ALL THE CLT'S

SAMPLE TOTAL	SAMPLE AVERAGE	SAMPLE TOTAL SUCCESS	SAMPLE PERCENTAGE
$E(T_0) = n\mu$	$E(\bar{X}) = \mu$	$E(X) = np$	$E(\hat{p}) = p$
$\sigma_{T_0} = \sigma\sqrt{n}$	$\sigma_{\bar{X}} = \sigma/\sqrt{n}$	$\sigma_X = \sqrt{np(1-p)}$	$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$

SAMPLE TOTAL and SAMPLE AVERAGE: FOLLOWS NORMAL DISTRIBUTION IF N IS LARGE

SAMPLE TOTAL SUCCESS: ↓ ACTUALLY BINOMIAL (DON'T USE NORMAL APPOX. WHEN COMPUTING PROBABILITY)

EXAMPLE 1: FINDING TOTAL APPROXIMATION

THIRTY-SIX JETS WAIT TO TAKE OFF FROM AN AIRPORT. THE AVERAGE TAXI + TAKE-OFF TIME FOR EACH JET IS 8.5 MINUTES, WITH AN SD OF 2.5 MINUTES. WHAT IS THE PROBABILITY THAT THE TOTAL TAXI + TAKE OFF TIME FOR THE 36 JETS IS LESS THAN 320 MIN?

WHAT IS THE VARIABLE? TIME

WHAT DISTRIBUTION DOES IT FOLLOW? WE DON'T KNOW, BUT WE CAN FIND TOTAL

IT IS CONTINUOUS

TOTAL = SAMPLE SUM

V_{ar} = time (taxi + take off)

$T_0 = X_1 + X_2 + \dots X_{36}$ (X_i = time per jet)

$\mu = 8.5$ $\sigma = 2.5$ $n = 36$

$E(T_0) = \mu_{T_0} = 8.5 * 36$

$= 306$

$SD(T_0) = \sigma_{T_0} = \sqrt{n}\,\sigma = \sqrt{36}\cdot 2.5$

$= 15$

WE KNOW TOTAL HAS NORMAL DISTRIBUTION

$P(T_0 < 320) = \text{normalcdf}(-1000, 320, 306, 15) \approx 0.8247$

EXAMPLE 2:

THE COOKIE MACHINE AT CHIPS AHOY ADDS A RANDOM # OF CHIPS TO EACH COOKIE. THE NUMBER OF CHIPS IS A RANDOM # WITH AVERAGE 28.5 AND SD 5.3. FIND THE PROBABILITY THAT, IN A BAG OF 50 COOKIES, THE AVERAGE # OF CHIPS PER COOKIE IS AT LEAST 30.

WHAT IS THE VARIABLE? # OF CHIPS

WHAT DISTRIBUTION DOES IT FOLLOW? WE DON'T KNOW

IT IS DISCRETE THO.

$\mu = 28.5$

$\sigma = 5.3$

$n = 50$ ← each cookie is a sample

$E(\bar{X}) = \mu = 28.5$

$SD = \sigma/\sqrt{n} = 5.3/\sqrt{50}$

$\bar{X} \sim \text{NORMAL}(28.5, 5.3/\sqrt{50})$

$P(\bar{X} \geq 30) = \text{NORMALCDF}(30, 1000, 28.5, 5.3/\sqrt{50})$ ← PUT IN EXACT, NO ROUNDING ERROR

$= 0.0227$

WK 8-1

$np \geq 10$

$.4(500) \geq 10$ = YES CAN USE CLT!

EXAMPLE 3: BINOMIAL SUCCESS

↗ TALKING ABOUT POPULATION AS A WHOLE: PERCENTAGE / ONE PERSON: PROBABILITY

IT IS KNOWN THAT 40% OF PEOPLE IN THE CITY ARE INFECTED BY COVID-19 (SOME NO, SOME MILD, SOME SEVERE SYMPTOMS).

A RANDOM SAMPLE OF 50 PEOPLE ARE SELECTED.

