Chapter 2 Homework

Jeremy Morris

February 6, 2006

year(i)	observations (j)										$T_{i\cdot}$	$s^2_{i\cdot}$
1968	5.0	4.9	5.9	5.0	7.2	7.0	4.9	6.4	6.3	7.3	59.9	0.9788
1969	5.4	3.3	4.5	5.5	5.0	7.6	7.6	6.4	5.3	7.1	57.7	1.9557
1970	4.3	5.1	4.0	6.0	4.4	4.6	4.8	6.4	5.2	4.3	49.1	0.6077
1971	5.2	7.8	3.0	5.0	5.8	4.0	4.5	4.4	4.4	3.9	48.0	1.7000
1972	5.0	5.8	6.6	4.1	3.8	4.5	3.8	4.5	5.3	3.8	47.2	0.9040

Table 1: Highway Safety Data

2. *A government committee on highway safety was interested in whether roads were getting safer. Ten states were randomly selected out of the 48 contiguous states and, for each of the years 1968 through 1972, the fatality rate (= deaths per 100 million vehicle miles) was calculated. This data is shown in Table 1*

   (a) *Use each of the three tests for homogeneity on this data. For consistency in calculating the* $\ln s^2$ *ANOVA test, use the first five and the last five observations in each state.*

Bartlett's Test for Homogeneity			
Source	DF	Chi-Square	Pr > ChiSq
Year	4	3.8867	0.4216

Levene's Test for Homogeneity					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Year	4	10.5638	2.6410	0.97	0.4329
Error	45	122.4	2.7206		

O'Brien's Test for Homogeneity					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Year	4	13.0418	3.2604	0.86	0.4953
Error	45	170.6	3.7918		

Brown and Forsythe's Test for Homogeneity					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Year	4	1.2788	0.3197	0.60	0.6627
Error	45	23.8750	0.5306		

These tests indicate that we cannot reject the hypothesis that the data is homogenous.

(b) *Test for the normality of the data collected in 1970.*

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.90363	Pr < W	0.2400
Kolmogorov-Smirnov	D	0.156109	Pr > D	>0.1500
Cramer-von Mises	W-Sq	0.069106	Pr > W-Sq	>0.2500
Anderson-Darling	A-Sq	0.435813	Pr > A-Sq	0.2409

These tests agree that the null hypothesis that the data is normally distributed cannot be rejected. The data does not need to be transformed.

(c) *Run an ANOVA on this data.*

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	14.1628	3.5407	2.88	0.0331
Error	45	55.315	1.2292		
Corrected Total	49	69.4778			

(d) *Is there a significant difference in fatality rates using* $\alpha = 0.05$*?*

Yes, there is a difference.

3. (a) *For the data listed in Table 1, what is the state to state (also called within year or error) standard deviation*

$\sigma = \sqrt{MSE} = \sqrt{1.2292} = 1.1087$

(b) *Assuming* $\beta = 0.10$*, what is the minimal difference in the annual death rate that can be detected with this data?*

$$\delta = \sigma\Delta = 2.051$$

Where Δ is determined from the table in Appendix 10 with five treatments and ten repeats.

4. *Use Duncan's test on the data in Table 1 to determine which years differ. Interpret the results. Are the results consistent with question 3? Why or why not?*

Means with the same letter are not significantly different.				
Duncan Grouping		Mean	N	Year
	A	5.9900	10	1968

Means with the same letter are not significantly different.				
Duncan Grouping		Mean	N	Year
B	A	5.7700	10	1969
B		4.9100	10	1970
B		4.8000	10	1971
B		4.7200	10	1972

These results are not inconsistent with the the answer for question 3. What we see is Duncan's test detecting a smaller difference than the minimal difference we calculated in question 3. This is because a difference smaller than $\delta = 2.051$ can be detected with probablility < 0.90.

5. *Use orthogonal polynomials to find the order of the best fitting polynomial for the data in Table 1. Give the ANOVA table in its final form.*

   The ANOVA and lack of fit tables follow :

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	14.1628	3.5407	2.88	0.0331
Error	45	55.315	1.2292		
Corrected Total	49	69.4778			

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Linear	1	12.3201	12.3201	10.02	0.0028
Quadratic	1	0.7577	0.75778	0.62	0.4365
Lack of Fit	2	1.08491	0.54245	0.44	0.6460

We conlude that there is a significant linear effect and there is not a significant lack of fit. So we say that a linear trend adequately describes the data.