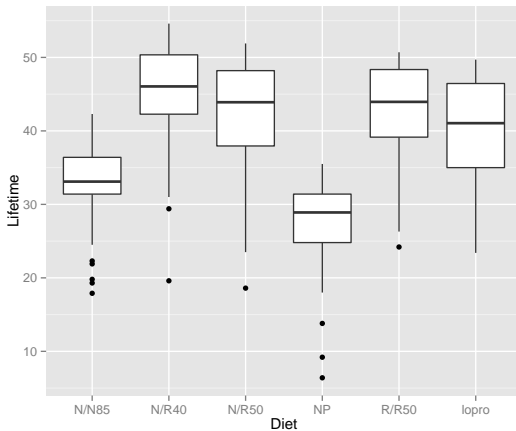# STAT 401A - Statistical Methods for Research Workers
One-way ANOVA (contrasts and multiple comparisons)

Jarad Niemi (Dr. J)

Iowa State University

last updated: October 8, 2014

Mice lifetimes

Simple hypothesis

Consider the one-way ANOVA model: $Y_{ij} \sim N(\mu_j, \sigma^2)$ where $j = 1, \ldots, J$.

Here are a few simple alternative hypotheses:

1. Mean lifetimes for N/R50 and R/R50 diet are different.
2. Mean lifetimes for N/R40 is different than for N/R50 and R/R50 combined.
3. Mean lifetimes for high calorie (NP and N/N85) diets is different than for low calorie diets combined.

$H_0 : \gamma = 0 \qquad H_1 : \gamma \neq 0 :$

$$\begin{aligned} \gamma_1 &= \mu_{R/R50} - \mu_{N/R50} \\ \gamma_2 &= \mu_{N/R40} - \tfrac{1}{2}(\mu_{N/R50} + \mu_{R/R50}) \\ \gamma_3 &= \tfrac{1}{4}(\mu_{N/R50} + \mu_{R/R50} + \mu_{N/R40} + \mu_{lopro}) - \tfrac{1}{2}(\mu_{NP} + \mu_{N/N85}) \end{aligned}$$

Contrasts

Definition
A linear combination of group means has the form

$$\gamma = C_1\mu_1 + C_2\mu_2 + \ldots + C_J\mu_J$$

where C_j are known coefficients and μ_j are the unknown population means.

Definition
A linear combination with $C_1 + C_2 + \cdots + C_J = 0$ is a contrast.

Remark Contrast interpretation is usually best if
$|C_1| + |C_2| + \cdots + |C_J| = 2$, i.e. the positive sum to 1 and the negative coefficients sum to -1.

Inference on contrasts

$$\gamma = C_1\mu_1 + C_2\mu_2 + \cdots + C_J\mu_J$$

Estimated by

$$g = C_1\overline{Y}_1 + C_2\overline{Y}_2 + \cdots + C_J\overline{Y}_J$$

with standard error

$$SE(g) = s_p\sqrt{\frac{C_1^2}{n_1} + \frac{C_2^2}{n_2} + \cdots + \frac{C_J^2}{n_J}}$$

t-statistic (compare to t_{n-J}) and CI:

$$t = \frac{g}{SE(g)} \qquad g \pm t_{n-J}(1 - \alpha/2)SE(g)$$

Contrasts for mice lifetime dataset

For these contrasts:

1. Mean lifetimes for N/R50 and R/R50 diet are different.
2. Mean lifetimes for N/R40 is different than for N/R50 and R/R50 combined.
3. Mean lifetimes for high calorie (NP and N/N85) diets is different than for low calorie diets combined.

