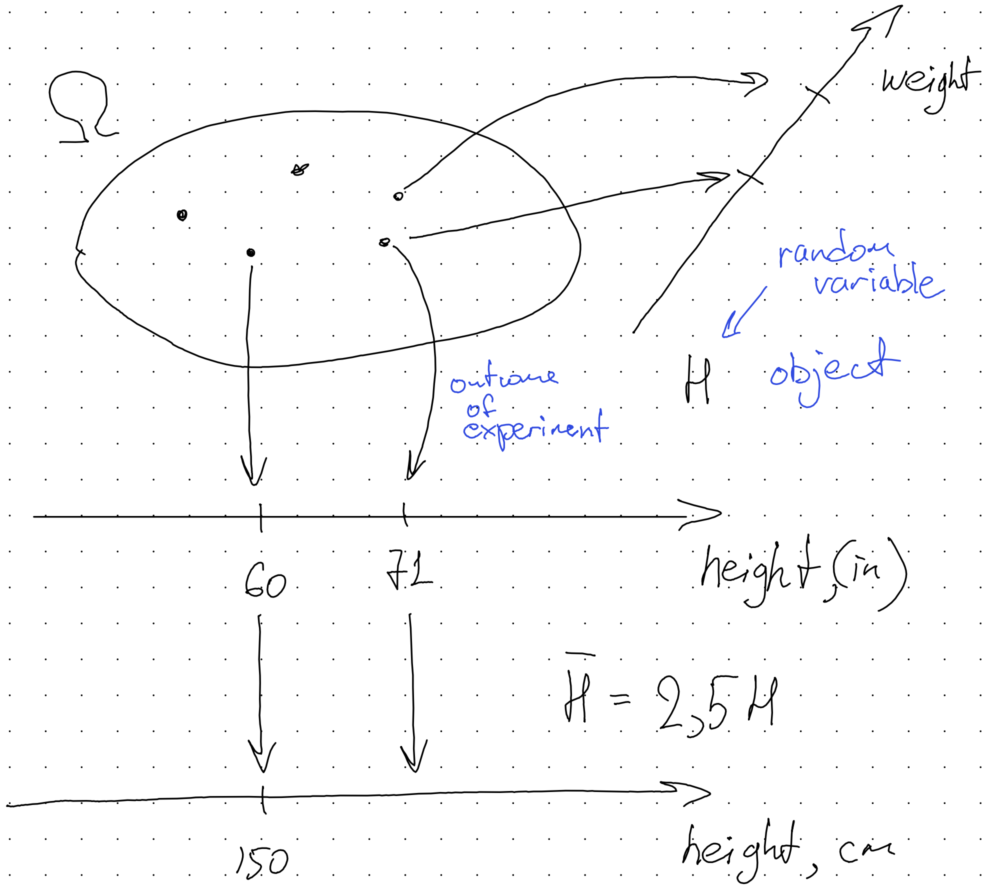


Random variable

- an assignment of a value (number) to every possible outcome
- Mathematically: A function from the SS Ω to the real number

$$RV \begin{cases} \nearrow \text{descrete} \\ \searrow \text{continuous} \end{cases}$$

