Stat 134 lec 13

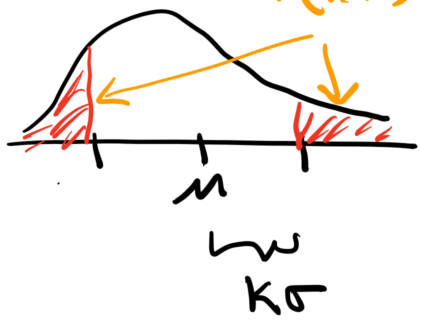
Quiz 3 - Wednesday sec 2.6, 3.1-3.3

Last time sec 3.3 $\sigma = \sqrt{E(X-\mu)^2}$

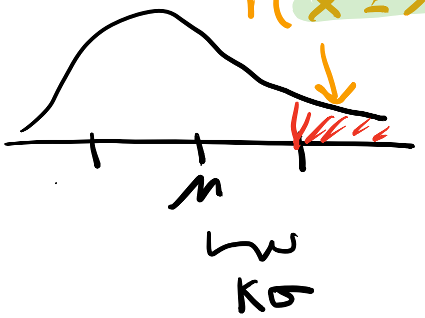
Tail bounds

Chebyshev's inequality

$P(|X-\mu| \ge k\sigma) \le \frac{1}{k^2}$

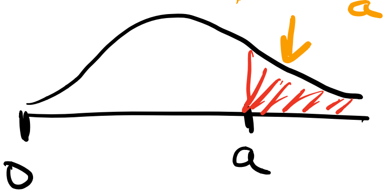


$P(X \ge \mu + k\sigma) \le \frac{1}{k^2}$

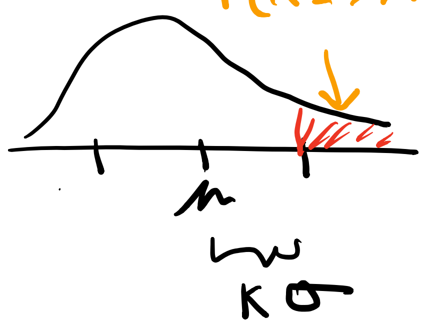


Markov's inequality

$P(X \ge a) \le \frac{\mu}{a}$



$P(X \ge \mu + k\sigma) \le \frac{\mu}{\mu + k\sigma}$



Today

Finish sec 3.3

another formula for $Var(X)$, $E(X^2)$

linear change of scale

$Var(X+Y)$ X, Y indep.

CLT

Sec 3.3 Another formula for Var(x).

Recall $E(cX) = cE(X)$

so $E(E(X)X) = E(X)E(X)$

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

$$\overset{\text{FOIL}}{=} E(X^2 - 2E(X)X + E(X)^2)$$

$E(cX)=cE(X)$

$$= E(X^2) \underbrace{- 2E(X)E(X) + E(X)^2}_{= -E(X)^2}$$

$$\Rightarrow \boxed{Var(X) = E(X^2) - E(X)^2}$$

(or) $\boxed{E(X^2) = Var(X) + E(X)^2}$

ex Let X be a nonnegative RV such that $E(X) = 100 = Var(X)$

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

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

$$= \boxed{100 + 10{,}000}$$

b) Can you find $P(70^2 < X^2 < 130^2)$ exactly? If not what can we say?

You could use Markov:

$$P(X^2 \geq 130^2) \leq \frac{E(X^2)}{130^2} = \frac{10100}{130^2} = .60$$

$$\Rightarrow P(70^2 < X^2 < 130^2) \leq P(X^2 < 130^2) \geq 1 - .60 = .40$$

or

You could use Chebyshev:

We dont know $SD(X^2)$ so we will need to apply Chebyshev to X

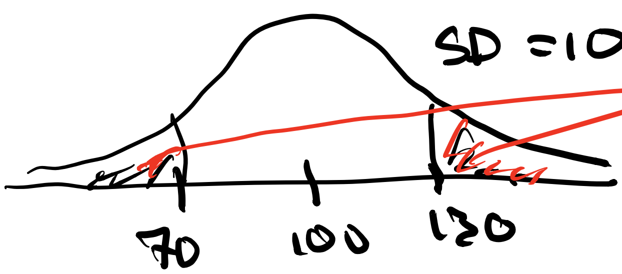
Note that:

$$70^2 < X^2 < 130^2 \iff 70 < |X| < 130$$

$|X| = X$ since $X \geq 0$

So,

$$P(70^2 < X^2 < 130^2) = P(70 < X < 130)$$



$K = 3$

$P(|X - 100| \geq 30) \leq \frac{1}{9}$

$$\Rightarrow P(70 < X < 130) \geq 1 - \frac{1}{9} = .89$$

Clearly Chebyshev is a better lower bound.

Stat 134

Chapter 3 Friday September 21 2018

1. X is a non negative random variable with $E(X) = 3$ and $SD(X) = 2$. True or False:

$$P(X^2 \geq 40) \leq \frac{1}{3}$$

a True

b False

Markov

$X^2 \geq 0$, $E(X^2) = Var(X) + (E(X))^2$
$= 4 + 9 = \boxed{13}$

$E(X^2 \geq 40) \leq \frac{13}{40} < \frac{13}{39} = \frac{1}{3}$ $\boxed{\text{TRUE}}$

Chebyshev

Don't know $SD(X^2)$

$P(X^2 \geq 40) = P(X \geq \sqrt{40})$

$E(X) + KSD = 3 + K \cdot 2$

$\sqrt{40} = 3 + K \cdot 2 \Rightarrow K = \frac{\sqrt{40} - 3}{2} = \boxed{1.66}$

$\Rightarrow P(X \geq \sqrt{40}) \leq \frac{1}{(1.66)^2} = \frac{1}{2.75} > \frac{1}{3}$

so Chebyshev says $\boxed{\text{Maybe}}$

We see Markov provides a better lower bound.

Change of Scale

Let $Y = -3X$

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

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

$$= E(9X^2) - (E(-3X))^2 \quad \leftarrow (-3E(X))^2$$

$$= 9E(X^2) - 9(E(X))^2$$

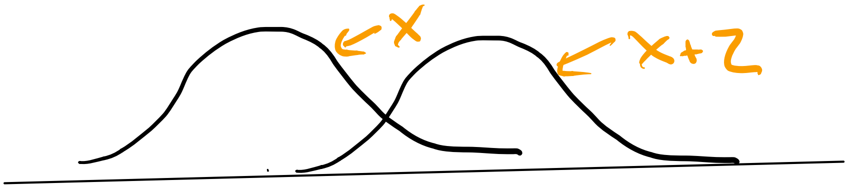
$$= (-3)^2 Var(X)$$

$$SD(Y) = \sqrt{Var(Y)} = \sqrt{(-3)^2 Var(X)} = |-3|\, SD(X)$$

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

Let $Y = X + 2$

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



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

$Var(X+Y)$

ex $\left.\begin{array}{l} X \sim Bin(n, p) \\ Y \sim Bin(m, p) \end{array}\right\}$ independent $\quad \begin{array}{l} Var(X) = npq \\ Var(Y) = mpq \end{array}$

$X+Y \sim Bin(n+m, p) \qquad Var(X+Y) = (n+m)pq$

It looks like: $\qquad = Var(X) + Var(Y)$

$$\boxed{\begin{array}{l} Var(X+Y) = Var(X) + Var(Y) \\ \qquad\qquad\qquad \text{if } X, Y \text{ are indep} \end{array}}$$

ex X = # hrs a student is awake in a day
Y = # hrs a student is asleep in a day

$X + Y = 24$

$Var(X+Y) = 0 \neq Var(X) + Var(Y)$

So the formula $Var(X+Y) = Var(X) + Var(Y)$ isnt true if X, Y are dependent.

ex

Let X, Y be independent RVs

$E(X) = 12 \qquad E(Y) = 2$

$SD(X) = 4 \qquad SD(Y) = 3$

let $W = 3X - 7Y - 4$

Find $E(W)$, $SD(W)$.

