Soc3811 Second Midterm Exam

SEMI-OPEN NOTE:

One sheet of paper, signed

& turned in with exam booklet

Bring Your Own Pencil with Eraser

and a Hand Calculator!

Standardized Scores & Probability

If we know the mean and standard deviation of any population, then we can standardize the score of the *i*th case in the population:

$$Z_i = \frac{Y_i - \mu_Y}{\sigma_Y}$$

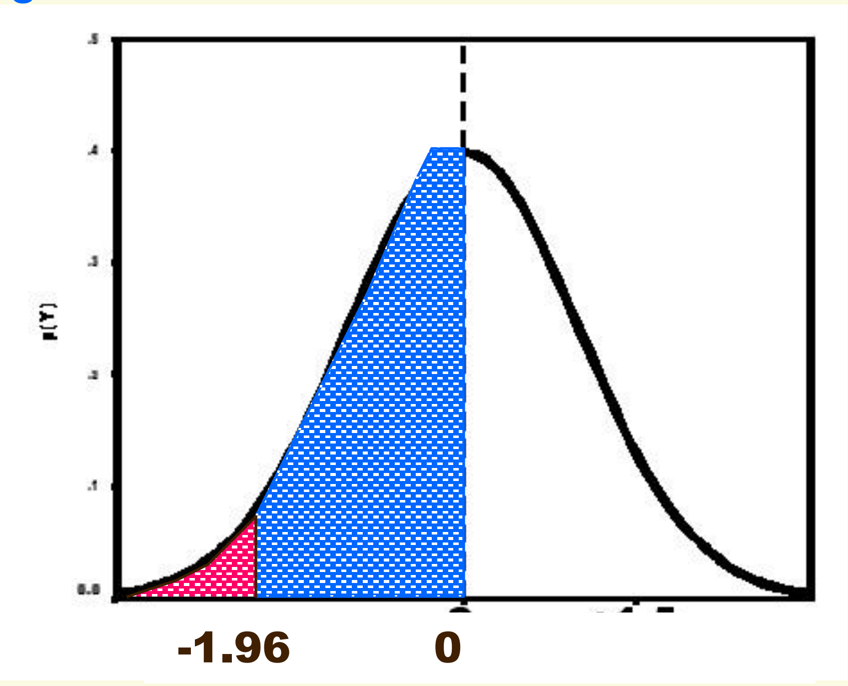
If the population has a normal distribution, we can use the known relation between Z scores and areas under the standardized normal curve to find probabilities of different segments.

Use the standardized Z-score table

(Appendix C) to look up the area(s):

(a) from the mean (Z=0) to ± Z_i

(b) from Z_i to ± ∞

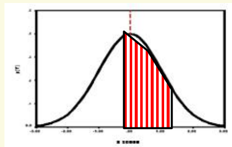
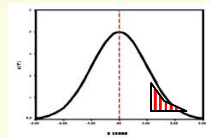


Convert Y_i to Z_i in a Population

In a population where $\mu_Y = 15$ and $\sigma_Y^2 = 4$, change these Y_i scores to Z_i :

Y_i	Z_i
28	________
18	________
12	________
8	________

Probabilities in the Z-Score Table

Z-score	Area from 0 to Z	Area from Z to ∞
		
1.50	0.4332	0.0668
...		
2.00	0.4772	0.0228
2.10	0.4821	0.0179
2.20	0.4861	0.0139
2.30	0.4893	0.0107
2.40	0.4918	0.0082
2.50	0.4938	0.062
2.60	0.4953	0.0047
2.70	0.4965	0.0035

A: What is the probability of a Z score between 0 and -2.10?

B: What is the probability of a Z score between -2.10 to -∞?

C: What is the probability of a Z score between +2.60 to +∞?

Sampling distribution mean & std. error

Even if a population is not normally distributed, the Central Limit Theorem assures that, for large *N* samples, the sampling distribution of sample means will approximate a normal distribution, whose mean and standard error have these relations to the population mean and standard deviation:

$$\mu_{\bar{Y}} = \mu_Y \qquad \sigma_{\bar{Y}} = \frac{\sigma_Y}{\sqrt{N}}$$

For a population with $\mu_Y = 38$ and $\sigma_Y = 3$, what are the sampling distribution means and standard errors for samples of these size Ns?

