Stat 134 lec 14

warmup 11:00-11:10

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

a) write X as a sum of indicators

$X = I_1 + \dots + I_7$ where $I_2 = \begin{cases} 1 & \text{if } 2^{nd} \text{ roll is } 6 \\ 0 & \text{else} \end{cases}$

$p = 1/6$

b) Find $Var(X)$

$Var(I_1 + \dots + I_7) = Var(I_1) + Var(I_2) + \dots + Var(I_7)$

$= 7 Var(I_1) = \boxed{7pq}$

c) let $X \sim Bin(n,p)$

$Var(X) = \boxed{npq}$

d) let $X \sim Bin(n,p)$ with n large and p small and $np \to \mu$.

Then X is approx. $Pois(\mu)$.

$Var(X) \approx \underbrace{np}_{\mu}\underbrace{q}_{1} = \boxed{\mu}$

Last time

Sec 3.3 $\text{Var}(X) = E\left((X - E(X))^2\right)$

or $\text{Var}(X) = E(X^2) - (E(X))^2$

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

$\text{var}(I) = pq$

$\text{var}(X+Y) = \text{var}(X) + \text{var}(Y)$ if X and Y are independent.

ex $X \sim \text{Bin}(n,p)$

$\text{Var}(X) = npq$

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

Today

(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.

① Sec 3.3

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,

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

$\approx$ approximately

"large" n: often ≥ 10

ex

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

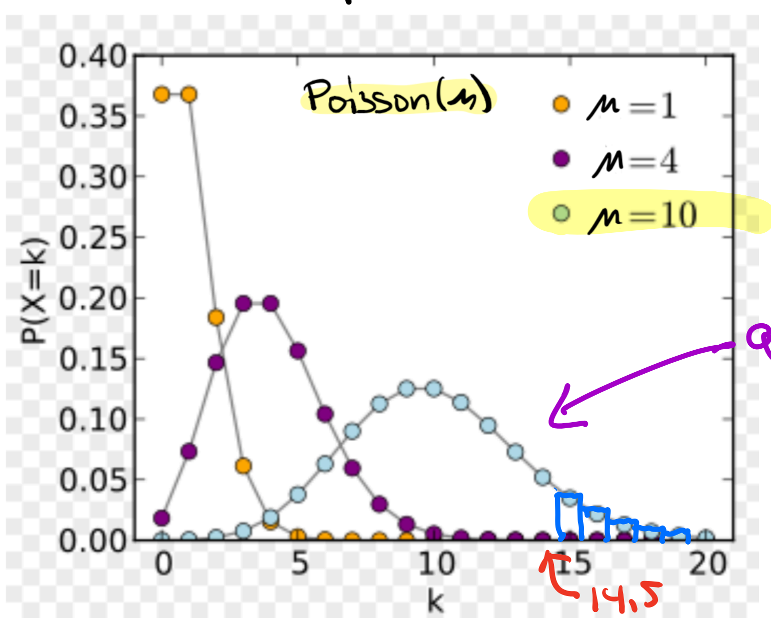
Let $S_{10} = X_1 + \cdots + X_{10}$

Facts

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

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

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



approx $N(10 \cdot 1, 10 \cdot 1)$

$E(X_1)$ → $10 \cdot 1$; $Var(X_1)$ → $10 \cdot 1$

Approximate $P(S_{10} \geq 15)$ with continuity correction:

$$P(S_{10} \geq 15) \approx \boxed{1 - \Phi\left(\frac{14.5 - 10}{\sqrt{10}}\right)}$$

(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?

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

X = number of elevator stops.

a) Find $E(X)$ $\quad X = I_1 + \dots + I_{10}$ where $I_z = \begin{cases} 1 & \text{if stop at } z^{\text{nd}} \text{ floor} \\ 0 & \text{else} \end{cases}$

$E(X) = 10p_1$

b) Find $Var(X)$. $\quad X = I_1 + \dots + I_{10}$ indicators are dependent.

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

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

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

$(AB)^c = A^c \cup B^c$

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

$p_{12} = 1 - P(\text{don't stop } 1^{\text{st}} \text{ flr. or don't stop } 2^{\text{nd}} \text{ floor}) = 1 - \left[\left(\frac{9}{10}\right)^{12} + \left(\frac{9}{10}\right)^{12} - \left(\frac{8}{10}\right)^{12}\right]$

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

$\quad \parallel$

I_{12}

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

$$\begin{array}{c|cccc} & I_1 & \cdots & \cdots & I_{10} \\ \hline I_1 & I_{11} & I_{12} & \cdots & \\ \vdots & \vdots & I_{22} & & \\ \vdots & \vdots & & \ddots & \\ I_{10} & & & & I_{1010} \end{array}$$

$I_{1010} = I_{10}$

$\leftarrow$ 100 indicators symmetric since $I_1 I_2 = I_2 \cdot I_1$

$E(X^2) = \underbrace{10\,E(I_1)}_{\text{diagonal}} + \underbrace{9 \cdot 10\,E(I_{12})}_{\text{non diagonal}} = 10p_1 + 10 \cdot 9\, p_{12}$

$(E(X))^2 = (10p_1)^2$

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

Summary

Identically Distributed

Variance of sum of dependent i.d. indicators

$X = I_1 + \dots + I_n$

$p_1 = E(I_1)$

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

$E(X) = np_1$

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

Variance of sum of i.d. independent indicators

$X = I_1 + \dots + I_n$

$p_1 = E(I_1)$

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

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

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 $Var(D)$.

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

a) $D = I_1 + \cdots + I_s$

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

$$E(D) = sP_1$$

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

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

Note: you could also find P_{12} using inclusion exclusion but more complicated,

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