Outline

One-way ANOVA

ANOVA = ANalysis Of VAriance.

to compare the means of any # of treatments (2 or more)

extends the 2 independent samples t-test assuming equal variances.

Key idea: break up the variation, i.e. sum of squares

$$\sum (y_i - \bar{y})^2$$

into variation explained by differences among treatments and variation within treatments.

First reconsider the independent two-sample case, then generalize the idea to independent multiple samples.

Two independent samples: simple example

X: 4, 12, 8 on drug A, and Y: 17, 8, 11 on drug B.

Summary statistics:

$\bar{x} = 8, \quad \sum_{i=1}^{3}(x_i - \bar{x})^2 = 32,$ so $s_x^2 =$

$\bar{y} = 12, \quad \sum_{i=1}^{3}(y_i - \bar{y})^2 = 42,$ so $s_y^2 =$, $s_p^2 =$ $= 18.5$

T-test for $H_0 : \mu_1 = \mu_2$ vs. $H_A : \mu_1 \neq \mu_2$:

$t =$ $= 1.14$ on df $=$

p-value: $2\,\mathbb{P}\{T \geq 1.14\} > 0.10$. No evidence against H_0.

Next: ANOVA with same data: partition the variation.

Sums of squares (SS)

Total SS: Total variation. Pretend all obs. form single sample.
Overall mean: $= 10$ and SSTotal is

$= 98$

on df =

Treatment SS: amount of the total variation explained by differences between groups. Replace each observation by its group mean.

`X: 8, 8, 8` and `Y: 12, 12, 12`

Overall mean: still
and SSTrt is

$= 24$

on df = 1.

Sums of squares (SS)

Error SS: amount of the total variation explained by differences within each group:

$= 74$

on df = 4.

note: SSError / dfError = $= s_p^2$.

also:

$$\text{SSTotal} = \text{SSTrt} + \text{SSError} \quad (\qquad)$$
$$\text{df Total} = \text{df Trt} + \text{df Error} \quad (\qquad)$$

An ANOVA table summarizes the information. Here MS = Mean Square = SS/df

Source	df	SS	MS
Trt	1	24	24
Error	4	74	18.5
Total	5	98	–

F-test

1. $H_0 : \mu_1 = \mu_2$ vs $H_A : \mu_1 \neq \mu_2$
2. If H_0 is true, then

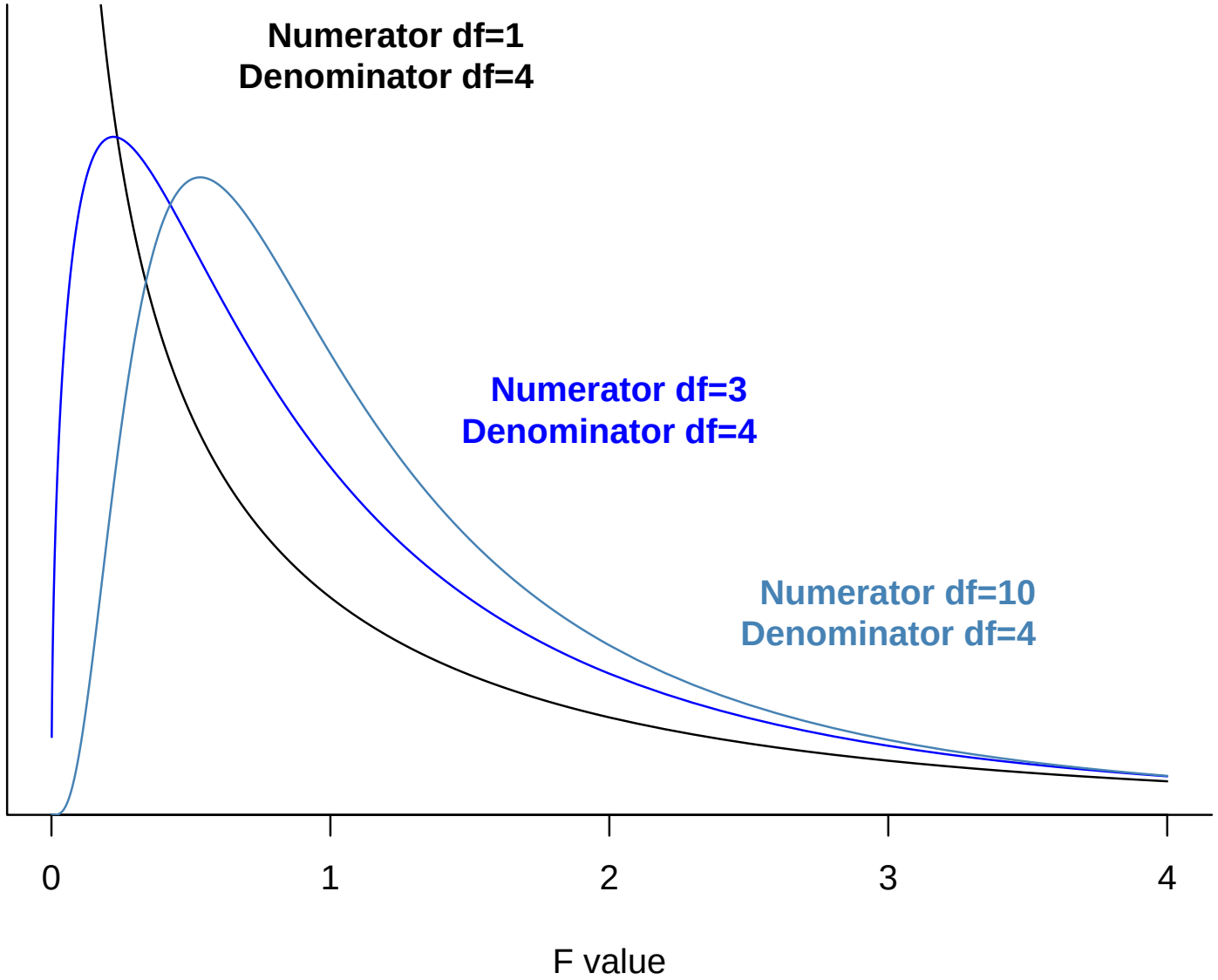
$$F = \frac{\text{MSTrt}}{\text{MSError}} \sim \mathcal{F}_{\text{dfTrt, dfError}}$$

3. In the example, the observed $f =$ $= 1.30$. Compare this to an F-distribution with 1 df numerator and 4 df denominator using Table D. p-value: $\mathbb{P}\{F_{1,4} \geq 1.30\} > 0.10$.
4. No evidence against H_0. Do not reject H_0 at the 10% level.

Note: $1.30 = (1.14)^2$ i.e $f = t^2$. This is special to 2 groups: ANOVA = t-test when only 2 groups.

The p-value is from one tail of the F distribution, even though H_A is two-sided.

The F distribution



F-values are always ≥ 0

The F-distributions is located around 1, for all df's.

Recap

SSTotal: total variation
SSTrt : variation due to treatment differences
SSErr : residual variation, within treatment groups

$$F = \frac{\text{SSTrt}/\text{dfTrt}}{\text{SSErr}/\text{dfErr}}$$

Small difference between group means relative to variability $\rightarrow$ $f \rightarrow$ p-value, and accept H_0.

