Statistics 19b_practice.pdf	Michael Hallstone, Ph.D. hallston@hawaii.edu

Lecture 19b: Practice Problems for Lecture 19 "ANOVA"

Everything that appears in these lecture notes is fair game for the test. They are the best "study guide" I can provide. It is impossible to provide a "list" that is more comprehensive than the lecture notes above. However, here are a few additional practice exercises or practice concepts.

Practice problems with step 4: setting the critical rejection region

There are practice poblems with the "seven steps" below, but there are only two different step 4 critical rejection regions. To provide some practice below are six data sets and your task is do step 4 on each one. This will provide practice using the F distribution table. You will do this 6 times for the six data sets.

Recall Step 4

4. **Define the rejection regions. And draw a picture!**

$df_{numerator}$=k-1 $df_{denominator}$=T-k

k= # of samples T=total # of items in all samples (or data observations)

for k or "samples" think number of groups or categories for which you will compare means.

Practice problems: Come up with the critical rejection region for the following data sets: use α=.05 (there are six problems and answers are below)

1

group1	
n	data (x)
x1	1
x2	2

group2	
n	data (x)
x1	0
x2	0

group3	
n	data (x)
x1	0
x2	0

group4
n
x1
x2

2

group1	
n	data (x)
x1	1
x2	2
x3	3

group2	
n	data (x)
x1	0
x2	0
x3	2

group3	
n	data (x)
x1	0
x2	0
x3	0

group4
n
x1
x2
x3

3

group1	
n	data (x)
x1	1
x2	2
x3	3
x4	2

group2	
n	data (x)
x1	0
x2	0
x3	2
x4	1

group3	
n	data (x)
x1	0
x2	0
x3	0
x4	1

4

group1	
n	data (x)
x1	1
x2	2
x3	3
x4	2
x5	1

group2	
n	data (x)
x1	0
x2	0
x3	2
x4	1
x5	1

group3	
n	data (x)
x1	0
x2	0
x3	0
x4	1
x5	1

5

group1	
n	data (x)
x1	1
x2	2
x3	3
x4	2
x5	1
x6	1

group2	
n	data (x)
x1	0
x2	0
x3	2
x4	1
x5	1
x6	1

group3	
n	data (x)
x1	0
x2	0
x3	0
x4	1
x5	1
x6	1

6

group1	
n	data (x)
x1	1
x2	2
x3	3
x4	2
x5	1
x6	1
x7	1

group2	
n	data (x)
x1	0
x2	0
x3	2
x4	1
x5	1
x6	1
x7	1

group3	
n	data (x)
x1	0
x2	0
x3	0
x4	1
x5	1
x6	1
x7	1

answers

1

group1	
n	data (x)
x1	1
x2	2

group2	
n	data (x)
x1	0
x2	0

group3	
n	data (x)
x1	0
x2	0

group4	
n	data (x)
x1	0
x2	0

Answer for α=.05

$df_{numerator}$=k-1 = 4-1=3

$df_{denominator}$=T-k =8-4=4

F= 6.59

2

group1	
n	data (x)
x1	1
x2	2
x3	3

group2	
n	data (x)
x1	0
x2	0
x3	2

group3	
n	data (x)
x1	0
x2	0
x3	0

group4	
n	data (x)
x1	0
x2	0
x3	0

Answer for α=.05

$df_{numerator}$=k-1 = 4-1=3

$df_{denominator}$=T-k =12-4=8

F= 4.07

3

group1	
n	data (x)
x1	1
x2	2
x3	3
x4	2

group2	
n	data (x)
x1	0
x2	0
x3	2
x4	1

group3	
n	data (x)
x1	0
x2	0
x3	0
x4	1

Answer for α=.05

$df_{numerator}$=k-1 = 3-1=2

$df_{denominator}$=T-k =12-3=9

F= 4.26

4

group1	
n	data (x)
x1	1
x2	2
x3	3
x4	2
x5	1

group2	
n	data (x)
x1	0
x2	0
x3	2
x4	1
x5	1

group3	
n	data (x)
x1	0
x2	0
x3	0
x4	1
x5	1

Answer for α=.05

$df_{numerator}$=k-1 = 3-1=2

$df_{denominator}$=T-k =15-3=12

F= 3.89

5

group1	
n	data (x)
x1	1
x2	2
x3	3
x4	2
x5	1
x6	1

group2	
n	data (x)
x1	0
x2	0
x3	2
x4	1
x5	1
x6	1

group3	
n	data (x)
x1	0
x2	0
x3	0
x4	1
x5	1
x6	1

Answer for α=.05
$df_{numerator}$=k-1 = 3-1=2
$df_{denominator}$=T-k =18-3=15
F= 3.68

