

Lecture 12 – ANOVA 方差分析

- Outline for today
 - Recall Experimental Design 回顾试验设计
 - Sample size 样本大小
 - Comparing more than two means
 - ANOVA 方差分析
 - Planned & unplanned comparisons 多重比较
 - Fixed & random effects 固定& 随机效应
 - Summary
 - R Lab & Discussion

生物统计学

李 勤

生态与环境科学学院



1. 回顾——Correction

- Poisson distribution formula 泊松分布

- 在任何给定的时间或空间块中，成功事件发生X次的概率为：

$$\Pr[X \text{ success}] = \frac{e^{-\mu} \mu^{-X}}{X!} \longrightarrow \Pr[X \text{ success}] = \frac{e^{-\mu} \mu^X}{X!}$$

- μ : 时间或空间中独立成功的平均数量
 - rate: 表示每单位时间或每单位空间的计数
- e: 自然对数的底
 - 一个约等于2.718的常数
- $X!$: X的阶乘



1. 回顾——试验设计

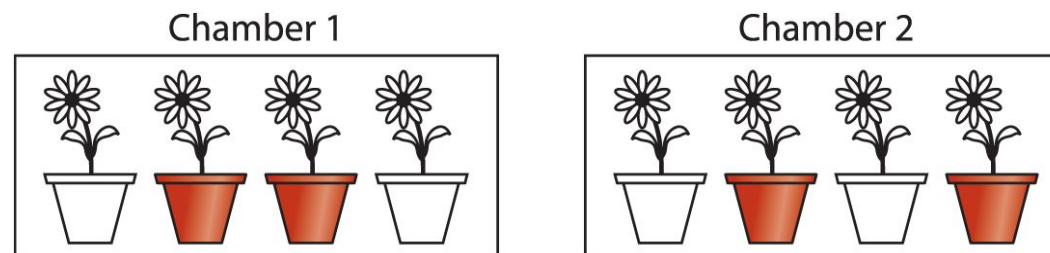
- 原则/目标

- 减少偏差

- 同时设计控制组;
 - 随机设计;
 - 单盲/双盲;

- 减少抽样误差

- 重复的试验单元;
 - 权衡样本大小;
 - 设置组别 (block);
 - 其它变异条件;



Whitlock & Schluter, *The Analysis of Biological Data*, 3e © 2020 W. H. Freeman and Company

1.1 试验设计——课堂练习

6. Identify the consequences (i.e., increase, decrease, or none) that the following procedures are likely to have on both bias and sampling error in an observational study.

- Matching sampling units between treatment and control
- Increasing sample size
- Ensuring that the frequency distribution of subject ages is the same in the two treatments
- Using a balanced design

2. 试验设计之一——样本大小

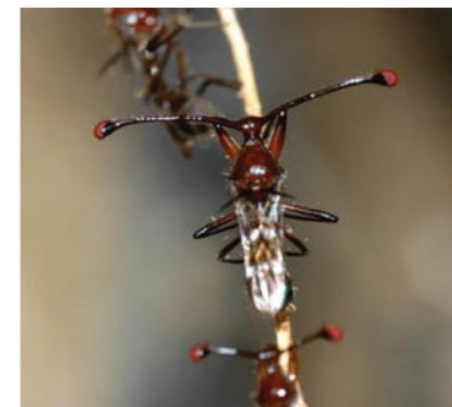
- 试验单元的多少？
 - 影响因素：
 - 成本
 - 是否能检测出预期效果
 - 置信区间的估计
- 探讨样本大小的目标
 - 达到预定的精确度；
 - 达到预定的功效；

2.1 达到预期的精确度 plan for precision

- 设定预期的精确度
 - 即设定可以接受的不确定性（uncertainty to tolerate）
- 对两个处理组之间差异的估计
 - 总体的差异： $\mu_1 - \mu_2$
 - 样本的估计（95%的置信区间）： $(Y_1 - Y_2) \pm \text{margin error}$
- 样本量的估计
 - $n = 8\left(\frac{\sigma}{\text{margin error}}\right)^2$
 - σ : 两个总体的标准差（方差和标准差均相等）；
 - n : 两个样本大小相等；
 - margin error: 置信区间的一半（2SE法则）；

2.1 达到预期的精确度 plan for precision

- 样本量的估计
 - σ 通常未知，需要通过前期经验/研究或预试验来设定；
- 例子： diet effects on stalk-eyed flies 饮食对杆状眼蝇的影响
 - 假定我们期望获得的误差边际margin error为0.1mm；
 - 依据EXAMPLE 11.2: Eye to eye 的估计： $s = 0.398$ ；
 - $$n = 8\left(\frac{\sigma}{\text{margin error}}\right)^2 = 8 \times \left(\frac{0.4}{0.1}\right)^2 = 128$$
 - 如若两个处理组，总样本为 $2n = 256$

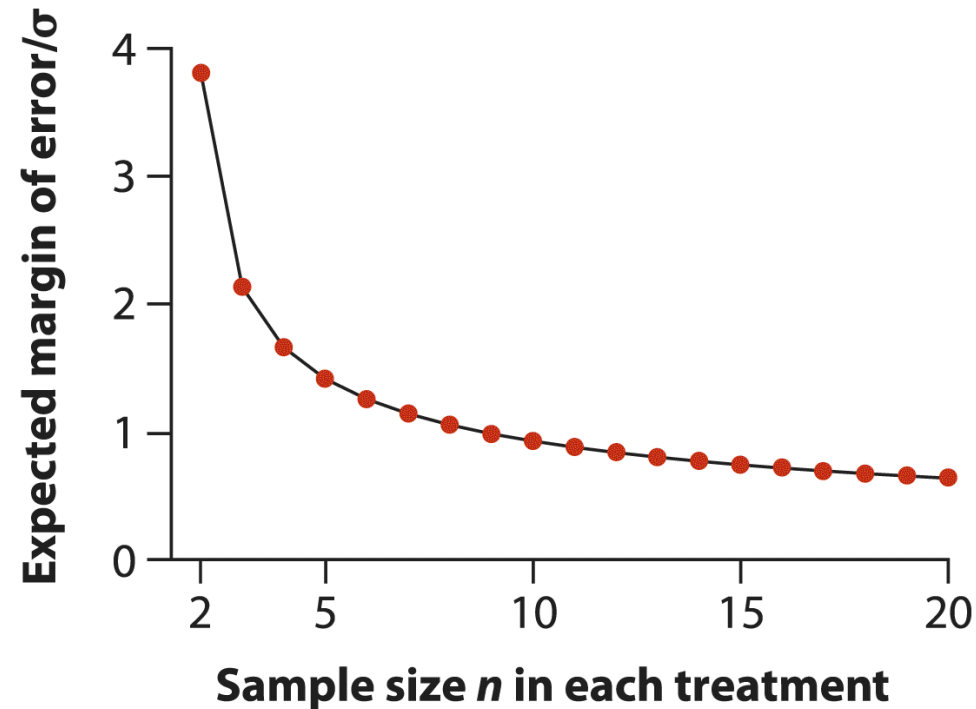


Zimmer, C. *Evo Edu Outreach* (2008) 1: 487. <https://doi.org/10.1007/s12052-008-0089-9>.
Photo courtesy of Gerald Wilkinson, University of Maryland

