Chapter 9: Sampling Distributions
9.1 Sampling Distributions - Day 1
Activity 9A - Young Women's Heights

- The height of a young woman varies approx. according to $N(64.5, 2.5)$ Distributions.
- Simulate the heights of 100 Randomly selected women and store in L_1
  - Stat / CLR List L_1
  - $L_1 = \underbrace{\text{randNorm}}_{\text{Math Prb - 6}}(64.5, 2.5, 100)$
  - Histogram $X \underbrace{[57, 72]_{2.5}}_{\pm 3\sigma \text{ from } \mu}$ $Y[-10, 45]_5$ Symm or skewed?
  - Compare $\bar{x}$ w/ $\mu = 64.5$
- Repeat 3 more times + record $\bar{x}$ on 3 separate P.I.N.'s
- Scale @ Bottom of Board $[63, 66]_{.25}$ (Tick marks 1" wider than post-it-notes)
- Build Post it Note Histogram
- Questions:
  - What is the approx. shape of the distribution of $\bar{x}$.
  - What is the center " " " " "
- Have someone call out the values of $\bar{x}$ from the post it notes + Enter values in L_2.
- 1-var stats for L_2 + Look @ $\bar{x}$ and compare w/ μ.

9.1 Sampling Distributions - Day 2

<u>OBJ</u>: You will know the difference between a parameter & statistic & the properties of each

<u>Parameter</u> - A number that describes the <u>population</u>
(μ, σ, P)

<u>Statistic</u> - A number that describes the <u>Sample</u>.
($\bar{x}, S, \hat{p}$)

Example 9.1

Sampling Variability - Bottom of pg 488
- $\bar{x}$ Δ's from sample to sample - Think of the post-it from the P.I.N. Histogram!

Example 9.2

9.1, 9.2, 9.3, 9.4

Post it note Histogram

How do you overcome sampling variability? - Top pg. 490
- Take a large # of samples from the same pop.
- Calculate $\bar{x}$ or $\hat{p}$ for each sample.
- Construct a histogram of $\bar{x}$ or $\hat{p}$.
- Examine for shape, center, spread, gaps & outliers.

Figure 9.1 - pg 491

Sampling Distribution - Middle of pg 491 + Next ¶.

9.5

<u>HW</u>: Read Example 9.5
Do 9.6
Tomorrow pgs (497-499)

9.1 Sampling Distributions - Day 3

<u>OBJ</u>: You will know the difference between bias and variability.

Figure 9.8 and Above

<u>Bias</u> concerns the <u>center</u> of the sampling distribution.

<u>Unbiased Statistic</u> - If the <u>mean</u> of the sampling dist. is equal to the <u>true mean</u> of the parameter being estimated. (μ, or p)

Larger Samples give smaller spread (Fig 9.8)

Example 9.6 and ¶ above

Last 2 ¶ @ Bottom of pg 498.

Variability of a Statistic - Top of pg 499
- Concerns the <u>spread</u> of the Sampling Dist.

Figure 9.9
9.10

9.8a $\longrightarrow$

9	$\frac{9}{200}$ = .045	1
13	$\frac{13}{200}$ = .065	3
14	$\frac{14}{200}$ = .070	2
15	⋮ .075	5
16	.080	11
17	.085	12
18	.090	12
19	.095	9
20	.100	7
21	.105	5
22	.110	6
23	.115	7
24	.120	10
25	.125	4
26	.130	1
27	.135	2
28	.140	2
30	.150	1

<u>HW</u>: 9.8 b-e
9.11

9.2 Sample Proportions

OBJ: You will use the mean + St. Dev of a sample prop.

$\hat{p} = \frac{x}{n} = \frac{\text{\# of "successes" in a sample}}{\text{Sample Size}}$ — x and $\hat{p}$ will vary, $\therefore$ x and $\hat{p}$ are random variables.

Mean + St. Dev. of a Sampling Dist. of a Sample Prop.

$\mu_{\hat{p}} = p$
$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$

* on F.S. Proof on pg. 505..... Have Fun ☺
Don't Forget $\boxed{\text{Pop} \geq 10n}$ if you want to use $\sqrt{\frac{p(1-p)}{n}}$

You Know why $\mu_{\hat{p}} = p$. (Don't You? – P.I.N. Histogram!)
You also Know that as your sample size (n) inc, the st. dev. decreases (Don't You? – Fig. 9.8 from yesterday)
But why does it decrease? Good Question!
As n inc. σ dec because n appears in the denom.
also... to cut σ in half, the sample size must be four times as large (n is under the radical)

Rule of Thumb 1: pop $\geq 10n$? Yes, then $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$
No, then Caution! and continue.