6

group1	
n	data (x)
x1	1
x2	2
x3	3
x4	2
x5	1
x6	1
x7	1

group2	
n	data (x)
x1	0
x2	0
x3	2
x4	1
x5	1
x6	1
x7	1

group3	
n	data (x)
x1	0
x2	0
x3	0
x4	1
x5	1
x6	1
x7	1

Answer for α=.05
$df_{numerator}$=k-1 = 3-1=2
$df_{denominator}$=T-k =21-3=18
F= 3.55

ANOVA practice problems: do the seven steps to ANOVA for the following

1: prisoner groups and age α =.05

A prison administrator has to design health care systems for inmates at three different prison sites. As part of this process she is trying to figure out if the mean age is different at each of the three prison sites. She takes random samples of mean age (in years) from the three groups of prisoners and comes up with the following data:

Test the theory or hypothesis that the mean age of the three populations of inmates in the populations that correspond to the three prison sites are equal. Conversely do an ANOVA test to prove that the mean age of inmates of the three populations that correspond to the three different prison sites are not all equal.

Data

group 1	
n	data (x)
x1	22
x2	23
x3	24
x4	33
x5	34
x6	55
sum	**191**
mean	**31.83333333**

group 2	
n	data (x)
x1	34
x2	35
x3	44
x4	33
x5	55
x6	49
sum	**250**
mean	**41.66666667**

group 3	
n	data (x)
x1	36
x2	34
x3	35
x4	32
x5	33
x6	38
sum	**208**
mean	**34.66666667**

grand mean= 36.1

Answer:

1. **State the null and alternative hypothesis (H_0 and H_1).**

H_0: $\mu_1 = \mu_2 = \mu_3$

H_1: not all population means are equal

2. **State level of significance or α "alpha."**

For this example we'll use alpha =.05

3. **Determine the test distribution to use – ANOVA tests use F distribution.**

5. **Define the rejection regions. And draw a picture!**

$df_{numerator}$=k-1 $df_{denominator}$=T-k

k= # of samples T=total # of items in all samples (or data observations)

for k or "samples" think number of groups or categories for which you will compare means. We have three groups and 18 total data observations

in this case $df_{numerator}$=3-1 =2 & $df_{denominator}$=18-3=15

$F_{(2,15,\alpha=.05)}$=3.68 (draw it out)

5. **State the decision rule.**
Reject the null if the TR >3.68, otherwise FTR.

6. **Perform necessary calculations on data and compute TR value.**

$$TR = \frac{\hat{\sigma}^2 between}{\hat{\sigma}^2 within}$$

$$\sigma^2_{between} = \frac{n_1(\text{x-bar}_1\text{-grand mean})^2 + n_2(\text{x-bar}_2\text{-grand mean})^2 + n_3(\text{x-bar}_3\text{-grand mean})^2}{k-1}$$

$$\sigma^2_{between} = \frac{109.23 + 185.93 + 12.33}{2} = \frac{307.5}{2} = 153.74$$

$$\sigma^2_{within} = \frac{\Sigma d^2_1 + \Sigma d^2_2 + \Sigma d^2_3 + \Sigma d^2_k}{T - k}$$

Σd^2_1= sum of squared differences for first sample: $\sum (x - \bar{x})^2$

Σd^2_2= sum of squared differences for second sample: $\sum (x - \bar{x})^2$ etc. etc etc.

