Stat 50 - Elementary Statistics — Spring 2019

Exam 3 – Normal and Sampling Dist., CLT, Confidence Intervals

Study Guide & Review — **Dr. Jorge Basilio**

Chapters 6 & 7 & 8.1 — Pasadena City College — **gbasilio@pasadena.edu**

General Exam Info

Exams are a way for you to show me what you have learned (and please show all your steps so I can see this!) and to give you a sense of accomplishment! They are meant to be challenging and not just homework problems with the numbers changed. I really want to prepare you for university level math classes–so some exams may be longer or more challenging than others. Remember that I do grade fairly and my goal is to push you to succeed and excel in this class.

- Attendance required for all exams–I do NOT drop the lowest exam score.
- **"Make-up Exams"** are given only in extreme cases and at instructor's discretion; a student is allowed at most one make-up exam. (Documentation must be provided for the instructor to even consider a make up exam. This means you would need a doctor's note, etc.) A "Make-Up Exam" means you will be allowed to replace the missing score with the percentage you earn on the final exam. Please contact your instructor as soon as possible should there be a problem.
- **Your student ID is required for all exams.**
- During the exams–you will be required to leave your backpack and all non-test items at the front of the room, including cell phones and smart watches. Only your pencil/eraser and calculator will be allowed during the exam, and there will be a calculator check. Should you need to leave during the exam please ask for permission first before leaving and leave your cell phone with me. Not doing these things could result in a 0 on your exam.
- Once the exam is graded and returned, any problem you would like me to revisit must be brought to my attention by the next class session.
- Always keep your exams!

Exam 1 Date & Time		
• Exam 3	Thursday, May 9	1:15 pm – 2:35 pm

ALSO, after the exam:

- As the test is only 80 minutes long, we will have a 10 minutes break, then continue with new material.
- I take attendance at the end of class on test days.

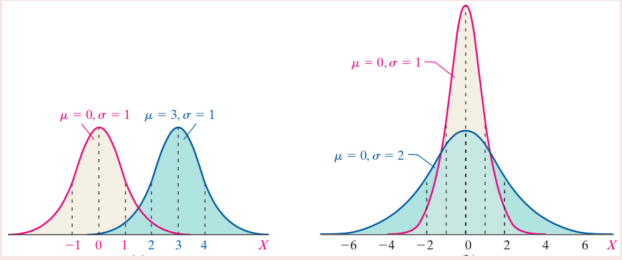
Exam 3 Specific Info

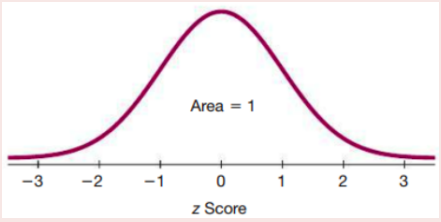
- Almost all questions have multiple parts
- You will need a calculator (only Ti83/84 allowed)–you won't be able to use your phone.
- You need to know what all the various terms in **bold** mean, but you don't need to memorize definitions. I'm not going to ask you to "Define ... ". Instead I ask you questions that use those terms.
- Material covered: Sections 6.1, 6.2, 6.3, 6.4, 7.1, 7.2, & 8.1
- Old material: You still need to know Chapters 1-5, especially Chapter 4 and 5 material. Basically: don't forget to study Exam 1 and Exam 2.
- I may ask you multiple choice/true-false/circle the right answer/fill-in the blanks/short response types of questions (you may need to provide examples).
- $(M \rightarrow E)$ **"Math to English"**If you see this, it just means that I would like you to write your answer in a complete sentence and include all relevant **units**.