$H_0 : \gamma = 0 \qquad H_1 : \gamma \neq 0 :$

$$\gamma_1 = \mu_{R/R50} - \mu_{N/R50}$$
$$\gamma_2 = \mu_{N/R40} - \frac{1}{2}(\mu_{N/R50} + \mu_{R/R50})$$
$$\gamma_3 = \frac{1}{4}(\mu_{N/R50} + \mu_{R/R50} + \mu_{N/R40} + \mu_{lopro}) - \frac{1}{2}(\mu_{NP} + \mu_{N/N85})$$

	N/N85	N/R40	N/R50	NP	R/R50	lopro
early rest - none @ 50kcal	0.00	0.00	-1.00	0.00	1.00	0.00
40kcal/week - 50kcal/week	0.00	1.00	-0.50	0.00	-0.50	0.00
lo cal - hi cal	-0.50	0.25	0.25	-0.50	0.25	0.25

Mice liftime examples

	Diet	n	mean	sd
1	N/N85	57	32.69	5.13
2	N/R40	60	45.12	6.70
3	N/R50	71	42.30	7.77
4	NP	49	27.40	6.13
5	R/R50	56	42.89	6.68
6	lopro	56	39.69	6.99

Contrasts:

	g	SE(g)	t	p	L	U
early rest - none @ 50kcal	0.59	1.19	0.49	0.62	-1.76	2.94
40kcal/week - 50kcal/week	2.53	1.05	2.41	0.02	0.46	4.59
lo cal - hi cal	12.45	0.78	15.96	0.00	10.92	13.98

```
DATA case0501;
   INFILE 'case0501.csv' DSD FIRSTOBS=2;
   INPUT lifetime diet $ ;

PROC MEANS DATA=case0501;
   CLASS diet;
   VAR lifetime;
   RUN;
```

                              The MEANS Procedure
                          Analysis Variable : lifetime

diet	N Obs	N	Mean	Std Dev	Minimum	Maximum
N/N85	57	57	32.6912281	5.1252972	17.9000000	42.3000000
N/R40	60	60	45.1166667	6.7034058	19.6000000	54.6000000
N/R50	71	71	42.2971831	7.7681947	18.6000000	51.9000000
NP	49	49	27.4020408	6.1337010	6.4000000	35.5000000
R/R50	56	56	42.8857143	6.6831519	24.2000000	50.7000000
lopro	56	56	39.6857143	6.9916945	23.4000000	49.7000000

```
PROC GLM;
  CLASS diet;
  MODEL lifetime = diet / CLPARM;
  ESTIMATE 'early rest - none @ 50kcal' diet  0  1 -1  0  0  0 ;
  ESTIMATE '40kcal/week - 50kcal/week'  diet  0  2 -1  0 -1  0 / DIVISOR = 2;
  ESTIMATE 'lo cal - hi cal'            diet -2  1  1 -2  1  1 / DIVISOR = 4 ;
  RUN;
  QUIT;
```

The GLM Procedure

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	12733.94181	2546.78836	57.10	<.0001
Error	343	15297.41532	44.59888		
Corrected Total	348	28031.35713			

Parameter	Estimate	Standard Error	t Value	Pr > |t|
early rest - none @ 50kcal	0.5885312	1.19355007	0.49	0.6223
40kcal/week - 50kcal/week	2.5252180	1.04854904	2.41	0.0166
lo cal - hi cal	12.4496851	0.78001425	15.96	<.0001

