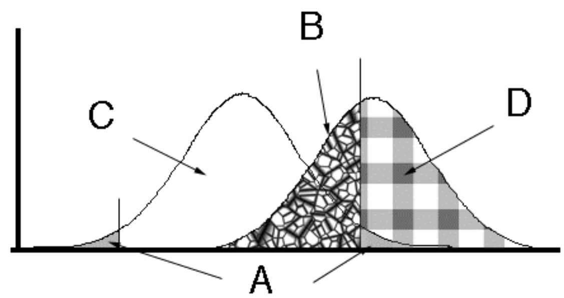
Exam 1 PS 217, Spring 2009

1. Label the graph below appropriately:



The distribution on the left represents H_0 True. The distribution on the right represents H_0 not true. So, in terms of hypothesis testing, what do A, B, C, and D represent? [2 pts]

Letter	Label	Letter	Label
A	**Type I Error (α)**	C	**Correctly retain H_0**
B	**Type II Error (β)**	D	**Correctly reject H_0 (1-β, power)**

2. You should recognize a few similarities between the curves in hypothesis testing (as above) and the curves in a signal detection paradigm. Indicate three similarities below: [3 pts]

In Hypothesis Testing Terms	In Signal Detection Terms
effect size, d (the distance separating the two distributions)	d'
D, power	Hits
B, Type II Error	Misses
A, Type I Error	False Alarms
α-level, value that leads you to reject H_0 (if larger) or retain H_0 (if smaller)	Criterion

3. Given that SAT math scores are normally distributed with $\mu = 500$ and $\sigma = 100$, answer the following questions: [10 pts]

a. If all the SAT math scores were transformed to *z*-scores, the distribution of *z*-scores would have parameters as follows (fill in the box below each parameter):

μ	σ
0	**1**

b. What proportion (or percentage) of people would have SAT math scores between 600 and 700? ***z*-scores of 1 and 2: .1359 or 13.59%**

c. What SAT math scores would determine the middle 50% of the distribution? (In other words, 25% above and 25% below the mean.) ***z*-scores of -0.67 and +0.67: 433 to 567**

d. What proportion of groups of 25 students taking the SAT-M exam would have a mean greater than 600? ***z*-score of 5, so almost 0 (extremely small proportion)**

4. Suppose that you take a sample of n = 16 students and determine the GPA for each of them, as seen below:

	GPA	GPA^2
	3.5	12.25
	3.0	9.0
	3.8	14.44
	2.9	8.41
	2.5	6.25
	3.5	12.25
	3.1	9.61
	3.0	9.0
	2.9	8.41
	3.4	11.56
	3.5	12.25
	2.1	4.41
	3.4	11.56
	3.8	14.44
	3.2	10.24
	2.7	7.29
Sum	50.3	161.37

a. Estimate the parameters of the population from which the sample was drawn. [5 pts]

$$\overline{X} = M = \hat{\mu} = \frac{50.3}{16} = 3.14$$

$$SS = 161.37 - \frac{50.3^2}{16} = 3.24$$

$$s^2 = \hat{\sigma}^2 = \frac{3.24}{15} = .216$$

$$s = \hat{\sigma} = .46$$

b. Educators often talk about grade inflation. Suppose that the mean GPA in 1960 was 2.6 (i.e., μ = 2.6). How likely is it that the sample above (problem 4) was drawn from a population with μ = 2.6? What might this evidence say about grade inflation? [10 pts]

H_0: μ = 2.6 H_1: μ ≠ 2.6 $t_{Critical}(15) = 2.131$

$$t = \frac{3.14 - 2.6}{.46 / \sqrt{16}} = \frac{.54}{.115} = 4.69$$

Decision: Reject H_0, because $t_{Obtained} \geq t_{Critical}$.
Conclusion: There does appear to be grade inflation, in that the current mean GPA (3.14) in your sample is significantly higher than the mean GPA in 1960.

5. Miscellaneous questions

a. Define the term *standard error*, both in a formula and in words that talk about its distribution. [2 pts]

$\sigma_{\bar{X}} = \sigma_M = \sigma / \sqrt{n}$, **it's the standard deviation of the sampling distribution of the mean**

b. Under which conditions is the median the preferred measure of central tendency? [1 pt]

When the distribution is skewed, the median provides a better sense of central tendency.

c. If you were to add a constant to all the scores in a sample, what would happen to the sample variance? [1 pt]

It would remain unchanged.

d. Why is the *SS* not a good measure of variability? [1 pt]

Because, the *SS* would typically increase with increases in n (number of scores in the distribution). (Only if scores equal to the mean were added would that not be true.)

e. In a positively skewed population, with $\mu = 85$ and $\sigma = 10$, what percentage of scores would be above 85? [1 pt]

You could either respond that you can't say, because the distribution isn't normal, or you could say that the percentage would be less than 50%.

f. What does the Central Limit Theorem tell you about the nature of the sampling distribution of the mean as sample size increases? [2 pts]

As the sample size (*n*) increases, the standard error will get smaller and the sampling distribution of the mean will approach a normal shape (as *n* approaches infinity).

g. Suppose that you are computing a *t*-test with $t_{Crit} = 1.96$. What can you tell me about that *t*-distribution? [1 pt]

Because of the large sample size, the *t*-distribution will be normal. That is, a two-tailed *t*-test with $\alpha = .05$ would be the same as conducting a *z*-test.

h. In the above situation (question 5g), if your *t* statistic came out as 1.94, what would you decide and what would you do next? [1 pt]

You would retain H_0, but you'd recognize that you had just missed the critical region, so you'd look for ways to increase power (e.g., increasing *n*).

i. In NHST, the probability of making a Type I Error is typically set to **5** %. [1 pt]

j. Interpret the SPSS output below as completely as you can. (In other words, what is the statistic, what are the scores, how would you interpret the results, etc.) [4pts]

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
IQ	22	131.2273	9.22694	1.96719

One-Sample Test

	Test Value = 120					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
IQ	5.707	21	.000	11.22727	7.1363	15.3183

First of all, you should recognize that the SPSS analysis is a single-sample *t*-test.

The analysis is testing H_0: μ = 120 vs. H_1: μ ≠ 120.

The $t_{Obtained}$ = 5.707, so the decision would be to reject H_0, with $p < .001$. Thus, you'd conclude that the IQ scores were likely drawn from a population with μ > 120.

k. In a short sentence, define power (1 – β). Then, tell me a strategy for making your statistic more powerful…being careful to articulate how that strategy works. [5 pts]

Power is the probability of correctly rejecting H_0. You could increase power by increasing the sample size, which decreases the standard error.