

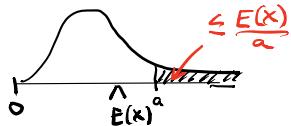
Stat 134, Dec 12

Last time

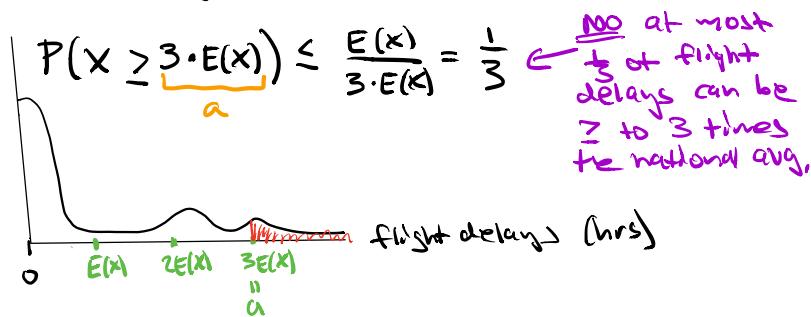
Sec 3.2 Markov's Inequality

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

Picture



ex Is it possible that half of all US flights have delay times greater than 3 or more times the national average?



Discrete Distributions

- (1) $\text{Ber}(p)$
- (2) $\text{Bin}(n, p)$
- (3) $\text{hyper}(n, 6, N)$
- (4) $\text{Pois}(\mu)$
- (5) $\text{Unif}\{1, \dots, n\}$
- (6) $\text{Geom}(p)$ on $\{1, 2, \dots\}$

Geometric RV

trials until first success

ex $X =$ number of p coin tosses until your first head

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

$$P(X=k) = q^{k-1} p \quad \text{Geom}(p) \text{ formula}$$

on $\{1, 2, \dots\}$

Note trials are independent

Today

- (6) Finish Markov problem from last time.
- (1) review student responses to concert test
- (2) Sec 3.2 $E(g(X, Y))$
- (3) Sec 3.3 $SD(X), Var(X)$
chebyshev inequality.

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

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

Draw the histogram of S and approximate an upperbound for $P(S \geq 3)$ using Markov's inequality.

Soln $S = X_1 + X_2 + \dots + X_{100} \sim \text{Pois}(100 \cdot .01) = \text{Pois}(1)$,

$$S \sim \text{Pois}(1)$$

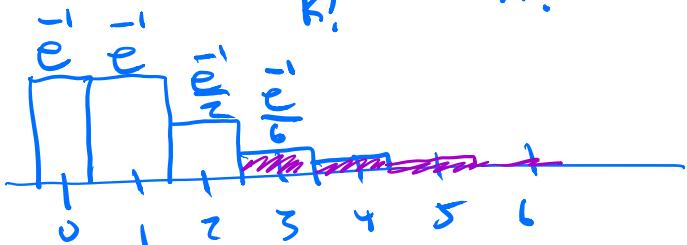
$$E(S) = 1$$

$$\text{By Markov's inequality, } P(S \geq 3) \leq \left(\frac{1}{3}\right)$$

Note Exact: $P(S \geq 3) = 1 - P(0) - P(1) - P(2)$

$$= 1 - \frac{e^{-1}}{0!} - \frac{e^{-1}}{1!} - \frac{e^{-1}}{2!}$$

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



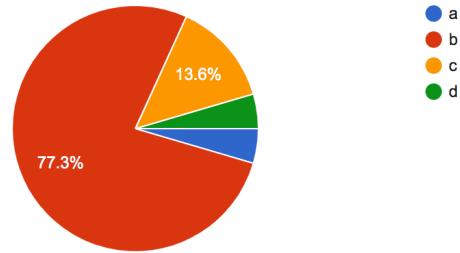
$$= 1 - \bar{e}^{(1+1+\frac{1}{2})}$$

$$= 0.08$$

① Concept test responses

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



b Think of slots: _ A _ A _ A _ A _ \rightarrow 5 slots.
Every non-ace card had a $1/5$ chance of being before the first ace
 $\Rightarrow 48 * 1/5 = 48/5 = B$

c spread out aces over 48 non-aces, means on average there should be one after every ~~12~~⁴⁸ theoretically

$$\frac{48}{4} = \frac{48}{3} = 16$$

c because expected value of geometric is $1/p$

Clever idea but geometric requires drawing cards with replacement.