Large difference between group means relative to variability $\rightarrow$ $f \rightarrow$ p-value, and reject H_0.

Russell *et al.* (2007) Science 317:941-943

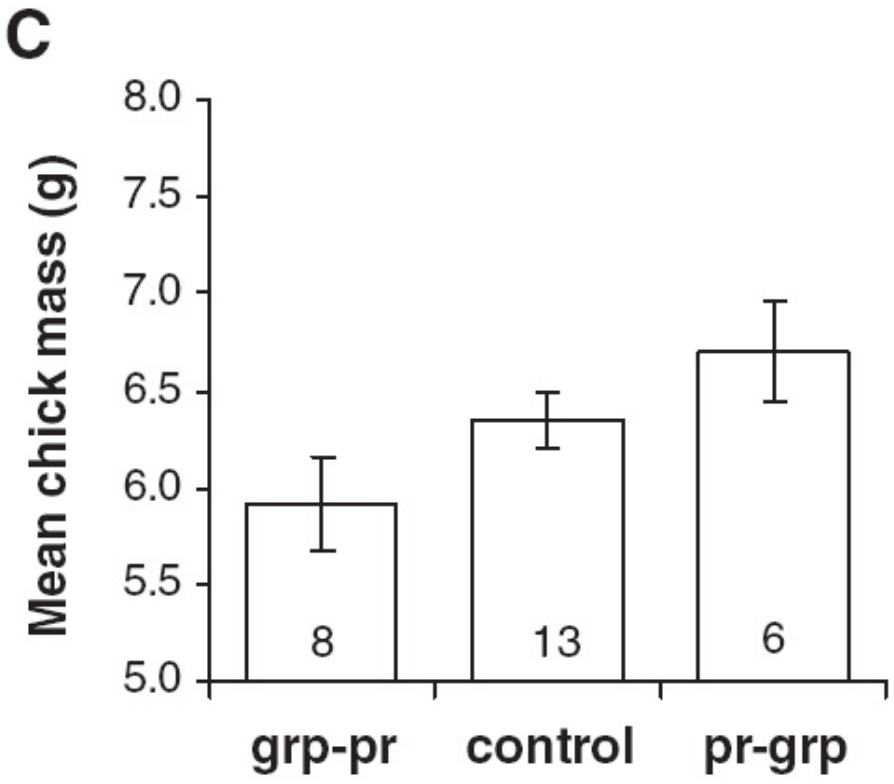
Reduced Egg Investment Can Conceal Helper Effects in Cooperatively Breeding Birds

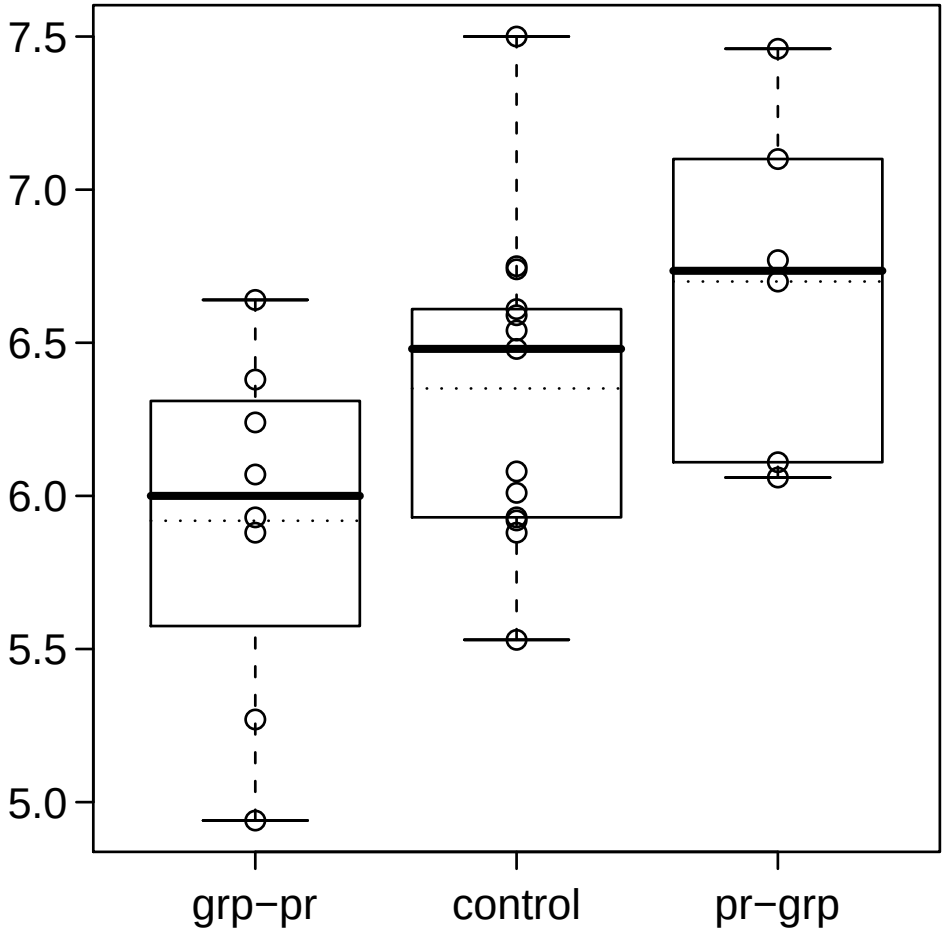
A. F. Russell,[1,2]*† N. E. Langmore,[3] A. Cockburn,[3,4] L. B. Astheimer,[5] R. M. Kilner[6]*

Cooperative breeding systems are characterized by nonbreeding helpers that assist breeders in offspring care. However, the benefits to offspring of being fed by parents and helpers in cooperatively breeding birds can be difficult to detect. We offer experimental evidence that helper effects can be obscured by an undocumented maternal tactic. In superb fairy-wrens (*Malurus cyaneus*), mothers breeding in the presence of helpers lay smaller eggs of lower nutritional content that produce lighter chicks, as compared with those laying eggs in the absence of helpers. Helpers compensate fully for such reductions in investment and allow mothers to benefit through increased survival to the next breeding season. We suggest that failure to consider maternal egg-investment strategies can lead to underestimation of the force of selection acting on helping in avian cooperative breeders.

Russell *et al.* (2007) Science 317:941-943

Mass (g) of chicks, 6-8 days after hatching.
`grp-pr`: laid in groups, reared in pairs.
`control`: laid and reared by their own parents (and helpers)
`pr-grp`: laid in paired, reared in groups





Chick mass

	Parenting		
grp-pr	control	pr-grp	
6.24	5.53	6.77	
4.94	6.74	6.11	
5.27	6.61	6.06	
5.93	7.50	6.70	
5.88	6.59	7.10	
6.38	6.01	7.46	
6.07	6.54		
6.64	5.92		
	5.93		
	6.48		
	5.88		
	6.75		
	6.08		
sum	47.36	82.55	40.2	170.11
mean	5.92	6.35	6.70	6.30

k independent samples

k treatments, n_i observations for treament i.