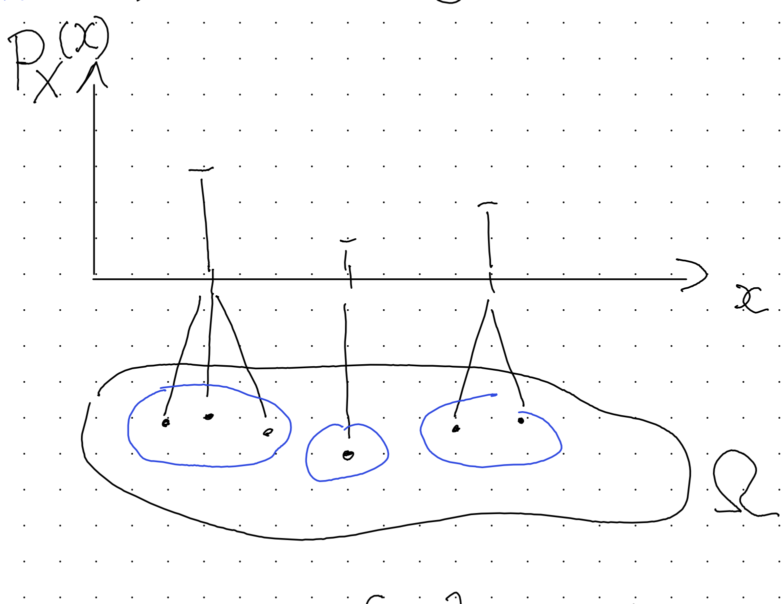
if we round height $\Rightarrow$ descrete

if measure in $\mathbb{R}^+$ $\Rightarrow$ continuous

RV X (function) $\Omega \to \mathbb{R}^+$

numerical value $x \in \mathbb{R}^+$

(PMF) - probability mass function



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

RV

$P_X(x) \geq 0$ $\qquad$ $\underbrace{\sum_x P_X(x) = 1}_{\text{from additivity and normalization axioms}}$

Example

X – number of coin tosses until first head

H $\qquad X = 1$

T H $\qquad X = 2$

$\underbrace{T\,T \ldots T}_{k-1}\,H \qquad X = k$

$$P_X = (1-p)^{k-1} \cdot p$$

Bernoulli Random Variable

Bernoulli RV takes values $(1, 0)$

$$X = \begin{cases} 1 & \text{if a head} \\ 0 & \text{if a tail} \end{cases}$$

Its PMF is

$$P_X(k) = \begin{cases} p & \text{if } k = 0 \\ 1 - p & \text{if } k = 0 \end{cases}$$

Binomial PMF

- X : number of heads in n independent coin tosses
- $P(H) = p$
- let $n = 4$

$$P_X(2) = P(HHTT) + P(HTHT) + P(HTTH) + P(THTH) + P(THHT) + P(TTHH)$$

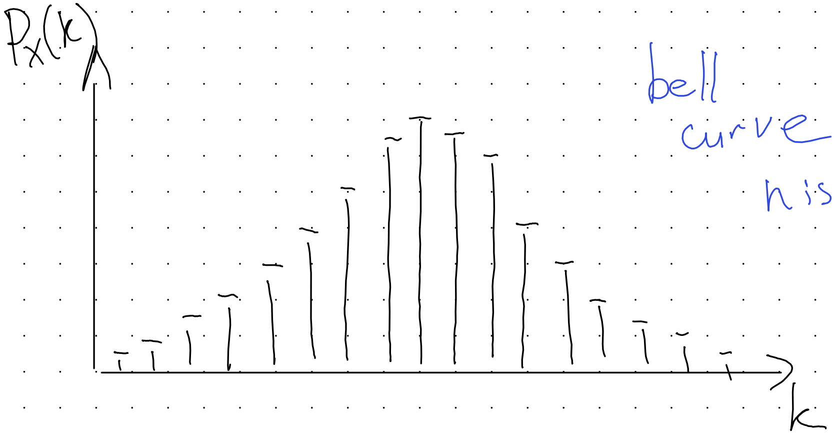
combinations $\binom{4}{2} = \frac{4!}{2!\,2!} = \frac{\cancel{4}^{2} \cdot 3}{\cancel{2}} = 6$

$$= \binom{4}{2} p^2 (1-p)^2$$

General:

