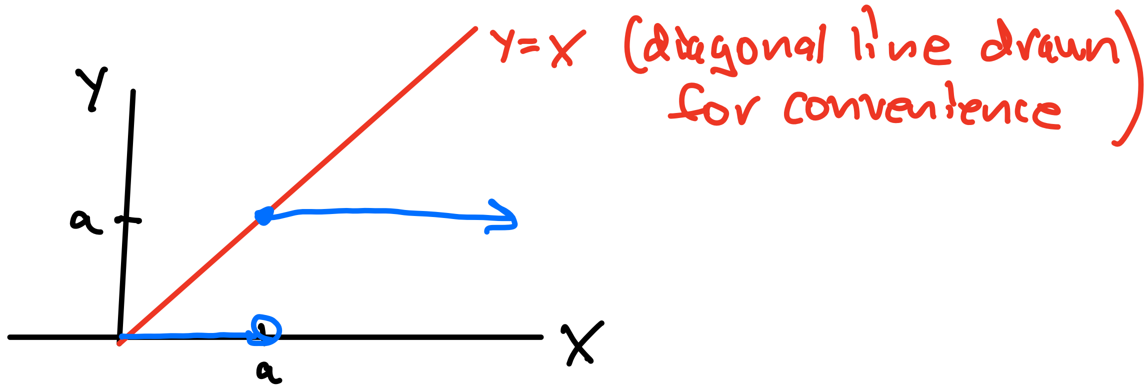
Warmup: 9:00–9:10

This question asks you to graph

Let $Y = \begin{cases} a & \text{if } X \geq a \\ 0 & \text{else} \end{cases}$



$a \cdot P(X \geq a) + 0 \cdot P(X < a)$ avg value of blue graph,

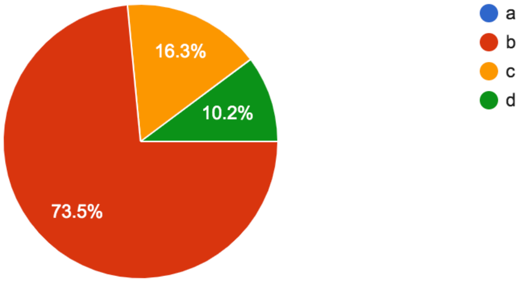
$E(X)$ is avg value of red graph,

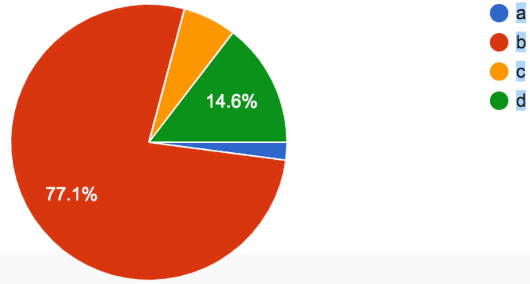
red $\geq$ blue

$E(X) \geq a \cdot P(X \geq a)$

$\Rightarrow \boxed{P(X \geq a) \leq \frac{E(X)}{a}}$ Markov's Inequality, assumes $X \geq 0$ and $a > 0$

Announcement: Q2 in section next Monday 9/26. Coverage: Sections 2.1, 2.2, 2.4, 2.5, 3.1, 3.2





Stat 134

Chapter 3 Friday February 15 2019

1. Consider a well shuffled deck of cards. The expected number of cards before the first ace is?

   a 52/5

   b 48/5

   c 48/4

   d none of the above

d	The p of getting an ace is 4/52 or 1/13, therefore, the EV should be 52/13 or np.

b	There are 48 possible cards to pull before the first Ace (52 cards - 4 Aces), which makes n = 48. For each individual card, there are 5 places it could go: before all the Aces, after the 1st Ace, after the 2nd Ace, after the 3rd Ace, and after the 4th Ace, making p = 1/5. E(X) = np = 48/5

