

Model Adequacy

Design of Experiments - Montgomery
Sections 3-4 and 14-1.1

5

Model Checking Diagnostics

- Assumptions
 - 1 Model is correct
 - 2 Independent observations
 - 3 Errors normally distributed
 - 4 Constant variance

$$\begin{aligned} y_{ij} &= (\bar{y}_{..} + (\bar{y}_{i.} - \bar{y}_{..})) + (y_{ij} - \bar{y}_{i.}) \\ y_{ij} &= \hat{y}_{ij} + \hat{\epsilon}_{ij} \\ \text{observed} &= \text{predicted} + \text{residual} \end{aligned}$$

- Diagnostics use predicted and residual values

5-1

Diagnostics

- Normality
 - Histogram of residuals
 - Normal probability plot / QQ plot
 - Shapiro-Wilks/Kolmogorov-Smirnov Test
- Variance
 - Plot $\hat{\epsilon}_{ij}$ vs \hat{y}_{ij} (residual plot)
 - Bartlett's or Levene's Test
- Independence
 - Plot $\hat{\epsilon}_{ij}$ vs time/space
 - Plot $\hat{\epsilon}_{ij}$ vs variable of interest
- Outliers
 - Is it influential? With and without analysis
 - Formal tests (e.g. standardized residuals)
 - Investigate why result may occur, don't try to eliminate

5-2

Normality Assumption

- Histogram
 - Is histogram of residuals bell-shaped?
- Normal Probability Plot
 - Ordered residuals vs cumulative probability points
 - On normal probability paper
 - * $P_t = (t - .5)/N$ for $t = 1, 2, \dots, N$
 - * Plot P_t vs $\hat{\epsilon}_{(t)}$
 - Is it approximately linear?
- QQ plot
 - Similar plot but compute associated quantile values
 - Find z_t such that $P(Z \leq z_t) = P_t$
 - On regular graph paper (z_t vs $\hat{\epsilon}_{(t)}$)

5-3

Constant Variance

- Often experiments with non-constant variance
- Size of residual associated with predicted value
- Residual plot
 - Plot $\hat{\epsilon}_{ij}$ vs \hat{y}_{ij}
 - Is the range constant for different levels of \hat{y}_{ij}
- Bartlett's Test (page 81-82)
 - More formal test
 - Compares pooled with a sample variances
 - Often unnecessary
 - Very sensitive to Normality assumption
- Levene's Test
 - Compute $|y_{ij} - m_i|$ where m_i is median
 - Compare samples using F-test

5-4

```
options ls=80 ps=65;

title1 'Diagnostics Example';

data one;
  infile 'c:\saswork\data\tensile.dat';
  input percent strength time;

proc glm data=one;
  class percent;
  model strength=percent;
  means percent / hovtest=bartlett hovtest=levene;
  output out=diag p=pred r=res;

proc sort; by pred;
symbol1 v=circle i=sm50; title1 'Residual Plot';
proc gplot; plot res*pred/frame; run;

proc univariate data=diag noprint;
  var res; qqplot res / normal (L=1 mu=est sigma=est);
  histogram res / normal; run;

proc sort; by time;
symbol1 v=circle i=sm75;
title1 'Plot of residuals vs time';
proc gplot; plot res*time / vref=0 vaxis=-6 to 6 by 1;
run;

symbol1 v=circle i=sm50;
title1 'Plot of residuals vs time';
proc gplot; plot res*time / vref=0 vaxis=-6 to 6 by 1;
run;
```

5-6

Independence

- Plot of the residuals over time
 - Is there a drift or pattern as trials proceed?
- Plot residuals versus relevant variables
 - Often variables omitted from analysis
 - Experimental conditions (e.g., temp)
 - May result in inclusion of factor in next exp

5-5

Diagnostics Example

The GLM Procedure

Dependent Variable: strength

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	475.7600000	118.9400000	14.76	<.0001
Error	20	161.2000000	8.0600000		
Corrected Total	24	636.9600000			

R-Square	Coeff Var	Root MSE	strength Mean
0.746923	18.87642	2.839014	15.04000

Source	DF	Type I SS	Mean Square	F Value	Pr > F
percent	4	475.7600000	118.9400000	14.76	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
percent	4	475.7600000	118.9400000	14.76	<.0001

