variable $X+10$. It is the r.v. that takes the value $a+10$, whenever the r.v. takes value a.

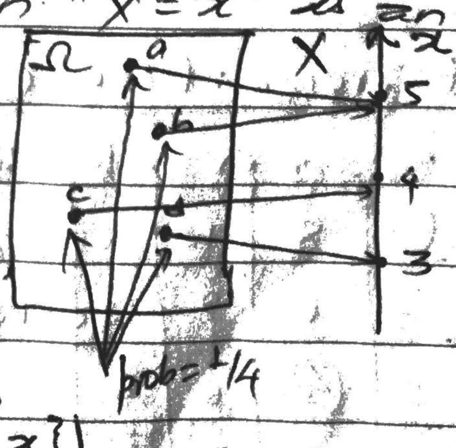
b) Is it always true that $X-x=0$?

→ No. Think like in (a), its like $X-10$.

Probability Mass Functions (PMF)

* It is the "probability law" or "probability-distribution" of X
* If we fix some x, then "$X=x$" is an event.

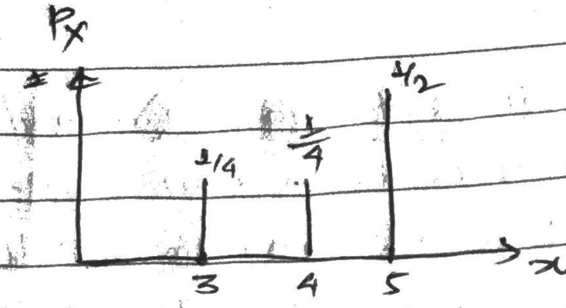
$x=5,\ X=5$ $\{\omega : X(\omega)=5\}$
$= \{a,b\}$.



$p_X(5) = \frac{1}{2}$

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

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



PMF Calculation

- Two rolls of a tetrahedral die
- Probability law: let every possible outcome have probability 1/16 (Assumption)

Y = second roll				
4	5	6	7	8
3	4	5	6	7
2	3	4	5	6
1	2	3	4	5
	1	2	3	4

X = First Roll

← $Z = X+Y$

Find $p_Z(z)$ for all z.

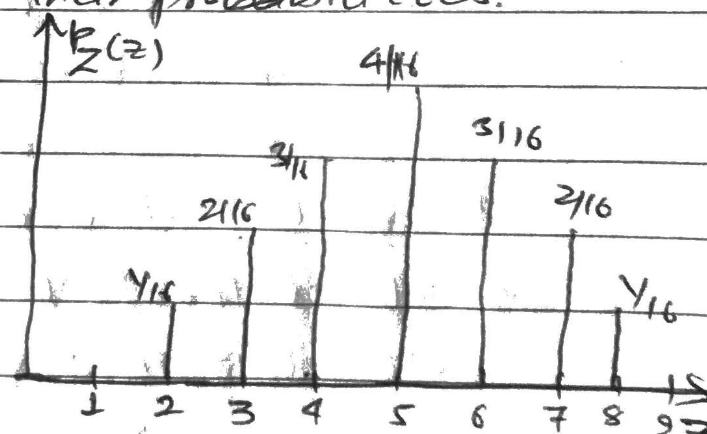
- repeat for all z:
  - collect all possible outcomes for which Z is equal to z.
  - add their probabilities.

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

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

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

and so on.



Exercise: PMF Calculation

As in example before, $W = XY$ (r.v. → W)

Y = 2nd roll				
4	4	8	12	16
3	3	6	9	12
2	2	4	6	8
1	1	2	3	4
	1	2	3	4

X = 1st roll

← $W = XY$

(a) $p_W(4) = p(W=4) = 3/16$

(b) $p_W(5) = p(W=5) = 0$

Exercise: Random Variables versus Numbers

Let X be a r.v. that takes integer values with PMF $p_X(x)$. Let Y be another integer-valued r.v. and y be a number.

a) Is $p_X(y)$ a r.v. or a number?

→ 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.

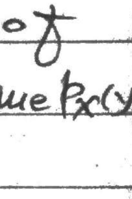
b) Is $p_X(Y)$ a r.v. or a number?

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

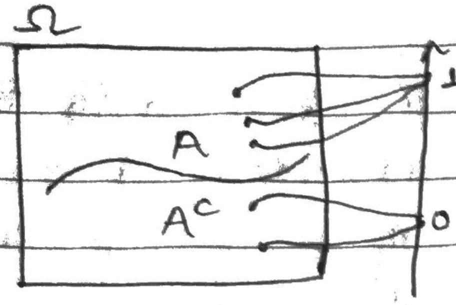
Bernouli and Indicator Random Variables

The simplest r.v: Bernoulli with parameter $p \in [0,1]$

$$X = \begin{cases} 1, & \text{with probability } p \\ 0, & \text{w.p. } 1-p \end{cases} \quad p_X(0) = 1-p \quad p_X(1) = p$$



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



$P_{I_A}(1) = P(I_A = 1) = P(A)$

Exercise: Indicator Variables

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.

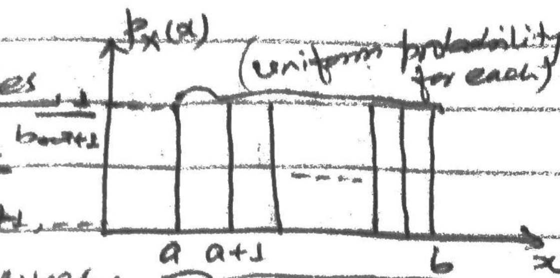
For each of the two cases below, select one statement that is true.

a) $I_A + I_B$

→ 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 $1+1=2$. But indicator r.vs. can take only 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 iff $I_A = 1$ (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 iff both A and B occur, and so is the indicator r.v. of the event $A \cap B$.

Uniform Random Variables



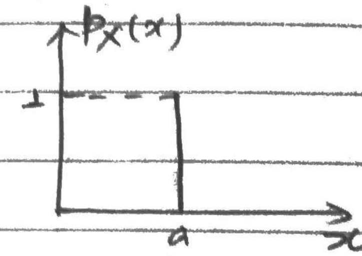
- parameters: a, b integers: $a<b$
- Experiment: Pick one of $a, a+1, \ldots, b$ at random; all equally likely
- sample space: $\{a, a+1, \ldots, b\}$
- Random variables X: $X(w) = w$

(integers between a and b w/ $b-a+1$ possible integers)

Real life Modeling: There is no reason to believe that one outcome is likely than others.

Special case: $a=b$

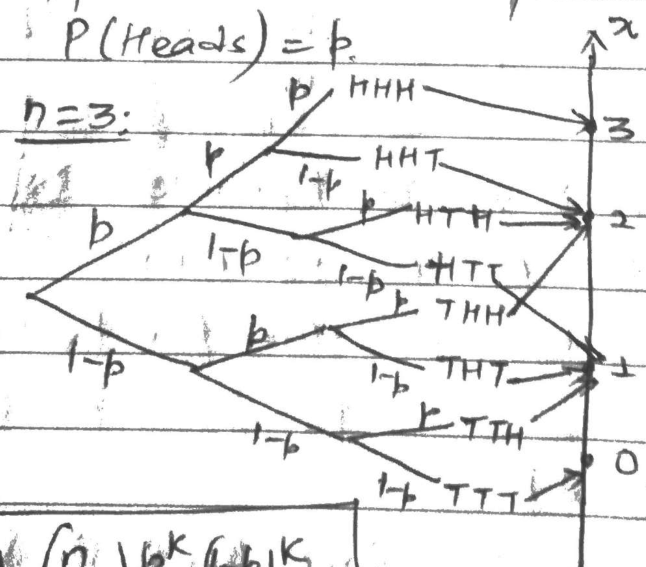
constant / deterministic r.v.



Binomial Random Variables

- parameters: positive integer n; $p \in [0,1]$
- Experiment: n independent tosses of a coin with $P(\text{Heads}) = p$.

$n=3$:



Random variable X: number of Heads observed.

Model of: number of successes in a given number of independent trials

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

$p_X(2) = P(X=2)$ (two heads)

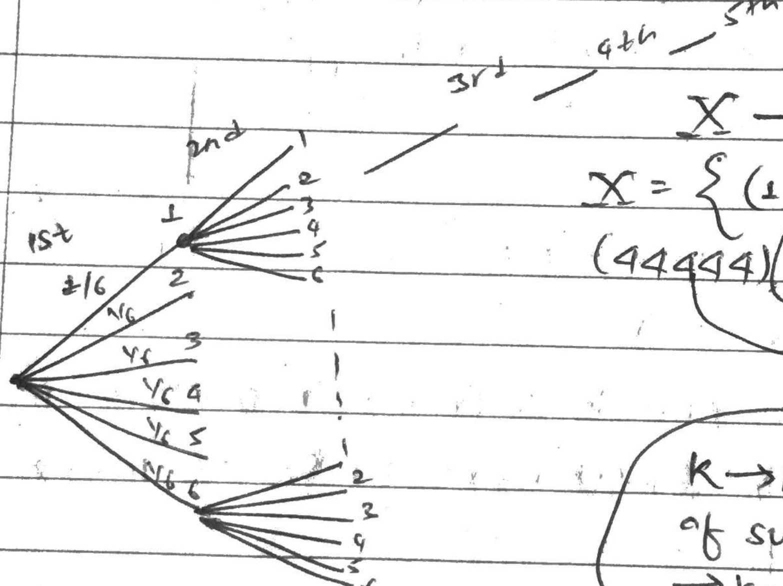
$= P(HHT) + P(HTH) + P(THH)$

$= 3p^2(1-p) = \binom{3}{2} p^2 (1-p)$

Exercise: The Binomial PMF

You roll a fair six-sided die (all of the 6 possible results of a die 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:

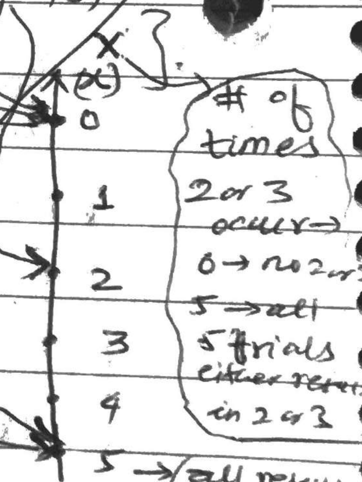
a) $p_X(2.5) = \mathbb{P}(X = 2.5)$



$X \rightarrow 6^5 \rightarrow$ total possible cases

$X = \{(11111), (22222), (33333), (44444), (55555), (12345) \ldots$

k → number of successes
→ $k = 0, \ldots, n$
$= 0, 1, 2, 3, 4, 5$



→ 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) $p_X(1) = \mathbb{P}(X = 1)$

