Stat 134 lec 10

warm up

(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 $P(X \geq 22{,}500) \leq \frac{15\,000}{22{,}500} = 2/3$

C $P(X \geq 22{,}500) \leq \frac{1}{(1.5)^2} = \boxed{4/9}$ ← smaller upper bound,

15000 + 1.5 (5000)

$\Rightarrow k = 1.5$

Last time

Sec 3.3

Markov Inequality:

let $X \geq 0$

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

need $X \geq 0$ and $E(X)$

Chebyshev Inequality:

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

need $E(X)$, $SD(X)$

Proof of Chebyshev

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

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

By Markov

$P(Y \geq a) \leq \frac{E(Y)}{a}$ for $Y \geq 0$, $a > 0$

Note that $P(A^2 < B^2) = P(|A| < |B|)$ since $A^2 < B^2$ iff $|A| < |B|$ for numbers A, B.

Hence,

$$P(|X-E(X)| \geq K\,SD(X))$$

(Here $A = X-E(X)$
$B = K\,SD(X)$)

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

$$= \frac{\cancel{Var(X)}}{K^2\,\cancel{Var(X)}} = \frac{1}{K^2}$$

$\square$

Sec 3.3

1. another formula for variance
2. Properties of variance
3. Central Limit Thm (CLT)

①

Sec 3.3 Another formula for Var(X).

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

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

$$\mathrm{Var}(X) = E\left((X - E(X))^2\right)$$

FOIL

$$= E\left(X^2 - 2E(X)X + (E(X))^2\right)$$

$E(cX) = cE(X)$

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

$$\Rightarrow \boxed{\mathrm{Var}(X) = E(X^2) - E(X)^2}$$

$(E(X))^2$

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

$(E(X))^2$

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 $E(X) = p$

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

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?

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

$3 \cdot SD$

$k = 3$

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

$$P(70^2 < X^2 < 130^2) = P(70 < \pm X < 130)$$
$$= \underbrace{P(-130 < X < -70)}_{=0} + P(70 < X < 130)$$

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

$P(X^2 \geq 40) = P(X \geq \sqrt{40})$ since X nonneg.

