17.10.2022

YZV 231E

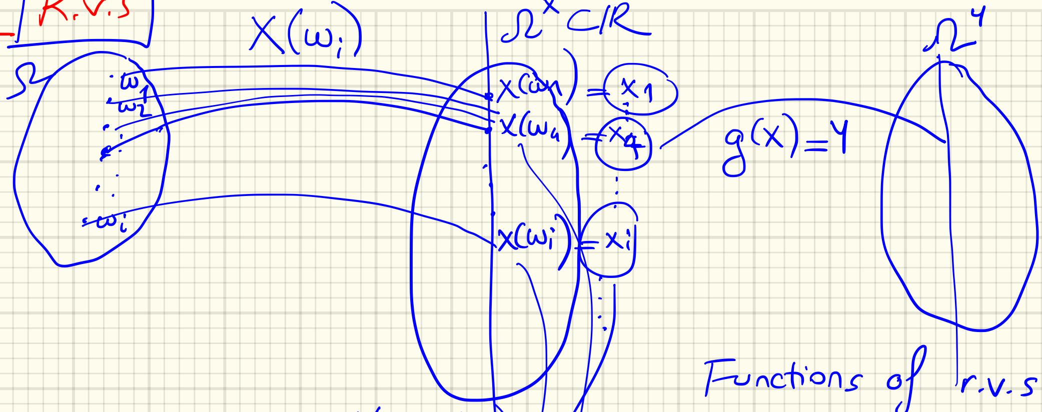
Probability Theory & Stats

Week 5

Gü.

Recap: $\boxed{\text{R.v.s}}$

Random experiment

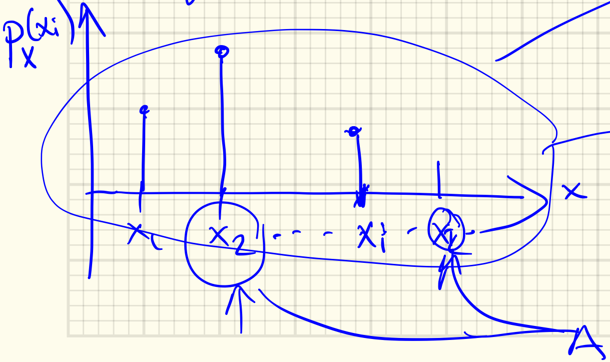


$X = X(\omega)$

↑ r.v.

Functions of r.v.s are also r.v.s.

pmf : prob. mass fn.



pmf properties

1) $P_X(x) \geq 0$

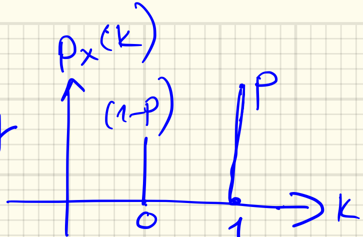
2) $\sum_i P_X(x_i) = 1$

3) $P(A) = \sum_{i:\, x_i \in A} P_X(x_i)$

Recap:

1) Uniform r.V. pmf

2) Bernouilli pmf (Bernouilli exp) : p : success probability, $(1-p)$: failure prob.

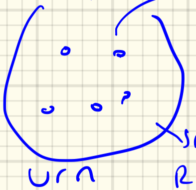


3) Binomial pmf (Bernouilli exp independent) : p : success prob.

seq of independent

k successes out of M (coin tosses) Bernouilli trials.

sampling



M balls

R are success

B are failure

$$P_X(X=k) = \binom{M}{k} p^k (1-p)^{(M-k)}$$

$k = 0, 1, 2, \ldots$

4) Geometric pmf ($\boxed{\text{indep}}$ Bernouilli experiments) : Success at the k^{th} trial

$$P(X=k) = (1-p)^{k-1} \cdot p.$$

$k = 1, 2, \ldots, \infty$

$\underbrace{T\ T\ T \ldots}_{(k-1)\text{ failures}}\ H$

k : success!!

5) Poisson pmf: Important in Queueing / Allocation of resources network modelling, traffic modeling.

$$p_X[k] = e^{-\lambda} \frac{\lambda^k}{k!} \quad , k \geq 0 \quad , \quad \lambda > 0 \text{ (real number)}.$$

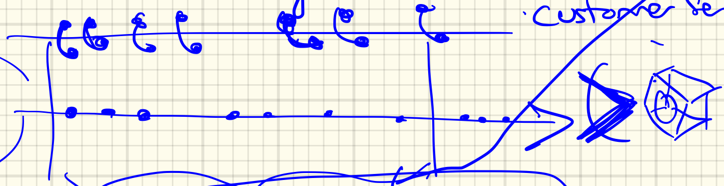
parameter $\longrightarrow$ λ: Average # arrivals (events) / unit time $(\equiv$ rate of arrivals$)$

X: no. of arrivals (requests)

X: no of events

customer service $\rightarrow$ Define λ $\rightarrow$ estimate of λ.

