

Stat 642 Spring 2022 - Solutions for Homework 5

1. (35 points) Accelerated Life Test Experiment:

- a. Let F_{ij} be the time to failure of the j th TV under temperature i . The normality of the distribution of the F_{ij} 's is evaluated using the residuals from the fitted model, $\hat{e}_{ij} = F_{ij} - \bar{F}_i$. From the normal probability plot and the stem-leaf plot, the distribution of \hat{e}_{ij} is nearly symmetric. The p-value from the Shapiro-Wilk p-value = .0896. Therefore, we conclude that the distribution of the residuals appears to be approximately normally distributed. The four temperatures have the following values for the sample means and standard deviations:

TEMP	N	FAILURE TIME MEAN	FAILURE TIME STD DEV
40	6	3955.66667	2021.46300
45	6	1925.50000	689.20149
55	6	1043.00000	344.64881
70	6	647.00000	73.42752

The p-value from the BFL test is p-value < .0001 which would imply that there is significant evidence of a difference in the four populations variances which is confirmed by examining the very large differences in the four sample standard deviations, S_i , 73.4 to 2021.5.

Brown and Forsythe's Test for Homogeneity of Variances

Source	DF	Sum of Squares	Mean Square	F Value	p-value
T	3	10612311	3537437	23.43	<.0001
Error	20	3019636	150982		

TESTS OF NORMALITY

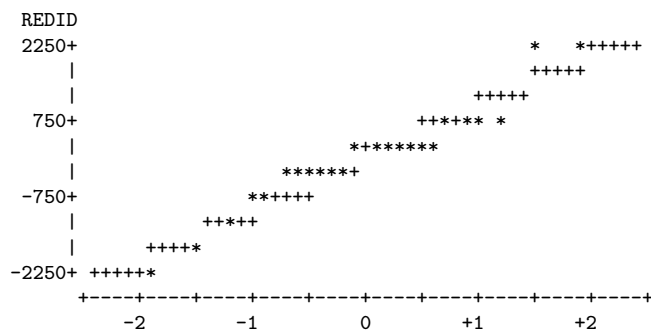
Test	Statistic	p Value
Shapiro-Wilk	W 0.92835	Pr < W 0.0896

STEM-LEAF PLOT OF RESIDUALS

Stem Leaf	#	Boxplot
2 24	2	0
1		
1		
0 5689	4	
0 0000123	7	---+---+
-0 442210	6	+-----+
-0 76	2	
-1		
-1 85	2	0
-2 0	1	0

-----+-----+-----+-----+

NORMAL REFERENCE DISTRIBUTION PLOT OF RESIDUALS



- b. Using the following R code, the estimated slope from the regression of $\log(S)$ on $\log(\bar{F}_i)$ is obtained.

```
mean = c(3955.7,1925.5,1043,647)
sd    = c(2021.46,689.20,344.65,73.43)
lm(log(sd)~log(mean))
```

```
Coefficients:
(Intercept)    log(mean)
      -6.583         1.729
```

Therefore, $\hat{\beta} = 1.73$ which yields the transformation: $y = F^{1-1.73} = F^{-.73}$. In most cases, the transformation $y = F^{-1} = 1/F$ works nearly as well and the reciprocal transformation is easier for non-mathematical consultees to understand.

- c. Using the following R code, the boxcox transformation is obtained.

```
d41 = matrix(0,24,2)
f = matrix(0,24,1)

f = c(1953, 2135, 2471, 4727, 6134, 6314, 1190, 1286, 1550, 2125, 2557, 2845,
      651, 817, 848, 1038, 1361, 1543, 511, 651, 651, 652, 688, 729)

tmp = c(rep("40",6),rep("45",6),rep("55",6),rep("70",6))
tmp = as.factor(tmp)
d41 = data.frame(f,t)
summary(lm(f~t))
anova(lm(f~tmp))

library(MASS)
boxcox(f~tmp,lambda=seq(-3, 3,.01))
boxcox(f~tmp,lambda=seq(-1.5, 0,.01))
boxcox(f~tmp,lambda=seq(-.8,-.5,.01))
```

From the R output $\theta = -.635$ with a 95% confidence interval which includes -1

- d. Using the boxcox transformation $y = F^{-.635}$ and the reciprocal transformation, $y = 1/F$, we obtain the following results:

Transform	SW p-value	Normal Reference Distribution Plot for Residuals	BFL p-value
$y = F^{-.635}$.5719	Symmetric with plotted points near a straight line	.1106
$y = 1/F$.8021	Symmetric with plotted points near a straight line	.1929

Both transformations yield data which appears to have normally distributed residuals with equal variances based on the Shapiro-Wilk test, Brown-Forsythe-Levene test, and the residual plots.

