Lecture 3. Experiments with a Single Factor: ANOVA

Montgomery Sections 3.1 through 3.5 and Section 15.1.1

Tensile Strength Experiment

Investigate the tensile strength of a new synthetic fiber. The factor is the weight percent of cotton used in the blend of the materials for the fiber and it has five levels.

percent	tensile strength						
of cotton	1	2	3	4	5	total	average
15	7	7	11	15	9	49	9.8
20	12	17	12	18	18	77	15.4
25	14	18	18	19	19	88	17.6
30	19	25	22	19	23	108	21.6
35	7	10	11	15	11	54	10.8

Data Layout for Single-Factor Experiments

treatment	observations				totals	averages
1	y_{11}	y_{12}	$\cdots$	y_{1n}	$y_{1.}$	$\bar{y}_{1.}$
2	y_{21}	y_{22}	$\cdots$	y_{2n}	$y_{2.}$	$\bar{y}_{2.}$
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$	$\vdots$
a	y_{a1}	y_{a2}	$\cdots$	y_{an}	$y_{a.}$	$\bar{y}_{a.}$

Analysis of Variance

- Statistical Model (Factor Effects Model):

$$y_{ij} = \mu + \tau_i + \epsilon_{ij} \quad \begin{cases} i = 1, 2 \ldots, a \\ j = 1, 2, \ldots, n_i \end{cases}$$

μ - grand mean; τ_i - ith treatment effect; $\epsilon_{ij} \overset{iid}{\sim} N(0, \sigma^2)$ - error

Constraint: $\sum_{i=1}^{a} \tau_i = 0$ (Conceptual Approach; SAS: $\tau_a = 0$).

- Estimates for parameters:

$$\hat{\mu} = \overline{y}_{..}$$
$$\hat{\tau}_i = (\overline{y}_{i.} - \overline{y}_{..})$$
$$\hat{\epsilon}_{ij} = y_{ij} - \overline{y}_{i.} \qquad (\text{ residual })$$

- Basic Hypotheses:

$H_0 : \tau_1 = \tau_2 = \ldots = \tau_a = 0$ vs $H_1 : \tau_i \neq 0$ for at least one i

Analysis of Variance (ANOVA) Table

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0
Between	$\text{SS}_{\text{Treatment}}$	$a-1$	$\text{MS}_{\text{Treatment}}$	F_0
Within	SS_{E}	$N-a$	MS_{E}	
Total	SS_{T}	$N-1$		

- If balanced: $N = n \times a$

  $\text{SS}_{\text{T}} = \sum\sum y_{ij}^2 - y_{..}^2/N$; $\text{SS}_{\text{Treatment}} = \frac{1}{n}\sum y_{i.}^2 - y_{..}^2/N$

  SS_{E}=SS_{T} - $\text{SS}_{\text{Treatment}}$

- If unbalanced: $N = \sum_{i=1}^{a} n_i$

  $\text{SS}_{\text{T}} = \sum\sum y_{ij}^2 - y_{..}^2/N$; $\text{SS}_{\text{Treatment}} = \sum \frac{y_{i.}^2}{n_i} - y_{..}^2/N$

  SS_{E}=SS_{T} - $\text{SS}_{\text{Treatment}}$

- $\text{SS}_{\text{Treatments}} = \sum_{i=1}^{a} n_i\hat{\tau}_i^2$ and $\text{SS}_{\text{E}} = \sum_i \sum_j \hat{\epsilon}_{ij}^2$.

- The Expected Mean Squares (EMS) are

  E($\mathrm{MS_E}$)=σ^2

  E($\mathrm{MS_{Treatment}}$) = $\sigma^2 + \sum n_i \tau_i^2/(a-1)$

- Test Statistic

$$F_0 = \frac{\mathrm{SS_{Treatments}}/(a-1)}{\mathrm{SS_E}/(N-a)} = \frac{\mathrm{MS_{Treatments}}}{\mathrm{MS_E}}$$

- Under H_0:

$$F_0 = \frac{\mathrm{SS_{Treatment}}/\sigma^2(a-1)}{\mathrm{SS_E}/\sigma^2(N-a)} = \frac{\chi^2_{a-1}/(a-1)}{\chi^2_{N-a}/(N-a)} \sim F_{a-1,N-a}$$

- **Decision Rule**: If $F_0 > F_{\alpha,a-1,N-a}$ then reject H_0

- When $a = 2$, the square of the t-test statistic $t_0^2 = \frac{\mathrm{MS_{Treatment}}}{\mathrm{MS_E}} = F_0$.
  - F-test and two-sample two-sided test are equivalent.

Example

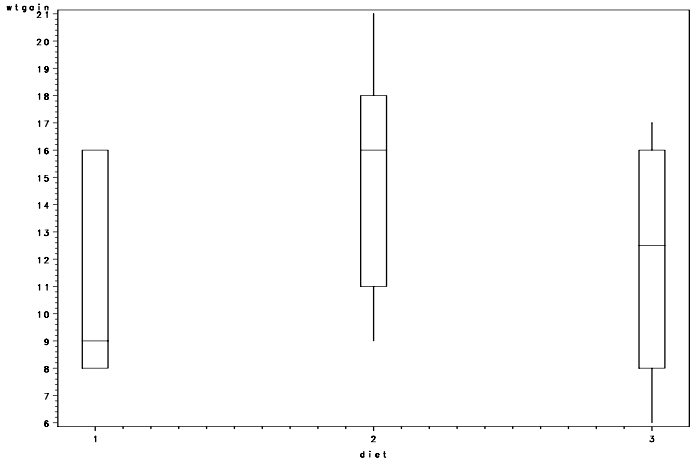
Twelve lambs are randomly assigned to three different diets. The weight gain (in two weeks) is recorded. Is there a difference between the diets?

Diet 1	8	16	9		
Diet 2	9	16	21	11	18
Diet 3	15	10	17	6	

- $N = 12, \ \sum\sum y_{ij} = 156$ and $\bar{y}_{..} = 156/12 = 13.$
- $n_1 = 3, y_{1.} = 33, \bar{y}_{1.} = 11; n_2 = 5, y_{2.} = 75, \bar{y}_{2.} = 15; n_3 = 4, y_{3.} = 48$ and $\bar{y}_{3.} = 12.$
- $\hat{\tau}_1 = \bar{y}_{1.} - \bar{y}_{..} = 11 - 13 = -2$; Similarly, $\hat{\tau}_2 = 15 - 13 = 2$ and $\hat{\tau}_3 = 12 - 13 = -1.$
- $\mathrm{SS_T} = \sum_i \sum_j (y_{ij} - \bar{y}_{..})^2 = 246.$
- $\mathrm{SS_{Treatment}} = 3 * (-2)^2 + 5 * (2)^2 + 4 * (-1)^2 = 36.$
- $\mathrm{SS_E} = 246 - 36 = 210; \ \mathrm{MS}_E = \hat{\sigma}^2 = 210/(12 - 3) = 23.33$
- $F_0 = (36/2)/(210/9) = 0.77$; P-value $>$ 0.25;
- Fail to reject $H_0 : \tau_1 = \tau_2 = \tau_3 = 0.$

Using SAS (lambs.sas)

```
option nocenter ps=65 ls=80;
data lambs;
  input diet wtgain @@;
  datalines;
  1  8  1 16  1  9  2  9  2 16  2 21
  2 11  2 18  3 15  3 10  3 17  3  6
;

symbol1 bwidth=5 i=box; axis1 offset=(5);
proc gplot; plot wtgain*diet / frame haxis=axis1; run; quit;
```



```
proc glm;
  class diet;
  model wtgain=diet;
  output out=diag r=res p=pred; run; quit;
```