2.1 达到预期的精确度 plan for precision

- 样本量的估计

- σ 通常未知，需要通过前期经验/研究或预试验来设定；
- 如果总体的 σ 大，则需要大样本来获取高精确度；



Whitlock & Schluter, *The Analysis of Biological Data*, 3e © 2020 W. H. Freeman and Company

2.2 达到预期的功效 plan for power

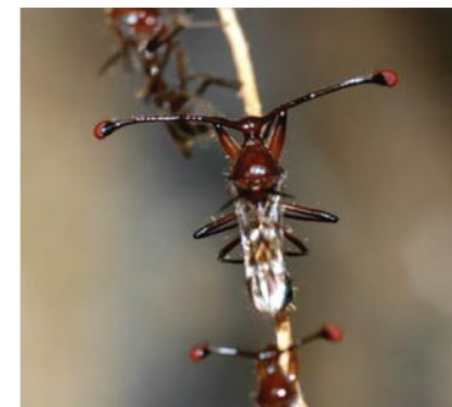
- 功效和样本量

- $H_0: \mu_1 = \mu_2; \quad H_A: \mu_1 \neq \mu_2;$
- 当 $|\mu_1 - \mu_2| \geq D$ 时, 我们认为达到了足够的功效;
 - D 不是真实的组间差异, 而是我们关心的某个临界值;
- 统计检验中, 通常需要功效达到0.8以上;

- 样本量的估计

- $n \approx 16\left(\frac{\sigma}{D}\right)^2$
 - (更多请查看 [Section 14.8 Quick Formula Summary](#))
- 若设定 $D = 0.2 \text{ mm}$, 则

- $n \approx 16\left(\frac{0.4}{0.2}\right)^2 = 64$



Zimmer, C. *Evo Edu Outreach* (2008) 1: 487. <https://doi.org/10.1007/s12052-008-0089-9>.
Photo courtesy of Gerald Wilkinson, University of Maryland

2.3 其它情景

- A great deal of caution
 - We assumed we know σ
- Sample sizes for desired precision and power are available for one- and two-sample means, proportions, and odds ratios in the Quick Formula Summary ([Section 14.8](#)).
- A variety of computer programs are available to calculate sample sizes when planning for power and precision.

3. 方差分析 ANOVA

- ANOVA: the analysis of variance
- 目的: 同时比较2组以上的组间均值;
 - 多次两组间比较 (t -test) 会造成犯I类错误的概率急剧增大;
- One-way ANOVA vs. Two-way ANOVA
 - 一个解释变量 vs. 两个解释变量



3.1 ANOVA方差分析的步骤



Oscar Burriel/Science Source

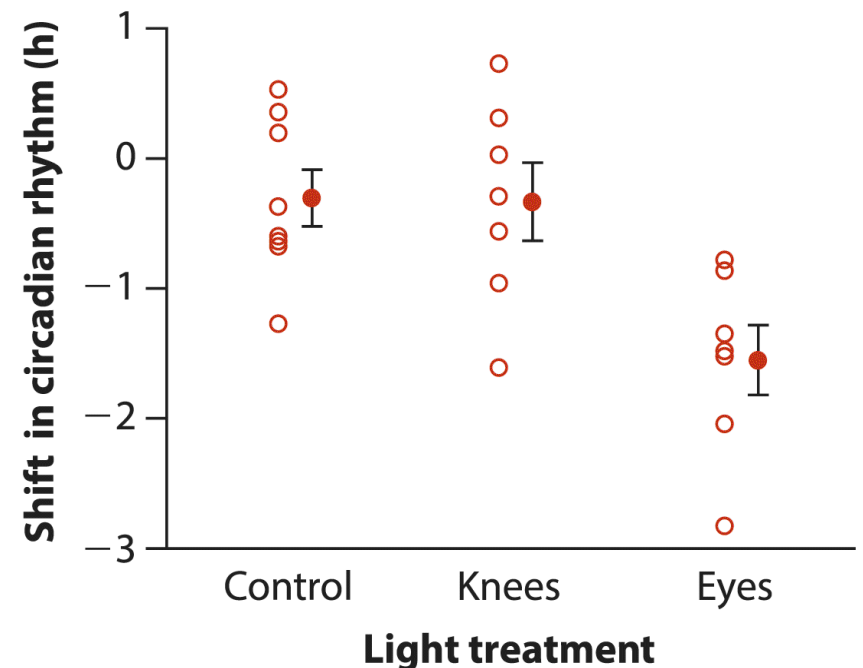
- 例子: [EXAMPLE 15.1: The knees who say night](#)

- 昼夜节律的调整 (circadian clock)

- 有研究提出对膝盖后侧的强光照射有助于调整时差;
- 重新的调查: 三个处理组间是否存在差异?
 - 强光对身体特定部位的照射 (3hrs);
 - 测量褪黑素分泌周期的相位移动;

```
> circadian %>% group_by(treatment) %>%  
+   summarise(mean = mean(shift),  
+             sd = sd(shift),  
+             n = n())  
# A tibble: 3 x 4  
  treatment    mean    sd      n  
  <chr>      <dbl> <dbl> <int>  
1 control   -0.309 0.618     8  
2 eyes     -1.55  0.706     7  
3 knee     -0.336 0.791     7
```

如何进行假设检验?



3.1 ANOVA方差分析的步骤

- 例子: [EXAMPLE 15.1: The knees who say night](#)

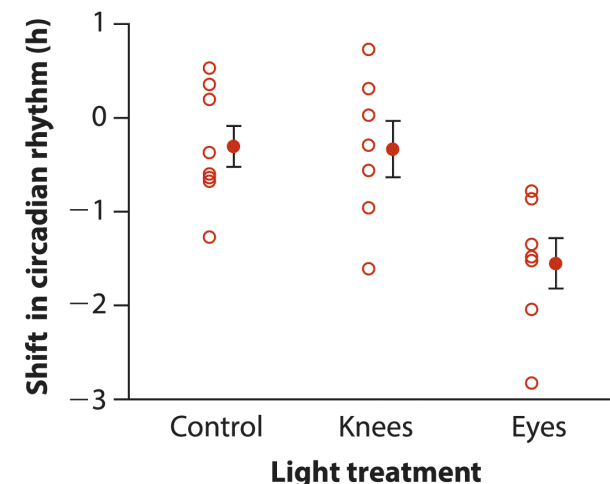
- 昼夜节律的调整 (circadian clock)

- 三个处理组间是否存在差异?

