STA 6127: Solutions of Exercises for ANOVA (Chap. 12)

1. **a)** Let $\mu_1, \mu_2, ..., \mu_{12}$ be the population mean number of true friends of individuals belonging to the 12 astrological signs. The Null hypothesis is given by $H_0 : \mu_1 = \mu_2 = ... = \mu_{12}$ and the alternative hypothesis is H_a: at least two of the population means are different.
b) The value of the F statistic is pretty small, keeping in mind that the mean of the F distribution is about 1. So, there seems to be only weak evidence against the Null hypothesis.
c) Since the P-value is large, there is not much evidence against the null. If the null were true, the probability would be 0.82 of getting an F statistic of .61 or larger.

4. **a)**(i) The Null hypothesis is given by $H_0 : \mu_1 = \mu_2 = \mu_3$ and the alternative hypothesis is H_a: at least two population means differ. Here μ_i is the population mean number of good friends of individuals who belong to the ith class of the variable measuring the frequency of visits to a bar or tavern, i =1, 2, 3.
(ii) The test statistic is F = 3.03.
(iii) The P value is 0.049, giving fairly strong evidence against the null hypothesis.
b) The standard deviations are much larger than the means, which suggests the response distributions may be very highly skewed. The test is robust to non-normality, but with extremely skewed distributions the mean is not the must useful measure of center. It may be better to compare the medians.

6. To do this enter the data as two columns viz. "scores" and "Groups". The first column contain the scores of the students i.e 4, 6, ... , 5. The second column contains indicators of the groups these scores belong to. So the first three elements of this column are 1, the next two are 2 and the last three are 3. Now go to "Analyze", then "Compare Means", then "One Way ANOVA". The dependent variable would be "Scores" and the independent one would be "Groups". Then click on "Options" and then on "Descriptives". Click O.K. You should get the result shown in the text.

7. You can do this either by hand (or calculator) or using software.
a) The sample means are respectively : 2, 3 and 13.
b) The within-group sum of squares is 12. The variance estimate is 12/3 = 4.
c) The between-group sum of squares is 148. The variance estimate is 148/2 = 74.
d) The Null hypothesis is $H_0 : \mu_1 = \mu_2 = \mu_3$. The Alternative hypothesis is H_a: at least two population means differ. Here, μ_i denotes the mean repair cost of the ith bumper. The test statistic is F = 74/4 = 18.5 with $df_1 = 2$, $df_2 = 3$. From the F table, the P-value lies between 0.01 and 0.05. There is fairly strong evidence against then null, and it seems that there do exist differences between the mean repair costs of at least one pair of bumpers.
e) The required ANOVA table is given by:

Table 1:

Source	Sum of Squares	df	Mean Square	F	P
Between	148	2	74	18.5	< 0.05
Within	12	3	4		
Total	160	5			

8. **a)** The required 95% confidence interval is $(2.0 - 3.0) \pm 3.182\sqrt{4.0(1/2 + 1/2)} = -1.0 \pm 6.4 =$ (-7.4, 5.4). Since this C.I. contains 0, there is no significant difference between the mean repair costs of bumpers A and B.

b) Here $\hat{\sigma}$ is 2, $\bar{y_1} = 2$, $\bar{y_2} = 3$ and $\bar{y_3} = 13$. To achieve overall confidence level of at least 0.95, the Bonferroni method uses error probability for each interval of $0.05/3 = 0.0167$. Thus, we use a t-score with two tail probability of 0.0167, or single tail probability 0.0083. From the table, with df=3, we have $t_{0.0083} \simeq 4.7$. So, the 95% Bonferroni simultaneous CI for:
(i) $\mu_1 - \mu_2$ is $(2-3) \pm 4.7(2)\sqrt{1/2+1/2} =$ (-10.4, 8.4).
(ii) $\mu_1 - \mu_3$ is $(2-13) \pm 4.7(2)\sqrt{1/2+1/2} =$ (-20.4, -1.6).
(iii) $\mu_2 - \mu_3$ is $(3-13) \pm 4.7(2)\sqrt{1/2+1/2} =$ (-0.6, -19.4).
We can interpret the C.I.s as follows: For (ii) we infer that the mean repair cost of bumper 1 is less than that of bumper 3 by at least 160 dollars and at most 2040 dollars. The other two C.I.s can be interpreted similarly. The 95% confidence applies to the entire set of three C.I.s rather than to each one. Since the last two C.I.s do not contain 0 (and have negative values), we conclude that there is significant evidence that the mean repair cost of bumper 3 is more than that of both bumper 1 and 2. There is no evidence though that the mean repair costs of bumpers 1 and 2 differ.

