Stat 50 - Elementary Statistics Winter 2020

Exam 3

Ch 6, 7, 8.1 **Dr. Jorge Basilio**

Jan_31 **gbasilio@pasadena.edu**

Honesty Pledge

On my honor, by printing and signing my name below, I vow to neither receive nor given any unauthorized assistance on this examination:

NAME (PRINT): Solutions SIGNATURE: ____________________

Directions

- YOU ARE ALLOWED TO USE A CALCULATOR ON THIS EXAM. (Ti83/Ti83+/Ti84/Ti84+/Ti84+CE-T, or scientific calculator)
- You have 80 minutes to complete this exam.
- The exam totals **105 points** but will be graded out of 100 only. (So it is possible to get 105% on this exam ☺).
- There are 9 problems, many of them with multiple parts.
- Place all of your belongings in the front of the classroom and I will assign you a seat. Bring with you your writing utensils.
- Cell phones must be turned off and put away in with your items in the front of the classroom.
- Handwriting should be neat and legible. If I cannot read your writing, zero points will be given.
- Make sure to ALWAYS SHOW YOUR WORK; you will not receive any partial credits unless work is clearly shown. *If in doubt, ask for clarification.*
- Leave answers in exact form (as simplified as possible), unless told otherwise.
- Put a $\boxed{\text{box around your final answer}}$ where applicable.
- **PLEASE INCLUDE UNITS** where applicable
- **PLEASE CHECK YOUR WORK!!!**
- If you need extra space, there is extra space on the back of the cover page and clearly indicate that you are continuing your work there in the original location.
- If you finish early, you may take a break but you must come back to class by 2:45 and we will have class.
- I will take attendance at the end of class
- Some questions contain multiple-parts which you must do individually and the parts are denoted by (a), (b), (c), etc. Some questions are multiple-choice and the choices are denoted with (A), (B), (C), (D), and (E).

Score	Grade

This page is intentionally blank. It may be used for scratch paper. If you wish for me to grade your work on this page, please (i) label the problem you are working on, (ii) box your answer, (iii) indicate in the original problem's location that you will continue your work on this page.

Problem 1: 10 pts (1 pts each)

TRUE or **FALSE** (please spell out/write the entire word for credit). *(No work needed)*

(a) TRUE For continuous random variables X, $P(a \le X \le b) = P(a < X < b)$.

(b) FALSE $\mu_{\bar{x}}$ and $\sigma_{\bar{x}}$ of a sampling distribution are ~~statistics~~ parameters. (see CLT)

(c) TRUE As n grows larger, the mean of the sampling distribution of $\bar{x}$ gets closer to μ.

(d) TRUE The standard deviation of the distribution of $\bar{x}$ (sampling) decreases as n increases. $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$ n inc, $\sigma_{\bar{X}}$ dec.

(e) TRUE $\bar{x}, \hat{p}, s$, and s^2 are point estimates.

(f) TRUE All else being equal, the higher the confidence level, the wider the confidence interval.

if we have 100% confidence (w/o knowing pop. para), we need large/wider interval

(g) TRUE A t-distribution is bell-shaped like the z- distribution but fatter tails.

(h) FALSE If the null hypothesis is not rejected, there is strong statistical evidence that the null hypothesis is ~~proved~~ true. basically correct, but in stats we never prove anything

(i) TRUE A type II error is made by failing to reject a false null hypothesis.

	H_0 T	H_0 F
FTR H_0		II
R H_0	I	

$P(\text{FTR } H_0 \mid H_0 \text{ F}) = \beta$

(j) TRUE The level of significance α of a test is the probability of making a type I error, given that the null hypothesis is true.

Problem 2: 12 pts (1 pt each blank)

Fill in the blanks:

(a) All other things being equal, choosing a **smaller** value of α (Prob making I Error) will increase the probability of making a type II error.

(b) For any random variable X, the shape of the sampling distribution of $\bar{x}$ will be approximately normal provided that the sample size is greater than 30.

