Examples of Hypothesis Testing

Dr Tom Ilvento
Department of Food and Resource Economics

Overview

- Let's continue with some examples of hypothesis tests
  - introduce computer output
  - compare hypothesis test to confidence intervals
  - see what happens if we use a t versus a z for the Critical Value
  - See what happens with an outlier
  - Introduce hypothesis tests for proportions

Example: Test for Humerus Bones

- Humerus bones from the same species have the same length to width ratio, so they are often used as a means to identify bones by archeologists
- It is known that a Species A exhibits a mean ratio of 8.5
- Suppose 41 fossils of humerus bones were unearthed in East Africa
- Test whether the mean ratio from this sample differs from Species A ($\mu = 8.5$).
- Use $\alpha = .01$

Humerus Bones Example

- The length to width ratio was calculated for the sample and resulted in the following univariate statistics
  - n= 41
  - Mean = 9.26
  - s= 1.20
  - Min value = 6.23
  - Max value=12.00

Humerus Bones Hypothesis Test

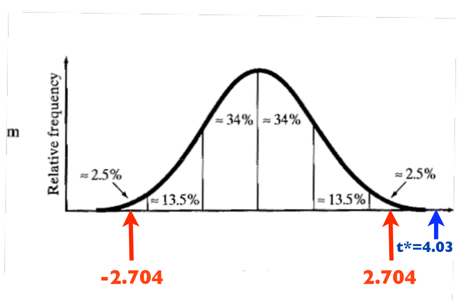
- **Set up the Null Hypothesis**
  - Ho: μ = ???
  - **Ho: μ = 8.5**
- **Set up the Alternative Hypothesis**
  - It takes up one of three forms
  - The problem asked to “Test whether the mean ratio from this sample differs from Species A”
  - **Ha: μ ≠ 8.5 Two-tailed**
- **Assumptions?**
  - **If large sample, > 30, use s as estimate of sigma and use a t-value**

The Components of a Hypothesis Test

• **Ho:**	• **Ho: μ = 8.5**
• **Ha:**	• **Ha: μ ≠ 8.5 2-tailed**
• **Assumptions**	• **n= 41, σ unknown, use t**
• **Test Statistic**	• **t* = (9.258 – 8.5)/.188**
• **Rejection Region**	• **α = .01, .01/2, 40 d.f., t = ± 2.704**
• **Calculation:**	• **t* = 4.032**
• **Conclusion:**	• **$t^* > t_{.01/2,\ 40\ df}$**
	• **4.032 > 2.704**
	• **Reject Ho: μ = 8.5**

Here’s how it Looks in Pictures

- Our critical values were **-2.704** and **2.704**
- Our test statistic was **4.032**
- the test statistic is in the rejection region on the right hand side
- I can also see it from the output from JMP



What would have happened if I used a z-test in place of the t-test?

• **Ho:**	• **Ho: μ = 8.5**
• **Ha:**	• **Ha: μ ≠ 8.5 2-tailed**
• **Assumptions**	• **n= 41, σ unknown, large sample use z**
• **Test Statistic**	• **z* = (9.258 – 8.5)/.188**
• **Rejection Region**	• **α = .01, .01/2, z = ± 2.575**
• **Calculation:**	• **z* = 4.032**
• **Conclusion:**	• **$z^* > z_{.01/2}$**
	• **4.032 > 2.575**
	• **Reject Ho: μ = 8.5**

Peanut Package Problem

- A peanut company sells a package product of 16 oz of salted peanuts through an automated process
- Not all packages contain exactly 16 oz of peanuts – they shoot for an average of 16 oz with a standard deviation of .8 oz.
- They routinely take random samples of 40 packages and weigh them
- They want to see if each sample is different from the package claim at α=.1

Peanut Package Problem

- If the manufacturing process overfills the packages, even by a little, they lose profit
- If the manufacturing process under-fills the packages they risk angry customers and fines from government
- They are interested in a **two-tailed test**, a priori
- Let's say they take a sample of 40 packages and get a mean value of **16.42**
- Does this sample result warrant checking the manufacturing process?
- Note: this is a problem where we can view σ as being known:
  - **σ = .8**
  - **SE= $.8/40^{.5}$ = .1265**
  - **Use the z-distribution**