- e. Using the original data and the transformation $y = 1/F$, we obtain the following results:

Dependent Variable: F - Time to Failure

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	39183995.13	13061331.71	11.15	0.0002
Error	20	23427428.83	1171371.44		
Corrected Total	23	62611423.96			

Source	DF	Type III SS	Mean Square	F Value	Pr > F
T	3	39183995.13	13061331.71	11.15	0.0002

Dependent Variable: Y = 1/F - Reciprocal of Time to Failure

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	5.4071936E-6	1.8023979E-6	32.85	<.0001
Error	20	1.0972338E-6	5.4861689E-8		
Corrected Total	23	6.5044274E-6			

Source	DF	Type III SS	Mean Square	F Value	Pr > F
T	3	5.4071936E-6	1.8023979E-6	32.85	<.0001

Based on the two AOV's, we would reach the same conclusion, there is significant evidence that the four temperatures have a difference in their mean time to failure. However, the conclusion is somewhat stronger from the reciprocal transformation than from the original data, a p-value < .0001 vs a p-value of .0002. The

validity of the p-value from the original data is questionable because there was significant evidence that the original data did not have equal variances.

- f. Both transformations, $y = 1/F$ and $y = F^{-.635}$, generated Tukey HSD groupings of the four treatment means as follows:

$$G_1 = \{40^\circ C, 45^\circ C\}; \quad G_2 = \{55^\circ C\}; \quad G_3 = \{70^\circ C\}$$

- The original data had considerably different groupings:

$G_1 = \{40^\circ C\}; \quad G_2 = \{45^\circ C, 55^\circ C, 70^\circ C\}$ but of course, these groupings are of questionable validity because the original data did not have equal variances.

- g. Using $\alpha_{PC} = 1 - (1 - .05)^{1/3} = .01695$ and the data from the reciprocal transformation, $y = 1/F$, there is significant evidence (p-value < .0001) of a positive linear trend in the mean wrinkle resistances across the four temperatures. The quadratic and cubic trends were not significant with p-values of .3737 and .9589, respectively.
- Using the data from the Box-Cox transformation, $y = F^{-.635}$, similar p-values were obtained for the three trends: Linear p-value < .0001, Quadratic p-value = .1335, and Cubic p-value = .8437.
 - Evaluating the trends using the **untransformed data**, the p-values for the three trends were Linear p-value < .0001, Quadratic p-value = .0143, Cubic p-value = .1821, thus there was significant evidence of both a linear and quadratic trend. These conclusions are different from the conclusions that were observed in the transformed data.
 - The above illustrates the importance of checking whether the conditions for applying HSD and tests of contrasts are valid. The conclusions from the nontransformed data are considerably different from the conclusions from the transformed data.

2. (15 points)

- a. Using the following R code or using the SAS program from Handout 5,

```
d41 = matrix(0,24,2)
f = matrix(0,24,1)
f = c( 1953, 2135, 2471, 4727, 6134, 6314, 1190, 1286, 1550, 2125, 2557, 2845,
651, 817, 848, 1038, 1361, 1543, 511, 651, 651, 652, 688, 729)
f = f-3000
tmp = c(rep("40",6),rep("45",6),rep("55",6),rep("70",6))
tmp = as.factor(tmp)
d41 = data.frame(f,t)
kruskal.test(f,tmp,f~tmp)
```

Kruskal-Wallis chi-squared = 18.2785, df = 3, p-value = 0.0003853

We can thus conclude there is very significant evidence (p-value = 0.0004) of a location difference in the population of times to failure for the four temperatures.

- b. The rank based multiple comparison procedure yields the following results: The p-value from the Kruskal-Wallis test was p-value = .0004.

The mean ranks are $\bar{R}_1 = 20.33$ for $40^\circ C$, $\bar{R}_2 = 16.00$ for $45^\circ C$, $\bar{R}_3 = 9.50$ for $55^\circ C$ and $\bar{R}_4 = 4.177$ for $70^\circ C$.

Using the Miller procedure, two pairs of Temperature Failure Times are said to be different if

$$|\bar{R}_i - \bar{R}_h| > q(.05, t, \infty) \sqrt{\frac{t(n+1)}{12}} = 3.633 \sqrt{\frac{4(24+1)}{12}} = 10.49$$

with $q(.05, 3, \infty) = qtukey(.95, 4, 100000) = 3.633$ or use the value from Table VII, Studentized Range Table.