(c) A statistic is a characteristic of a sample.

(d) The **standard normal distribution** has a mean of 0 and standard deviation of 1.

(e) List two **unbiased** estimators: $\bar{X}$ and $\hat{p}$. (also: s^2)

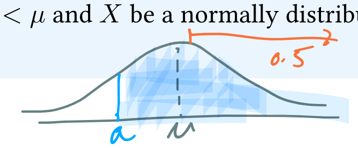
(f) List two **biased** estimators: Med and s. (also: Range)

(g) Which type of sampling must be used to select samples used for constructing sampling distributions and confidence intervals?

Simple Random Samples. (SRS)

(h) Let $a < \mu$ and X be a normally distributed random variable. Then $P(a < X) = 0.5 +$ $P(a < X < \mu)$.

(or $P(a \le X \le \mu)$
(or normalcdf(a, μ, μ, σ)))



Problem 3: 4 pts (1 pts each)

Multiple-choice. Select the correct answer:

(a) Which of the following is NOT true of both the **z-distribution** (normal) and the **t-distribution** (Student's)?

(A) The total area under their curve is 1.0. ✓

(B) The mean is zero. ✓

(C) The curve is bell shaped. ✓

D $\boxed{\text{(D) The standard deviation is 1.}}$ ✗

(E) The area under the curve is the probability. ✓

(b) A confidence interval is an interval that is used to estimate a:

(A) population parameter based on information from a population.

(B) sample statistic based on the information from a population.

C $\boxed{\text{(C) population parameter based on information from a sample.}}$

(D) sample statistic based on the information from a sample.

(E) sample parameter based on the information from a population statistic.

(c) A 99% confidence interval for μ can be interpreted to mean that if we take 100 samples of the same size and construct 100 confidence intervals for μ, then

(A) 99 of them will not include the true population μ. ✗

B $\boxed{\text{(B) 99 of them will include the true population } \mu.}$

(C) 99 of them will include the true sample $\bar{x}$. ✗

(D) 99 of them will not include the true sample $\bar{x}$. ✗

(E) 99% of the μ will be in 100 confidence intervals. ✗

(d) The width of the confidence interval depends on the size of the

(A) population proportion. ✗

(B) sample proportion. ✗

(C) population mean. ✗

(D) sample mean. ✗

E $\boxed{\text{(E) margin of error.}}$

$(\hat{p} - E, \hat{p} + E)$ or $(\bar{x} - E, \bar{x} + E)$

Problem 4: 6 pts

The owners of the Burger Emporium are looking for new supplier of tomatoes for their famous hamburgers. It is important that the tomato slice be roughly the same diameter as the hamburger patty. After careful analysis, they determine that they can only use tomatoes with diameters between 9 and 10 cm. $9 < x < 10$

Company A provides tomatoes with diameters that are approximately normally distributed with mean 10.5 cm and standard deviation of 1.1 cm. Company B provides tomatoes with diameters that are approximately normally distributed with mean 10.3 cm and standard deviation of 0.8 cm.

Which company provides the higher proportion of usable tomatoes? *Justify your choice with an appropriate statistical argument.* $(M \to E)$

Company A

- $\mu = 10.5$ cm
- $\sigma = 1.1$ cm
- x_A = diameter of tomato from Comp. A
- only tomatos $9 < x_A < 10$

$P(9 < x_A < 10) = \text{normalcdf}(9, 10, 10.5, 1.1)$

$\doteq 0.238$

Company B

- $\mu = 10.3$ cm, $\sigma = 0.8$ cm
- x_B = diameter of tomato from Comp. B
- only tomatos $9 < x_B < 10$.
- $P(9 < x_B < 10) = \text{normalcdf}(9, 10, 10.3, 0.8) = 0.302$

$(M \to E)$ "Company B provides a higher proportion of usable tomatoes since 30.2% will have a diameter b/w 9 & 10 cm."

Read instructions carefully!