- 假设检验

- 零假设和备择假设

- $H_0: \mu_1 = \mu_2 = \mu_3$;
 - H_A : at least one μ_i is different;
 - 至少有一个均值显著不同, 并非 $\mu_1 \neq \mu_2 \neq \mu_3$;



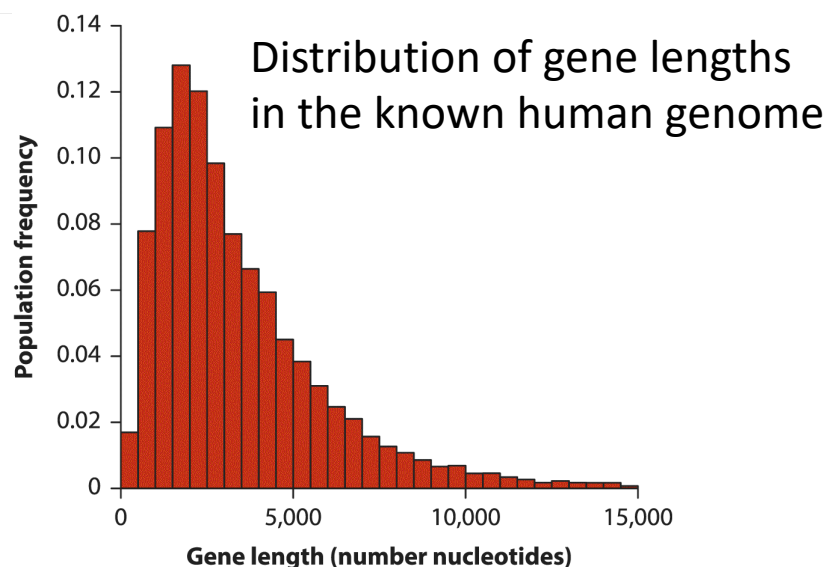
Whitlock & Schluter, *The Analysis of Biological Data*, 3e © 2020 W. H. Freeman and Company

3.1 ANOVA方差分析的步骤

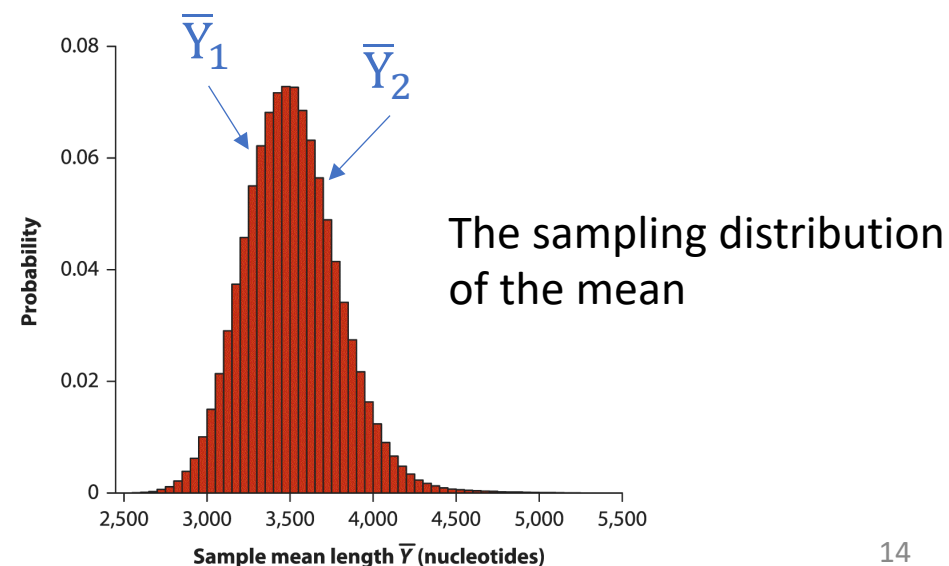
- 数据中的变异 variance

- 总体 vs 样本

- 即使不同总体的均值相等，但由于抽样误差的存在，基于样本计算的均值可能呈现为不同的值；
 - 意味着数值变异的来源之一是抽样误差；



Whitlock & Schluter, *The Analysis of Biological Data*, 3e © 2020 W. H. Freeman and Company



Whitlock & Schluter, *The Analysis of Biological Data*, 3e © 2020 W. H. Freeman and Company

3.1 ANOVA方差分析的步骤

- 数据中的变异

- 组间变异和组内变异：变异的分解

- 如果不同总体（不同处理组）的均值相等，也就是零假设成立时：

- 组间变异和组内变异一样，即组间差异无显著不同；

- within-treatments/groups variance = among-treatments/groups variance

- 如果不同总体（不同处理组）的均值不相等，也就是零假设不成立时：

- 组间变异和组内变异不一样，即组间差异显著不同；

- within-treatments/groups variance \neq among-treatments/groups variance

3.1 ANOVA方差分析的步骤

- 方差分析的基本思想：组间变异和组内变异的比较
 - 不同处理组的样本均值之间的变异是否大于抽样误差导致的变异？
 - 后者即因偶然性所预期的差异；
 - 如果组间变异 > 组内变异，那么可推断不同总体的均值间存在真正的差异；即拒绝零假设；

3.1 ANOVA方差分析的步骤

- 方差分析的基本思想：组间变异和组内变异的比较
 - 不同处理组的样本均值之间的变异是否大于抽样误差导致的变异？
 - 后者即因偶然性所预期的差异；
 - 如果组间变异 > 组内变异，那么可推断不同总体的均值间存在真正的差异；即拒绝零假设；
- 变异的衡量：均方（mean squares）
 - 组间变异——组间均方： MS_{groups} (group mean squares)
 - 衡量了由组间的差异所造成的数据中的变异；
 - the variation among the sampled subjects that belong to different groups;
 - 组内变异——组内均方： MS_{error} (the error mean squares)
 - 衡量了各组内的数据变异；
 - the variance among subjects that belong to the same group;

3.1 ANOVA方差分析的步骤

- 变异分解 (Partitioning the sum of squares)

