BIOS 662 Fall 2018

Analysis of Variance, Part III

David Couper, Ph.D.

david_couper@unc.edu

or

couper@bios.unc.edu

https://sakai.unc.edu/portal

Outline

- Diagnostics
- Nonparametric alternative: Kruskal-Wallis

ANOVA: Diagnostics

- Diagnostics discussed in section 10.6 of the text
- Assumptions

  1. Homogeneity of variance
  2. Normality of residual error
  3. Independence of residual error
  4. Linearity

ANOVA: Diagnostics

- Homogeneity of variance
  - Inspect plot of raw data or standard deviations by group means
  - Hartley's and Cochran's test

$$F_{\mathrm{MAX}} = \frac{s^2_{\max}}{s^2_{\min}}, \; C = \frac{s^2_{\max}}{\sum s^2_i}$$

    Tables are given in the Web appendix of the text as "Maximum F Tables" and "Cochran Test Tables", respectively
  - These tests require equal sample size and are sensitive to the normality assumption

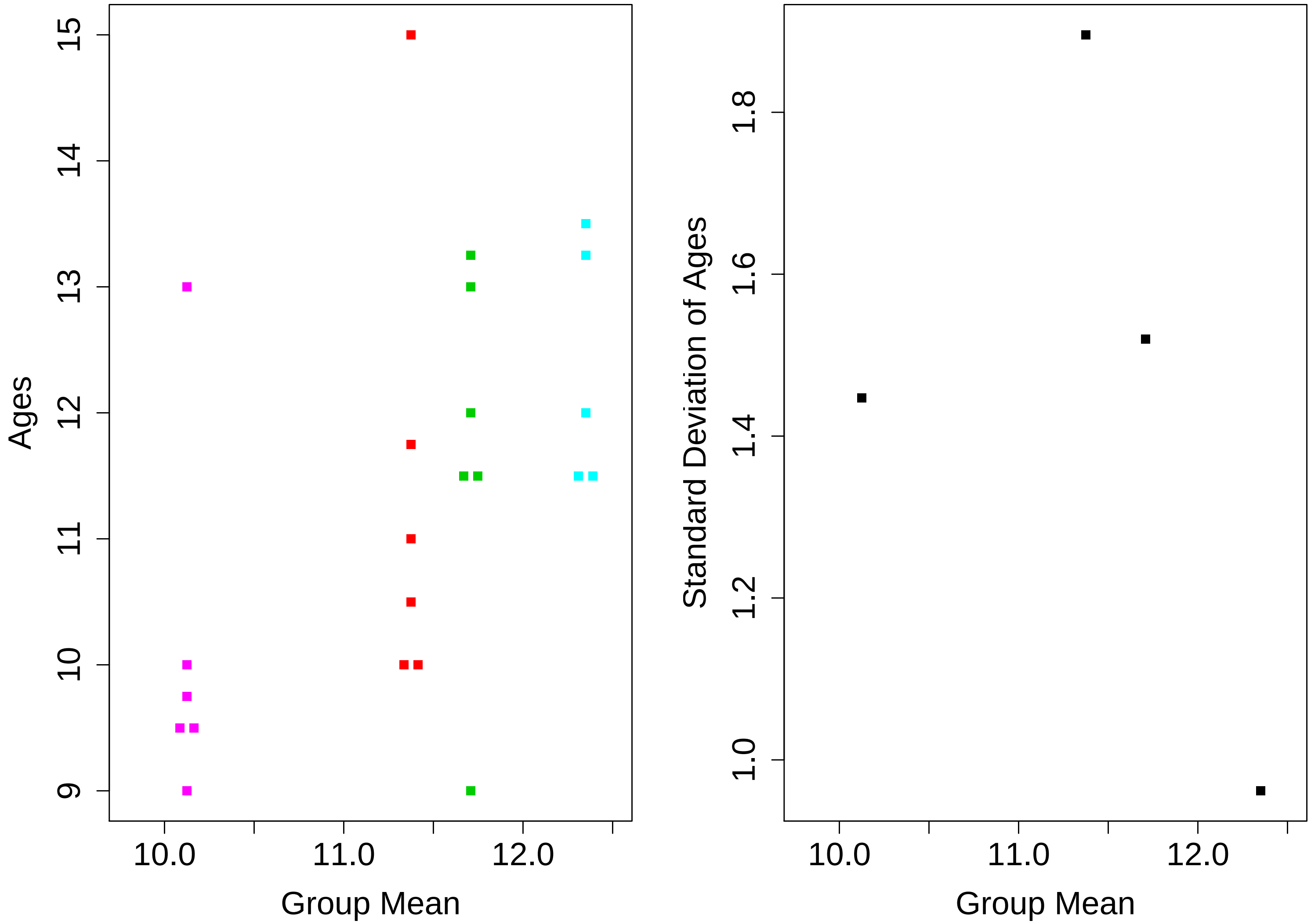
ANOVA: Diagnostics

- Homogeneity of variance
  - Modified Levene test (Brown-Forsythe test): apply ANOVA to the absolute deviations from group medians

    $$d_{ij} = |Y_{ij} - \tilde{Y}_{i\cdot}|$$

    use usual F test; rejection indicates lack of homogeneity (Ordinary Levene test uses means, not medians)
  - Robust to normality; does not require equal sample sizes
  - Cf. Chapter 18.2 of Kutner et al. *Applied Linear Statistical Models*, 5th Edition, 2005

Homogeneity of Variance Plot



Modified Levene Test: SAS