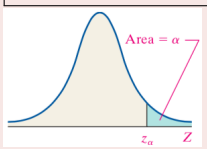
Chapter 6

§6.1

- X a continuous, RV.
- **density curve** associated to X
- **Important:** area under density curve = probability
- **Normal Distribution**
  - bell-shaped, symmetric, center at μ, standard deviation σ
  - there are lots of normal distributions (depending on μ and σ)



- **Standard Normal Distribution**
  - $\mu = 0$ and $\sigma = 1$
  - 

- **Finding probabilities given z-scores:**
  - $\boxed{P(a < X < b) = \text{normalcdf}(a, b, \mu, \sigma)}$
  - Can replace "$<$" or "$>$" with "$\leq$" or "$\geq$" for normal distributions (also, any continuous distribution)
  - Symmetry Trick:
    $\boxed{P(X < \mu) = 0.5}$ and $\boxed{P(X > \mu) = 0.5}$
  - If $a < \mu$, compute using symmetry trick:
    $P(X < a) = 0.5 - P(a < X < \mu)$
  - If $a < \mu$, compute using symmetry trick:
    $P(X > a) = P(a < X < \mu) + 0.5$
  - There's many more cases, but draw a picture!
- **Finding z-score given area/probability:**
  - $\boxed{z_\alpha = \text{invNorm}(\alpha, \mu, \sigma, \text{RIGHT})}$
  - 

  - If use older TI calculators: you must use left tail:
    $\boxed{z_\alpha = \text{invNorm}(1 - \alpha, \mu, \sigma)}$

§6.2

- Examples of Normal Distributions
- Important section, uses all the theory from §6.1
- Will need the formulas for z-scores from earlier:
  $\boxed{z = \dfrac{x - \mu}{\sigma}}$ and $\boxed{x = \mu + z \cdot \sigma}$

§6.3

- **Sampling Distribution**
  - **Sampling distribution of proportion**: 1) approximately normal, 2) sample proportion **targets** the value of the true population proportion
  - **Sampling distribution of mean**: 1) approximately normal, 2) sample mean **targets** the value of the true population mean
- **Estimators**
  - **Unbiased:** sample stat targets pop parameter
    Examples: $\hat{p}, \bar{x}, s^2$
  - **Biased:** sample stat DOESN'T target pop parameter
    Examples: Med, Range, s

§6.4

- **Central Limit Theorem:**
  - **Requirements:** X original distribution is normal or sample size $n > 30$.
  - CLT says: sampling distribution of sample means (samplING dist) can be approximated by normal distribution with $\boxed{\mu_{\bar{x}} = \mu}$ and $\boxed{\sigma_{\bar{x}} = \sigma/\sqrt{n}}$
- **WHEN TO USE CLT vs Normal Dist?**
  - when computing probability of **one random sample**: use §6.1, 6.2
  - when computing probability of **many random samples**: use CLT and samplING dist. Must check if $n > 30$ or original X is normal. USE $\mu_{\bar{x}} = \mu$ and $\sigma_{\bar{x}} = \sigma/\sqrt{n}$

Chapter 6

For the following I may ask you multiple choice/true-false/circle the right answer types of questions but mostly you'll see free-response questions similar to the worksheet problems:

- Study all of the examples from the worksheets carefully!
- Know what a density curve is and the properties of probability distributions
- Know the important fact that area=probability under a density curve
- Know the important properties of normal distributions, including the empirical rule (from previous chapters)
- Be able to compute z-scores
- Be able to compute probabilities given z-scores for many examples including a picture of the correct shaded area (there will lots of problems on this–see worksheet problems and hw problems). You will use your calculator with normalcdf. Especially using the symmetry properties discussed in class.
- Be able to compute a z-score given an area/probability. You will use your calculator with invNorm.
- Know what a sampling distribution is and how it is different than a regular distribution.
- Know what biased and unbiased estimators are.
- Be able to state the requirements, the givens, and the conclusion of the CLT (as in the worksheet).
- Be able to recognize when to use the CLT vs regular normal distribution when solving probability problems.