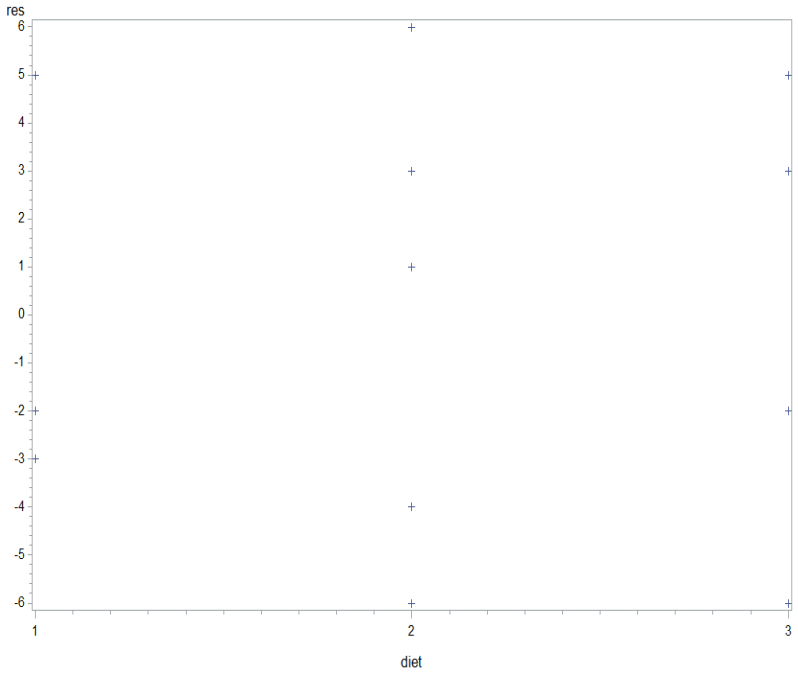
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	36.0000000	18.0000000	0.77	0.4907
Error	9	210.0000000	23.3333333		
Corrected Total	11	246.0000000			

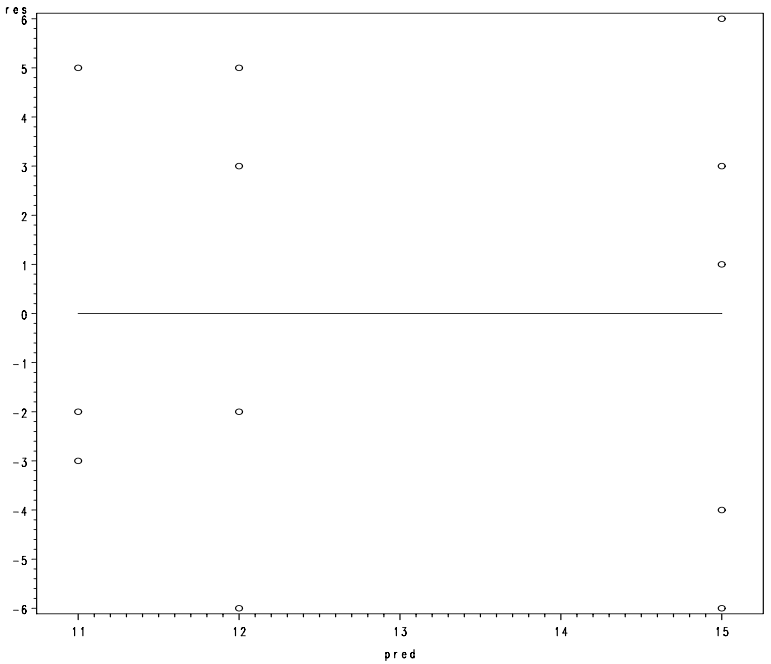
R-Square	Coeff Var	Root MSE	wtgain Mean
0.146341	37.15738	4.830459	13.00000

Source	DF	Type I SS	Mean Square	F Value	Pr > F
diet	2	36.00000000	18.00000000	0.77	0.4907

Source	DF	Type III SS	Mean Square	F Value	Pr > F
diet	2	36.00000000	18.00000000	0.77	0.4907

```
proc gplot; plot res*diet /frame haxis=axis1;

proc sort; by pred;
symbol1 v=circle i=sm50;
proc gplot; plot res*pred / haxis=axis1;
run; quit;
```





Model Checking and Diagnostics

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

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

- Note that the predicted response at treatment i is $\hat{y}_{ij} = \bar{y}_{i.}$
- Diagnostics use predicted responses and residuals.

Diagnostics

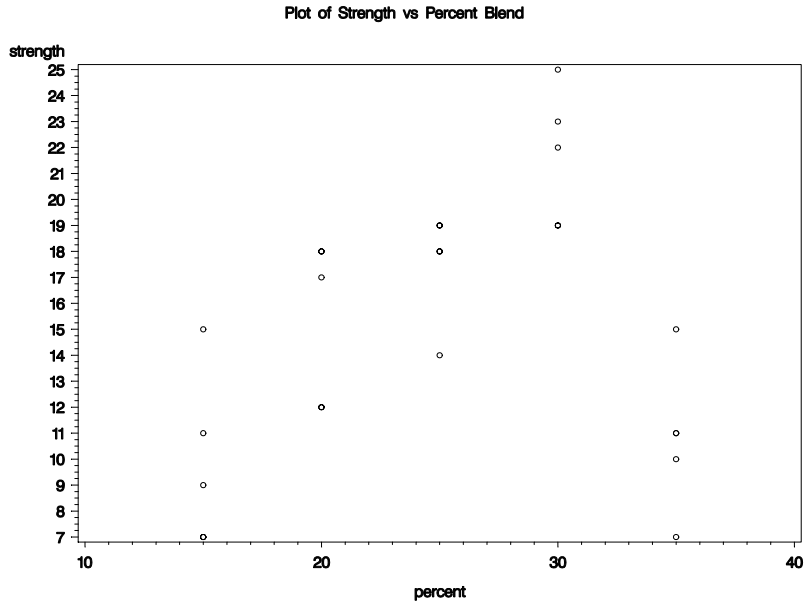
- Normality
  - Histogram of residuals
  - Normal probability plot / QQ plot
  - Shapiro-Wilk Test
- Constant Variance
  - Plot $\hat{\epsilon}_{ij}$ vs $\hat{y}_{ij}$ (residual plot)
  - Bartlett's or Modified Levene's Test
- Independence
  - Plot $\hat{\epsilon}_{ij}$ vs time/space
  - Plot $\hat{\epsilon}_{ij}$ vs variable of interest
- Outliers

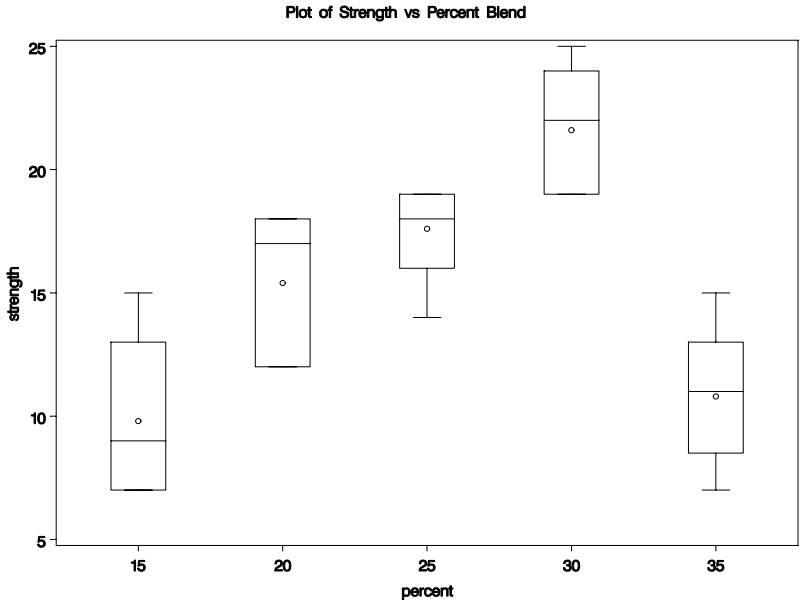
Diagnostics Example: Tensile Strength Experiment

```
options ls=80 ps=60 nocenter;
goptions device=win target=winprtm rotate=landscape ftext=swiss
  hsize=8.0in vsize=6.0in htext=1.5 htitle=1.5 hpos=60 vpos=60
  horigin=0.5in vorigin=0.5in;
data one;
  infile 'c:\saswork\data\tensile.dat';
  input percent strength time;

title1 'Tensile Strength Example';
proc print data=one; run;
```

