MATH 136 - Statistics — Spring 2020

Exam 3

Ch 7, 8, 9

May_13

Dr. Jorge Basilio

jbasilio@glendale.edu

SOLUTIONS

Directions

1. **At the top of your first page:** Please hand-write the statement provided below in quotes; print your name; put the date; and sign your name below it that acknowledges the **honor code**:

   "On my honor, by printing and signing my name, I vow to neither receive nor give any unauthorized assistance on this examination. I understand what my professor has deemed appropriate and inappropriate for this test and vow to follow these rules."

2. The exam is written to last 65 minutes, however, you have one hour and 45 minutes to submit this exam without penalty.
3. The exam will be available on Canvas at 6 PM. You will need to submit your hand-written solutions by 7:45 PM.
4. How to submit: upload a **single PDF file** of your solutions to Canvas no later than 7:45 pm to avoid penalties.
5. Write your solutions to the exam on one side of the page (front side only, do NOT write double-sided). You do NOT need to copy the questions on your piece of paper. However, you must submit the test problems in the order given and you must clearly label each problem and part. If I cannot identify which problem you are working on, no points will be given.
6. Make sure to ALWAYS SHOW YOUR WORK; you will not receive any partial credits unless work is clearly shown unless told otherwise. *If in doubt, ask for clarification.* Correct answers with little to no work will receive no points. Students might be randomly selected to have a 1-1 conference where you are asked to defend your work and explain to me all your steps on certain questions or problems that are similar to a test question.
7. **Penalties for late submissions:** Exams received between 7:46 PM and 7:55 PM will have 10 points deducted from their score. Exams received after 7:55 PM will not be graded and be given a score of 0.
8. **Allowed Materials**
   - You may use your calculator during the test (TI-83, 84, 84+, or 84+CET)
   - Blank pieces of paper to write your solutions. Writing utensils, erasers, etc
   - Formula Sheet (posted on Canvas and also at the end of this document)
   - Chi-Squared Table of Critical Values (posted on Canvas and also at the end of this document)
9. **Materials NOT Allowed**
   - Do not use your cell phone (for any reason: do not send or receive texts or calls or use the internet, etc)
   - Do not use your textbook (either digital or physical)
   - Do not use digital or printed out notes: the slides, the study guides, etc
   - Do not consult your HW
   - Do not give or receive any outside help (no getting help from a family member, friend, or any person either in person, via chat, message board, text message or any form of communication–again– you will be on camera the entire time so I will be looking for suspicious behavior)
   - Do not use your computer to look up anything using the internet (don't google; don't consult homework help websites, etc)

Continued on the next page...

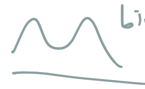
Directions Continued:

10. The exam totals **100 points**.
11. There are 10 problems; many of them with multiple parts.
12. Handwriting should be neat and legible. If I cannot read your writing, zero points will be given.
13. 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). For True/False questions, you must spell out the entire word "true" or "false" in your answer.
14. Leave answers in exact form (as simplified as possible), unless told otherwise.
15. Put a $\boxed{\text{box around your final answer}}$ where applicable.
16. **PLEASE INCLUDE UNITS** where applicable
17. **PLEASE CHECK YOUR WORK!!!**
18. GOOD LUCK!!!!

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: 8 pts (1 pts each)

TRUE or **FALSE** (please spell out/write the entire word for credit). *(No work needed)*  bimodal → not normal

(a) FALSE The distribution of ~~all~~ values of ~~any~~ only some random variable is called a normal distribution.

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

(c) TRUE The standard normal distribution has a mean of 1 and standard deviation of 0.

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

(e) TRUE The standard deviation of the distribution of $\bar{x}$ (sampling) decreases as n increases.

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

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

(h) FALSE A confidence interval for the variance σ^2 is of the form $(s^2 - E,\ s^2 + E)$, where E is the error.