The Miller rank procedure finds

$$\begin{aligned} \bar{R}_1 - \bar{R}_2 &= 20.33 - 16.00 = 4.33 < 10.49 \Rightarrow \text{Not significant} \\ \bar{R}_1 - \bar{R}_3 &= 20.33 - 9.50 = 10.83 > 10.49 \Rightarrow \text{Significant} \\ \bar{R}_1 - \bar{R}_4 &= 20.33 - 4.17 = 16.16 > 10.49 \Rightarrow \text{Significant} \\ \bar{R}_2 - \bar{R}_3 &= 16.00 - 9.50 = 6.50 < 10.49 \Rightarrow \text{Not significant} \\ \bar{R}_2 - \bar{R}_4 &= 16.00 - 4.17 = 11.83 > 10.49 \Rightarrow \text{Significant} \\ \bar{R}_3 - \bar{R}_4 &= 9.50 - 4.17 = 5.33 < 10.49 \Rightarrow \text{Not significant} \end{aligned}$$

The following groupings of the temperatures are based on similar location parameters for the times to failure

$$G_1 = \{40^\circ C, 45^\circ C\}; \quad G_2 = \{45^\circ C, 55^\circ C\}; \quad G_3 = \{55^\circ C, 70^\circ C\}$$

- c. The groupings using the ranks are different from the groupings obtained from the untransformed data but are nearly the same as the groupings from the transformed data with the exception that the rank procedure did not separate the mean from 55°C, from the means of 45°C and 70°C.

3. (30 points) Female Moth Egg Count Problem:

- a. The three Strains of moth have the following values for the sample means and standard deviations:

STRAIN	N	Mean	Variance
USDA	15	368.000000	70554.71
FIELD	15	181.266667	40799.96
RESIST	15	90.800000	13949.17

The Poisson distribution has the relationship of $\sigma_i^2 = \mu_i$. From the sample means and variances for the three strains, the sample variances are much larger than the sample means for all three strains. Therefore, an overdispersed Poisson may be possible but the regular Poisson distribution would not be a reasonable fit to the data.

- b. The following SAS program will fit an overdispersed Poisson model to the data:

```
ods html;   ods graphics on;
option ls=80 ps=50 nocenter nodate;
title 'Spring 2014 - HW 5 - Problem 3';
data count; array Y Y1-Y15;
input STRAIN $ Y1-Y15; do over Y; N=Y; output; end;
      drop Y1-Y15;

label N = 'Number of Eggs' ;
cards;
USDA    448 906 28 277 634 48 369 137 29 522 319 242 261 566 734
FIELD   211 276 415 787 18 118 1 151 0 253 61 0 275 0 153
RESIST  0 9 143 1 26 127 161 294 0 348 0 14 21 0 218
RUN;
title "Poisson Regression on Moth Egg Data";
proc means mean std;
var N;
class STRAIN;
run;
proc genmod data=count;
class STRAIN;
model N = STRAIN/Dist=P link=log;
run;
Title "Overdispersed Poisson Regression on MOTH EGG Data";
proc genmod data=count ORDER=DATA;
class STRAIN;
model N = STRAIN/dist=P link=log scale = pearson;
contrast 'FIELD vs USDA' STRAIN -1 1 0;
contrast 'FIELD vs RESIST' STRAIN 0 1 -1;
contrast 'RESIST vs USDA' STRAIN -1 0 1;
run;
proc genmod data=count;
class STRAIN;
model N = STRAIN/dist=P link=log scale = pearson;
contrast 'S2 vs S1' STRAIN 1 -1 0;
contrast 'S2 vs S3' STRAIN 0 1 -1;
contrast 'S3 vs S1' STRAIN 1 0 -1;
run;
ods graphics off;   ods html close;
```

From the SAS output for the fit of a Poisson model to the data, the "Scaled Deviance/DF = 200.3217" which is not very close to 1.0. Therefore, the results of the Poisson analysis would not be valid.