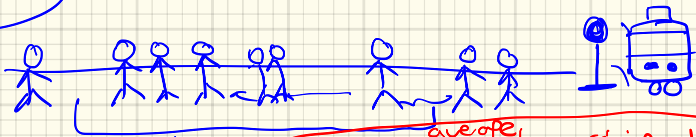
ex; Photon source



t sec. : integration range

α; arrival rate (photons/sec) $\rightarrow$ $\boxed{\lambda = \alpha . t}$

Passengers



t.

λ : # average people arriving at a bus stop / minute,

$$p_X[k] = e^{-\lambda} \frac{\lambda^k}{k!} \quad , k = 0, 1 \ldots$$

prob that k people arrive in a minute.

Binomial pmf can be approximated by Poisson pmf under certain cond.

$\underline{M, p}$; ∴ $M \to \infty$, $p \to 0$: $\boxed{M \cdot p} \to \lambda$. (constant)

M is large ↑↑
p is small ↓↓

$$\text{Binomial}(M, p) \longrightarrow \text{Poisson}(\lambda)$$

$\lambda \approx M \cdot p$ average # successes in M Bernoulli trials.

⟹ Binomial & Poisson are very close.

Ex: optical communication system. Bit arrival rate is $\boxed{10^9 \text{ bits/sec}}$

Probability of having ~~one~~ error bit is 10^{-9}.

Xmitter 001101....1110(1)0 Rcvr

In 1 second : 10^9 bits.

Q: Prob of having 5 error bits in 1 sec?

Soln: $M = 10^9$
$p : 10^{-9}$
$\lambda = M \cdot p = 1$.

$$P_X[X=5] = \binom{10^9}{5}(10^{-9})^5(1-10^{-9})^{10^9-5}$$

can Use Poisson approx. to binomial for $k=5$, $\lambda = 1$

$$p_X[k=5] = \frac{e^{-1}(1)^k}{5!} \approx 0.003$$

Servicing Customers: (Chapter 5 of Skay).

1 "express" lane (: in $\boxed{5-6\text{ pm}}$) range, services each customer ~ 1 min.

Q. Manager wants to know how many extra express lanes should be opened?

Requirement: No more than 1 person waiting in the line 95% of the time

$\equiv$ no more than 2 persons arriving at the express lane in a $\boxed{1\text{-minute}}$ interval.

use Poisson pmf. $\longrightarrow$ need to estimate λ: avg # arrivals /min

avg # customers / min

Go to the store / 1 week make observations

Record

Monday	Tue	Wed	Thu. .	Sun.
71 customers	69	68	72 ⋯	69

in 5-6pm = 60 minute interval

70 arrivals in a 60 minute interval on average

$$\lambda \simeq \frac{70}{60\text{ min}}$$

avg customers / min

$$\boxed{\lambda = \frac{7}{6}}$$ avg # customers/min,

$$p_X[k] = P[k \text{ arrivals in 1 minute}] = e^{-\lambda}\frac{\lambda^k}{k!}, \quad k = 0, 1, \ldots$$

$$\Rightarrow P[X \leq 2] = \sum_{k=0}^{2} p_X[k] = \sum_{k=0}^{2} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda}\left(1 + \lambda + \frac{\lambda^2}{2!}\right)$$

$P[k \leq 2]$

w/ $\lambda = \frac{7}{6} \rightarrow P[X \leq 2] \approx 0.88 < 0.95$

Solution: Consider opening a 2nd express lane:

Two lanes / Two sets of arrivals $\rightarrow$ $\boxed{\text{Independent}}$ $\left(P\left(\bigcap_i A_i\right) = \bigcap_i P(A_i)\right)$

$P[X \leq 2] = P[2 \text{ or fewer arrivals at both lanes}]$

$= \underbrace{P[k \leq 2 \text{ in Lane } 1]}_{\text{identical}} \cdot \underbrace{P[k \leq 2 \text{ in Lane } 2]}_{\text{distributions}}$: Poisson w/ λ^{new}