Parameter	95% Confidence Limits	
early rest - none @ 50kcal	-1.7590676	2.9361299
40kcal/week - 50kcal/week	0.4628224	4.5876136
lo cal - hi cal	10.9154718	13.9838985

```
library(multcomp)
m = lm(Lifetime~Diet-1, case0501) # The -1 indicates no intercept (see Ch 7)
summary(m)




Call:
lm(formula = Lifetime ~ Diet - 1, data = case0501)

Residuals:
    Min      1Q  Median      3Q     Max
-25.517  -3.386   0.814   5.183  10.014

Coefficients:
          Estimate Std. Error t value Pr(>|t|)
DietN/N85   32.691      0.885    37.0   <2e-16 ***
DietN/R40   45.117      0.862    52.3   <2e-16 ***
DietN/R50   42.297      0.793    53.4   <2e-16 ***
DietNP      27.402      0.954    28.7   <2e-16 ***
DietR/R50   42.886      0.892    48.1   <2e-16 ***
Dietlopro   39.686      0.892    44.5   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6.68 on 343 degrees of freedom
Multiple R-squared: 0.972,  Adjusted R-squared:  0.972
F-statistic: 2.01e+03 on 6 and 343 DF,  p-value: <2e-16


K


                        N/N85 N/R40 N/R50   NP R/R50 lopro
early rest - none @ 50kcal   0.0  0.00 -1.00  0.0  1.00  0.00
40kcal/week - 50kcal/week    0.0  1.00 -0.50  0.0 -0.50  0.00
lo cal - hi cal             -0.5  0.25  0.25 -0.5  0.25  0.25
```

```
t = glht(m, linfct=K)
summary(t)


 Simultaneous Tests for General Linear Hypotheses

Fit: lm(formula = Lifetime ~ Diet - 1, data = case0501)

Linear Hypotheses:
                           Estimate Std. Error t value Pr(>|t|)
early rest - none @ 50kcal == 0     0.589      1.194    0.49   0.946
40kcal/week - 50kcal/week == 0      2.525      1.049    2.41   0.049 *
lo cal - hi cal == 0               12.450      0.780   15.96   <1e-04 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Adjusted p values reported -- single-step method)


confint(t, calpha=univariate_calpha())


 Simultaneous Confidence Intervals

Fit: lm(formula = Lifetime ~ Diet - 1, data = case0501)

Quantile = 1.967
95% confidence level


Linear Hypotheses:
                           Estimate lwr      upr
early rest - none @ 50kcal == 0  0.589   -1.759   2.936
40kcal/week - 50kcal/week == 0   2.525    0.463   4.588
lo cal - hi cal == 0            12.450   10.915  13.984
```

Summary

- Contrasts are linear combinations that sum to zero
- t-test tools are used to calculate pvalues and confidence intervals

SAS code and output for one-way ANOVA

```
DATA mice;
  INFILE 'case0501.csv' DSD FIRSTOBS=2;
  INPUT lifetime diet $;

PROC GLM DATA=mice;
  CLASS diet;
  MODEL lifetime = diet;
  LSMEANS diet / ADJUST=T;
  RUN;
```

                          The GLM Procedure

Dependent Variable: lifetime

| | | Sum of | | | |
Source	DF	Squares	Mean Square	F Value	Pr > F
Model	5	12733.94181	2546.78836	57.10	<.0001
Error	343	15297.41532	44.59888		
Corrected Total	348	28031.35713			

SAS code and output for pairwise comparisons

```
                        The GLM Procedure
                       Least Squares Means

                              lifetime        LSMEAN
                  diet          LSMEAN         Number

                  N/N85       32.6912281            1
                  N/R40       45.1166667            2
                  N/R50       42.2971831            3
                  NP          27.4020408            4
                  R/R50       42.8857143            5
                  lopro       39.6857143            6


            Least Squares Means for effect diet
            Pr > |t| for H0: LSMean(i)=LSMean(j)

                   Dependent Variable: lifetime
```

i/j	1	2	3	4	5	6
1		<.0001	<.0001	<.0001	<.0001	<.0001
2	<.0001		0.0166	<.0001	0.0731	<.0001
3	<.0001	0.0166		<.0001	0.6223	0.0293
4	<.0001	<.0001	<.0001		<.0001	<.0001
5	<.0001	0.0731	0.6223	<.0001		0.0117
6	<.0001	<.0001	0.0293	<.0001	0.0117	

```
NOTE: To ensure overall protection level, only probabilities associated with pre-planned
      comparisons should be used.
```

Statistical testing errors

Definition
A type I error occurs when a true null hypothesis is rejected.

Definition
A type II error occurs when a false null hypothesis is not rejected. Power is one minus the type II error probability.

Remark We set α to control the type I error probability. If we set $\alpha = 0.05$, then we will incorrectly reject a true null hypothesis 5% of the time.