- Using the Overdispersed Poisson model, the SAS output has the "Scaled Deviance/DF = 1.0170" which is very close to 1.0. Therefore, the results of the Overdispersed Poisson analysis would appear to be valid.
- From the SAS output for the overdispersed Poisson

Analysis Of Maximum Likelihood Parameter Estimates

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits	Wald Chi-Square	Pr > ChiSq
Intercept	1	4.5087	0.3803	3.7633, 5.2540	140.56	<.0001
STRAIN USDA	1	1.3994	0.4246	0.5672, 2.2317	10.86	0.0010
STRAIN FIELD	1	0.6913	0.4659	-0.2219, 1.6045	2.20	0.1379
STRAIN RESIST	0	0.0000	0.0000	0.0000, 0.0000	.	.
Scale	0	14.0350	0.0000	14.0350, 14.0350		

Note: The scale parameter was estimated by the square root of Pearson's Chi-Square/DOF.

Contrast Results

Contrast	Num DF	Den DF	F Value	Pr > F	Chi-Square	Pr > ChiSq	Type
FIELD vs USDA	1	42	4.93	0.0318	4.93	0.0264	LR
FIELD vs RESIST	1	42	2.33	0.1340	2.33	0.1265	LR
RESIST vs USDA	1	42	13.67	0.0006	13.67	0.0002	LR

Based on the SAS output using $\alpha_{PC} = .05/3 = .0167$, there is significant evidence of a difference in the mean egg count of the USDA and RESIST strains but not between USDA and FIELD nor between FIELD and RESIST.

Note that there is a difference in the two methods of comparing the three types of Strains: MLE of the parameters versus Contrasts. Most of the time, even though the exact p-values will differ between the two methods, the conclusions will be the same.

- c. Because there was an indication that the original data was not normally distributed, the Runs Test will be used to evaluate correlation in the data. The following R code will yield the results for the runs test:

```
N1=c(448,906, 28,277,634 ,48,369,137,29,522,319,242,261,566,734 )
N2=c(211,276,415,787, 18,118, 1,151, 0,253, 61, 0,275, 0,153 )
N3=c(0,9, 143, 1, 26,127,161,294, 0,348, 0, 14, 21 , 0,218 )
N = c(N1,N2,N3)
data3 = matrix(N,nrow=3,byrow=T)
S = matrix(0,3,15)
meansN = matrix(0,3,15)
for (i in 1:3) {
  meansN[i] = mean(data3[i,])
  S[i,] = data3[i,]-meansN[i]
}

n.neg = rep(0,3)
n.pos = rep(0,3)
for (i in 1:3) {
  n.neg[i] = length(S[i,][S[i,]<0])
  n.pos[i] = length(S[i,][S[i,]>0])
}

numb.runs = rep(1,3)
for (i in 1:3) {
  for (j in 2:15) {
    if (sign(S[i,j]) != sign(S[i,j-1])) {numb.runs[i] = numb.runs[i] + 1}
  }
}

Sruns.result = as.data.frame(cbind(numb.runs, n.pos, n.neg))
names(Sruns.result) = c("No. runs", "N+", "N-")
```

The critical values n_{Lower}, n_{Upper} are from Table A.30(a) on page 54 in HO 5:

With N equal to the number of runs, N+ the number of positive differences $y_i - \bar{y}$, and N- the number of negative differences, we obtain the following results

N	N+	N-	n_{Lower}	n_{Upper}
9	7	8	4	13
6	6	9	4	13
8	6	9	4	13

For all three strains, there is not significant evidence of correlation in the data because N fell between n_{Lower} and n_{Upper} in all three cases.

4. (20 points, 4 each) Short Answer Questions

- a. The power will be increased but also the probability of a Type I Error will be much larger than the nominal value.

(Hence the procedure would too often state that there is strong evidence that the research hypothesis is true when in fact it is false.)

- b. Regression of S_i on \bar{y}_i yields $R^2 = .92$ and p-value=.0023, which is an indication of nonequal variances. Use transformation $X_{ij} = y^{1-\hat{\beta}_1} = y_{ij}^{-.5} = 1/\sqrt{y_{ij}}$

- c. The ordering of X_{ij} values is reversed from the ordering of y_{ij} values. Apply Hsu's procedure to X_{ij} 's values with "best is smallest".

- d. The t treatment population distributions have same location-scale family with equal scale but different location parameters. Data iid within treatments and independent between treatments. Conditions are identical with standard AOV model except non-Normal distributions are allowed.

- e. Declare a data value an outlier, if

$$|\hat{e}_{ij}| \geq \hat{\sigma}_e \sqrt{1 - \frac{1}{n_i}} t_{.0005, df_E} = \sqrt{9} \sqrt{1 - \frac{1}{10}} (3.52) = 10.02$$

Data value is outlier if corresponding residual is less than -10.02 or greater than 10.02. Thus, no outliers are identified in the 50 data values.