Problem 5: 30 pts

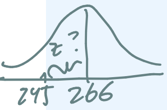
To earn full credit: sketch the appropriate distribution curve, indicate the probability by shading, label the points on the axes, and show what you enter into the calculator (as we do on our worksheets).

The lengths of human pregnancies are approximately normally distributed, with mean $\mu = 266$ days and standard deviation $\sigma = 16$ days.

Let RV X = length (in days) of human pregnancy.

(3 pt) (a) Becky's baby was born 245 days after gestation. How many standard deviations away from the mean was Becky's pregnancy?

→ asking z-score



$x = 245$ $\quad z = ?$ $\quad \boxed{z = -1.31}$

$\mu = 266$

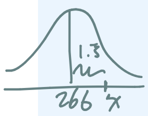
$\sigma = 16$

$z = \frac{245-266}{16} = -1.31$

Becky's pregnancy is 1.31 standard dev below the mean.

(3 pt) (b) Aria's baby was born 1.3 standard deviations above the mean. How long was Aria's gestation period?

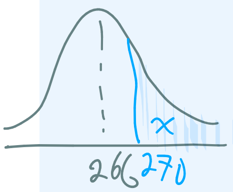
↳ $z = 1.3$ $\quad$ ↳ $x = ?$



$z = 1.3$

$x = \mu + z \cdot \sigma = 266 + (1.3)16 = \boxed{286.8 \text{ days}}$ Aria's gestation period was 286.8 days.

(4 pt) (c) Find the **probability** that one randomly selected expecting mother has a pregnancy that lasts more than 270 days?

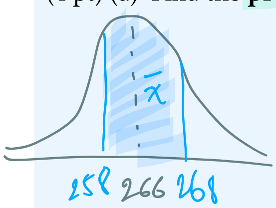


$P(x > 270) = P(x > 266) - P(266 < x < 270)$

$= 0.5 - \text{normalcdf}(266, 270, 266, 16) = \boxed{0.401}$

(4 pt) (d) Find the **probability** that fifty randomly selected expecting mothers have pregnancies lasting between 258 and 268 days?

↳ use sampling dist $n = 50$



Sampling

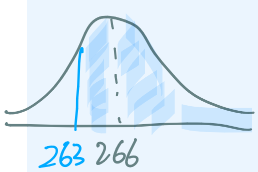
$\mu_{\bar{x}} = \mu = 266$

$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{16}{\sqrt{50}} = 2.26$ (but better to use exact)

$P(258 < \bar{x} < 268) = \text{normalcdf}(258, 268, 266, \frac{16}{\sqrt{50}}) = \boxed{0.811}$

if used $\sigma_{\bar{x}} = 2.26 \longrightarrow 0.812$

(4 pt) (e) What **percentage** of pregnancies last more than 263 days?



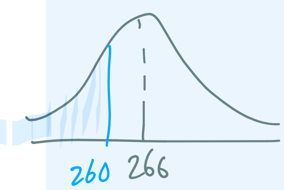
$P(x > 263) = P(263 < x < 266) + P(x > 266)$

$= \text{normalcdf}(263, 266, 266, 16) + 0.5 = 0.574$

But asking percentage: $\boxed{57.4\%}$

(4 pt) (f) Find the **probability** that fifteen randomly selected expecting mothers have pregnancies that lasts no more than 260 days?

↳ sampling $n = 15$



Sampling

$\mu_{\bar{x}} = \mu = 266$

$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{16}{\sqrt{15}} = 4.13$

$P(\bar{x} < 260) = P(x < 266) - P(260 < \bar{x} < 266)$

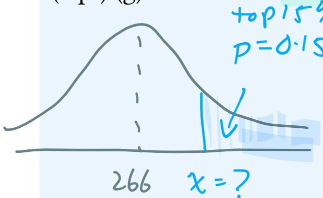
$= 0.5 - \text{normalcdf}(260, 266, 266, 16/\sqrt{15})$ $\quad$ or use 4.13