T= total # of all items in the all samples ($n_1 + n_2 + n_3 + \ldots + n_k$)

k= number of samples (think "groups" or "categories")

Computing w/in group variance

group1

n	data (x)	mean	x - mean	(x-mean)2
x1	22	31.83333333	-9.8	96.7
x2	23	31.83333333	-8.8	78.0
x3	24	31.83333333	-7.8	61.4
x4	33	31.83333333	1.2	1.4
x5	34	31.83333333	2.2	4.7
x6	55	31.83333333	23.2	536.7
sum	191			778.8
mean	31.8			

group2

n	data (x)	mean	x - mean	(x-mean)2
x1	34	41.7	-7.7	58.8
x2	35	41.7	-6.7	44.4
x3	44	41.7	2.3	5.4
x4	33	41.7	-8.7	75.1
x5	55	41.7	13.3	177.8
x6	49	41.7	7.3	53.8
sum	250			415.3
mean	41.7			

group3

n	data (x)	mean	x - mean	(x-mean)2
x1	36	34.66666667	1.3	1.8
x2	34	34.66666667	-0.7	0.4
x3	35	34.66666667	0.3	0.1
x4	32	34.66666667	-2.7	7.1
x5	33	34.66666667	-1.7	2.8
x6	38	34.66666667	3.3	11.1
sum	208			23.3
mean	34.7			

$$\sigma^2_{within} = \frac{778.8333333 + 415.3333333 + 23.33333333}{15} = \frac{1217.5}{15} = 81.16666667$$

TR=

$$\frac{153.74}{81.16666667} = 1.89$$

SPSS output

ANOVA

age

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	307.444	2	153.722	1.894	.185
Within Groups	1217.500	15	81.167		
Total	1524.944	17			

7. **CompareTR value with the decision rule and make a statistical decision. (Write out decision in English! -- my addition)**

Comparing our TR value to the critical value stated in step #5, we see that TR< 3.68. Therefore, fail to reject the null. There is insufficient evidence to reject the theory that mean age of inmates of the three populations that correspond to the three prison sites are equal.

2: prisoner groups and mean sentence length α =.05

A prison administrator has to design health care systems for inmates at three different prison sites. As part of this process she is trying to figure out how long inmates will be at each prison; thus she wonders if the mean length of prison sentence (number years an inmate will serve) is different at each of the three prison sites. She takes random samples of mean length of prison sentence (in years) from the three groups of prisoners and comes up with the following data:

Test the theory or hypothesis that the mean length of sentence (in years) of the three populations of inmates that correspond to the three prison sites are equal. Conversely do an ANOVA test to prove that the mean length of sentence of inmates in the three populations that correspond to the three prison sites are not all equal.

Data

group1	
n	data (x)
x1	5
x2	6
x3	7
x4	5
x5	6
x6	6
sum	**35**
mean	**5.833333333**

group2	
n	data (x)
x1	5
x2	6
x3	7
x4	8
x5	7
x6	8
sum	**41**
mean	**6.833333333**

group3	
n	data (x)
x1	3
x2	3
x3	4
x4	5
x5	6
x6	4
sum	**25**
mean	**4.166666667**

grand mean= 5.6

Answer

1. State the null and alternative hypothesis (H_0 and H_1).

H_0: $\mu_1 = \mu_2 = \mu_3$

H_1: not all population means are equal

2. State level of significance or α "alpha."

For this example we'll use alpha =.05

3. Determine the test distribution to use – ANOVA tests use F distribution.

4. Define the rejection regions. And draw a picture!

$df_{numerator}$=k-1 $df_{denominator}$=T-k

k= # of samples T=total # of items in all samples (or data observations)

for k or "samples" think number of groups or categories for which you will compare means. We have three groups and 18 total data observations

in this case $df_{numerator}$=3-1 =2 & $df_{denominator}$=18-3=15

$F_{(2,15,\alpha=.05)}=3.68$ (draw it out)

5. State the decision rule.

Reject the null if the TR >3.68, otherwise FTR.

6. Perform necessary calculations on data and compute TR value.

$$TR=\frac{\hat{\sigma}^2 between}{\hat{\sigma}^2 within}$$

$$\sigma^2_{between}=\frac{n_1(\text{x-bar}_1\text{-grand mean})^2+n_2(\text{x-bar}_2\text{-grand mean})^2+n_3(\text{x-bar}_3\text{-grand mean})^2}{k-1}$$

$$\sigma^2_{between}=\frac{0.30+8.96+12.52}{2}=\frac{21.8}{2}=10.88888889$$

$$\sigma^2_{within}=\frac{\Sigma d^2_1+\Sigma d^2_2+\Sigma d^2_3+\Sigma d^2_k}{T-k}$$

Σd^2_1= sum of squared differences for first sample: $\sum(x-\bar{x})^2$

Σd^2_2= sum of squared differences for second sample: $\sum(x-\bar{x})^2$ etc. etc etc.