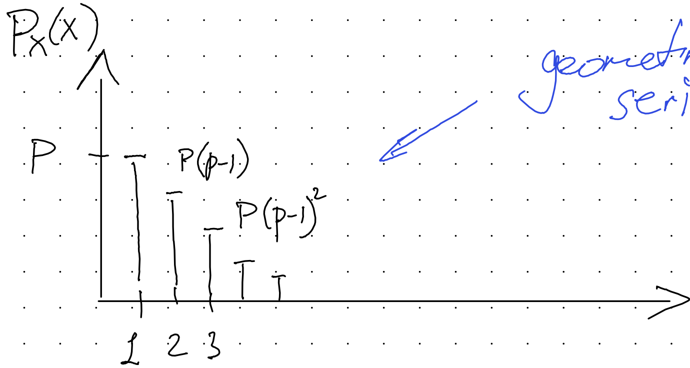
$$P_X(k) = \binom{n}{k} p^k (1-p)^{n-k} \qquad k = 1, 2, \dots n$$

$$\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = 1$$ - normalization property



bell curve
n is big

Geometric RV



geometric series

Geometric random v. is number X of tosses needed for a head to come up for the first time

$$P_X(k) = (1-p)^{k-1}\, p$$

head

$$\sum_{k=1}^{\infty} P_X(k) = \sum_{k=1}^{\infty} (1-p)^{k-1} p = p \sum_{k=0}^{\infty} (1-p)^k = p \frac{1}{1-(1-p)} = 1$$

tosses is just an insight, we can interpret geometric RV as repeated independent trials until success

pass exam 10% $p = 0.1$ $\quad P_X(1) = 0.9^0 \cdot 0.1$

$P_X(3) = 0.9^2 \cdot 0.1$

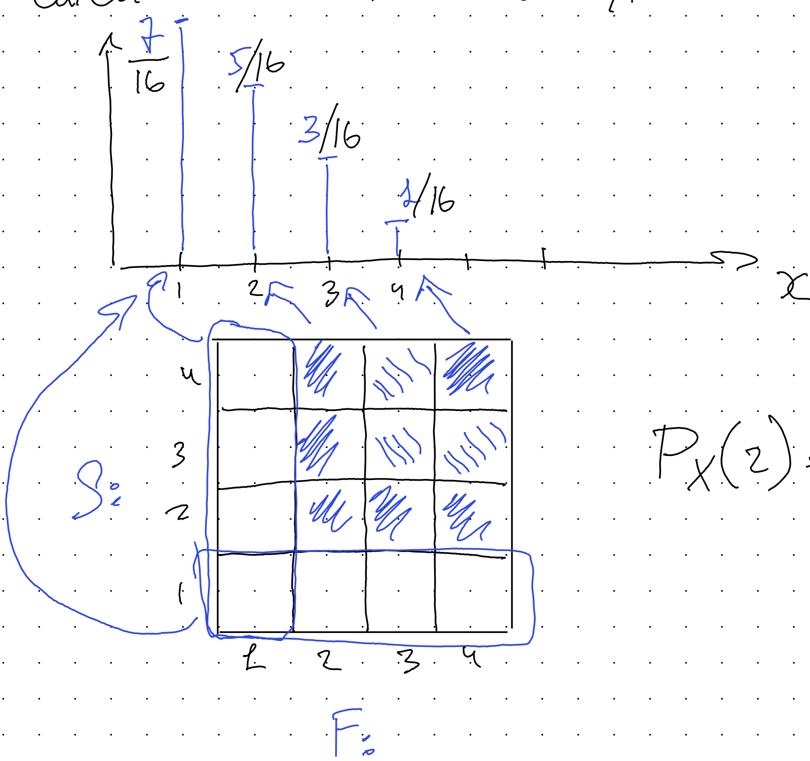
Example 2

Experiment two independent rolls tetrahedral die

F: outcome of first throw
S: outcome of second throw

X: min (F, S)

