AMAT 362 - Probability for Statistics

TODAY: LECTURE 16

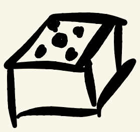
IMPORTANT DISTRIBUTIONS & THEIR PROPERTIES

BUILDING BLOCK

↳ **BERNOULLI RV** (= INDICATOR RV)

$$X = \begin{cases} 1 & \text{if event occurred w/ prob } p \\ 0 & \text{if didn't occurr w/ prob } q = 1-p \end{cases}$$

Ex: Flip a coin $X = \begin{cases} 1 & \text{if heads } p = 1/2 \\ 0 & \text{if tails } q = 1/2 \end{cases}$



Roll a die $X = \begin{cases} 1 & \text{if face 5 is up } p = 1/6 \\ 0 & \text{o.w.} \end{cases}$

Event = 5

$\boxed{\text{MEAN} = p}$

$\boxed{\text{S.DEV} = \sqrt{pq}}$

$\text{S.DEV} = \sqrt{\text{Var}}$

BINOMIAL RV

$S_n = X_1 + \dots + X_n$ = Sum of n IID Bernoullis

= counts # times the event of interest occurs

$\boxed{P(S_n = k) = \binom{n}{k} p^k q^{n-k}}$ PMF

event of interest = SUCCESS

MEAN of BINOMIAL

$$E(S_n) = E(X_1 + \cdots + X_n)$$

Linearity of Expectation $= E(X_1) + \cdots + E(X_n)$

$$= \boxed{np}$$

S.DEV of BINOMIAL

RECALL

$$VAR(X+Y) = VAR(X) + VAR(Y)$$

IF X & Y INDEP

$$VAR(S_n) = npq$$

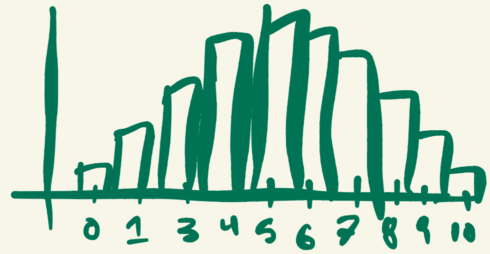
$$\Rightarrow \boxed{SD(S_n) = \sqrt{npq}}$$

Ex Flip a coin 10 times

$$P(3 \text{ heads}) = \binom{10}{3}\left(\tfrac{1}{2}\right)^3\left(\tfrac{1}{2}\right)^7 = \frac{1}{2^{10}}\binom{10}{3}$$

$$E(S_{10}) = 10 \cdot \tfrac{1}{2} = 5$$

$$SD(S_{10}) = \sqrt{10 \cdot \tfrac{1}{2} \cdot \tfrac{1}{2}} = \sqrt{\tfrac{10}{4}}$$



NORMAL APPROX TO BINOMIAL

(BELL CURVE)

As long as $\sqrt{npq} \geq 3$ then

$$P(a < S_n < b) \approx \int_{a^*}^{b^*} \frac{e^{-z^2/2}}{\sqrt{2\pi}}\,dz = \Phi\Big(\underbrace{\frac{b-np}{\sqrt{npq}}}_{b^*}\Big) - \Phi\Big(\underbrace{\frac{a-np}{\sqrt{npq}}}_{a^*}\Big)$$

$$\Phi(b^*) - \Phi(a^*)$$

BONUS

$$\boxed{P(a \leq S_n \leq b) \approx \Phi\left(\frac{b + \frac{1}{2} - np}{\sqrt{npq}}\right) - \Phi\left(\frac{a - \frac{1}{2} - np}{\sqrt{npq}}\right)}$$

CONTINUITY CORRECTION

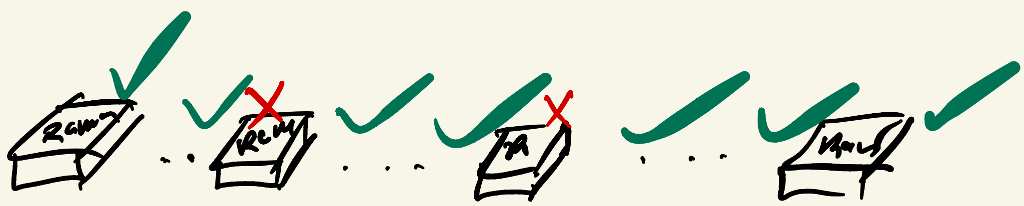
CDF of φ PDF unit normal

Ex Ramen Noodle Co

produces defective noodle packets @ 5% rate. I open a crate containing 1000 ramen noodle packets and find 60 defective ones.