Trt	1	2	$\cdots$	k	
Obs	y_{11}	y_{21}	$\cdots$	y_{k1}	
	y_{12}	y_{22}	$\cdots$	y_{k2}	
	$\vdots$	$\vdots$		$\vdots$	
	y_{1n_1}	y_{2n_2}	$\cdots$	y_{kn_k}	
Sum	$y_{1\cdot}$	$y_{2\cdot}$	$\cdots$	$y_{k\cdot}$	$y_{\cdot\cdot}$
Mean	$\bar{y}_{1\cdot}$	$\bar{y}_{2\cdot}$	$\cdots$	$\bar{y}_{k\cdot}$	$\bar{y}_{\cdot\cdot}$

Sum for the i^{th} trt: $y_{i\cdot} = \sum_{j=1}^{n_i} y_{ij}$

Mean for the i^{th} trt: $\bar{y}_{i\cdot} = y_{i\cdot}/n_i$

Grand sum: $y_{\cdot\cdot} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} y_{ij} = \sum_{i=1}^{k} y_{i\cdot}$

Grand mean: $\bar{y}_{\cdot\cdot} = y_{\cdot\cdot}/N$ where the total # of obs is:

$$N = \sum_{i=1}^{k} n_i = n_1 + n_2 + \cdots + n_k.$$

Partitionning the variability (Sums of Squares)

SS Total = SS Trt + SS Error
df Total = df Trt + df Error

$$\text{SS Total} = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{..})^2 = \sum_{i=1}^{k}\sum_{j=1}^{n_i} y_{ij}^2 - \frac{y_{..}^2}{N}$$

$$\text{on df Total} = N - 1$$

$$\text{SS Trt} = \sum_{i=1}^{k} n_i(\bar{y}_{i.} - \bar{y}_{..})^2 = \sum_{i=1}^{k} \frac{y_{i.}^2}{n_i} - \frac{y_{..}^2}{N} \text{ on df Trt} = k - 1$$

$$\text{SS Error} = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.})^2, \text{ also}$$

$$= (n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \cdots + (n_k - 1)s_k^2$$

$$\text{on df Error} = N - k = (n_1 - 1) + \cdots + (n_k - 1)$$

Chick mass: SS and ANOVA table

$$
\begin{aligned}
\text{using} \quad & \bar{y}_{1.} = 5.92, \bar{y}_{2.} = 6.35, \bar{y}_{3.} = 6.70, \ \bar{y}_{..} = 6.30 \\
\text{and} \quad & n_1 = 8, n_2 = 13, n_3 = 6 \ : \\
\text{SSTrt} &= = 2.1563 \\
\text{using} \quad & s_1 = 0.5656, s_2 = 0.5201, s_3 = 0.5477 \ : \\
\text{SSErr} &= = 6.9956 \\
\text{SSTotal} &= = 9.1519
\end{aligned}
$$

Source	df	SS	MS
Trt: Parenting		2.15	1.08
Error		7.00	0.29
Total		9.15	

Chick mass: the F test

H_0: "all population means are equal" vs.
H_A: "not all population means are equal".

Observed test statistic:

$$f = \frac{\text{MSTrt}}{\text{MSErr}} = \qquad = 3.70$$

Compare this with $F_{2,24}$ from Table D: at 5% $f_{2,24} = 3.40$, and at 1% $f_{2,24} = 5.61$, so

$$< \text{p-value} <$$

Reject H_0 at level $\alpha = 0.05$. Moderate evidence that there is a treatment effect on chick mass, i.e. that $\mu_{\text{gr-pr}}$, μ_{control} and $\mu_{\text{pr-grp}}$ are not all equal.

Remark on the F-distribution

The F distribution with degrees of freedom d_1 and d_2 is the distribution of

$$F = \frac{X^2_{d_1}/d_1}{X^2_{d_2}/d_2}$$

when $X^2_{d_1}$ and $X^2_{d_2}$ are independent and

$X^2_{d_1} \sim \chi^2$ distribution with df $= d_1$ and
$X^2_{d_2} \sim \chi^2$ distribution with df $= d_2$.

The pooled standard deviation

Source	df	SS	MS
Trt: Parenting	2	2.15	1.08
Error	24	7.00	0.29
Total	26	9.15	

Pooled standard deviation

MS Error = pooled estimate of variance s_p^2, so

$$s_p = \sqrt{\text{MSerror}}$$

For the 3 samples we had

$s_{\text{grp-pr}} = .566$ g on df$= 8 - 1 = 7$,

$s_{\text{control}} = .520$ g on df$= 13 - 1 = 12$,

$s_{\text{pr-grp}} = .548$ g on df$= 6 - 1 = 5$.

Here we get $s_p = \sqrt{0.292} = 0.540$ g: some kind of average.

$s_p =$ better estimate of the common standard variation σ within each group, based on higher df$= 24$.

Confidence interval based on pooled SD

s_p is useful to get confidence intervals for each treatment mean!

CI for treatment means based on pooled SD

For the population mean μ_1 in treatment 1, a 95% CI is

$$\bar{y}_1 \pm t_{.025,\text{dfErr}} * \text{SE}_{\bar{y}_1} \quad \text{where } \text{SE}_{\bar{y}_1} = \frac{s_p}{\sqrt{n_1}} = \frac{\sqrt{\text{MSerr}}}{\sqrt{n_1}}$$

Mean chick mass in the control group:

sample mean $\bar{y}_{2.} = 6.35$ g, from $n_2 = 13$ obs,
SE$= s_p/\sqrt{13} = 0.150$ g
dfErr$= 24$, multiplier $t = 2.06$ for 95% confidence, so interval:

$$6.35 \pm 2.06 * 0.15 = (6.04, 6.67) \text{ g}.$$

Assumptions

The F-test:

$H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$ versus

H_A: Not all μ_i's are equal.

Under H_0, $F = \dfrac{\text{MSTrt}}{\text{MSError}}$ has an F distribution: $\mathcal{F}_{\text{dfTrt, dfError}}$

This is assuming:

1. Independence of observations between and within samples. Complete Randomized Design (no blocks!)
2. In each treatment, observations come from a normal distribution **or** the sample size is **large**.
3. Equal variances. The population standard deviations of the observations are equal among all treatments:
   $\sigma_1 = \sigma_2 = \cdots = \sigma_k$.

Assumptions

How do we determine if those conditions hold in practice?

1. Assess **independendence** and randomness from the design
2. Look at the normal quantile plots, watch for outliers, look at the normal quantile plot of *residuals*.
3. Compare the sample **standard deviations**.
   $\leq$ 3-fold difference between the smallest and largest: okay.
   $\leq$ 2-fold difference is better.