```
proc anova;  class group;  model age=group;  means group/hovtest=bf;
```

```
                          The ANOVA Procedure

          Brown and Forsythe's Test for Homogeneity of age Variance
               ANOVA of Absolute Deviations from Group Medians

                             Sum of         Mean
     Source          DF     Squares       Square    F Value    Pr > F

     group            3      0.8003       0.2668       0.19    0.9001
     Error           19     26.3125       1.3849

            Level of          -------------age-------------
            group        N             Mean          Std Dev

            active       6       10.1250000       1.44697961
            eight        5       12.3500000       0.96176920
            no           6       11.7083333       1.52000548
            passive      6       11.3750000       1.89571886
```

Modified Levene Test: R

```
Levene <- function(y, group)
{
    group <- as.factor(group)  # precautionary
    medians <- tapply(y, group, median)
    resp <- abs(y - medians[group])
    anova(lm(resp ~ group))[1, 4:5]
}

> Levene(age,group)
      F value Pr(>F)
group  0.1926 0.9001


# Changing anova(lm(resp ~ group))[1, 4:5] to anova(lm(resp ~ group))

> Levene(age,group)
Analysis of Variance Table

Response: resp
          Df  Sum Sq Mean Sq F value Pr(>F)
group      3  0.8003  0.2668  0.1926 0.9001
Residuals 19 26.3125  1.3849
```