calculate PMF of X

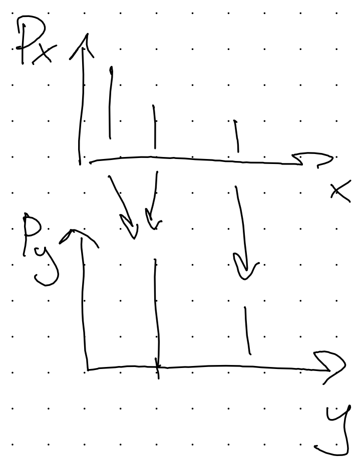


$$P_X(2) = \frac{5}{16}$$

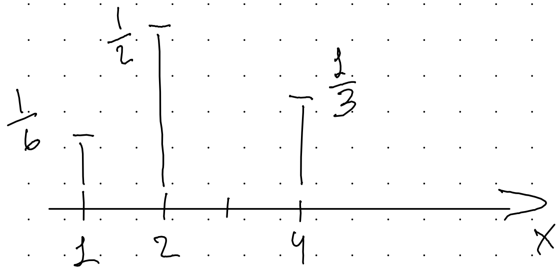
Poisson RV

Function of Random var.

$$P_y(y) = \sum_{\{x \mid g(x) = y\}} P_x(x)$$



Expected value of RV



Play the game multiple times.
Average payoff if we think of
probability as frequency

$$\frac{1}{6}\cdot 1 + \frac{1}{2}\cdot 2 + \frac{1}{3} 4 = 2{,}5$$

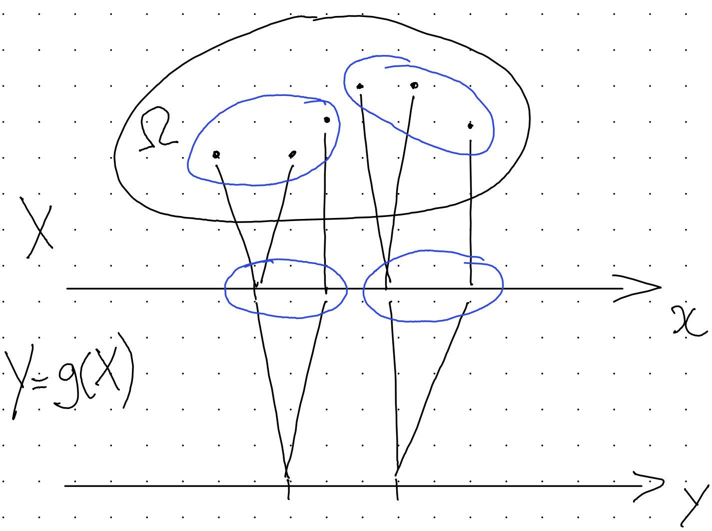
* Expectation is summary of all
PMF values. Weighted (in proportion of
probability) average of the possible
values of X

Expected value of RV X (also mean or expectation)

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

Interpretations:

- Center of gravity of PMF
- Average in large number of repetitions of the experiment



$$E[Y] = \sum_y y \cdot P_Y(y)$$

$$E[Y] = \sum_x g(x)\, P_X(x)$$ (easy)

Caution

in general: $E[g(X)] \neq g(E[X])$

$=$ if g is linear

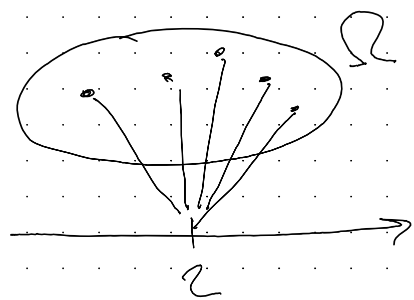
caution

Properties: if α, β are constants

$E[\alpha] = \alpha$ $\qquad$ $E[2] = 2$

$E[\alpha X] = \alpha E[X]$

example
measure student
instead of (in)
in cm



The average value
in experiment
that allways gives
2 is 2

$$E[\alpha X + \beta] = E[\alpha X] + \beta = \alpha E[X] + \beta$$

linearity

How far away from mean?

number

$$E[X - E[X]] = E[X] - E[E[X]] =$$

$$= E[X] - E[X] = 0$$

on average, signed distance from mean is zero.