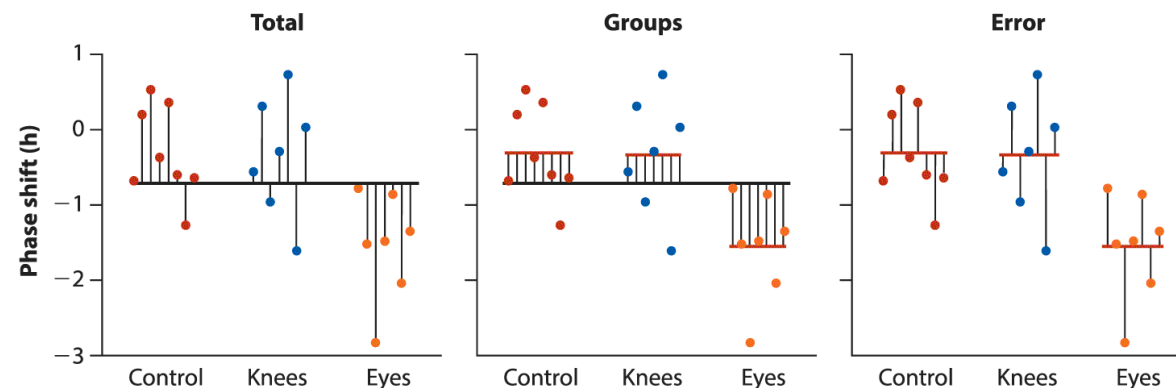
- SS: sum of squares

- $SS_{\text{total}} = SS_{\text{groups}} + SS_{\text{error}}$

- $SS_{\text{total}} = \sum_i \sum_j (Y_{ij} - \bar{Y})^2$

- $SS_{\text{groups}} = \sum_i n_i (\bar{Y}_i - \bar{Y})^2$

- $SS_{\text{error}} = \sum_i \sum_j (Y_{ij} - \bar{Y}_i)^2$



Whitlock & Schluter, *The Analysis of Biological Data*, 3e © 2020 W. H. Freeman and Company

- 其中:

- $\bar{Y} = \sum_i n_i \bar{Y}_i / N$: 总体均值/ grand mean; N 为总样本量;

- Y_{ij} : 各个观察值/测量值; i 为第*i*个处理组, j 为该组内第*j*个个体;

- Y_{ij} : 各个观察值/测量值; i 为第*i*个处理组, j 为该组内第*j*个个体;

- $\bar{Y}_i = \sum_j Y_{ij} / n_i$: 第*i*个处理组的均值; n_i 为第*i*个处理组的样本量;

3.1 ANOVA方差分析的步骤

- 变异分解: 均方 mean squares

- 组间均方 (the group mean squares)

- $MS_{\text{groups}} = \frac{SS_{\text{groups}}}{df_{\text{groups}}}$

- $df_{\text{groups}} = k - 1$, 其中: k 为组数 (number of groups)

- 衡量了由组间的差异所造成的数据中的变异;

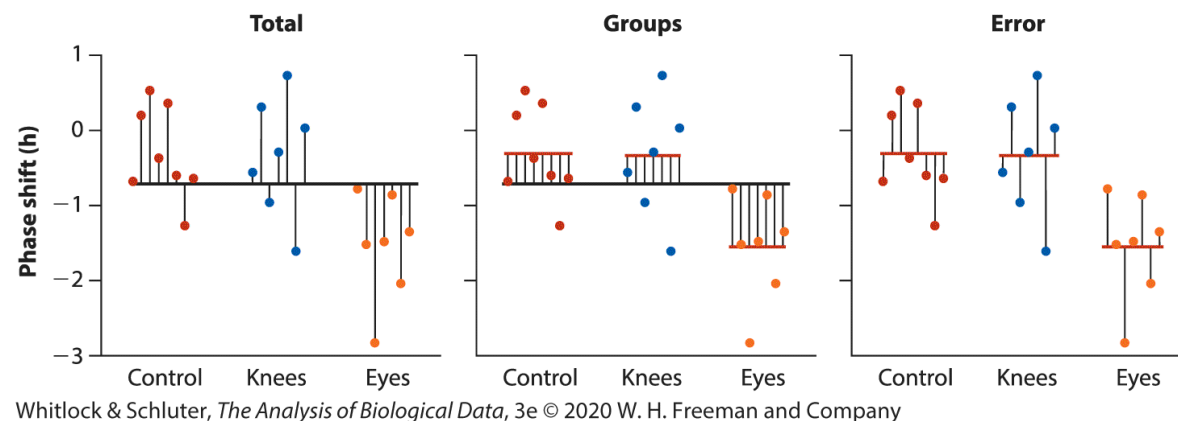
- 组内均方 (the error mean squares)

- $MS_{\text{error}} = \frac{SS_{\text{error}}}{df_{\text{error}}}$

- $df_{\text{error}} = \sum_i (n_i - 1) = N - k$

- 衡量了各组内的数据变异, 类似 pooled sample variance;

- Assumption: $\sigma_1 = \sigma_2 = \sigma_3$



3.1 ANOVA方差分析的步骤

- 例子: [EXAMPLE 15.1: The knees who say night](#)

- 昼夜节律: 三个处理组间是否存在差异?

- 假设检验

- $H_0: \mu_1 = \mu_2 = \mu_3;$

- H_A : at least one μ_i is different; 至少有一个总体的均值与其它组不同;

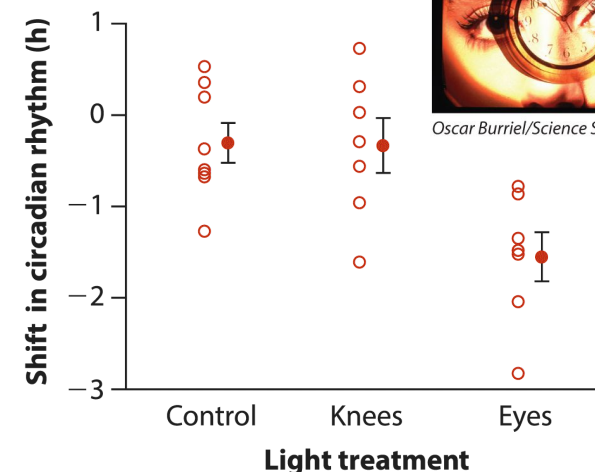
- 变异的衡量

- $MS_{\text{groups}} = \frac{7.224}{3-1} = 3.6122$

- $MS_{\text{error}} = \frac{9.415}{22-3} = 0.4955$

- 检验统计量:

- $F = \frac{MS_{\text{groups}}}{MS_{\text{error}}} = \frac{3.6122}{0.4955} = 7.29$