$= \boxed{0.0732}$

(4 pt) (g) What is the **length** of a human pregnancy that separates the top 15% of all pregnancy lengths?

top 15%

$p = 0.15$

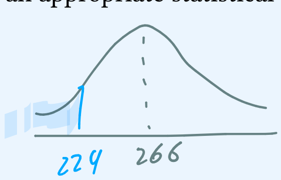


$x = \text{invNorm}(0.15, 266, 16, \text{RIGHT}) = \boxed{282.6 \text{ days}}$

(4 pt) (h) A "very preterm" baby is one whose gestation period is less than 224 days. Are very preterm babies **unusual**? Justify your answer with an appropriate statistical argument.

unusual

z score $\pm 2\sigma$ $\quad$ probability $< 5\%$



$P(\text{"very preterm"}) = P(x < 224) = 0.5 - \text{normalcdf}(224, 266, 266, 16)$

$= 0.004$

(MOE) "The probability of a very preterm baby is 0.004 or 0.4% which is much less than 5%, so yes they are considered unusual."

Problem 6: 16 pts

In the Parent-Teen Cell Phone Survey conducted by Princeton Survey Research Associates International, 800 randomly sampled 16- to 17-year-olds living in the United States were asked whether they have ever used their cell phone to text while driving. Of the 800 teenagers surveyed, 272 indicated that they text while driving. ← sample

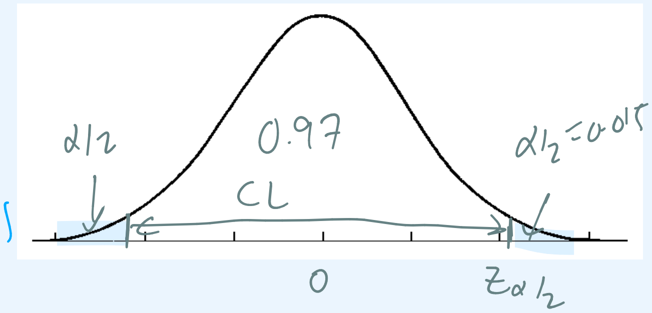
Obtain a 97% confidence interval for the proportion of 16- to 17-year-olds who text while driving. *Round final answers to three decimal places where appropriate* .000

↳ p

(2 pt) (a) Identify the **point estimate**:

$\hat{p} = \frac{272}{800} = \boxed{0.34}$ recall: proportions don't have units

$\hat{p} = \frac{x}{n}$ $x = 272$ $n = 800$



(3 pt) (b) Determine the **critical value**:

$CL = 0.97$

$\alpha = 1 - CL = 0.03$

$\alpha/2 = 0.015$

$z_{\alpha/2} = \text{invNorm}(0.015, 0, 1, \text{RIGHT})$

$= \boxed{2.17}$

(3 pt) (c) Find the **margin of error**:

$E = z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}\hat{q}}{n}}$ $E = (2.17^*)\sqrt{\frac{(0.34)(0.66)}{800}} = 0.0363\ldots$

$\hat{p} = 0.34$

$\hat{q} = 1 - \hat{p} = 0.66$ $n = 800$

$\boxed{E = 0.036}$ (Note: no units)

(3 pt) (d) Contsruct the **confidence interval**:

$CI: (\hat{p} - E, \hat{p} + E)$

$\hat{p} - E = 0.34 - 0.036 = 0.304$

$\hat{p} + E = 0.34 + 0.036 = 0.376$

$\boxed{CI: (0.304, 0.376)}$

(5 pt) (e) $(M \to E)$ **Interpretation of CI**:

"We are 97% confident that the proportion of all 16- to 17-year olds in the US who have texted while driving is between 0.304 and 0.376, or 30.4% and 37.6%."

Problem 7: 4 pts

A faculty member in the math department wants to construct a 98% confidence interval for the proportion of students at PCC who have children. What sample size is needed so that the confidence interval will have a margin of error of 0.1? *(Show work to receive full credit)* ↗ p

(A) 105 students.