9. **a)** Since the standard deviations are the same, the within-groups variability would remain the same. The between-groups variability would decrease, and there is less evidence of differences among the means, so the F test statistic would be smaller. (So, the P-value would be larger.)
b) Since the standard deviations have decreased, the within-groups variability would decrease. So, the F-statistic value would increase (and thus the P-value would decrease).
c) With larger sample sizes, effects of a particular size are more highly significant (with any significance test). The F statistic would be larger.
d) Larger in (a), smaller in (b) and (c).

13. **a)** Since the P-value is small, there is strong evidence that there exist differences between the mean number of friends (of individuals belonging to the three groups). The F-test is just a global test of independence between the response and explanatory variables. The conclusion of this test does not specify which means are different or how different are they. So, it is not possible to know the exact nature of dependence between the explanatory variable (nature of happiness) and response variable (number of friends) just by using the F test.
b) This implies that we are 95% confident that the population mean number of friends of "very happy" individuals is between 0.7 and 5.3 higher than that of "pretty happy" individuals.
c) The Bonferroni intervals are generally wider than the ordinary C.I. because the multiple comparison approach uses a higher confidence level for each separate interval to ensure achieving the simultaneous confidence level.
d) Since only the first interval does not contain 0, we conclude that only the very happy and pretty happy groups are significantly different in the population mean number of friends.

20. **a)** The required regression model is given by : $E(Y) = \alpha + \beta_1 Z_1 + \beta_2 Z_2$, where $Z_1 = 1$ for bumper A and 0 otherwise while $Z_2 = 1$ for bumper B and 0 otherwise.
b) In ANOVA, H_0 is $\mu_1 = \mu_2 = \mu_3$. In regression, it is $\beta_1 = \beta_2 = 0$.
c) The prediction equation $\hat{y} = 13 - 11z_1 - 10z_2$ would give the three sample means of 2, 3, and 13 (when plugging in the dummy variable values).

23. **a)** Let's use the 0.05 significance level to make decisions. (i) Since the p-value corresponding to Gender is greater than 0.05, we conclude that there is no significant evidence of a Gender effect.
(ii)Since the p-value corresponding to Race is 0.000, we conclude that there is significant evidence of a Race effect.
b) Controlling for race, the sample mean number of hours of TV watched by males and females are very similar (2.79 and 2.71 for whites, 4.16 and 4.13 for blacks). Thus, the sample does not show any strong evidence of a gender effect.
Controlling for gender, there is a large difference between the sample mean number of hours of TV watched by whites and blacks (2.79 and 4.16 for males, 2.71 and 4.13 for females). So, the sample does seem to have strong evidence of a race effect.

25. **a)** Here, the response variable is "hourly wage" and the two factors are "Type of job" and "gender".
b) We form the following 2×2 table between job type and gender:

Table 2:

		Type of job		
		White collar	Blue collar	Service
Gender	Male	22	14	11
	Female	15	10	8

c) The difference between males and females is 7 dollars for white collar jobs and 4 dollars for blue collar jobs. There is interaction between the two factors because the difference in the hourly wage between each pair of categories of "type of job" varies between males and females and the difference between males and females varies for the different job types. For example, the difference between males and females in hourly wage is 7 for white collar jobs, 4 for blue collar jobs and 3 for service jobs.

28. The estimated means for Females (s = 1) are 3.79, 3.96 and 4.50 and those for males(s = 0) are 3.87, 4.04 and 4.58. (For example, for female Democrats, $\hat{y} = 4.58 - .71(1) - .54(0) - .08(1) = 3.79$). The difference between the estimated means for females and males is -0.08 for each party. Moreover, it can be shown easily that the difference between the estimated means of each pair of parties is the same for males and females, thus indicating a lack of interaction.

44. **a)** The ANOVA test can be easily done using "Prices" as the dependent variable and "bathrooms" as the independent or explanatory variable. The F test statistic value is 20.3 with $\text{df}_1 = 3$ and $\text{df}_2 = 96$. The P-value is 0.000. Thus there is very strong evidence against the null hypothesis of equality of mean selling price of homes with 1, 2 or 3 bathrooms. At the usual significance levels such as 0.05, we conclude that the mean price of homes depends on the number of bathrooms they have.
c) The F test of independence of selling price and number of bathrooms in (a) tests the null hypothesis $H_0 : \mu_1 = \mu_2 = \mu_3$ where μ_i denotes the mean prices of houses having i bathrooms (i = 1, 2, 3). It can detect any type of difference among the true means. By contrast, the t test for the slope β in a regression model tests equality of means against an alternative by which the means change in a linear manner. It will tend to be a more powerful test if that in fact is the case. It would be less appropriate than the ANOVA F test if actually the means differ but do not tend to increase or tend to decrease continuously as the number of bathrooms increases. For example, if the population mean selling price were $100,000 with 1 bedroom, $200,000 with 2 bedrooms, but $100,000 with 3 bedrooms, the straight line regression model would be inappropriate, and tests designed to detect linear effects would do poorly.