Obs	percent	strength	time
1	15	7	15
2	15	7	19
3	15	15	25
4	15	11	12
5	15	9	6
6	20	12	8
:	:	:	:
24	35	15	16
25	35	11	23

```
symbol1 v=circle i=none;
title1 'Plot of Strength vs Percent Blend';
proc gplot data=one;  plot strength*percent/frame; run;

proc boxplot;
 plot strength*percent/boxstyle=skeletal pctldef=4; run;
```





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

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			

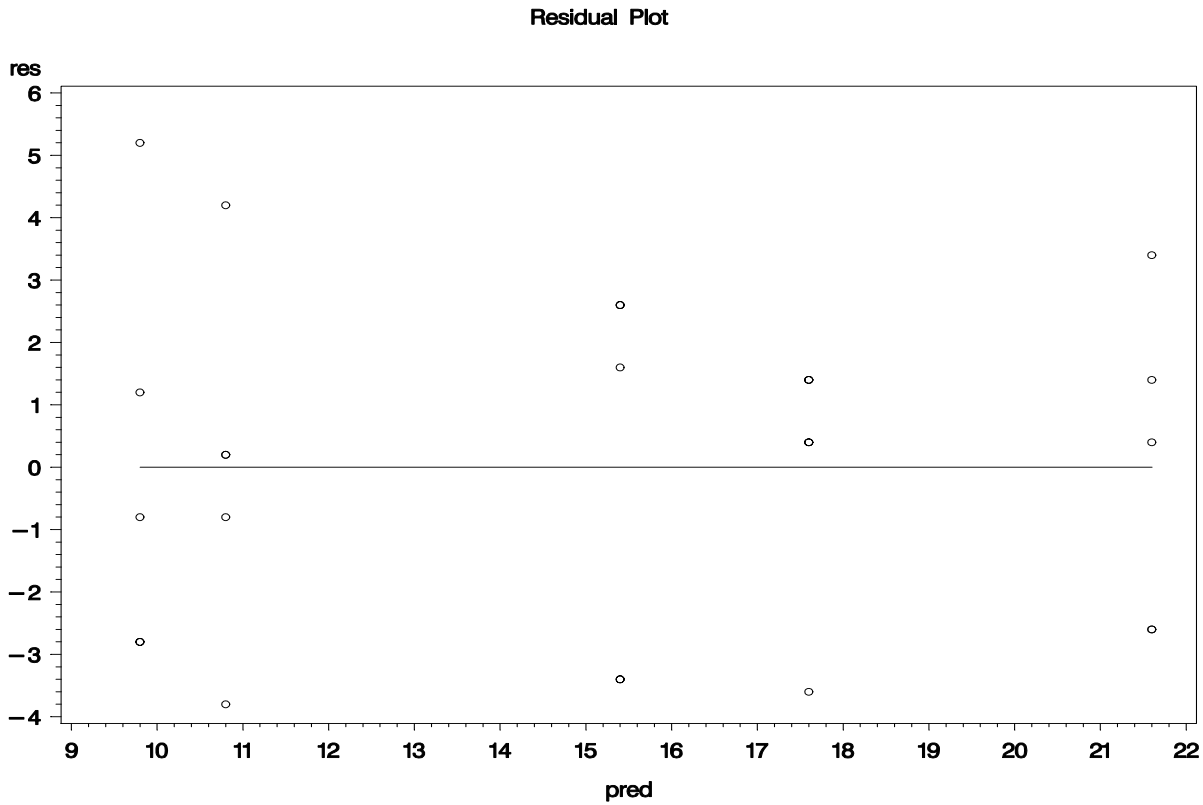
Levene's Test for Homogeneity of strength Variance
ANOVA of Squared Deviations from Group Means

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

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



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

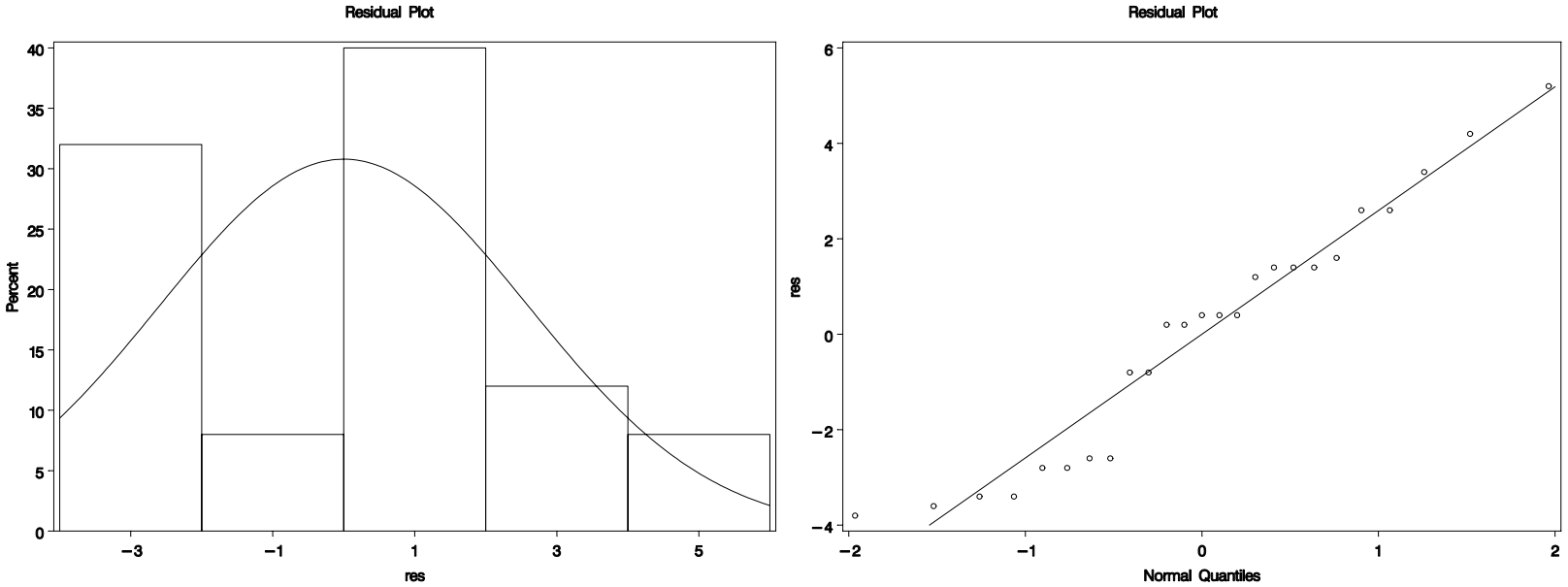
Moments

N	25	Sum Weights	25
Mean	0	Sum Observations	0
Std Deviation	2.59165327	Variance	6.71666667
Skewness	0.11239681	Kurtosis	-0.8683604
Uncorrected SS	161.2	Corrected SS	161.2
Coeff Variation	.	Std Error Mean	0.51833065