Note Same answer if between 1st and 2nd ace.

② Sec 3.2 Expectation of a function of a RV.

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

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

Ex Suppose $X \sim \text{Geom}(p)$ on $\{1, 2, \dots\}$ with $p > 2/3$
 Find $E(3^X)$.

Picture

$$\begin{array}{ccc} p & 1 & \xrightarrow{y=3^x} 3^1 \\ q & 2 & \longrightarrow 3^2 \\ q^2 & 3 & \longrightarrow 3^3 \\ & \vdots & \end{array}$$

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

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

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

$$\frac{1}{1-3q} \quad \text{if } 3q < 1$$

yes since
 $p > 2/3$

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

$$\text{Ex } X \sim \text{Pois}\left(\frac{1}{3}\right) \quad \left(\begin{array}{l} X \sim \text{Pois}(m) \\ P(X=k) = \frac{e^{-m} m^k}{k!} \end{array} \right)$$

Find $E(X!)$

Soln

$$E(X!) = \sum_{k=0}^{\infty} k! P(X=k) = e^{-\frac{1}{3}} \sum_{k=0}^{\infty} \frac{\left(\frac{1}{3}\right)^k}{k!} = e^{-\frac{1}{3}} \sum_{k=0}^{\infty} \left(\frac{1}{3}\right)^k$$

$$= e^{-\frac{1}{3}} \left[1 + \frac{1}{3} + \left(\frac{1}{3}\right)^2 + \dots \right] = \boxed{\frac{3}{2} e^{-\frac{1}{3}}}$$

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

Several Variables

(X, Y) joint distribution

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

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

\equiv (X, Y) joint distribution

		$\frac{1}{2}$	$\frac{1}{2}$	$P(X)$
		$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$
		$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$
		$\frac{1}{4}$	0	$\frac{1}{4}$
		0	$\frac{1}{4}$	$\frac{1}{4}$
		0	0	0

y

x

Find $E(X)$:

$$E(X) = \sum_{\text{all } x, y} x P(X=x, Y=y) = 0 \cdot \frac{1}{4} + 0 \cdot \frac{1}{4} + 0 \cdot 0 + 1 \cdot 0 + 1 \cdot \frac{1}{4} + 1 \cdot \frac{1}{4} = \frac{1}{2}$$

$g(x, y) = x$

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

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

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

$$\begin{aligned} E(X+Y) &= \sum_{\text{all } x, y} (x+y) P(X=x, Y=y) \\ &= \underbrace{\sum_{\text{all } x, y} x P(X=x, Y=y)}_{E(X)} + \underbrace{\sum_{\text{all } x, y} y P(X=x, Y=y)}_{E(Y)}. \end{aligned}$$

□

or Thm if X and Y are independent

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

a) Write the formula for $E(XY)$

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

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

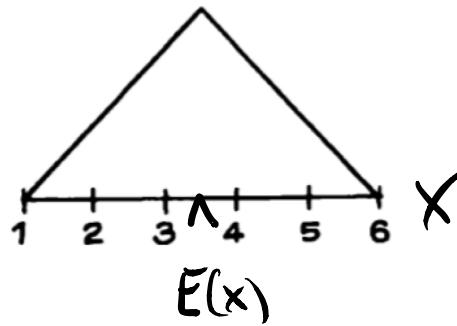
finish the proof. *order of addition doesn't matter*

$$\begin{aligned} E(XY) &= \sum_{\text{all } x, y} xy P(X=x, Y=y) = \sum_{\text{all } x} \sum_{\text{all } y} xy P(X=x) P(Y=y) \\ &= \left(\sum_{\text{all } x} x P(X=x) \right) \left(\sum_{\text{all } y} y P(Y=y) \right) = E(X)E(Y) \end{aligned}$$