SUCCESS = INFECTED (WHAT WE ARE INTERESTED IN)

PROF ADDED INCLUSIVE

SHOULD BE FOLLOWED BY INCLUSIVE / NOT INCLUSIVE

a) WHAT IS THE PROBABILITY THAT BETWEEN 200 + 210 ARE INFECTED? (USE BINOMIAL THEN NORMAL APPROX. TO SEE HOW ACCURATE)

LOOKING FOR SAMPLE TOTAL SUCCESSES

WE CAN USE BINOMIAL (ACCURATE) + NORMAL (APPROXIMATE) →

① BINOMIAL:

$X \sim X_1 + X_2 + \dots X_{500}$

$X_i = \begin{cases} 1 & \text{success: infected by covid} \\ 0 & \text{failure: not infected} \end{cases}$

$X \sim \text{Binomial}(n = 500, p = 0.4)$

bc DISCRETE

$P(200 \leq X \leq 210) = P(X \leq 210) - P(X \leq 199)$

$= \text{BINOMIAL}(500, 0.4, 210) - \text{BINOMIAL}(500, 0.4, 199)$

≈ 0.3481

WHEN DOING NORMAL + CLT WAY, IT WOULD BE CONTINUOUS DIST. + $\leq$ + $<$ DON'T MATTER

② NORMAL APPROXIMATION

USING CLT → $X \sim \text{normal}(np, \sqrt{np(1-p)})$

$np = .4(500) = 200$ $\sqrt{np(1-p)} = \sqrt{500(.4\cdot.6)} = \sqrt{120}$

WHEN WE DO THIS X BECOMES CONTINUOUS!

↳ CLT $\leq$ / $<$ don't matter

normalcdf(200, 210, $\overbrace{200}^{np}$, $\overbrace{\sqrt{120}}^{\sqrt{np(1-p)}}$)

≈ 0.3193

b) WHAT IS THE PROBABILITY THAT THE SAMPLE PROPORTION (= PERCENTAGE) IS BETWEEN 39% AND 43%?

$\hat{p} = \frac{X_1 + X_2 + \dots X_{500}}{500} \overset{\text{approx}}{\sim} \text{normal}\left(\underset{0.4}{p}, \sqrt{\frac{p(1-p)}{n}}\right)$ ($\frac{0.4\cdot 0.6}{500}$)

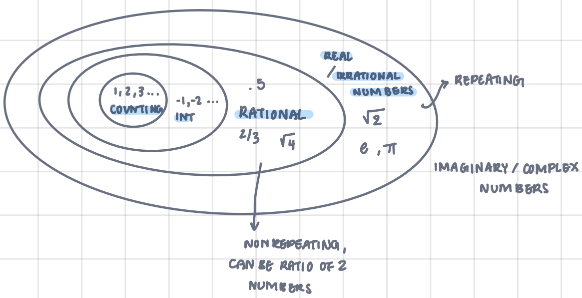
→ $P(0.39 \leq \hat{p} \leq 0.43) = \text{normalcdf}(.39, 0.43, 0.4, \sqrt{.24/500}) \approx 0.5905$

QUIZ:

- ROUNDING → under 5 = ignore, over 5 = up ← ALWAYS DO REGULAR ROUNDING UNLESS SPECIFIED
  - ↳ ROUND UP → always up } only specified
    (TRUNCATE 4 ⊕ #)
  - ↳ ROUND DOWN → always down }
  - ↳ TRUNCATE = JUST THROW AWAY THE DECIMALS

  ~~$1/3 = 0.3$~~ WRONG

- IRRATIONAL NUMBERS



4.6 ON HW:

T_1, T_2, T_3 PARALLE

Total Time → E(T)

T = time to get job time = $T_1 + T_2 + T_3$

sol. 1: $ET = ET_1 + ET_2 + ET_3 = 15$ ✗ ← parallel, don't add!