ANOVA: Diagnostics for Normality

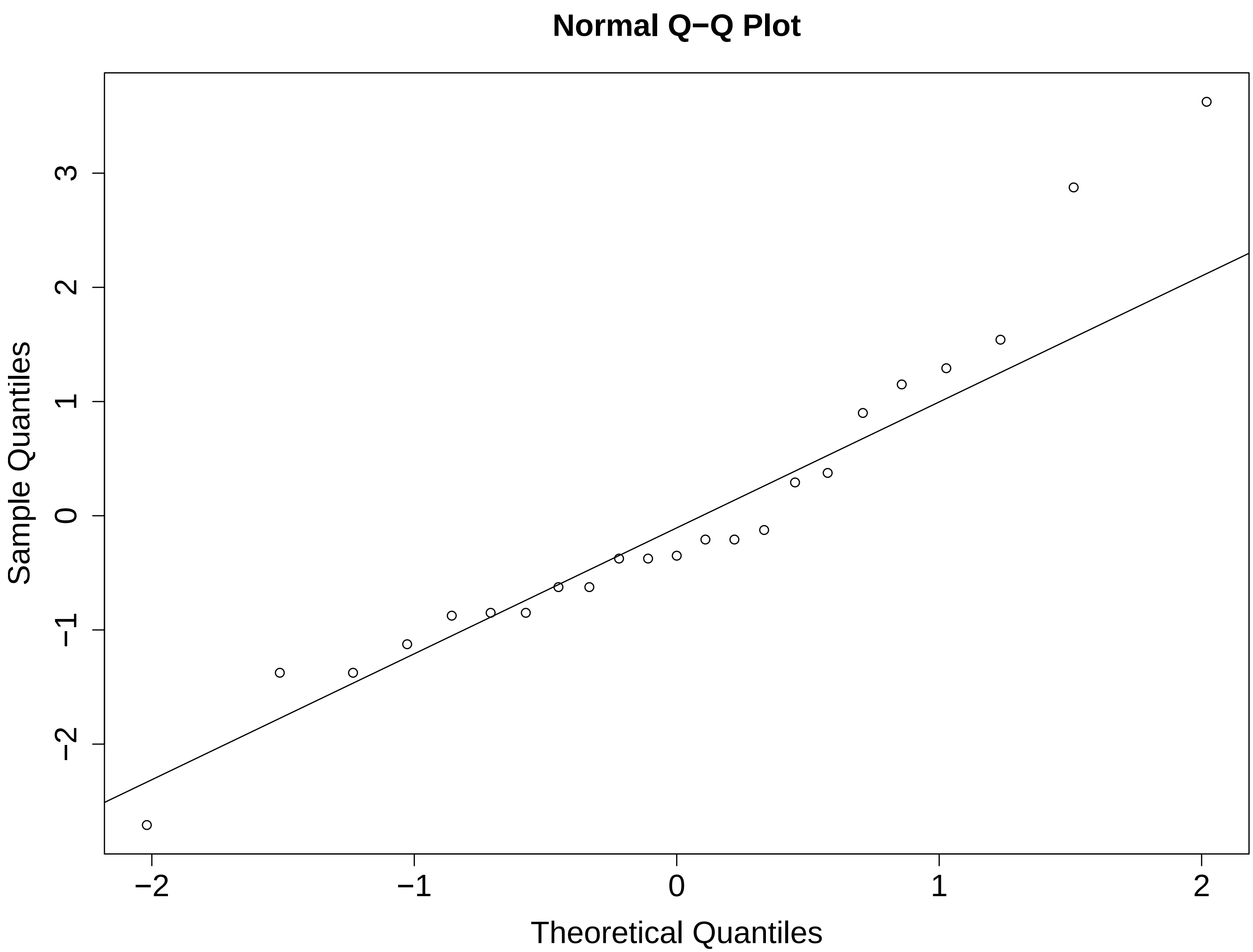
- QQ plot
- K-S GOF test
- Pearson correlation coefficient test:
  - Ordered residuals and expected values under normality
  - Assumption of normality in question if observed correlation is less than or equal to the critical value on the next page

ANOVA: Diagnostics for Normality

- Critical values for $\alpha = 0.05$

N	Crit. val.	N	Crit. val.	N	Crit. val.
5	0.88	10	0.92	24	0.96
6	0.89	12	0.93	30	0.96
7	0.90	15	0.94	40	0.97
8	0.91	20	0.95	50	0.98
9	0.91	22	0.95	100	0.99

ANOVA: Diagnostics for Normality



ANOVA: Diagnostics for Normality in R

```
> group <- as.factor(group)
> av <- aov(age ~ group)
> qq <- qqnorm(av$residuals)
> cor.test(qq$x,qq$y)

        Pearson's product-moment correlation

data:  qq$x and qq$y
t = 15.7572, df = 21, p-value = 4.146e-13
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.9070150 0.9832468
sample estimates:
      cor
0.9602173
```