48. Here the response variable is the "number of firefights" and the explanatory variable the group. The hypothesis tested was $H_0 : \mu_1 = \mu_2 = \mu_3$, where μ_1 refers to the mean number of firefights involving soldiers deployed in Afghanistan, μ_2 refers to the mean for soldiers in Iraq and, μ_3 refers to the mean for Marines in Iraq.
The P-value is 0.001. It can be simply explained as follows: If the true mean number of firefights in the three groups were the same, it would have been extremely unlikely to observe results such as were observed by this study.

55. Suppose the sample means for A, B and C are respectively 15, 21 and 27. If the margin of error (the $\pm$ part of a C.I.) is 10 then it is easy to check that the C.I. of the difference of the means of A and B and of B and C contains 0 but that of A and C does not contain 0. So, A is not significantly different from B and B is not significantly different from C.

58. The four examples are respectively :

Table 3:

a.	20	20	b.	10	20	c.	10	20	d.	20	20
	30	30		30	40		30	60		20	20

61. We first fit the model of interaction i.e we test the null hypothesis of no interaction between religious beliefs and church attendance of the students. The SPSS output is given below in Table 4. Since the P-value corresponding to the interaction term is quite large, we conclude that there is no significant evidence of interaction between religious beliefs and church attendance.

Next we proceed to fit the model with no interaction and test the null hypotheses that (i) The population mean attitude toward abortion scores of the students are identical for the two religious beliefs controlling for church attendance and (ii) The population mean attitude toward abortion scores of the students are identical across church attendance controlling for religious beliefs. These are the main effects. The SPSS output is given below in Table 5. Since the p-values corresponding to both the main effects are smaller than 0.05, we conclude that both church attendance and religious beliefs have significant influence on the attitude toward abortion scores of the students.

Looking at the output for the regression model (shown below in Table 6) corresponding to the no-interaction ANOVA gives information about the sizes of the effects. The mean response is estimated to be 2.5 lower (95% CI of -1.3 to -3.8) for those who are frequent attenders of church than for those who are infrequent attenders and it is estimated to be 1.4 lower (95% CI of -2.6 to -0.1) for those who are fundamentalist than for those who are nonfundamentalist.

66. The correct responses are a, b, c, d.

67. The correct response is c.

68. The correct responses are c, e, f.

69. The correct responses are c) and d).

Table 4:

Analysis Variable : Abortion

CHURCH = freq RELIGION = fund

N	Mean	Variance
9	1.5555556	2.2777778

CHURCH = freq RELIGION = nonfund

N	Mean	Variance
6	2.6666667	1.8666667

CHURCH = infreq RELIGION = fund

N	Mean	Variance
3	3.6666667	0.333333

CHURCH = infreq RELIGION = nonfund

N	Mean	Variance
8	5.500000	2.8571429

Model with interaction:

Source	DF	Sum of Squares	F Value	Pr > F
Model	3	68.89316239	10.48	0.0002
Error	22	48.22222222		
Corrected Total	25	117.11538462		

Source	DF	Sum of Squares	F Value	Pr > F
CHURCH	1	33.21174004	15.15	0.0008
RELIGION	1	11.77777778	5.37	0.0301
CHURCH*RELIGION	1	0.70859539	0.32	0.5754

Table 5:

Source	DF	Sum of Squares	F Value	Pr > F
Model	2	68.18456701	16.03	0.0001
Error	23	48.93081761		
Corrected Total	25	117.11538462		

Source	DF	Sum of Squares	F Value	Pr > F
CHURCH	1	36.84299191	17.32	0.0004
RELIGION	1	11.06918239	5.20	0.0321

Table 6:

Model Summary

R	R-squared	Adj. R-squared	Standard Error
0.763	0.582	0.546	1.459

ANOVA

	S.S	df	M.S	F	Sig.
Regression	68.185	2	34.092	16.025	0
Residual	48.931	23	2.127		
Total	117.115	25			

Coefficients

	Unstandardized		Standardized			95% C.I for B	
	B	Std.Error	Beta	t	Sig	L.B	U.B
Const.	5.377	0.470		11.445	0.000	4.405	6.349
Religion	-1.384	0.607	-0.325	-2.281	0.032	-2.638	-0.129
Church Attd.	-2.547	0.612	-0.593	-4.162	0.000	-3.813	-1.281