③ 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 — too small
- b 1** — about right (I am not expecting you to calculate this)
- c 2 — too large

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

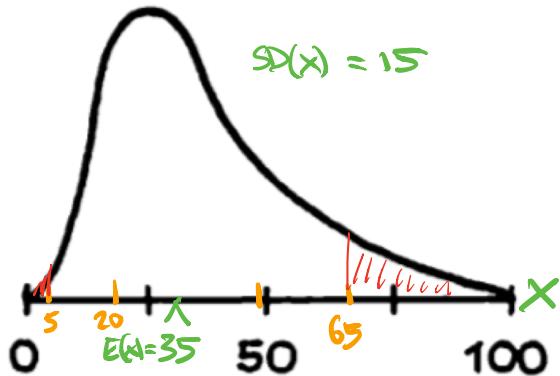
$$Var(x) = (SD(x))^2 = E((x - E(x))^2)$$

Chesnyshev's Inequality

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

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

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



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

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

What can you say about $P(X \geq 65) \leq \frac{1}{4}$

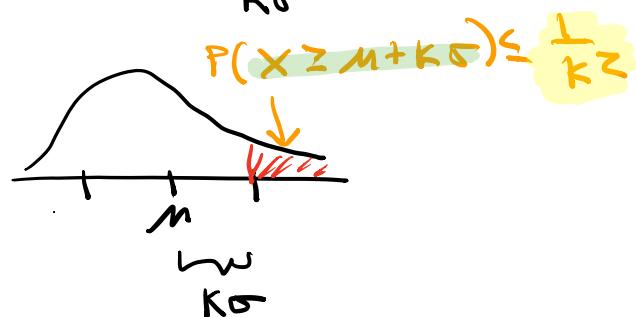
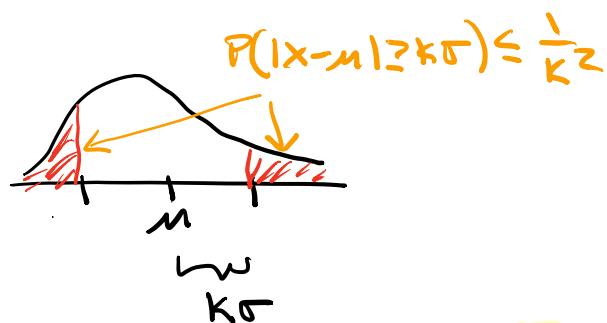
Since the distribution isn't symmetric or normal.

$$\mu = E(x)$$

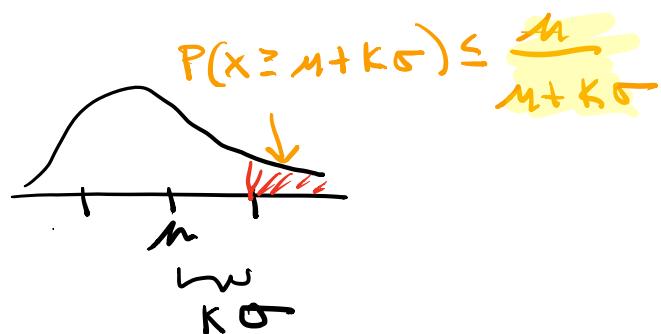
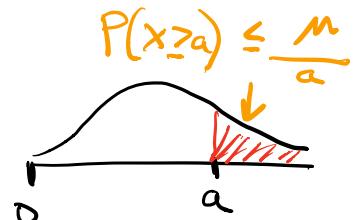
$$\sigma = SD(x).$$

Summary of
Tail bounds

Chebychev's inequality



Markov's inequality



Stat 134

Wednesday February 20 2019

1. A list of non negative numbers has an average of 1 and an SD of 2. Let p be the proportion of numbers over 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

$$P(X \geq 5) \leq \frac{1}{5}$$

smaller

$$P(X \geq 5) \leq \frac{1}{k^2} = \frac{1}{4}$$

\uparrow

$$\left. \begin{array}{l} m + k\sigma \\ = 1 + 2 \cdot 2 \\ = 5 \end{array} \right\} \Rightarrow k = 2$$