N	$\mu_{\bar{Y}}$	$\sigma_{\bar{Y}}$
400	______	______
800	______	______

Confidence Intervals

Another application of the Central Limit Theorem is to construct a confidence interval, a range of scores around a sample point estimate. If we either know or can estimate the standard error of a sampling distribution, then we can construct 95% and/or 99% confidence intervals around the point estimate of any sample statistic (such as a sample mean or proportion). For all the samples of size = N, the population parameter will fall inside the CI – the range of values between lower and upper confidence limits – 95% or 99% of the time, respectively.

$$\bar{Y} \pm (Z_{\alpha/2})(\sigma_{\bar{Y}})$$

Upper confidence limit, UCL:

$$\bar{\mathbf{Y}} + (\mathbf{Z}_{\alpha/2})(\boldsymbol{\sigma}_{\bar{\mathbf{Y}}})$$

Lower confidence limit, LCL:

$$\bar{\mathbf{Y}} - (\mathbf{Z}_{\alpha/2})(\boldsymbol{\sigma}_{\bar{\mathbf{Y}}})$$

Compute & Interpret CIs

A recent poll of 553 consumers finds that their mean optimism about the economy is $\overline{Y} = 68$. If the estimated standard error is $\sigma_{\overline{Y}} = 4$, what are the lower & upper limits of the 95% and 99% CIs?

$$\overline{\mathbf{Y}} \pm (\mathbf{Z}_{\alpha/2})(\boldsymbol{\sigma}_{\overline{\mathbf{Y}}}) = 68 \pm (1.96)(4)$$

LCL = ____________ UCL = ________________

$$\overline{\mathbf{Y}} \pm (\mathbf{Z}_{\alpha/2})(\boldsymbol{\sigma}_{\overline{\mathbf{Y}}}) = 68 \pm (2.58)(4)$$

LCL = ____________ UCL = ________________

INTERPRETATIONS: We can be 95% confident that the mean consumer optimism in the population is inside the interval from 60.2 to 75.8. We can be 99% confident that the population's mean consumer optimism falls into the range from 57.7 to 78.3.

Steps in Testing Hypotheses

1. Write the research hypothesis (H_1) & null hypothesis (H_0) in English
2. Restate the hypothesis pair in symbolic form. For two-tailed tests, rearrange to show the expected difference in parameter values:

$$H_0: \mu_1 = \mu_2 \qquad H_0: \mu_1 - \mu_2 = 0$$
$$H_1: \mu_1 \neq \mu_2 \qquad H_1: \mu_1 - \mu_2 \neq 0$$

3. Chose an α-level for H_0 (i.e., set the probability of Type I, or false rejection, error). Alternatively, after calculating the t-test (#5), determine the smallest α-level at which you can reject H_0
4. In the normal *(Z)* table, find <u>critical value(s)</u> (c.v.) of *t* for the α-level
5. Calculate the <u>*t*-test statistic</u>, using the sample size(s) and standard deviation(s) to estimate the sampling distrbution's standard error
6. Compare this *t*-test statistic to the c.v. to see if it falls into or outside the region(s) of rejection & decide whether to reject H_0 in favor of H_1
7. If you decide to reject the null hypothesis, H_0, then state the probability that you made a Type I, or false rejection, error (p = α)

Useful Formulas

Estimate the standard error using two sample values.

For a single sample:

$$\hat{\sigma}_{\bar{Y}} = \sqrt{s^2 / N} = s / \sqrt{N}$$

For two samples:

$$\hat{\sigma}_{(\bar{Y}_1 - \bar{Y}_2)} = \sqrt{s_1^2 / N_1 + s_2^2 / N_2}$$

Use this estimated standard error to compute the *t*-test.

