Stat 134 lec 9

Warmup:

The formula for the expectation of the function of a random variable is:

$$E(g(X)) = \sum_{x \in X} g(x) P(X=x)$$

$1 + q + q^2 + \cdots$ geometric sum $|q|<1$

$= \frac{1}{1-q}$

Find:

$X \sim Pois(\frac{1}{3})$

Find $E(X!)$

$X \sim Pois(\mu)$

$P(X=k) = \frac{e^{-\mu}\mu^k}{k!}$ $k = 0, 1, 2, \ldots$

$g(x) = x!$

$$E(g(x)) = \sum_{\text{all } k} g(k) P(X=k)$$

$\frac{1}{1-\frac{1}{3}} = \frac{3}{2}$

$$E(X!) = \sum_{k=0}^{\infty} \cancel{k!} \frac{e^{-\mu}\mu^k}{\cancel{k!}} = \sum_{k=0}^{\infty} e^{-\frac{1}{3}} \left(\tfrac{1}{3}\right)^k = e^{-\frac{1}{3}}\left(1 + \tfrac{1}{3} + \left(\tfrac{1}{3}\right)^2 + \cdots\right)$$

$$= \boxed{\frac{3}{2} e^{-\frac{1}{3}}}$$

Last time

$E(X) = P(X\geq 1)+P(X\geq 2)+P(X\geq 3)+\cdots$ 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 Expectation of a function of a RV
2. Sec 3.2 Markov inequality
3. Sec 3.3 $SD(X)$, $Var(X)$, Chebyshev's Inequality

② Sec 3.2 Expectation of a function of a RV.

$$E(X) = \sum_{x \in X} x P(X=x)$$

$$E(g(X)) = \sum_{x \in X} g(x) P(X=x)$$

<u>ex</u> Suppose $X \sim \text{Geom}(p)$ on $\{1, 2, \cdots\}$ with $p > 2/3$

Find $E(3^X)$.