sol. 2: $T = T_i$, $ET = ET_i = 5$ ✗ THEY'RE RANDOM

sol. 3: $T = \max\{T_1, T_2, T_3\}$

E(T) = ? (A LOT OF MATH I DON'T WANT TO DO RIGHT NOW) OK FINE I DO IT

E(T) = 5, but through a lot of $\int$ + $\frac{d}{dx}$

THEY'RE PARALLEL, SO NOT GAMMA

$X \sim \lambda e^{-\lambda x}, x > 0$

$F(x) = \int_0^x \lambda e^{-\lambda t}\,dt = 1 - e^{-\lambda x}$

we want pdf of T, then $\int$

$T = \max\{T_1, T_2, T_3\}$

CDF: $F_T(x) = P(T < x) = P(T_1 < x)\,P(T_2 < x)\,P(T_3 < x)$

$= [F(x)]^3 = [1 - e^{-\lambda x}]^3$

PDF: $f_T(x) = \frac{d}{dx}[1 - e^{-\lambda x}]^3 = 3\lambda\,[e^{-\lambda x} - 2e^{-2\lambda x} + e^{-3\lambda x}]$

$E(T) = \int_0^\infty x \cdot 3\lambda\,[e^{-\lambda x} - 2e^{-2\lambda x} + e^{-3\lambda x}]\,dx$

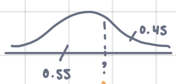
$= 1/\lambda = 5$

EXAMPLE FOR HW 6 #2 (4.18 d)

$P(|X+3| \geq 2.39) = .45$ μ: $E(X) = -3$ $\text{Var}(X) = 4$ σ: $SD = 2$

THE IDEA:

$P(X > ?) = 0.45$



#1 THEY FIND THIS z^*

INVNORM C(0.55, -3, 2) (μ, SD)

PROF EXPLAINS TEXTBOOK WAY:

#2 THEY USE $x = \mu + z^*\sigma$ FORMULA

BUT WE DON'T USE THIS.

- conceptual + calculation!

WK 8-2

Ch. 6 STOCHASTIC PROCESSES

6.1 + 6.2

" we skip chapter 5 "

→ this chapter is still computing probability, just a different process

RANDOM VARIABLE AS A FUNCTION OF TIME!

STOCHASTIC PROCESS

IN THE PAST: X = R.V., and it had a possible value.

Now $X(t)$, X is a diff. random variable at different times.

↳ TIME CHANGES

IS A RANDOM VARIABLE THAT ALSO DEPENDS ON TIME. IT IS WRITTEN AS $X(t)$. $t \in T$ WITH T BEING A SET OF POSSIBLE TIMES, USUALLY $[0, \infty)$, $(-\infty, \infty)$, $\{0, 1, 2, 3\}$, or $\{\ldots, -2, -1, 0, 1, 2, \ldots\}$

↳ **STATES** POSSIBLE VALUES OF $X(t)$ →

$t = 0$: X 0 1 2, P(X) . . .

$t = 1$: X 0 1 2, P(X) . . .

$t = 2$: X 0 1 2, P(X) . . .

X'S DISTRIBUTION CHANGES AT DIFFERENT TIMES

↳ **REALIZATION** / AKA SAMPLE PATH / TRAJECTORY OF A PROCESS OF $X(t)$ = RECORDING OF WHAT HAPPENED IN THE PAST

↳ $X(t)$ is a **DISCRETE-STATE** PROCESS IF: X is discrete for each time t

a **CONTINUOUS STATE** PROCESS IF: X is continuous

↳ $X(t)$ is a **DISCRETE-TIME** PROCESS IF THE SET OF TIME T IS DISCRETE

is a **CONTINUOUS-TIME** PROCESS IF T IS AN TIME INTERVAL

WE DON'T DO THE GENERAL STOCHASTIC PROCESS, JUST THE SPECIAL CASE **MARKOV** PROCESS/CHAINS

↳ STOCHASTIC PROCESS IS MARKOV IF P(FUTURE | PRESENT + PAST) = P(FUTURE | PRESENT)