Definition
The familywise error rate is the probability of rejecting at least one true null hypothesis.

Type I error for all pairwise comparisons of J groups

How many combinations when choosing 2 items out of J?

$$\binom{J}{2} = \frac{J!}{2!(J-2)!}.$$

If $J = 6$, then there are 15 different comparison of means. If we set $\alpha = 0.05$ as our significance level, then individually each test will only incorrectly reject 5% of the time.

If we have 15 tests and use $\alpha = 0.05$, what is the familywise error rate?

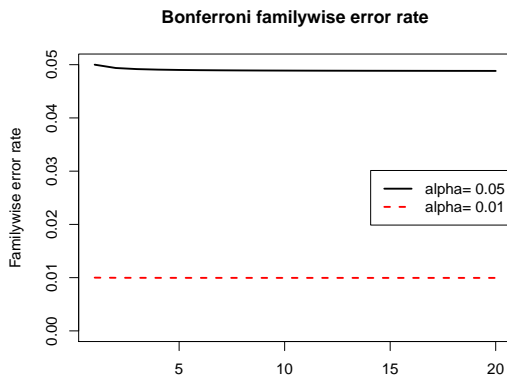$$1 - (1 - 0.05)^{15} = 1 - (0.95)^{15} = 1 - 0.46 = 0.54$$

So there is a greater than 50% probability of falsely rejecting a true null hypothesis!

Bonferroni correction

Definition

If we do m tests and want the familywise error rate to be α, the Bonferroni correction uses α/m for each individual test. The familywise error rate, for independent tests, is $1 - (1 - \alpha/, m)^m$.

Bonferroni familywise error rate

SAS code and output for pairwise comparisons

Compare the unadjusted pvalues to $\alpha/15 = 0.05/15 = 0.0033$.

```
                      Least Squares Means for effect diet
                      Pr > |t| for H0: LSMean(i)=LSMean(j)

                          Dependent Variable: lifetime

 i/j        1           2           3           4           5           6

  1                   <.0001      <.0001      <.0001      <.0001      <.0001
  2     <.0001                    0.0166      <.0001      0.0731      <.0001
  3     <.0001      0.0166                    <.0001      0.6223      0.0293
  4     <.0001      <.0001      <.0001                    <.0001      <.0001
  5     <.0001      0.0731      0.6223      <.0001                    0.0117
  6     <.0001      <.0001      0.0293      <.0001      0.0117
```

NOTE: To ensure overall protection level, only probabilities associated with pre-planned
comparisons should be used.

Now 2-3, 3-6, and 5-6 are no longer significant.

SAS code and output for one-way ANOVA

If you use SAS to do the adjustment, compare pvalues to $\alpha = 0.05$.

```
DATA mice;
  INFILE 'case0501.csv' DSD FIRSTOBS=2;
  INPUT lifetime diet $;

PROC GLM DATA=mice;
  CLASS diet;
  MODEL lifetime = diet;
  LSMEANS diet / ADJUST=BON;
  RUN;
```

```
                    Least Squares Means for effect diet
                    Pr > |t| for H0: LSMean(i)=LSMean(j)

                        Dependent Variable: lifetime

   i/j          1          2          3          4          5          6

     1                 <.0001     <.0001     0.0009     <.0001     <.0001
     2      <.0001                0.2488     <.0001     1.0000     0.0002
     3      <.0001     0.2488                <.0001     1.0000     0.4402
     4      0.0009     <.0001     <.0001                <.0001     <.0001
     5      <.0001     1.0000     1.0000     <.0001                0.1751
     6      <.0001     0.0002     0.4402     <.0001     0.1751
```

Comments on the Bonferroni correction

Remark The Bonferroni correction can be used in any situation. In particular, it can be used on unadjusted pvalues reported in an article that has many tests by comparing their pvalues to α/m where m is the number of tests they perform.

Remark The Bonferroni correction is (in general) the most conservative multiple comparison adjustment we will discuss, i.e. it will lead to the least null hypothesis rejections.

