

# STAT 88: Lecture 22

## Contents

Section 6.3: Markov's Inequality

Section 6.4: Chebyshev's Inequality

Section 7.1: Sums of Independent Random Variables

Warm up:

(a) State Markov's inequality.

(b) Is it possible that half of US flights have delay times at least 3 times the national average?

(a)  $X$  non negative RV,

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

(b)  $X =$  delay time of a flight  $\leftarrow$  random number.

$E(X) =$  national average.

$$P(X \geq \underbrace{3 \cdot E(X)}_c) \leq \frac{E(X)}{3 \cdot E(X)} = \frac{1}{3}$$

So it's not possible since  $\frac{1}{3}$  less than  $\frac{1}{2}$ .

Last time

"Expectation & Var put constraint on tail probability"

Upper bounds for tail probability:

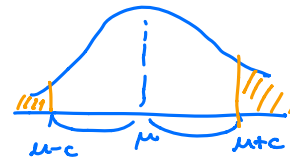
**Markov's inequality** For a non-negative random variable  $X$  and a positive constant  $c > 0$ ,

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



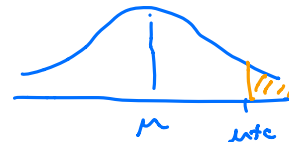
**Chebyshev's inequality** For a random variable  $X$  with mean  $\mu$  and SD  $\sigma$  and a positive constant  $c > 0$ ,

$$P(|X - \mu| \geq c) \leq \frac{\sigma^2}{c^2} = \frac{\text{Var}(X)}{c^2}.$$



One tail bound:

$$P(X - \mu \geq c) \leq \frac{\sigma^2}{c^2} = \frac{\text{Var}(X)}{c^2}.$$



Alternative formula for  $\text{Var}(X)$ :

- $\text{Var}(X) = E((X - \mu)^2)$ .
- $\text{Var}(X) = E(X^2) - E(X)^2$ . This implies  $E(X^2) = E(X)^2 + \text{Var}(X)$ .  
 $\geq (E(X))^2$  (Jensen's Inequality)

$$\text{Var}(X) = 0$$

$$\hookrightarrow E((X - \mu)^2) = 0$$

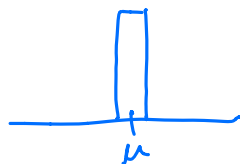
"

$$\sum_{\text{all } x} (x - \mu)^2 P(X = x) = 0$$

$$\hookrightarrow (x - \mu)^2 = 0 \text{ for all possible } x$$

$$\hookrightarrow x = \mu \text{ for all possible } x.$$

$X$  takes value  $\mu$  w/ probability 1.



Example: 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.

$$X = \# \text{ admits in 2019}, \quad E(X) = 15000, \quad SD(X) = 5000$$

$$P(X \geq 22500) ?$$

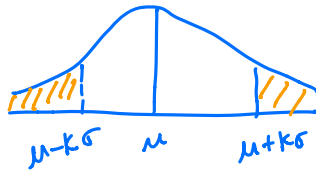
$$\textcircled{1} \quad P(X \geq \underbrace{22500}_c) \leq \frac{E(X)}{c} = \frac{15000}{22500} = \frac{2}{3} \quad (\text{Markov's})$$

$$\begin{aligned} \textcircled{2} \quad P(X \geq 22500) &= P(X - \underbrace{15000}_\mu \geq \underbrace{7500}_c) \\ &\leq \frac{\text{Var}(X)}{c^2} \\ &= \frac{5000^2}{7500^2} \\ &= \boxed{\frac{4}{9}} \quad (\text{Chebyshev's}) \end{aligned}$$

Example: Suppose a list of numbers  $x = \{x_1, \dots, x_n\}$  has mean  $\mu$  and standard deviation  $\sigma$ . Let  $k$  be the smallest number of standard deviations away from  $\mu$  we must go to ensure the range  $(\mu - k\sigma, \mu + k\sigma)$  contains at least 50% of the data in  $x$ . What is  $k$ ?