For each die, there is a probability $2/6 = 1/3$ of obtaining 2 or 3. Hence a r.v. X is binomial with $n = 5$ and $p = 1/3$, so that $P(X=1) = \binom{5}{1} \cdot \left(\frac{1}{3}\right)^1 \left(\frac{2}{3}\right)^4 \approx 0.32922$

Geometric Random Variables

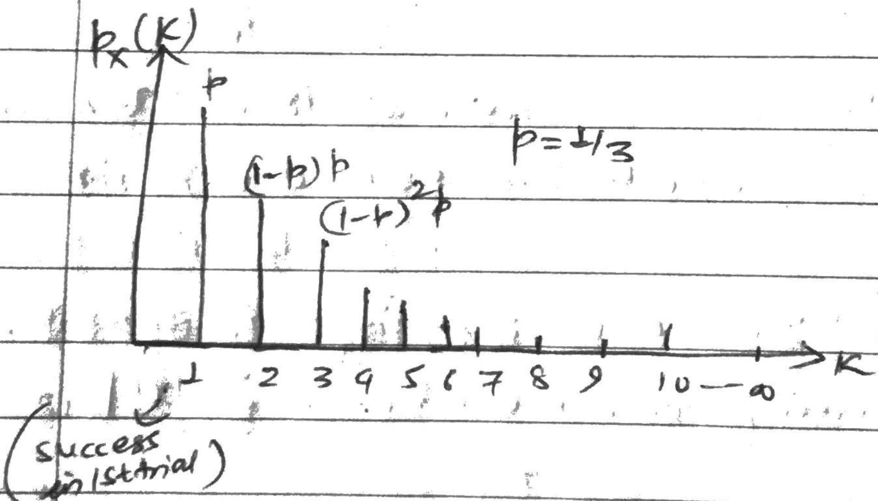
parameter p: $0 < p \le 1 \rightarrow p \in (0, 1]$

- Experiment: infinitely many independent tosses of a coin: $P(\text{Heads}) = p$ (finite subset of tosses are independent)

- Sample space: Set of infinite sequences of H and T
- Random variable X: number of tosses until the first Heads. eg. $\underbrace{TTTTH}_{X=5}HT\text{-----}$

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

$$p_X(k) = P(X=k) = P(T \ldots TH) = (1-p)^{k-1}p \qquad k = 1, 2, 3, \ldots$$



$P(\text{no Heads ever})$

$$P(X=\infty) \le P(\underbrace{T \ldots T}_{k}) = (1-p)^k \rightarrow 0 \text{ as } k \rightarrow \infty.$$

Exercise: Geometric Random Variables

Let X be a geometric r.v. with parameter p. Find the probability that $X \ge 10$. Express your answer in terms of p. → [equivalent to saying 1st 9 attempts fail → $(1-p)^9$] (k starts from 1)

Soln → $P(X \ge 10) = 1 - P(X<10) = 1 - [P(X=1) + \ldots + P(X=9)] = 1 - [p + (1-p)p + \ldots + p(1-p)^8]$

(# of trials until first (a) success)

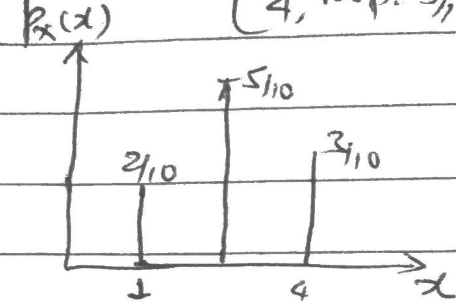
$$= 1 - \frac{p(1-(1-p)^9)}{1-(1-p)} = (1-p)^9$$

Expectation/Mean of a random variable

- Motivation: Play a game 1000 times.

Random gain at each play described by:
(win money)

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



- "Average" gain:

Using frequency:

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

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

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

- 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.

$$X = \begin{cases} 1, & \text{w.p. } p \\ 0, & \text{w.p } 1-p \end{cases} \longrightarrow E[X] = 1 \cdot p + 0 \cdot (1-p) = p$$

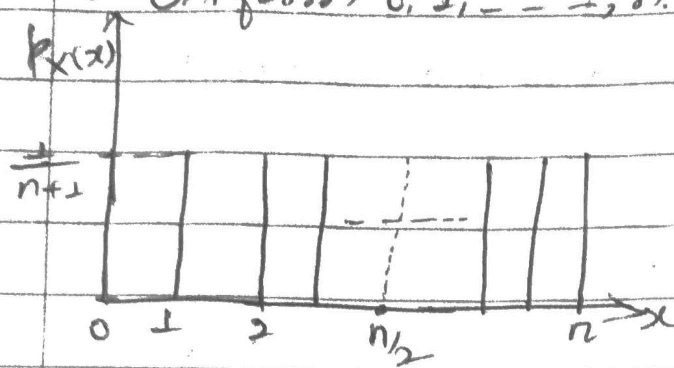
If X is the indicator of an event A, $X = I_A$:

$X = 1$ iff A occurs. $p = P(A)$

$E[I_A] = P(A)$

Expectation of a uniform r.v.

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



Def: $E[X] = \sum_x x\, p_X(x)$

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

$= n/2$

$\Rightarrow$ PMF is symmetric around $\frac{n}{2} \Rightarrow E[X] = \frac{n}{2}$
[Symmetry of PMF]

$\Rightarrow$ If we do not have symmetry of PMF, $E[X]$ turns out to be the center of gravity of PMF.

Expectation as a population average

- n students
- Weight of ith student : x_i
- Experiment: pick a student at random, all equally like
- Random variable X : weight of a selected ~~student~~
  - assume that x_i are distinct

$p_X(x_i) = P(X = x_i) = \frac{1}{n}$

$$\boxed{E[X] = \sum_i x_i \frac{1}{n} = \frac{1}{n} \sum_i x_i \longrightarrow \text{Average over a particular population}}$$

Exercise: Expectation Calculation

The PMF of the r.v. Y satisfies $p_Y(-1) = 1/6$, $p_Y(2) = 2/6$, $p_Y(5) = 3/6$, and $p_Y(y) = 0$ for all other values of y.

The expected value of Y is:

Soln: $E[Y] = \sum_y y \cdot p_Y(y)$

$$= -1 \cdot \frac{1}{6} + 2 \cdot \frac{2}{6} + 5 \cdot \frac{3}{6} + y \cdot 0$$

$$= \frac{1}{6}(-1+4+15) = 3$$

Elementary Properties of Expectations

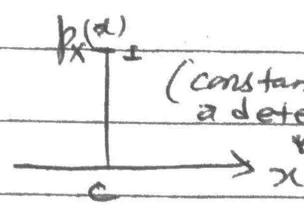
- If $X \geq 0$ then $E[X] \geq 0$ $[\forall w : X(w) \geq 0]$
- If $a \leq X \leq b$, then $a \leq E[X] \leq b$

$\rightarrow \forall w; a \leq X(w) \leq b$ $\quad E[X] = \sum_x x\, p_X(x) \geq a \underbrace{\sum_x p_X(x)}_{1}$

$\|$ly, $E[X] \leq b \Rightarrow a \leq E[X] \leq b$.

- If c is a constant, $E[c] = c$

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



(constant is a deterministic r.v.)

Exercise: Random Variables with Bounded Range

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

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

Soln → smallest value of $X-Y$ is obtained if X takes its smallest value -1 and Y takes its largest value 3 which is -4. Similarly, largest value is obtained if X takes largest value 2, and Y takes its smallest value -2 resulting in 4. → $a=-4$, $b=4$.

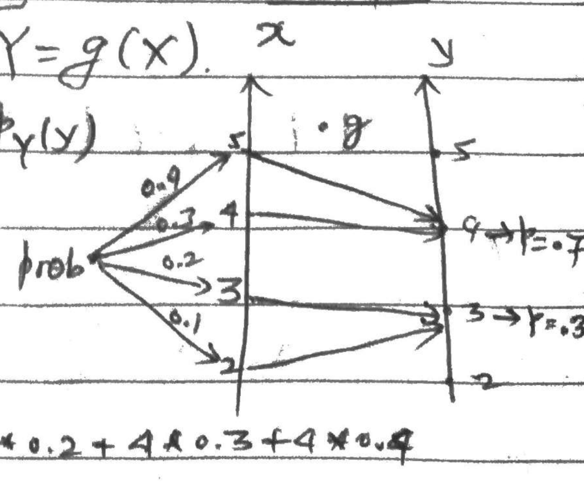
b) Can the expected value of $X+Y$ be equal to 6?

Soln → No matter what the outcome of the experiment is, the value of $X+Y$ will be at most 5, and so the expected value can be at most 5.

The Expected Value Rule

Example

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



$$E[Y]=\sum_y y\, p_Y(y) = 3(.3)+4(.7)$$
$$=$$

- Alt; Averaging over x: $3*0.1+3*0.2+4*0.3+4*0.4$

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

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

$= \sum_y \sum_{x: g(x)=y} y\, p_X(x)$ for a given $x \to g(x) = y$

$= \sum_y y \underbrace{\sum_{x: g(x)=y} p_X(x)}_{p_Y(y)} = E[Y]$

$*$ $E[X^2] = \sum_x x^2\, p_X(x)$

Caution: In general, $E[X^2] \neq (E[X])^2$

Exercise: The Expected Value Rule

Let X be a uniform r.v. on the range $\{-1, 0, 1, 2\}$.
Use the expected value rule to calculate $E[Y]$.

soln: X Uniform r.v. $\longrightarrow$ all four events are equally likely

$p_X(x) = \frac{1}{4}$ for $X = \{-1, 0, 1, 2\}$

$E[Y] = E[X^4] = \sum_x x^4\, p_X(x) = \frac{1}{4}\left((-1)^4 + (0)^4 + (1)^4 + (2)^4\right)$

$= \frac{9}{2}$

Linearity of Expectations

* $E[aX+b] = aE[X]+b$

Ex: X = salary $E[X]$ = average salary

Y = new salary $= \underbrace{2X}_{\text{(double salary)}} + \underbrace{100}_{\text{(bonus)}}$

$$E[Y] = E[2X+100] = 2E[X]+100$$

proof: $E[aX+b] = \sum_x \overbrace{(ax+b)}^{g(x)} p_X(x)$

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

$$= aE[X]+b$$

\# $E[g(X)] = g(E[X]) \rightarrow$ True iff g is a linear function.

