Difference of Means and ANOVA Problems

Dr. Tom Ilvento
FREC 408

Accounting Firm Study

- An accounting firm specializes in auditing the financial records of large firm
- It is interested in evaluating its fee structure,particularly in relation to charges by the size of the firm
- It takes a random sample of 10 companies from three size classes
  - Sales over $250 million
  - Sales of $100 to $250 million
  - Sales of Less than $100 million

Stem and Leaf Plot

STEM	Leaf
0	20 52 55 75 76 80 80 88 92
1	00 00 10 25 32 41 50 50 50 60 86
2	00 00 30 33 50 75
3	25
4	75
5	
6	00
7	
8	00

Descriptive Statistics with 95% C.I.

	$250+	$100-$250	< $100	Total
Mean	335.50	129.50	106.00	190.33
Standard Error	70.84	18.03	18.02	30.87
Median	262.50	121.00	90.00	145.50
Mode	150	#N/A	#N/A	150
Standard Deviation	224.02	57.02	56.99	169.08
Sample Variance	50185.83	3251.17	3247.78	28587.08
Kurtosis	0.84	-0.48	-0.51	5.79
Skewness	1.16	0.57	0.41	2.30
Range	700	178	180	780
Minimum	100	55	20	20
Maximum	800	233	200	800
Sum	3355	1295	1060	5710
Count	10	10	10	30
Confidence Level(95.0%)	160.26	40.79	40.77	63.13

Things to note

- The mean cost for the largest firms is much larger than the other two firm classes
- While there is a difference in the means for the lower two classes, when we look at the BOE for the 95% C.I. there is overlap
- The variances of the lower two firms is very similar – but both are considerably lower than that for the largest firms

Difference of Means Test

- Let's test to see if there is a significant difference between firms of sales between
  - $100 to $250 million
  - Less than $100 million
- Is this reasonable?
  - Ratio of variances is
  - 3251.17/3247.78 = 1.001

Decision Tree for Two Means

Target	Assumptions	Test Statistic
H_0: $\mu_1-\mu_2=D$	Independent random samples Large sample size (n_1, n_2>30)	z, using sample variance
	Independent random samples Small sample size Populations appr. normal **Equal variances**	t, using pooled variance S_p^2

POOLED ESTIMATE OF THE VARIANCE

- **The Our formula will be a weighted average of s_1 and s_2**

$$s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{(n_1+n_2-2)}$$

$$s_p^2 = \frac{(10-1)3251.17 + (10-1)3247.78}{(10+10-2)}$$

$$s_p^2 = \frac{58490.55}{18} = 3249.475 \qquad s_p = 57.004$$

NOTE

- **Since the sample sizes were equal, we could have simply taken the average of the two variances**
- **(3251.17 + 3247.78)/2 = 3249.475**

Use the Pooled Estimate of the Variance to calculate the standard error

$$\hat{\sigma}_{(\bar{x}_1-\bar{x}_2)} = s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

$$\hat{\sigma}_{(\bar{x}_1-\bar{x}_2)} = 57.004\sqrt{\frac{1}{10}+\frac{1}{10}} = 25.493$$

Accounting Firm Problem

Null hypothesis	**H_0: $(\mu_1-\mu_2) = 0$**
Alternative	**H_a: $(\mu_1-\mu_2) \neq 0$ two-tailed test**
Assumptions	**Small independent samples, approx normal, variances are equal**
Test Statistic	**$t^* = (129.50-106.00 - 0)/25.493$**
Rejection Region	**$t_{.05/2,\ 18\ d.f.} = 2.101$**
Calculation	**$t^* = .922$**
Conclusion	**$t^* < -t_{.05/2,\ 18\ d.f.}$** **.922 < 2.101** **We cannot reject H_0: $(\mu_1-\mu_2) = 0$**

EXCEL Output

t-Test: Two-Sample Assuming Equal Variances		
	$100-$250	*< $100*
Mean	129.500	106.000
Variance	3251.167	3247.778
Observations	10	10
Pooled Variance	3249.472	
Hypothesized Mean Difference	0	
df	18	
t Stat	0.922	
P(T<=t) one-tail	0.184	
t Critical one-tail	1.734	
P(T<=t) two-tail	0.369	
t Critical two-tail	2.101	

Let's shift to ANOVA

- **ANOVA would allow us to test the difference of means across all classes of firm size**
- **We need to assume equal variances – is this reasonable?**

Review Degrees of freedom

- **k = 3**
- **n = 30**

- **SS(Total) df = n-1 = 30 - 1 = 29**
- **SST df = k-1 = 3 – 1 = 2**
- **SSE df = n-k = 30 – 3 = 27**

Excel ANOVA Output

Anova: Single Factor						
SUMMARY						
Groups	Count	Sum	Average	Variance		
$250+	10	3355	335.50	50185.83		
$100-$250	10	1295	129.50	3251.17		
< $100	10	1060	106.00	3247.78		
ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	318862	2	159430.83	8.44	0.00	3.35
Within Groups	510163	27	18894.93			
Total	829025	29				

ANOVA Hypothesis Test

- **Null hypothesis**
- **Alternative**
- **Assumptions**
- **Test Statistic**
- **Rejection Region**
- **Conclusion**

- **H_0: $\mu_1 = \mu_2 = \mu_3$**
- **H_a: At least two means differ**
- **Equal variances, normal distribution**
- **$F^* = 8.44$**
- **$F_{.05,\ 2,\ 27\ d.f.} = 3.35$**
- **$F^* > F$**
- **$8.44 > 3.35$**
- **Reject H_0: $\mu_1 = \mu_2 = \mu_3$**

Brake Test

- A firm makes disc brakes for the automobile industry
- The R&D department tested four different brake systems
- In the test, they used 40 identical mid-sized cars, 10 each for the four brake systems
- The cars were driven on a test track and stopped electronically
- They measured the distance in feet to bring the car to a stop
- **Are there differences in stopping distance across the different brakes?**

The data

	Brake A	*Brake B*	*Brake C*	*Brake D*	*Total*
Mean	272.300	271.200	262.300	265.100	267.725
Standard Error	2.290	2.525	1.476	3.250	1.360
Median	274.500	270.000	264.000	259.000	267.500
Mode	275	267	264	259	257
Standard Deviation	7.243	7.983	4.668	10.279	8.602
Sample Variance	52.456	63.733	21.789	105.656	73.999
Kurtosis	0.192	0.270	-1.646	-0.641	-0.785
Skewness	-0.722	0.714	-0.008	0.945	0.399
Range	24	26	12	28	32
Minimum	259	261	257	255	255
Maximum	283	287	269	283	287
Sum	2723	2712	2623	2651	10709
Count	10	10	10	10	40
Confidence Level(95.0%)	5.181	5.711	3.339	7.353	2.751