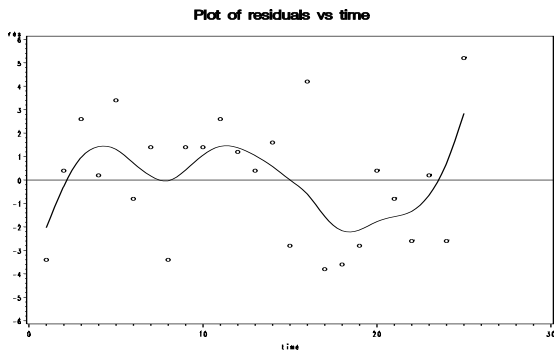
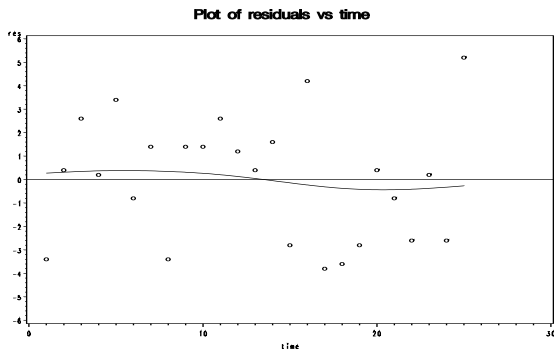
Levene's Test for Homogeneity of strength Variance ANOVA of Squared Deviations from Group Means

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
percent	4	91.6224	22.9056	0.45	0.7704
Error	20	1015.4	50.7720		

Bartlett's Test for Homogeneity of strength Variance

Source	DF	Chi-Square	Pr > ChiSq
percent	4	0.9331	0.9198

5-7



5-8

Constant Variance

- Often experiments with non-constant variance
- Does not affect F-test dramatically (balanced)
- Why concern?
 - Comparison of treatments depends on MSE
 - Incorrect intervals and comparison results
- Variance-Stabilizing Transformations
 - Common transformations
 \sqrt{x} , $\log(x)$, $1/x$, $\arcsin(\sqrt{x})$, and $1/\sqrt{x}$
 - Box-Cox transformations
 uses maximum likelihood procedures
 can approximate using relationship $\sigma_i = \theta\mu_i^\beta$
 transformation is $X^{1-\beta}$
- Dist often more “normal” after transformation

5-9

Transformations

- Consider response X with $E(X) = \mu_x$ and $\text{Var}(X) = \sigma_x^2$
- Define $Y = f(X)$; What is the mean and var of Y ?

Delta Method

Consider $f(X)$ where $f'(\mu_x) \neq 0$
 $f(X) \approx f(\mu_x) + (X - \mu_x)f'(\mu_x)$

$$E(Y) = E(f(X)) \approx E(f(\mu_x)) + E((X - \mu_x)f'(\mu_x)) = f(\mu_x)$$

$$\text{Var}(Y) \approx [f'(\mu_x)]^2 \text{Var}(X) = [f'(\mu_x)]^2 \sigma_x^2$$

- Suppose σ_x^2 depends on $\mu_x \rightarrow \sigma_x^2 = g(\mu_x)$
- Want to find $Y = f(X)$ such that $\text{Var}(Y) \approx c$
- Have shown $\text{Var}(f(X)) \approx [f'(\mu_x)]^2 \sigma_x^2$
- Want to choose f such that $[f'(\mu_x)]^2 g(\mu_x) \approx c$

Examples

$g(\mu) = \mu$	(Poisson)	$f(X) = \int \frac{1}{\sqrt{\mu}} d\mu \rightarrow f(X) = \sqrt{X}$
$g(\mu) = \mu(1 - \mu)$	(Binomial)	$f(X) = \int \frac{1}{\sqrt{\mu(1-\mu)}} d\mu \rightarrow f(X) = \text{asin}(\sqrt{X})$
$g(\mu) = \mu^{2\beta}$	(Box-Cox)	$f(X) = \int \mu^{-\beta} d\mu \rightarrow f(X) = X^{1-\beta}$
$g(\mu) = \mu^2$	(Box-Cox)	$f(X) = \int \frac{1}{\mu} d\mu \rightarrow f(X) = \log X$