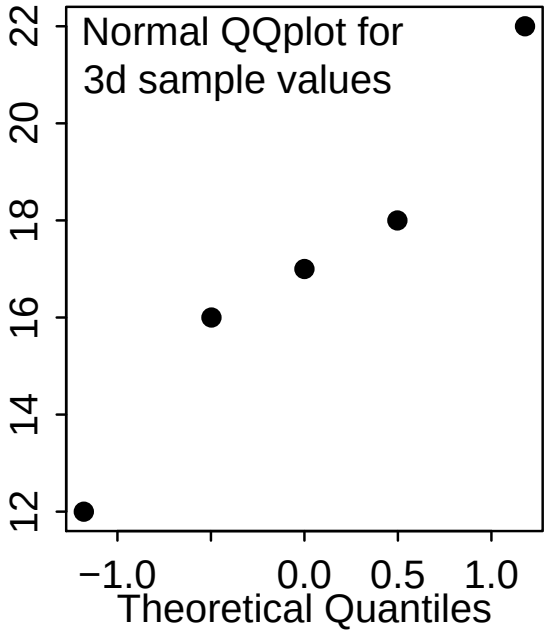
Detecting non-normality: plot of residuals

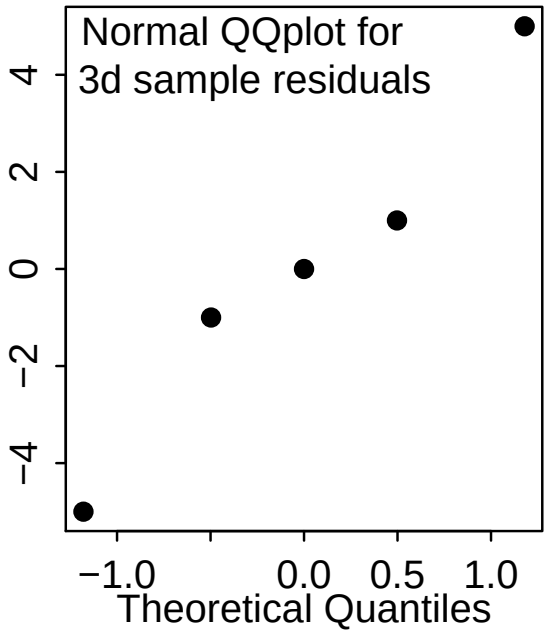
Residuals

deviations from sample means

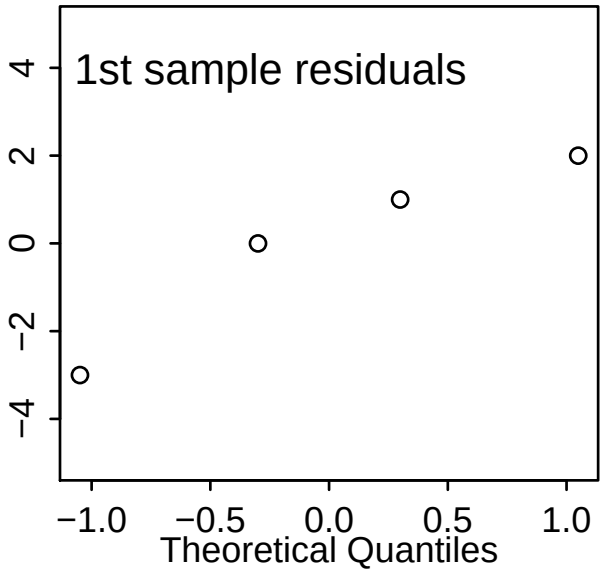
$r_{ij} = y_{ij} - \bar{y}_{i\cdot}$

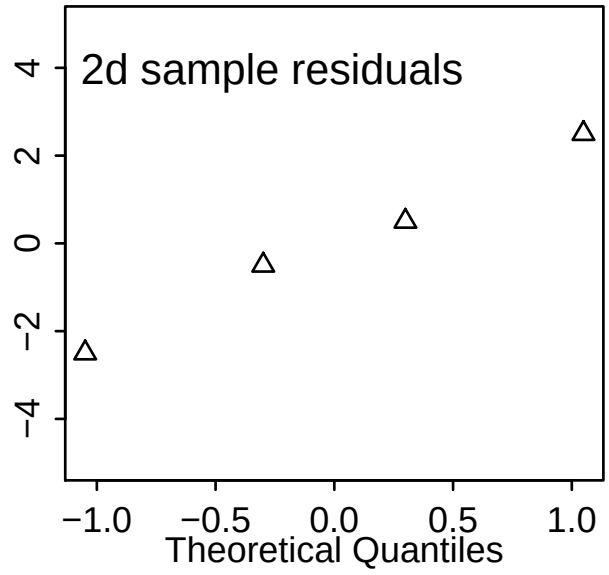
Y_1	res.	Y_2	res.	Y_3	res.
3	-3	10	-2.5	12	-5
6	0	12	-0.5	16	-1
7	1	13	0.5	17	0
8	2	15	2.5	18	1
				22	5
$\bar{y}_{1\cdot} = 6$		12.5		17	

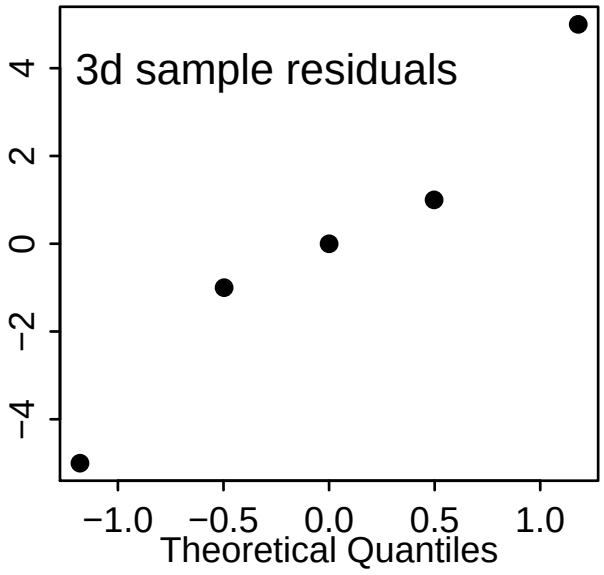


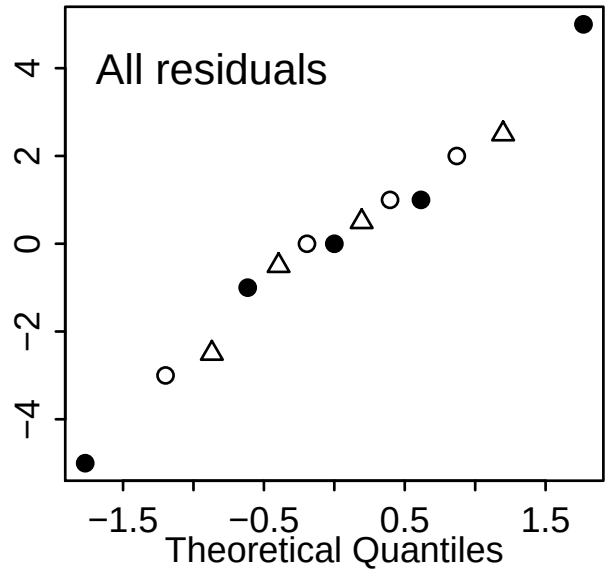


Normal quantile plot of residuals









Normal quantile plot of residuals

Take home message: To address the normality assumption,

1. Do all the calculations, get the residuals.
2. Combine all residuals and do one normal quantile plot.
3. Check its linearity.