1 lane | 2nd lane

$\lambda = \frac{7}{6}$

$\lambda^{new} = \frac{70}{2 \cdot 60} = \frac{7}{12}$

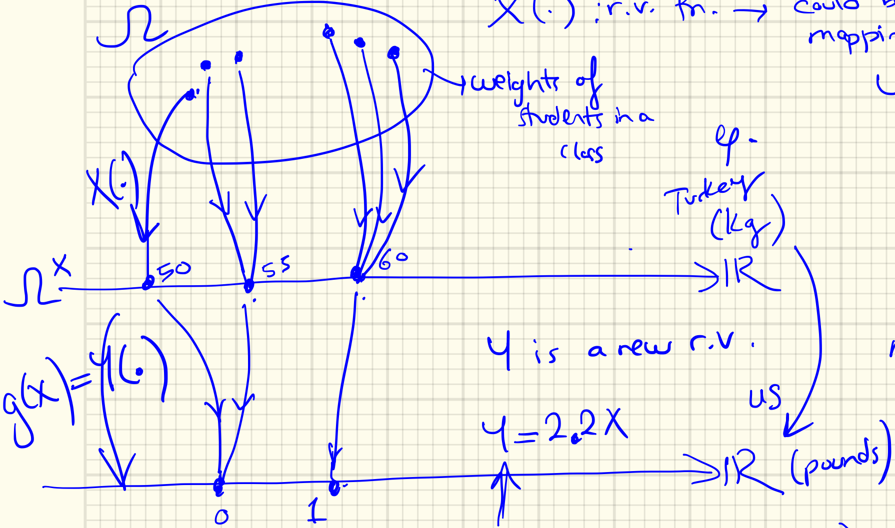
$= P[k \leq 2 \text{ in Lane } i]^2$

$= \left(\sum_{k=0}^{2} e^{-\lambda^{new}} \frac{{\lambda^{new}}^k}{k!} \right)^2 = \left[e^{-\lambda^{new}} \left(1 + \lambda^{new} + \frac{(\lambda^{new})^2}{2}\right) \right]^2$

$\approx \underline{\underline{0.957}}$ ✓ meets the requirement

Transformation of r.v.s : $g : X \longrightarrow Y$

$X(\cdot)$: r.v. fn. $\rightarrow$ could be one-to-one or many-to-one mapping

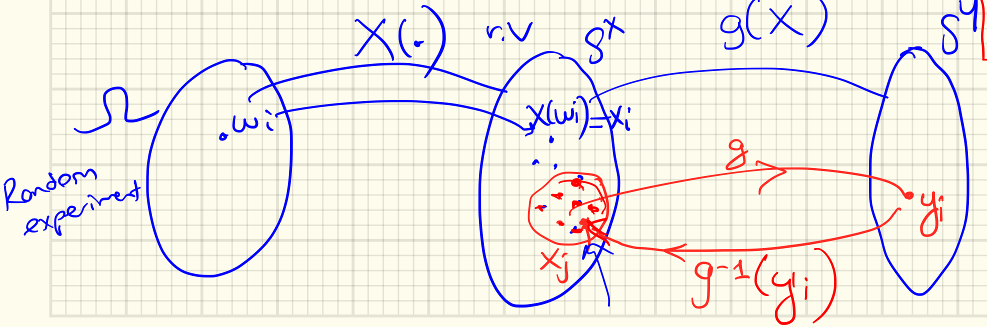


$Y = g(X)$ $\rightarrow$ an r.v.

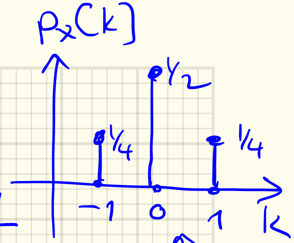
pmf of Y ?

$$\underbrace{P_Y}_{\text{new xformed pmp}}[y_i] = \sum_{\substack{x_j \in g^{-1}(y_i) \\ \equiv \{j \,:\, g(x_j) = y_i\}}} P_X[x_j]$$

$$\boxed{P_Y[y] = \sum_{x \in g^{-1}(y)} P_X[x]}$$



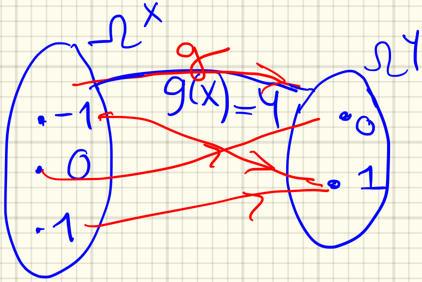
Ex: X is an r.v. w/ pmf $p_X[k] = \begin{cases} \frac{1}{4}, & k=-1 \\ \frac{1}{2}, & k=0 \\ \frac{1}{4}, & k=1 \end{cases}$



