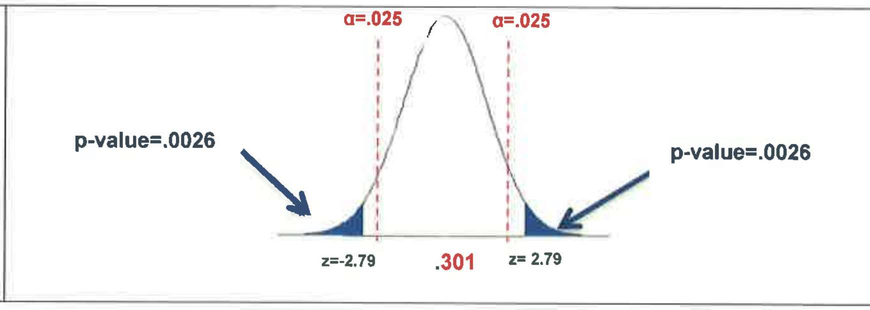
AP Statistics – 9.2b	Name:
Goal: 2-Sided TOH and CI for Population Proportion (p)	Date:

I. 2-Sided Test of Significance for Proportions -- Example #3 *"Benford's law and fraud"*

When the accounting firm AJL and Associates audits a company's financial records for fraud, they often use a test based on Benford's law. Benford's law states that the distribution of first digits in many real-life sources of data is not uniform. In fact, when there is no fraud, about 30.1% of the numbers in financial records begin with the digit 1. However, if the proportion of first digits that are 1 is significantly different from 0.301 in a random sample of records, AJL and Associates does a much more thorough investigation of the company. Suppose that a random sample of 300 expenses from a company's financial records results in only 68 expenses that begin with the digit 1. Should AJL and Associates do a more thorough investigation of this company?

- **Parameter of Interest** p = the true proportion of expenses that begin with the digit 1
- **Level of Significance** $\alpha = 0.05$ significance level
- **Choice of Test** one-sample z test for p
- **Null Hypothesis** $H_0: p = 0.301$
- **Alternative Hypothesis** $H_a: p \neq 0.301$
- **Conditions of Test**
  - Random: A random sample of expenses was selected.
  - Independent: It is reasonable to assume that there are more than 10(300) = 3000 expenses in this company's financial records.
  - Normal: $np_0 = (300)(0.301) = 90.3 \geq 10$, $n(1-p_0) = (300)(1-0.301) = 209.7 \geq 10$.
- **Sampling Distribution** (Sketch of the sampling distribution of the sample statistic under the null hypothesis, indicating the mean)

$\hat{p} = 68/300 = 0.227$
n=300



- **Test Statistic** (clearly show calculation)

$$\text{Test statistic: } z = \frac{0.227 - 0.301}{\sqrt{\frac{0.301(1-0.301)}{300}}} = -2.79$$

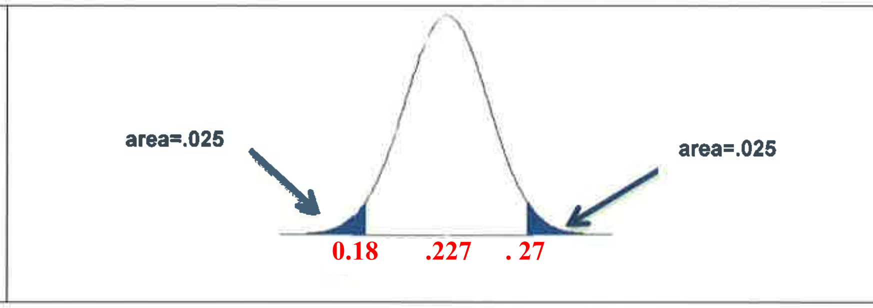
- **P-value** (Use correct probability notation.) P-value: $P(z < -2.79) = 0.0026$
  normalcdf(–e99,–2.79,0,1) = 0.0026 * 2 = 0.0052
- **Meaning of the P-value** (Reject or Fail to reject null hypothesis) **AND Conclusions** (in context)
  - Since the P-value (0.0052) is less than α=.05, **we reject the null hypothesis.**
  - **There is convincing evidence that the proportion of expenses that have first digit of 1 is not 0.301.** Therefore, AJL and Associates should do a more thorough investigation of this company.

II. Confidence Interval for Proportions -- Example #4 *"Benford's law and fraud(continued)"*

Problem: Find and interpret an appropriate confidence interval for the true proportion of expenses that begin with the digit 1 for the company in the previous alternate example. Use your interval from (a) to decide whether this company should be investigated for fraud.

- **Parameter of Interest** p = the true proportion of expenses that begin with the digit 1
- **Level of Significance** $\alpha = 0.05 \rightarrow$ 95% confidence level
- **Choice of Test** 1-sample z interval for p
- **Conditions of Test**
  - Random: A random sample of expenses was selected.
  - Normal: $n\hat{p} = 68 \geq 10$ and $n(1-\hat{p}) = 232 \geq 10$
  - Independent: It is reasonable to assume that there are more than 10(300) = 3000 expenses in this company's financial records.
- **Sampling Distribution** (Sketch of the sampling distribution of the sample statistic under the null hypothesis, indicating the mean)

$\hat{p}$= 68/300 = 0.227
n=300



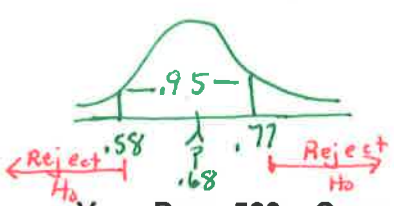
- **Confidence Interval** (clearly show calculation)