The Components of a Hypothesis Test

- **Ho:**
- **Ha:**
- **Assumptions**
- **Test Statistic**
- **Rejection Region**
- **Calculation:**
- **Conclusion:**

- **Ho: μ = 16.0**
- **Ha: $\mu \neq$ 16.0 2-tailed**
- **n= 40, σ =.8, use z**
- **z^* = (16.42 – 16.00)/.1265**
- **α = .10, .10/2, z = ± 1.645**
- **z^* = 3.32**
- **$z^* > z_{.10/2}$**
- **3.32 > 1.645**
- **Reject Ho: μ = 16.0**

90% Confidence Interval

- Calculate the 90% confidence interval for this problem
  - $16.42 \pm 1.645[.8/(40)^{.5}]$
  - 16.42 ± .208
  - 16.21 to 16.63
- **Note that 16 is NOT in the 90% C.I.**
- A similar (1- α) C.I. will generate the same result as a two-tailed hypothesis test
  - **If the the Null value is in the C.I.**
  - **You cannot reject Ho**

Another way to approach this problem using C.I.

- Another way to use a confidence interval:
  - Calculate the C.I. Around 16 oz
    - **$16 \pm 1.645[.8/(40)^{.5}]$**
    - **$16 \pm .208$**
    - **15.792 to 16.208**
- Any sample that falls outside of this interval will cause them to reject the null hypothesis (based on two-tailed test and α = .1)
- **Note: Type I Error = .1 They can expect to wrongly reject Ho: 10 of 100 times**

Let's try a problem together

- The Body mass index (BMI) is a measure of body fat based on height and weight that applies to both adult men and women.
- A BMI > than 30 is considered obese.
- A random sample of adults participated in a health study, and 13 of them had a BMI > 30.
- We will look at the systolic blood pressure reading, which represents the maximum pressure exerted when the heart contracts.
- Assume the systolic blood pressure follows something like a normal distribution and an unhealthy reading is greater than 120.
- **We want to test to see if people with BMI > 30 tend to have a systolic blood pressure reading greater than 120.**
- **Use α = .10**

Systolic Blood Pressure for patients with BMI >30

- Here are the results from JMP
- You take the relevant information

SYS BP

Quantiles

100.0%	maximum	181.00
99.5%		181.00
97.5%		181.00
90.0%		170.60
75.0%	quartile	132.00
50.0%	median	125.00
25.0%	quartile	113.50
10.0%		108.20
2.5%		107.00
0.5%		107.00
0.0%	minimum	107.00

Moments

Mean	127.615
Std Dev	20.304
Std Err Mean	5.631
upper 95% Mean	139.885
lower 95% Mean	115.346
N	13.000
Sum Wgt	13.000
Sum	1659.000
Variance	412.256
Skewness	1.780
Kurtosis	3.424
CV	15.910
N Missing	0.000

Stem and Leaf

Stem	Leaf	Count
18	1	1
17		
17		
16		
16		
15	5	1
15		
14		
14		
13		
13	13	2
12	556	3
12	3	1
11	6	1
11	034	3
10	7	1

10|7 represents 107

Hypothesis Test for Sys BP

- **Ho:**
- **Ha:**
- **Assumptions**
- **Test Statistic**
- **Rejection Region**
- **Calculation:**
- **Conclusion:**

Hypothesis Test for Sys BP

- **Ho:** — **Ho: μ = 120**
- **Ha:** — **Ha: μ > 120   1-tailed upper**
- **Assumptions** — **n= 13, σ unknown, use t**
- **Test Statistic** — **t* = (127.615 – 120)/5.631**
- **Rejection Region** — **α = .10, 12 d.f., t = 1.356**
- **Calculation:** — **t* = 1.352**
- **Conclusion:** — **t* < $t_{.10,\ 12\ df}$**
- **1.352 < 1.356**
- **Cannot Reject Ho: μ = 120**

JMP output for the Hypothesis Test

- JMP shows the same output, but not the t-value for the Critical Value
- Instead it gives a p-value
- This is the probability of finding a value greater than the test statistic into the tail
- as either a one-tail or two-tail test
- We would compare the p-value for the appropriate test to α

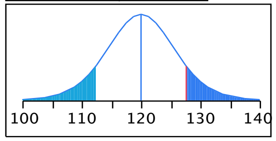
SYS BP

Test Mean=value

Hypothesized Value	120
Actual Estimate	127.615
df	12
Std Dev	20.3041

t Test

Test Statistic	1.3523
Prob > |t|	0.2012
Prob > t	0.1006
Prob < t	0.8994



One value was Extreme, 181, what happens if we remove it?