$X \sim \text{Geom}(p)$

trials to 1st success

$P(X=k) = q^{k-1} p$

<u>Picture</u>

$$1 \xrightarrow{y=3^x} 3^1$$

$$2 \longrightarrow 3^2$$

$$3 \longrightarrow 3^3$$

$\vdots$

$$E(3^X) = \sum_{k=1}^{\infty} 3^k P(X=k) = \sum_{k=1}^{\infty} 3^k q^{k-1} p$$

$$= 3p + 3^2 q p + 3^3 q^2 p + \cdots$$

$$= 3p\left(1 + 3q + (3q)^2 + \cdots\right)$$

$$1 + 3q + (3q)^2 + \cdots = \frac{1}{1-3q} \quad \text{if } 3q < 1$$

yes since $p > 2/3$

$$\boxed{E(3^X) = 3p\left(\frac{1}{1-3q}\right)}$$

Several variables

(X,Y) joint distribution

$$E(g(X)) = \sum_{\text{all } x} g(x) P(X=x)$$

$$E(g(X,Y)) = \sum_{\text{all } x,y} g(x,y) P(X=x, Y=y)$$

see appendix to notes

Thm $E(X+Y) = E(X) + E(Y)$

$X = aA+b \quad Y = cB+d$

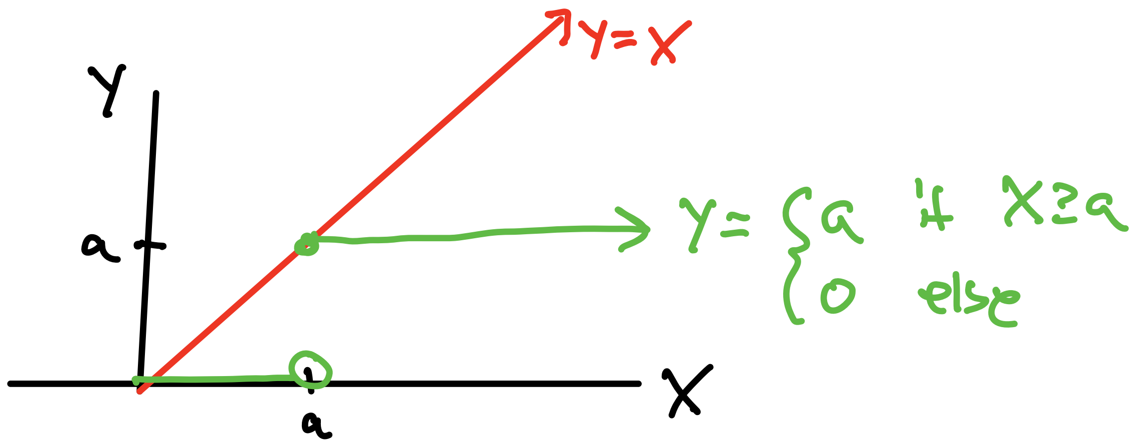
see appendix to notes

Thm if X and Y are independent

$E(XY) = E(X)E(Y)$

1. Sec 3.2 <u>Markov Inequality</u>

Compare the red and green RVs,



Both red and green RVs are functions of a nonnegative RV X.

Notice that red ≥ green (i.e the red graph is above the green graph,

It follows that E(red) ≥ E(green)

i.e $E(X) \geq 0 \cdot P(X<a) + a \cdot P(X \geq a)$

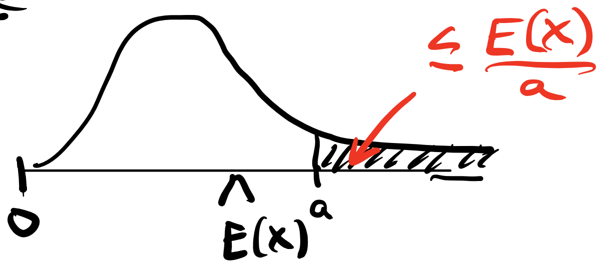
$\Rightarrow E(X) \geq a P(X \geq a)$

$\Rightarrow \boxed{P(X \geq a) \leq \frac{E(X)}{a}}$ ← Markov's inequality, Proved above

Markov's inequality:

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

Picture



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

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

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

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

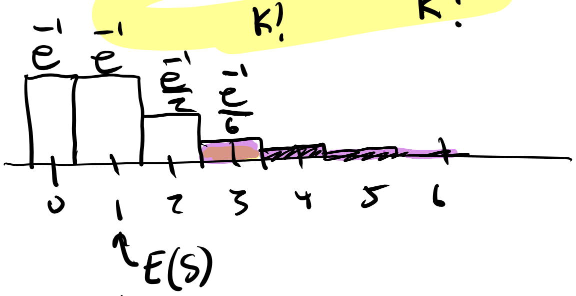
a) What distribution is S? $S \sim \text{Pois}(\underbrace{100(.01)}_{=1})$

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

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

Note Exact! $P(S \geq 3) = 1 - P(0) - P(1) - P(2)$
$\qquad\qquad\qquad\qquad\qquad \; = e^{-1} \quad = e^{-1} \quad = \frac{e^{-1}}{2}$

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



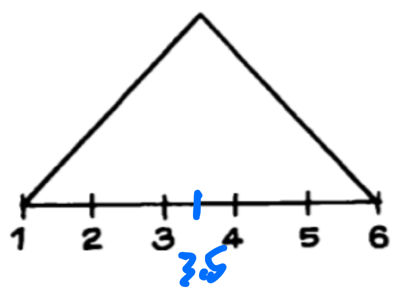
$= 1 - e^{-1}(1 + 1 + \frac{1}{2})$

$= .08$

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?