Exercise: Linearity of Expectations

The r.v. X is known to satisfy $E[X]=2$ and $E[X^2]=7$. Find the expected value of $8-X$ and of $(X-3)(X+3)$.

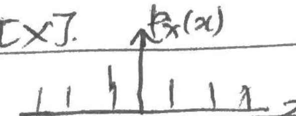
a) $E[8-X] = 8-E[X] = 1$ $[\because E[aX+b] = aE[X]+b]$

b) $E[(X+3)(X+3)] = E[X^2-9] = E[X^2]-9 = -2$

Variance; Conditioning on an event; Multiple r.v.'s

Variance — a measure of the spread of a PMF

* Random Variable X, with mean $\mu = E[X]$.



* Distance from the mean: $X - \mu$
* Average distance from the mean?

  $E[X-\mu] = E[X] - \mu = \mu - \mu = 0.$
* Definition of variance; $\text{var}(X) = E[\underbrace{(X-\mu)^2}_{g(X)}] \geq 0$
* Calculation:

$$\text{Var}(X) = E[g(X)] = \sum_x (x-\mu)^2 p_X(x)$$

Properties of Variance

$$\text{Var}(aX+b) = a^2\,\text{var}(X) + b$$

* Let $Y = X + b \rightarrow \text{Var}(Y) = E[(Y - \overbrace{E[Y]}^{\mu+b})^2]$

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

$$= E[(X-\mu)^2] = \text{var}(X)$$

$\Rightarrow$ Intuitively, adding constant to X just moves the entire PMF left ($b<0$) or right ($b>0$) w/o changing its shape.

* let $Y = aX$, $\mu \Rightarrow E[X] \Rightarrow a\mu$

  $\text{var}(Y) = E[a^2(X-\mu)^2] = a^2\,\text{var}(X)$

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

Exercise: Variance Calculation

Suppose that $var(X) = 2$. The variance of $2-3X$ is:

$var(2-3X) = +9\,var(X) = +18$ $\quad$ $(var(X) \geq 0)$

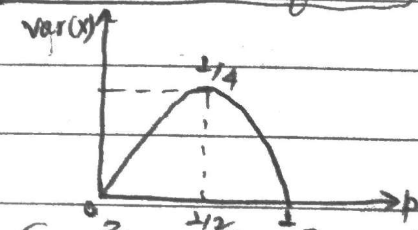
Exercise: Variance Properties

Is it always true that $E[X^2] \geq (E[X])^2$?

Sol $\Rightarrow$ $var(X) = E[X^2] - (E[X])^2 \geq 0 \Rightarrow E[X^2] \geq (E[X])^2$

Variance of the Bernoulli and the Uniform

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



$$var(X) = \sum_x (x - E[X])^2 p_X(x) = (1-p)^2 \cdot p + (0-p)^2(1-p) = p(1-$$

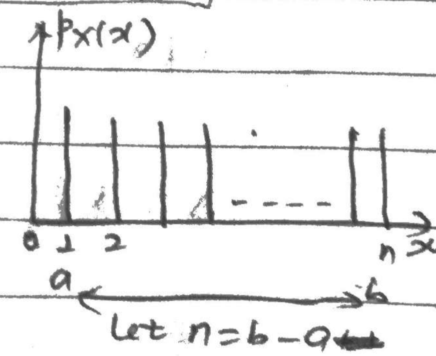
OR

$$var(X) = E[X^2] - (E[X])^2 = p(1-p)$$

Variance of a Uniform Discrete R.V.

$$var(X) = E[X^2] - (E[X])^2$$

$$= \frac{1}{n+1}(0^2 + 1^2 + 2^2 + \dots + n^2)$$



$$= \frac{n(n+2)}{12}$$

$$var(X) = \frac{(b-a)(b-a+2)}{12}$$

Exercise: Variance of the Uniform

$X \xrightarrow{r.v} \{0, 2, 4, 6, \dots, 2n\}$

Let $Y = \frac{X}{2} \xrightarrow{r.v.} \{0, 1, 2, 3, \dots, n\} \rightarrow var(\{0, 1, \dots, n\})$
$= n(n+2)/12$

$Var(X) = Var(2Y) = 4\, Var(Y) = \frac{n(n+2)}{3}$

Conditional PMF and Expectation, given an event

- Condition on an event $A \Rightarrow$ use conditional probabilities

$p_X(x) = P(X = x)$ $\qquad$ $p_{X|A} = P(X = x \mid A)$ assume $P(A) > 0$

$\sum_x p_X(x) = 1$

$\sum_x p_{X|A}(x) = 1$

$E[X \mid A] = \sum_x x\, p_{X|A}(x)$

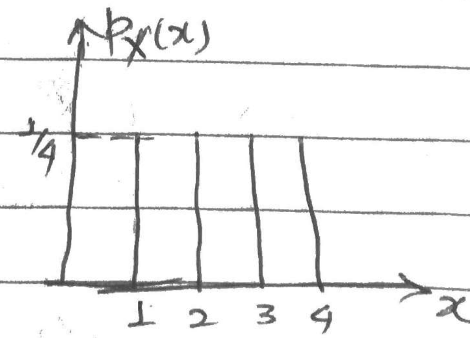
$E[g(X) \mid A] = \sum_x g(x)\, p_{X|A}(x)$

Example:

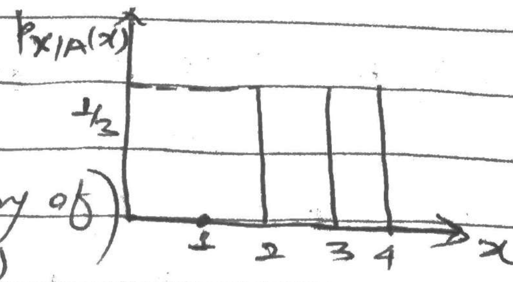
$E[X] = 2.5$ (by symmetry of PMF)

$var(X) = \frac{1}{12}(b-a)(b-a+2)$

$= \frac{5}{4}$



Let $A = \{X \geq 2\}$



$E[X|A] = 3$ (by symmetry of PMF (X/A))

$$var(X|A) = \underbrace{\frac{1}{3}(4-3)^2 + \frac{1}{3}(3-3)^2 + \frac{1}{3}(2-3)^2}_{\sum_x (x - E[X])^2 p_X(x)}$$

$$= \frac{2}{3}$$

Exercise: Conditional Variance

In the last example, we saw that the conditional distribution of X, which was a uniform over a smaller range (and in some sense, less uncertain), had a smaller variance, i.e, $Var(X|A) \leq Var(X)$. Here is an example where this is not true. Let Y be uniform on $\{0, 1, 2\}$ and B be the event that Y belongs to $\{0, 2\}$.

a) What is the variance of Y?

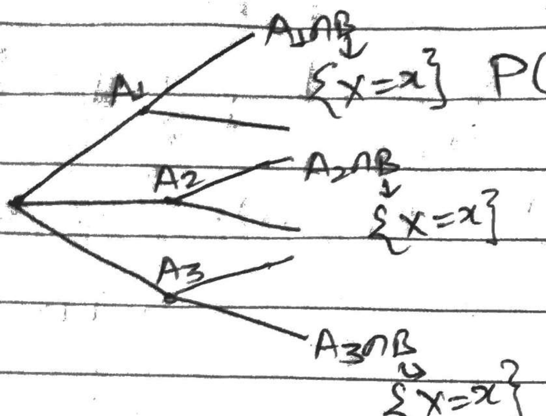
Soln → $Var(Y) = \sum_y (y - E[Y])^2 p_Y(y) = \frac{1}{3}\{(0-\overset{E[Y]}{1})^2 + (1-1)^2 + (2-1)^2\} = \frac{2}{3}$

b) In the conditional model, the conditional mean is

$E[Y|B] = \sum_y y\, p_{Y|B}(y) = \sum_y y \underbrace{p(Y=y|B)}_{1/2} = \frac{1}{2}(0+2) = 1$

$var(Y|B) = \sum_y [y - E(Y|B)]^2 p_{Y|B}(y) = \frac{1}{2}[(0-1)^2 + (2-1)^2] = 1$

Total Expectation Theorem



$$P(B) = P(A_1 \cap B) + P(A_2 \cap B) + P(A_3 \cap B)$$
$$= P(A_1)\, P(B/A_1) + P(A_2)\, P(B/A_2) + P(B/A_3)$$
$$B = \{X = x\}$$

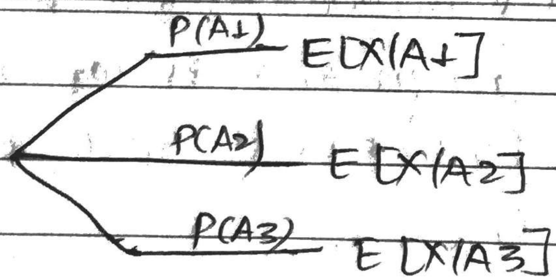
Translate into PMF Notation

$$p_X(x) = P(A_1)\, p_{X/A_1}(x) + P(A_2)\, p_{X/A_2} + \dots + P(A_n)\, p_{X/A_n}(x) \quad \forall x.$$

$$\sum_x x\, p_X(x) = P(A_1) \sum_x p_{x/A_1}(x) + \dots + P(A_n) \sum_x x\, p_{X/A_n}(x)$$

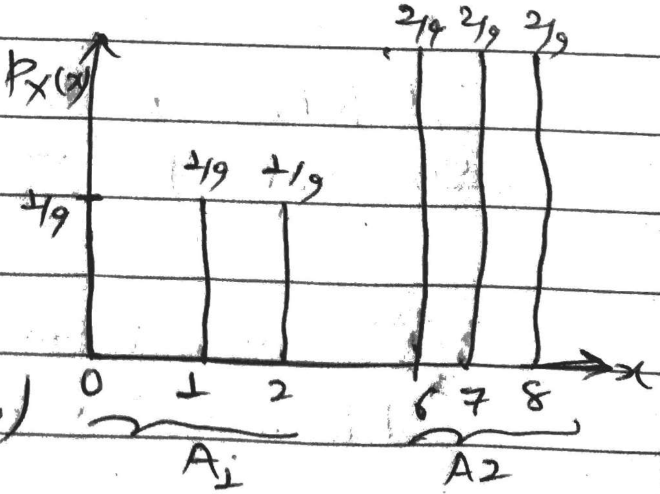
$$\Rightarrow \boxed{E[X] = P(A_1)\, E[X/A_1] + \dots + P(A_n)\, E[X/A_n]}$$