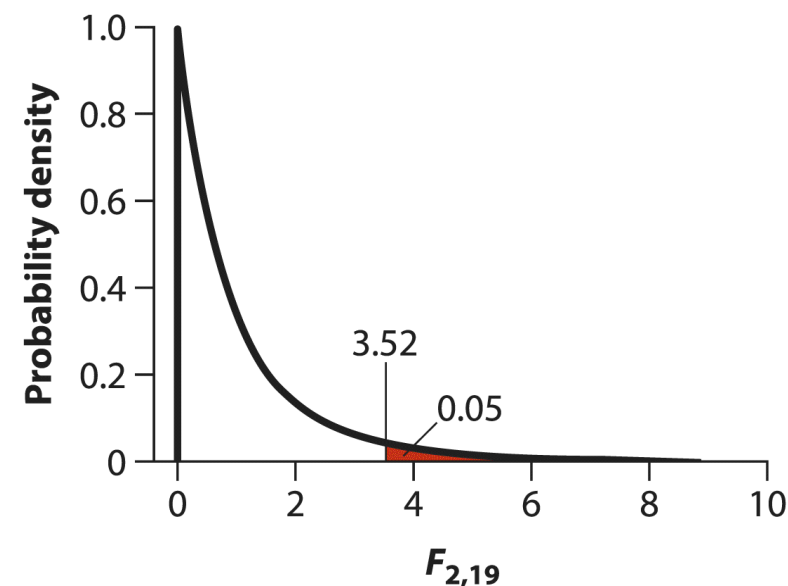
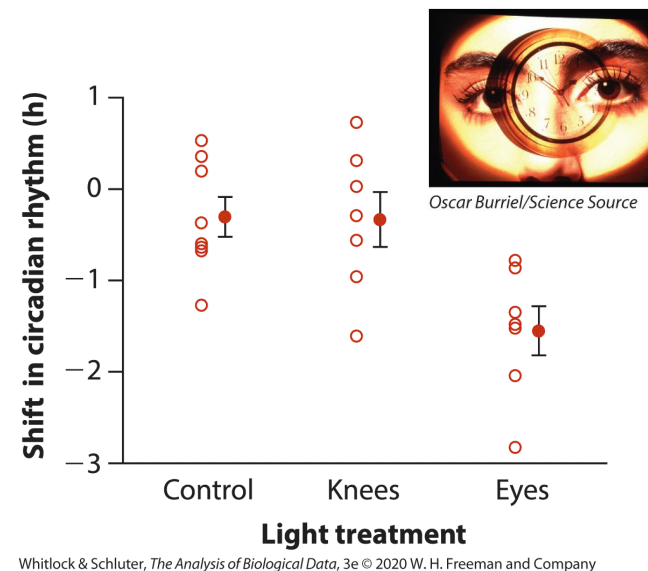
Whitlock & Schluter, *The Analysis of Biological Data*, 3e © 2020 W. H. Freeman and Company

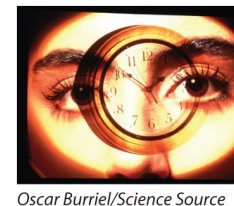
```
circadianAnova <- lm(shift ~ treatment, data = circadian)
anova(circadianAnova)
```

```
## Analysis of Variance Table
##
## Response: shift
##           Df Sum Sq Mean Sq F value Pr(>F)
## treatment  2  7.2245   3.6122  7.2894 0.004472 **
## Residuals 19  9.4153   0.4955
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

3.1 ANOVA方差分析的步骤

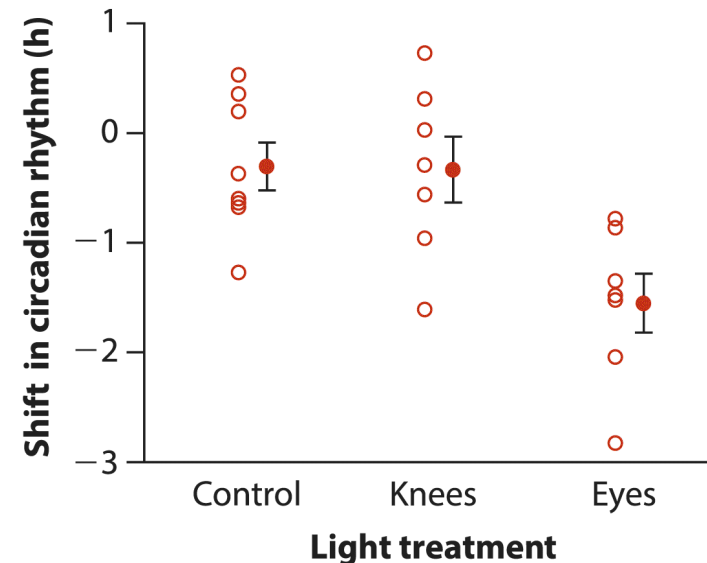
- 例子: [EXAMPLE 15.1: The knees who say night](#)
 - 昼夜节律: 三个处理组间是否存在差异?
 - 假设检验
 - $H_0: \mu_1 = \mu_2 = \mu_3$;
 - H_A : at least one μ_i is different;
 - 检验统计量:
 - $F = \frac{MS_{\text{groups}}}{MS_{\text{error}}} = \frac{3.6122}{0.4955} = 7.29$
 - P值: 与F分布进行比较
 - $F \sim F_{\alpha(1), df_{\text{group}}, df_{\text{error}}} = F_{0.05(1), 2, 19}$
 - F是否显著大于1?
 - “(1)”表示单侧检验 (查看F分布的右侧)





3.1 ANOVA方差分析的步骤

- 例子: [EXAMPLE 15.1: The knees who say night](#)
 - 昼夜节律: 三个处理组间是否存在差异?
 - 假设检验
 - $H_0: \mu_1 = \mu_2 = \mu_3$;
 - H_A : at least one μ_i is different;
 - 检验统计量:
 - $F = \frac{MS_{\text{groups}}}{MS_{\text{error}}} = \frac{3.6122}{0.4955} = 7.29$
 - P值: 与F分布进行比较
 - $P = 0.0045 < 0.05$
 - 结论: 拒绝零假设
 - 至少有一组的总体均值与其它组不同;



Whitlock & Schluter, *The Analysis of Biological Data*, 3e © 2020 W. H. Freeman and Company

3.2 ANOVA方差分析的前提假设

- ANOVA的前提假设 assumptions
 - 随机样本 (random samples)
 - 变量在每组总体中均呈正态分布 (normal distributions)
 - 每组总体的方差相等 (same variance)
- ANOVA的稳健性
 - 当样本量较大时，变量偏离正态分布时ANOVA也可以得出可靠结论；
 - 可进行数据转化来查看是否满足正态分布；
 - 当样本量较大且各组样本接近时，方差差异不超过10倍时ANOVA也可以得出可靠结论；

3.3 ANOVA方差分析相关的其它衡量指标

- R^2 (R-squared)

- $R^2 = \frac{SS_{\text{groups}}}{SS_{\text{total}}}$

- 理解为：变异被“不同组”解释的比例（variation explained）

- the “fraction of the variation in Y that is explained by groups”

- 例子：昼夜节律的三个处理组

- $R^2 = \frac{7.224}{16.639} = 0.43$

- 即43%的变异可以被处理组间的差异解释；

3.4 两组间的ANOVA方差分析

- 当 $k = 2$ 时,
 - ANOVA 和 t -test 会得到相同的结论;
- 不同点
 - t -test 可以被扩展到更泛的检验中, 比如:
 - $\mu_1 = \mu_2$, 或
 - $\mu_1 - \mu_2 = 10$
 - 当两组间的方差极为不一致时可以应用 Welch's t -test ;

3.5 ANOVA的非参方法替代

- Kruskal-Wallis test

- 在有两组以上的情况下，相当于Mann-Whitney U 检验
 - 基于秩的方差分析
- 检验统计量: H
 - 基于每组的秩和 R_i
 - H 的抽样分布近似于自由度为 $k - 1$ 的 χ^2 分布
 - 当 $H > \chi^2_{\alpha, k-1}$ 时，拒绝零假设

