STA261 Exam III (Fall 2001) Name Key Row: ____

Instructions: *You must show all work to receive full credit, except in cases where calculator stat functions are used. Be sure your answers are in the context of the problems. Completeness and conciseness will benefit you. When performing any hypothesis test, be sure to give all elements of the test. Point values in <>.*

1. At a certain university there are 25,000 students. You collect a random sample of 425 students from this school and find that 126 of them are nearsighted.

a] Estimate the true proportion of the entire student body at this University that is nearsighted, using a 95% confidence interval. Fully interpret the CI in context. <10>

$\tilde{p} = \frac{126+2}{425+4} = \frac{128}{429} = .298$

CI: $.298 \pm 1.96\sqrt{\frac{.298(.702)}{425+4}} \rightarrow .298 \pm .043 \rightarrow (.255, .341)$

We can be 95% confident that between 25.5% to 34.1% of this student body is nearsighted.

b] The prevalence of nearsightedness in the general population is about 44%, based on empirical information. It is now your goal to find the sample size necessary to get a 95% confidence interval for p (the true proportion of the entire student body at this University that is nearsighted) that has a margin of error of ±0.03. However, you need a 'guessed' value of p in order to calculate n. Briefly explain, in real terms, why you should *not* use 0.44 as your guessed value of p in this case. <5>

Firstly, we have a real estimate that would be the best choice ($\tilde{p} = .298$). Using $\tilde{p} = .44$ may produce a bias because it is reflective of the general pop'n (which will include older people); the eyesight of the younger Univ. pop'n is likely better than this!

c] Refer to part b]. Use your best 'guessed' value of p to find the sample size n necessary to get a 95% confidence interval for p that has a margin of error of ±0.03. <5>

Using $\tilde{p} = .298$, $n = \frac{.298(1-.298)}{\left(\frac{0.03}{2}\right)^2} - 4 = 926$

2. Offer a concise, yet complete, one or two sentence critique of each of the following statements. <5+5>

a] *"An experiment compared two treatments (a study treatment and a placebo). Because the p-value in the hypothesis test for the efficacy of the study treatment was 0.0001, we can conclude that there is a strong effect due to the treatment."*

A small p-value only exhibits strong evidence of the existence of an effect, but nothing about the strength of the effect.

b] *"The study treatment appears to be effective in causing a positive response in individuals."*

Even though this was a designed experiment, one must be very cautious about concluding 'causation' without a full assessment of the effects of any lurking variables in the study.

3. To investigate the relationship between intracellular calcium and blood pressure, researchers measured the free calcium concentration (measured in nM) in the blood platelets of 38 people with normal blood pressure and 45 people with high blood pressure. The results are given in the table.

BP group	*n*	*sample mean*	*sample SD*
Normal	38	107.9	16.1
High	45	168.2	31.7

a] Is this an observational study or an experiment? Why? <6>

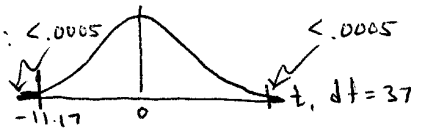
Observational – the BP of an individual is a pre-existing condition, not imposed by the researchers.

b] Use a *t* test to compare the means. Be sure to show all elements of the test, and use the 'conservative' approach to assessing df. Let $\alpha = 0.05$. <20>

$H_0: \mu_N - \mu_H = 0$

$H_A: \mu_N - \mu_H \neq 0$

TS: $t_s = \dfrac{(107.9 - 168.2) - 0}{\sqrt{\frac{16.1^2}{38} + \frac{31.7^2}{45}}} = -11.17$ $(df = \min(37, 44) = 37)$



$2 \times P(t > |-11.17|) < .0010$.

Conclusion: Since p-value < .05, reject H_0. The true mean free calcium concentration is signif. higher, on avg, in persons with high BP.

c] Construct a 95% confidence interval for the true mean difference in free calcium concentration between individuals with normal blood pressure and high blood pressure. Fully interpret your interval in context. <10>

$(168.2 - 107.9) \pm 2.042(5.399) \rightarrow 60.3 \pm 11.025$

$\rightarrow (49.27, 71.32)$

We are 95% confident that the true mean free calcium concentration is between 49.27 nM to 71.32 nM higher in persons with high BP than it is for persons w/ normal BP.

4. Match the following terms with their correct description by filling the blank with the correct letter. <2 ea>

A	alternative hypothesis	a]	*A statement about some population characteristic which a hypothesis test is generally trying to show is true*
D	*p*-value	b]	*A probability that defines the highest risk you are willing to take when rejecting Ho in favor of H_A.*
C	test statistic	c]	*A quantity used to measure the discrepancy between sample evidence and an "assumed" truth*
E	null hypothesis	d]	*A probability that measures the risk you will incur if you reject Ho in favor of H_A*
B	α	e]	*A statement about some population characteristic which is assumed to be true when running a hypothesis test*

Instructions: *Clearly write the letter corresponding to the most appropriate response in the space provided.* Each multiple choice problem is worth 3 points.

B **5.** Which of the following is **false**? "In a courtroom trial, ...

a] ...the prosecutor is like the researcher: the burden of proof is on them to substantiate their claim."
b] ...the evidence either substantiates guilt beyond a reasonable doubt, or it does not. If it does not, then the assumption of innocence held during the trial must be true."
c] ...there is a chance of mistakenly convicting an innocent person, but by requiring a "preponderance of the evidence", we manage to keep this risk low."
d] ...the evidence presented by the prosecution is much like the test statistic in a research scenario."

A **6.** When running a hypothesis test or confidence interval using Student's *t*-distribution, one must:

a] check that the parent populations are normally distributed, especially if the sample sizes are small.
b] check that the sample sizes are small, because *t*-based procedures are designed for small samples.
c] use the Central Limit Theorem to validate any conclusion.
d] assume nothing; there are no assumptions underlying a *t*-based inference.

A **7.** If the *p*-value in a hypothesis test is large, this means that:

a] a sample such as we actually saw would be likely if in fact *Ho* were true; so, we should not reject *Ho*.
b] a sample such as we actually saw would be unlikely if in fact *Ho* were true; so, we should reject *Ho*.
c] a sample such as we actually saw would be likely if in fact H_A were true; so, we should not reject H_A.
d] a sample such as we actually saw would be unlikely if in fact H_A were true; so, we should reject H_A.

Students in a large statistics class were randomly divided into two groups. The first group took a midterm exam with music playing in the background while the second group took the exam with no music. The scores of the two groups on the exam were compared. Questions 8 through 12 refer to this information.

C **8.** In this experiment, the experimental units are:

a] the songs played during the exam.
b] the scores on the midterm exam.
c] the students taking the exam.
d] the questions on the midterm exam.

A **9.** In this experiment, the treatments are:

a] whether or not music was playing during the exam.
b] the different scores on the midterm exam between the two groups.
c] lurking variables.
d] the kinds of music played for the first group.

C **10.** The study design for this experiment is called:

a] a simple random sample design.
b] a randomized blocks design.
c] a completely randomized design.
d] an observational study.

B **11.** This experiment was <u>not</u> *blind* because:

a] students were allowed to keep their eyes open while taking the exam.
b] the students knew whether or not music was playing while they were taking the exam.
c] some of the students did not study for the exam.
d] students were randomized into the two groups.

C **12.** To make this a good experiment, one part of it should be done in a blind way. To do this, we could:

a] make the students take the exam in a dark room.
b] let the students in the music group have extra time for the exam if they need it.
c] make sure that the exam grader doesn't know which students listened to music and which did not.
d] not allow the students to find out their grades on the exam.