T= total # of all items in the all samples ($n_1+n_2+n_3+\ldots+n_k$)

k= number of samples (think "groups" or "categories")

Computing w/in group variance

group1

n	data (x)	mean	x - mean	(x-mean)2
x1	5	5.833333333	-0.8	0.7
x2	6	5.833333333	0.2	0.0
x3	7	5.833333333	1.2	1.4
x4	5	5.833333333	-0.8	0.7
x5	6	5.833333333	0.2	0.0
x6	6	5.833333333	0.2	0.0
sum	35			2.8
mean	5.8			

group2

n	data (x)	mean	x - mean	(x-mean)2
x1	5	6.8	-1.8	3.4
x2	6	6.8	-0.8	0.7
x3	7	6.8	0.2	0.0
x4	8	6.8	1.2	1.4
x5	7	6.8	0.2	0.0
x6	8	6.8	1.2	1.4
sum	41			6.8
mean	6.8			

group3

n	data (x)	mean	x - mean	(x-mean)2
x1	3	4.166666667	-1.2	1.4
x2	3	4.166666667	-1.2	1.4
x3	4	4.166666667	-0.2	0.0
x4	5	4.166666667	0.8	0.7
x5	6	4.166666667	1.8	3.4
x6	4	4.166666667	-0.2	0.0
sum	25			6.8
mean	4.2			

$$\sigma^2_{within} = \frac{2.833333333 + 6.833333333 + 6.833333333}{15} = \frac{16.5}{15} = 1.1$$

TR (or F) $\frac{10.88888889}{1.1} = 9.90$

SPSS output

ANOVA

years

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	21.778	2	10.889	9.899	.002
Within Groups	16.500	15	1.100		
Total	38.278	17			

7. CompareTR value with the decision rule and make a statistical decision. (Write out decision in English! -- my addition)

Comparing our TR value to the critical value stated in step #5, we see that TR> 3.68. Therefore, reject the null. We are at least 95% confident that the mean length of sentence in the three populations that correspond to the three prison sites are not all equal to each other. At least one of the prison sites has a different mean age than the other prison sites, but additional research is needed to define the nature of this relationship between prison site and mean sentence length.

3: prisoner groups and mean number of prior drug convictions α =.05

A prison administrator has to design drug treatment programs for inmates at three different prison sites. Because she knows that drug treatment takes "longer" for those with more chronic drug abuse histories, she is trying to figure out how many prior drug convictions prisoners have; thus she wonders if the mean number of prior drug convictions is different at each of the three prison sites. She takes random samples of mean number of prior drug convictions from the three groups of prisoners and comes up with the following data:

Test the theory or hypothesis that the mean number of prior drug convictions for the three populations of inmates that correspond to the three prison sites are equal. Conversely do an ANOVA test to prove that the mean number of prior drug convictions of inmates in the three populations that correspond to the three different prison sites are not all equal.

Data

group1	
n	data (x)
x1	2
x2	2
x3	1
x4	2
x5	1
x6	0
sum	**8**
mean	**1.333333333**

group2	
n	data (x)
x1	1
x2	2
x3	0
x4	0
x5	4
x6	1
sum	**8**
mean	**1.333333333**

group3	
n	data (x)
x1	0
x2	0
x3	0
x4	1
x5	2
x6	0
sum	**3**
mean	**0.5**

Answer

1. State the null and alternative hypothesis (H_0 and H_1).

H_0: $\mu_1 = \mu_2 = \mu_3$

H_1: not all population means are equal

2. State level of significance or α "alpha."

For this example we'll use alpha =.05

3. Determine the test distribution to use – ANOVA tests use F distribution.

4. Define the rejection regions. And draw a picture!

$df_{numerator}$=k-1 $df_{denominator}$=T-k

k= # of samples T=total # of items in all samples (or data observations)

for k or "samples" think number of groups or categories for which you will compare means. We have three groups and 18 total data observations

in this case $df_{numerator}$=3-1 =2 & $df_{denominator}$=18-3=15

$F_{(2,15,\ \alpha=.05)}$=3.68 (draw it out)

5. State the decision rule.

Reject the null if the TR >3.68, otherwise FTR.