R commands

First: have all the data in one file, with one **column** to indicate the **treatment**, and one **column** for the numerical **outcome**

```
> chickmass = read.table("chickmass.dat", header=T)

> chickmass
   mass parents
1  6.24  grp-pr
2  4.94  grp-pr
...
8  6.64  grp-pr
9  5.53 control
10 6.74 control
11 6.61 control
...
26 7.10  pr-grp
27 7.46  pr-grp
```

R commands: aov and anova

First let R do all the calculations with `aov( )`, save them:

```
> fit = aov(mass~parents, data=chickmass)
```

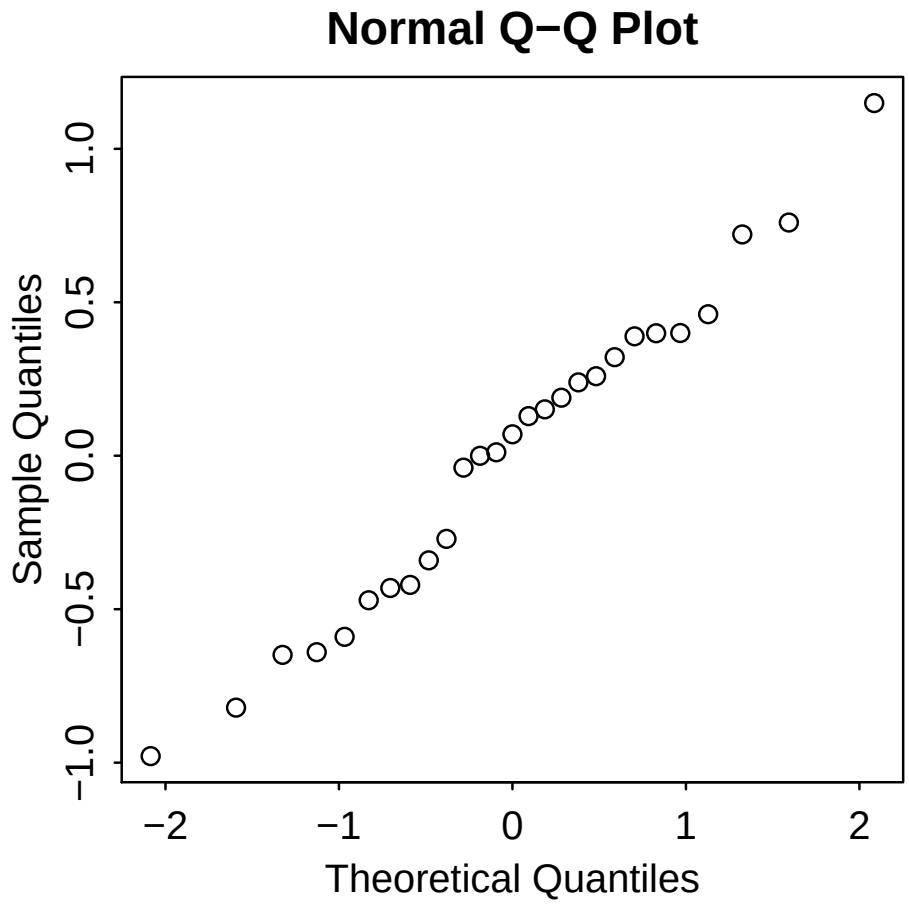
Then ask for the ANOVA table, or residuals, or residual plot:

```
> anova(fit)
Analysis of Variance Table
Response: mass
          Df Sum Sq Mean Sq F value  Pr(>F)
parents    2 2.1563  1.0782  3.6989 0.03979 *
Residuals 24 6.9956  0.2915

> residuals(fit)
    1     2     3     4     5     6     7     8     9    10    11    1
 0.32 -0.98 -0.65  0.01 -0.04  0.46  0.15  0.72 -0.82  0.39  0.26  1.1
   14    15    16    17    18    19    20    21    22    23    24    2
-0.34  0.19 -0.43 -0.42  0.13 -0.47  0.40 -0.27  0.07 -0.59 -0.64  0.0
   27
 0.76
> qqnorm(residuals(fit))
> plot(fit)
```