For a single sample:

$$t = \frac{\bar{Y} - \mu}{s / \sqrt{N}}$$

For two samples:

$$t = \frac{(\bar{Y}_1 - \bar{Y}_2) - (\mu_1 - \mu_2)}{\sqrt{s_1^2 / N_1 + s_2^2 / N_2}}$$

Here are the critical values of *t* for the three conventional regions of rejection:

α (alpha)	One-tail c.v. for *t*	Two-tail c.v. for *t*
.05	1.65	±1.96
.01	2.33	±2.58
.001	3.07	±3.29

Write Hypotheses Pairs

To encourage you to invest in the stock market, your broker predicts that the Dow-Jones Industrial Average will rise past 12,000 sometime before the end of this year. Write her research and null hypothesis in both English language and symbolic form:

H_0: ______________________________

H_1: ______________________________

H_0: ________________

H_1: ________________

Is this a one-tail or two-tailed test? ____________________

Why? ____

____

Test a Null Hypothesis about One Mean

An economist believes that the average tax rebate was more than $300.

Sample statistics: Mean = $325; s.d. = $200; *N* = 400

Write the hypothesis pair:

H_0: __________

H_1: __________

Set $\alpha = .001$ and find c.v. for *t*-test: ________

Estimate standard error and the *t*-test:

$$t = \frac{\bar{Y} - \mu_Y}{s_Y / \sqrt{N}} = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_$$

Compare *t*-score to c.v., decide H_0:____________________

What is probability of Type I error? ____________________

Conclusion: ____

Test Another One

The mean patient stay in hospital this year differs from last year's 4.5 days.

Sample statistics: Mean = 4.7; s.d. = 1.7; *N* = 874

Write the hypothesis pair:

H_0: _______

H_1: _______

Set $\alpha = .05$ and find c.v. for *t*-test: _______

Estimate standard error and the *t*-test:

$$t = \frac{\bar{Y} - \mu_Y}{s_Y / \sqrt{N}} = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_$$

Compare *t*-score to c.v., decide H_0: _________________

What is probability of Type I error? _________________

Conclusion: ____

Test a Null Hypothesis about a Proportion

More than 80% of UM students graduate within six years.

Sample statistics: p = 0.85; *N* = 200

Write the hypothesis pair:

H_0: ________

H_1: __________

Set α = .01 and find c.v. for *t*-test: __________

Estimate standard error and the *t*-test:

$$t = \frac{p - \rho}{\sqrt{pq / N}}$$ **=** ____________________________

Compare *t*-score to c.v., decide H_0: ____________________

What is probability of Type I error? ____________________

Conclusion: ____

Test a Mean Difference Hypothesis

Students studying for the exam score higher than those who don't.

H_0: ________

H_1: ________

	Study	None
N	77	53
Mean	93.5	88.2
Variance	174.6	213.2

$$t = \frac{(\overline{Y}_1 - \overline{Y}_2) - (\mu_1 - \mu_2)}{\sqrt{s_1^2 / N_1 + s_2^2 / N_2}}$$

= ______________________________

Decision about null hypothesis: ______________

Probability of Type I error: ______________

Conclusion: ______________________________

Test a Proportion Difference Hypothesis

Anti-war attitudes differ between Independents and Democrats.

H_0: ________

H_1: ________

	Inds	Dems
N	189	277
Prop. p	.28	.34
Prop. q	.72	.66

$$t = \frac{(p_I - p_D) - (\rho_I - \rho_D)}{\sqrt{p_I q_I / N_I + p_D q_D / N_D}}$$

= ____________________________________

Decision about null hypothesis: ____________________

Probability of Type I error: ____________________

Conclusion: ____________________________________

Test a Paired Means Hypothesis

A family sociologist hypothesizes than husbands and wives differ in the mean number of household decisions they make.

H_0: ______

H_1: ______

$$t = \frac{\bar{Y}_D - \mu_D}{s_D / \sqrt{N}}$$

	Wives	Husbands
Mean	8.8	7.4
Sample N	238	
s_D	11.0	

= ______________________________

Decision about null hypothesis: ______________________

Probability of Type I error: ______________________

Conclusion: ______________________________