$\Downarrow$



Ex: $P(A_1) = \frac{1}{9} + \frac{1}{9} + \frac{1}{9} = \frac{1}{3}$

$P(A_2) = \frac{2}{9} + \frac{2}{9} + \frac{2}{9} = 2/3$



$E[X/A_1] = 1$ (By symmetry)

$E[X/A_2] = 7$

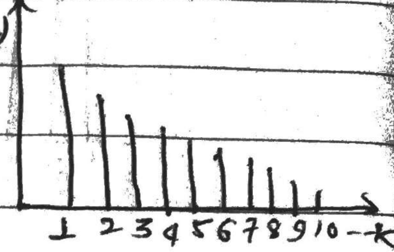
$E[X] = \frac{1}{3} \cdot 1 + \frac{2}{3} \cdot 7 = 5$

Geometric PMF, Memorylessness and Expectation

* Conditioning a Geometric Random Variable

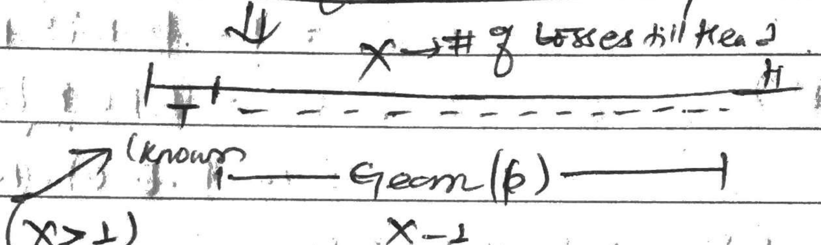
X: number of independent coin tosses until first head: $P(H) = p$

$$p_X(k) = (1-p)^{k-1} p : \quad k = 1, 2, \ldots$$



Memorylessness:

Number of remaining coin tosses, conditioned on Tails in the first toss, is Geometric with parameter "p"



⇒ Conditioned on $X > 1$, $X-1$ is geometric with parameter p.

For a coin:

eg. $P(X-1=3 \mid X>1) = P(T_2 T_3 H_4 \mid T_1) = P(T_2 T_3 H_4) = (1-p)^2 p$

$p_{X-1|X>1}(3)$

(Occurrence of 1st toss result tell nothing about remaining tosses / independence)

$= p_X(3) = P(X=3)$

We see, $p_{X-1|X>1}(3) = p_X(3)$

In general, $p_{X-1|X>1}(k) = p_X(k)$ → first n tosses were wasted

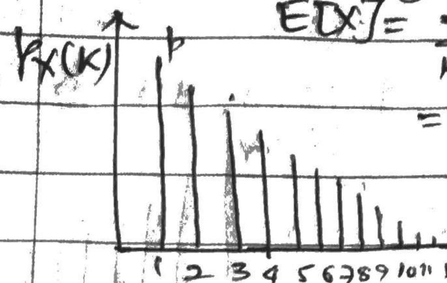
Conditioned on $X > n$, $X-n$ is geometric with parameter p.

$$\boxed{p_{X-1|X>1}(k) = p_X(k) = p_{X-n|X>n}(k)}$$

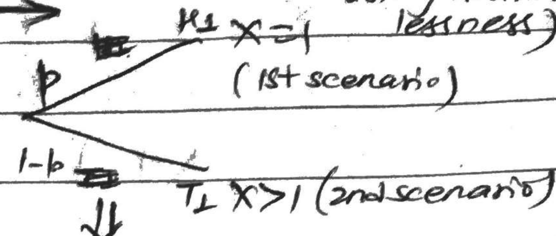
The mean of the Geometric

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

$$= \sum_{k=1}^{\infty} k(1-p)^{k-1} p$$



Com (Calculating $E(X)$ using memorylessness)



$$E[X] = \underbrace{1 + E[X-1]}_{E[1+X-1]}$$

(we've first toss nomatterwhat)

$$= 1 + p \cdot E[X-1 \mid X=1] + (1-p)\, E[X-1 \mid X>1]$$

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

$E[X-1 \mid X>1] = E[X]$

(memorylessness)

$$\Rightarrow E[X] = 1 + (1-p)\, E[X]$$

$$\Rightarrow E[X]\,[1-1+p] = 1$$

$$\Rightarrow \boxed{E[X] = \frac{1}{p}}$$

NB: $\text{var}(X) = \dfrac{1-p}{p^2}$ $\quad (X \sim \text{Geom}(p))$

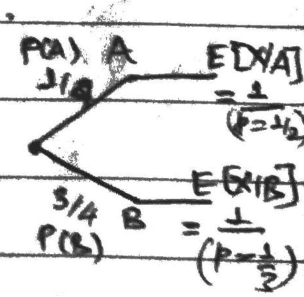
Exercise: Total Expectation Calculation

We have two coins A and B. For each toss of coin A, we obtain Heads with probability $\frac{1}{2}$; for each toss of coin B, we obtain Heads with probability $1/3$. All tosses of the same coin are independent. We select a coin at random, where the probability of selecting coin A is $1/4$, and then toss it until Heads is obtained for the first time.

The expected number of tosses until the first Heads is: 2.75 $\left(\text{or } \frac{11}{4}\right)$

Solution: Let X be the number of tosses until the first Heads. Once a coin is selected, the conditional distribution of X is geometric, with a mean of $1/p$, where p is the probability of Heads for the selected coin. Let A and B be the events of selecting coin A and B.

$$E[X] = P(A)\,E[X|A] + P(B)\,E[X|B]$$



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

$$\boxed{E[X] = \frac{11}{4}}$$

Exercise: Memorylessness of the Geometric

Let X be a geometric r.v., and assume $Var(X)=5$.

a) What is the conditional variance $Var(X-4 \mid X>4)$?

Soln: The conditional distribution of $X-4$ given $X>4$ is the same geometric PMF that describes the distribution of X. Hence $Var(X-4 \mid X>4) = Var(X) = 5$.

b) What is the conditional variance $Var(X-8 \mid X>4)$?

Soln: In the conditional model (i.e. given $X>4$), the r.v.s. $X-4$ and $X-8$ differ by a constant. Hence they have the same variance and the answer is again 5.

Joint PMFs and the Expected Value Rule

Multiple random variables and joint PMFs

$X: p_X$ $\qquad P(X=Y)=?$

$Y: p_Y$ $\qquad \to$ To answer this we need joint PMF

(Marginal PMFs) $\qquad : p_{X,Y}(x,y) = P(X=x \text{ and } Y=y)$

* $\sum_x \sum_y p_{X,Y}(x,y) = 1$

* $p_X(x) = \sum_y p_{X,Y}(x,y)$

* $p_Y(y) = \sum_x p_{X,Y}(x,y)$

More than two Random Variables

$$p_{X,Y,Z}(x,y,z) = P(X=x \text{ and } Y=y \text{ and } Z=z)$$

$$\sum_x \sum_y \sum_z p_{X,Y,Z}(x,y,z) = 1$$

$$p_X(x) = \sum_y \sum_z p_{X,Y,Z}(x,y,z)$$

$$p_{X,Y}(x,y) = \sum_z p_{X,Y,Z}(x,y,z)$$

Functions of Multiple Random Variables

$$Z = g(X,Y)$$

PMF: $p_Z(z) = P(Z=z) = P(g(X,Y)=z)$

$$= \sum_{(x,y):\, g(x,y)=z} p_{X,Y}(x,y)$$

Expected Value Rule:

$$E[g(X,Y)] = \sum_x \sum_y g(x,y)\, p_{X,Y}(x,y)$$

Exercise: Joint PMF Calculation

The r.v. V takes values in the set $\{0,1\}$ and the r.v. W takes in the set $\{0,1,2\}$. Their joint PMF is of the form $p_{V,W}(v,w) = c\cdot(v+w)$, where c is some constant, for v and w in their respective ranges, and is zero everywhere else.

a) Find the value of c.

Sol: $\sum_{v}\sum_{w} p_{V,W}(v,w) = 1$

$V \uparrow$			
1	(1,0) 1c	(1,1) 2c	(1,2) 3c
0	(0,0) 0c	(0,1) 1c	(0,2) 2c
	0	1	2
	$W \rightarrow$		

$c.(0+0) + c.(0+1) + c.(0+2) + c.(1+0) + c.(1+1) + c.(1+2) = 1$

$\Rightarrow c = \frac{1}{9}$

b) Find $p_V(1)$?

$p_V(1) = \sum_{w} p_{V,W}(1,w)$

$= p(1,0) + p(1,1) + p(1,2)$

$= 6c = \frac{6}{9}$

Exercise: Expected Value Rule

Let X and Y be discrete r.vs. For each one of the formulas below, state whether it is True or False.

a) $E[X^2] = \sum_{x} x\, p_X(x^2) \longrightarrow$ False

b) $E[X^2] = \sum_{x} x^2 p_X(x) \longrightarrow$ True

c) $E[X^2] = \sum_{x} x^2 p_{X,Y}(x) \longrightarrow$ False

d) $E[X^2] = \sum_{x} \cancel{\sum_{y}} x^2 p_{X,Y}(x,y) \longrightarrow$ False

e) $E[X^2] = \sum_x \sum_y x^2 p_{X,Y}(x,y) \longrightarrow$ True

f) $E[X^2] = \sum_z z\, p_{X^2}(z) \longrightarrow$ True $[\overset{g(x,y)}{\overset{\uparrow}{z}} = X^2]$

Linearity of Expectations and mean of the Binomial

Linearity of Expectations

* $E[aX+b] = a\,E[X]+b$

* $E[X+Y] = E[X]+E[Y]$

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

eg. $E[2X+3Y-Z] = 2E[X]+3E[Y]-E(Z)$

EX: X: binomial with parameter n, p
- number of successes in n independent trials

$$E[X] = \sum_x x\,p(x) = \sum_{k=0}^{n} k \underbrace{\binom{n}{k} p^k (1-p)^{n-k}}_{p_X(k)}$$

Use indicator Variable: $X_i = 1$ if ith trial is a success $\rightarrow p$

$X_i = 0$ otherwise. $(1-p)$

$X = X_1 + \cdots + X_n$

$E[X] = E[X_1]+\cdots+E[X_n]$ $\quad X_i \sim Ber(p)$

$= p + \cdots + p$

$= np$

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

Exercise: Linearity of Expectations Drill

Suppose that $E[X_i] = i$ for every i. Then,

$$E[X_1 + 2X_2 - 3X_3] = E[X_1] + 2E[X_2] - 3E[X_3]$$
$$= 1 + 2\cdot 2 - 3\cdot 3$$
$$= -4$$

Exercise: Using Linearity of Expectations.

We have two coins, A and B. For each toss of coin A, we obtain Heads with probability $1/2$; for each toss of coin B, we obtain Heads with probability $1/3$. All tosses of the coin are independent.