→ PAST DOESN'T AFFECT FUTURE

MARKOV CHAIN

IS A DISCRETE-TIME, DISCRETE-STATE MARKOV STOCHASTIC PROCESS

$T = \{0, 1, 2\}$, $X(t)$ can be written as a random sequence $\{X(0), X(1), X(2) \ldots\}$

sometimes we use $1, 2, \ldots, n$ to represent the states, n could be ∞

TRANSITION PROBABILITY

$P_{ij}(t) = P(X(t+1) = j \mid X(t) = i)$ PROBABILITY MAKING A TRANSITION FROM STATE i TO STATE j AT TIME t

STATE: SUNNY 1, RAINY 2, CLOUDY 3 — POSSIBLE VALUES OF RV

TIME:	M	T	W	TH	F	S	SU	M
	0	1	2	3	4	5	6	7

$P_{23}(0) = P(X(1) = 3 \mid X(0) = 2)$ GIVEN 2 (RAIN) ON 0 (MON.), THE PROB OF 3 (CLOUDY) ON 1 (TUE)

monday

1 STEP PROB.

WE HAVE 3 STATES, WE CAN HAVE 9 COMBOS.

$P_{31}(1) = P(X(2) = 1 \mid X(1) = 3)$

DON'T NEED t, t does not affect anything

HOMOGENEOUS

IF ALL TRANSITION PROBABILITIES ARE INDEPENDENT OF t.

$P_{ij}(t) = P_{ij}$, $P_{ij}^{(h)}(t) = P_{ij}^{(h)}$

THE DISTRIBUTION OF HOMOGENEOUS MARKOV CHAIN IS COMPLETELY DETERMINED BY THE INITIAL DISTRIBUTION OF P_0 + ONE-STEP TRANSITION PROB. P_{ij}

$P_{12}^{(2)}(1) = P_{12}^{(2)}(2) = P_{12}^{(2)}(3) \ldots$

SO = $P_{12}^{(2)}$

H-STEP TRANSITION PROBABILITY

$P_{ij}^{(h)}(t) = P(X(t+h) = j \mid X(t) = i)$ PROBABILITY MAKING A TRANSITION FROM STATE i TO STATE j FROM TIME t to TIME t+h

MEANS TWO STEPS $P_{12}^{(2)}(1) = P(X(3) = 2 \mid X(1) = 1)$

+ ONE MORE DAY

WE WANT TO COMPUTE THE

① TRANSITION PROBABILITY MATRIX P (n by n matrix if n states w/ element p_{ij})

↳ we represent the states through the matrix!

$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix}$$

sometimes this is given, sometimes it needs to be found

② n-step transition probability matrix $P^{(n)}$ (n by n matrix w/ element $p_{ij}^{(n)}$)

③ P_n the distribution at time n (1 by n matrix)

④ limit of $p_{ij}^{(n)}$ + P_n as n goes to infinite which is our long term forecast.

→ we are still finding probability

EXAMPLE 1:

IN SOME TOWN, EACH DAY IS EITHER SUNNY OR RAINY. A SUNNY DAY IS FOLLOWED BY ANOTHER SUNNY DAY W/ PROBABILITY 0.7, WHEREAS A RAINING DAY IS FOLLOWED BY SUNNY DAY W/ A PROBABILITY OF 0.4

GIVEN IT RAINS ON MONDAY, MAKE FORECAST FOR TUESDAY, WEDNESDAY, N THURSDAY.

THE WEATHER IS HOMOGENEOUS MARKOV CHAIN W/ TWO STATES: 1 = SUNNY 2 = RAINY

$T = \{0, 1, 2, \dots\}$, $X(0)$ = weather on monday, $X(1)$ = weather on tuesday, . . .