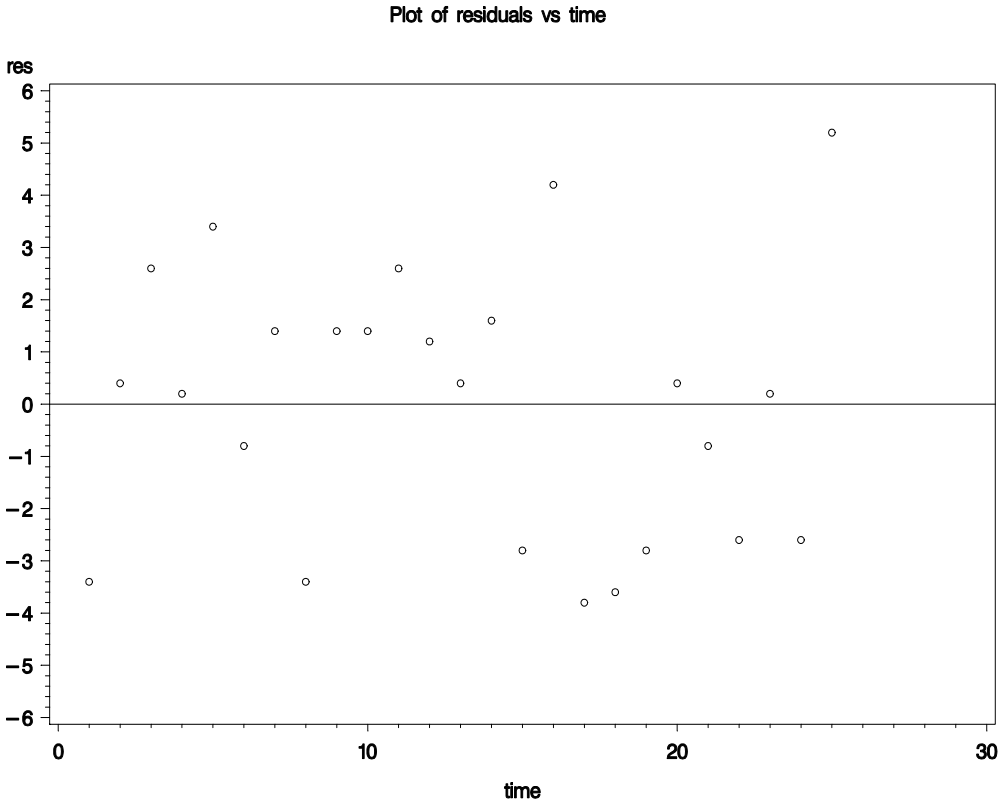
Tests for Normality

Test	--Statistic---		-----p Value------	
Shapiro-Wilk	W	0.943868	Pr < W	0.1818
Kolmogorov-Smirnov	D	0.162123	Pr > D	0.0885
Cramer-von Mises	W-Sq	0.080455	Pr > W-Sq	0.2026
Anderson-Darling	A-Sq	0.518572	Pr > A-Sq	0.1775

Histogram of Residuals & QQ Plot



```
/* Time Serial Plot */
symbol1 v=circle i=none;
title1 'Plot of residuals vs time';
proc gplot; plot res*time / vref=0 vaxis=-6 to 6 by 1;
run;
```



Non-Constant Variance: Impact and Remedy

- Does not affect F-test dramatically when the experiment is balanced
- Why concern?
  - Comparison of treatments depends on MSE
  - Incorrect intervals and comparison results
- Variance-Stabilizing Transformations
  - Ideas for Finding Proper Transformations ($E[Y] = \mu$, Var$(Y) = \sigma^2$)

$$\begin{aligned} f(Y) &\approx f(\mu) + (Y - \mu) f'(\mu) \\ \Longrightarrow \text{Var}(f(Y)) &\approx [f'(\mu)]^2 \text{Var}(Y) = [f'(\mu)]^2 \sigma^2 \end{aligned}$$

    * Find f such that Var$(f(Y))$ does not depend on μ anymore. So, $\tilde{Y} = f(Y)$ has constant variance for different $f(\mu)$.
  - Common transformations

    $\sqrt{x}$, $\log(x)$, $1/x$, $\arcsin(\sqrt{x})$, and $1/\sqrt{x}$

Transformations

- Suppose σ^2 is a function of μ, that is $\sigma^2 = g(\mu)$
- Want to find transformation f such that $\tilde{Y} = f(Y)$ has constant variance: Var($\tilde{Y}$) does not depend on μ.
- Have shown Var($\tilde{Y}$)$\approx [f'(\mu)]^2\sigma^2 \approx [f'(\mu)]^2 g(\mu)$
- Want to choose f such that $[f'(\mu)]^2 g(\mu) \approx c$

Examples

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

Identify Box-Cox Transformation Using Data: Approximate Method

- From the previous slide, if $\sigma = \theta\mu^{\beta}$, the transformation is

$$f(Y) = \begin{cases} Y^{1-\beta} & \beta \neq 1; \\ \log Y & \beta = 1 \end{cases}$$

  So it is crucial to estimate β based on data y_{ij}, $i = 1, \ldots, a$.

- We have $\log\sigma_i = \log\theta + \beta\log\mu_i$

- Let s_i and $\bar{y}_{i.}$ be the sample standard deviations and means. Because $\hat{\sigma}_i = s_i$ and $\hat{\mu}_i = \bar{y}_{i.}$, for $i = 1, \ldots, a$,

$$\begin{aligned} & \log s_i \approx \text{ constant } + \beta\log\bar{y}_{i.}, \\ \Longrightarrow \quad & \log s_i = \text{ constant } + \beta\log\bar{y}_{i.} + \text{error}_i. \end{aligned}$$

- We can plot $\log s_i$ against $\log\bar{y}_{i.}$, fit a straight line and use the slope to estimate β.

Identify Box-Cox Transformation: Formal Method

1 . For a fixed λ, perform analysis of variance on

$$y_{ij}(\lambda) = \begin{cases} \frac{y_{ij}^{\lambda}-1}{\lambda \dot{y}^{\lambda-1}} & \lambda \neq 0 \\ \\ \dot{y} \log y_{ij} & \lambda = 0 \end{cases} \quad \text{where } \dot{y} = \left(\prod_{i=1}^{a} \prod_{j=1}^{n_i} y_{ij} \right)^{1/N}.$$

2 . Step 1 generates a transformed data $y_{ij}(\lambda)$. Apply ANOVA to the new data and obtain SS_E. Because SS_E depends on λ, it is denoted by $\text{SS}_E(\lambda)$.

- Repeat 1 and 2 for various λ in an interval, e.g., [-2,2], and record $\text{SS}_E(\lambda)$

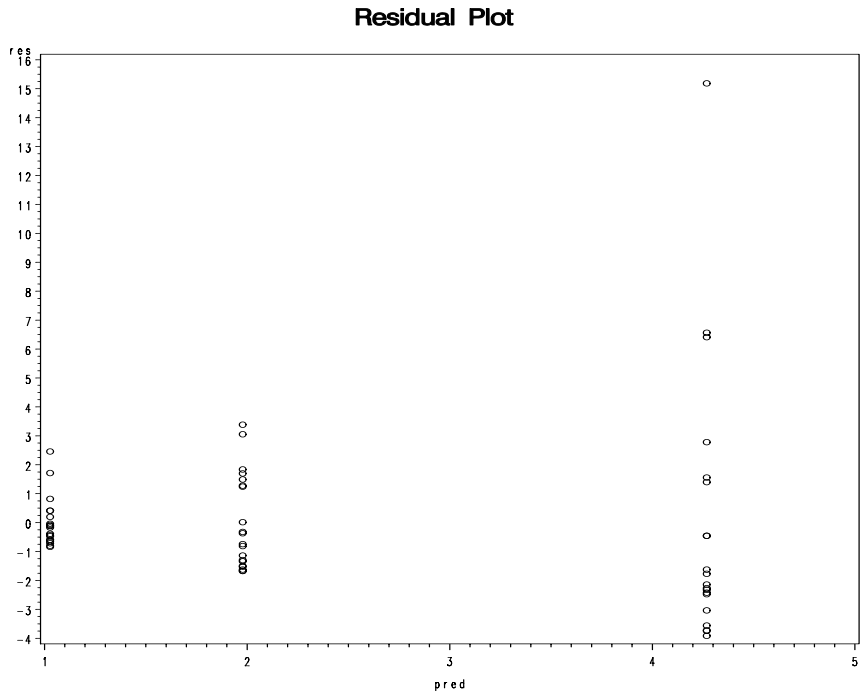
3 Find λ_0 which minimizes $\text{SS}_E(\lambda)$ and pick up a meaningful λ in the neighborhood of λ_0. Denote it again by λ. (Maximum Likelihood Principle)