Chapter 7

§7.1

- **point estimate**
- **Confidence Interval for Population Proportion**:
  - **CI**: $(\hat{p} - E, \hat{p} + E)$
  - point estimate for population proportion is $\hat{p}$
  - $\hat{p} = \dfrac{x}{n} = \frac{\#\text{success}}{\#\text{ sample size}}$
  - **CL**: $CL = 1 - \alpha$, $\alpha = 1 - CL$, $\alpha/2$
  - **Critical Value**:
    $z_{\alpha/2} = \text{invNorm}(\alpha/2, 0, 1, \text{RIGHT})$
    uses standard normal distribution
    draw a picture!!
  - **Error**: $E = z_{\alpha/2}\sqrt{\dfrac{\hat{p} \cdot \hat{q}}{n}}$, where $\hat{q} = 1 - \hat{p}$.
- **Requirements for CI for pop. prop.:**
  1. SRS: simple random sample
  2. Conditions for binomial distributions: 4 conditions to check!
  3. $nq \geq 5$, i.e. "at least 5 successes and 5 failures"
- **MEANING**:
  A CI of 95% tells us that *the process* of finding the CIs should in the long run result in CIs that contain the true pop. prop. 95% of the time, that is 95% of the CIs will contain the true pop. prop.
- **Calculator**
  - STAT → TESTS → **1-PropZInt**
- **Determining sample size**:
  - $\hat{p}$ **known** → use $n = \dfrac{[z_{\alpha/2}]^2 \hat{p}\hat{q}}{E^2}$
  - $\hat{p}$ **unknown** → use $n = \dfrac{[z_{\alpha/2}]^2 0.25}{E^2}$
  - round up!

§7.2

- **point estimate**
- **Confidence Interval for Population Mean**:
  - **CI**: $(\bar{x} - E, \bar{x} + E)$
  - point estimate for population mean is $\bar{x}$
  - $\bar{x} = \dfrac{\sum x}{n}$, or using calculator **1-Var-Stats**
  - **CL**: $CL = 1 - \alpha$, α, $\alpha/2$
  - **Critical Value**: $t_{\alpha/2} = \text{invT}(1 - \alpha/2, \text{df})$
    *uses **student t-distribution***
    - bell-shaped, centered at $\mu = 0$, fatter tails
    - depends on **degrees of freedom**, df
      $df = n - 1$
    - As n increases, t-distribution approaches standard normal distribution
    - For $n > 30$, can assume t-distribution is standard normal distribution

    draw a picture!!
  - **Error**: $E = t_{\alpha/2}\dfrac{s}{\sqrt{n}}$.
- **Requirements for CI for pop. mean:**
  1. normality
  2. sample size $n > 30$
- **MEANING**:
  A CI of 95% tells us that *the process* of finding the CIs should in the long run result in CIs that contain the true pop. mean. 95% of the time, that is 95% of the CIs will contain the true pop. mean.
- **Calculator**
  - STAT → TESTS → **1-TInterval**
  - Either summary statistics given, or enter data into a list
- **Determining sample size**:
  - σ **known** → use $n = \left[\dfrac{z_{\alpha/2} \cdot \sigma}{E}\right]^2$
  - round up!

Chapter 7

For the following I may ask you multiple choice/true-false/circle the right answer types of questions but mostly you'll see free-response questions similar to the worksheet problems:

- Study all of the examples from the worksheets carefully!
- Be able to solve problems involving confidence intervals in **5 steps** (as in worksheets):
  1. Find **point estimate** (either $\hat{p}$ or $\bar{x}$)
  2. Find **critical value** (either $z_{\alpha/2}$ or $t_{\alpha/2}$)
  3. Find the **error** (either $E = z_{\alpha/2}\sqrt{\frac{\hat{p}\cdot\hat{q}}{n}}$ or $E = t_{\alpha/2}\frac{s}{\sqrt{n}}$)
  4. Find the **confidence interval** (either $(\hat{p} - E, \hat{p} + E)$ or $(\bar{x} - E, \bar{x} + E)$)
  5. $^\star(M \to E)^\star$ **Interpretation**
- Pay close attention to step 5 and practice writing your answers in a complete sentence.
- Pay close attention to the units given in the problem.
- Be able to determine the sample size needed for a confidence interval either for population proportion ($\hat{p}$ known or unknown) or mean (σ known).

