



上海交通大学  
SHANGHAI JIAO TONG UNIVERSITY



# 生物医学统计概论(BI148-2)

Fundamentals of Biomedical Statistics

授课：林关宁

2022 春季



# 课程内容安排

| 上课日期 | 章节 | 教学内容            | 教学要点  | 作业           | 随堂测          | 学时 |
|------|----|-----------------|---|--------------|--------------|----|
| 2.16 | 1  | 数据可视化, 描述性统计    | 1. 课程介绍 & 数据类型                                      | 作业1<br>(8%)  | 测试1<br>(8%)  | 2  |
| 2.23 |    |                 | 2. 描述性统计Descriptive Statistics & 数据常用可视化            |              |              | 2  |
| 3.2  |    |                 | 3. 大数定理 & 中心极限定理                                    |              |              | 2  |
| 3.9  |    |                 | 4. 常用概率分布   |              |              | 2  |
| 3.16 | 2  | 推断性统计, 均值差异检验   | 5. 统计推断基础-1: 置信区间 Confidence Interval *             | 作业2<br>(10%) | 测试2<br>(10%) | 2  |
| 3.23 |    |                 | 6. 统计推断基础-2: 假设检验, I及II类错误, 统计量, p-值                |              |              | 2  |
| 3.30 |    |                 | 7. 数值数据的均值比较-1: 单样本及双样本t-检验, 效应量, 功效                |              |              | 2  |
| 4.6  |    |                 | 8. 数值数据的均值比较-2: One-Way ANOVA, 正态性检验                |              |              | 2  |
| 4.13 |    |                 | 9. 数值数据的均值比较-3: Two-Way ANOVA                       |              |              | 2  |
| 4.20 | 3  | 比例差异检验          | 10. 样本和置信区间预估 *                                     | 作业3<br>(6%)  | 测试3<br>(6%)  | 2  |
| 4.27 |    |                 | 11. 类别数据的比例比较-1: 单样本比例推断                            |              |              | 2  |
| 5.7  |    |                 | 12. 类别数据的比例比较-2: 联立表的卡方检验                           |              |              | 2  |
| 5.11 | 4  | 协方差, 相关分析, 回归分析 | 13. 相关分析 (Pearson r, Spearman rho, Kendal' s tau) * | 作业4<br>(6%)  | 测试4<br>(6%)  | 2  |
| 5.18 |    |                 | 14. 简单回归分析  |              |              | 2  |
| 5.25 |    |                 | 15. 多元回归Multiple Regression                         |              |              | 2  |
| 6.1  | 5  | Course Summary  | 16. 课程总结 *  |              |              | 2  |
|      |    |                 | Total   | 30%          | 30%          | 32 |

\* 随堂测试

# 上节课回顾

## • 什么是正态性检验？有哪几种方法？

### • 正态性检验：偏度和峰度

1. 偏度 (Skewness)：描述数据分布不对称的方向及其程度

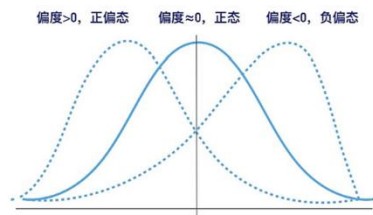


图1. 数据分布的偏态特征

2. 峰度 (Kurtosis)：描述数据分布形态的陡缓程度

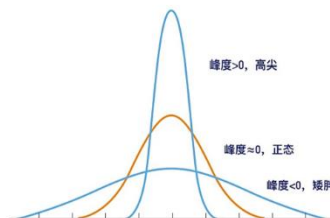
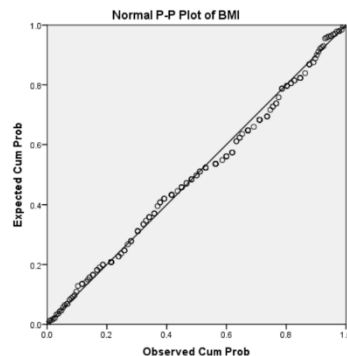
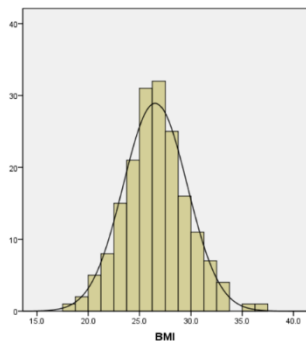