Q: What's the probability of finding at least 60 defectives?

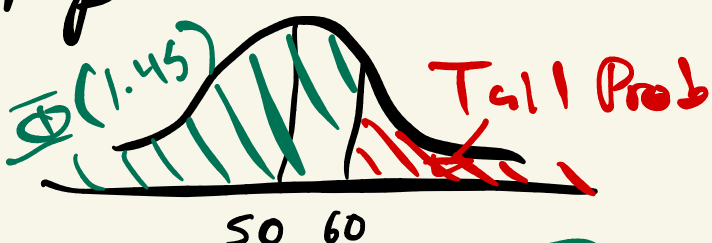
A:



$$X_1 + X_{20} + \dots + X_{600} + \dots + X_{1000}$$

$$X_i = \text{indicator RV w/} \begin{cases} 1 & (\text{is defective}) \; p = \frac{1}{20} \\ 0 & (\text{is fine}) \; q = \frac{19}{20} \end{cases}$$

$$E(S_n) = 1000 \cdot \frac{1}{20} = n \cdot p$$
$$= 50$$



$$P(S_n \geq 60) = 1 - \Phi\left(\frac{60 - np}{\sqrt{npq}}\right) = 1 - \Phi\left(\frac{10}{\sqrt{50 \cdot \frac{19}{20}}}\right)$$

1.45

Z-Table for $\Phi(1.45)$

$$= 1 - .9265$$
$$= 7.35\%$$

UNIFORM RV $\{a, a+1, \ldots, b\}$

$X \sim \text{Unif}\{a, \ldots, b\}$ $P(X=k) = \begin{cases} \frac{1}{b-a+1} & \text{for } a \le k \le b \\ 0 & \text{o.w.} \end{cases}$

integer

MEAN

$$E(X) = \frac{b+a}{2}$$

SD

$$SD(X) = \sqrt{\frac{(b-a+1)^2 - 1}{12}}$$

Ex: Rolling a die, let X = face value

6-sided die $P(X=k) = \frac{1}{6}$ for $1 \le k \le 6$

$= 0$ o.w.

$$E(X) = \frac{6+1}{2} = \frac{7}{2} = 3.5$$

$$SD(X) = \sqrt{\frac{35}{12}}$$

Def. not mode = most probable

Empirical Mean $S_n = \frac{X_1 + \cdots + X_n}{n} = \bar{X}$

Die rolls

$S_n \to E(X_i) = 7/2$

Law of Averages or Large Numbers

NOT AS ACCURATE AS

TRUE CLT

If $X_1, \ldots, X_n$ are IID RVs w/ $E(X_i) = \mu$ $SD(X_i) = \sigma$

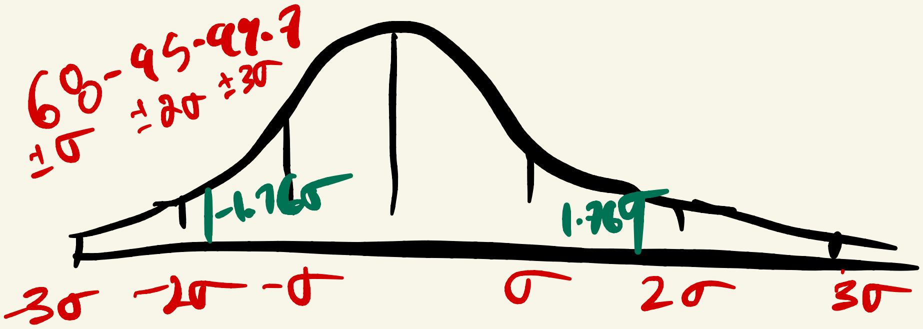
$$P(a < S_n < b) \underset{\text{large } n}{\approx} \Phi\left(\frac{b - n\mu}{\sqrt{n\sigma^2}}\right) - \Phi\left(\frac{a - n\mu}{\sqrt{n\sigma^2}}\right)$$