Rule of Thumb 2: $np \geq 10$, $n(1-p) \geq 10$ — Yes, then you will use the normal approx.
No, then Caution! And do it anyway!

pg 506
507

Example 9.7 – Read 1st P then close book + do problem

$p = .35$
$n = 1500$
$\therefore \mu_{\hat{p}} = p = .35$

$P(.35 - .02 \leq \hat{p} \leq .35 + .02)$
$P(.33 \leq \hat{p} \leq .37)$

Pop $\geq 10n$
Pop $\geq 10(1500)$
Pop $\geq 15{,}000$ ✓

$np \geq 10$ — $(1500)(.35) \geq 10$ — $525 \geq 10$ ✓
$n(1-p) \geq 10$ — $(1500)(1-.35) \geq 10$ — $975 \geq 10$ ✓

$\therefore \sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$
$\sigma_{\hat{p}} = \sqrt{\frac{(.35)(1-.35)}{1500}}$

$\therefore P\left(\frac{\hat{p} - \mu_{\hat{p}}}{\sigma_{\hat{p}}} \leq z \leq \frac{\hat{p} - \mu_{\hat{p}}}{\sigma_{\hat{p}}}\right)$

$P\left(\frac{.33 - .35}{.0123} \leq z \leq \frac{.37 - .35}{.0123}\right)$

$P(-1.63 \leq z \leq 1.63)$

$\therefore$ 90% of all samples will give a result w/in 2% pts of the truth about the po

Also:
normalcdf(.33, .37, .3
.01

$= .8961$

9.3 Sample Means - Day 1

OBJ: You will use the mean & St Dev. of a Sample Mean.

Example 9.9

Averages are less variable than indiv. obs. (P.I.N. Hist!)

Averages are more NORMAL than indiv. obs.

Mean + St. Dev of a Sample Mean

Suppose $\bar{x}$ is the mean of an SRS of size n drawn from a Lg. Pop. w/mean μ and St. Dev. σ.

On* F.S.

The mean of the Samp. Dist. of $\bar{x}$ is $\mu_{\bar{x}} = \mu$.

Think: $\mu_{\hat{p}} = p$

The St. Dev is $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ } ÷ by $\sqrt{n}$ b/c as n inc., $\sigma_{\bar{x}}$ dec., ∴ less variable. (Pop ≥ 10n)

The Behavior of X... See Middle of pg 516.

* The orig. Dist. may be skewed, but if you keep inc. the Sample size (n → 5 → 8 → 10) Etc, the dist. will look more and more normal b/c you are only graphing the means!

Example 9.10

Sample Dist. of a Sample Mean from a Normal Pop.

Draw an SRS of size n from a pop. that has the NORM Dist. w/mean μ and St. dev σ. Then the sample mean $\bar{x}$ has the norm dist $N(\mu, \sigma/\sqrt{n})$ w/ Mean μ and St. Dev $\sigma/\sqrt{n}$.

Example 9.11 Do, But Read 1st ¶ and Finish on Board

Figure 9.16

9.31

HW: 9.32, 9.33
pgs (521-523)

9.3 Sample Means - Day 2

OBJ: You will apply the Central Limit Theorem.

Central Limit Theorem - pg 521 + TP After.

Simulation: Student scores on the A.C.T. Entrance Exam Are $N(18.6, 5.9)$

Don't Do.... Just Look @ Results

- Mode to Float
- Generate a List of 25 Sample Means of size n ($n = 5 \rightarrow 15 \rightarrow 30 \rightarrow 100 \rightarrow 500$)
- Seq (Mean (Randnorm (18.6, 5.9, n)), X, 1, 25, 1) → L1
  - Seq = 2nd / LIST / OPS
  - Mean = 2nd / LIST / MATH
  - RANDNORM = MATH / PROB
- Histogram (OPTIONAL) $[10, 30]_1$, $[-2, 15]_5$ FOR $n=5$
- 1-Var Stats - NOTICE Mean + St. Dev.

25 Samples of

Sample size n	$\bar{X}$	S_x
5	18.93	3.14
15	18.82	1.45
30	18.74	.86
100	18.64	.52
Take A NAP ← 500	18.61	.26

Example 9.12

9.35

HW: 9.37, 9.38, 9.39

Day 3: Class: 9.47, 9.48, 9.53

HW: 9.49, 9.50