Stat 134 lec 11

Last time:

sec 3.2 Expectation

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

If X is a count, X can be written as a sum of indicators

$$X = I_1 + I_2 + \cdots + I_n \quad , \quad I_j = \begin{cases} 1 & \text{prob } p \\ 0 & \text{prob } 1-p \end{cases} \quad 1 \le j \le n$$

$$E(I_j) = 1 \cdot p + 0 \cdot (1-p) = p.$$

Idea Even if indicators are dependent the expectation of each indicator is an unconditional probability.

Then $E(X) = n \cdot p$

ex $X = \#$ aces in a poker hand from a deck of cards

$0, 1, 2, 3, 4$

Find $E(X)$.

$$X = I_1 + I_2 + I_3 + I_4 + I_5 \qquad I_j = \begin{cases} 1 & \text{if } j^{th} \text{ card is ace} \\ 0 & \text{else} \end{cases} \quad p = \frac{4}{52}$$

$$E(X) = 5 \cdot \left(\frac{4}{52}\right)$$

Today

Almost finish sec 3.2
- example indicators
- tail sum
- Markov's Inequality.

Sec 3.2 Another indicator example

ex Consider a well shuffled deck. What is the expected number of cards before the ace of hearts?

X = expected number of cards before the ace of hearts.

$0, 1, 2, \dots, 51$

Intuitively it should be half the deck (minus the ace of hearts card), by symmetry. There is no reason the expected number of cards before the ace of hearts should be different than the expected number of cards after the Ace of hearts.

Here is how to solve the problem using indicators:

There are 51 cards without the ace of hearts. Write

$$X = I_1 + I_2 + \dots + I_{51}, \quad I_j = \begin{cases} \overset{p}{1} & \text{if } j^{th} \text{ card before Ace of hearts} \\ 0 & \text{else} \end{cases}$$

To find p note that it is equally likely that the j^{th} card goes before or after the ace of hearts. Hence $p = 1/2$.

Ace of hearts

non ace of heart so $p = \frac{1}{2}$ $\Rightarrow$ $\boxed{E(X) = 51 \cdot \frac{1}{2}}$

Stat 134

Chapter 3 Monday September 17 2018

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

X = # cards before the 1^{st} ace

$0, 1, 2, \ldots, 48$

$4 \times 12 = 48$ nonaces

$$\underbrace{\overset{1}{2}\ \overset{2}{3}\ \overset{3}{4}\ \overset{4}{5}\ \overset{5}{6} \cdots \overset{10}{J}\ \overset{11}{Q}\ \overset{12}{K}}_{\text{hearts}}\ \underbrace{\overset{13}{2}\ \overset{14}{3}\ \overset{15}{4} \cdots}_{\text{spades}}\ \underbrace{\qquad}_{\text{clubs}}\ \underbrace{\cdots \overset{47}{Q}\ \overset{48}{K}}_{\text{diamonds}}$$

$X = I_1 + I_2 + \cdots + I_{48}$

$$I_j = \begin{cases} 1 & \text{if } j^{th} \text{ nonace before } 1^{st} \text{ ace} \quad p \\ 0 & \text{else} \end{cases}$$

$$\_\_\, A_1 \,\_\_\, A_2 \,\_\_\, A_3 \,\_\_\, A_4 \,\_\_$$

$P = 1/5 \Rightarrow \boxed{E(X) = 48\left(\frac{1}{5}\right)}$

Sec 3.2 Tail sum formula for expectation

Suppose X is a count $1, 2, 3, \ldots$

$$E(X) = 1 \cdot \underbrace{P(X=1)}_{p_1} + 2 \cdot \underbrace{P(X=2)}_{p_2} + \cdots$$

$$= \begin{array}{lcccc} & P(X\geq 1) & P(X\geq 2) & P(X \geq 3) & \\ & p_1 & & & \\ & p_2 & p_2 & & \\ & p_3 & p_3 & p_3 & \\ & p_4 & p_4 & p_4 & p_4 \\ + & \vdots & & \ddots & \end{array}$$