6. Perform necessary calculations on data and compute TR value.

$$TR = \frac{\hat{\sigma}^2 between}{\hat{\sigma}^2 within}$$

$$\sigma^2_{between} = \frac{n_1(\text{x-bar}_1\text{-grand mean})^2 + n_2(\text{x-bar}_2\text{-grand mean})^2 + n_3(\text{x-bar}_3\text{-grand mean})^2}{k-1}$$

$$\sigma^2_{between} = \frac{0.46 + 0.46 + 1.85}{2} = \frac{2.8}{2} = 1.388888889$$

$$\sigma^2_{within} = \frac{\Sigma d^2_1 + \Sigma d^2_2 + \Sigma d^2_3 + \Sigma d^2_k}{T - k}$$

Σd^2_1= sum of squared differences for first sample: $\sum (x - \bar{x})^2$

Σd^2_2= sum of squared differences for second sample: $\sum (x - \bar{x})^2$ etc. etc etc.

T= total # of all items in the all samples ($n_1 + n_2 + n_3 + \ldots + n_k$)

k= number of samples (think "groups" or "categories")

Computing w/in group variance

group1

n	data (x)	mean	x - mean	(x-mean)2
x1	2	1.333333333	0.7	0.4
x2	2	1.333333333	0.7	0.4
x3	1	1.333333333	-0.3	0.1
x4	2	1.333333333	0.7	0.4
x5	1	1.333333333	-0.3	0.1
x6	0	1.333333333	-1.3	1.8
sum	8			3.3
mean	1.3			

group2

n	data (x)	mean	x - mean	(x-mean)2
x1	1	1.3	-0.3	0.1
x2	2	1.3	0.7	0.4
x3	0	1.3	-1.3	1.8
x4	0	1.3	-1.3	1.8
x5	4	1.3	2.7	7.1
x6	1	1.3	-0.3	0.1
sum	8			11.3
mean	1.3			

group3

n	data (x)	mean	x - mean	(x-mean)2
x1	0	0.5	-0.5	0.3
x2	0	0.5	-0.5	0.3
x3	0	0.5	-0.5	0.3
x4	1	0.5	0.5	0.3
x5	2	0.5	1.5	2.3
x6	0	0.5	-0.5	0.3
sum	3			3.5
mean	0.5			

$$\sigma^2_{within} = \frac{3.333333333 + 11.33333333 + 3.5}{15} = \frac{18.16666667}{15} = 1.211111111$$

TR (or F) $= \frac{1.388888889}{1.211111111} = 1.15$

SPSS output

ANOVA

prior drug convictions

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2.778	2	1.389	1.147	.344
Within Groups	18.167	15	1.211		
Total	20.944	17			

7. CompareTR value with the decision rule and make a statistical decision. (Write out decision in English! -- my addition)

Comparing our TR value to the critical value stated in step #5, we see that TR< 3.68. Therefore, fail to reject the null. There is insufficient evidence to reject the theory that the mean number of prior drug convictions are equal in the three populations that correspond to the three different prison sites.

4: prisons sites and mean number of sick days α =.05

A prison administrator oversees three different prison sites around the state. It seems as if the guards at one prison are calling in sick more often than usual. Thus she wonders if the mean number of sickdays is different at each of the three prison sites. She takes random samples of mean number of sick days from the three groups of prisoners and comes up with the following data:

Test the theory or hypothesis that the mean number of sick days in the three populations that correspond to the three prison sites are equal. Conversely do an ANOVA test to prove that the mean sick days in the three populations that correspond to the three prison sites are not all equal.