↳ No! It has a totally different form. All others do this this form

Problem 2: 8 pts (1 pt each blank)

Fill in the blanks. *(No work needed)*

(a) The **probability density function (pdf)** of a random variable X must have area under the curve equal to 1 and every point on the curve must have vertical height greater than zero.

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

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

Simple random samples (SRS).

(d) If we are told the confidence interval for the variance is $(17, 23)$ then the point estimate is 20.

→ midpoint: $\frac{17+23}{2} = 20$

(e) The critical values that separate the middle 98% of a χ^2-distribution from each tail with $\boxed{\text{16 degrees of freedom}}$ are:

$\chi^2_L = \chi^2_{1-\alpha/2} = \chi^2_{0.99} = 5.812$ and $\chi^2_R = \chi^2_{\alpha/2} = \chi^2_{0.01} = 32.000$.

$CL = .98$
$\alpha = 0.02$
$\alpha/2 = 0.01$

Use table of critical values

 ← χ^2-dist

(f) The data for the weekly rate of return (in percent) of Microsoft for 16 randomly selected weeks shows that the standard deviation is 4.697. The 98% confidence interval for the standard deviation (risk) of Microsoft stock is:

(3.4235, 8.0331). *(Hint: use part (e))*

Use $n = 17$, $s^2 = 4.697^2 = 22.0618$

Formula sheet: standard deviation $\sqrt{\frac{(n-1)s^2}{\chi^2_R}} < \sigma < \sqrt{\frac{(n-1)s^2}{\chi^2_L}} \Rightarrow \sqrt{\frac{17(22.0618)}{32.000}} < \sigma < \sqrt{\frac{17(22.0618)}{5.812}}$

use calculator!

Problem 3: 8 pts (2 pts each)

Multiple-choice. Select the correct answer: *No work needed*

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

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

(B) The area under the curve is the probability.

(C) The curve is bell shaped.

(D) The mean is zero.

E (E) The standard deviation is 1.

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

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

(D) 99% of the $\cancel{\mu}$ will be in 100 confidence intervals.

E (E) 99 of them will include the true population μ.

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

(A) population proportion.

B (B) margin of error.

(C) sample proportion.

(D) population mean.

(E) sample mean.

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 7 and 8 cm.

μ_A

Company A provides tomatoes with diameters that are approximately normally distributed with mean 7.2 cm and standard deviation σ_A of 0.8 cm. Company B provides tomatoes with diameters that are approximately normally distributed with mean 7.7 cm and standard deviation of 1.1 cm. σ_B μ_B

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

next page

Problem 4

Usable tomatoes are between 7 and 8 cm, or $(7, 8)$.

Company A

x_A = tomato diameter from Company A

$\mu_A = 7.2$ cm

$\sigma_A = 0.8$ cm

Normally distributed

Probability = proportion of allowed tomatoes!

$$P(7 < x_A < 8) = \text{normalcdf}(7, 8, 7.2, 0.8)$$
$$= 0.440$$

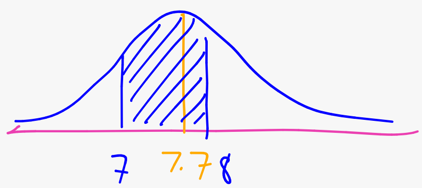
Company B

x_B = tomato diameter from Company B.

$\mu_B = 7.7$

$\sigma_B = 1.1$

Normally distributed



Probability = proportion of allowed tomatoes

$$P(7 < x_B < 8) = \text{normalcdf}(7, 8, 7.7, 1.1)$$
$$= 0.345$$

(M→E) "Company A provides a higher proportion of useable tomatoes since 44.0% will have a diameter between 7 & 8 cm, whereas only 34.5% from Company B will make the cut."

Problem 5: 26 pts

Read instructions carefully!

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).

normal cdf

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

units

z score

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

$$x = 245, \quad \mu = 266, \quad \sigma = 16 \longrightarrow z = \frac{x-\mu}{\sigma} = \frac{245-266}{16} = \boxed{-1.31}$$

z-score rounding rule!