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

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

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

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

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

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

$= 4 + 9 = 13$

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

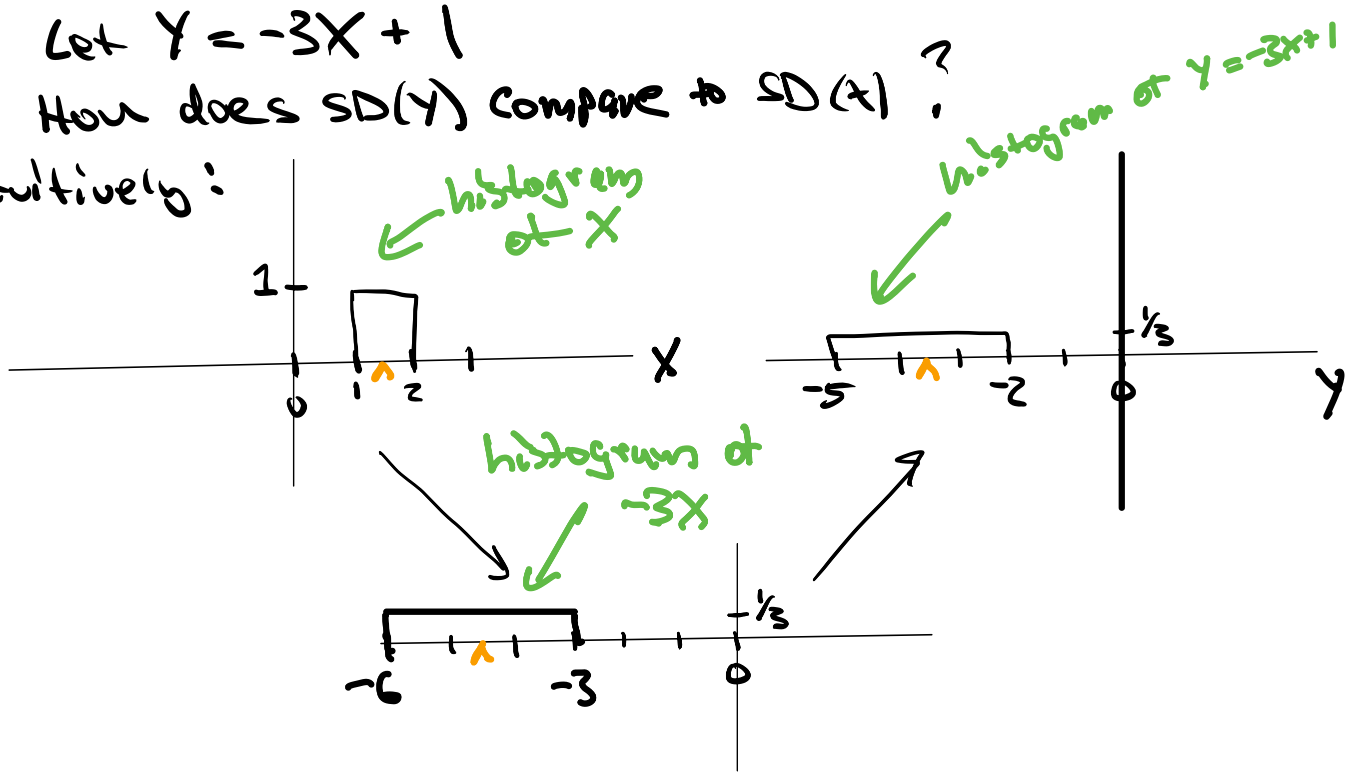
true

② Property of Variance

Let $Y = -3X + 1$

How does $SD(Y)$ compare to $SD(X)$?

intuitively:



$$\boxed{\begin{aligned} SD(aX+b) &= |a|\,SD(X) \\ Var(aX+b) &= a^2\,Var(X) \end{aligned}}$$

③

Central Limit Thm (CLT)

Let $S_n = X_1 + \cdots + X_n$ where $X_1, \ldots, X_n$ are iid RVs, $E(X) = \mu$, $Var(X) = \sigma^2$.

Then,

$$S_n \approx N(n\mu, n\sigma^2) \quad \text{for "large" } n.$$

approximately

often ≥ 10

ex

Let $X_1, X_2, \ldots X_{10}$ be i.i.d. Poisson(1).