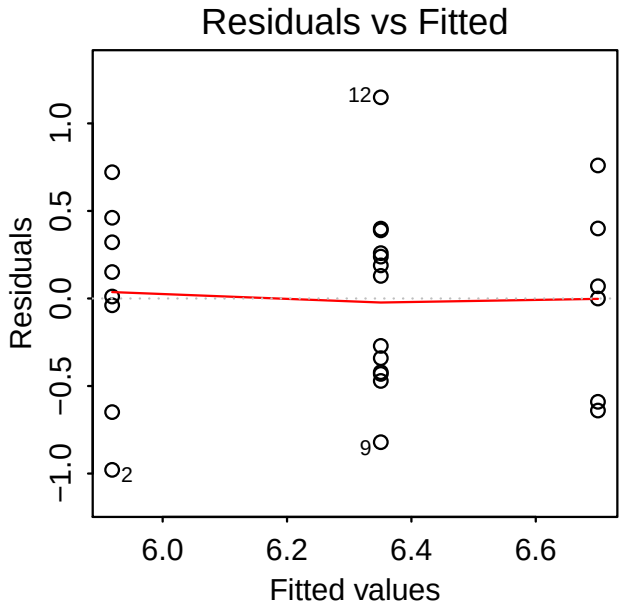
Normal quantile plot of residuals

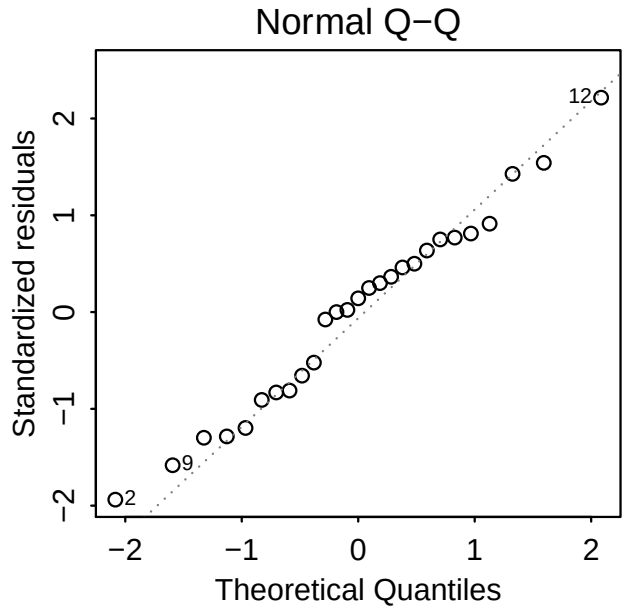
```
qqnorm(residuals(fit))
```

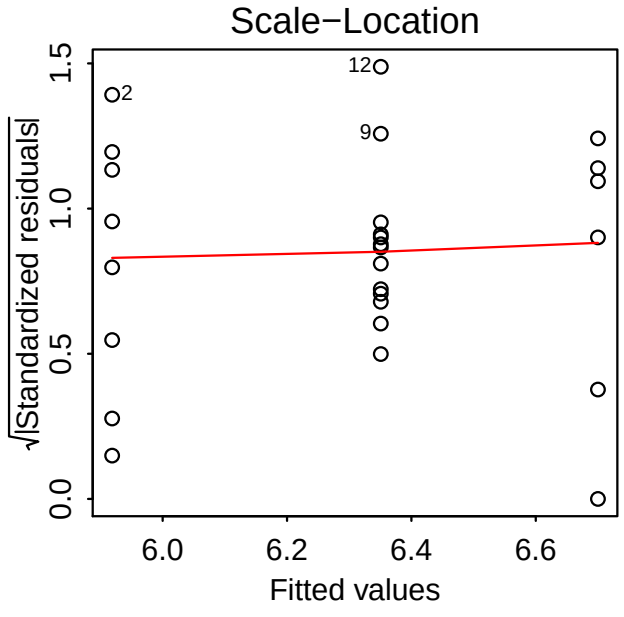


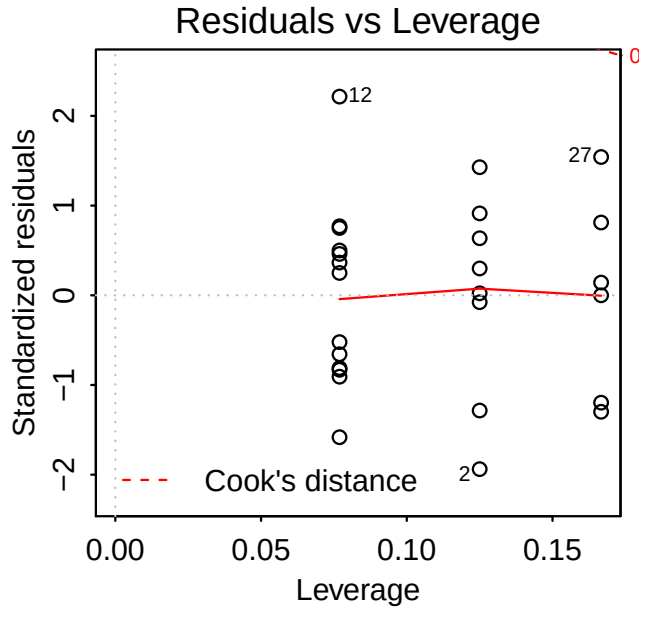
Looks very nice: normality assumption is met.

Residual plots: `plot(fit)`









Corrective actions

If data not normally distributed and/or variances too different:

1. try transforming the data. A log (or square-root) transformation *might* fix both issues.
2. if not, non-parametric alternative (Kruskal-Wallis test), but not covered.

Outline

ANOVA: if we reject H_0, we know that not all treatment means are the same. Then what?

This may not be informative enough. Now consider particular comparisons of treatment means. Which pairs of treatments have significantly different means?

Pairwise comparisons among means

Problem if we went with many standard 2-sample t-tests.

6 t-tests with 4 groups:

- group 1 vs. group 2 (5% chance of type I error)
- group 1 vs. group 3 (5% chance of type I error)
- group 1 vs. group 4 (5% chance of type I error)
- group 2 vs. group 3 (5% chance of type I error)
- group 2 vs. group 4 (5% chance of type I error)
- group 3 vs. group 4 (5% chance of type I error)

If the truth is $\mu_1 = \mu_2 = \mu_3 = \mu_4$, these errors accumulate: up to **20% chance** that **at least 1** type I error is made out of these 6 tests.

That's why the F-test in ANOVA is so useful! Ensures a 5% type I error rate overall. New tool needed for pairwise comparisons.

Concerns with Multiple Comparisons: HIV vaccine trial

Science 318:1048 (13 November 2007)

AIDS RESEARCH

Did Merck's Failed HIV Vaccine Cause Harm?

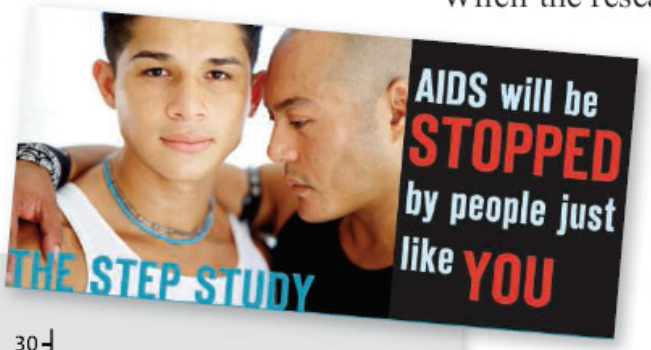
3,000 subjects enrolled in 2004/2005, "at high-risk of becoming infected with HIV". 62% men, 38% women. Control group had placebo: saltwater injection.

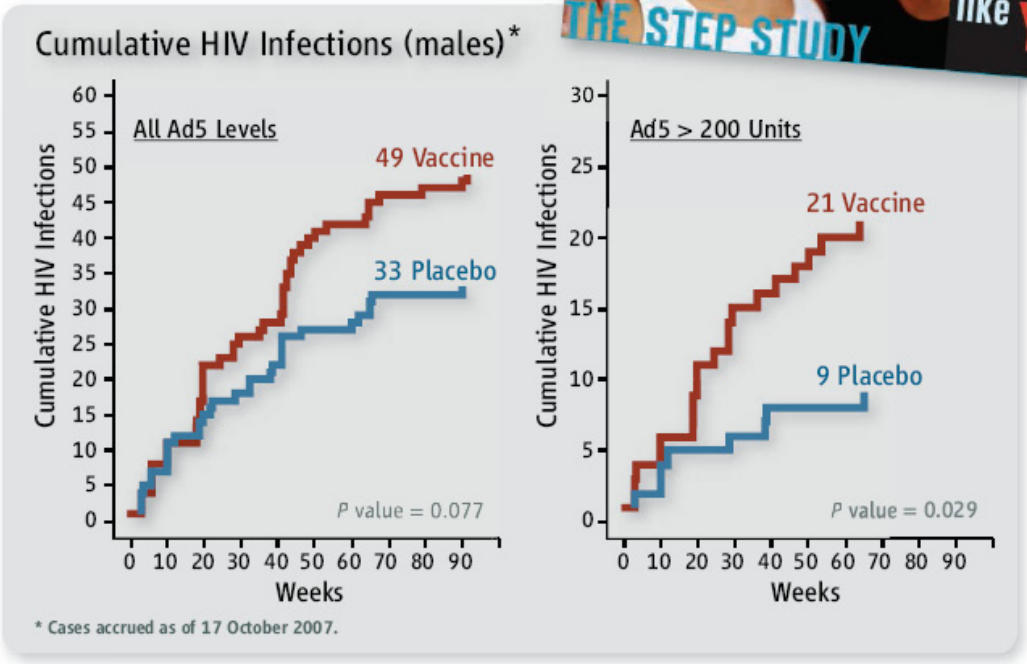
The vaccine had worked on monkeys.

ing from the unexpected failure in September of the most promising vaccine candidate in clinical trials, met here last week to explore an even more alarming finding: The vaccine, made by Merck and Co., may actually have increased the risk of HIV infection in some study participants.