Data

group1			**group2**			**group3**	
n	data (x)		n	data (x)		n	data (x)
x1	4		x1	4		x1	0
x2	5		x2	2		x2	0
x3	6		x3	1		x3	0
x4	3		x4	0		x4	1
x5	1		x5	1		x5	2
x6	0		x6	1		x6	0
sum	**19**		**sum**	**9**		**sum**	**3**
mean	**3.166666667**		**mean**	**1.5**		**mean**	**0.5**

grand mean= 1.7

Answer

1. State the null and alternative hypothesis (H_0 and H_1).

H_0: $\mu_1 = \mu_2 = \mu_3$

H_1: not all population means are equal

2. State level of significance or α "alpha."

For this example we'll use alpha =.05

3. Determine the test distribution to use – ANOVA tests use F distribution.

4. Define the rejection regions. And draw a picture!

$df_{numerator}=k-1$ $df_{denominator}=T-k$

k= # of samples T=total # of items in all samples (or data observations)

for k or "samples" think number of groups or categories for which you will compare means. We have three groups and 18 total data observations

in this case $df_{numerator}=3-1=2$ & $df_{denominator}=18-3=15$

$F_{(2,15,\alpha=.05)}=3.68$ (draw it out)

5. State the decision rule.

Reject the null if the TR >3.68, otherwise FTR.

6. Perform necessary calculations on data and compute TR value.

$$TR = \frac{\hat{\sigma}^2 between}{\hat{\sigma}^2 within}$$

$$\sigma^2_{between} = \frac{n_1(\text{x-bar}_1\text{-grand mean})^2 + n_2(\text{x-bar}_2\text{-grand mean})^2 + n_3(\text{x-bar}_3\text{-grand mean})^2}{k-1}$$

$$\sigma^2_{between} = \frac{12.52 + 0.30 + 8.96}{2} = \frac{21.8}{2} = 10.88888889$$

Computing w/in group variance

group1

n	data (x)	mean	x - mean	(x-mean)2
x1	4	3.166666667	0.8	0.7
x2	5	3.166666667	1.8	3.4
x3	6	3.166666667	2.8	8.0
x4	3	3.166666667	-0.2	0.0
x5	1	3.166666667	-2.2	4.7
x6	0	3.166666667	-3.2	10.0
sum	19			26.8
mean	3.2			

group2

n	data (x)	mean	x - mean	(x-mean)2
x1	4	1.5	2.5	6.3
x2	2	1.5	0.5	0.3
x3	1	1.5	-0.5	0.3
x4	0	1.5	-1.5	2.3
x5	1	1.5	-0.5	0.3
x6	1	1.5	-0.5	0.3
sum	9			9.5
mean	1.5			

group3

n	data (x)	mean	x - mean	(x-mean)2
x1	0	0.5	-0.5	0.3
x2	0	0.5	-0.5	0.3
x3	0	0.5	-0.5	0.3
x4	1	0.5	0.5	0.3
x5	2	0.5	1.5	2.3
x6	0	0.5	-0.5	0.3
sum	3			3.5
mean	0.5			

$$\sigma^2_{within} = \frac{\Sigma d^2_1 + \Sigma d^2_2 + \Sigma d^2_3 + \Sigma d^2_k}{T - k}$$

Σd^2_1= sum of squared differences for first sample: $\sum (x - \bar{x})^2$

Σd^2_2= sum of squared differences for second sample: $\sum (x - \bar{x})^2$ etc. etc etc.

T= total # of all items in the all samples ($n_1 + n_2 + n_3 + \ldots + n_k$)

k= number of samples (think "groups" or "categories")

$$\sigma^2_{within} = \frac{26.83333333 + 9.5 + 3.5}{15} = \frac{39.83333333}{15} = 2.655555556$$

$$\text{TR (or F)} \quad \frac{10.88888889}{2.655555556} = 4.10$$

SPSS output

ANOVA

days absent

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	21.778	2	10.889	4.100	.038
Within Groups	39.833	15	2.656		
Total	61.611	17			

7. CompareTR value with the decision rule and make a statistical decision. (Write out decision in English! -- my addition)

Comparing our TR value to the critical value stated in step #5, we see that TR> 3.68. Therefore, reject the null. We are at least 95% confident that the mean number of sick days in the three populations that correspond to the three prison sites are not all equal to each other. At least one of the prison sites has a different mean number of sick days than the other prison sites, but additional research is needed to define the nature of this relationship between prison site and mean number of sick days.