We toss coin A until Heads is obtained for the first time. ~~We~~ then toss coin B until Heads is obtained for the first time with coin B.

The expected value of total number of tosses is: $\boxed{3+2=5}$

Soln $\Rightarrow$ let X_1 and X_2 be the number of tosses of coins A and B, respectively. We know that T_A is geometric with parameter $p = 1/2$, so that $E[X_1] = 1/p = 2$. Similarly, $E[X_2] = 3$. The total number of coin tosses is $X_1 + X_2$. Using linearity $E[X_1 + X_2] = E[X_1] + E[X_2] = 2+3$
$= 5.$

Conditioning on a Random Variable; Independence of r.v.'s

* Conditional PMFs
- Conditional Expectations
- Total expectation Theorem

* Independence of r.v.'s
- Expectation properties
- Variance properties

* The Variance of the Binomial

Conditional PMFs

$p_{X|A}(x|A) = P(X=x|A)$ $\quad A = \{Y=y\}$

(conditioned on event)

$p_{X|Y}(x|y) = P(X=x|Y=y) = \dfrac{P(X=x, Y=y)}{P(Y=y)}$

(conditioned on another r.v.)

$p_{X|Y}(x|y) = \dfrac{p_{X,Y}(x,y)}{p_Y(y)}$ defined for y such that $p_Y(y) > 0$

$$\boxed{\sum_x p_{X|Y}(x|y) = 1}$$

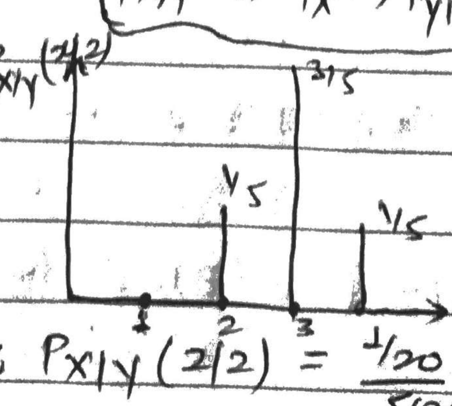
$p_{X,Y}(x,y) = p_Y(y)\, p_{X|Y}(x|y)$

$p_{X,Y}(x,y) = p_X(x)\, p_{Y|X}(y|x)$

$y \backslash x$	1	2	3	4
4	1/20	2/20	2/20	0
3	2/20	4/20	1/20	2/20
2	0	1/20	3/20	1/20
1	0	1/20	0	0

let $Y = 2$

$p_Y(2) = \dfrac{5}{20}$



$p_{X|Y}(1|2) = 0$; $p_{X|Y}(2|2) = \dfrac{1/20}{5/20} = \dfrac{1}{5}$

$p(x|Y=2)$

0 $\quad$ 1/20 $\quad$ 3/20 $\quad$ 1/20

$\xrightarrow{\text{rescale}}$ 0 $\quad$ 4/20 $\quad$ 12/20 $\quad$ 4/20

Conditional PMFs involving more than two r.v.'s

- Self-explanatory notation

$$p_{X|Y,Z}(x|y,z) = P(X=x|Y=y, Z=z)$$

$$= \frac{P(X=x, Y=y, Z=z)}{P(Y=y, Z=z)} = \frac{p_{X,Y,Z}(x,y,z)}{p_{Y,Z}(y,z)}$$

- Multiplication Rule

$$P(A\cap B\cap C) = P(A)\,P(B|A)\,P(C|A\cap B)$$

$$A=\{X=x\} \quad B=\{Y=y\} \quad C=\{Z=z\}$$

$$p_{X,Y,Z}(x,y,z) = p_X(x)\,p_{Y|X}(y|x)\,p_{Z|X,Y}(z|x,y)$$

Exercise: Conditional PMFs

For each of the formulas below, state whether it is True or False.

a) $p_{X,Y,Z}(x,y,z) = p_Y(y)\,p_{Z|Y}(z|y)\,p_{X|Y,Z}(x|y,z) \rightarrow$ True

b) $p_{X,Y|Z}(x,y|z) = p_X(x)\,p_{Y|Z}(y|z) \longrightarrow$ False

$\boxed{p_{X,Y,Z}(x,y,z) \neq p_X(x)\,p_{Y,Z}(y,z)}$

c) $\sum_x p_{X,Y|Z}(x,y|z) = 1 \longrightarrow$ True

d) $\sum_x p_{X,Y|Z}(x,y|z) = 1 \longrightarrow$ False

e) $\sum_x \sum_y p_{X,Y|Z}(x,y|z) = 1 \rightarrow$ True

f) $p_{X,Y|Z}(x,y|z) = \dfrac{p_{X,Y,Z}(x,y,z)}{p_Z(z)}$ $\longrightarrow$ True

g) $p_{X|Y,Z}(x|y,z) = \dfrac{p_{X,Y,Z}(x,y,z)}{p_{Y,Z}(y,z)}$ $\longrightarrow$ True

Conditional Expectation and the Total Expectation Theorem

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

$$A = \{Y = y\}$$

$$E[X|A] = \sum_x x\, p_{X|A}(x) \qquad E[X|Y=y] = \sum_x x\, p_{X|Y}(x|y)$$

- Expected value rule

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

$$E[g(X)|A] = \sum_x g(x)\, p_{X|A}(x)$$

$$E[g(X)|Y=y] = \sum_x g(x)\, p_{X|Y}(x|y)$$

- Total probability and Expectation Theorems

* $A_1, \ldots, A_n$: partition of Ω
* $p_X(x) = P(A_1)\, p_{X|A_1}(x) + \cdots + P(A_n)\, p_{X|A_n}(x)$

$$Y = \{y_1, \ldots, y_n\}, \quad A_i = \{Y = y_i\}$$

$$p_X(x) = \sum_y p_Y(y)\, p_{X|Y}(x|y)$$

* $E[X] = P(A_1)\, E[X|A_1] + \cdots + P(A_n)\, E[X|A_n] \overset{?}{=} \sum_y p_Y(y)\, E[X|Y=y]$

Exercise: The Expected Value Rule With Conditioning.
For each of the formulas below, state whether it is True or False.

1) $E[g(X,Y)|Y=2] = \sum_x g(x,y)\, p_{X,Y}(x,y) \longrightarrow$ False

2) $E[g(X,Y)|Y=2] = \sum_x g(x,y)\, p_{X,Y}(x,2) \longrightarrow$ False

3) $E[g(X,Y)|Y=2] = \sum_x g(x,2)\, p_{X,Y}(x,2) \longrightarrow$ False

4) $E[g(X,Y)|Y=2] = \sum_x g(x,2)\, p_{X|Y}(x|2) \longrightarrow$ True

5) $E[g(X,Y)|Y=2] = \sum_x g(x,2) \frac{p_{X,Y}(x,2)}{p_Y(2)} \rightarrow$ True

6) $E[g(X,Y)|Y=2] = \sum_x \sum_y g(x,y)\, p_{X,Y|Y}(x,y|2)$
$\rightarrow$ True

$p_{X,Y}(x,y|2) = 0 \quad \forall y \neq 2.$

Independence of Random Variables.

- Independence of two events: $P(A \cap B) = P(A).P(B)$
  $P(A|B) = P(A)$
- (Independence) of a r.v. and an event: $P(X=x \text{ and } A) = P(X=x).P(A)$ $\forall x.$
  $P_{X|A}(x) = P_X(x), \forall x.$
  $P_{A|X}(A) = P(A) \ \forall x.$
- (Independence) of two r.v.s': $P(X=x \text{ and } Y=y)$
  $= P(X=x).P(Y=y) \ \forall x,y$

$P_{X|Y}(x|y) = P_X(x)$ $\quad$ $P_{X,Y}(x,y) = P_X(x)P_Y(y)$
$P_{Y|X}(y|x) = P_Y(y)$ $\quad$ $\forall\, x, y.$

X, Y, Z are independent if:
$p_{X,Y,Z}(x,y,z) = p_X(x)\,p_Y(y)\,p_Z(z) \;\forall\, x,y,z.$

Exercise: Independence

Let X, Y, and Z be discrete r.vs.

a) Suppose that Z is identically equal to 3, i.e. $P(Z=3)=1$. Is X guaranteed to be independent of Z?

Soln: Since Z is deterministic, the value of Z does not provide any information, and, so intuitively, we have independence. For a formal argument, suppose that $z \neq 3$. Then $p_{X,Z}(x,z) = 0 = p_X(x)\,p_Z(z)$. And for $z=3$, $p_{X,Z}(x,3) = P(X=x, Z=3) = P(X=x)$ $= P(X=x)\cdot 1 = p_X(x)\,p_Z(3)$, so the definition of independence is satisfied.

b) Would either of the following be an appropriate definition of independence of the pair (X,Y) from Z?

- $p_{X,Y,Z}(x,y,z) = p_X(x)\,p_Y(y)\,p_Z(z) \;\forall\, x,y,z \longrightarrow$ No
  (It forces X & Y to be independent also)
- $p_{X,Y,Z}(x,y,z) = \underbrace{p_{X,Y}(x,y)}_{\text{pair}(X,Y)}\,\underbrace{p_Z(z)}_{Z} \;\forall\, x,y,z \longrightarrow$ Yes.

c) Suppose that X, Y, Z are independent. Is it true that X and Y are independent.

soln → Intuitively, X, Y, Z are independent means none of the r.vs provide information about the others.

$(X,Y),(Y,Z),(Z,X)$ pairswise independence $\not\Rightarrow$ independence X, Y, Z

but X, Y, Z independence $\Rightarrow (X,Y),(Y,Z),(Z,X)$ are pairswise independent.

d) Suppose that X, Y, Z are independent. Is is true that (X, Y) is independent from Z?

⟶ Yes: $p_{X,Y,Z}(x,y,z) = p_X(x)\, p_Y(y)\, p_Z(z) = p_{X,Y}(x,y)\, p_Z(z)$

Exercise: A Criterion for Independence

Suppose that the conditional PMF of X, given Y = y, is the same for every y for which $p_Y(y) > 0$. Is this enough to guarantee independence?

soln Yes: The condition given means that when I tell you the value of Y, the conditional PMF of X will be the same. Thus, the value of Y makes no difference, and, intuitively, we have independe[nce]

For a formal argument, let $c(x) = p_{X|Y}(x|y)$: we can define $c(x)$ this way (w/o a dependence on y) since we are assuming that $p_{X|Y}(x|y)$ is the same for all y. $p_{X|Y}(x,y) = p_Y(y)\, p_{X|Y}(x|y) = p_Y(y)\, c(x)$

Summing over all y we obtain

$$p_X(x) = \sum_y p_{X,Y}(x,y) = \sum_y p_Y(y)\, c(x) = c(x)$$

Therefore, $p_X(x) = c(x)$. It follows that $p_{X,Y}(x,y) = p_{X|Y}(x|y)\, p_Y(y) = c(x)\, p_Y(y) = p_X(x)\, p_Y(y)$ which establishes independence.