$E(X) = P(X \geq 1) + P(X \geq 2) + P(X \geq 3) + \cdots$ Tail Sum Formula

This is useful when X = min or max,

Discrete Distributions

1. Ber(p)
2. Bin(n,p)
3. HG(n,N,G)
4. Pois(μ)
5. Unif$\{1,\ldots,n\}$
6. Geom(p) on $\{1,2,\ldots\}$

Geometric RV

\# trials until first success

ex X = number of p coin tosses until your first heads

$X=1$	H	p
$X=2$	TH	qp
$X=3$	TTH	q^2p

$$P(X=k) = q^{k-1}p$$ Geom(p) formula on $\{1,2,\ldots\}$

Note trials are independent

Today

1. Sec 3.2 Markov Inequality
2. Sec 3.3 $SD(X)$, $Var(X)$, Chebyshev's Inequality
3. Sec 3.2 $E(g(X,Y))$

① Sec 3.2 Markov Inequality

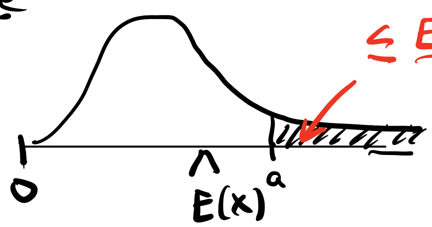
Proved in warmup

Markov's inequality:

If $X \geq 0$, then $P(X \geq a) \leq \frac{E(X)}{a}$ for every $a > 0$.

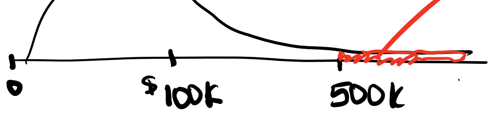
Picture

$\leq \frac{E(X)}{a}$



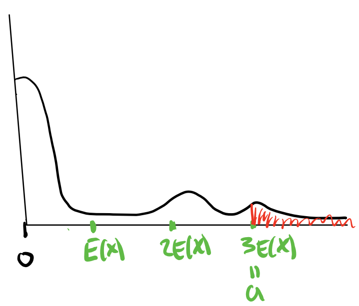
ex Let X be the yearly income of Bay area residents. $E(X) = \$100K$. Find an upper bound for $P(X \geq 500K)$

$P(X \geq 500K) \leq \frac{E(X)}{500} = \boxed{\frac{1}{5}}$



ex Give an upper bound for the fraction of all US flights that have delay times greater than 3 or more times the national average.

$P(X \geq \underbrace{3E(X)}_{a}) \leq \frac{E(X)}{3E(X)} = \boxed{1/3}$



ex Let $X_1, X_2, \dots, X_{100}$ be independent and identically distributed (iid) Pois(.01).

Let $S = X_1 + X_2 + \dots + X_{100}$

$X \sim \text{Pois}(\mu)$

$$P(X=k) = \frac{e^{-\mu}\mu^k}{k!}$$

a) What distribution is S? Pois(1)

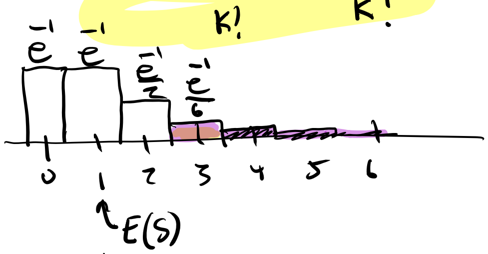
b) Find an upperbound for $P(S \geq 3)$ using Markov's inequality.

$E(S) = 1$

$$P(S \geq 3) \leq \frac{E(S)}{3} = \boxed{1/3}$$

Note Exact: $P(S \geq 3) = 1 - \underbrace{P(0)}_{=e^{-1}} - \underbrace{P(1)}_{=e^{-1}} - \underbrace{P(2)}_{=\frac{e^{-1}}{2}}$

$$P(S=k) = \frac{e^{-1} 1^k}{k!} = \frac{e^{-1}}{k!}$$



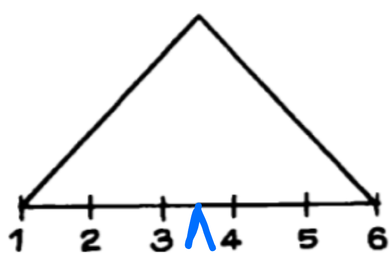
$= 1 - e^{-1}(1 + 1 + \frac{1}{2})$

$= .08$

② Sec 3.3 Standard deviation (SD)

SD is the average spread of your data around the mean.