a 0.5

b 1

c 2

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

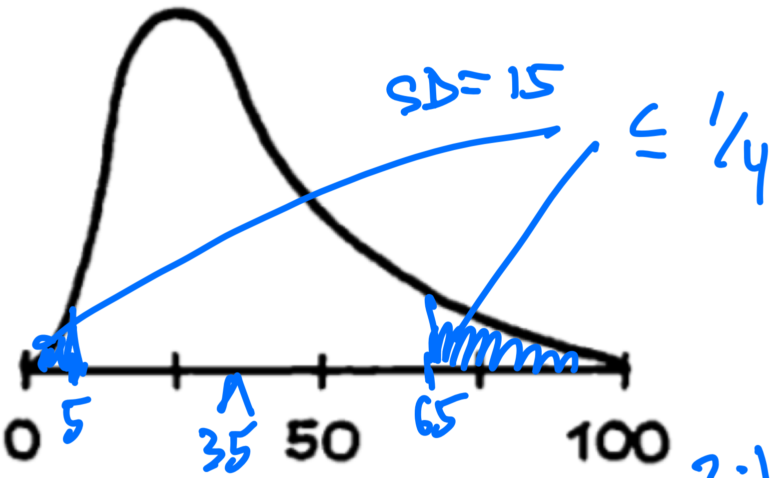
$$Var(x) = (SD(x))^2 = E\left((x-E(x))^2\right)$$

Chebyshev's Inequality

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

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

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



$2\cdot 15$ $\quad k=2$

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

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

$\le 1/4$

What can you say about $P(X \ge 65)$?

better upper bound.

M! $\quad P(X \ge 65) \le \dfrac{E(X)=35}{65}$

tinyurl.com/feb10-2023

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

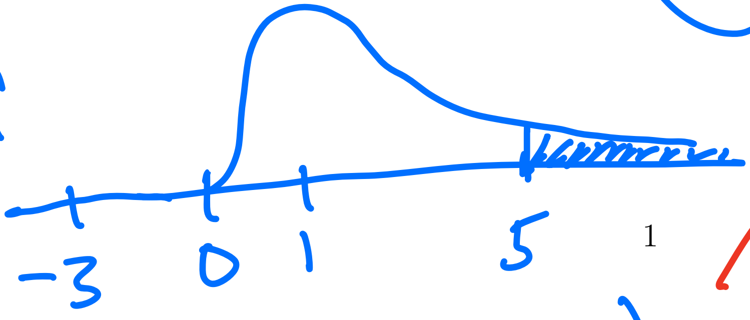
   c Use Chebyshev's inequality

   d none of the above

We should test both Markov and Chebyshev and see which has a smaller upper bound

M: $P(X \geq 5) \leq 1/5$

C:



since $P(X \leq -3) = 0$

$P(X \geq 5) = P(X-1 \geq 4) = P(|X-1| \geq 4) \leq 1/4$

$4 = 2 \cdot 2$

$\Rightarrow k = 2$

Appendix

Thm $E(X+Y) = E(X) + E(Y)$

Pf/ $E(X) = \sum_{\text{all } x,y} x P(X=x, Y=y)$

$E(Y) = \sum_{\text{all } x,y} y P(X=x, Y=y)$

$E(X+Y) = \sum_{\text{all } x,y} (x+y) P(X=x, Y=y)$

$= \underbrace{\sum_{\text{all } x,y} x P(X=x, Y=y)}_{\substack{\shortparallel \\ E(X)}} + \underbrace{\sum_{\text{all } x,y} y P(X=x, Y=y)}_{\substack{\shortparallel \\ E(Y).}}$ $\square$

ex Thm if X and Y are independent $E(XY) = E(X)E(Y)$

Pf/ $E(XY) = \sum_{\text{all } x,y} xy P(X=x, Y=y)$ $= P(X=x)P(Y=y)$ by independence

$= \sum_{\text{all } x,y} x P(X=x) y P(Y=y)$

$= \sum_{\text{all } x} x P(X=x) \sum_{\text{all } y} y P(Y=y) = E(X)E(Y)$ $\square$