Practice Q

Roll a die 100 times, S_{100} = sum face values

Estimate $P(320 < S_{100} < 380)$

$$\approx \Phi(-1.76, 1.76) = 2\underbrace{\Phi(1.76)}_{96\%} - 1$$



$192 - 100 \approx 92\%$

SMALL SUMS, i.e. $Z = X + Y$

"CONVOLUTION" of PMFs/PDFs

If JDF or JMF of X, Y is $P(x, y)$ = prob of x AND y

then

$$P(Z = k) = \sum_{\substack{i,j \\ \text{s.t. } i+j=k}} P(X = i, Y = j)$$

N.B. Sum of Uniforms $\neq$ Uniform

Ex Roll a 6-sided die 2x, $S_2 = X_1 + X_2$

$$P(S_2 = 5) = P(X_1 = 4, X_2 = 1) + P(X_1 = 3, X_2 = 2) + P(X_1 = 2, X_2 = 3) + P(X_1 = 1, X_2 = 4)$$

$= 4/36$

SUM OF NORMALS IS NORMAL

$X_1 \sim Norm(\mu_1, \sigma_1^2)$
$X_2 \sim Norm(\mu_2, \sigma_2^2)$] $\xrightarrow{\text{If indep then}}$ $X_1 + X_2 \sim$ Normal w/

$E(X_1 + X_2) = \mu_1 + \mu_2$

$Var(X_1 + X_2) = \sigma_1^2 + \sigma_2^2$

POISSON RV

= Another Approximation of $Bin(n, p)$ where p is small and $np \to \lambda$

Say X is Poisson if
$$P(X=k) = \begin{cases} \frac{\lambda^k}{k!} e^{-\lambda} & \text{for } k \geq 0 \\ 0 & \text{o.w.} \end{cases}$$

MEAN

$E(\text{Poisson}(\lambda)) = \lambda$

$E(\text{Bin}) = np$

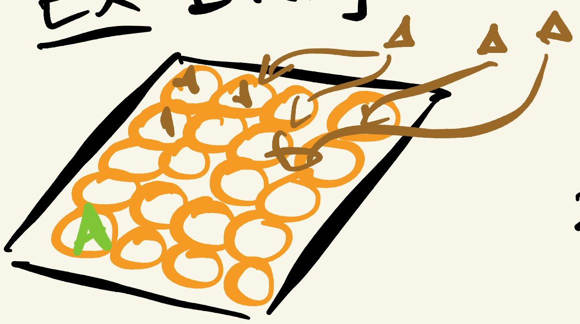
SD

$SD(\text{Poisson}) = \sqrt{\lambda}$

$SD(\text{Binom}) = \sqrt{npq}$

$np \to \lambda$
$q \to 1$

Ex Baking 50 cookies w/ 400 choc. chips



1) Prob A gets no choc. chips $\sim e^{-8}$
$\lambda = np = 8$
$\approx .0003$

2) Prob A gets more than 15 choc. chips?

