Stat 134 lec 14

Last time

<u>sec 3.3</u> $Var(X+Y) = Var(X) + Var(Y)$ if X, Y independent

<u>Central Limit Thm</u> (CLT)

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

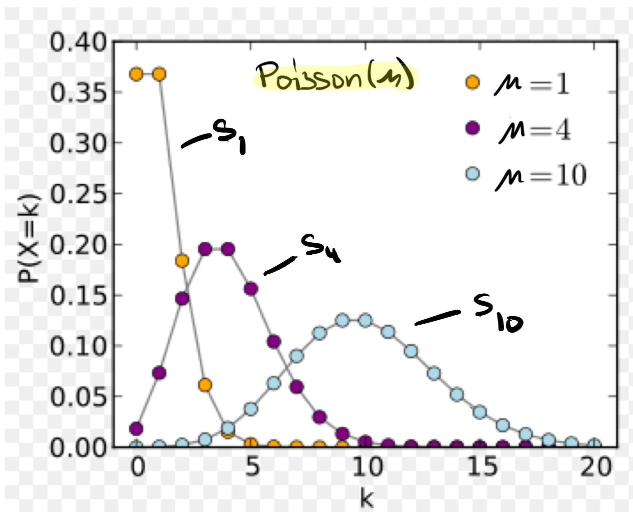
Then,

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

approximately

often ≥ 10

ex



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

<u>tail bounds</u> Markov $X \geq 0$, $a > 0$ $P(X \geq a) \leq \frac{E(X)}{a}$

Chebyshev $P(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}$

Today

1. Go over concept test responses from last time
2. <u>Sec 3.6</u> (next time sec 3.4) Calculating variance of a sum of dependent indicators
3. <u>sec 3.6</u> Hypergeometric distribution

1. X is 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

b	using chebyshev on P(X >= 40) we get k = 1.66. thus, it is less than 1/1.66^2, which is greater than 1/3

Note that we can't use Chebyshev since

$P(X^2 \geq 40) = P(X \geq \sqrt{40})$ if X not nonnegative,

(but if X was nonneg) $\quad \| \quad 3 + K \cdot 2$

$$K = \frac{\sqrt{40} - 3}{2} = 1.66$$

$$\leq \frac{1}{(1.66)^2} = .36 \quad \Rightarrow \text{maybe.}$$

a	Var(X) = (SD(X))^2 = 2^2 = 4 E(X^2) = Var(X) + E(X)^2 = 4 + 3^2 = 13 P(X^2 >= 40) <= 13/40 < 1/3 Thus, the statement is true.

Sec 3.6 Var of sum of dependent indicators.

But first:
The variance of a sum of independent indicators

ex $I = \begin{cases} 1 & \text{Prob } p \\ 0 & \text{prob } 1-p \end{cases}$

$Var(I) = E(I^2) - E(I)^2$

$E(I^2) = 1^2 \cdot p + 0^2(1-p) = p$

$E(I) = p$

$\Rightarrow Var(I) = p - p^2 = p(1-p)$

ex $X \sim Bin(n,p)$

$X = I_1 + I_2 + \dots + I_n$

$E(X) = n E(I_1) = \boxed{np}$

$Var(X) = n Var(I_1) = \boxed{np(1-p)}$

$SD(X) = \sqrt{np(1-p)}$

Next we will look at an example where X is a sum of dependent indicators ex

14. A building has 10 floors above the basement. If 12 people get into an elevator at the basement, and each chooses a floor at random to get out, independently of the others, at how many floors do you expect the elevator to make a stop to let out one or more of these 12 people?

X = number of elevator stops.

$p_1 = 1 - \left(\frac{9}{10}\right)^{12}$

a) Find $E(X)$

$X = I_1 + \cdots + I_{10}$ $\quad I_2 = \begin{cases} 1 & \text{if stop } 2^{nd} \text{ floor} \\ 0 & \text{else} \end{cases}$

$E(X) = 10p_1$

b) Find $Var(X)$.