Chapter 8

§8.1

- **Hypothesis Testing**:
  - **Null Hypothesis** H_0
    a null hypothesis will almost always be an "="
  - **Alternative Hypothesis** H_1 (or H_A)
    alt hypothesis will be: $<$, $>$, or $\neq$
- **Type of Test**:
  - **Left-Tailed**
  - **Two-Tailed**
  - **Right-Tailed**

- **Type of Errors**:
  - **Type I Error**
  - **Type II Error**
  -

		Reality	
		H_0 Is True	H_0 *is False* H_1 Is True
Conclusion	Do Not Reject H_0	Correct Conclusion	Type II Error
	Reject H_0	Type I Error	Correct Conclusion

- **Levels of Significance**:
  - $\boxed{\alpha = P(\text{Type I Error}) = P(\text{Reject } H_0 | H_0 \text{is True})}$
  - $\boxed{\beta = P(\text{Type II Error}) = P(\text{Fail to Reject } H_0 | H_0 \text{is False})}$
  - α and β are **inversely related**: if $\alpha \downarrow$ then $\beta \uparrow$, and if $\alpha \uparrow$ then $\beta \downarrow$

Chapter 8.1

For the following I may ask you multiple choice/true-false/circle the right answer types of questions but mostly you'll see free-response questions similar to the worksheet problems:

- Study all of the examples from the worksheets carefully in Section 8.1!
- Be able to identify and state the null and alternative hypotheses in problems using correct math notation.
- Be able to write correct complete sentences based on conclusions of hypothesis tests.
- Be able to identify type I or II errors in word problems and explain them in complete sentences.

Calculator Skills you must know

- STATS Edit lists, clear lists
- 1-Var Stats use to compute mean, standard deviation, 5 number summary
- 2nd+VARS Distributions
  - normalcdf(a,b,μ, σ)
  - invNorm(α, μ, σ,LEFT/CENTER/RIGHT) older Ti: can only use left tails. Use invNorm($1-\alpha, \mu, \sigma$)
- 2nd+ENTER Trick!
- Confidence Intervals:
  - for pop. proportion use: STAT $\rightarrow$ TESTS $\rightarrow$ **1-PropZInt**
  - for pop. mean use: STAT $\rightarrow$ TESTS $\rightarrow$ **1-TInterval**
    Either summary statistics given, or enter data into a list

Exam 3 Practice Problems

- Chapter 6 Quiz Review: p.291-292: # 1-9 all;
- Chapter 6 Review Exersizes: p.292-294: # 1-4 all, 5a,b, 6, 7a,b;

- Chapter 7 Quiz Review: p.350: # 2,-6 all, 9;
- Chapter 7 Review Exersizes: p.351-352: # 1-4 all, 5a,e, 7;

- Chapter 8 Quiz Review: p.407-408: #3;
- Chapter 8 Review Exersizes: p.408-409: #$1a, b, e, 5, 8a$;

Formula Sheet for Exam 3

- normalcdf(a, b, μ, σ)
- invNorm$(\alpha, \mu, \sigma, \text{TAIL})$
- $z = \dfrac{x-\mu}{\sigma}$ and $x = \mu + z \cdot \sigma$
- $\mu_{\bar{x}} = \mu$ and $\sigma_{\bar{x}} = \sigma/\sqrt{n}$

- $CL = 1-\alpha$, $\alpha = 1-CL$, $\alpha/2$
- $E = \mathbf{z}_{\alpha/2}\sqrt{\dfrac{\hat{p}\cdot\hat{q}}{n}}$
- $n = \dfrac{[z_{\alpha/2}]^2\hat{p}\hat{q}}{E^2}$
- $n = \dfrac{[z_{\alpha/2}]^2 0.25}{E^2}$
- invT$(1-\alpha/2, \text{df})$
- $E = \mathbf{t}_{\alpha/2}\dfrac{s}{\sqrt{n}}$
- $n = \left[\dfrac{z_{\alpha/2}\cdot\sigma}{E}\right]^2$