图2. 数据分布的峰态特征

### • 正态性检验：图形判断

直方图+QQ plot



正态性检验属于非参数检验，零假设为“样本来自的总体与正态分布**无显著性差异**，即符合正态分布”，也就是说 $P > 0.05$ 才能说明数据符合正态分布

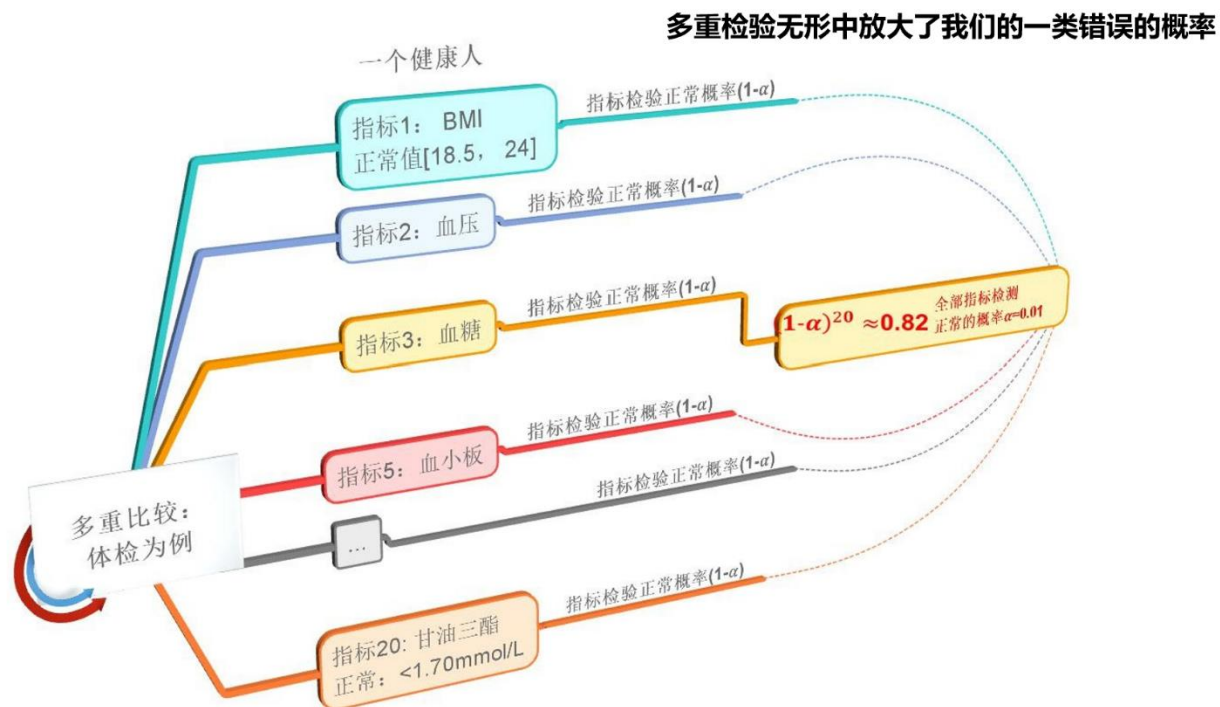
通常正态分布的检验方法有两种，一种是Shapiro-Wilk检验，适用于小样本资料（样本量 $\leq 5000$ ），另一种是Kolmogorov-Smirnov检验，适用于大样本资料（ $> 5000$ ）

# 上节课回顾

- 什么是正态性检验？有哪几种方法？
- 什么是多重比较？能否举个例子？

也就是说，**一个小效率事件就在多次反复尝试后，变成了一个多次出现的事件**（“常在河边走，怎能不湿鞋”）。如果这20,000个基因中有200个基因真实存在差异的，在  $p$  value 为1%的阈值标准下，我们可能会得出399个基因有差异的结论（阳性结果的错误率，即FDR值约为50%）。

可以看到，在进行多次检验后（也就是所说的多重检验，multiple test），基于单次比较的检验标准将变得过于宽松，使得阳性结果中的错误率（FDR值）已经大到令人不可忍受的地步。

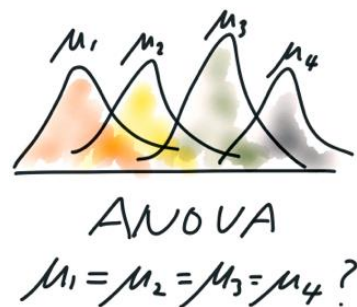


# 上节课回顾

- 什么是正态性检验？有哪几种方法？
- 什么是多重比较？能否举个例子？
- 方差分析和假设检验的区别是什么？

## 方差分析 Vs 假设检验

- 假设检验：一次只能研究2个样本
  - 需要比较的次数随因素的数量增多而增加
  - Type I error 发生的可能性增大
- 方差分析：**同时分析多个样本**
  - 提高检验效率
  - 将所有信息结合在一起，增加了分析的可靠性