Hence we have variance

Variance

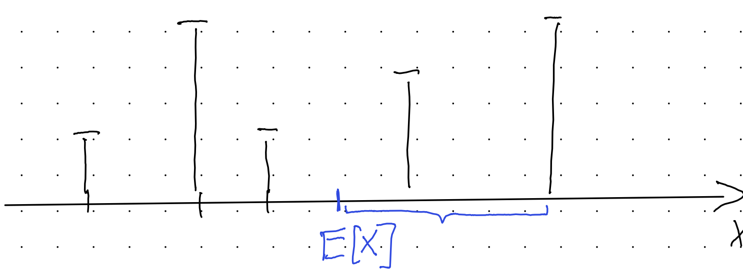
second moment of RV

$$E[X^2] = \sum_x x^2 P_X(x)$$

$$\operatorname{var}(X) = E[(X - E[X])^2] = \sum_x (x - E[X])^2 P_X(x) =$$

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

the problem with units



varience is measure of dispersion of X around its mean.

Properties

$$\mathrm{var}(X) \geq 0$$

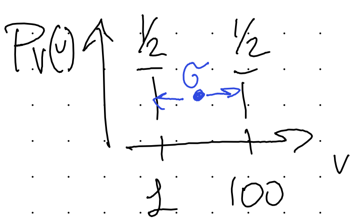
$$\mathrm{var}(\alpha X + \beta) = \alpha^2 \,\mathrm{var}(X)$$

Standard deviation

$$\sigma(X) = \sqrt{\mathrm{var}(X)}$$

Example:

traverse 200 mile distance at constant but random speed V



$$E[V] = 1 \cdot \frac{1}{2} + 100 \frac{1}{2} = 50{,}5$$

$$\mathrm{var}(V) = E[V^2] - E[V]^2 = \left(1^2 \cdot \frac{1}{2} + (100)^2 \frac{1}{2}\right) - (50{,}5)^2 =$$

$$= 5000{,}5 - 2550{,}25 = 2450{,}25$$

$$\sigma_V = \sqrt{2450{,}25} \approx 50$$

$$T = t(V) = \frac{200}{V}$$

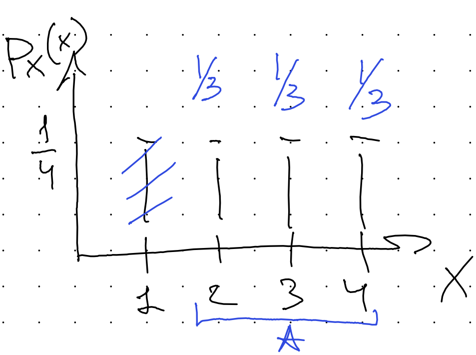
$$E[T] = E[t(V)] = \sum_v t(v)\, p_V(v) =$$

$$= \frac{1}{2}200 - \frac{1}{2}2 = 101$$

$$E[TV] = 200 \neq E[T] \cdot E[V]$$

$$E[200/V] = E[T] \neq 200/E[V]$$

Conditional PMF and expectation



Let $A = X \geq 2$

$$p_{X|A}(x) = P(X = x \mid A)$$

Event happened

* you can argue by gravity $E[X|A] = 3$

$$p_{X|A}(x) = \frac{1}{3} \quad x = 2, 3, 4$$

$$E[X|A] = \sum_{x} x \, P_{X|A}(x)$$

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

Geometric PMF

first person

Y count until Head appears | TT...TH $\quad P_Y(k) = (1-p)^{k-1} \cdot p$

$k \geq 0$

second person

$X \quad \underbrace{TT \;|\; TT\ldots TH}_{X}$

$X > 2$ (condition)