NEW WAY TO COMPUTE: USE MATRIX

TRANSITION MATRIX P: $\begin{bmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & \ddots & \vdots \\ p_{n1} & \cdots & p_{nn} \end{bmatrix}$ $\quad p_{i1} + p_{i2} + \dots + p_{in} = 1$

N-STEP TRANSITION MATRIX: $P^{(n)} = \begin{bmatrix} p_{11}^{(n)} & \cdots & p_{1n}^{(n)} \\ \vdots & \ddots & \vdots \\ p_{n1}^{(n)} & \cdots & p_{nn}^{(n)} \end{bmatrix}$

$$P^{(n)} = P^n$$

n-step trans. prob. matrix — trans. matrix raised to power of n

$$P_n = P_0 P^{(n)}$$

PROB. DIST. OF $X(n)$ ← 1xn matrix — INITIAL PROB. DIST. 1 x n matrix

OLD WAY TO COMPUTE:

GIVEN:

→ $p_{11} = 0.7$ $\quad p_{12} = .3$ $\quad p_{22} = 0.6$

FORECAST 4 TUESDAY = $P\{X(1) = 1\} = p_{21} = 0.4$

$P\{X(1) = 2\} = p_{22} = 0.6$

FORECAST 4 WEDNESDAY: USE LAW OF TOTAL PROB.

$P\{X(2) = 1\} =$

USING TUE. TO FIND WED.

$P\{X(2) = 1 \mid X(1) = 1\} \, P\{X(1) = 1\} +$

$P\{X(2) = 1 \mid X(1) = 2\} \, P\{X(1) = 2\} =$

$\overset{.7}{p_{11}} \cdot 0.4 + \overset{.4}{p_{21}} \cdot 0.6 = 0.52$

$P\{X(2) = 2\} = 1 - 0.52 = 0.48$

FORECAST ON THURSDAY:

$P\{X(3) = 1\} = P\{X(3) = 1 \mid X(2) = 1\} \, P\{X(2) = 1\} +$ USE WED. TO CALCULATE THURSDAY

$P\{X(3) = 1 \mid X(2) = 2\} \, P\{X(2) = 2\}$

$= \overset{.7}{p_{11}} \cdot .52 + \overset{.4}{p_{21}} \cdot .48 = 0.5556$

$P\{X(3) = 2\} = 1 - 0.556 = 0.444$

USING THE TI-84 TO OPERATE W/ MATRICES!

- MATRIX : EDIT → OPERATE
- X^{-1} : FIND INVERSE MATRIX
- MATH > ENTER TO CONVERT DEC → FRACTION

→ SOLUTION W/ MATRIX

$$P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} = \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix}$$ = INITIAL DISTRIBUTION

$P_0 = [0 \quad 1]$ →

X	1	2
P(X)	0	1

TUESDAY:

$$P_1 = P_0 P^{(1)} = [0 \quad 1] \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix} = [0.4 \quad 0.6]$$

WEDNESDAY:

$$P_2 = P_0 P^{(2)} = [0 \quad 1] \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix}^2 = [0 \quad 1] \begin{bmatrix} 0.61 & 0.39 \\ 0.52 & 0.48 \end{bmatrix} = [0.52 \quad 0.48]$$

THURSDAY:

$$P_3 = P_0 P^{(3)} = [0 \quad 1] \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix}^3 = [0.556 \quad 0.444]$$

MINI LIN ALG REVIEW:

$$\begin{cases} 2x_1 - 3x_2 = 1 \\ x_1 + 2x_2 = -1 \end{cases} \rightarrow \begin{bmatrix} 2 & -3 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \Rightarrow AX = B$$

$$X = A^{-1}B = \begin{bmatrix} -1/7 \\ -3/7 \end{bmatrix}$$

$$\begin{cases} 2x_1 - 3x_2 = 1 \\ x_1 + 2x_2 = -1 \\ 3x_1 - x_2 = 0 \end{cases} \rightarrow \underbrace{\begin{bmatrix} 2 & -3 \\ 1 & 2 \\ 3 & -1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}_{\vec{x}} = \underbrace{\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}}_{\vec{b}}$$

$$A = \begin{bmatrix} 2 & -3 & 1 \\ 1 & 2 & -1 \\ 3 & -1 & 0 \end{bmatrix}$$

$x = \text{rref}(A) = \begin{bmatrix} -1/7 \\ -3/7 \end{bmatrix}$

BUTTON IN MATRIX > MATH