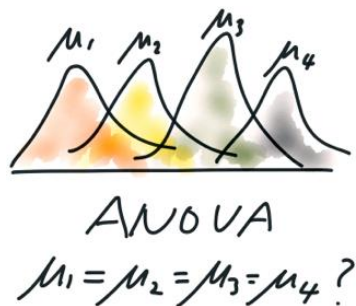


# 上节课回顾

- 什么是正态性检验？有哪几种方法？
- 什么是多重比较？能否举个例子？
- 方差分析和假设检验的区别是什么？方差分析的统计量？怎么计算的？

## 方差分析 Vs 假设检验

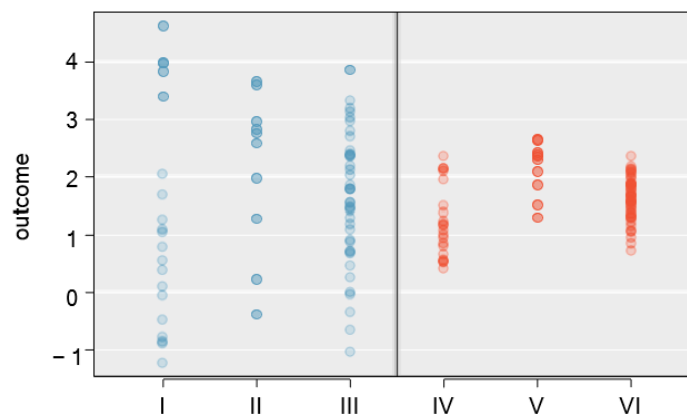
- 假设检验：一次只能研究2个样本
  - 需要比较的次数随因素的数量增多而增加
  - Type I error 发生的可能性增大
- 方差分析：**同时分析多个样本**
  - 提高检验效率
  - 将所有信息结合在一起，增加了分析的可靠性



The *F*-statistic is the test statistic for ANOVA. **F统计量是方差分析的检验统计量**

$$F = \frac{\text{variance between groups}}{\text{variance within groups}} = \frac{MSG}{MSE} \quad \begin{array}{l} \text{组间差异} \\ \text{组内差异} \end{array}$$

- When the population means are equal, the *F*-statistic is approximately 1. **当总体均值相等时，F统计量约为1**
- When the population means differ, *F* will be larger than 1. Larger values of *F* represent stronger evidence against the null. **当总体均值不同时，F将大于1。F的值越大，则表示拒绝零假设的证据越有力。**



哪个F值大？

# 上节课回顾

- 什么是正态性检验？有哪几种方法？
- 什么是多重比较？能否举个例子？
- 方差分析和假设检验的区别是什么？方差分析的统计量？怎么计算的？
- 方差分析的使用前提？

It is important to check whether the assumptions for conducting ANOVA are reasonably satisfied: 进行方差分析前需要考虑，数据假设是否合理满足前提要求

独立性

1. Observations independent within and across groups 组内和组间的观察值都是独立的观察值
  - Think about study design/context

正态性

2. Data within each group are nearly normal 每组内的观察值基本符合正态分布
  - Look at the data graphically, such as with a histogram 以图形方式查看数据，例如使用直方图
  - Normal Q-Q plots can help... 也可以使用 Q-Q图

方差齐性

3. Variability across groups is about equal 每组的组内差异类似
  - Look at the data graphically 直接画图观察
  - Numerical rule of thumb: ratio of largest variance to smallest variance  $< 3$  is considered “about equal”

经验：最大方差与最小方差之比 $<3$ 被认为 “大致相等”



# 课程内容安排

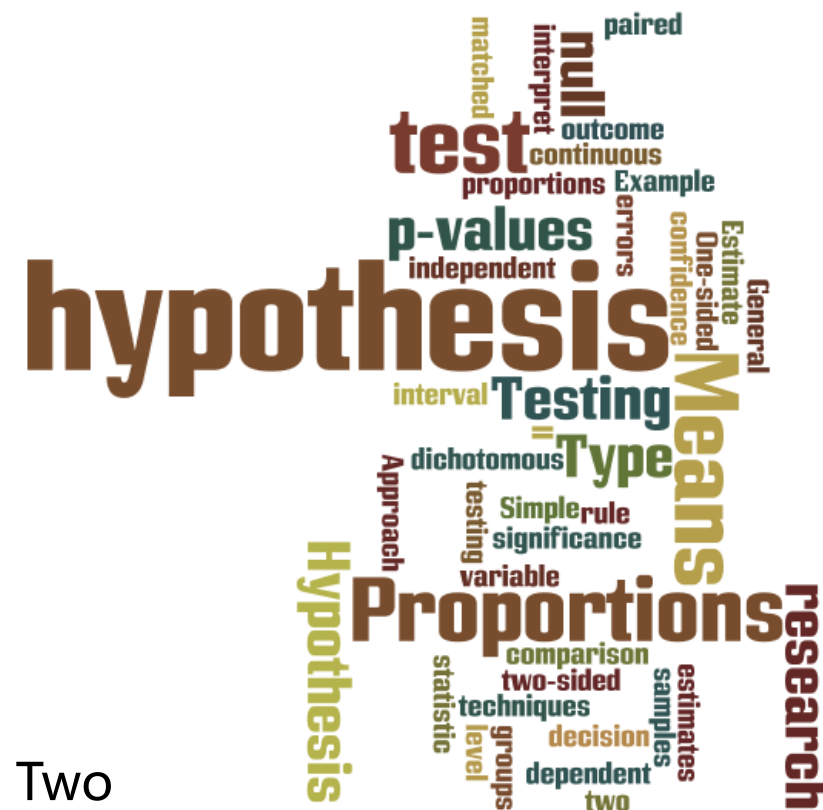
| 上课日期 | 章节 | 教学内容            | 教学要点  | 作业           | 随堂测          | 学时 |
|------|----|-----------------|---|--------------|--------------|----|
| 2.16 | 1  | 数据可视化, 描述性统计    | 1. 课程介绍 & 数据类型                                      | 作业1<br>(8%)  | 测试1<br>(8%)  | 2  |
| 2.23 |    |                 | 2. 描述性统计Descriptive Statistics & 数据常用可视化            |              |              | 2  |
| 3.2  |    |                 | 3. 大数定理 & 中心极限定理                                    |              |              | 2  |
| 3.9  |    |                 | 4. 常用概率分布   |              |              | 2  |
| 3.16 | 2  | 推断性统计, 均值差异检验   | 5. 统计推断基础-1: 置信区间 Confidence Interval *             | 作业2<br>(10%) | 测试2<br>(10%) | 2  |
| 3.23 |    |                 | 6. 统计推断基础-2: 假设检验, I及II类错误, 统计量, p-值                |              |              | 2  |
| 3.30 |    |                 | 7. 数值数据的均值比较-1: 单样本及双样本t-检验, 效应量, 功效                |              |              | 2  |
| 4.6  |    |                 | 8. 数值数据的均值比较-2: One-Way ANOVA, 正态性检验                |              |              | 2  |
| 4.13 |    |                 | 9. 数值数据的均值比较-3: Two-Way ANOVA                       |              |              | 2  |
| 4.20 | 3  | 比例差异检验          | 10. 样本和置信区间预估 *                                     | 作业3<br>(6%)  | 测试3<br>(6%)  | 2  |
| 4.27 |    |                 | 11. 类别数据的比例比较-1: 单样本比例推断                            |              |              | 2  |
| 5.7  |    |                 | 12. 类别数据的比例比较-2: 联立表的卡方检验                           |              |              | 2  |
| 5.11 | 4  | 协方差, 相关分析, 回归分析 | 13. 相关分析 (Pearson r, Spearman rho, Kendal' s tau) * | 作业4<br>(6%)  | 测试4<br>(6%)  | 2  |
| 5.18 |    |                 | 14. 简单回归分析  |              |              | 2  |
| 5.25 |    |                 | 15. 多元回归Multiple Regression                         |              |              | 2  |
| 6.1  | 5  | Course Summary  | 16. 课程总结 *  |              |              | 2  |
|      |    |                 | Total   | 30%          | 30%          | 32 |