Working with the academic-based HIV Vaccine Trials Network (HVTN) and the U.S. National Institutes of Health (NIH) in Bethesda, Maryland, Merck researchers stopped the multicountry study after an interim analysis revealed that the vaccine did not work (*Science*, 5 October, p. 28). Now further analysis suggests that the vaccine may have helped HIV infect a subset of participants who at the trial's start had high levels of antibody to adenovirus 5 (Ad5), which causes the common cold and is also a component of the vaccine. "This is the worst possible outcome in a vaccine trial," said AIDS researcher Eric Hunter of Emory University in Atlanta, Georgia, one of

trial results, Merck researchers and their partners reported that, as of 17 October, HIV had infected 83 people in the placebo-controlled trial. Of these, 49 were vaccinated and 34 received saltwater injections. This differ-





Double trouble. The vaccine clearly failed (*left*), but in men with high Ad5 antibodies (*right*), it may have increased their risk of infection. (Women were excluded from this analysis because only one became infected during the study.)

When the researchers subsequently examined the high–Ad5-antibody group, they were startled to find 21 infections in vaccinees versus nine in the placebo group.

The statistical analysis is ambiguous. Typically, researchers deem a difference as significant if it has a 95% probability of not being due to chance—a *P* value of less than 0.05. By these standards, the finding, with a *P* value of 0.029, was significant. But Steven Self, HVTN's head statistician at the University of Washington (UW), Seattle, cautioned that this comparison merits a more stringent cutoff for significance, between 0.025 and 0.0025, because the study was not designed to assess potential harm, nor did investigators plan to evaluate a subset of the study population. Still, Self said this "trend" deserves close examination.

(by the way, notice the wrong interpretation of the p-value...)

Making all pairwise comparisons

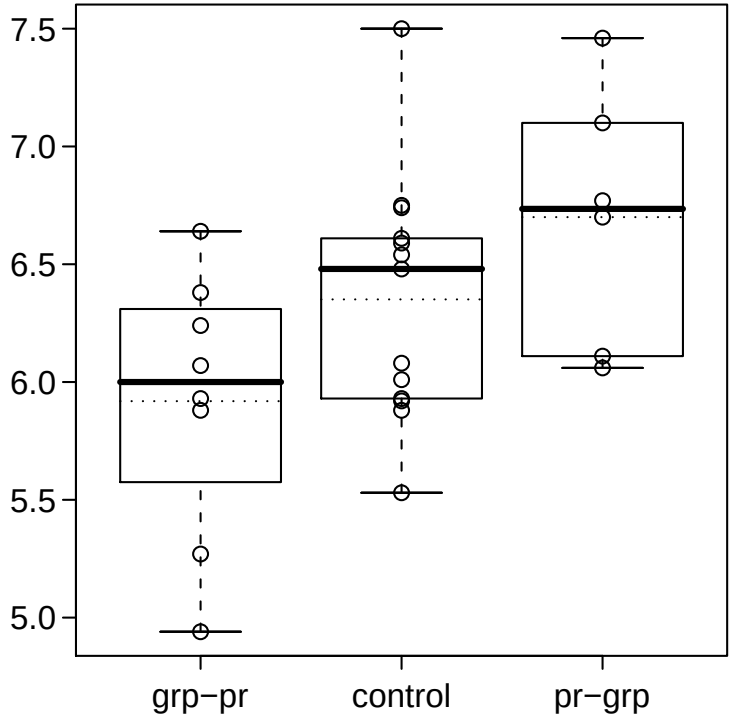
Mass of chicks reared by parents with/without helpers: F-test from ANOVA gave $.01 < \text{p-value} < .05$. Now: compare all pairs of group means.

Group	grp-pr	control	pr-grp
Mean (g)	5.92	6.35	6.70

grp-pr: laid in groups (with helpers), reared in pair (without helpers)

control: laid & reared by own parents (with or without helpers)

pr-grp: laid by pair (without helpers), reared in group (with helpers)



The Tukey-Kramer method

Among many methods, we consider Tukey-Kramer:

- used most widely,
- best for balanced data ($n_1 = n_2 = \cdots = n_k$) but can still be applied to non-balanced studies,
- exact same assumptions as ANOVA.
- also known as **studentized range**, or Q-method, or HSD for honestly significant difference.

The Tukey-Kramer method

Do all the pairwise tests but compare to the Q-distribution

1. Do ANOVA first.

   Stop if p-value > 0.05: none of the pairs are significantly different.
   Otherwise, keep $s_p = \sqrt{\text{MSerror}}$ and its df=dfError.

2. To compare treatment i and j, calculate $\bar{y}_i - \bar{y}_j$ and its standard error:

SE for the difference between 2 means

$$\text{SE}_{\bar{y}_i - \bar{y}_j} = s_p \sqrt{\frac{1}{n_i} + \frac{1}{n_j}}$$

3. compare the t-value $T_{ij} = \dfrac{\bar{y}_i - \bar{y}_j}{\text{SE}_{\bar{y}_i - \bar{y}_j}}$ to the Q-distribution in Table F for df=dfError and k groups,
   to see if the p-value is < 0.05 or not.

The Tukey-Kramer method

From ANOVA we had .01 $<$ p-value $<$ .05 and MSerror$=$ 0.29, i.e. $s_p = 0.54$ g on dfErr$=$ 24.

Group	grp-pr	control	pr-grp
Mean (g)	5.92	6.35	6.70
n	8	13	6

Table F: critical value: 2.50 for 3 groups, df=24.

To compare grp-pr and control: SE$= 0.54 * \sqrt{\frac{1}{8} + \frac{1}{13}} = 0.243$.

	$\bar{y}_i - \bar{y}_j$	SE	T value	critical value	significantly different?
grp-pr vs. control	0.43	0.243	1.77	2.50	No
grp-pr vs. pr-grp	0.78	0.292	2.67	2.50	Yes
control vs. pr-grp	0.35	0.266	1.32	2.50	No

Conclusion

There is evidence that the average chick mass is higher when laid in pairs and reared with helpers (pr-grp) than when laid in groups (parents + helpers) and reared by parents only (grp-pr).

There is *not enough evidence* of differences in average chick mass between the control group and the other 2 groups.

Warning: failing to reject H_0 is not accepting H_0! We do not conclude that the control group has the same average chick mass as both other groups. Contradiction otherwise...