$$X = \text{number from } x$$

$$\begin{aligned} P(X \in (\mu - k\sigma, \mu + k\sigma)) \\ = 1 - P(X \notin (\mu - k\sigma, \mu + k\sigma)) &\geq \boxed{1 - \frac{1}{k^2}} = \frac{1}{2} \quad \leadsto k = \sqrt{2} \\ &\quad \underbrace{\hspace{1.5cm}}_{\substack{\text{"} \\ P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2} \quad (\text{Chebyshev's})}} \end{aligned}$$



Example: A list of non negative numbers has an average of 1 and an SD of 2. Let  $p$  be the proportion of numbers over 4. To get an upper bound for  $p$ , you should:

(a) Assume a binomial distribution.

$x$  = number from the list.

(b) Use Markov's inequality.

$E(X)=1, SD(X)=2.$

(c) Use Chebyshev's inequality.

$p = P(X \geq 5)$  Need upper bound.

(d) None of the above.

⊢ Markov's.  $p = P(X \geq 5) \leq \frac{E(X)}{5} = \boxed{\frac{1}{5}}$

⊢ Chebyshev's.  $p = P(X \geq 5)$

$$= P(X - \underset{\mu}{1} \geq \underset{c}{4})$$

$$\leq \frac{\text{Var}(X)}{c^2} = \frac{4}{16} = \frac{1}{4}$$

Example: Let  $X$  be a non negative random variable such that  $E(X) = 100 = \text{Var}(X)$ .

(a) Can you find  $E(X^2)$  exactly? If not what can you say?

(b) Can you find  $P(70 < X < 130)$  exactly? If not what can you say?

$$(a) \quad \text{Var}(X) = E(X^2) - (EX)^2$$

$$\begin{aligned} \rightarrow E(X^2) &= \text{Var}(X) + (EX)^2 \\ &= 100 + (100)^2 \\ &= 10100 \end{aligned}$$

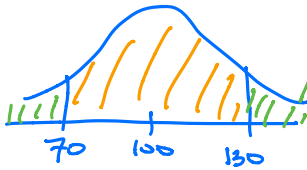
$$(b) \quad P(70 < X < 130) = 1 - P(X \notin (70, 130))$$

$$= 1 - P(|X - \underbrace{100}_{\mu}| \geq \underbrace{30}_c) \geq 1 - \frac{1}{9} = \frac{8}{9}$$

$$\textcircled{1} \quad P(|X - \mu| \geq c) \leq \frac{\text{Var}(X)}{c^2}$$

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

$$\leq \frac{\text{Var}(X)}{c^2} = \frac{100}{900} = \frac{1}{9}$$



## 7.1. Sums of Independent Random Variables

We know that  $E(X + Y) = E(X) + E(Y)$  but does  $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$ ?

*Not always true.*

Let  $X$  be the number of hours a student is awake a day and let  $Y$  be the number of hours a student is asleep a day. Then  $X + Y = 24$ , so trivially

$$\text{Var}(X + Y) = \text{Var}(24) = 0 \neq \text{Var}(X) + \text{Var}(Y).$$

So when  $X$  and  $Y$  are dependent, it is possible that  $\text{Var}(X + Y) \neq \text{Var}(X) + \text{Var}(Y)$ .