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.227 \pm 1.96\sqrt{\frac{0.227(1-0.227)}{300}} = 0.227 \pm 0.047 = (0.180, 0.274)$$

- **Interpret the confidence interval**
  - ***Conclude:*** **We are 95% confident that the interval from 0.180 to 0.274 captures the true proportion of expenses at this company that begin with the digit 1.**
- **Use your interval to decide whether this company should be investigated for fraud.**
  - ***Since 0.301 is not in the interval from, 0.301 is not a plausible value for the true proportion of expenses that begin with the digit 1, we reject Ho. Thus, this company should be investigated for fraud.***

III. CYU – page 558 – 2-sided Test (recommend using yellow sheet)– FAIL TO Reject Ho

IV. CYU – page 561 – Interpret Confidence Interval



We are 95% confident that the true proportion is between .58 and .77. Since this confidence interval includes our population parameter (.75), our decision would be the same "Fail to Reject".

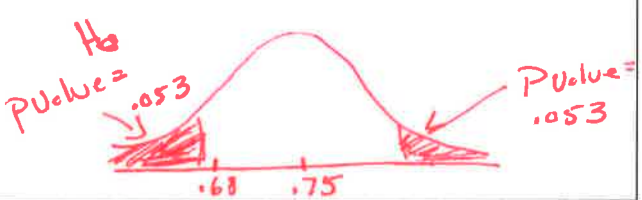
V. Page 563 – Compare the differences between doing a TOH(#50) versus CI (#52)

① We can use both TOH and CI to make significance decisions

② CI gives more information since it provides all the plausible values for the population parameter.

9.2 CYU page 558

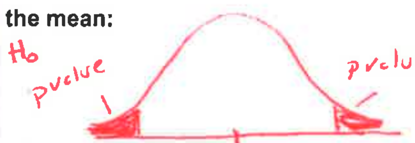
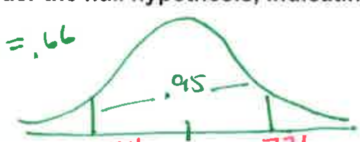
Test of Significance Template

Parameter of Interest	p = actual proportion of restaurant workers who say work stress has a negative impact
Choice of Test	1-Sample Z Test for a proportion
Level of Significance	$\alpha = .05$ (since not given)
Null Hypothesis	$H_0: p = .75$
Alternative Hypothesis	$H_A: p \neq .75$
Conditions of Test	Random - SRS $n = 100$ Independent - It is very likely there at least 10(100) = 1,000 restaurant workers Normal - $np = 68 \geq 10$ ✓ $n(1-p) = 32 \geq 10$ ✓
Sampling Distribution	**Sketch of the sampling distribution of the sample statistic under the null hypothesis, indicating the mean:** $n = 100$ $\hat{p} = \frac{68}{100} = .68$ 
Test Statistic	**Formula:** **Plug-ins & Value:** $Z = \frac{.68 - .75}{\sqrt{\frac{(.75)(.25)}{100}}} = -1.62$
P-value	**Use correct probability notation.** Pvalue = $P(Z < -1.62) = .053 * 2 = .106$
Meaning of the P-value	Since the pvalue (.106) is large and greater than $\alpha = .05$, WE FAIL TO REJECT H_0
Conclusions	☐ Reject null hypothesis ☒ Fail to reject null hypothesis **English:** We do not have sufficient evidence to say that the employees at this restaurant differs from the national average

9.2 page 563

Compare TOH vs CI

Test of Significance Template

	#50 TOH	#52 CI
Parameter of Interest	p = true proportion of 1st year college students being very well off financially	
Choice of Test	1 sample Z test for a proportion	1 sample z-interval for a proportion
Level of Significance	$\alpha = .05$	95% confidence level
Null Hypothesis	Ho: $p = .73$	
Alternative Hypothesis	H_A: $p \neq .73$	
Conditions of Test	Random: SRS n = 200 Independent: Reasonable there are more than 10(200) = 2,000 students at the university Normal: $np = .73(200) = 146 > 10$ ✓ $n(1-p) = .27(200) = 54 > 10$ ✓	Random: same Independent: same Normal: $np = 132 > 10$ ✓ $n(1-p) = 68 > 10$ ✓
Sampling Distribution	Sketch of the sampling distribution of the sample statistic under the null hypothesis, indicating the mean:  Ho pvalue pvalue $\hat{p} = \frac{132}{200} = .66$ .73	 $\hat{p} = .66$ .95 .594 .66 .726 $z^* = 1.96$
Test Statistic	Formula: $z = \dfrac{.66 - .73}{\sqrt{\frac{(.73)(.27)}{200}}} = -2.23$ Plug-ins & Value:	$.66 \pm 1.96\sqrt{\frac{(.66)(.34)}{200}}$ $.66 \pm .066$ (.594, .726)
P-value	Use correct probability notation. $P(z \leq -2.23) = .013 \times 2 = .026$	
Meaning of the P-value	Since the pvalue (.026) is less than $\alpha = .05$, we reject the null hypothesis	
Conclusions	☒ Reject null hypothesis ☐ Fail to reject null hypothesis	* Reject Ho
	English: We reject H_0 and conclude that the student at this college think that being well off differs from the national average.	We are 95% confident that the true population parameter is between .594 and .726. Since the population parameter (.73) falls outside the confidence interval, we can reject Ho.