\* 随堂测试



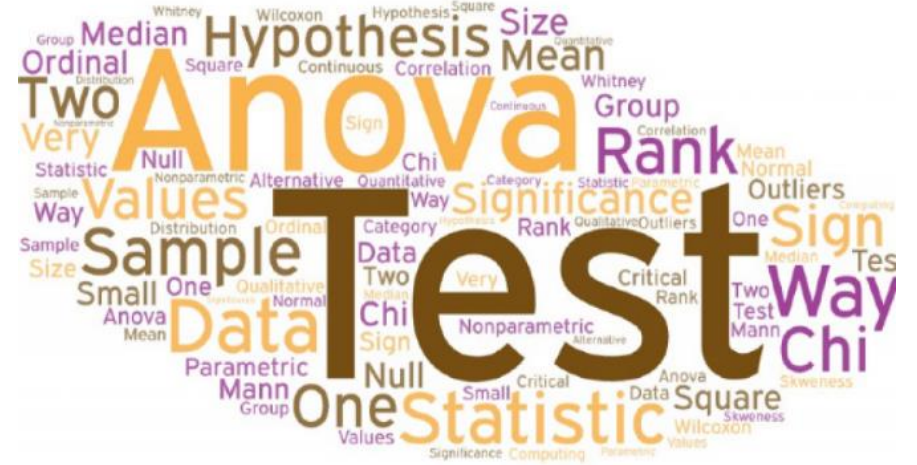
# 单元2内容 (week 5-9)

- 总体、样本
  - 参数、统计量
- 置信区间
  - 区间估计、总体方差
- 均值差异
  - One-sample 均值估计
  - Two-sample 均值差异
- 零假设检验
  - $H_0$  vs  $H_a$ , Type I & II errors, P-value, One-sample vs Two sample test
- 样本量、效应量、功效 (Sample size, Effect size, Power)
- 方差分析
  - ANOVA test



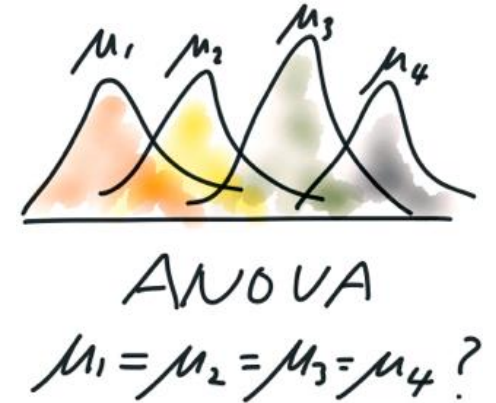


**back**



# 方差分析 (Analysis of Variance, ANOVA)

- 研究分类型自变量对数值型应变量的影响
  - 1个或多个分类型自变量
  - 1个数值型因变量
- 通过检验多个总体均值是否相等来判断是否有显著影响
  - 通过分析数据的误差 -> 判断各总体均值是否相等
- 方差分析 Vs 假设检验
  - 假设检验：一次只能研究2个样本
    - 需要比较的次数随因素的数量增多而增加
    - Type I error 发生的可能性增大
  - 方差分析：同时分析多个样本
    - 提高检验效率
    - 将所有信息结合在一起，增加了分析的可靠性