ANOVA: Diagnostics

- Remedial measures

  1. Normality: appeal to CLT
  2. Transformations
     - Plot $(\bar{y}_{i\cdot}, s_i)$, $(\bar{y}_{i\cdot}, s_i^2)$, $(\bar{y}_{i\cdot}^2, s_i)$; linearity suggests $\log(y)$, $\sqrt{y}$, $1/y$ transformations, respectively
     - Box-Cox family: minimize SSE (that is, within group SS)
  3. Nonparametrics, e.g., Kruskal-Wallis

Box-Cox Transformations

- Family of transformations indexed by λ

$$Y_\lambda = \begin{cases} k_1(Y^\lambda - 1) & \text{for } \lambda \neq 0 \\ k_2 \log(Y) & \text{for } \lambda = 0 \end{cases}$$

where

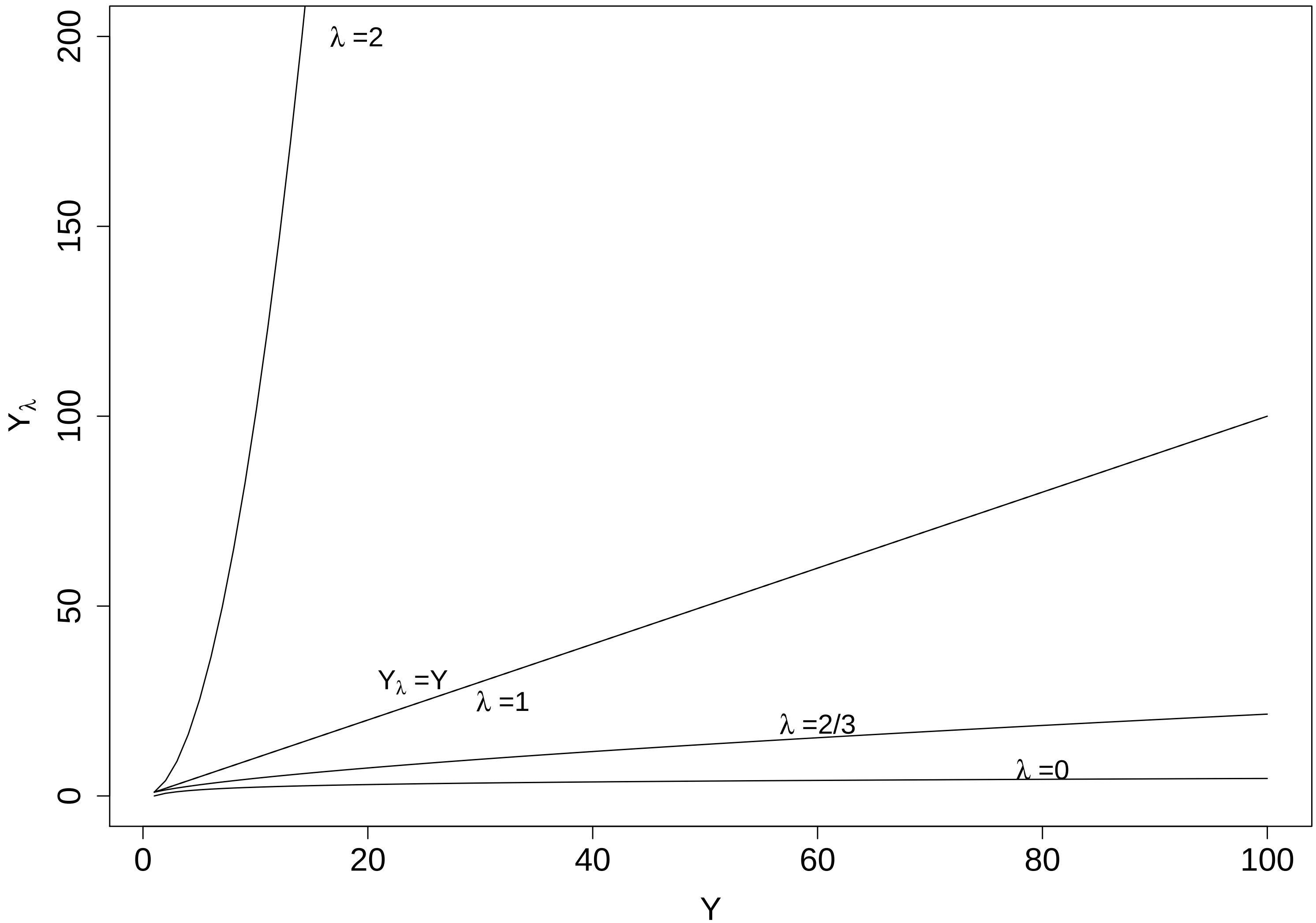
$$k_2 = \Big(\prod_{i,j} Y_{ij}\Big)^{1/N} \quad \text{and} \quad k_1 = \frac{1}{\lambda k_2^{\lambda - 1}}$$