(B) 137 students.

(C) 49 students.

(D) 48 students.

E (E) 136 students.

Givens: $CL = 0.98$, $E = 0.1$

Want: $n = ?$

Parameter: proportion p

Do we know sample $\hat{p}$? No

↳ use $n = \frac{[z_{\alpha/2}]^2 \, 0.25}{E^2}$

Can Find:

$\alpha = 1 - CL = 0.02$

$\alpha/2 = 0.01$

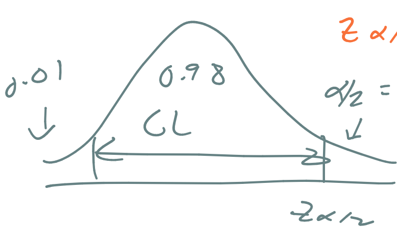
$z_{\alpha/2} = \text{invNorm}(0.01, 0, 1, \text{RIGHT})$

$= 2.33^*$

"Store → A"

$n = \frac{[2.33^*]^2 \cdot 0.25}{(0.1)^2}$

$n = 135.297\ldots$ (round up!)



Not normally dist!!!

Problem 8: 6 pts

The mean weight gain during pregnancy is 30 pounds, with a standard deviation of 12.9 pounds. Weight gain during pregnancy is skewed right. An obstetrician obtains a random sample of 35 low-income patients and determines their mean weight gain during pregnancy was 36.2 pounds.

↳ n = 35 sampling

What is the **probability** that a random sample of 35 low-income patients have a mean weight gain during pregnancy of 36.2 pounds or higher? Does this result suggest anything unusual? $(M \to E)$

- X = mean weight gain during pregnancy
- Asking for sampling dist.

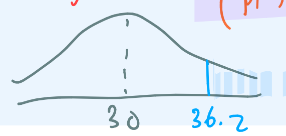
$n = 35 > 30$ ✓

$\mu_{\bar{x}} = \mu = 30$ lbs

$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{35}} = \frac{12.9}{\sqrt{35}} = 2.18^*$

- $\underbrace{P(\bar{x} > 36.2)}_{\text{prob that 35 women mean weigh gain 36.2 or higher}} = 0.5 - \text{normalcdf}\left(30, 36.2, 30, \frac{12.9}{\sqrt{35}}\right) = 0.00223$

$\boxed{P(\bar{x} > 36.2) = 0.00223}$



(M→E) The probability that a random sample of 35 women have a mean weight gain of 36.2 pounds is 0.00223, or 0.223%. This is indeed unusual. "This suggests the mean weight gain by low-income women is higher than 30 lbs."

Problem 9: 17 pts

The website fueleconomy.gov allows drivers to report the miles per gallon (mpg) of their vehicle. The data shown below shows the reported miles per gallon of 2011 Ford Focus automobiles for 16 different owners. Treat the sample as a simple random sample of all 2011 Ford Focus automobiles and assume that the mpgs are normally distributed.

Mr. T is Mean → use t-dist

mpg is average.

n = 16

35.7	37.2	34.1	38.9	32.0	41.3	32.5	37.1
37.3	38.8	38.2	39.6	32.2	40.9	37.0	36.0

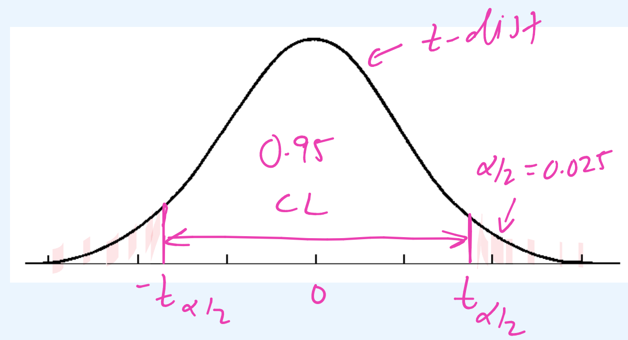
→ enter into a List & use 1-VARS STATS

$\bar{x} = 36.8 \quad s = 2.917$

Construct a 95% confidence interval for the mean miles per gallon of a 2011 Ford Focus. *Round final answers to two decimal places.* (where applicable)