- ✓ • 有单因素方差分析和双因素方差分析
  - 单因素方差分析 (One-way ANOVA)
    - 涉及1个分类型自变量对数值因变量影响
  - 双因素方差分析 (Two-way ANOVA)
    - 涉及2个自变量对数值因变量影响
    - 它分为只考虑主效应 (main effect) 的双因素方差分析和考虑交互效应 (interaction) 的双因素方差分析

Tests of Between-Subjects Effects

Dependent Variable: Score on Beck's Depression Inventory

| Source            | Type III Sum of Squares | df | Mean Square | F        | Sig. | Partial Eta Squared |
|-------------------|-------------------------|----|-------------|----------|------|---------------------|
| Corrected Model   | 3687.775 <sup>a</sup>   |    |             |          |      |                     |
| Intercept         | 117629.949              | 1  | 117629.949  | 3199.098 | .000 | .972                |
| gender            | 332.202                 | 1  | 332.202     | 9.035    | .003 | .089                |
| medicine          | 2519.647                | 3  | 839.882     | 22.842   | .000 | .427                |
| gender * medicine | 567.784                 | 3  | 189.261     | 5.147    | .002 | .144                |
| Error             | 3382.815                | 92 | 36.770      |          |      |                     |
| Total             | 128105.000              |    |             |          |      |                     |
| Corrected Total   | 7070.590                | 99 |             |          |      |                     |

MAIN EFFECTS: GENDER AND MEDICINE

INTERACTION EFFECT: GENDER \* MEDICINE

a. R Squared = .522 (Adjusted R Squared = .485)

© 2018 www.spss-tutorials.com

# 方差分析的理念

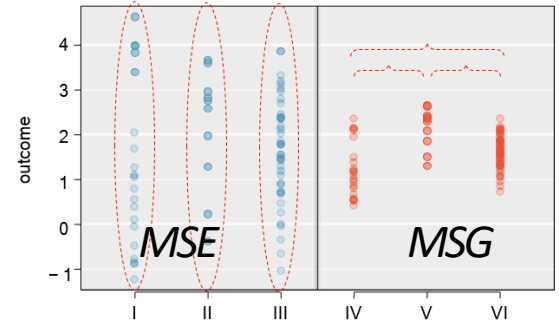
Under the null hypothesis, there is no real difference between the groups; thus, any observed variation in group means is due to chance. 在零假设成立下，组间是没有真正的差异；因此，观察到的总体平均数的任何变化都是偶然的

- Think of all observations as belonging to a single group. 所有的观察值都看作是属于同一个总体
- Variability between group means should equal variability within groups 组间有差异意味着组内是不存在差异的

The *F-statistic* is the test statistic for ANOVA. F统计量是方差分析的检验统计量

$$F = \frac{\text{variance between groups}}{\text{variance within groups}} = \frac{MSG}{MSE}$$

组间差异  
组内差异



当总体均值相等时，F统计量约为1

- When the population means are equal, the *F*-statistic is approximately 1.
- When the population means differ, *F* will be larger than 1. Larger values of *F* represent stronger evidence against the null. 当总体均值不同时，F将大于1。F的值越大，则表示拒绝零假设的证据越有力。
- The *F* statistic follows an *F* distribution, with two degrees of freedom,

$$df_1 \text{ and } df_2; df_1 = n_{\text{groups}} - 1, df_2 = n_{\text{obs}} - n_{\text{groups}}.$$

F统计量遵循F分布，有两个自由度 $df_1$ 和 $df_2$

- The *p*-value for the *F*-statistic is the probability *F* is larger than the *F*-statistic.

p值是得到的F值 > 给定的F统计量的概率





# One-way ANOVA 单因素方差分析对样本的要求

- **独立性 (统计学检验: 无)**

- 各样本必须是相互独立的随机样本
- 样本的个体数量尽可能相等或相差不大 (最大数量/最小数量  $\leq 1.5$ )

assumptions of ANOVA:

- Independence of Observations
  - Cannot be tested statistically, is determined by research methodology only
- Normally Distributed Data
  - Shapiro-Wilk's  $W$  statistic, if significant, indicates significant non-normality in data

- **正态性 (统计学检验: Shapiro-Wilk test)**

- 样本的总体符合正态分布, 偏态分布不适用于方差分析 (**Normality test**)
- 对偏态分布应考虑用对数转换、平方根变换、倒数变换、平方根反正弦变换等变量变换方法变为正态或接近正态分布后再进行方差分析
- 但一般来说, 只要数据量  $> 30$ , 就可以认为是正态分布

- **方差齐性 (统计学检验: Levene's test)**

- 各组样本具有相同的方差 (**homogeneity of variance**)
- Note: Robust to heterogeneity of variance - the largest variance is not more than **4 times** the smallest variance

- **Note**

- 方差分析对独立性要求高, 若独立性得不到满足, 方差分析结构往往受到较大影响
- 对正态性和方差齐性要求比较宽松, 当正态性得不到满足和方差略有参差不齐时, 对影响不是很大。