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

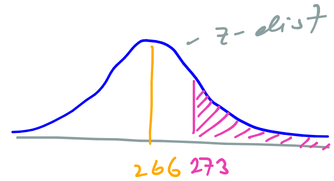
x-value

don't forget units!

$$z = 1.3, \quad \mu = 266, \quad \sigma = 16 \longrightarrow x = \mu + 1.3\sigma = 266 + (1.3)16 = \boxed{286.8 \text{ days}}$$

(4 pt) (c) Find the **probability** that one randomly selected expecting mother has a pregnancy that lasts more than 273 days?

$$P(x > 273) = \text{normalcdf}(273, 1E99, 266, 16) = \boxed{0.331}$$



(4 pt) (d) Find the **probability** that fifty randomly selected expecting mothers have pregnancies lasting between 261 and 271 days?

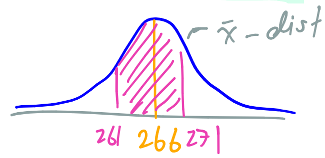
sampling! $n = 50$

Sampling

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

$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{16}{\sqrt{50}}$

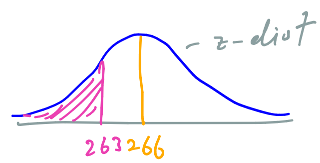
$$P(261 < \bar{x} < 271) = \text{normalcdf}\left(261, 271, 266, \frac{16}{\sqrt{50}}\right) = \boxed{0.973}$$



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

$$P(x < 263) = \text{normalcdf}(-1E99, 263, 266, 16) = 0.426$$

percentage! $\boxed{42.6\%}$



(4 pt) (f) Find the **probability** that fifteen randomly selected expecting mothers have pregnancies that lasts no more than 262 days?

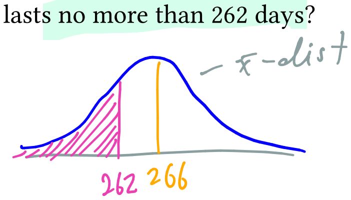
sampling! $n = 15$

Sampling

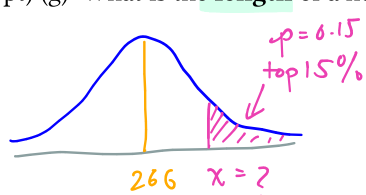
$\mu_{\bar{x}} = \mu = 266$

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

$$P(\bar{x} \le 262) = \text{normalcdf}\left(-1E99, 262, 266, \frac{16}{\sqrt{15}}\right) = \boxed{0.166}$$



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



Use inverse norm!

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

$$= \text{invNorm}(\underbrace{0.85}_{1-0.15}, 266, 16)$$

with old TI83 calc.

Problem 6: 15 pts

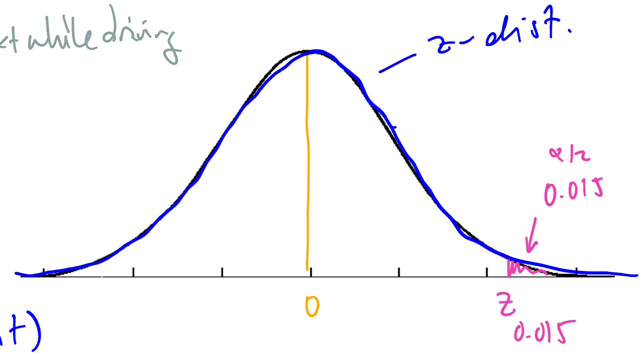
In the Parent-Teen Cell Phone Survey conducted by Princeton Survey Research Associates International, 800 randomly sampled 16- to 17-year-olds living in the United States were asked whether they have ever used their cell phone to text while driving. Of the 800 teenagers surveyed, 280 indicated that they text while driving.

Obtain a 97% confidence interval for the proportion of 16- to 17-year-olds who text while driving.