Let $Y = X^2 = g(X)$

defined on the sample space $\Omega^X = \{-1, 0, 1\}$

What is $\Omega^Y = \{0, 1\}$



$g(x_j) = x_j^2 = 0$ only for $x_j = 0$

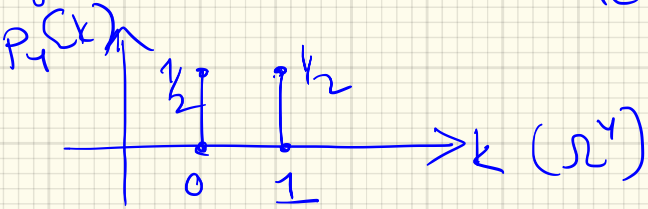
$p_Y[0] = p_X[0]$

$g(x_j) = \underbrace{x_j^2}_{y_i} = 1$ $\rightarrow$ $x_j = -1$, $x_j = +1$

$p_Y[1] = p_X[-1] + p_X[1]$

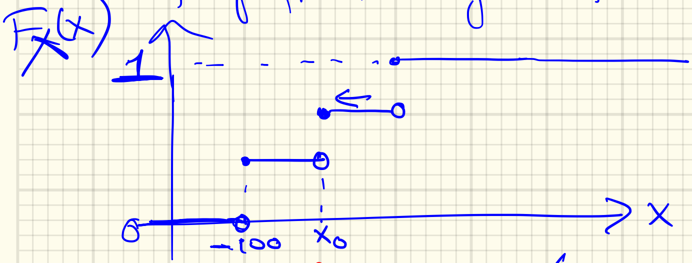
$= \frac{1}{2}$

pmf for Y



CDF : $\boxed{F_X(x) = P[X \le x]}$

pmf from cdf = $P_X[k] = F_X(x^+) - F_X(x^-)$



Properties of the CDF (Section 5.8 S Kay Book)

1) $0 \le F_X(x) \le 1$

2) $\underbrace{\lim_{x\to\infty} F_X(x)}$ $= 1$, $\lim_{x\to -\infty} F_X(x) = 0$

$P(\underbrace{X \le \infty}_{\Omega}) = 1$ $\quad$ $P(\underbrace{X \le -\infty}_{\emptyset}) = 0$ $\quad$ (from Axioms of Probability)

3) CDF is right-continuous $\rightarrow \equiv$ as we approach x_0 from the right the limiting value of CDF is $F(x_0)$.

$$\lim_{x\to x_0^+} F_X(x) = F_X(x_0)$$

4) CDF is monotonically non decreasing.

$$\underbrace{F_X(x_1)}_{P(X \leq x_1)} \leq \underbrace{F(x_2)}_{P(X \leq x_2)} \quad \text{if } \underline{x_1 \leq x_2}.$$

Show this.

$$\underbrace{P(X \leq x_1)}_{A} \subset \underbrace{P(X \leq x_2)}_{B}$$

$$A = \{-\infty < x < x_1\}$$
$$B = \{-\infty < x < x_2\}$$

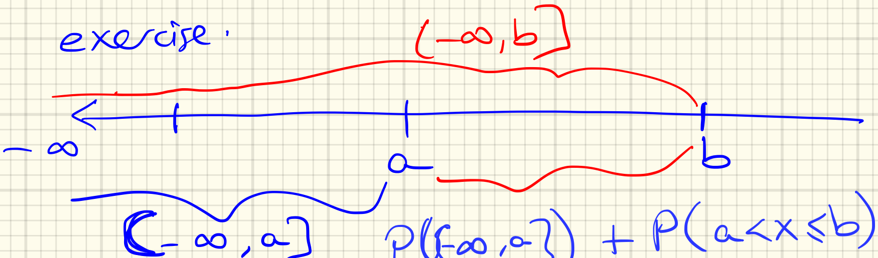
$x_1 \leq x_2 \quad A \subset B$

$\rightarrow P(A) < P(B)$

5) Intervals $\boxed{P[a < x \leq b]} = \underbrace{F_X(b)}_{P\{-\infty < x \leq b\}} - \underbrace{F_X(a)}_{P(-\infty < x \leq a)}$

Show this for $\boxed{a < b}$.

exercise:

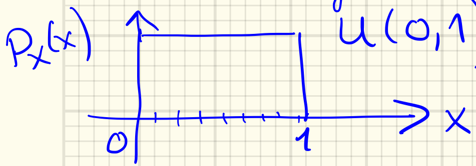


$P((-\infty, a]) + P(a < x \leq b) = F_X(b)$

$P(-\infty < x < b)$

Probability Integral Transformation ;

Recall uniform r.v. $U(0,1)$



(1) If an r.v. is transformed according to its CDF ;

$$U = F_X(x)$$

r.v. $X \xrightarrow{g(\cdot)} Y = g(X)$; $\boxed{g = \text{CDF of } X}$

Start w/ an arbitrary r.v. $X \rightarrow F_X(\cdot) \rightarrow U$: uniform.