4 The transformation is:

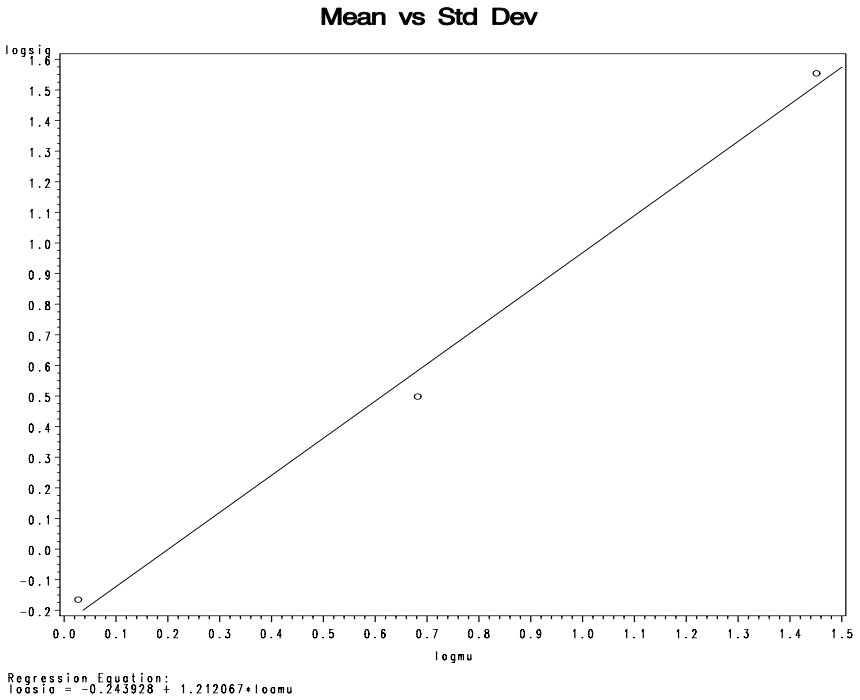
$$\tilde{y}_{ij} = y_{ij}^{\lambda_0} \text{ if } \lambda_0 \neq 0;$$
$$\tilde{y}_{ij} = \log y_{ij} \text{ if } \lambda_0 = 0.$$

Example: Approximate Method (trans.sas)

```
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; run;

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



```
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=rl;
proc gplot; plot logsig*logmu / regeqn; run;
```



Example: Formal Method (trans1.sas)

```
data one;
  infile 'c:\saswork\data\boxcox.dat';
  input trt resp;
  logresp = log(resp);

proc univariate data=one noprint;
 var logresp; output out=two mean=mlogresp;

data three;
  set one; if _n_ eq 1 then set two;
  ydot = exp(mlogresp);
  do l=-2.0 to 2.0 by .25;
    den = l*ydot**(l-1);   if abs(l) eq 0 then den = 1;
    yl=(resp**l -1)/den;   if abs(l) < 0.0001 then yl=ydot*log(resp);
    output;
  end;
  keep trt yl l;

proc sort data=three out=three;  by l;
```

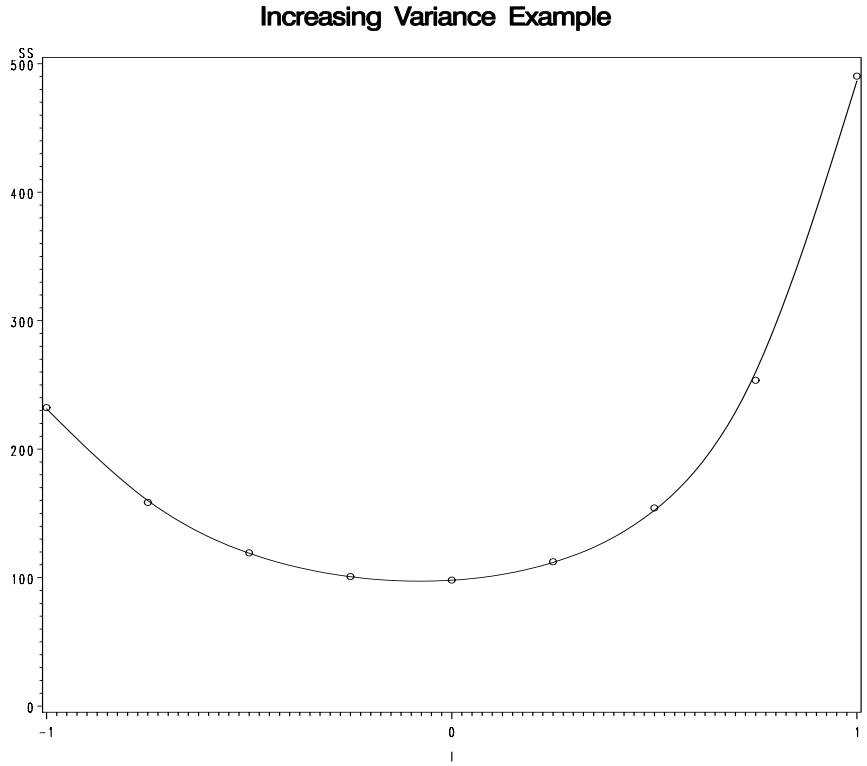
```
proc glm data=three noprint outstat=four;
  class trt;  model yl=trt;  by l;

data five;  set four;
  if _SOURCE_ eq 'ERROR';  keep l SS;

proc print data=five; run;
```

$SS_E(\lambda)$ and λ

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			

```
symbol1 v=circle i=sm50;
proc gplot; plot SS*l; run;
```

Plot of $SS_E(\lambda)$ vs λ



Using PROC TRANSREG

```
*PBOXCOXTABLE: Print likelihood table;
proc transreg data=one PBOXCOXTABLE;
  model boxcox(resp/lambda=-2.0 to 2.0 by 0.1)=class(trt); run;
```

Box-Cox Transformation Information
for resp

Lambda		R-Square	Log Like	
-2.0		0.10	-108.906	
:		:	:	
-0.5		0.18	-22.154	
-0.4		0.19	-19.683	
-0.3		0.20	-17.814	*
-0.2		0.20	-16.593	*
-0.1		0.21	-16.067	<
0.0	+	0.21	-16.284	*
0.1		0.22	-17.289	*
0.2		0.22	-19.124	
0.3		0.22	-21.820	
:		:	:	
2.0		0.10	-174.641	

< Best Lambda
* 95% Confidence Interval
+ Convenient Lambda

Kruskal-Wallis Test: a Nonparametric Alternative for Nonnormality

a treatments, H_0: a treatments are not different.

- Rank the observations y_{ij} in ascending order
- Replace each observation by its rank R_{ij} (assign average for tied observations), and apply one-way ANOVA to R_{ij}.
- Test statistic ($a = 2 \Longrightarrow$ Wilcoxon rank-sum test)

$$H = \frac{1}{S^2}\left[\sum_{i=1}^{a}\frac{R_{i.}^2}{n_i} - \frac{N(N+1)^2}{4}\right] \approx \chi^2_{a-1} \text{ under } H_0$$

  where $S^2 = \frac{1}{N-1}\left[\sum_{i=1}^{a}\sum_{j=1}^{n_i} R_{ij}^2 - \frac{N(N+1)^2}{4}\right]$

- Decision Rule: reject H_0 if $H > \chi^2_{\alpha,a-1}$.
- Let F_0 be the F-test statistic in ANOVA based on R_{ij}. Then

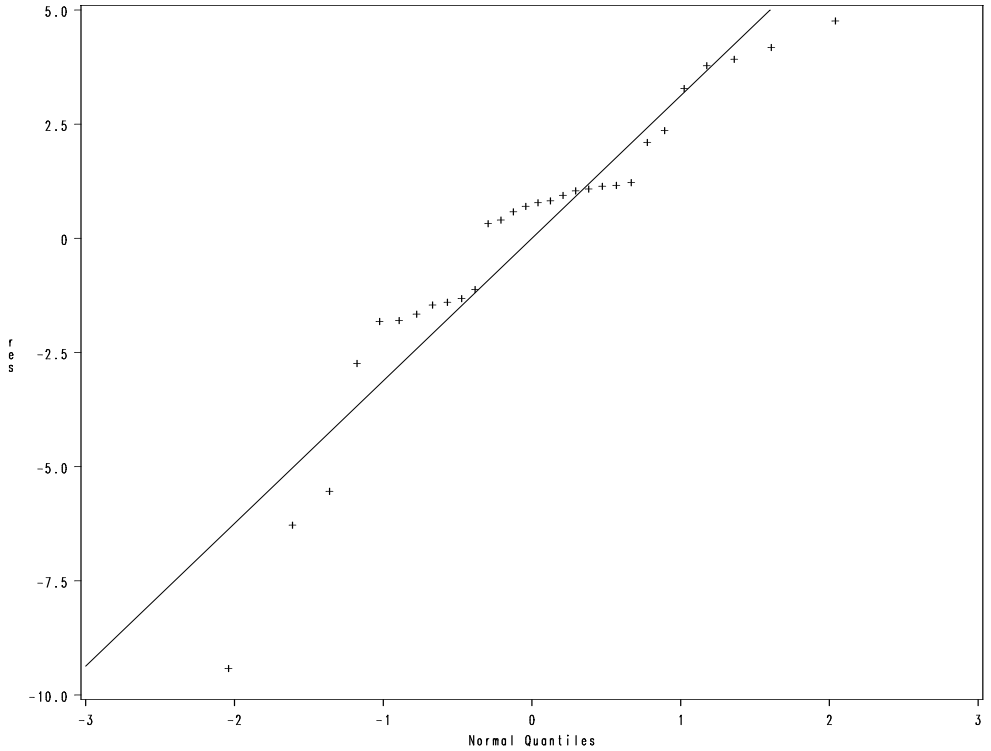
$$F_0 = \frac{H/(a-1)}{(N-1-H)/(N-a)}$$

A Nonnormality Example