- Choose λ that minimizes SSW

- SAS: macro on course website or proc transreg

  R: MASS library, function boxcox()

Box-Cox Transformations



Kruskal-Wallis

- Assume

$$Y_{ij} = \mu_i + \epsilon_{ij}$$

for $i = 1, \ldots, K$; $j = 1, \ldots, n_i$.

- ϵ_{ij} are independent and identically distributed with mean zero, but not necessarily normal

Kruskal-Wallis

- Same hypotheses

$$H_0 : \mu_1 = \cdots = \mu_K \quad \text{vs.} \quad H_A : \text{ at least one inequality}$$

- Pool all N observations and rank from smallest to largest
- Let R_{ij} be the rank of the j^{th} obs in the i^{th} group
- Let $\bar{R}_i = \sum_{j=1}^{n_i} R_{ij}/n_i$ equal the average rank in the i^{th} group
- Let $\bar{R}$ denote the overall average rank. What must this equal?

Kruskal-Wallis

- The Kruskal-Wallis test statistic is

$$T_{\text{KW}} = \frac{12\sum_{i=1}^{K} n_i(\bar{R}_i - \bar{R})^2}{N(N+1)}$$

- Equivalently

$$T_{\text{KW}} = \frac{12\sum_{i=1}^{K}(\sum_{j=1}^{n_i} R_{ij})^2/n_i}{N(N+1)} - 3(N+1)$$

- Reject H_0 for large values of T_{KW}

Kruskal-Wallis

- Under H_0, if the n_i are moderately large (rule of thumb: $n_i \geq 5$), then

$$T_{\mathrm{KW}} \sim \chi^2_{K-1}$$

- If the n_i are small, the exact distribution of T_{KW} can be computed

Kruskal-Wallis: Exact

- There are

$$\binom{N}{n_1 n_2 \cdots n_K} = \frac{N!}{n_1!\, n_2!\, n_3! \, \cdots \, n_K!}$$

possible ways to assign n_1 ranks to group 1, n_2 ranks to group 2, ...

- Under H_0 each occurs with equal probability

- Suppose $n_1 = 2$, $n_2 = n_3 = 1$. Then

$$\binom{N}{n_1 n_2 \cdots n_K} = \frac{4!}{2!\, 1!\, 1!} = 12$$

Kruskal-Wallis: Exact

R_{1j}	R_{2j}	R_{3j}	$\sum_i R_{i\cdot}^2/n_i$	T_{KW}
1 2	3	4	9/2+9+16=29.5	2.7
1 3	2	4	28	1.8
1 4	2	3	25.5	0.3
2 3	1	4	29.5	2.7
2 4	1	3	28	1.8
3 4	1	2	29.5	2.7

k	$\Pr[T_{\text{KW}} = k]$
0.3	1/6
1.8	1/3
2.7	1/2

Kruskal-Wallis with Ties

- If there are ties among the ranks, we use the midrank method as in the Wilcoxon tests
- The KW statistic adjusted for ties is:

$$T_{\text{KWadj}} = \frac{T_{\text{KW}}}{1 - \sum_{i=1}^{q}(t_i^3 - t_i)/(N^3 - N)}$$

  where q is the number of sets of tied observations and t_i is the number of observations in the i^{th} set