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

$E(X^2) = E\left((I_1 + \cdots + I_{10})^2\right) = \sum_{1 \le i,j \le 10} E(I_i I_j)$

Picture

$$\begin{array}{c|ccccc} & I_1 & I_2 & I_3 & \cdots & I_{10} \\ \hline I_1 & I_1I_1 & I_1I_2 & & & \\ I_2 & I_2I_1 & I_2I_2 & & & \\ \vdots & & & \ddots & & \\ I_{10} & & & & & I_{10}I_{10} \end{array}$$

$I_1 = \begin{cases} 1 & \text{if stop } 1^{st} \text{ floor} \\ 0 & \text{else} \end{cases}$

$I_2 = \begin{cases} 1 & \text{if stop } 2^{nd} \text{ floor} \\ 0 & \text{else} \end{cases}$

p_{12}

$I_{12} = I_1 I_2 = \begin{cases} 1 & \text{if stop at } 1^{st} \text{ \underline{and} } 2^{nd} \text{ floor} \\ 0 & \text{else} \end{cases}$

Picture

$$\begin{array}{c|cccc}
 & I_1 & I_2 & I_3 \cdots & I_{10} \\ \hline
I_1 & I_1 & I_{12} & & \\
I_2 & I_{12} & I_2 & & \\
\vdots & & & \ddots & \\
I_{10} & & & & I_{10}
\end{array}$$

$= I_{10\,10}$

$9 \cdot 10$ nondiag entries

$$P_{12} = 1 - \text{Prob}\left(\begin{array}{l}\text{dont stop at } 1^{st} \text{ floor or} \\ \text{dont stop at } 2^{nd} \text{ floor}\end{array}\right)$$

$$= 1 - \left[\left(\frac{9}{10}\right)^{12} + \left(\frac{9}{10}\right)^{12} - \left(\frac{8}{10}\right)^{12}\right]$$

$$= \boxed{1 - 2\left(\frac{9}{10}\right)^{12} + \left(\frac{8}{10}\right)^{12}}$$

Prob dont stop at 1^{st} and 2^{nd} floor

Note $P_{11} = P_1$ since P_{11} is chance you get off at 1^{st} floor and 1^{st} floor.

$$E(X^2) = \underbrace{10\,E(I_1)}_{\text{diagonals}} + \underbrace{9 \cdot 10\,E(I_{12})}_{\text{off diagonals}}$$

$P_1 = E(I_1)$, $I_{11} = I_1$, $P_{12} = E(I_{12})$

$$(E(X))^2 = (10P_1)^2$$

$$\boxed{\text{Var}(X) = 10P_1 + 9 \cdot 10 P_{12} - (10P_1)^2}$$

Recall multinomial distribution

roll a die 10 times

$$P(2 \text{ ones}, 3 \text{ twos}) = \binom{10}{2,3,5}\left(\frac{1}{6}\right)^2\left(\frac{1}{6}\right)^3\left(\frac{4}{6}\right)^5$$

$$\binom{10}{2,3,5} = \frac{10!}{2!\,3!\,5!}$$

Stat 134

Monday February 24 2019

1. A fair die is rolled 14 times. Let X be the number of faces that appear exactly twice. Which of the following expressions appear in the calculation of $Var(X)$

   a $14 * 13 * \binom{14}{2,2,10}(1/6)^2(1/6)^2(4/6)^{10}$

   b $\binom{14}{2}(1/6)^2(5/6)^{12}$

   c more than one of the above

   d none of the above

$P_1 = \binom{14}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^{12}$

X = # face that appear twice

$X = I_1 + I_2 + \cdots + I_6$

$I_2 = \begin{cases} 1 & 2^{nd} \text{ face twice} \\ 0 & \end{cases}$

$P_{12} = \binom{14}{2,2,10}\left(\frac{1}{6}\right)^2\left(\frac{1}{6}\right)^2\left(\frac{4}{6}\right)^{10}$

$I_{12} = \begin{cases} 1 & 1^{st} \text{ and } 2^{nd} \text{ face twice} \\ 0 & \text{else} \end{cases}$

$E(X^2) = 6P_1 + 6 \cdot 5 P_{12}$

$Var(X) = E(X^2) - (E(X))^2 = 6P_1 + 30P_{12} - (6P_1)^2$

sec 3.6 Hypergeometric (N, G, n)

ex

A deck of cards has G aces.

X = # aces in n cards drawn without replacement from a deck of N cards.

$N = 52$
$G = 4$
$n = 5$

a) Find $E(X)$

$X = I_1 + I_2 + \cdots + I_5$

$I_2 = \begin{cases} 1 & \text{if } 2^{nd} \text{ card ace} \\ 0 & \text{else} \end{cases}$

$p_1 = \frac{G}{N}$

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

$= \boxed{n\left(\frac{G}{N}\right)}$

b) Find $Var(X)$.

$p_{12} = G/N \, \frac{G-1}{N-1}$

$I_{12} = \begin{cases} 1 & \text{if } 1^{st} \text{ and } 2^{nd} \text{ ace} \\ 0 & \text{else} \end{cases}$

$Var(X) = \underbrace{np_1 + n(n-1)p_{12}}_{E(X^2)} - \underbrace{(np_1)^2}_{E(X)^2}$