```
kruskal.test(shift ~ treatment, data = circadian)
```

```
##  
##  Kruskal-Wallis rank sum test  
##  
## data:  shift by treatment  
## Kruskal-Wallis chi-squared = 9.4231, df = 2, p-value = 0.008991
```

3.5 ANOVA的非参方法替代

- Kruskal-Wallis test

- 在有两组以上的情况下，相当于Mann-Whitney U 检验

- 基于秩的方差分析

- 检验统计量: H

- 基于每组的秩和 R_i

- H 的抽样分布近似于自由度为 $k - 1$ 的 χ^2 分布

- 当 $H > \chi_{\alpha, k-1}^2$ 时，拒绝零假设

- 前提假设

- 随机样本 & 变量的分布在各总体中的形状是一样的;

- 注意事项

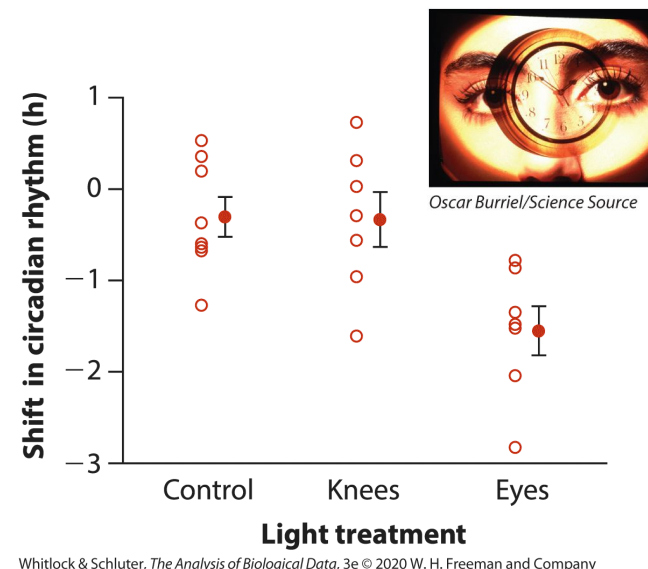
- 样本量很小时，该检验的功效低；样本量大时，该检验与ANOVA功效一致；

4. Planned & Unplanned comparisons

- 配对比较/多重比较
 - Planned comparison（计划比较）
 - 试验设计时明确的感兴趣的组间比较（特定的），需要充足的理由；
 - A comparison between means identified as being of crucial interest during the design of the study, identified prior to obtaining the data.
 - Unplanned comparison（事后多重比较）
 - 多组间的两两配对比较，来找到有差异的处理组（探索性的）；
 - 需要明确对犯I类错误的概率的矫正；

4.1 计划比较Planned comparisons

- 计划比较：对进行比较的处理组有明确的考量；
 - 例子： [EXAMPLE 15.1: The knees who say night](#)
 - 一开始的目标是要看对Knee强光照射是否影响昼夜节律；
 - 一组比较： Control vs. Knee ($\mu_1? = \mu_2$)
 - 基于 t 统计量来进行检验和计算置信区间，其中SE有矫正；
 - 优势：比两样本 t 检验有更高精确度（更窄的置信区间）



$$SE = \sqrt{MS_{\text{error}} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

```
circadianPlanned <- contrast(circadianPairs, method = "pairwise", adjust = "none")
```

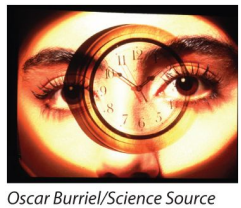
```
circadianPlanned
```

##	contrast	estimate	SE	df	t.ratio	p.value
##	control - knee	0.027	0.364	19	0.074	0.9418
##	control - eyes	1.243	0.364	19	3.411	0.0029
##	knee - eyes	1.216	0.376	19	3.231	0.0044

```
confint(circadianPlanned)
```

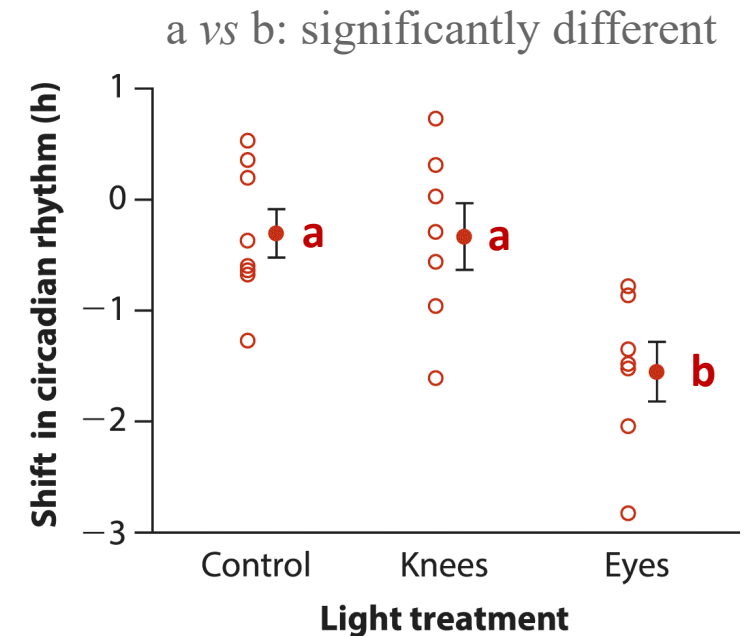
##	contrast	estimate	SE	df	lower.CL	upper.CL
##	control - knee	0.027	0.364	19	-0.736	0.79
##	control - eyes	1.243	0.364	19	0.480	2.01
##	knee - eyes	1.216	0.376	19	0.428	2.00
##						
##	Confidence level used: 0.95					

（其中contrast()命令会给出所有配对比较结果，但我们只关注control-knee）

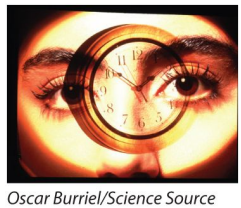


4.2 多重比较 Unplanned comparisons

- 多重配对比较: unplanned/非计划的
 - E.g., 所有组间的两两配对比较;
 - Tukey-Kramer test是最常用的一种方法
 - Tukey's HSD test
 - 检验统计量: q
 - 自由度 $N - k$
 - (具体计算可查看 Section 15.8 Formula Summary)
 - 对I类错误概率的矫正
 - 所有配对比较的检验中, 出现至少一次I类错误的概率不超过显著性水平 α 。
 - 其它方法
 - Bonferroni, Dunnett's Test, etc.



Whitlock & Schluter, *The Analysis of Biological Data*, 3e © 2020 W. H. Freeman and Company



4.2 多重比较 Unplanned comparisons

- 多重配对比较: unplanned/非计划的
 - E.g., 所有组间的两两配对比较;
 - Tukey-Kramer test是最常用的一种方法
 - 检验统计量: q
 - 基于 MS_{error} , 自由度 $N - k$

$$q = \frac{\bar{Y}_i - \bar{Y}_j}{SE}$$

$$SE = \sqrt{MS_{\text{error}} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