Memoryless property: Given that $X > 2$ the r.v. $X-2$ has same geometric PMF

Conditional PMF

$$p_{X|A}(x) = P(X=x \mid A)$$

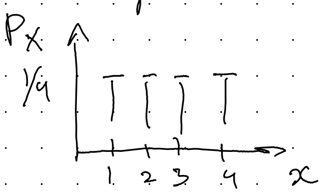
assume $P(A) > 0$

Conditional PMF like ordinary PMF $\Rightarrow$

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

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

Example



$E[X] = 2.5$

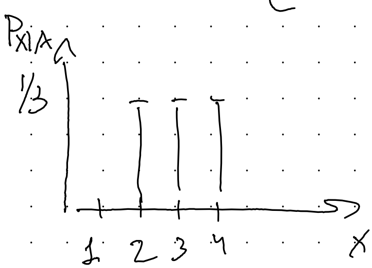
! you can apply formula for uniform distribution from lecture

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

$$= \frac{1}{4}(1+4+9+16) - \frac{25}{4} =$$

$$= \frac{30-25}{4} = \frac{5}{4}$$

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



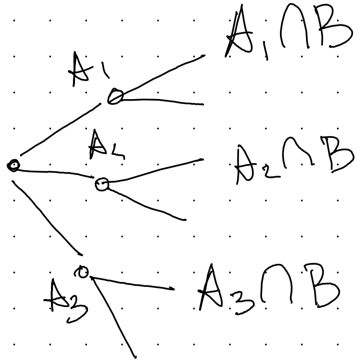
$$E[X \mid A] = 3$$

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

$$= \frac{1}{3}(4+9+16) - 9 = \frac{29}{3} - \frac{27}{3} = \frac{2}{3}$$

Total expectation theorem

Reminder of total prob. theorem



$$P(B) = P(A_1)P(B|A_1) + \dots + P(A_n)P(B|A_n)$$

Let's bring RV to the picture:

$$B = \{X = x\}$$

total probability theorem translated to PMF notation

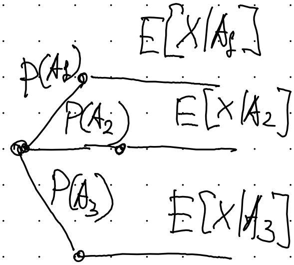
we replaced B to $\{X = x\}$

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

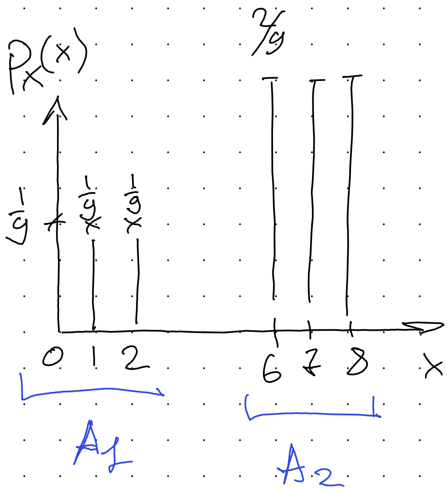
conditional pmfs

true for all x $\Rightarrow$

$$\underbrace{\sum_x x\, p_X(x)}_{E[X]} = P(A_1) \underbrace{\sum_x x\, p_{X|A_1}(x)}_{E[X|A_1]} + \dots +$$



$$E[X] = \sum_{i=1}^{n} \underbrace{P(A_i)\, E[X|A_i]}_{\text{weighted linear combination}}$$



$$P(A_1)=\frac{1}{3}$$

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

$$E[X|A_1]=1$$

* by middle point

$$E[X|A_2]=7$$

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

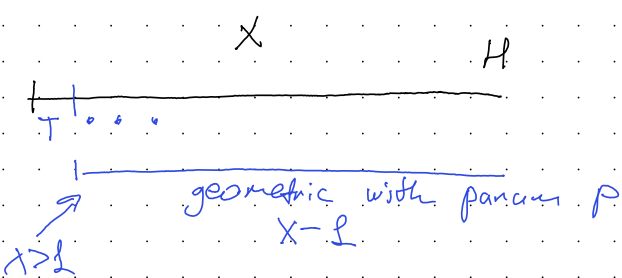
Geometric example

Example

X : number of indep. coin tosses until first Head $\quad P(H)=p$

$$p_X(x)=(1-p)^{k-1}p \qquad k=1\ldots n$$

Introduce memorylesness property. Past coin tosses doesn't affect future tosses



Remaining number of tosses

X−1

$$p_{X-1|X>1}(k)=P(X-1=3|X>1)=P(T_2T_3H_4|T_1)=P(T_2T_3H_4)=$$
$$=(1-p)^2p=p_X(3)$$