5: prisons sites and mean number of injuries α =.05

NOTE: In all the problems above, the total number of observations was 18. It has changed for this problem; the total number of observations is now 15 (3 groups and 5 in each group= 3x5=15) this will change your critical rejection value for step 4

A prison administrator oversees three different prison sites around the state. It seems as if the guards at one prison are getting injured more often than usual. Thus she wonders if the mean number of injuries is different at each of the three prison sites. She takes random samples of mean number of injuries from the three groups of prisoners and comes up with the following data:
Test the theory or hypothesis that the mean number of injuries in the three populations that correspond to the three prison sites are equal. Conversely do an ANOVA test to prove that the mean number of injuries in the three populations that correspond to the three prison sites are not all equal.

Data

group1	
n	data (x)
x1	0
x2	1
x3	1
x4	2
x5	1
sum	**5**
mean	**1**

group2	
n	data (x)
x1	3
x2	2
x3	3
x4	2
x5	3
sum	**13**
mean	**2.6**

group3	
n	data (x)
x1	0
x2	0
x3	0
x4	1
x5	2
sum	**3**
mean	**0.6**

grand mean= 1.4

Answer

1. State the null and alternative hypothesis (H_0 and H_1).

H_0: $\mu_1 = \mu_2 = \mu_3$
H_1: not all population means are equal

2. State level of significance or α "alpha."

For this example we'll use alpha =.05

3. Determine the test distribution to use – ANOVA tests use F distribution.

4. Define the rejection regions. And draw a picture!

$df_{numerator}$=k-1 $df_{denominator}$=T-k
k= # of samples T=total # of items in all samples (or data observations)
for k or "samples" think number of groups or categories for which you will compare means. We have three groups and 15 total data observations

in this case $df_{numerator}$=3-1 =2 & $df_{denominator}$=15-3=12
$F_{(2,12,\ \alpha=.05)}=3.89$ (draw it out)

5. State the decision rule.

Reject the null if the TR >3.89, otherwise FTR.

6. Perform necessary calculations on data and compute TR value.

$$TR = \frac{\hat{\sigma}^2 between}{\hat{\sigma}^2 within}$$

$$\sigma^2_{between} = \frac{n_1(\text{x-bar}_1\text{-grand mean})^2 + n_2(\text{x-bar}_2\text{-grand mean})^2 + n_3(\text{x-bar}_3\text{-grand mean})^2}{k-1}$$

$$\sigma^2_{between} = \frac{0.80 + 7.20 + 3.20}{2} = \frac{11.2}{2} = 5.6$$

Computing w/in group variance

group1

n	data (x)	mean	x - mean	(x-mean)2
x1	0	1	-1.0	1.0
x2	1	1	0.0	0.0
x3	1	1	0.0	0.0
x4	2	1	1.0	1.0
x5	1	1	0.0	0.0
sum	5			2.0
mean	1.0			

group2

n	data (x)	mean	x - mean	(x-mean)2
x1	3	2.6	0.4	0.2
x2	2	2.6	-0.6	0.4
x3	3	2.6	0.4	0.2
x4	2	2.6	-0.6	0.4
x5	3	2.6	0.4	0.2
sum	13			1.2
mean	2.6			

group3

n	data (x)	mean	x - mean	(x-mean)2
x1	0	0.6	-0.6	0.4
x2	0	0.6	-0.6	0.4
x3	0	0.6	-0.6	0.4
x4	1	0.6	0.4	0.2
x5	2	0.6	1.4	2.0
sum	3			3.2
mean	0.6			

$$\sigma^2_{within} = \frac{2 \quad + \quad 1.2 \quad + \quad 3.2}{12} = \frac{6.4}{12} = 0.533333333$$

$$\text{TR (or F)} \quad \frac{5.6}{0.533333333} = 10.50$$

SPSS output

ANOVA

injuries

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	11.200	2	5.600	10.500	.002
Within Groups	6.400	12	.533		
Total	17.600	14			

7. CompareTR value with the decision rule and make a statistical decision. (Write out decision in English! -- my addition)

Comparing our TR value to the critical value stated in step #5, we see that TR> 3.89. Therefore, reject the null. We are at least 95% confident that mean number of injuries in the three populations that correspond to the three prison sites are not all equal. At least one of the prison sites has a different mean number of injuries than the other prison sites, but additional research is needed to define the nature of this relationship between prison site and mean number of injuries.