```
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 glm data=new;
  class strain; model nitrogen=strain;
  output out=newres r=res; run;

proc univariate data=newres normal;
  var res;  qqplot res / normal (L=1 mu=est sigma=est);
  run; quit;
```

Tests for Normality

Test		Statistic		p Value
Shapiro-Wilk	W	0.910027	Pr < W	0.0149
Kolmogorov-Smirnov	D	0.174133	Pr > D	0.0205
Cramer-von Mises	W-Sq	0.155870	Pr > W-Sq	0.0198
Anderson-Darling	A-Sq	0.908188	Pr > A-Sq	0.0194



```
proc npar1way data=new;  /* May need option: wilcoxon */
  class strain;  var nitrogen; run;
```

```
Analysis of Variance for Variable nitrogen
      Classified by Variable strain

Source    DF    Sum of Squares    Mean Square  F Value  Pr > F
-------------------------------------------------------------
Among      5        847.046667     169.409333  14.3705  <.0001
Within    24        282.928000      11.788667

Kruskal-Wallis Test
Chi-Square         21.6593
DF                       5
Pr > Chi-Square    0.0006

Median One-Way Analysis
Chi-Square         13.5333
DF                       5
Pr > Chi-Square     0.0189
```

Linear Combinations of Treatment Means

- ANOVA Model:

$$y_{ij} = \mu + \tau_i + \epsilon_{ij} \quad (\tau_i\text{: treatment effect})$$

$$= \mu_i + \epsilon_{ij} \quad (\mu_i\text{: treatment mean})$$

- Linear combination with given coefficients $c_1, c_2, \ldots, c_a$:

$$L = c_1\mu_1 + c_2\mu_2 + \ldots + c_a\mu_a = \sum_{i=1}^{a} c_i\mu_i,$$

- Want to test: $H_0 : L = \sum c_i\mu_i = L_0$
- Examples:

1. Pairwise comparison: $\mu_i - \mu_j = 0$ for all possible i and j.
2. Compare treatment vs control: $\mu_i - \mu_1 = 0$ when treatment 1 is a control and $i = 2, ..., a$ are new treatments.
3. General cases such as $\mu_1 - 2\mu_2 + \mu_3 = 0$, $\mu_1 + 3\mu_2 - 6\mu_3 = 0$, etc.

- Estimate of L:

$$\hat{L} = \sum c_i \hat{\mu}_i = \sum c_i \bar{y}_{i.}$$

$$\text{Var}(\hat{L}) = \sum c_i^2 \text{Var}(\bar{y}_{i.}) = \sigma^2 \sum \frac{c_i^2}{n_i} \left(= \frac{\sigma^2}{n} \sum c_i^2 \right)$$

- Standard Error of $\hat{L}$

$$\text{S.E.}\{\hat{L}\} = \sqrt{\text{MSE} \sum \frac{c_i^2}{n_i}}$$

- Test statistic

$$t_0 = \frac{\hat{L} - L_0}{S.E.\{\hat{L}\}} \sim t(N-a) \text{ under } H_0$$

Example: Lambs Diet Experiment

- Denote the treatment means of three diets by μ_1, μ_2 and μ_3. Suppose one wants to test $H_0 : L = 60$ with
  $L = \mu_1 + 2\mu_2 + 3\mu_3 = 6\mu + \tau_1 + 2\tau_2 + 3\tau_3$.

```
proc glm data=lambs;
  class diet;
  model wtgain=diet;
  means diet;
  estimate 'l1' intercept 6 diet 1 2 3;
run;
```

Parameter	Estimate	Standard Error	t Value	Pr > |t|
l1	77.0000000	8.88506862	8.67	<.0001

- $t_0 = (77.0 - 60)/8.89 = 1.91$

$$P - \text{value} = P(t \leq -1.91 \text{ or } t \geq 1.91 | t(12 - 3)) = .088$$

- Fail to reject $H_0 : \mu_1 + 2\mu_2 + 3\mu_3 = 60$ at $\alpha = 5\%$.

Contrasts

- $\Gamma = \sum_{i=1}^{a} c_i \mu_i$ is a contrast if $\sum_{i=1}^{a} c_i = 0$.

  Equivalently, $\Gamma = \sum_{i=1}^{a} c_i \tau_i$.

- Examples

  1. $\Gamma_1 = \mu_1 - \mu_2 = \mu_1 - \mu_2 + 0\mu_3 + 0\mu_4$,
     $c_1 = 1, c_2 = -1, c_3 = 0, c_4 = 0$
     Comparing μ_1 and μ_2.
  2. $\Gamma_2 = \mu_1 - 0.5\mu_2 - 0.5\mu_3 = \mu_1 - 0.5\mu_2 - 0.5\mu_3 + 0\mu_4$
     $c_1 = 1, c_2 = -0.5, c_3 = -0.5, c_4 = 0$
     Comparing μ_1 and the average of μ_2 and μ_3.

- Estimate of Γ:

$$C = \sum_{i=1}^{a} c_i \bar{y}_{i.}$$

- Contrast Sum of Squares

$$\text{SS}_C = \left(\sum c_i \overline{y}_{i.}\right)^2 / \sum (c_i^2/n_i)$$

  SS_C represents the amount of variation attributable Γ.

- Test: $H_0 : \Gamma = 0$

$$t_0 = \frac{C}{\text{S.E.}_C} \overset{H_0}{\sim} t(N-a)$$

$$t_0^2 = \frac{(\sum c_i \bar{y}_{i.})^2}{\text{MSE} \sum \frac{c_i^2}{n_i}} = \frac{\text{SS}_C/1}{\text{MSE}} \overset{H_0}{\sim} F_{1,N-a}$$

Tensile Strength Example (cont.sas)