5-10

Box-Cox Transformation

- Perform analysis of variance on

$$y^\lambda = \begin{cases} \frac{y^\lambda - 1}{\lambda y^{\lambda-1}} & \lambda \neq 0 \\ y \log y & \lambda = 0 \end{cases}$$

\dot{y} is the geometric mean of the observations

$$\dot{y} = \left(\prod_{i=1}^a \prod_{j=1}^{n_i} y_{ij} \right)^{1/N}$$

- y to the power of λ rescaled for direct comparison
- Find λ which minimizes SS_E
- trans1.sas does these calculations for you

5-11

```

trans.sas

options nocenter ps=65 ls=80;
title1 'Increasing Variance Example';
data one;
  infile 'c:\saswork\data\boxcox.dat'; input trt resp;

proc glm data=one; class trt;
  model resp=trt; output out=diag p=pred r=res;

title1 'Residual Plot'; symbol1 v=circle i=none;
proc gplot data=diag; plot res*pred /frame;

proc univariate data=one noprint;
  var resp; by trt; output out=two mean=mu std=sigma;

data three;
  set two; logmu = log(mu); logsig = log(sigma);

proc reg; model logsig = logmu;

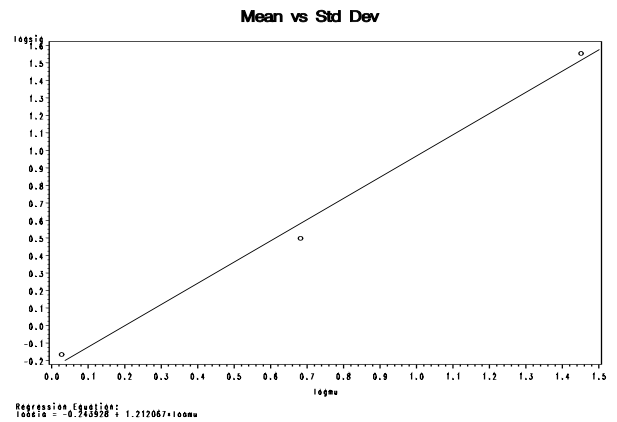
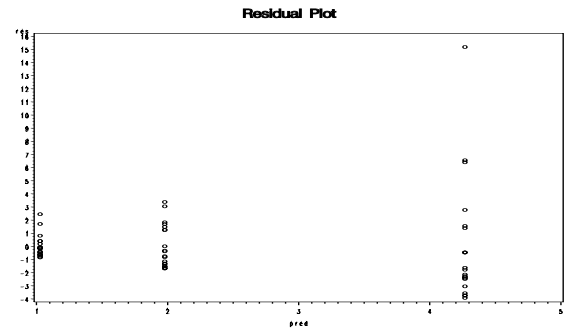
title1 'Mean vs Std Dev'; symbol1 v=circle i=r1;
proc gplot; plot logsig*logmu / regeqn;

proc transreg data=one;
  model boxcox(resp / lambda=-2 to 2 by .2) = class(trt);

run;

```

5-12



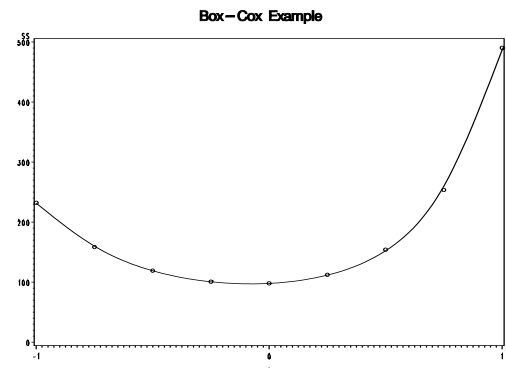
5-13

Transformation Information for BoxCox(resp)

Lambda	R-Square	Log Like
-2.0	0.10	-108.906
-1.8	0.10	-93.460
-1.6	0.11	-78.920
-1.4	0.12	-65.416
-1.2	0.13	-53.103
-1.0	0.14	-42.153
-0.8	0.16	-32.767
-0.6	0.17	-25.184
-0.4	0.19	-19.683
-0.2	0.20	-16.593 *
0.0 +	0.21	-16.284 <
0.2	0.22	-19.124
0.4	0.22	-25.398
0.6	0.22	-35.200
0.8	0.20	-48.378
1.0	0.18	-64.563
1.2	0.17	-83.272
1.4	0.15	-104.019
1.6	0.13	-126.377
1.8	0.12	-150.006
2.0	0.10	-174.641