What is the SD of the following figure?



informal def^n

$SD = E(|x - E(x)|)$

a 0.5

b 1

c 2

$SD(x) = \sqrt{E((x-E(x))^2)}$

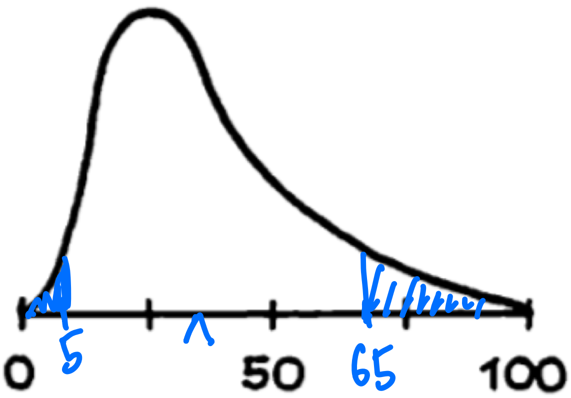
$Var(x) = (SD(x))^2 = E((x-E(x))^2)$

Chebyshev's Inequality

For any random variable X, and any $k>0$,

$$P\left(|X - E(X)| \geq k \cdot SD(X)\right) \leq \frac{1}{k^2}$$

ex Let X have distribution with $E(X) = 35$, $SD(X) = 15$.



$2 \cdot SD \Rightarrow \boxed{k=2}$

Find $P(|X-35| \geq 30)$?

$\leq \frac{1}{2^2} = \frac{1}{4}$

What can you say about $P(X \geq 65)$? $\leq 1/4$

$P(X \geq 65) = P(X \geq \underset{\underset{35}{\shortparallel}}{\mu} + \underset{\underset{2}{\shortparallel}}{k}\underset{\underset{15}{\shortparallel}}{\sigma}) \leq (1/2)^2$

tinyurl.com/sept19-2022

Stat 134

1. A list of non negative numbers has an average of 1 and an SD of 2. Let p be the proportion of numbers greater than or equal to 5. To get an upper bound for p, you should:

   a Assume a normal distribution

   b Use Markov's inequality

   c Use Chebyshev's inequality

   d none of the above

X = nonnegative number

$E[X] = 1$

$a = 5$ M: $P(X \ge 5) \le 1/5$

C: $P(X \ge 5) \le 1/4$

$1 + 2 \cdot 2 \Rightarrow k = 2$

$1/5 \le 1/4$

Proof of Chebyshev

For any random variable X, and any $k>0$

$$P\left(|X-E(X)| \geq k\,SD(X)\right) \leq \frac{1}{k^2}$$

By Markov

$$P(Y \geq a) \leq \frac{E(Y)}{a} \quad \text{for } \begin{matrix} Y \geq 0 \\ a > 0 \end{matrix}$$

$Y = (X-E(X))^2 \leftarrow$ non negative

$a = (k\,SD(X))^2 \leftarrow$ positive.

M:

$$P\left((X-E(X))^2 \geq (k\,SD(X))^2\right) \leq \frac{E\left((X-E(X))^2\right)}{k^2 (SD(X))^2} = \frac{1}{k^2}$$

$$E\left((X-E(X))^2\right) = \cancel{Var(X)}, \quad (SD(X))^2 = \cancel{Var(X)}$$

$$\parallel$$

$$P\left(\sqrt{(X-E(X))^2} \geq \sqrt{(k\,SD(X))^2}\right)$$

$$\sqrt{(X-E(X))^2} = |X-E(X)|, \quad \sqrt{(k\,SD(X))^2} = k\,SD(X)$$

$$|X-E(X)| \geq k\,SD(X)$$

So $\boxed{P\left(|X-E(X)| \geq k\,SD(X)\right) \leq \frac{1}{k^2}}$ $\square$