Soln

$$E(W) = E(3X - 7Y - 4) = 3\overset{12}{\overset{\shortparallel}{E(X)}} - 7\overset{2}{\overset{\shortparallel}{E(Y)}} - 4 = 18$$

$$Var(W) = 9\overset{16}{\overset{\shortparallel}{Var(X)}} + 49\overset{9}{\overset{\shortparallel}{Var(Y)}} = 585$$

$SD(W) = \sqrt{585}$.

Thm (Addition rule for variances)

If X, Y are independent random variables, $Var(X+Y) = Var(X) + Var(Y)$.

We will follow the proof in the book. First notice that $E(X - E(X))$ is 0.

so $E(X - E(X)) = 0$

translate of X by $-E(X)$.

Here is the proof (please read carefully)!

p193 If X, Y independent then $Var(X+Y) = Var(X) + Var(Y)$

Proof of the addition rule for variances. Let $S = X + Y$. Then $E(S) = E(X) + E(Y)$, so

$$S - E(S) = [X - E(X)] + [Y - E(Y)]$$

Note ① $E(S-E(S)) = 0$ by above

② $Var(S-E(S)) = Var(S)$

③ $Var(S-E(S)) = E((S-E(S))^2) - \underbrace{[E(S-E(S))]^2}_{0}$

$= E[(S-E(S))^2]$

Now square both sides and then take expectations to get

$$[S - E(S)]^2 = [X - E(X)]^2 + [Y - E(Y)]^2 + 2[X - E(X)][Y - E(Y)]$$

$$Var(S) = Var(X) + Var(Y) + 2E\{[X - E(X)][Y - E(Y)]\}$$

If X and Y are independent, then so are $X - E(X)$ and $Y - E(Y)$. So by the rule for the expectation of a product of independent variables, the last term above is the product of $E[X - E(X)]$ and $E[Y - E(Y)]$. This is zero times zero which equals zero, giving the addition rule for two independent variables. Apply this addition rule for two variables repeatedly to get the result for n variables. □

$$E[(X-E(X))(Y-E(Y))] = E\underbrace{(X-E(X))}_{0}\,E\underbrace{(Y-E(Y))}_{0}$$

Recall $E(AB) = E(A)E(B)$ if A, B indep.

Central Limit theorem (CLT)

$X_1, \ldots, X_n$ indep. $E(X) = \mu$
$SD(X) = \sigma$

$E(X_1 + \cdots + X_n) = n\mu$

$Var(X_1 + \cdots + X_n) = n\sigma^2 \Rightarrow SD(X_1 + \cdots + X_n) = \sqrt{n}\,\sigma$

P196

The Normal Approximation (Central Limit Theorem)

Let $S_n = X_1 + \cdots + X_n$ be the sum of n independent random variables each with the same distribution over some finite set of values. For large n, the distribution of S_n is approximately normal, with mean $E(S_n) = n\mu$, and standard deviation $SD(S_n) = \sigma\sqrt{n}$, where $\mu = E(X_i)$ and $\sigma = SD(X_i)$. That is to say, for all $a \leq b$

$$P\left(a \leq \frac{S_n - n\mu}{\sigma\sqrt{n}} \leq b\right) \approx \Phi(b) - \Phi(a)$$