```
circadianPairs <- emmeans(circadianAnova, "treatment")
circadianUnplanned <- contrast(circadianPairs, method = "pairwise", adjust = "tukey")
circadianUnplanned
```

				q	
##	contrast	estimate	SE	df	t-ratio p-value
##	control - knee	0.027	0.364	19	0.074 0.9970
##	control - eyes	1.243	0.364	19	3.411 0.0079
##	knee - eyes	1.216	0.376	19	3.231 0.0117
##					
##	P value adjustment: tukey method for comparing a family of 3 estimates				

4.2 多重比较 Unplanned comparisons

- 多重配对比较: unplanned/非计划的
 - E.g., 所有组间的两两配对比较;
 - Tukey-Kramer test是最常用的一种方法
 - 检验统计量: q
 - 基于 MS_{error} , 自由度 $N - k$
 - 前提假设
 - 和ANOVA一致;
 - 相对保守 conservative
 - 在各组样本量不一致时, 犯I类错误的概率比 α 要低;
 - (更难拒绝零假设)

$$q = \frac{\bar{Y}_i - \bar{Y}_j}{SE}$$

$$SE = \sqrt{MS_{\text{error}} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

5. Fixed and Random effects

- One-way ANOVA: 单因子方差分析
 - 单因子: 也是固定效应 = fixed effect
 - 也是我们感兴趣的, 在收集和分析数据前所设计的处理组 (direct interest);
 - 例子: 临床研究中不同的药物、人群中的年龄/性别等等;
 - 一般不会把结果推演到其它 (未设定) 的组中;
- Random-effects ANOVA: 随机效应方差分析
 - 非固定处理, 从总体中随机选择的 (random chosen)
 - 组间差异结果一般可推演到不同组所在的总体;
 - 一般不重复使用的分组;
 - 例子: 人群中的家庭、总体中的个体;

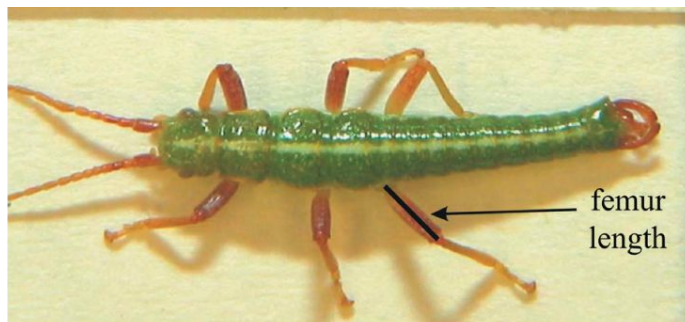
5.1 随机效应方差分析 random-effects ANOVA

- 主要应用场景

- 估计方差成分（variance components）来评估测量误差的重要性；
- 量化：随机抽取的分组间的方差 vs. 组内的方差；

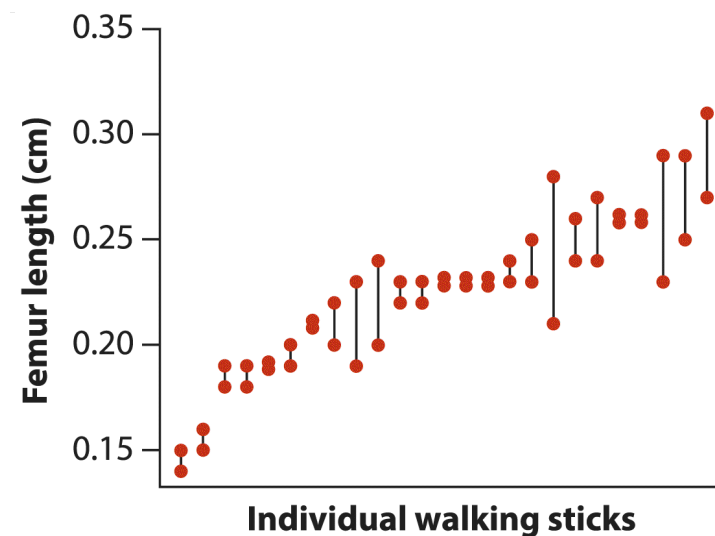
- 例子：竹节虫的股骨 [EXAMPLE 15.6: Walking-stick limbs](#)

- 数据：25个标本，两次股骨长度测量；
- 分组：每个竹节虫个体为一组；



© Patrick Nosil

##	specimen	femurLength	meanFemur
## 1	1	0.26	0.26
## 2	1	0.26	0.26
## 3	2	0.23	0.21
## 4	2	0.19	0.21
## 5	3	0.25	0.24
## 6	3	0.23	0.24



5.1 随机效应方差分析 random-effects ANOVA

- 例子：竹节虫的股骨 [EXAMPLE 15.6: Walking-stick limbs](#)
- 方差成分：包括两个水平的随机效应
 - 组内变异 within-group variance
 - σ^2
 - ANOVA assumption假设：各组的组内变异是一致的；
 - 该例子中是因测量误差导致的（measurement error）；
 - 组间变异 among-group variance
 - σ_A^2
 - 每组有自己的均值；
 - 各组的均值服从某个正态分布；
 - 该正态分布服从 $N(\mu_A, \sigma_A)$

组间 {
组内1 {
组内2 {
组内3 {

##	specimen	femurLength	meanFemur
## 1	1	0.26	0.26
## 2	1	0.26	0.26
## 3	2	0.23	0.21
## 4	2	0.19	0.21
## 5	3	0.25	0.24
## 6	3	0.23	0.24

5.1 随机效应方差分析 random-effects ANOVA

- 例子：竹节虫的股骨 [EXAMPLE 15.6: Walking-stick limbs](#)
- 方差成分：包括两个水平的随机效应
 - 组内变异 within-group variance
 - σ^2
 - 最佳估计为： MS_{error}
 - 组间变异 among-group variance
 - σ_A^2
 - 最佳估计为： $s_A^2 = \frac{MS_{\text{groups}} - MS_{\text{error}}}{n}$
 - n 为每组内的样本量；
 - 前提：每组内样本量一样；

ANOVA table

```
> anova(lm(femurLength ~ specimen, data = walkingstick))
Analysis of Variance Table

Response: femurLength
          Df    Sum Sq   Mean Sq F value    Pr(>F)
specimen   24  0.059132   0.0024638   6.9209 4.077e-06 ***
Residuals  25  0.008900   0.0003560
```

$$MS_{\text{error}} = 0.000356 \text{ cm}^2$$

$$s_A^2 = \frac{0.002464 - 0.000356}{2} = 0.00105 \text{ cm}^2$$

5.1 随机效应方差分析 random-effects ANOVA

- 例子：竹节虫的股骨 [EXAMPLE 15.6: Walking-stick limbs](#)
- 方差成分：包括两个水平的随机效应
 - 组内变异 within-group variance
 - σ^2
 - 最佳估计为： MS_{error}
 - 组间变异 among-group variance
 - σ_A^2
 - 最佳估计为： $s_A^2 = \frac{MS_{\text{groups}} - MS_{\text{error}}}{n}$
 - n 为每组内的样本量；
 - 前提：每组内样本量一样；

$$MS_{\text{error}} = 0.000356 \text{ cm}^2$$

$$s_A^2 = \frac{0.002464 - 0.000356}{2} = 0.00105 \text{ cm}^2$$