$$= P(X \geq 1) + P(X \geq 2) + P(X \geq 3) + \cdots$$

Tail sum formula,

ex A fair die is rolled 10 times

let $X = \max(X_1, \ldots, X_{10})$ — $1,2,3,4,5,6$

Find $E(X)$

$$P(X \geq k) = 1 - P(X < k)$$

$$= 1 - P(X_1 < k, X_2 < k, \ldots, X_{10} < k)$$

independence

$$= 1 - P(X_1 < k)P(X_2 < k) \cdots P(X_{10} < k)$$

$$= 1 - \left(P(X_1 < k)\right)^{10}$$

$$= 1 - \left(\frac{k-1}{6}\right)^{10}$$

$$E(X) = P(X\geq 1) + P(X\geq 2) + P(X \geq 3) + \cdots + P(X\geq 6)$$

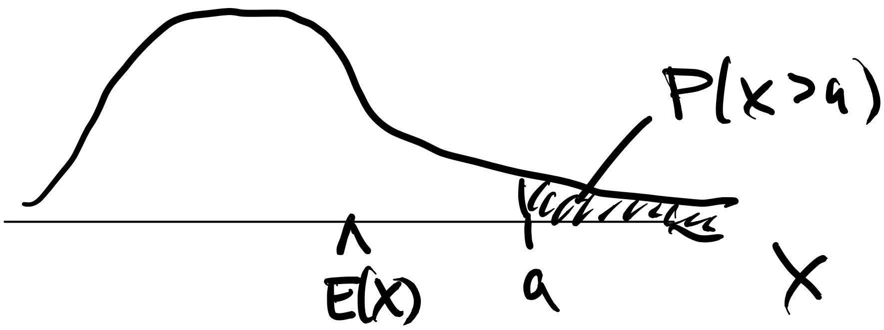
$$= 1 + \left[1-\left(\tfrac{1}{6}\right)^{10}\right] + \left[1-\left(\tfrac{2}{6}\right)^{10}\right] + \cdots + \left[1-\left(\tfrac{5}{6}\right)^{10}\right]$$