(3 pt) (a) Identify the **point estimate**:

$\boxed{\bar{x} = 36.8 \text{ mpg}}$



(3 pt) (b) Determine the **critical value**:

$CL = 0.95$ $\qquad df = n - 1 = 15$

$\alpha = 1 - CL = 0.05$

$\alpha/2 = 0.025$

$t_{\alpha/2} = \text{invT}(\underbrace{1 - 0.025}_{1-\alpha/2}, \underbrace{15}_{df}) = \boxed{2.13^*}$ No units!

(3 pt) (c) Find the **margin of error**:

$E = t_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$

$= \frac{(2.13^*)(2.917)}{\sqrt{16}} = 1.5543\ldots \quad \boxed{E = 1.55 \text{ mpg}}$

Recall Error has same units as point estimate

(3 pt) (d) Contsruct the **confidence interval**:

$CI: (\bar{x} - E, \bar{x} + E)$

$\boxed{CI: (35.25 \text{ mpg}, 38.35 \text{ mpg})}$

$\bar{x} - E = 36.8 - 1.55 = 35.25$

$\bar{x} + E = 36.8 + 1.55 = 38.35$

(5 pt) (e) $(M \to E)$ **Interpretation of CI**:

"We are 95% confident that the mean miles per gallon of all 2011 Ford Focus cars is between 35.25 and 38.35 mpg."

Formula Sheet for Exam 3

- $\boxed{\text{normalcdf}(a, b, \mu, \sigma)}$
- $\boxed{z = \dfrac{x-\mu}{\sigma}}$ $\boxed{x = \mu + z \cdot \sigma}$
- $\boxed{CL = 1-\alpha}$, $\boxed{\alpha = 1 - CL}$, $\alpha/2$
- $\boxed{\text{invT}(1-\alpha/2, \text{df})}$
- $\boxed{E = \mathbf{t}_{\alpha/2}\dfrac{s}{\sqrt{n}}}$
- $\boxed{E = \mathbf{z}_{\alpha/2}\sqrt{\dfrac{\hat{p}\cdot\hat{q}}{n}}}$
- $\boxed{n = \left[\dfrac{z_{\alpha/2}\cdot\sigma}{E}\right]^2}$
- $\boxed{\text{invNorm}(\alpha, \mu, \sigma, \text{TAIL})}$
- $\boxed{\mu_{\bar{x}} = \mu}$ $\boxed{\sigma_{\bar{x}} = \sigma/\sqrt{n}}$
- $\boxed{n = \dfrac{[z_{\alpha/2}]^2\hat{p}\hat{q}}{E^2}}$
- $\boxed{n = \dfrac{[z_{\alpha/2}]^2 0.25}{E^2}}$

Post Exam Survey

Now that you have finished the exam, please take a few minutes to reflect on how you prepared for the exam and how you think you did. Then answer these questions.

1. When taking the exam I felt
   (a) Rushed. I wanted more time.
   (b) Relaxed. I had enough time.
   (c) Amazed. I had tons of extra time.
2. The week before the test I did all my homework on time: YES NO
3. The week before the test, in addition to the homework I followed a study plan. YES NO
   (a) I think this helped: YES NO
4. The day before the test I spend __________ hours studying and reviewing.
   (a) I think that was enough time: YES NO
5. The night before the test:
   (a) I stayed up very late cramming for the test
   (b) I stayed up very late, but I wasn't doing math
   (c) I didn't need to cram because I was prepared
   (d) I got a good night's sleep so my brain would function well.
6. I think I got the following grade on this test: __________
7. Strategies that worked well for me were (please elaborate):
8. Next time I will do an even better job preparing for the test by: