Stat 134 lec 13

warmup 9:00-9:10

Let X = number of sixes in 7 tosses of a fair die.

a) write X as a sum of indicators

$X = I_1 + I_2 + \cdots + I_7$

b) Find $Var(X)$

$I_2 = \begin{cases} 1 & \text{if } 2^{nd} \text{ roll is a } 6 \\ 0 & \text{else} \end{cases}$ 1/6

$Var(X) = Var(I_1 + \cdots + I_7) = Var(I_1) + \cdots + Var(I_7)$

since indicators are independent.

recall, $Var(I_1) = p(1-p) = pq$

$\Rightarrow Var(X) = 7 \cdot pq = \boxed{7(1/6)(5/6)}$

$X \sim Bin(n,p)$

$Var(X) = npq$

If n is large p small and $np \to \mu$

then $X \sim Pois(\mu)$

$Var(\underset{Pois(\mu)}{X}) = \mu$

Last time

Sec 3.3 $\quad Var(X) = E\left(\left(X - E(X)\right)^2\right)$

or $\quad Var(X) = E(X^2) - \left(E(X)\right)^2$

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

$Var(I) = pq$

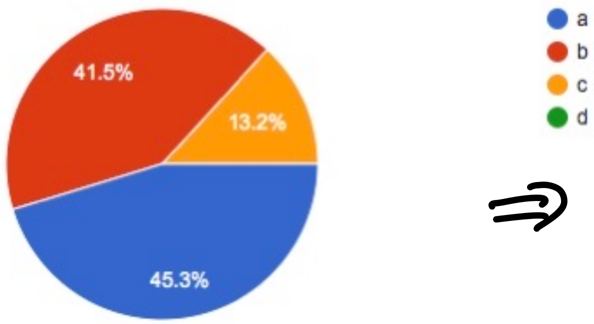
see p 193 Pitman for the proof.

Thm $\quad Var(X+Y) = Var(X) + Var(Y)$ if X, Y are independent.

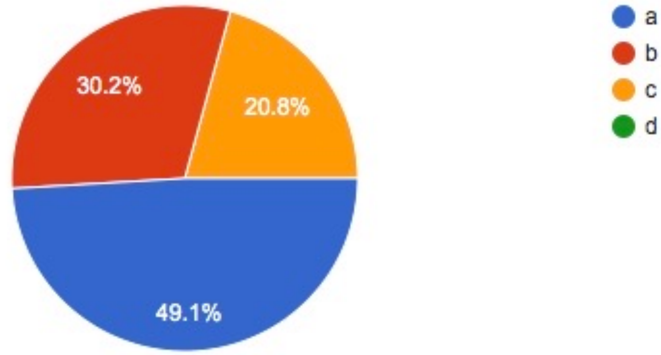
ex $\quad X \sim Bin(n,p)$

$Var(X) = npq$

$SD(X) = \sqrt{npq}$



$\Rightarrow$



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

b	Both Markov's and Chebyshev's give values bigger than 1/3 when we think about P(X>= sqrt(40))

M: $P(X \geq \sqrt{40}) \leq 3/\sqrt{40} \Rightarrow$ maybe.

a	We can solve for E[x^2] and using Markov Inequality, we go 1/4 and it is true for both that it is less than 1/4 and 1/3.

Today

⓪ Property of variance

① Sec 3.3 Central Limit theorem (CLT)

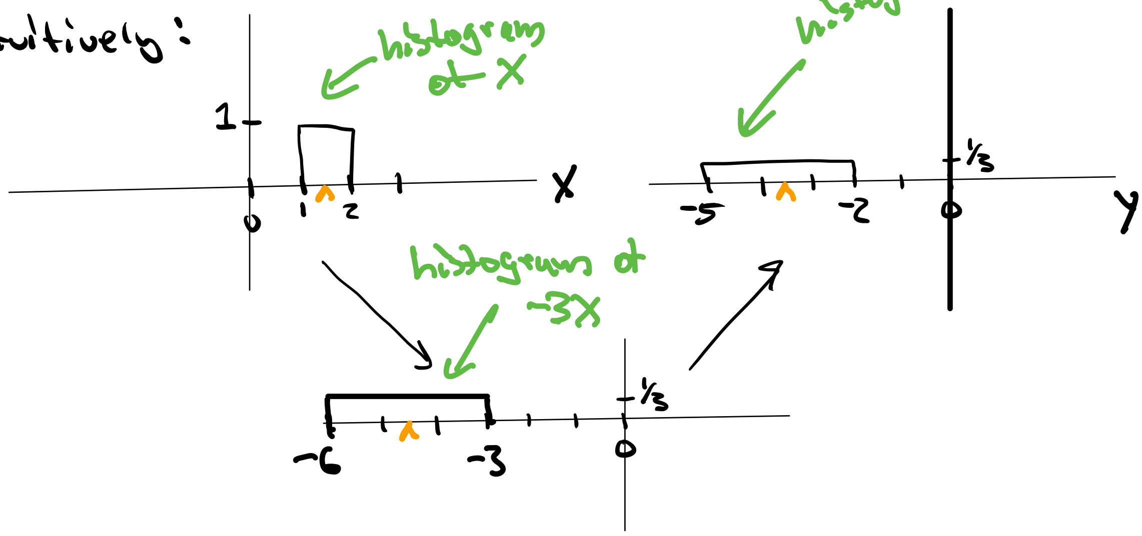
② Sec 3.6 (next time sec 3.4) Calculating the variance of a sum of dependent indicators.

⓪ Property of Variance

Let $Y = -3X + 1$

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

intuitively:



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

① **Sec 3.3**

See appendix to these notes for an example

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,

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

approximately

often ≥ 10

(2) Sec 3.6 Var of 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.

$p_1 = 1 - (9/10)^{12}$

$I_2 = \begin{cases} 1 & \text{if at least 1 person gets off at 2nd floor} \\ 0 & \text{else} \end{cases}$

a) Find $E(X)$ $\quad X = I_1 + \cdots + I_{10}$

$E(X) = 10 p_1$

b) Find $Var(X)$. $\quad X = I_1 + \cdots + I_{10}$

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

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

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

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

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

$\parallel$

I_{12}

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

j

$p_{12} = 1 -$ Prob no one gets off at 1st or 2nd floor,

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

$E(X^2) = 10 E(I_1) + 9 \cdot 10 E(I_{12}) = 10 p_1 + 9 \cdot 10 p_{12}$

$Var(X) = \boxed{10 p_1 + 9 \cdot 10 p_{12} - (10 p_1)^2}$

Summary

Identically Distributed

Variance of sum of dependent i.d. indicators

$X = I_1 + \cdots + I_n$

$p_1 = E(I_1)$

$p_{12} = E(I_{12}) = E(I_1 I_2)$

$E(X) = np_1$

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

Variance of sum of i.d. independent indicators

$X = I_1 + \cdots + I_n$

$p_1 = E(I_1)$

$p_{12} = p_1 \cdot p_2 = p_1^2$

$Var(X) = \underbrace{np_1 + n(n-1)p_1^2}_{E(X^2)} - \underbrace{(np_1)^2}_{E(X)^2} = np_1 - np_1^2$

$= np_1(1-p_1)$

tinyurl.com/sept23-2022

Stat 134

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 $Var(X)$

should be 6.5

a $14 * 13 * \binom{14}{2,2,10}(1/6)^2(1/6)^2(4/6)^{10}$ — P_{12}

b $\binom{14}{2}(1/6)^2(5/6)^{12}$ — P_1

c more than one of the above

d none of the above

X = # faces that appear twice

$X = I_1 + \cdots + I_6$

P_1

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

P_{12}

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

$$Var(X) = \underbrace{nP_1 + n(n-1)P_{12}}_{E(X^2)} - \underbrace{(nP_1)^2}_{E(X)^2}$$

Extra Problem:

6. A drawer contains s black socks and s white socks, where s is a positive integer. I pull two socks out at random without replacement and call that my first pair. Then I pull two socks out at random without replacement from the remaining socks in the drawer, and call that my second pair. I proceed in this way till I have s pairs and the drawer is empty.

Let D be the number of pairs in which the two socks are of different colors.

a) Find $E(D)$.

b) Find $Var(D)$.

$$P_1 = \frac{\overset{w}{\binom{s}{1}}\overset{B}{\binom{s}{1}}}{\binom{2s}{2}}$$

Soln

a) $D = I_1 + \cdots + I_s$ where $I_2 = \begin{cases} 1 & \text{if } 2^{nd} \text{ pair different} \\ 0 & \end{cases}$

$$\Rightarrow E(D) = s \cdot \frac{\binom{s}{1}\binom{s}{1}}{\binom{2s}{2}}$$

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

$$P_{12} = \frac{\binom{s}{1}\binom{s}{1}}{\binom{2s}{2}} \cdot \frac{\binom{s-1}{1}\binom{s-1}{1}}{\binom{2s-2}{2}}$$

then

$$Var(D) = sP_1 + s(s-1)P_{12} - (sP_1)^2$$

Appendix

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,

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

approximately

often ≥ 10

ex

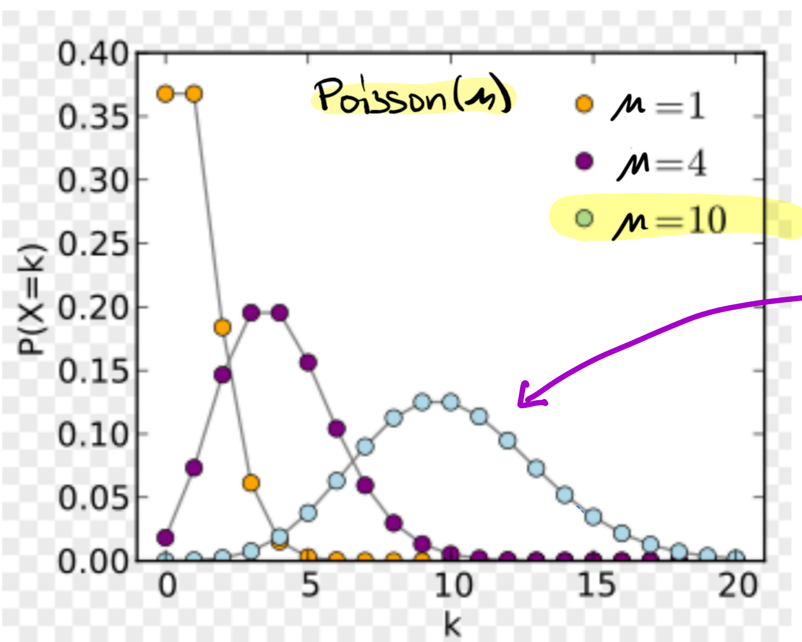
Let $X_1, X_2, \ldots X_{10}$ be i.i.d. Poisson(1).

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