$$= 6 - \frac{1}{6^{10}}\left[1^{10} + 2^{10} + \cdots + 5^{10}\right] = 5.82$$

Markov's Inequality

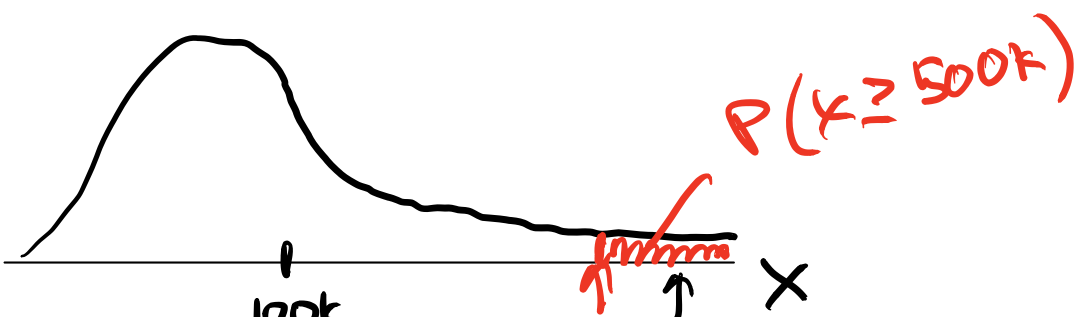
Let $X \geq 0$, $a > 0$

$E(X)$ is the center of distribution



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

ex let X be yearly income of Bay area residents. $E(X) = \$100k$



500k google gugs

Find an upper bound for $P(X \geq 500k)$

$$P(X \geq 500k) \leq \frac{100}{500} = \frac{1}{5}$$

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

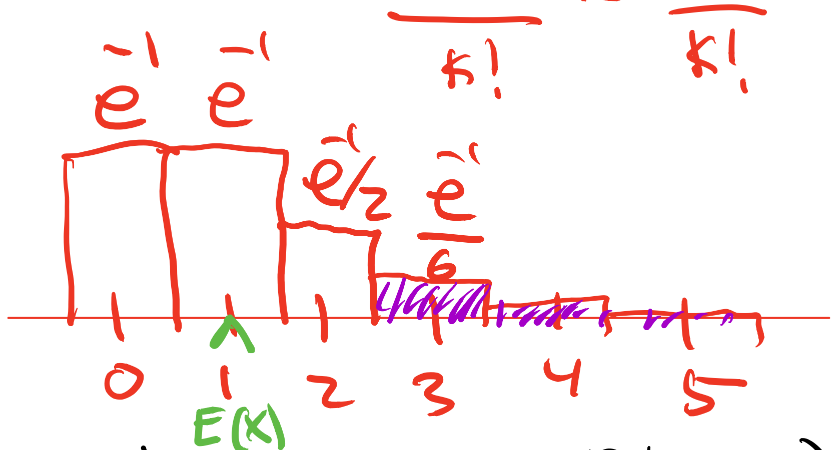
Let $S = X_1 + X_2 + \dots + X_{100}$

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

soln Remember that $S = X_1 + \dots + X_n \sim \text{Pois}(100 \cdot \frac{1}{100}) = \text{Pois}(1)$, since the sum of independent Pois(μ) is Po

$$S \sim \text{Pois}(1), \quad E(S) = 1$$

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



By Markov's inequality $P(S \geq 3) \leq \boxed{\frac{1}{3}}$ E(S)

One can of course find $P(S \geq 3)$ exactly:

$$P(S \geq 3) = 1 - P(S=0) - P(S=1) - P(S=2) = 1 - e^{-1}(1 + 1 + \tfrac{1}{2})$$

$$= .08 < 1/3 \checkmark$$

Property of expectation

if $X \le Y$ then $E(X) \le E(Y)$.

eg let $X \ge 0$, and $Y = 2X$

$$X \le Y \Rightarrow E(X) \le E(Y) = E(2X) = 2E(X).$$

Markovs inequality

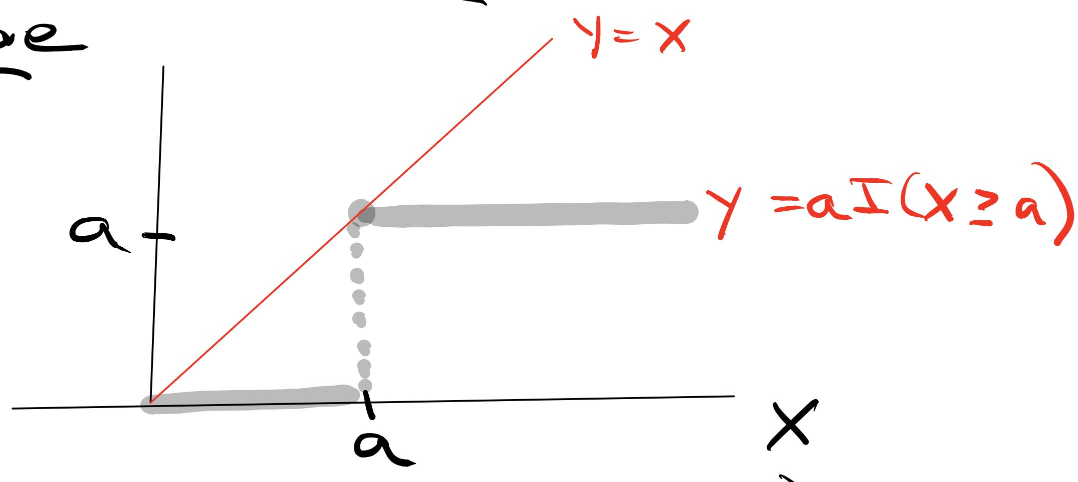
If $X \ge 0$, then $P(X \ge a) \le \dfrac{E(X)}{a}$ for every $a > 0$.

Pf/

Let $I(X \ge a)$ be an indicator equal to 1 when $X \ge a$, Then

$$aI(X \ge a) = \begin{cases} a & \text{if } X \ge a \\ 0 & \text{else.} \end{cases}$$

Picture



$$aI(X \ge a) \le X \Rightarrow E(aI(X \ge a)) \le E(X)$$

$$E(aI(X \ge a)) = aE(I(X \ge a)) = aP(X \ge a)$$

$$\Rightarrow P(X \ge a) \le \frac{E(X)}{a} \qquad \square$$