Graphical representations: underlying bars

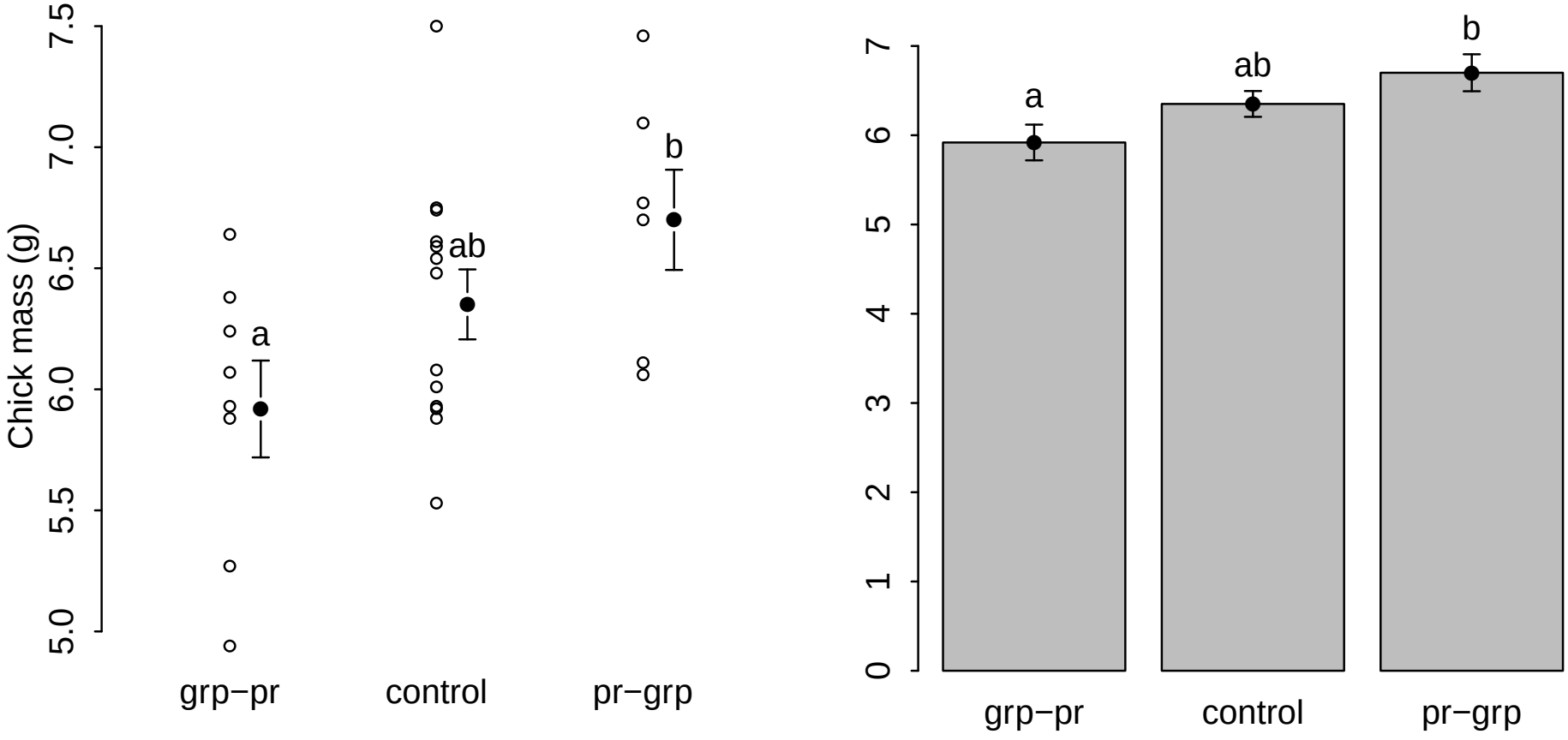
1. order groups by increasing means,
2. underlie groups whose means are not significantly different:

```
Group:  grp-pr control pr-grp
Mean:    5.92   6.35    6.70
         -------------
               --------------
```

Best if sample sizes are close to equal.

Graphical representations: with letters

most widely used, but not as visually convincing.



Bars showing standard errors. Prefer graph on the left: shows both the full data (including sample size) and the conclusion.

Full analysis with R

```
> chickmass = read.table("../data/chickmass.dat", header=T)
> chickmass
   mass parents
1  6.24  grp-pr
2  4.94  grp-pr
3  5.27  grp-pr
4  5.93  grp-pr
5  5.88  grp-pr
6  6.38  grp-pr
...          ...
20 6.75 control
21 6.08 control
22 6.77  pr-grp
23 6.11  pr-grp
24 6.06  pr-grp
25 6.70  pr-grp
26 7.10  pr-grp
27 7.46  pr-grp
```

aov and model.tables

```
> fit = aov(mass ~ parents, data=chickmass)
> anova(fit)
          Df Sum Sq Mean Sq F value  Pr(>F)
parents    2 2.1563 1.07816  3.6989 0.03979 *
Residuals 24 6.9956 0.29148

> model.tables(fit, type="means")
Tables of means

Grand mean
6.30037

 parents
    control grp-pr pr-grp
      6.351  5.919    6.7
rep  13.000  8.000    6.0
```

TukeyHSD on aov object

```
> TukeyHSD(fit, ordered=T)

  Tukey multiple comparisons of means
    95% family-wise confidence level
    factor levels have been ordered

                     diff         lwr      upr     p adj
control-grp-pr 0.4320192 -0.17383437 1.037873 0.1973407
pr-grp-grp-pr  0.7812500  0.05310459 1.509395 0.0338181
pr-grp-control 0.3492308 -0.31620204 1.014664 0.4031720
```

outputs:

- observed differences $\bar{y}_i - \bar{y}_j$
- Tukey-Kramer confidence interval for each $\mu_i - \mu_j$
- p-value for testing H_0: $\mu_i = \mu_j$. Difference declared significant if p-value $< .05$.

Barley root example: step 1 = ANOVA

5 varieties of barley. Weight of roots recorded for $n = 7$ plants per variety. Observed group means:

$\bar{y}_{1\cdot}$	$\bar{y}_{2\cdot}$	$\bar{y}_{3\cdot}$	$\bar{y}_{4\cdot}$	$\bar{y}_{5\cdot}$
16.3	19.3	14.7	20.3	18.5

ANOVA table:

Source	df	SS	MS	F	p-value
Trt		145.94	36.48	5.09	< 0.01
Error		214.74	7.16	–	
Total		360.68	–	–	

Ingredients needed for next step, Tukey-Kramer: $k = 5$ groups (varieties), $n = 7$ in each group, $s_p =$ $= 2.68$, dfErr $=$

Barley root example, step 2 = Tukey-Kramer

Now compare pairs of varieties: $k = 5$ groups (varieties), $n = 7$, $s_p = 2.68$, dfErr $= 30$, and $Q_{5,30,0.05} = 2.90$ at $\alpha = 0.05$.

For each comparison $\bar{y}_i - \bar{y}_j$, SE$= s_p \sqrt{\frac{1}{7} + \frac{1}{7}} = 1.43$, so compare $t = (\bar{y}_i - \bar{y}_j)/1.43$ to the critical value 2.90.

Examples:

	$\bar{y}_i - \bar{y}_j$	SE	T value	critical value	significantly different?
1 vs. 3	1.6	1.43	1.12	2.90	No
5 vs. 3	3.8	1.43	2.66	2.90	No
5 vs. 1	2.2	1.43	1.54	2.90	No
2 vs. 3	4.6	1.43	3.22	2.90	Yes

```
Variety:   3    1    5    2    4
Mean:    14.7 16.3 18.5 19.3 20.3
          ------------
                ---------------------
```