- T_{KWadj} will also be approximately χ^2_{K-1}

Kruskal-Wallis: Example

- A study was conducted to compared three doses of aspirin in the treatment of fever in children with the flu
- 15 children with a fever between 100.0 and 100.9 F were randomly assigned to each dose $(n_1 = n_2 = n_3 = 5;$ $N = 15)$
- Temperature was measured three hours later
- Let μ_i denote the mean temperature change for dose i
- $H_0 : \mu_1 = \mu_2 = \mu_3$

Kruskal-Wallis: Example cont.

- Distribution of T_{KW} (Owen 1962, page 422; Kruskal, Wallis 1952, JASA, Table 6.1)

k	$\Pr[T_{\text{KW}} \geq k]$
4.50	0.102
4.56	0.100
5.66	0.051
5.78	0.049
7.98	0.010
8.00	0.009

- $C_{0.05} = \{T_{\text{KW}} \geq 5.78\}$

Kruskal-Wallis: Example cont.

Low		Med		High	
ΔTemp	R	ΔTemp	R	ΔTemp	R
2.0	14	0.6	8	1.1	10
1.6	13	1.2	11	-1.0	1
2.1	15	0.5	7	-0.2	3
0.7	9	0.2	4	0.4	6
1.3	12	-0.4	2	0.3	5

- $R_{1\cdot} = 63, \;\; R_{2\cdot} = 32, \;\; R_{3\cdot} = 25$

Kruskal-Wallis: Example cont.

- Therefore

$$T_{\text{KW}} = \frac{12(63^2/5 + 32^2/5 + 25^2/5)}{15(16)} - 3(16) = 8.18$$

- Asymptotic p-value

$$\Pr[\chi^2_2 > 8.18] = 0.0167$$

- From Owen table, expect exact p-value < 0.009

Kruskal-Wallis: SAS

```
proc npar1way;  class dose;  var temp_change;  exact wilcoxon;
```

```
      Wilcoxon Scores (Rank Sums) for Variable temp_change
                   Classified by Variable dose

                      Sum of     Expected       Std Dev          Mean
dose          N       Scores     Under H0      Under H0         Score
-----------------------------------------------------------------
Low           5         63.0         40.0      8.164966         12.60
Medium        5         32.0         40.0      8.164966          6.40
High          5         25.0         40.0      8.164966          5.00


                       Kruskal-Wallis Test

              Chi-Square                         8.1800
              DF                                      2
              Asymptotic Pr >  Chi-Square        0.0167
              Exact      Pr >= Chi-Square        0.0081
```

Kruskal-Wallis: R

```
> kruskal.test(change,dose)

        Kruskal-Wallis rank sum test

data:  change and dose
Kruskal-Wallis chi-squared = 8.18, df = 2, p-value = 0.01674
```

Kruskal-Wallis

- Suppose we perform ANOVA with the Y_{ij} replaced by their ranks
- Resulting F test

$$F_{\mathrm{R}} = \frac{(N-K)T_{\mathrm{KW}}}{(K-1)(N-1-T_{\mathrm{KW}})}$$

- If $K=2$, the KW test is equivalent to the Wilcoxon rank sum test
- ARE is $3/\pi = 0.955$ compared to the F-test under normality
- For multiple comparisons of means, use Wilcoxon rank sum tests with Bonferroni correction