$X \leftarrow \boxed{F_X^{-1}(\cdot)} \leftarrow U$ start w/ uniform r.v.

Inverse Probability Integral Transform (Thm 10.9.1 SKay) : read the proof.

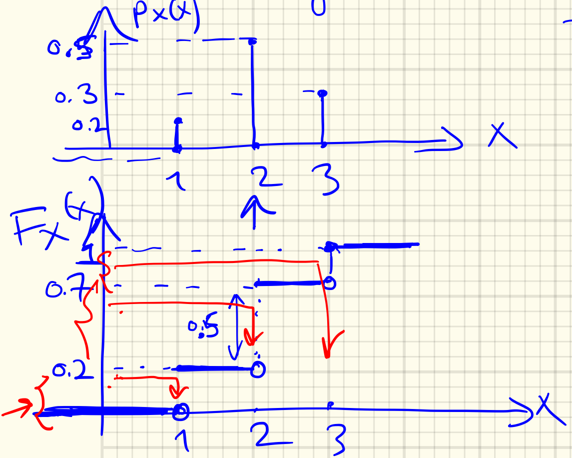
$$\boxed{X \xrightarrow{F_X(\cdot)} U \xrightarrow{F_Y^{-1}(\cdot)} Y}$$

matching distrib. of X to distribution of Y.

Ex: How to simulate an r.v. on a computer?

Say X takes on values $\Omega^X = \{1, 2, 3\}$

w/ a pmf



Q. Write a code that generates $M = 1000$ realizations of X.

A. We have $U[0,1]$ r.v. (uniform) generator; $u_1, u_2, \ldots, u_{1000}$ in the interval $(0,1)$.

We sample $\boxed{u_1, u_2, u_3, \ldots u_{M=1000}}$ → distributed uniformly. $\boxed{\text{rand}(\cdot) \text{ function}}$

$0 \le u < 0.2 \rightarrow x = 1$

$0.2 \le u < 0.7 \rightarrow x = 2$

$0.7 \le u \le 1 \rightarrow x = 3$

Using $F_X^{-1}(u)$

$x_1, x_2, x_3, \ldots, x_{1000}$: $\boxed{N = 1000 \text{ realizations of } X.}$

Now these are distributed according $p_X(x)$.

Thm: Given the CDF of X, want to generate random numbers distributed as X:

$x_1, x_2, \ldots, x_{10{,}000}$

I can generate u's: $\underset{\approx}{U} \longrightarrow F_X^{-1}(u) = X \longrightarrow$ from the distribution X.

$g = F_X^{-1}(\cdot)$ cdf inverse.

or

$X \longrightarrow U$

$g = \underbrace{F_X(\cdot)}_{\text{CDF itself.}}$

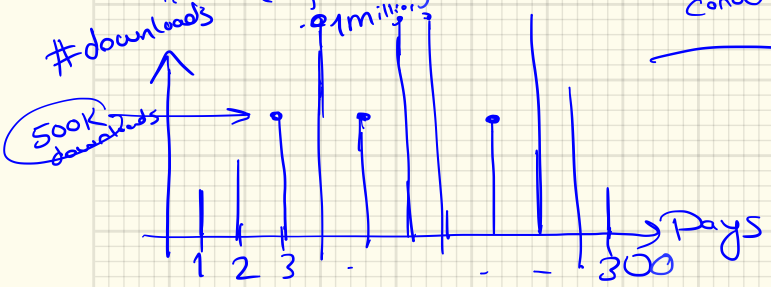
plot their pmf. it should look like $\underline{P_X(x)}$

Expectation of an r.V.

pmf, cdf.