< - Best Lambda
 * - Confidence Interval
 + - Convenient Lambda

OBS	L	SS	OBS	L	SS
1	-2.00	2150.06	10	0.25	112.37
2	-1.75	1134.83	11	0.50	154.23
3	-1.50	628.94	12	0.75	253.63
4	-1.25	369.35	13	1.00	490.36
5	-1.00	232.32	14	1.25	1081.29
6	-0.75	158.56	15	1.50	2636.06
7	-0.50	119.28	16	1.75	6924.95
8	-0.25	100.86	17	2.00	19233.39
9	0.00	98.09			



5-14

5-15

Unusual Observations

- Can use residual plot to detect unusual observations
- Unusual = large $|\hat{\epsilon}_{ij}|$
- Sometimes typographical error
- Otherwise worth investigation
 - Helpful to have detailed lab notes
- Don't immediately look for excuses
- Formal test to see if unusual
- But does not answer if in error
- Does observation make a difference if excluded?
- Perform analysis with and without observations(s)

5-16

Near-Zero/Truncated Values

- Can have heavily skewed observations (near zero)
 - Concentration of rare contaminant
 - Number of defects in assembly line
 - Number of birds at a given site
- Measurements may be truncated
 - Concentration of rare contaminant (not detectable)
 - Lifetime of component (does not fail)
- Transformations sometimes successful
 - $\log(x+.001)$, $\sqrt{x+.001}$
- Non-parametric analysis (Kruskal-Wallis)
- More advanced
 - Assume non-Normal errors (GLIM models)
 - EM algorithm (iterative MLE approach)
 - Bayesian methods

5-17

```
options nocenter ps=65 ls=80;
```

```
data new;
input strain nitrogen @@;
cards;
1 19.4 1 32.6 1 27.0 1 32.1 1 33.0
2 17.7 2 24.8 2 27.9 2 25.2 2 24.3
3 17.0 3 19.4 3 9.1 3 11.9 3 15.8
4 20.7 4 21.0 4 20.5 4 18.8 4 18.6
5 14.3 5 14.4 5 11.8 5 11.6 5 14.2
6 17.3 6 19.4 6 19.1 6 16.9 6 20.8
;
```

```
proc npar1way;
class strain;
var nitrogen;
```

```
run;
```

N P A R 1 W A Y P R O C E D U R E

Analysis of Variance for Variable NITROGEN

STRAIN	N	Mean	Among MS	Within MS
1	5	28.820000	169.409333	11.7886667
2	5	23.980000		
3	5	14.640000	F Value	Prob > F
4	5	19.920000	14.371	0.0001
5	5	13.260000		
6	5	18.700000		

5-18

N P A R 1 W A Y P R O C E D U R E

Wilcoxon Scores (Rank Sums) for Variable NITROGEN

STRAIN	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
1	5	130.0	77.5000000	17.9618849	26.0000000
2	5	111.0	77.5000000	17.9618849	22.2000000
3	5	40.0	77.5000000	17.9618849	8.0000000
4	5	88.0	77.5000000	17.9618849	17.6000000
5	5	23.0	77.5000000	17.9618849	4.6000000
6	5	73.0	77.5000000	17.9618849	14.6000000

Average Scores Were Used for Ties

Kruskal-Wallis Test (Chi-Square Approximation)

CHISQ = 21.659 DF = 5 Prob > CHISQ = 0.0006

Median Scores (Number of Points Above Median) for Variable NITROGEN

STRAIN	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
1	5	5.0	2.5000000	1.03806850	1.0000000
2	5	4.0	2.5000000	1.03806850	0.8000000
3	5	1.0	2.5000000	1.03806850	0.2000000
4	5	3.0	2.5000000	1.03806850	0.6000000
5	5	0.0	2.5000000	1.03806850	0.0000000
6	5	2.0	2.5000000	1.03806850	0.4000000

Average Scores Were Used for Ties

Median 1-Way Analysis (Chi-Square Approximation)

CHISQ = 13.533 DF = 5 Prob > CHISQ = 0.0189

5-19