## Homogeneity of Variance

- Assumes equal variance of scores across groups
  - i.e. *homogeneity of variance*
- *Levene's Test for Equality of Variances* tests whether two sample variances are approximately equal
- As long as  $N > 30$  and  $n_1 \approx n_2 \approx n_3 \dots$ , ANOVA is robust to violations of homogeneity of variance

# Levene's test for equality of variance

- To perform Levene's Test:
  1. Calculate each  $z_{ij} = |y_{ij} - \bar{y}_i|$ .
  2. Run an ANOVA on the set of  $z_{ij}$  values.
  3. If  $p\text{-value} \leq \alpha$ , reject  $H_o$  and conclude the variances are not all equal.
- Levene's Test is robust because the true significance level is very close to the nominal significance level for a large variety of distributions.
- It is not sensitive to symmetric heavy-tailed distributions (such as the double exponential and student's  $t$  distributions).





# ANOVA 类型1：独立样本 one-way ANOVA（单因素方差分析）

单因素方差分析（ANOVA）用于确定三个或三个以上独立（不相关）组的均值之间是否存在统计显著性差异

单因素方差分析比较我们感兴趣的组之间的平均值，并确定这些平均值中是否有任何一个在统计学上显著不同。具体来说，它检验了零假设：

$$H_0: \mu_1 = \mu_2 = \mu_3 = \cdots = \mu_k$$

其中  $\mu$ =组平均值， $k$ =组数。然而，如果单向方差分析返回了具有统计学意义的结果，我们接受备择假设（ $H_A$ ），即至少有两组均值在统计学上存在显著差异。

重要的是：单向方差分析是一种综合测试统计，不能告诉你哪些特定组在统计学上彼此存在显著差异，只有至少两组存在显著差异。要确定哪些特定组彼此不同，需要使用事后测试（post hoc test）



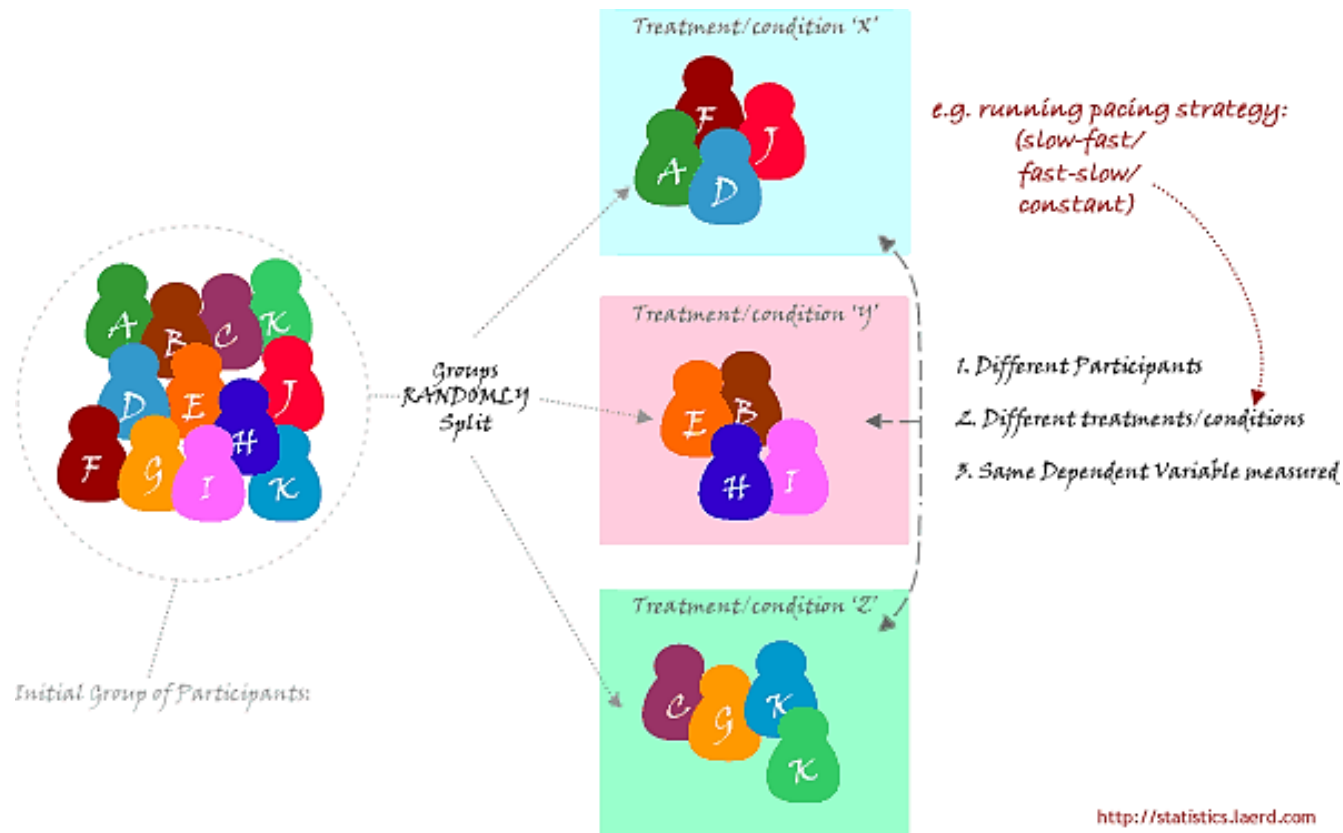
# ANOVA 类型1：独立样本 one-way ANOVA (单因素方差分析)