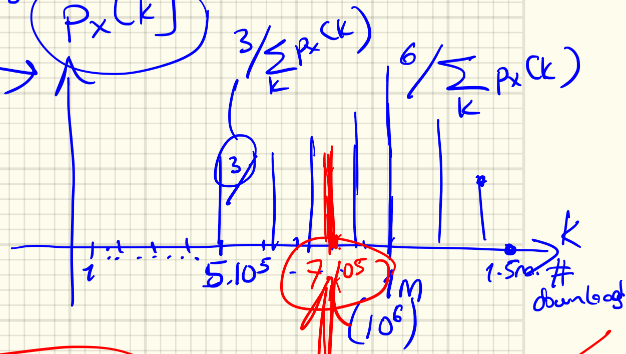
pmf (cdf) is a complete description of an r.V → it gives us all required probabilities of data values.

eg. You collected data, eg. server download requests in the past year.



convert into pmf

$P_X[k]$



To interpret probability distribution: want more compact values from the pmf like the average:

: average value of the outcomes of a "large" # of experimental trials

Def: (Expected Value of an r.v.)

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

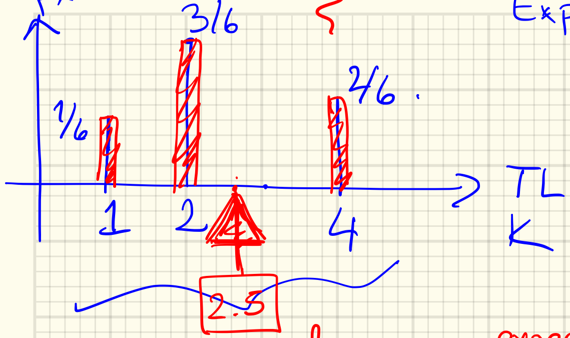
$$E[X] = \sum_k k . p_X[k]$$

Sample Mean: An r.v. W/ N realizations.

$x_i \Rightarrow i = 1, 2 \ldots N$

$$\bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i$$

pmf : mass fn



Expected value (TL) one would earn ?

$$E[X] = \sum_{k=1,2,4} k \, P_X[k]$$

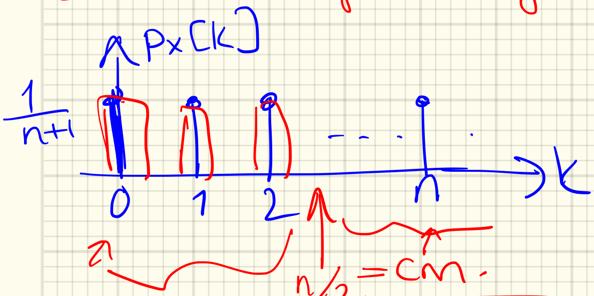
$$1 \cdot \frac{1}{6} + 2 \cdot \frac{3}{6} + 4 \cdot \frac{2}{6} = \boxed{2.5}$$

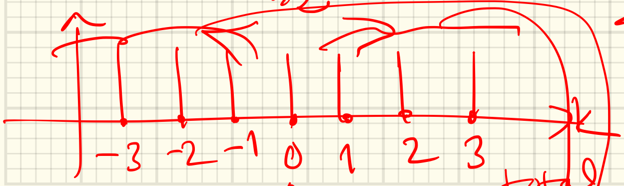
Center of mass = expected value

Discrete Uniform Pmf : $E[X] = ? \sum_{k=0}^{n} k \cdot \frac{1}{n+1}$

$$= 0 \cdot \frac{1}{n+1} + 1 \cdot \frac{1}{n+1} + 2 \cdot \frac{1}{n+1} + \dots + n \cdot \frac{1}{n+1}$$

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

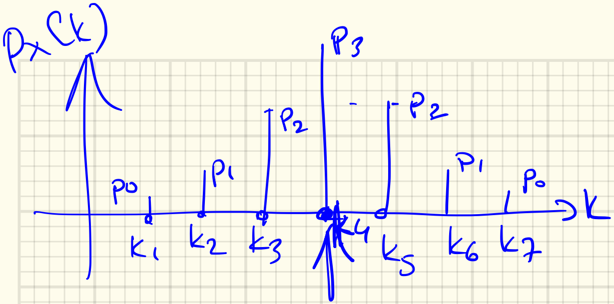




↑ cm

total prob on both sides are equal.

★ For a symmetric pmf around a certain value ⇒ the expected value of the pmf.

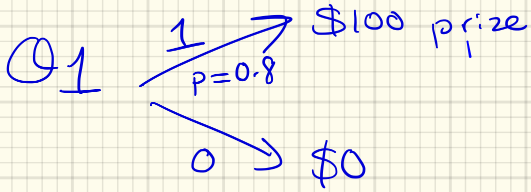


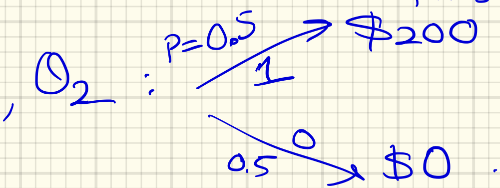
$2p_0 + 2p_1 + 2p_2 + p_3 = 1$.