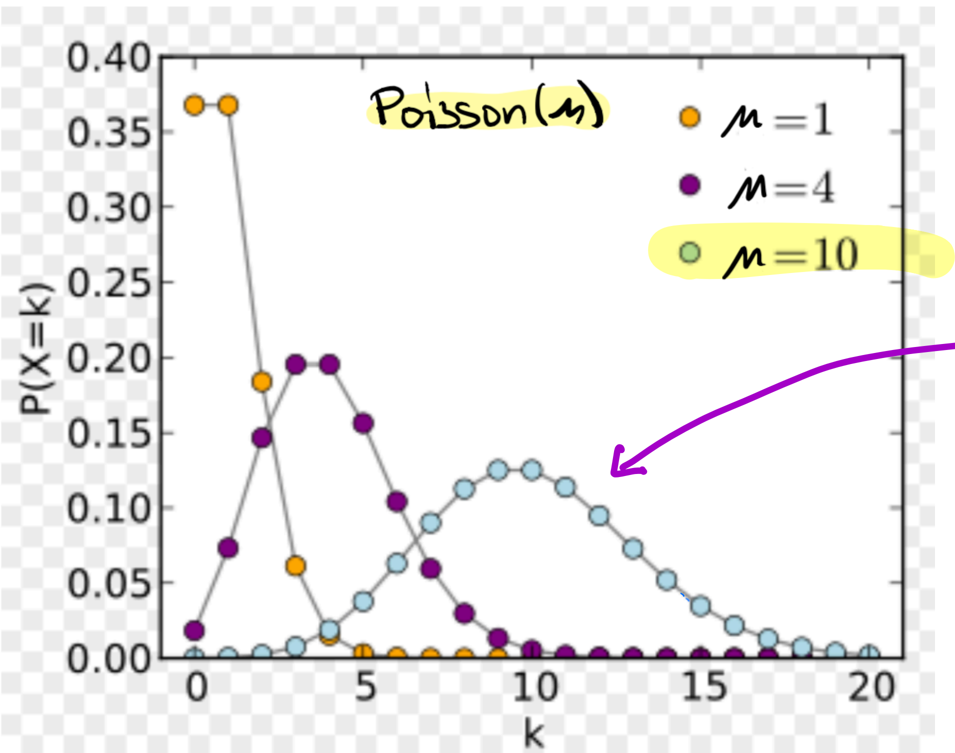
Let $S_{10} = X_1 + \cdots + X_{10}$

Facts

if $X \sim Pois(1)$, $E(X) = 1$, $Var(X) = 1$

$$E(S_{10}) = E(X_1 + \cdots + X_{10}) = 10E(X_1) = 10$$

$$Var(S_{10}) = Var(X_1 + \cdots + X_{10}) = 10Var(X_1) = 10$$



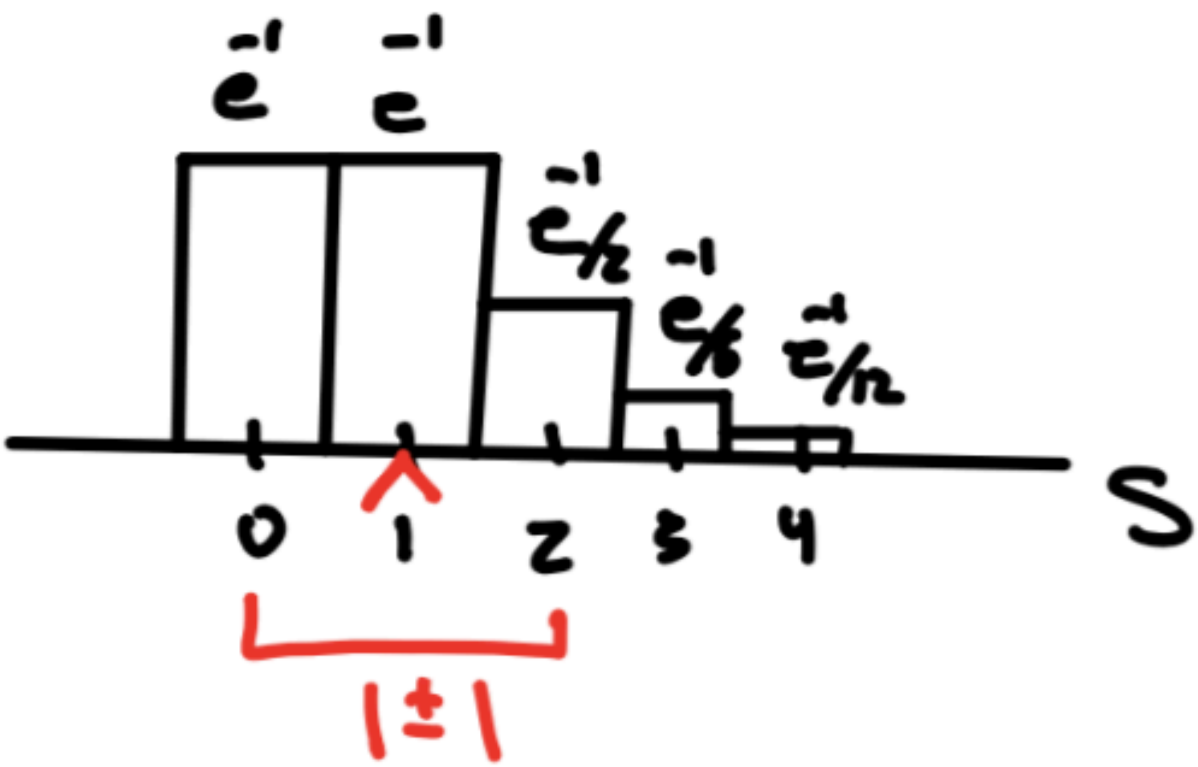
Pois(10) is a sum of 10 iid Pois(1) and is approx $N(10, 10)$

μ σ^2

extra example

1. [10 pts] Let $X_1, X_2, \ldots$ be independent and identically distributed, each with Poisson(0.01) distribution. Let $S = X_1 + X_2 + \cdots + X_{100}$. Sketch the probability histogram of S. Give values for $E(S)$ and $SD(S)$, and show the interval $E(S) \pm SD(S)$ on the horizontal axis. (Hint: what distribution is sum of independent poisson?)

Solution: $S \sim Pois(1)$ since the sum of independent Poisson is Poisson. $E(S) = 1$ and $var(S) = 1$ so $SD(S) = 1$ Using the Poisson formula we have $P(S=0) = e^{-1}, P(S=1) = e^{-1}, P(S=2) = e^{-1}/2$ etc.



Notice that $n=100$ isn't enough here for CLT to make S approx normal since $X_1, \ldots, X_n \overset{iid}{\sim} Pois(.01)$ is so not normal.