## 什么时候需要做这个测试？

如果研究对象是人的话，很可能在**两种不同类型**的研究设计中遇到这种情况：

(1) 一种研究设计是招募一组人，然后将该群体随机分为三个或更多更小的群体（即，每个参与者被分配到一个且仅一个群体）。然后，你让每组承担不同的任务（或将其置于不同的条件下），并在同一因变量上测量结果/反应。

例如，一位研究人员希望知道不同的配速策略是否会影响完成马拉松的时间。研究人员随机将一组志愿者分为两组：(a) 先慢后快，(b) 先快后慢，或 (c) 全程以稳定的速度跑步。完成马拉松的**时间是结果（因变量）**。该研究设计如右图所示：



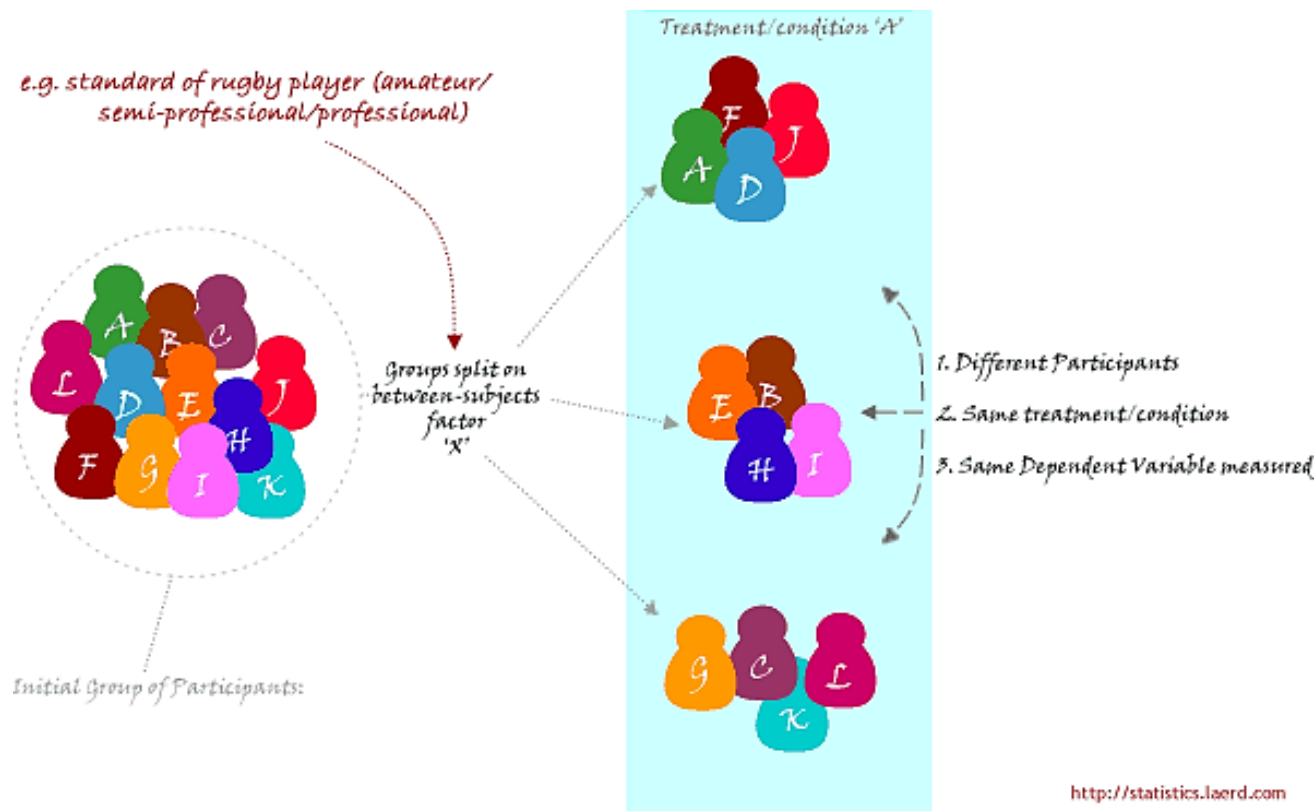
# ANOVA 类型1：独立样本 one-way ANOVA (单因素方差分析)

## 什么时候需要做这个测试？

如果研究对象是人的话，很可能在**两种不同类型**的研究设计中遇到这种情况：

(2) 第二个研究设计是招募一组人，然后根据一些自变量将他们分成若干组。同样，每个人只会被分配到一个小组。这个自变量有时被称为属性自变量，因为你是根据他们拥有的某些属性（例如，他们的教育水平；每个人都有一个教育水平，即使是“无”）。然后，每组在经历相同任务或条件（或完全没有）的同一因变量上进行测量。