Independence and Expectations

- In general, $E[g(X,Y)] \neq g(E[X], E[Y])$
- Exceptions: $E[aX+b] = aE[X]+b$

  $E[X+Y+Z] = E[X]+E[Y]+E[Z]$

* If X, Y are independent: $E[XY] = E[X] \cdot E[Y]$

* $g(X)$ and $h(Y)$ also independent:

$$E[g(X)\, h(Y)] = E[g(X)]\; E[g(Y)]$$

Exercise: Independence and Expectations.

Let X and Y be independent positive discrete r.vs. For each of the following statements, determine whether it is true (that is, always true) or false (that is not guaranteed to be always true).

1. $E[X/Y] = E[X]/E[Y] \longrightarrow$ False
2. $E[X/Y] = E[X]\, E[1/Y] \rightarrow$ True $\left[\because XY = X \cdot \frac{1}{Y}\right]$

Independence, Variances and the Binomial Variance

- Always true: $var(aX) = a^2 var(X)$; $Var(X+a) = var(X)$
- In general: $Var(X+Y) \neq Var(X) + Var(Y)$

If X and Y are independent,

Assume $E[X] = E[Y] = 0$ $\boxed{Var(X+Y) = var(X) + var(Y)}$

$$var(X+Y) = E[X^2 + 2XY + Y^2]$$
$$= E[X^2] + E[Y^2] + 2E[XY]$$
$$= Var(X) + var(Y) - \underset{0}{\overset{\shortparallel}{}} \rightarrow \text{since } E[X].E[Y] = 0$$

$$\boxed{Var(X+Y) = var(X) + var(Y) \text{ if } E[X] = E[Y] = 0 \text{ and } X, Y \text{ are independent}}$$

Ex:
- If $X = Y$: $var(X+Y) = var(2X) = 4\, var(X)$
- If $X = -Y$; $Var(X+Y) = var(0) = 0$
- If X, Y independent: $var(X - 3Y) = var(X) + 9\, var(Y)$

Variance of the Binomial

- X: binomial parameters w/ n, p
- number of successes in n independent trials.

$X_i = 1$ if ith trial is a success (indicator variable)

$X_i = 0$ otherwise

$X = X_1 + X_2 + \cdots + X_n$ (sum of independent r.v.)

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

Exercise: Independence and Variances

The pair of r.vs. (X,Y) is equally likely to take any of the four pairs of values $(0,1)$, $(1,0)$, $(-1,0)$, $(0,-1)$. Note that X and Y each have zero mean.

a) Find $E[XY]$.

Soln → $p_{XY}(x,y) = \frac{1}{4} \quad \forall (x,y)$ pairs

$$E[XY] = \sum_x \sum_y xy \, p_{XY}(x,y)$$

product of all xy terms are zero

b) For this pair of r.vs (X,Y), is it true that $Var(X+Y) = Var(X) + Var(Y)$?

Soln → Yes! Since $E[X]=0$, $E[Y]=0 \Rightarrow E[X+Y]=0$

$$Var(X+Y) = E[(X+Y)^2] - (E[X+Y])^2$$
$$= E[X^2 + 2XY + Y^2]$$
$$= E[X^2] + E[Y^2] + 2\cancel{E[X]E[Y]}^{\,0}$$
$$= Var(X) + Var(Y)$$

c) We know that if X and Y are independent, then $Var(X+Y) = Var(X) + Var(Y)$. Is the converse true? That is, does the condition $Var(X+Y) = Var(X) + Var(Y)$ imply independence?

Soln → Let $Var(X+Y) = Var(X) + Var(Y)$. But these r.vs are not independent. For example, the information that $X=1$ tells us that the value of Y must be zero.

The hat problem

- n people throw their hats in a box and then pick one at random.
  - All permutations equally likely $\rightarrow \frac{1}{n!}$
  - Equivalent to picking one hat at a time

people	hats
1	1
2	2
3	3

$\frac{1}{3} \cdot \frac{1}{2} \cdot 1 = \frac{1}{3!}$

- X: number of people who get their own hat
  - Find $E[X]$

$$X_i = \begin{cases} 1 & \text{if selects own hat} \\ 0 & \text{otherwise} \end{cases}$$

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

- $E[X_i] = E[X_1]$ (symmetric w.r.to all people)
  $= P(X_1 = 1) = \frac{1}{n}$

$$E[X] = E[X_1] + \cdots + E[X_n] = n \cdot \frac{1}{n} = 1$$

eg. $n=2$, (two people)

$X_1 = 1 \Rightarrow X_2 = 1$ (1st gets his own hat back) $\rightarrow$ (definitely 2nd gets his hat back too)

$X_1 = 0 \Rightarrow X_2 = 0$

(1st gets 2nd person hat) (2nd gets 1st person hats back)

$\Rightarrow$ r.vs. $X_1, X_2, \ldots, X_n$ dependent.

$$Var(X) = E[X^2] - (E[X])^2$$

$$X^2 = \overbrace{\sum_i X_i^2}^{n} + \overbrace{\sum_{i,j\; i\neq j} X_i X_j}^{n(n-1) \to \text{cross terms}}$$

$X_1 \sim$ either 0 or 1

$X_1^2 \sim$ either 0 or 1

- $E[X_i^2] = E[X_1^2]$ [X_i^2 has the same distribution as X_1^2]

$$= E[X_1]$$

$$= \frac{1}{n}$$

$\sum_{x_1,x_2} x_1 x_2\, p(x_1,x_2)$ → only exist $1, 1$

- For $i \neq j$: $E[X_iX_j] = E[X_1X_2] = P(X_1X_2 = 1) = P(X_1=1, X_2=1)$

$\Rightarrow X_1 = 1 \;\&\; X_2 = 1$

$$= P(X_1=1)\, P(X_2=1 \mid X_1=1)$$

$$= \frac{1}{n}\,\frac{1}{n-1}$$

$$E[X^2] = n \cdot \frac{1}{n} + n(n-1)\frac{1}{n(n-1)} = 1+1 = 2$$

$$Var(X) = E[X^2] - (E[X])^2 = 2-1 = 1$$

Exercise: The hat problem

Consider the problem with $n=100$. What is the expected value of $X_3X_6X_7$?

soln → By symmetry this is same as $E[X_1X_2X_3]$. Since the product $X_1X_2X_3$ is either zero or one, this is same as, $P(X_1X_2X_3=1)$ [$E(X_1X_2X_3) = \sum_{x_1,x_2,x_3} x_1x_2x_3\, p(x_1,x_2,x_3)$ only exist if $x_1=1, x_2=1, x_3=1$]

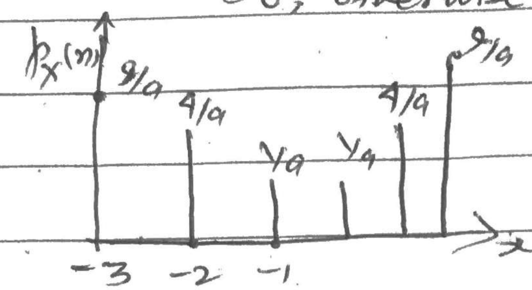
$= P(X_1=1)\, P(X_2=1 \mid X_1=1)\, P(X_3=1 \mid X_1=1, X_2=1)$

$= 1/10 \cdot 1/9 \cdot 1/8 = 1/720 = 0.00139$

Solve Problem: 1

Consider a r.v. X such that $p_X(x) = \begin{cases} x^2/a, & \text{for } x \in \{-3,-2,-1,1,2,3\} \\ 0, & \text{otherwise} \end{cases}$

where $a>0$ is a real parameter.



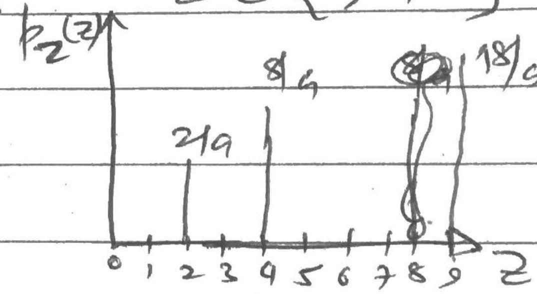
1. Find a.

$$\sum_x p_X(x) = 1 \Rightarrow a = 28$$

2. Find PMF of $Z = X^2$.

soln $\{Z=1\} = \{X=-1\} \cup \{X=1\}$, $Z \in \{9, 4, 1\}$

$$P(\{Z=1\}) = P(\{X=-1\}) + P(\{X=1\})$$
$$= 1/a + 1/a = 2/a$$



$$P(\{Z=4\}) = 8/a$$
$$P(\{Z=9\}) = 18/a \Rightarrow p_Z(k) = \begin{cases} 2/28, & k=1 \\ 9/28, & k=4 \\ 18/28, & k=9 \\ 0, & \text{otherwise.} \end{cases}$$

Conditional Independence 1

4/4 points (graded)

Suppose that we have a box that contains two coins:

1. A fair coin: $\mathbf{P}(H) = \mathbf{P}(T) = 0.5$.
2. A two-headed coin: $\mathbf{P}(H) = 1$.

A coin is chosen at random from the box, i.e. either coin is chosen with probability $1/2$, and tossed twice. Conditioned on the identity of the coin, the two tosses are independent.

Define the following events:

- Event A: first coin toss is H.
- Event B: second coin toss is H.
- Event C: two coin tosses result in HH.
- Event D: the fair coin is chosen.

For the following statements, decide whether they are true or false.

1. A and B are independent.

- ◯ True
- ◉ False

✓

2. A and C are independent.

- ◯ True
- ◉ False

✓

3. A and B are independent given D.

- ◉ True
- ◯ False

4. A and C are independent given D.

- ◯ True
- ◉ False

✓

Solution:

1. False. Since we do not know whether it is a fair coin or the two-headed one when the coin is being tossed, getting a Heads during one toss increases our belief the the coin is the two-headed one, so that also increases our belief that the other toss also results in a Heads. Or we can also verify by definition: $\mathbf{P}(A \cap B) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot 1 = \frac{5}{8} \neq \frac{9}{16} = \frac{3}{4} \cdot \frac{3}{4} = \mathbf{P}(A)\mathbf{P}(B)$.

2. False. $\mathbf{P}(A \cap C) = \mathbf{P}(C) \neq \mathbf{P}(A)\mathbf{P}(C)$.

3. True. Conditioned on D, A and B becomes the outcome of two independent fair coin tosses.

4. False. $\mathbf{P}(A \cap C|D) = \mathbf{P}(C|D) \neq \mathbf{P}(A|D)\mathbf{P}(C|D)$.

Conditional Independence 2

2/2 points (graded)

1. Suppose three random variables X, Y, Z have a joint distribution

$$\mathbf{P}_{X,Y,Z}(x, y, z) = \mathbf{P}_X(x)\,\mathbf{P}_{Z|X}(z \mid x)\,\mathbf{P}_{Y|Z}(y \mid z).$$

Then, X and Y are independent given Z.

- (selected) True
- False

✔

2. Suppose random variables X and Y are independent given Z, then the joint distribution must be of the form

$$\mathbf{P}_{X,Y,Z}(x, y, z) = h(x, z)\, g(y, z),$$

where h, g are some functions.

- (selected) True
- False

Solution:

1. True. Using $\mathbf{P}_{X,Y,Z}(x,y,z)=\mathbf{P}_X(x)\,\mathbf{P}_{Z|X}(z|x)\,\mathbf{P}_{Y|Z}(y|z)$, we have

$$\begin{aligned}
\mathbf{P}_{X,Y|Z}(x,y|z) &= \frac{\mathbf{P}_{X,Y,Z}(x,y,z)}{\mathbf{P}_Z(z)} \\
&= \frac{\mathbf{P}_X(x)\,\mathbf{P}_{Z|X}(z|x)\,\mathbf{P}_{Y|Z}(y|z)}{\mathbf{P}_Z(z)} \\
&= \frac{\mathbf{P}_X(x)\,\mathbf{P}_{Z|X}(z|x)}{\mathbf{P}_Z(z)}\mathbf{P}_{Y|Z}(y|z) \\
&= \mathbf{P}_{X|Z}(x|z)\,\mathbf{P}_{Y|Z}(y|z)\,,
\end{aligned}$$

   which shows X and Y are conditionally independent given Z.

2. True. Since X and Y are conditionally independent given Z, we have

$$\begin{aligned}
\mathbf{P}_{X,Y,Z}(x,y,z) &= \mathbf{P}_Z(z)\,\mathbf{P}_{X,Y|Z}(x,y|z) \\
&= \mathbf{P}_Z(z)\,\mathbf{P}_{X|Z}(x|z)\,\mathbf{P}_{Y|Z}(y|z) \\
&= h(x,z)\,g(y,z)\,,
\end{aligned}$$

   by letting $h(x,z):=\mathbf{P}_Z(z)\,\mathbf{P}_{X|Z}(x|z)$, $g(y,z):=\mathbf{P}_{Y|Z}(y|z)$. (In fact, by generalizing the argument for the part 1, we can show X and Y are conditionally independent given Z if and only if $\mathbf{P}_{X,Y,Z}(x,y,z)=h(x,z)\,g(y,z)$ for some h,g.)

Variance of Difference of Indicators

2.0/2.0 points (graded)

Let A be an event, and let I_A be the associated indicator random variable (I_A is 1 if A occurs, and zero if A does not occur). Similarly, let I_B be the indicator of another event, B. Suppose that $P(A) = p$, $P(B) = q$, and $P(A \cap B) = r$.

Find the variance of $I_A - I_B$, in terms of p, q, r.

$\text{Var}(I_A - I_B) =$ p+q-2*r-(p-q)^2

$p + q - 2 \cdot r - (p-q)^2$

✔ **Answer:** p-2*r+q-(p-q)^2

STANDARD NOTATION

Solution:

$$\begin{aligned}
\text{Var}(I_A - I_B) &= \mathbf{E}\left[(I_A - I_B)^2\right] - \left(\mathbf{E}\left[(I_A - I_B)\right]\right)^2 \\
&= \mathbf{E}\left[I_A^2 - 2I_A I_B + I_B^2\right] - \left(\mathbf{E}\left[I_A\right] - \mathbf{E}\left[I_B\right]\right)^2 \\
&= \mathbf{E}\left[I_A^2\right] - 2\mathbf{E}\left[I_A I_B\right] + \mathbf{E}\left[I_B^2\right] - \left(\mathbf{E}\left[I_A\right]\right) - \left(\mathbf{E}\left[I_B\right]\right)^2 \\
&= \mathbf{E}\left[I_A\right] - 2\mathbf{E}\left[I_A I_B\right] + \mathbf{E}\left[I_B\right] - \left(\mathbf{E}\left[I_A\right]\right) - \left(\mathbf{E}\left[I_B\right]\right)^2 \\
&= P(A) - 2P(A \cap B) + P(B) - \left(P(A) - P(B)\right)^2 \\
&= p - 2r + q - (p-q)^2
\end{aligned}$$

6.

Mid Term due Oct 6, 2020 19:59 EDT Completed

Bookmark this page

For all problems on this page, use the following setup:

Let N be a positive integer random variable with PMF of the form

$$p_N(n) = \frac{1}{2} \cdot n \cdot 2^{-n}, \qquad n = 1, 2, \ldots.$$

Once we see the numerical value of N, we then draw a random variable K whose (conditional) PMF is uniform on the set $\{1, 2, \ldots, 2n\}$.

Joint PMF

1/1 point (graded)

Write down an expression for the joint PMF $p_{N,K}(n,k)$.

For $n = 1, 2, \ldots$ and $k = 1, 2, \ldots, 2n$:

$p_{N,K}(n,k) =$ 2^(-(n+2)) ✔

STANDARD NOTATION

Solution:

We are given that:

$$p_{K|N}(k \mid n) = \frac{1}{2n}, \qquad k = 1, 2, \ldots, 2n. \tag{7.2}$$

By definition:

$$p_{N,K}(n,k) = p_{K|N}(k \mid n)\, p_N(n) = \frac{1}{2n}\frac{1}{2} \cdot n \cdot 2^{-n} = \left(\frac{1}{2}\right)^{n+2}, \qquad n = 1, 2, \ldots, \quad k = 1, 2, \ldots, 2n \tag{7.3}$$

Submit You have used 1 of 3 attempts

Show Answer

Answers are displayed within the problem

Marginal Distribution

1.5/1.5 points (graded)

Find the marginal PMF $p_K(k)$ as a function of k. For simplicity, provide the answer **only for the case when k is an even number**. (The formula for when k is odd would be slightly different, and you do not need to provide it).

Hint: You may find the following helpful: $\sum_{i=0}^{\infty} r^i = \frac{1}{1-r}$ for $0 < r < 1$.

For $k = 2, 4, 6, \ldots$:

$p_K(k) =$ 1/2^((k/2) + 1) ✔ **Answer:** (1/2)^(k/2+1)

$$\frac{1}{2^{\left(\frac{k}{2}\right)+1}}$$

STANDARD NOTATION

Solution:

Solution

Observe that in the infinite sum $p_K(k) = \sum_{n=1}^{\infty} p_{N,K}(n,k)$ only the terms from $n = k/2$ and above have non-zero probability. Indeed, $K = k = 4$ has probability 0 if $n < k/2 = 4/2 = 2$.

Hence:

$$
\begin{aligned}
p_K(k) &= \sum_{n=k/2}^{\infty} p_{N,K}(n,k) = \sum_{n=k/2}^{\infty} \left(\frac{1}{2}\right)^{n+2} \\
&= \sum_{n=k/2}^{\infty} \left(\frac{1}{2}\right)^{n+2} = \frac{1}{4} \sum_{n=k/2}^{\infty} \left(\frac{1}{2}\right)^{n} \\
&= \frac{1}{4}\left[\sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^{n} - \sum_{0}^{k/2-1} \left(\frac{1}{2}\right)^{n}\right] \\
&= \frac{1}{4}\left[\frac{1}{1-\frac{1}{2}} - \frac{1-\left(\frac{1}{2}\right)^{k/2-1+1}}{1-\frac{1}{2}}\right] \\
&= \left(\frac{1}{2}\right)^{k/2+1} \qquad \text{for} k = 2, 4, \ldots
\end{aligned}
$$

Submit

You have used 1 of 3 attempts

Show Answer

Answers are displayed within the problem

Discrete PMFs

2/2 points (graded)

Let A be the event that K is even. Find $P(A|N=n)$ and $P(A)$.

1. You have opened a timed exam with a **48 hours** time limit. Please use the timer to see the time remaining. If you had opened this exam too close to the exam **closing time, October 5 23:59UTC** , you will not have the full 48 hours, and the exam will close at the closing time.

2. This is an **open book exam** and you are allowed to refer back to all course material and use (online) calculators. However, **you must abide by the honor code and not ask for answers directly from any aide** .

3. As part of the honor code, you **must not share the exam content** with anyone in any way, i.e. **no posting of exam content anywhere on the internet** . Violators will be removed from the course. **If you find exam content anywhere outside of the this course site, you must still abide by the honor code and not refer to or use it in any way.**

4. You will be given **no feedback** during the exam. This means that unlike in the problem sets, you will not be shown whether any of your answers are correct or not. This is to test your understanding, to prevent cheating, and to encourage you to try your very best before submitting. Solutions will be available after the exam closes.

5. You will be given **3 attempts** for each (multipart) problem. Since you will be given no feedback, the extra attempts will be useful only in case you hit the "submit" button in a haste and wish to reconsider. **With no exception, your last submission will be the one that counts** . **DO NOT FORGET TO SUBMIT your answers to each question** . The "end your exam" button will **not** submit answers for you.

6. The exam will only be **graded 1 day after the due date** , and the **Progress Page will be disabled while the exam is open** .

7. **Error and bug reports:** While the exam is open, you are **not allowed to post on the discussion forum on anything related to the exam, except to report bugs/platform difficulties** . If you think you have found a bug, please **state on the forum only what needs to be checked on the forum** . You can still post questions relating to course material, but **the post must not comment on the exam** , and in particular **must not shed any light on the contents or concepts in the exam** . **Violators will receive a failing grade or a grade reduction in this exam** .

Let A be the event that K is even. Find $P(A|N=n)$ and $P(A)$.