```
walkingstickAnova <- lme(fixed = femurLength ~ 1,  
                        random = ~ 1|specimen, data = walkingstick)
```

```
walkingstickVarcomp <- VarCorr(walkingstickAnova)  
walkingstickVarcomp
```

```
## specimen = pdLogChol(1)  
##      Variance StdDev  
## (Intercept) 0.0010539182 0.03246411  
## Residual    0.0003559996 0.01886795
```

5.2 随机效应的重复性 Repeatability

- 例子：竹节虫的股骨 EXAMPLE 15.6: Walking-stick limbs

- 重复性 Repeatability

- $$\frac{s_A^2}{s_A^2 + MS_{\text{error}}}$$

- 即，组间成分所占比例
- 分母：the total amount of measurement variance in the population
- Repeatability 趋近 0，则表明数据变异主要由组内变异导致；
- Repeatability 趋近 1，则表明数据变异主要由组间变异导致；
- 反映了方差成分的大小，是对特定总体参数的估计；
 - 它只适用于随机效应；
 - 不要与 R^2 混淆（基于平方和，可用于固定/随机效应）；

$$MS_{\text{error}} = 0.000356 \text{ cm}^2$$

$$s_A^2 = \frac{0.002464 - 0.000356}{2} = 0.00105 \text{ cm}^2$$

$$\text{Repeatability} = \frac{0.00105}{0.002464 - 0.000356} = 0.75$$

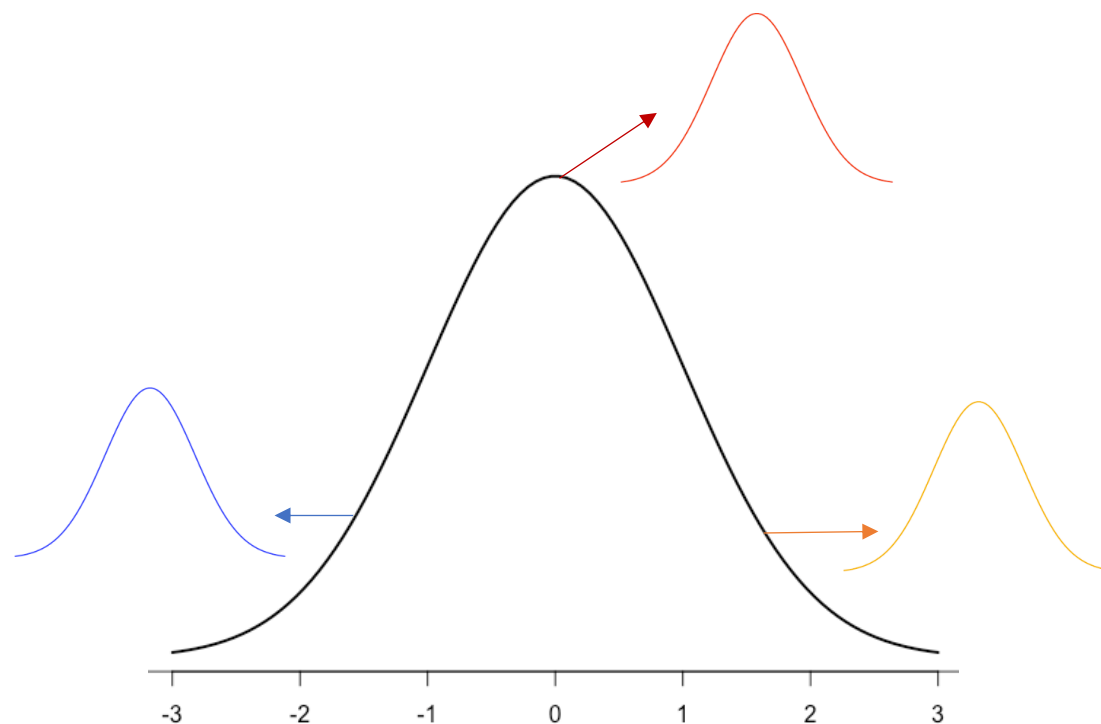
5.3 随机效应方差分析的前提假设

- Assumptions for random-effects ANOVA

- 各组组长内数据为随机样本;
- 各组组长内数据均服从正态分布;
- 各组的方差相等;

与 one-way ANOVA 一致

- 各组组长间数据也构成随机样本;
- 各组组长间数据也服从正态分布;



6. 小结

- 方差分析（ANOVA）检验多组均值之间的差异
 - 通过比较组内均方和组间均方来进行；
 - 统计检验量为： $F = \frac{MS_{\text{groups}}}{MS_{\text{error}}}$
 - 当 $F > 1$ 时，拒绝零假设，表明组间差异大于仅由抽样误差导致的差异；
 - R^2 衡量了组间设置对解释数据变异的百分比；
- 方差分析的前提假设包括
 - 随机样本、变量在各组内呈正态分布、不同组的总体的方差一致；
 - 方差分析对偏离正态分布和方差不一致具有稳健性；
- Kruskal-Wallis检验是方差分析的非参替代；
- 方差分析中的组间设置属于固定效应
 - 预先确定的、可重复的类别；因此结果只适用于研究中包含的组间设置；

6. 小结

- 计划比较的配对均值比较较少
 - 仅代表在数据收集和分析前确定为关键的比较;
- 非计划比较是为了寻找有趣的模式而进行的更全面的比较
 - 例如所有的配对比较组合;
 - 计划外比较需要特殊的方法来防止高I类错误率;
 - Tukey-Kramer 方法是最常用的对所有均值配对进行比较的方法;
- 在随机效应方差分析中各组是从总体中随机抽取的样本
 - 随机效应方差分析的结果可以推广到群体;
 - 重复性估计了随机效应方差分析中各组间总方差的比例, 常用于评估测量误差的重要性;

1.1 试验设计——课堂练习答案

6. Identify the consequences (i.e., increase, decrease, or none) that the following procedures are likely to have on both bias and sampling error in an observational study.

- Matching sampling units between treatment and control
 - Decrease bias (reduces effects of confounding variables);
 - decrease sampling error (by grouping similar units into pairs).
- Increasing sample size
 - No effect on bias; reduce sampling error.
- Ensuring that the frequency distribution of subject ages is the same in the two treatments
 - Decrease bias (corrects for effects of age, a possible confounding variable);
 - No effect on sampling error.
- Using a balanced design
 - No effect on bias; reduce sampling error.