SYS BP

Moments

Mean	123.167
Std Dev	13.002
Std Err Mean	3.753
upper 95% Mean	131.428
lower 95% Mean	114.905
N	12.000
Sum Wgt	12.000
Sum	1478.000
Variance	169.061
Skewness	1.242
Kurtosis	2.356
CV	10.557
N Missing	0.000

Stem and Leaf

Stem	Leaf	Count
15	5	1
15		
14		
14		
13		
13	13	2
12	556	3
12	3	1
11	6	1
11	034	3
10	7	1

10|7 represents 107

- The data are better behaved. The changes:
  - the mean is lower,
  - but so is the standard deviation and ultimately the standard error;
  - we lose a degree of freedom

- **Ho: μ = 120**
- **Ha: μ > 120   1-tailed upper**
- **n= 12, σ unknown, use t**
- **t* = (123.167 – 120)/3.753**
- **α = .10, 11 d.f., t = 1.363**
- **t* = .8439**
- **t* < $t_{.10,\ 11\ df}$**
- **.8439 < 1.363**
- **Cannot Reject Ho: μ = 120**

Hypothesis Tests for Proportions

- The Pepsi Challenge asked soda drinkers to compare Diet Coke and Diet Pepsi in a blind taste test.
- Pepsi claimed that more than ½ of Diet Coke drinkers said they preferred Diet Pepsi (P=.5)
- Suppose we take a random sample of 100 Diet Coke Drinkers and we found that 56 preferred Diet Pepsi.
- **Use α = .05 level to test if we have enough evidence to conclude that more than half of Diet Coke Drinkers will prefer Pepsi.**

Hypothesis Test for a Proportion

- Hypothesis test for proportions is the same
- It must be based on a large sample
- We have an estimate of the population parameter, P, from a sample - *p*
- We use the same strategy of comparing our sample estimate to the theoretical sampling distribution
- And the same formulas
- **But, with one slight twist!**

Remember, if we have additional information we should use it

- With proportions we have a slightly different approach to the standard error
- Remember, the variance, std dev, and standard error of a proportion is tied to P or *p*
- $\sigma^2 = PQ$
- $\sigma = (PQ/n)^{.5}$
- If we hypothesize that P = .5 under a null hypothesis
- **Then we ought to use the hypothesized P and Q as the components for the standard error of the sampling distribution**

$$\sigma_P = \sqrt{(.5 \cdot .5)/100}$$

$$\sigma_P = \sqrt{.25/100} = .05$$

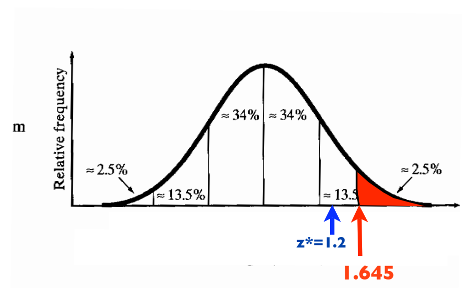
Pepsi Challenge Hypothesis Test

- **Ho:**
- **Ha:**
- **Assumptions**
- **Test Statistic**
- **Rejection Region**
- **Calculation:**
- **Conclusion:**

- **Ho: P = .5**
- **Ha: P > .5 1-tailed, upper**
- **n= 100, σ =.25, binomial = normal**
- **$z^* = (.56 - .5)/.05$**
- **$\alpha = .05$, $z = 1.645$**
- **$z^* = 1.20$**
- **$z^* < z_{.05}$**
- **1.20 < 1.645**
- **Cannot Reject Ho: P = .5**

Here's how it Looks in Pictures

- Our critical value was **1.645**
- Our test statistic was **1.20**
- the test statistic is not in the rejection region on the right hand side



Summary

- I hope you are getting more comfortable with the mechanics of a hypothesis test
- Take it step by step
- Determine if the problem is dealing with a proportion or a mean
- And if a mean,
  - do we know sigma?
  - Is the sample size large?
  - Can we reasonable assume the population variable follows a normal distribution?