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

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

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

$A_1 = \{X = 1\}$ $\quad A_2 = \{X > 1\}$

$$E[X] = \underset{p}{P(A_1)}\, E[X | A_1] + \overset{(1-p)}{P(A_2)}\, \underbrace{E[X | A_2]}$$

p

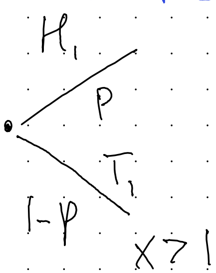
become 1
if known X=1
then X become
a number
Expectation of
number is the number

first toss

num of
remaining
tosses
That's E[X]

Expected
value of
remaining
coin flips
given first was T

$$E[X | X>1] = E[X-1 | X-1>0] + 1$$



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

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

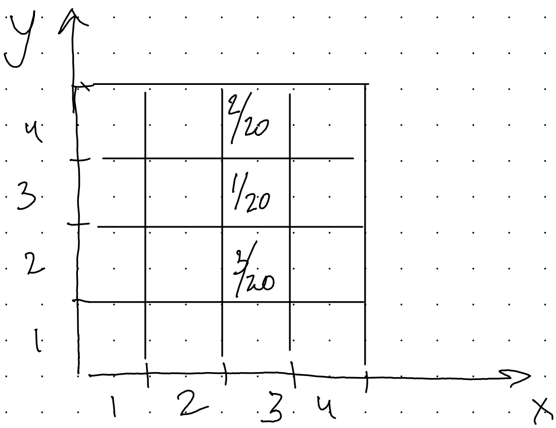
$$E[X](1 - (1-p)) = 1$$

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

Joint PMFs

$$P_{X,Y}(x,y) = P(X=x \text{ and } Y=y)$$

assotiation between two R.V



- $$\sum_x \sum_y P_{X,Y}(x,y) = 1$$

- $$P_X(x) = \sum_y P_{X,Y}(x,y)$$

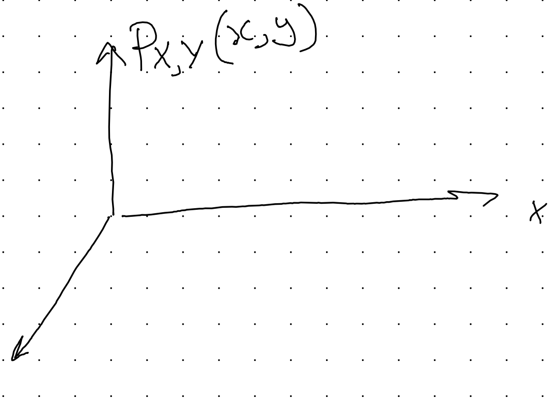
- $$P_{X|Y}(x|y) = P(X=x \mid Y=y) = \frac{P_{X,Y}(x,y)}{P_Y(y)}$$

Example

transmitter sending messages

X - travel time Y : message length

$$P_Y(y) = \begin{cases} 5/6 & \text{if } y = 10^2 \\ 1/6 & \text{if } y = 10^4 \end{cases}$$