```
proc glm data=one;
  class percent;
  model strength=percent;
  contrast 'C1' percent 0 0 0 1 -1;
  contrast 'C2' percent 1 0 1 -1 -1;
  contrast 'C3' percent 1 0 -1 0 0;
  contrast 'C4' percent 1 -4 1 1 1;
```

Dependent Variable: STRENGTH

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	475.76000	118.94000	14.76	0.0001
Error	20	161.20000	8.06000		
Corrected Total	24	636.96000			

Source	DF	Type I SS	Mean Square	F Value	Pr > F
PERCENT	4	475.76000	118.94000	14.76	0.0001

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
C1	1	291.60000	291.60000	36.18	0.0001
C2	1	31.25000	31.25000	3.88	0.0630
C3	1	152.10000	152.10000	18.87	0.0003
C4	1	0.81000	0.81000	0.10	0.7545

Orthogonal Contrasts

- Two contrasts $\{c_i\}$ and $\{d_i\}$ are **Orthogonal** if

$$\sum_{i=1}^{a} \frac{c_i d_i}{n_i} = 0 \quad (\sum_{i=1}^{a} c_i d_i = 0 \text{ for balanced experiments})$$

- Example (Balanced Experiment)

  $\Gamma_1 = \mu_1 + \mu_2 - \mu_3 - \mu_4$, So $c_1 = 1, c_2 = 1, c_3 = -1, c_4 = -1$.

  $\Gamma_2 = \mu_1 - \mu_2 + \mu_3 - \mu_4$. So $d_1 = 1, d_2 = -1, d_3 = 1, d_4 = -1$

  It is easy to verify that both Γ_1 and Γ_2 are contrasts. Furthermore,

  $c_1 d_1 + c_2 d_2 + c_3 d_3 + c_4 d_4 =$

  $1 \times 1 + 1 \times (-1) + (-1) \times 1 + (-1) \times (-1) = 0$. Hence, Γ_1 and Γ_2 are orthogonal to each other.

- A **complete set** of orthogonal contrasts $\mathcal{C} = \{\Gamma_1, \Gamma_2, \ldots, \Gamma_{a-1}\}$ if contrasts are mutually orthogonal and there does not exist a contrast orthogonal outside of $\mathcal{C}$ to all the contrasts in $\mathcal{C}$.

- If there are a treatments, $\mathcal{C}$ must contain $a-1$ contrasts.
- Complete set is not unique. For example, in the tensile strength example

$$\mathcal{C}_1 \text{ : includes : } \begin{array}{rcrrrrr} \Gamma_1 & = & (0, & 0, & 0, & 1, & -1) \\ \Gamma_2 & = & (1, & 0, & 1, & -1, & -1) \\ \Gamma_3 & = & (1, & 0, & -1, & 0, & 0) \\ \Gamma_4 & = & (1, & -4, & 1, & 1, & 1) \end{array}$$

$$\mathcal{C}_2 \text{ : includes : } \begin{array}{rcrrrrr} \Gamma_1' & = & (-2, & -1, & 0, & 1, & 2) \\ \Gamma_2' & = & (2, & -1, & -2, & -1, & 2) \\ \Gamma_3' & = & (-1, & 2, & 0, & -2, & 1) \\ \Gamma_4' & = & (1, & -4, & 6, & -4, & 1) \end{array}$$

- Suppose $C_1, C_2, \ldots, C_{a-1}$ are the estimates of the contrasts in a complete set of contrasts $\{\Gamma_1, \Gamma_2, \ldots, \Gamma_{a-1}\}$, then

$$\text{SS}_{\text{Treatment}} = \text{SS}_{C_1} + \text{SS}_{C_2} + \cdots + \text{SS}_{C_{a-1}}$$
$$F_0 = \frac{\text{MS}_{\text{Treatment}}}{\text{MSE}} = \frac{F_{10} + F_{20} + \cdots + F_{(a-1)0}}{a-1}$$

  where F_{i0} is the test statistic used to test contrast Γ_i.

- Orthogonal contrasts (estimates) are independent with each other
  - The results follow Cochran's Theorem, so comparisons are indepedent
  - Example on Slide 39

- Can also use orthogonal contrasts to study trend
  - Only interesting if treatments are quantitative (ordered): $X_1, \cdots, X_a$
  - For equally spaced treatments and $n_i = n$, c_i in Table IX
  - Breakdown of polynomial regression $\mu(x) = \beta_0 + \beta_1 x + \cdots + \beta_{a-1} x^{a-1}$
    $\Longrightarrow \mu(x) = \tilde{\beta}_0 + \tilde{\beta}_1 P_1(x) + \cdots + \tilde{\beta}_{a-1} P_{a-1}(x)$
  - $P_k(x)$ is of k-th order, and $\sum_{i=1}^{a} P_k(X_i) = 0$ (constrast)

Determining Orthogonal Polynomial Coefficients Using SAS

- Complete set: $\{(P_k(X_1), \cdots, P_k(X_a)) : k = 1, \cdots, a-1\}$
- Often the levels of the treatments are not equally spaced
- Can use `PROC IML` to determine coefficients $(P_k(X_1), \cdots, P_k(X_a))$