$\forall p_i \geq 0$.

$E[X] = k_4$

Ex 2.8 Quiz problem: Given 2 questions, which to answer first?

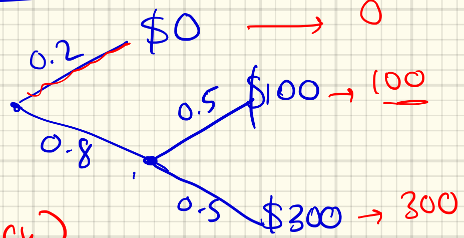




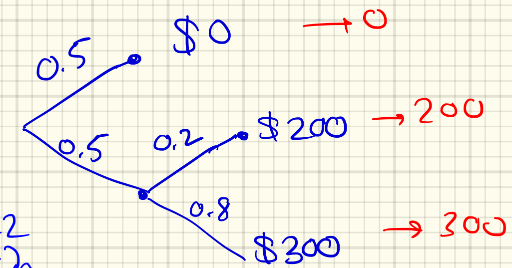
If the 1st attempted question is answered incorrectly, the quiz terminates.

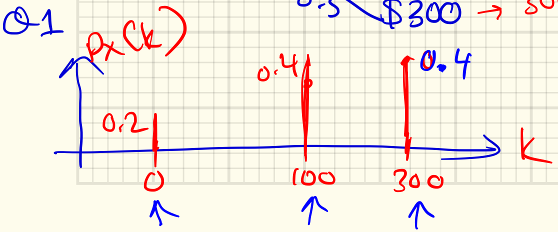
Q. Which question should be answered first to maximize the expected earning?

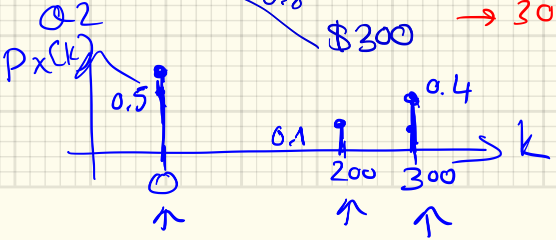
Q1 answered 1st case



Q2 answered 1st case







Q1: first $E[X] = 0 \cdot (0.2) + 100\,(0.4) + 300\,(0.4) = \$160$ ←.

Q2: first $E[X] = 0\,(0.5) + 200\,(0.1) + 300\,(0.4) = \$140$.

Expectation helps us in our decision making

→ You'd go w/ Q1 first.

Expected Values of Some Important r.v.s (Sec 6.4 Skay)

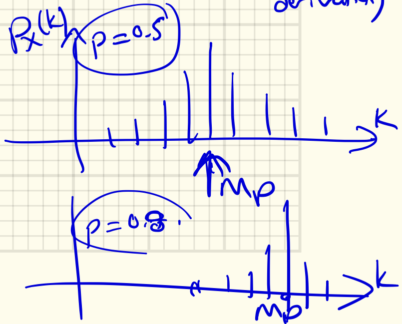
1) Bernoulli: $X \sim Ber(p)$

$$E[X] = \sum_{k=0,1} k\, p_X(k) = 0 \cdot (1-p) + 1 \cdot p = p.$$

2) Binomial: $X \sim bin(M, p)$

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

((Check Kay book for derivation))



3) Geometric: $X \sim geom(p)$

$$E[X] = \sum_{k=1}^{\infty} k \underbrace{(1-p)^{k-1} \cdot p}_{\text{pmf}} = p \sum_{k=1}^{\infty} k (\underbrace{1-p}_{\alpha})^{k-1}$$

$$E[X] = p \underbrace{\sum_{k=1}^{\infty} k(1-p)^{k-1}}_{\frac{1}{p^2}} = \frac{1}{p}$$

$$\frac{d}{d\alpha}\left(\sum_{k=0}^{\infty}\alpha^k\right) = \sum_k \frac{d}{d\alpha}\alpha^k = \boxed{\sum_k k\alpha^{k-1}}$$