6: prisons sites and mean number of overtime shifts α =.05

again the total number of observations = 15 not 18 in this problem

A prison administrator oversees three different prison sites around the state. It seems as if the guards at one prison are getting more overtime shifts than usual. Since overtime will "kill" an administrator's budget, she wonders if the mean number of overtime shifts is different at each of the three prison sites. She takes random samples of mean number of overtime shifts from the three groups of prisoners and comes up with the following data:

Test the theory or hypothesis that the mean number of overtime shifts in the three populations that correspond to the three prison sites are equal. Conversely do an ANOVA test to prove that the mean number of overtime shifts in the three populations that correspond to the three prison sites are not all equal.

Data

group1	
n	Data (x)
x1	1
x2	2
x3	3
x4	2
x5	1
sum	**9**
mean	**1.8**

group2	
n	data (x)
x1	0
x2	0
x3	2
x4	1
x5	1
sum	**4**
mean	**0.8**

group3	
n	data (x)
x1	0
x2	0
x3	0
x4	1
x5	1
sum	**2**
mean	**0.4**

grand mean= 1.0

Answer

1. State the null and alternative hypothesis (H_0 and H_1).

H_0: $\mu_1 = \mu_2 = \mu_3$
H_1: not all population means are equal

2. State level of significance or α "alpha."

For this example we'll use alpha =.05

3. Determine the test distribution to use – ANOVA tests use F distribution.

4. Define the rejection regions. And draw a picture!

$df_{numerator}$=k-1 $df_{denominator}$=T-k

k= # of samples T=total # of items in all samples (or data observations)

for k or "samples" think number of groups or categories for which you will compare means. We have three groups and 15 total data observations

in this case $df_{numerator}$=3-1 =2 & $df_{denominator}$=15-3=12
$F_{(2,12, \alpha=.05)}$=3.89 (draw it out)

5. State the decision rule.

Reject the null if the TR >3.89, otherwise FTR.

6. Perform necessary calculations on data and compute TR value.

$$TR = \frac{\hat{\sigma}^2 between}{\hat{\sigma}^2 within}$$

$$\sigma^2_{between} = \frac{n_1(\text{x-bar}_1\text{-grand mean})^2 + n_2(\text{x-bar}_2\text{-grand mean})^2 + n_3(\text{x-bar}_3\text{-grand mean})^2}{k-1}$$

$$\sigma^2_{between} = \frac{3.20 + 0.20 + 1.80}{2} = \frac{5.2}{2} = 2.6$$

Computing w/in group variance

group1

n	data (x)	mean	x - mean	(x-mean)2
x1	1	1.8	-0.8	0.6
x2	2	1.8	0.2	0.0
x3	3	1.8	1.2	1.4
x4	2	1.8	0.2	0.0
x5	1	1.8	-0.8	0.6
sum	9			2.8
mean	1.8			

group2

n	data (x)	mean	x - mean	(x-mean)2
x1	0	0.8	-0.8	0.6
x2	0	0.8	-0.8	0.6
x3	2	0.8	1.2	1.4
x4	1	0.8	0.2	0.0
x5	1	0.8	0.2	0.0
sum	4			2.8
mean	0.8			

group3

n	data (x)	mean	x - mean	(x-mean)2
x1	0	0.4	-0.4	0.2
x2	0	0.4	-0.4	0.2
x3	0	0.4	-0.4	0.2
x4	1	0.4	0.6	0.4
x5	1	0.4	0.6	0.4
sum	2			1.2
mean	0.4			

$$\sigma^2_{within} = \frac{2.8 + 2.8 + 1.2}{12} = \frac{6.8}{12} = 0.566666667$$

$$\text{TR (or F)} \quad \frac{2.6}{0.566666667} = 4.59$$

SPSS output

ANOVA

\# of overtime shifts

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	5.200	2	2.600	4.588	.033
Within Groups	6.800	12	.567		
Total	12.000	14			

7. CompareTR value with the decision rule and make a statistical decision. (Write out decision in English! -- my addition)

Comparing our TR value to the critical value stated in step #5, we see that TR> 3.89. Therefore, reject the null. We are at least 95% confident that the mean number of overtime shifts in the three populations that correspond to the three prison sites are not all equal. At least one of the prison sites has a different mean number of overtime shifts than the other prison sites, but additional research is needed to define the nature of this relationship between prison site and mean number of overtime shifts.