```
proc iml;
  levels={1 2 5 10 20};   /* Consider these 5 levels */
  print levels;
  coef=ORPOL(levels,4);   /* Gives all coefs up through 4th order */
  coef=t(coef);           /* Puts coefs in rows instead of cols */
  coef=coef[2:5,];        /* Eliminates the first row of coef matrix*/
  print coef;   run;
```

LEVELS

1	2	5	10	20

COEF

-0.424967	-0.360578	-0.167411	0.1545335	0.798423
0.4348974	0.2072899	-0.325207	-0.711616	0.3946361
-0.433125	0.1365799	0.7252914	-0.510844	0.0820972
0.4926155	-0.779975	0.3743878	-0.093597	0.0065682

Testing Multiple Contrasts (Multiple Comparisons)

- One contrast:

$$H_0 : \Gamma = \sum c_i \mu_i = \Gamma_0 \text{ vs } H_1 : \Gamma \neq \Gamma_0 \text{ at } \alpha$$

100(1-α) Confidence Interval (CI) for Γ:

$$\text{CI} : \sum c_i \bar{y}_{i.} \pm t_{\alpha/2, N-a} \sqrt{MS_E \sum \frac{c_i^2}{n_i}}$$

$$P(\text{CI does not contain } L_0 | H_0) = \alpha (= \text{type I error rate})$$

- Decision Rule: Reject H_0 if CI does not contain Γ_0.

- Multiple contrasts, for $i = 1, 2, \cdots, m$,

$$H_0^{(i)} : \Gamma^{(i)} = \Gamma_0^{(i)}, \text{ vs } H_1^{(i)} : \Gamma^{(i)} \neq \Gamma_0^{(i)}$$

If we construct CI$_1$, CI$_2$,..., CI$_m$, each with 100(1-α) level, then for each CI$_i$,

$$P(\text{reject} H_0^{(i)} \mid H_0^{(i)}) = P(\text{CI}_i \text{ does not contain } \Gamma_0^{(i)} \mid H_0^{(i)}) = \alpha$$

- But the **overall Type I error rate** (probability of any type I error in testing $H_0^{(i)}$ vs $H_1^{(i)}$, $i = 1, \cdots, m$) is inflated and much larger than α, that is,

$$\begin{aligned} & P(\text{reject at least one of } \{H_0^{(i)}, i = 1, \cdots, m\} \mid H_0^{(1)}, \cdots, H_0^{(m)}) \\ = \quad & P(\text{at least one CI}_i \text{ do not contain } \Gamma_0^{(i)} \mid H_0^{(1)}, \cdots, H_0^{(m)}) \\ \gg \quad & \alpha \end{aligned}$$

- One way to achieve small overall error rate, we require much smaller error rate (α') of each individual CI$_i$.

Bonferroni Method for Testing Multiple Contrasts

- Bonferroni Inequality

$$\begin{aligned} & P(\text{reject at least one of } \{H_0^{(i)}, i = 1, \cdots, m\} \mid H_0^{(1)}, \cdots, H_0^{(m)}) \\ = & P(\text{reject } H_0^{(1)} \text{ or reject } H_0^{(2)} \text{ or ... or reject } H_0^{(m)} \mid H_0^{(1)}, \cdots, H_0^{(m)}) \\ \leq & P(\text{reject } H_0^{(1)} \mid H_0^{(1)}) + \cdots + P(\text{reject } H_0^{(m)} \mid H_0^{(m)}) = m\alpha' \end{aligned}$$

- In order to control overall error rate (or, overall confidence level), let

$$m\alpha' = \alpha, \text{ we have, } \alpha' = \alpha/m$$

- Bonferroni CIs:

$$\text{CI}_i : \sum c_{ij}\bar{y}_{j\cdot} \pm t_{\alpha/2m}(N-a)\sqrt{\text{MS}_E \sum \frac{c_{ij}^2}{n_j}}$$

- When m is large, Bonferroni CIs are too conservative (overall type II error too large).

Scheffe's Method for Testing All Contrasts

- Consider all possible contrasts: $\Gamma = \sum c_i \mu_i$

  Estimate: $C = \sum c_i \bar{y}_{i\cdot}$, St. Error: $\mathrm{S.E.}_C = \sqrt{\mathrm{MS}_E \sum \frac{c_i^2}{n_i}}$

- Critical value: $\sqrt{(a-1)F_{\alpha,a-1,N-a}}$

- Scheffe's simultaneous CI: $C \pm \sqrt{(a-1)F_{\alpha,a-1,N-a}}\ \mathrm{S.E.}_C$

- Overall confidence level and error rate for m contrasts

$$P(\text{CIs contain true parameter for any contrast}) \geq 1-\alpha$$

$$P(\text{at least one CI does not contain true parameter}) \leq \alpha$$

  Remark: Scheffe's method is also conservative, too conservative when m is small

Methods for Pairwise Comparisons

- There are $a(a-1)/2$ possible pairs: $\mu_i - \mu_j$ (contrast for comparing μ_i and μ_j). We may be interested in m pairs or all pairs.
- Standard Procedure:
  1. Estimation: $\bar{y}_{i.} - \bar{y}_{j.}$
  2. Compute a **Critical Difference (CD)** (based on the method employed)
  3. If

$$\mid \bar{y}_{i.} - \bar{y}_{j.} \mid > \mathrm{CD}$$

  or equivalently if the interval

$$(\bar{y}_{i.} - \bar{y}_{j.} - \mathrm{CD},\ \bar{y}_{i.} - \bar{y}_{j.} + \mathrm{CD})$$

  does not contain zero, declare $\mu_i - \mu_j$ significant.

Methods for Calculating CD

- Least significant difference (LSD):

$$\text{CD} = t_{\alpha/2, N-a}\sqrt{\text{MS}_E(1/n_i + 1/n_j)}$$

not control overall error rate

- Bonferroni method (for m pairs)

$$\text{CD} = t_{\alpha/2m, N-a}\sqrt{\text{MS}_E(1/n_i + 1/n_j)}$$

control overall error rate for the m comparisons.

- Tukey's method (for all possible pairs)

$$\text{CD} = \frac{q_\alpha(a, N-a)}{\sqrt{2}}\sqrt{\text{MS}_E(1/n_i + 1/n_j)}$$

$q_\alpha(a, N-a)$ from studentized range distribution (Table VII). Control overall error rate (exact for balanced experiments). (Example 3.7).

Comparing Treatments with Control (Dunnett's Method)

1. Assume μ_1 is a control, and $\mu_2, \ldots, \mu_a$ are (new) treatments
2. Only interested in $a-1$ pairs: $\mu_2 - \mu_1, \ldots, \mu_a - \mu_1$
3. Compare $\mid \bar{y}_{i.} - \bar{y}_{1.} \mid$ to

$$\mathrm{CD} = d_\alpha(a-1, N-a)\sqrt{\mathrm{MS}_E(1/n_i + 1/n_1)}$$

where $d_\alpha(p, f)$ from Table VIII: critical values for Dunnett's test.

4. Remark: control overall error rate. Read Example 3.9

For pairwise comparison, which method should be preferred? LSD, Bonferroni, Tukey, Dunnett or others?

Tensile Strength Example

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

proc glm data=one;
  class percent;
  model strength=precent;

/* Construct CI for Treatment Means*/
means percent /alpha=.05 lsd clm;
means percent / alpha=.05 bon clm;

/* Pairwise Comparison*/
means percent /alpha=.05 lines lsd;
means percent /alpha=.05 lines bon;
means percent /alpha=.05 lines scheffe;
means percent /alpha=.05 lines tukey;
means percent /dunnett;
run;
```

The GLM Procedure

t Confidence Intervals for y

Alpha	0.05
Error Degrees of Freedom	20
Error Mean Square	8.06
Critical Value of t	2.08596
Half Width of Confidence Interval	2.648434

trt	N	Mean	95% Confidence Limits	
30	5	21.600	18.952	24.248
25	5	17.600	14.952	20.248
20	5	15.400	12.752	18.048
35	5	10.800	8.152	13.448
15	5	9.800	7.152	12.448

The GLM Procedure

Bonferroni t Confidence Intervals for y

Alpha	0.05
Error Degrees of Freedom	20
Error Mean Square	8.06
Critical Value of t	2.84534
Half Width of Confidence Interval	3.612573

trt	N	Mean	Simultaneous 95% Confidence Limits	
30	5	21.600	17.987	25.213
25	5	17.600	13.987	21.213
20	5	15.400	11.787	19.013
35	5	10.800	7.187	14.413
15	5	9.800	6.187	13.413

t Tests (LSD) for y

NOTE: This test controls the Type I comparisonwise error rate, not the experimentwise error rate.

Alpha	0.05
Error Degrees of Freedom	20
Error Mean Square	8.06
Critical Value of t	2.08596
Least Significant Difference	3.7455

Means with the same letter are not significantly different.

t Grouping	Mean	N	trt
A	21.600	5	30
B	17.600	5	25
B	15.400	5	20
C	10.800	5	35
C	9.800	5	15

Bonferroni (Dunn) t Tests for y

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	20
Error Mean Square	8.06
Critical Value of t	3.15340
Minimum Significant Difference	5.6621

Means with the same letter are not significantly different.

Bon Grouping		Mean	N	trt
	A	21.600	5	30
B	A	17.600	5	25
B	C	15.400	5	20
	C	10.800	5	35
	C	9.800	5	15

Scheffe's Test for y

NOTE: This test controls the Type I experimentwise error rate.

Alpha	0.05
Error Degrees of Freedom	20
Error Mean Square	8.06
Critical Value of F	2.86608
Minimum Significant Difference	6.0796

Means with the same letter are not significantly different.

Scheffe Grouping	Mean	N	trt
A	21.600	5	30
A			
B A	17.600	5	25
B			
B C	15.400	5	20
C			
C	10.800	5	35
C			
C	9.800	5	15

Tukey's Studentized Range (HSD) Test for y

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Error Degrees of Freedom	20
Error Mean Square	8.06
Critical Value of Studentized Range	4.23186
Minimum Significant Difference	5.373

Means with the same letter are not significantly different.

Tukey Grouping		Mean	N	trt
	A	21.600	5	30
	A			
B	A	17.600	5	25
B				
B	C	15.400	5	20
	C			
D	C	10.800	5	35
D				
D		9.800	5	15

Dunnett's t Tests for y

NOTE: This test controls the Type I experimentwise error for comparisons of all treatments against a control.

Alpha	0.05
Error Degrees of Freedom	20
Error Mean Square	8.06
Critical Value of Dunnett's t	2.65112
Minimum Significant Difference	4.7602

Comparisons significant at the 0.05 level are indicated by ***.

trt Comparison	Difference Between Means	Simultaneous 95% Confidence Limits		
30 - 15	11.800	7.040	16.560	***
25 - 15	7.800	3.040	12.560	***
20 - 15	5.600	0.840	10.360	***
35 - 15	1.000	-3.760	5.760	