3 Possible Answers

1) Exact Answer $\sum_{k=16}^{\infty = 400} P(X=k)$

$$= \sum_{k=16}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda}$$

$\lambda = 8$

2) Markov's Inequality

$$\boxed{\text{If } X \geq 0 \quad P(X \geq a) \leq \frac{E(X)}{a} \quad \text{MARKOV'S INEQ}}$$

$P(X \geq 16) \leq \frac{8}{16} = \frac{1}{2}$

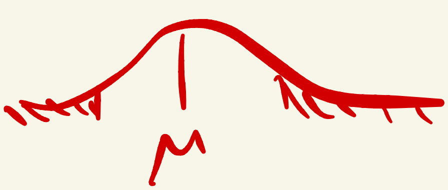
$= 50\%$

2) Chebyshev's Inequality

$$\boxed{\text{For any RV } X \quad P(|X-\mu| \geq c) \leq \frac{\sigma^2}{c^2} \quad \text{CHEBYSHEV'S INEQ}}$$

$P(|X-\mu| \geq c)$
$=$
$P(X \leq \mu - c)$
$+$
$P(X \geq \mu + c)$

TWO TAIL PROBABILITY



$\mu = 8 \quad c = 8$

$$\underbrace{P(X \leq 0)}_{e^{-8}} + \underbrace{P(X \geq 16)}_{\text{what we want}} \leq \frac{8}{8^2} = \frac{1}{8}$$

$= 12.5\%$

Since Poisson is a limiting distribution the sum of Poissons is Poisson

$X_1 \sim \text{Poisson}(\lambda_1)$ & $X_2 \sim \text{Poisson}(\lambda_2) \xrightarrow[\text{indep}]{} X_1 + X_2 \sim \text{Poisson}(\lambda_1 + \lambda_2)$

GEOMETRIC & NEGATIVE BINOMIAL

Geometric R.V.

T = Try # of first success

$$P(T=k) = \begin{cases} q^{k-1}p & \text{for } k \geq 1 \\ 0 & \text{o.w.} \end{cases}$$

MEAN

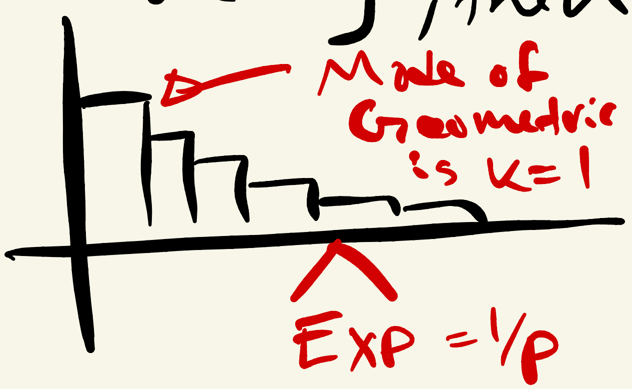
$$E(T) = \frac{1}{p}$$

SDEV

$$SD(T) = \frac{\sqrt{q}}{p}$$

$$= \sqrt{Var} \quad \text{so} \quad VAR = \frac{q}{p^2}$$

Ex I can land one out of 5 kickflips on avg, then



$$E(T) = \frac{1}{1/5} = 5$$

$$SD(T) = \sqrt{25 \cdot \frac{4}{5}} = \sqrt{20}$$

NEG BINOMIAL

T_r = # tries to the r^{th} success.

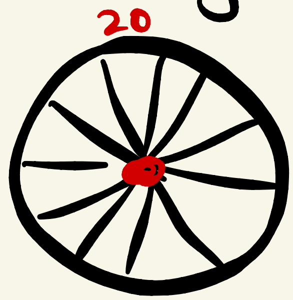
$$P(T_r = k) = \underbrace{\binom{k-1}{k-1} p^{k-1} q^{\overset{k-r}{\cancel{r-1-(k-1)}}}}_{\text{Prob of } r-1 \text{ successes in } k-1 \text{ tries}} \times \underbrace{p}_{\text{Prob } k^{th} \text{ try is a success}}$$

$$E(T_r) = \frac{r}{p}$$

$$SD(T_r) = \frac{\sqrt{rq}}{p}$$

Ex Cricket

= game in darts



Sps I can hit 20 w/ $p = 1/4$

What's the probability that my 3rd 20 occurs on $k=6$

$$\left[\binom{5}{2}\left(\frac{1}{4}\right)^2\left(\frac{3}{4}\right)^3\right] \times \frac{1}{4}$$

Prob of 2 successes in 3 throws

Prob 6th throw is success