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

parameter: proportion of 16- to 17 year olds who text while driving

$\hat{p} = \frac{280}{800} = \boxed{0.35}$



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

$CL = 0.97$

$\alpha = 1 - CL = 0.03$

$\alpha/2 = 0.015$

$z_{\alpha/2} = z_{0.015}$

$= \text{invNorm}(0.015, 0, 1, \text{Right})$

$\approx \boxed{2.17}$

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

$E = z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} = (2.17)\sqrt{\frac{(0.35)(0.65)}{800}} = \boxed{0.037}$

$\hat{p} = 0.35$

$\hat{q} = 1 - \hat{p} = 1 - 0.35 = 0.65$

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

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

$\boxed{CI: (0.313, 0.387)}$

$\hat{p} \pm E = 0.35 \pm 0.037$

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

" We are 97% confident that the true proportion of 16- to 17-year olds who text and drive is between 0.313 and 0.387. "

Problem 7: 16 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

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

$n = 16$ 1-VARSTATS
$\bar{x} = 36.80$ (stats law of rounding)
$s = 2.92$

Construct a 97% confidence interval for the mean miles per gallon of a 2011 Ford Focus.

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

parameter: mean (miles per gallon)

point estimate: $\boxed{\bar{x} = 36.8 \text{ mpg}}$

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

$CL = 0.97$

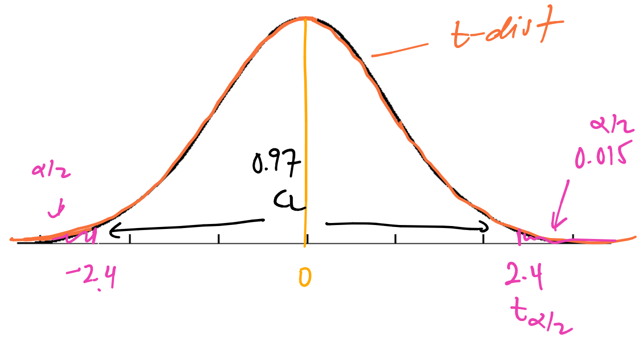
$\alpha = 1 - CL = 0.03$

$\alpha/2 = 0.015$

$t_{\alpha/2} = t_{0.015}$

$= \text{invT}(\underset{\alpha/2}{0.015}, \underset{df}{15})$

$= \boxed{2.40}$



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

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

$= 2.40 \cdot \frac{2.92}{\sqrt{16}} = \boxed{1.75 \text{ mpg}}$

Recall Error same units as the point estimate.

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

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

$\bar{x} \pm E = 36.8 \pm 1.75$

$\boxed{CI: (35.05, 38.55)}$

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

"We are 97% confident that the mean miles per gallon of all 2011 Ford Focus cars is between 35.05 mpg and 38.55 mpg."

Problem 8: 4 pts

A faculty member in the math department wants to construct a 96% confidence interval for the proportion of students at GCC who have children. What sample size is needed so that the confidence interval will have a margin of error of 0.05?
(Show work to receive full credit)

A (A) 421 students.

(B) 438 students.

(C) 420 students.

(D) 489 students.

(E) 500 students.

$\to \hat{p}$

Givens
- CL = 0.96
- E = 0.05

Want
- n = ?

Can find
- $\alpha = 1 - CL = 0.04$
- $\alpha/2 = 0.02$
- $z_{\alpha/2} = \text{invNorm}(0.02, 0, 1, R) = 2.05$

Formula proportion, don't know $\hat{p}$ $\to$ $n = \frac{(z_{\alpha/2})^2 \cdot 0.25}{E^2}$

$n = \frac{(2.05)^2 \cdot 0.25}{(0.05)^2} = 420.25 \to$ round up

421 students

Problem 9: 4 pts

A doctor wants to estimate the mean HDL cholesterol of all 20- to 29-year-old females. How many subjects are needed to estimate the mean HDL cholesterol within 4 percent with 96% confidence assuming a sample standard deviation equals 19.5 based on earlier studies?
(Show work to receive full credit)

E   CL   s

(A) 998,123 subjects.