$P(A \mid N=n) =$ 1/2 ✔ **Answer:** 1/2

$\frac{1}{2}$

$P(A) =$ 1/2 ✔ **Answer:** 1/2

$\frac{1}{2}$

STANDARD NOTATION

Solution:

Let A be then event that K is even. We need to check whether $P(A \mid N=n) = P(A)$ is true for the event A to be independent of N. Now because $p_{K|N}(k \mid n)$ is uniform over the $2n$-size set $\{1, 2, \ldots, 2n\}$ and there are exactly n even numbers in this set, we have that:

$$P(A \mid N=n) = \frac{n}{2n} = \frac{1}{2}, \qquad n \geq 1. \tag{7.4}$$

Intuitively, knowledge of n does not affect the beliefs about A, and we have independence. A full, formal argument goes as follows:

$$P(A) = \sum^{\infty} P(A \mid N=n)\, P(N=n)$$

$$= \frac{1}{2} \sum_{n=1}^{\infty} P(N = n) = \frac{1}{2},$$

where the last step follows because PMFs always sum to 1. So, $P(A \mid N = n) = P(A)$, for all n.
Equivalently, $P(A \text{ and } N = n) = P(A \mid N = n) \cdot P(N = n) = P(A) \cdot P(N = n)$, for all n, which is the defining property of independence.

Submit You have used 3 of 3 attempts

Show Answer

Answers are displayed within the problem

Independence 2

0.5/0.5 points (graded)

Is the event A independent of N?

- (selected) yes
- no
- not enough information to determine

1.

Mid Term due Oct 6, 2020 19:59 EDT Completed

Bookmark this page

True or False

4/4 points (graded)

Let A, B, and C be events associated with the same probabilistic model (i.e., subsets of a common sample space), and assume that $P(C) > 0$.

For each one of the following statements, decide whether the statement is True (always true), or False (not always true).

1. Suppose that $A \subset C$. Then, $P(A \mid C) \geq P(A)$.

- [x] True
- [ ] False

✔

2. Suppose that $A \subset B$. Then, $P(A \mid C) \leq P(B \mid C)$.

- [x] True
- [ ] False

3. Suppose that $P(A) \leq P(B)$. Then, $P(A \mid C) \leq P(B \mid C)$.

- ( ) True
- (•) False

✔

4. Suppose that $A \subset C$, $B \subset C$, and $P(A) \leq P(B)$. Then, $P(A \mid C) \leq P(B \mid C)$.

- (•) True
- ( ) False

✔

Solution:

1. Suppose that $A \subset C$. Then, $P(A \mid C) \geq P(A)$. This is **TRUE**:

$$P(A \mid C) = \frac{P(A \cap C)}{P(C)} = \frac{P(A)}{P(C)} \geq P(A), \tag{7.1}$$

since $P(C) \leq 1$.

2. Suppose that $A \subset B$. Then, $P(A \mid C) \leq P(B \mid C)$. This is **TRUE**.

$$P(A \mid C) = \frac{P(A \cap C)}{P(C)} \leq \frac{P(B \cap C)}{P(C)} = P(B \mid C)$$

where the inequality follows from $A \cap C \subset B \cap C$.

3. Suppose that $P(A) \leq P(B)$. Then, $P(A \mid C) \leq P(B \mid C)$.This is **FALSE**, with the following counter example:
Suppose that A and B are disjoint events with positive probability and that $C = A$. Then, $P(A \mid C) = P(A) > 0$, whereas $P(B \mid C) = 0$.

4. Suppose that $A \subset C, B \subset C$, and $P(A) \leq P(B)$. Then, $P(A \mid C) \leq P(B \mid C)$. This is **TRUE**:
Since $A, B \subset C$, we have $P(A \mid C) = \dfrac{P(A)}{P(C)}$ and similarly $P(B \mid C) = \dfrac{P(B)}{P(C)}$. Then, $P(A) \leq P(B)$ implies $P(A \mid C) \leq P(B \mid C)$.

Submit

You have used 3 of 3 attempts

Show Answer

Answers are displayed within the problem

2.

Mid Term due Oct 6, 2020 19:59 EDT Completed

Bookmark this page

A Drunk Person at the Theater

4.0/4.0 points (graded)

There are n people in line, indexed by $i = 1, \ldots, n$, to enter a theater with n seats one by one. However, the first person ($i = 1$) in the line is drunk. This person has lost her ticket and decides to take a random seat instead of her assigned seat. That is, the drunk person decides to take any one of the seats 1 to n with equal probability. Every other person $i = 2, \ldots, n$ that enters afterwards is sober and will take his assigned seat (seat i) unless his seat i is already taken, in which case he will take a random seat chosen uniformly from the remaining seats.

Suppose that $n = 3$. What is the probability that person 2 takes seat 2?

(Enter a fraction or a decimal accurate to at least 3 decimal places.)

2/3 ✔ **Answer:** 2/3

Suppose that $n = 5$. What is the probability that person 3 takes seat 3?

(Enter a fraction or a decimal accurate to at least 3 decimal places.)

3/4 ✔ **Answer:** 3/4

STANDARD NOTATION

Solution:

1. $\mathbf{P}(\text{Person 2 takes seat 2}) = \mathbf{P}(\text{Person 1 takes seat 1 or 3}) = \frac{2}{3}$.

2. $$\begin{aligned} & \mathbf{P}(\text{Person 3 takes seat 3}) \\ = \; & \mathbf{P}(\text{Person 1,2 does not take seat 3}) \\ = \; & \mathbf{P}(\text{Person 1 takes seat 1 or 4 or 5}) \cdot 1 + \mathbf{P}(\text{Person 1 takes seat 2}) \cdot \mathbf{P}(\text{Person 2 does not take seat 3}) \\ = \; & \frac{3}{5} \cdot 1 + \frac{1}{5} \cdot \frac{3}{4} = \frac{3}{4}. \end{aligned}$$

Submit You have used 3 of 3 attempts

Show Answer

Answers are displayed within the problem

Mid Term due Oct 6, 2020 19:59 EDT *Completed*

Bookmark this page

Expectation 1

1/1 point (graded)

Compute $\mathbf{E}(X)$ for the following random variable X:

$$X = \text{Number of tosses until getting 4 (including the last toss) by tossing a fair 10-sided die.}$$

$\mathbf{E}(X) =$ 10 ✓ **Answer:** 10

Solution:

This is just the mean of a geometric random variable with parameter $1/10$. Hence, $\mathbf{E}(X) = 10$.

Submit You have used 1 of 3 attempts

Show Answer

Answers are displayed within the problem

Expectation 2

1.3333333333333333/2.0 points (graded)

Compute $\mathbf{E}(X)$ for the following random variable X:

$$X = \text{Number of tosses until all 10 numbers are seen (including the last toss) by tossing a fair 10-sided die.}$$

To answer this, we will use induction and follow the steps below:

Let $\mathbf{E}(i)$ be the expected number of additional tosses until all 10 numbers are seen (including the last toss) **given i distinct numbers have already been seen**.

1. Find $\mathbf{E}(10)$.

   $\mathbf{E}(10) =$ 0 ✔ **Answer:** 0

2. Write down a relation between $\mathbf{E}(i)$ and $\mathbf{E}(i+1)$. Answer by finding the function $f(i)$ in the formula below.

   For $i = 0, 1, \ldots, 9$:

$$\mathbf{E}(i) = \mathbf{E}(i+1) + f(i)$$

   where $f(i) =$ 10/(10-i) **Answer:** 10/(10-i)

   $\frac{10}{10-i}$

3. Finally, using the results above, find $\mathbf{E}[X]$.

(Enter an answer accurate to at least 1 decimal place.)

$\mathbf{E}[X] =$ 29.289682539682538 ✔ **Answer:** 29.28968

Solution:

Recall $\mathbf{E}(i)$ is the expected number of additional tosses until all 10 numbers are seen (including the last toss) given i distinct numbers have already been seen.

1. $\mathbf{E}(10) = 0$

2. The induction step is as follows. For $i = 1, \ldots 9$:

$$\begin{aligned} \mathbf{E}(i) &= (\mathbf{E}(i) + 1) \times \frac{i}{10} + (\mathbf{E}(i+1) + 1) \times \left(1 - \frac{i}{10}\right) \\ \iff \mathbf{E}(i) &= \mathbf{E}(i+1) + \frac{10}{10 - i}. \end{aligned}$$

Using $\mathbf{E}(10) = 0$ and the induction step, we have

$$\mathbf{E}(0) = \frac{10}{10} + \frac{10}{9} + \ldots + \frac{10}{2} + \frac{10}{1} + 0 \approx 29.28968.$$

1. You have opened a timed exam with a **48 hours** time limit. Please use the timer to see the time remaining. If you had opened this exam too close to the exam **closing time, October 5 23:59UTC** , you will not have the full 48 hours, and the exam will close at the closing time.

2. This is an **open book exam** and you are allowed to refer back to all course material and use (online) calculators. However, **you must abide by the honor code and not ask for answers directly from any aide** .

3. As part of the honor code, you **must not share the exam content** with anyone in any way, i.e. **no posting of exam content anywhere on the internet** . Violators will be removed from the course. **If you find exam content anywhere outside of the this course site, you must still abide by the honor code and not refer to or use it in any way.**

4. You will be given **no feedback** during the exam. This means that unlike in the problem sets, you will not be shown whether any of your answers are correct or not. This is to test your understanding, to prevent cheating, and to encourage you to try your very best before submitting. Solutions will be available after the exam closes.

5. You will be given **3 attempts** for each (multipart) problem. Since you will be given no feedback, the extra attempts will be useful only in case you hit the "submit" button in a haste and wish to reconsider. **With no exception, your last submission will be the one that counts** . **DO NOT FORGET TO SUBMIT your answers to each question** . The “end your exam" button will **not** submit answers for you.

6. The exam will only be **graded 1 day after the due date** , and the **Progress Page will be disabled while the exam is open** .

7. **Error and bug reports:** While the exam is open, you are **not allowed to post on the discussion forum on anything related to the exam, except to report bugs/platform difficulties** . If you think you have found a bug, please **state on the forum only what needs to be checked on the forum** . You can still post questions relating to course material, but **the post must not comment on the exam** , and in particular **must not shed any light on the contents or concepts in the exam** . **Violators will receive a failing grade or a grade reduction in this exam** .

1.

Mid Term due Oct 6, 2020 19:59 EDT *Completed*

Bookmark this page

True or False

4/4 points (graded)

Let A, B, and C be events associated with the same probabilistic model (i.e., subsets of a common sample space), and assume that $P(C) > 0$.

For each one of the following statements, decide whether the statement is True (always true), or False (not always true).

1. Suppose that $A \subset C$. Then, $P(A \mid C) \geq P(A)$.

- (x) True
- ( ) False

2. Suppose that $A \subset B$. Then, $P(A \mid C) \leq P(B \mid C)$.

- (x) True
- ( ) False