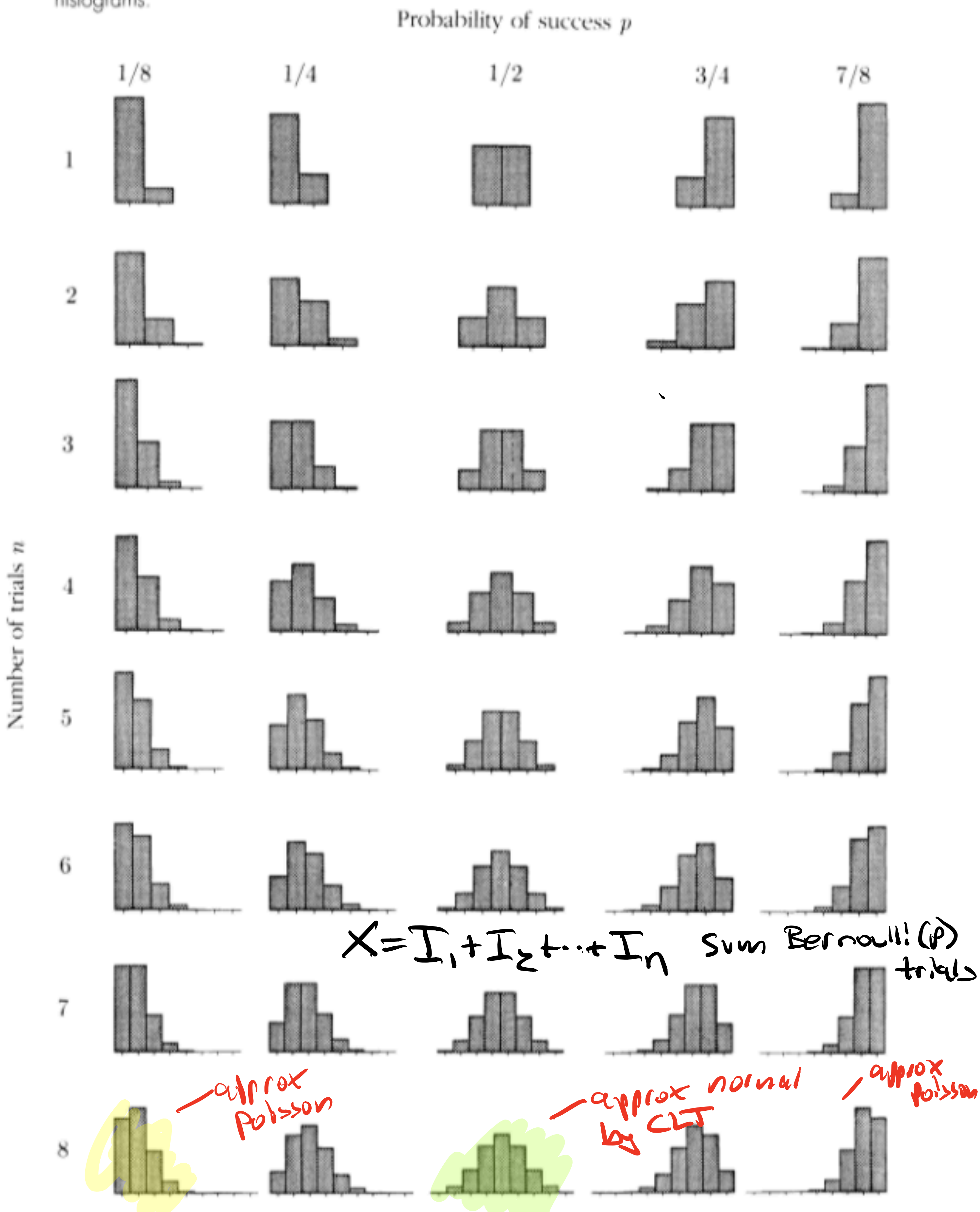
where Φ is the standard normal c.d.f. No matter what the distribution of the terms X_i, for every $a \leq b$ the error in using this normal approximation tends to zero as $n \to \infty$. The same result holds for X_i with an infinite range of possible values, provided the standard deviation is defined and finite.

The CLT says that if $X_1, \ldots, X_n$ are independent and identically distributed (mean μ, SD σ) then $S = X_1 + \cdots + X_n$ is approx normal for large n with mean $n\mu$ and SD $\sqrt{n}\,\sigma$.

ex $X_1, \ldots, X_n$ sum of n Bernoulli (p) trials:

P87

FIGURE 3. Histograms of some binomial distributions. The histogram in row n, column p shows the binomial (n, p) distribution for the number of successes in n independent trials, each with success probability p. In row n, the range of values shown is 0 to n. The horizontal scale changes from one row to the next, but equal probabilities are represented by equal areas, even in different histograms.



$X = I_1 + I_2 + \cdots + I_n$ sum Bernoulli (p) trials