B (B) 998,751 subjects.

(C) 998,903 subjects.

(D) 901,501 subjects.

(E) Cannot be determined.

Givens
- CL = 0.96
- E = 0.04
- s = 19.5

Want
- n = ?

Can find
- $\alpha = 1 - CL = 0.04$
- $\alpha/2 = 0.02$
- $z_{\alpha/2} = z_{0.02} = 2.05$

save time same as in Problem 8

Formula mean, don't know σ $\to$ use $n = \left[\frac{z_{\alpha/2} \cdot \sigma}{E}\right]^2$

$n = \left[\frac{(2.05)(19.5)}{0.04}\right]^2 = 998750.39\ldots$

$= 998751$ subjects

Problem 10: 5 pts

NOT NORMALLY DIST!

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.

μ σ $n = 35$ sampling

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

(Show work to receive full credit)

(A) 0.2;   No, this is not unusual.

B (B) 0.02;   Yes, this is unusual.

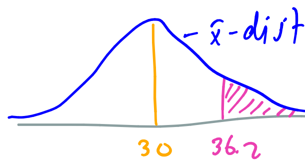
(C) 0.002;   Yes, this is unusual.

(D) 0.12;   No, this is not unusual.

(E) 0.05;   Yes, this is unusual.

$\mu_{\bar{x}} = \mu$ $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

- $P(\bar{x} \geq 36.2) = \text{normalcdf}\left(36.2, 1E99, 30, \frac{12.9}{\sqrt{35}}\right) = \boxed{0.00223}$



- Unusual if probability is less than 0.05 (or 5%): $\boxed{\text{Yes!}}$

Formula Sheet for Exam 3

- $\boxed{z = \dfrac{x-\mu}{\sigma}}$
  $\boxed{x = \mu + z \cdot \sigma}$
- $\boxed{\begin{cases} \mu_{\bar{x}} = \mu \\ \sigma_{\bar{x}} = \sigma/\sqrt{n} \end{cases}}$
- $\boxed{CL = 1 - \alpha}$
  $\boxed{\alpha = 1 - CL}$ $\alpha/2$
- $\boxed{\text{tcdf}(a, b, \text{df})}$
- $\boxed{E = \mathbf{z}_{\alpha/2}\sqrt{\dfrac{\hat{p}\cdot\hat{q}}{n}}}$
- $\boxed{|\text{invT}(\alpha/2, \text{df})|}$
- $\boxed{\text{normalcdf}(a, b, \mu, \sigma)}$
- $\boxed{\begin{cases} \mu_{\hat{p}} = p \\ \sigma_{\hat{p}} = \sqrt{\frac{\hat{p}\hat{q}}{n}} \end{cases}}$
- $\boxed{n = \dfrac{[z_{\alpha/2}]^2\hat{p}\hat{q}}{E^2}}$
- $\boxed{E = \mathbf{t}_{\alpha/2}\dfrac{s}{\sqrt{n}}}$
- $\boxed{\text{invNorm}(\alpha, \mu, \sigma, \text{TAIL})}$
- $\boxed{n = \dfrac{[z_{\alpha/2}]^2}{4E^2}}$
- $\boxed{n = \left[\dfrac{z_{\alpha/2}\cdot s}{E}\right]^2}$
- $\boxed{\dfrac{(n-1)s^2}{\chi^2_R} < \sigma^2 < \dfrac{(n-1)s^2}{\chi^2_L}}$
- $\boxed{\chi^2_L = \chi^2_{1-\alpha/2}}$ $\boxed{\chi^2_R = \chi^2_{\alpha/2}}$
- Reminder: Consult the provided table for critical values for the χ^2-distribution
- $\boxed{\sqrt{\dfrac{(n-1)s^2}{\chi^2_R}} < \sigma < \sqrt{\dfrac{(n-1)s^2}{\chi^2_L}}}$