**Theorem** If  $X$  and  $Y$  are independent,

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y).$$

Ex: Let  $X_1, X_2, \dots, X_n$  be a i.i.d. random sample with mean  $\mu$  and SD  $\sigma$ . Let  $S_n = \sum_{i=1}^n X_i$ . Then

$$\text{Var}(S_n) = \text{Var}(X_1) + \text{Var}(X_2) + \dots + \text{Var}(X_n) = n\sigma^2.$$

So,  $\text{SD}(S_n) = \sigma\sqrt{n}$ . *: grows  $\sim \sqrt{n}$*

*E(S<sub>n</sub>) =  $\mu \cdot n$  : grows  $\sim n$*

*$n = 10,000$*

*$\text{SD}(S_n) \sim 100$*

*$E(S_n) \sim 10,000$*

**SD of Binomial** Let  $X \sim \text{Bernoulli}(p)$ . Then

$$\text{Var}(X) = E(X^2) - E(X)^2 = p - p^2 = p(1 - p).$$

$x$	1	0
$P(X=x)$	$p$	$1-p$

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

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

Now let  $X \sim \text{Binomial}(n, p)$ . Write  $X$  as a sum of  $n$  Bernoulli random variables and find  $\text{SD}(X)$ .

$X = \# \text{ successes out of } n \text{ trials}$ ,  $I_j = \begin{cases} 1 & \text{if } j\text{th trial is success} \\ 0 & \text{o.w.} \end{cases}$

$$= I_1 + \dots + I_n$$

$$I_j \sim \text{Bernoulli}(p)$$

$$\text{Var}(X) = \text{Var}(I_1) + \dots + \text{Var}(I_n)$$

$$= n \cdot p \cdot (1-p)$$

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

**SD of Poisson** Recall that  $\text{Binomial}(n, p)$  can be approximated by  $\text{Poisson}(np)$  for large  $n$  and small  $p$ .

$$\text{Binomial}(n, p) \xrightarrow[np \rightarrow \mu]{\substack{n \rightarrow \infty \\ p \rightarrow 0}} \text{Poisson}(\mu = np)$$

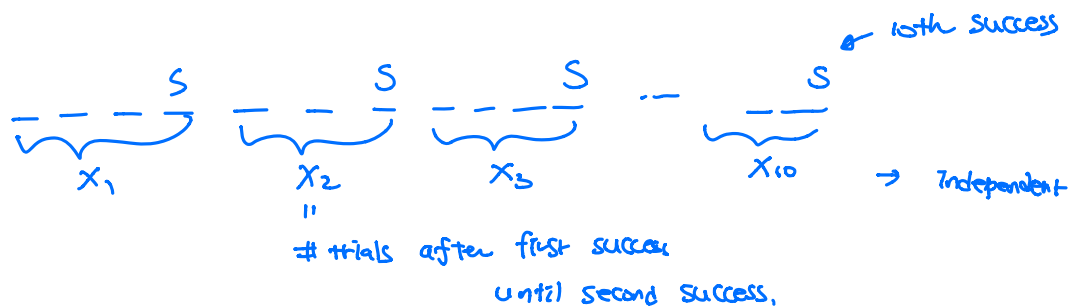
So we can find SD of  $\text{Poisson}(\mu)$  from limit of SD of  $\text{Binomial}(n, p)$  as  $n \rightarrow \infty, p \rightarrow 0, np \rightarrow \mu$ .

$$\begin{aligned} & \sqrt{np(1-p)} \xrightarrow[np \rightarrow \mu]{\substack{n \rightarrow \infty \\ p \rightarrow 0}} \sqrt{\mu} \\ \Rightarrow & \text{ If } X \sim \text{pois}(\mu), \\ & \text{SD}(X) = \sqrt{\mu} \end{aligned}$$

**SD of Geometric** Fact: If  $X \sim \text{Geom}(p)$ ,

$$\text{SD}(X) = \frac{\sqrt{1-p}}{p}.$$

Ex: (Waiting till the 10th success) Suppose you roll a die until the 10th success. Let  $R$  be the number of rolls required. Find  $\text{SD}(R)$ .



$$X_i \sim \text{Geom}(p)$$

$$R = X_1 + X_2 + \dots + X_{10}.$$

$$\text{Var}(R) = \text{Var}(X_1) + \dots + \text{Var}(X_{10})$$

$$= 10 \cdot \frac{1-p}{p^2}$$

$$\text{SD}(R) = \sqrt{\quad}$$