Stat 134 lec 12

warmup

(3 pts) Suppose that each year, Berkeley admits 15,000 students on average, with an SD of 5,000 students. Assuming that the application pool is roughly the same across years, find the smallest upper bound you can on the probability that Berkeley will admit at least 22,500 students in 2019. (Hint: you should be comparing two possible bounds.)

let X = # students Cal admits in 2019

M $\quad P(X \ge 22{,}500) \le \dfrac{15\,000}{22{,}500} = 2/3$

C $\quad P(X \ge 22{,}500) \le \dfrac{1}{(1.5)^2} = \boxed{4/9}$ ← smaller upper bound.

$15000 + 1.5(5000)$

$\Rightarrow k = 1.5$

last time

Sec 3.3 SD(X) is the average deviation from the mean

i.e $SD = \sigma = \sqrt{E((X-\mu)^2)}$

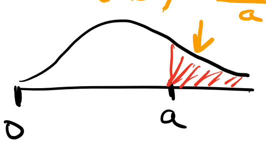
$Var = \sigma^2 = E((X-\mu)^2)$

often write $E(X-\mu)^2$

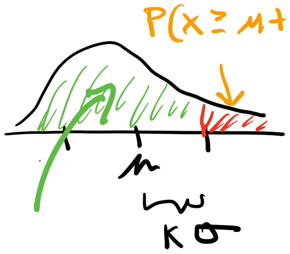
Tail bounds

Markov's inequality

$P(X \ge a) \le \frac{\mu}{a}$



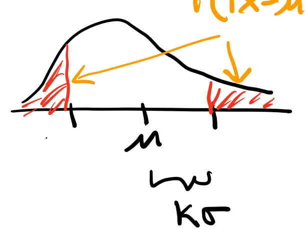
$P(X \ge \mu + k\sigma) \le \frac{\mu}{\mu + k\sigma}$



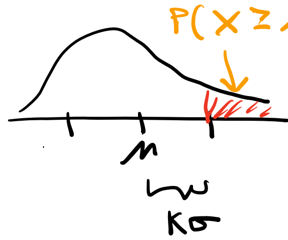
$P(X < \mu + k\sigma) \ge 1 - \frac{\mu}{\mu + k\sigma}$

Chebyshev's inequality

$P(|X-\mu| \ge k\sigma) \le \frac{1}{k^2}$



$P(X \ge \mu + k\sigma) \le \frac{1}{k^2}$



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 7 5. To get an upper bound for p, you should:

M: $\frac{1}{7}$

C: $7 = 1 + k \cdot 2$
$k = 3$
$\frac{1}{9}$

a Assume a normal distribution
b Use Markov's inequality
c Use Chebyshev's inequality
d none of the above

b	Markov E(x)/a, so the bound = 1/5. Chebyshev k = 2, so the bound is 1/4. Markov bound is lower/tighter, so we would prefer the Markov bound.

I don't understand the limitations of each inequality

M: $X \geq 0$ and know $E(X)$ and to be useful $a > \mu$

C: know $E(X)$, $SD(X)$ and to be useful $k > 1$.

Today

Sec 3.2

(1) Expectation of a function of a RV

Sec 3.3

(1) another formula for variance

(2) properties of variance

② <u>Sec 3.2</u> <u>Expectation of a function of a RV.</u>

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

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

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

<u>Picture</u>

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

$$2 \longrightarrow 3^2$$

$$3 \longrightarrow 3^3$$

$\vdots$

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

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

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

yes since $p > 2/3$

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

extra problem

ex

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

Find $E(X!)$

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

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

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

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

$E(X!) = \sum_{k=0}^{\infty} k!\,P(X=k) = \sum_{k=0}^{\infty} \cancel{k!}\frac{e^{-1/3}(1/3)^k}{\cancel{k!}} = e^{-1/3}\left(1+\frac{1}{3}+\frac{1}{3^2}+\dots\right)$

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

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

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

see appendix to notes

Thm if X and Y are independent

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

(3)

Sec 3.3 Another formula for Var(x).

Recall $E(cX) = cE(X)$

so $E(E(X)X) = E(X)E(X)$

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

$$\overset{\text{FOIL}}{=} E(X^2 - 2E(x)X + (E(x))^2)$$

$E(cX) = cE(X)$

$$= E(x^2) - \underbrace{2E(x)E(x) + (E(x))^2}_{= -E(x)^2}$$

$$\Rightarrow \boxed{Var(x) = E(x^2) - E(x)^2} \quad (E(x))^2$$

$$(\text{or}) \quad \boxed{E(x^2) = Var(x) + E(x)^2} \quad (E(x))^2$$

ex let $X = \begin{cases} 1 & \text{with prob } p \\ 0 & \text{with prob } q \end{cases}$

$$E(x^2) = \sum_{\text{all } x} x^2 P(X=x) = 1^2 \cdot p + \cancel{0 \cdot q} = \boxed{p}$$

Note $E(x) = p$

$$Var(x)\,? = p - p^2 = p(1-p) = \boxed{pq}$$

ex: Let X be a nonnegative RV such that

$E(X) = 100 = Var(X)$

a) Can you find $E(X^2)$ exactly? If not what can you say.

$$E(X^2) = Var(X) + (E(X))^2$$
$$= 100 + 10{,}000 = \boxed{10{,}100}$$

b) Can you find $P(70^2 < X^2 < 130^2) = P(70 < X < 130)$ exactly? If not what can you say?

$\le \frac{1}{9}$

3 SD

$k = 3$

$$P(70 < X < 130) \ge 1 - \frac{1}{9} = \boxed{8/9}$$

$$P(70^2 < X^2 < 130^2) = P(70 < \pm X < 130)$$
$$= P(-130 < X < -70) + P(70 < X < 130)$$

$\parallel$

0 since X is nonnegative

tinyurl.com/feb13-2023

Stat 134

1. X is nonnegative random variable with $E(X) = 3$ and $SD(X) = 2$. True, False or Maybe:

$$P(X^2 \geq 40) \leq \frac{1}{3}$$

a True

b False

c Maybe

$= P(X \geq \sqrt{40})$ since X nonneg.

maybe

M: $P(X \geq \sqrt{40}) \leq \frac{3}{\sqrt{40}} = .47$

C: $P(X \geq \sqrt{40}) \leq \frac{1}{k^2} = \frac{1}{\left(\frac{\sqrt{40}-3}{2}\right)^2} = .36$

$\sqrt{40} = 3 + k \cdot 2$

$\Rightarrow k = \frac{\sqrt{40}-3}{2}$

maybe

We know $E(X^2)$ so lets try Markov for X^2:

M: $E(X^2) = Var(X) + (E(X))^2$

$= 4 + 9 = 13$

$P(X^2 \geq 40) \leq \frac{E(X^2)}{40} = \frac{13}{40} = .325 < 1/3$

true

Appendix

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

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

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

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

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

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

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

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

$\square$