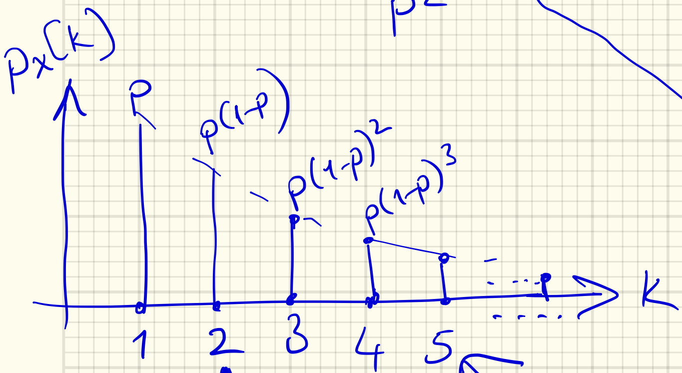
$$\frac{d}{d\alpha}\left(\frac{1}{1-\alpha}\right)$$

$|\alpha| < 1.$

$$= \underbrace{\frac{1}{(1-\alpha)^2}} = \sum_k k\alpha^{k-1}$$

$\alpha = 1-p$

$1-\alpha = p.$



$p = 0.5$

$E[X] = \frac{1}{p} = 2$

if $p = 0{,}2$ reduced the success prob.

$E(X) = 5$

4) Poisson : $E[X] = \lambda$

$k = 0, 1, \dots, \infty$.

$$E[X] = \sum_{k=0}^{\infty} k \cdot \underbrace{\left(e^{-\lambda} \frac{\lambda^k}{k!}\right)}_{\text{pmf. of Poisson}} = e^{-\lambda} \sum_{k=0}^{\infty} k \cdot \frac{\lambda^k}{k!}$$

$$= e^{-\lambda} \cdot \lambda \cdot e^{\lambda}$$

$$E[X] = \lambda \quad \checkmark$$

$$\frac{d}{d\lambda}\left(\underbrace{\sum_{k=0}^{\infty} \frac{\lambda^k}{k!}}_{= e^{\lambda}}\right) = e^{\lambda}$$

$$\sum_{k=0}^{\infty} k \frac{\lambda^{k-1}}{k!} \cdot \frac{\lambda}{\lambda} = e^{\lambda}$$

Variance: Second moment

Recall (Standard Deviation) $= \sqrt{\mathrm{Var}\, X}$

$$\mathrm{Var}(X) \triangleq E\left[(X - E[X])^2\right]$$

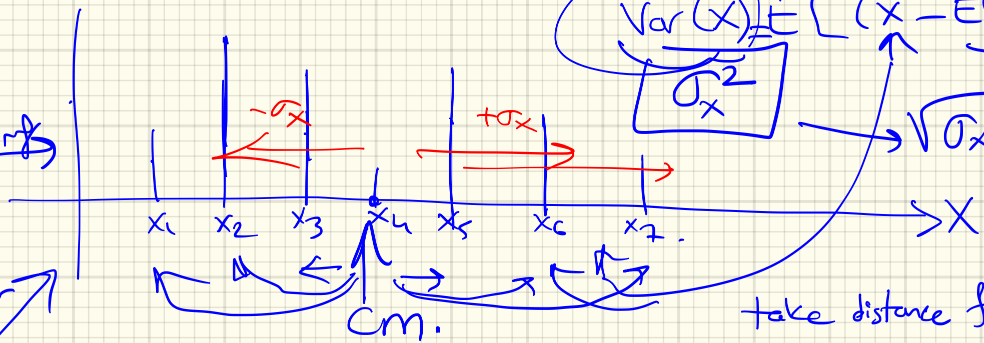
$$= \sum_x (x - E[x])^2 \cdot P_X(x)$$

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

Show this exercise at home.

$$\mathrm{Var}(X) = \underbrace{E\left[(X - \underbrace{E[X]}_{\text{cm.}})^2\right]}_{\boxed{\sigma_X^2}}$$

$$\sqrt{\sigma_X^2} = \sigma_X \text{ : std deviation}$$



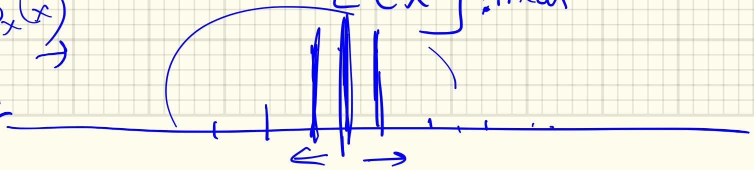
this pmf

variance of

smaller than

Variance of this pmf $P_X(x)$

$E[X]$: mean



take distance from the mean & square it & take average that is the variance.