Constructing multiple confidence intervals

Remark A $100(1 - \alpha)\%$ confidence interval should contain the true value $100(1 - \alpha)\%$ of the time.

Remark An error occurs if the confidence interval does not contain the true value.

Just like the Type I error and familywise error rate, we can ask what is the probability at least one confidence interval does not cover the true value.

The procedures we will talk about for confidence intervals have equivalent approaches for hypothesis testing (pvalues). Within these procedures we still have the equivalence between pvalues and CIs.

Constructing multiple confidence intervals

Confidence interval for the difference between group j and group j':

$$\overline{Y}_j - \overline{Y}_{j'} \pm M\, s_p \sqrt{\frac{1}{n_j} + \frac{1}{n_{j'}}}$$

where M is a multiplier that depends on the adjustment procedure:

Procedure	M	Use
LSD	$t_{n-J}(1 - \alpha/2)$	After significant F-test
		(no adjustment)
Dunnett	multivariate t	Compare all groups to control
Tukey-Kramer	$q_{J,n-J}(1 - \alpha/2)/\sqrt{2}$	All pairwise comparisons
Scheffé	$\sqrt{(J-1)F_{(J-1,n-J)}(1 - \alpha)}$	All contrasts
Bonferroni	$t_{n-J}(1 - \alpha/2m)$	m tests
		(most generic)

SAS code and output for one-way ANOVA

```
DATA mice;
  INFILE 'case0501.csv' DSD FIRSTOBS=2;
  INPUT lifetime diet $;

PROC GLM DATA=mice;
  CLASS diet;
  MODEL lifetime = diet;
  LSMEANS diet / CL ADJUST=TUKEY;
  RUN;
```

```
                          The GLM Procedure
                        Least Squares Means
            Adjustment for Multiple Comparisons: Tukey-Kramer


                       Dependent Variable: lifetime


   i/j          1           2           3           4           5           6
     1                  <.0001      <.0001      0.0008      <.0001      <.0001
     2      <.0001                  0.1565      <.0001      0.4684      0.0002
     3      <.0001      0.1565                  <.0001      0.9964      0.2460
     4      0.0008      <.0001      <.0001                  <.0001      <.0001
     5      <.0001      0.4684      0.9964      <.0001                  0.1168
     6      <.0001      0.0002      0.2460      <.0001      0.1168
```

SAS code and output for one-way ANOVA

```
                    The GLM Procedure
                  Least Squares Means
      Adjustment for Multiple Comparisons: Tukey-Kramer


            Least Squares Means for Effect diet

              Difference          Simultaneous 95%
               Between          Confidence Limits for
       i    j    Means          LSMean(i)-LSMean(j)
       1    2  -12.425439     -15.965442     -8.885435
       1    3   -9.605955     -13.009741     -6.202169
       1    4    5.289187       1.560626      9.017749
       1    5  -10.194486     -13.795557     -6.593416
       1    6   -6.994486     -10.595557     -3.393416
       2    3    2.819484      -0.536769      6.175736
       2    4   17.714626      14.029406     21.399846
       2    5    2.230952      -1.325223      5.787128
       2    6    5.430952       1.874777      8.987128
       3    4   14.895142      11.340571     18.449714
       3    5   -0.588531      -4.009133      2.832070
       3    6    2.611469      -0.809133      6.032070
       4    5  -15.483673     -19.227592    -11.739755
       4    6  -12.283673     -16.027592     -8.539755
       5    6    3.200000      -0.416969      6.816969
```

False Discovery Rate

Remark Not wanting to make a single mistake is pretty conservative.

In high-throughput fiels a more common multiple comparison adjustment is false discovery rate.

Definition
False discovery rate procedures try to control the expected proportion of incorrectly rejected null hypotheses.

How to incorporate multiple comparison adjustments

1. Determine what tests are going to be run (before looking at the data) or what confidence intervals are going to be constructed.
2. Determine which multiple comparison adjustment is the most relevant
3. Use/state that adjustment and interpret your results