$$p_X(x) = \sum_y P_{X,Y}(x,y) = \sum_y P_Y(y)\, P_{X|Y}(x,y)$$

$\boxed{\text{Lecture 7}}$ Conclude Ch 2

$$P_X(x) = P(X=x) \quad \text{Marginal}$$

$$P_{X,Y}(x,y) = P(X=x, Y=y) \quad \text{joint}$$

$$P_{X|Y}(x|y)$$

$$P_X(x) = \sum_y P_{X,Y}(x,y)$$

$$P_{X,Y}(x,y) = P_X(x)\, P_{Y|X}(y|x)$$

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

From multiplication rule

What happens if we have more RVs

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

Multiplication rule for 3 RV

$$P_{XYZ}(x,y,z) = P_X(x)\, P_{Y|X}(y|x)\, P_{Z|X,Y}(z|x,y)$$

Independence

RV X, Y, Z are independent $\Longleftrightarrow$

$$P_{X,Y,Z}(x,y,z) = P_X(x) \cdot P_Y(y)\, P_Z(z)$$

$$\forall\ x, y, z \in \mathcal{X}, \mathcal{Y}, \mathcal{Z}$$

$$P_{X|Y}(x|y) = P_X(x)$$

this definition make sense when conditional var. well defined

$P_Y(y) > 0$

Expectation

shortcut for expectation

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

Exceptions

linearity

$$E[\alpha X + \beta] = \alpha E[X] + \beta$$

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

- if X, Y independent

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

$$= E[X]\, E[Y]$$

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

Variance

- $E[((ax - E[ax]))^2] = var(aX) =$

$$= a^2 \, var(X)$$

- $Var(X + a) = Var(X)$

Let $z = X + Y$

$$var(X + Y) = var(X) + var(Y)$$

Binomial mean and variance

$X =$ # of success in n indep. trials

$$E[X] = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k}$$

Messy

messy and difficult to calc

Let's use indicator RV

$$X_i = \begin{cases} 1 & \text{if success in trial } i \\ 0 & \text{otherwise} \end{cases}$$

ith flip

$$\sum_i X_i = X$$

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

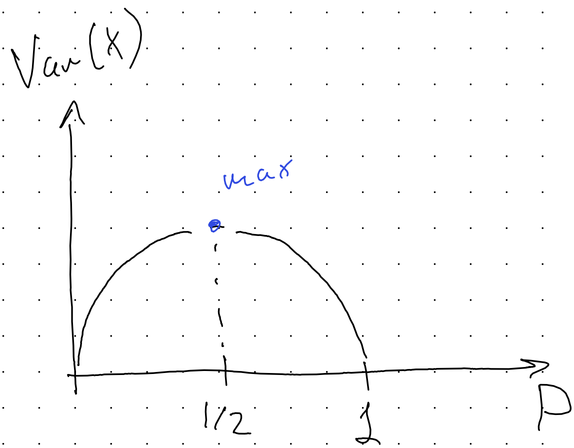
$$E[X] = n \cdot p$$

$$\text{var}(X_i) = (1-p)^2 p + (0-p)^2 (1-p)$$

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

$Var(X) = np(1-p)$

* for indep of RV varance of sum is sum of var-es



coin flips are more uncertain when coin is fair

Example The hat problem

n people throw their hats in a box and then pick one at random

X - number of people who get their own hat

Find $E[X]$

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

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

$$P(X_i = 1) = 1/n$$

$$E[X_i] = 1 \cdot \frac{1}{n} + 0 \cdot \left(1 - \frac{1}{n}\right) = \frac{1}{n}$$

RVs are dependent but

Expectations are linear disregard to dependence of events =>

$$E[X] = \sum_{i=1}^{n} E[X_i] = 1$$

Calculating variance

Sum of variances are not the same to variance of RV for dependent events (what about indepependent events?)

$$\mathrm{Var}(X) = E[X^2] - (E[X])^2 = E[X^2] - 1$$

$$X^2 = \left(\sum X_i\right)^2$$

$$X^2 = \sum_i X_i^2 + \overbrace{\sum_{ij\ i \neq j} X_i X_j}^{n^2 - n \text{ terms}}$$

$$E[X_i^2] = E[X_i^2] = \frac{1}{n}$$

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

$$= \frac{1}{n} \cdot \frac{1}{n-1} = E[X_i X_j]_{i \neq j}$$

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

$$\mathrm{Var}(X) = 2 - 1 = 1$$