例如，研究人员想确定业余、半职业和职业橄榄球运动员之间的腿部力量是否存在差异。在等速机器上测量的力/强度是因变量。此类研究设计右图所示：



# One-way ANOVA 计算

Treatments  
Error (or  
Residual)

|         | Degrees of freedom | Sum of squares   | Mean square        |                 |
|---------|--------------------|--|--------------------|-----------------|
| source  | Df                 | SS   | MSS                | F               |
| Between | $df_b = k - 1$     | $SSB = \sum_j n_j (\bar{x}_j - \bar{x})^2$             | $MSB = SSB / df_b$ | $F = MSB / MSW$ |
| Within  | $df_w = n - k$     | $SSW = \sum_j \sum_i (x_{ij} - \bar{x}_j)^2$           | $MSW = SSW / df_w$ |                 |
| Total   | $df_t = n - 1$     | $SST = SSB + SSW = \sum_j \sum_i (x_{ij} - \bar{x})^2$ |                    |                 |

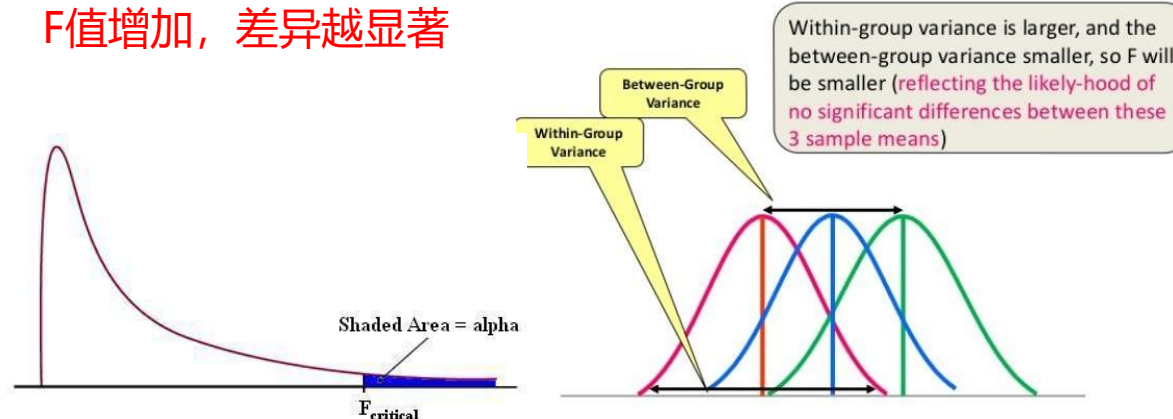
最后是统计决策将统计量的值F与给定的显著性水平 $\alpha$ 的临界值 $F_\alpha$ 进行比较，作出对原假设 $H_0$ 的决策。

K=组数，n=总样本数

- 根据给定的显著性水平 $\alpha$ ，在F分布表中查找与第一自由度 $df_1 = k - 1$ 、第二自由度 $df_2 = n - k$ 相应的临界值 $F_\alpha$ 。
- 若 $F > F_\alpha$ ，则拒绝原假设 $H_0$ ，表明均值之间的差异是显著的，所检验的因子对观察值有显著影响。
- 若 $F < F_\alpha$ ，则不能拒绝原假设 $H_0$ ，无证据支持表明所检验的因子对观察值有显著影响。

结论：

MSB组间 增大 and/or MSW组内减小，  
F值增加，差异越显著



做一个进一步的总结：

- (1) 提出假设；
- (2) 构造检验统计量；

均值：全部观察值的总均值、处理的均值。

平方和：总平方和SST，处理平方和SSR，误差平方和SSE。

均方：处理均方MSTR，误差均方MSE。

均方比：MSTR/MSE~F分布。

- (3) 统计决策。

手动计算的例子：<https://sixsigmastudyguide.com/anova-analysis-of-variation/>

# One-way ANOVA 结果

You will have calculated the following results or obtained them from SPSS Statistics or other tools

Structure of results:

| Source  | SS            | df    | MS     | F           | Sig.      |
|---------|---------------|-------|--------|-------------|-----------|
| Between | $SS_b$        | $k-1$ | $MS_b$ | $MS_b/MS_w$ | $p$ value |
| Within  | $SS_w$        | $N-k$ | $MS_w$ |             |           |
| Total   | $SS_b + SS_w$ | $N-1$ |        |             |           |

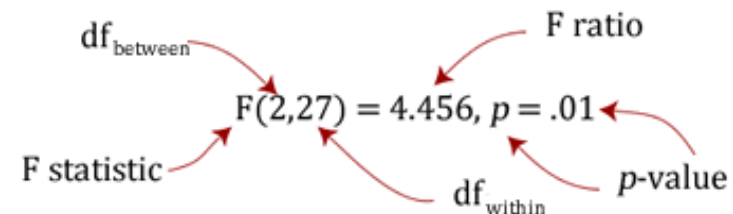
An example:

| Source  | SS      | df | MS     | F     | Sig. |
|---------|---------|----|--------|-------|------|
| Between | 91.476  | 2  | 45.733 | 4.467 | .021 |
| Within  | 276.400 | 27 | 10.237 |       |      |
| Total   | 367.867 | 29 |        |       |      |

## 结果汇报:

通过单因素方差分析

ANOVA ( $F(2, 27) = 4.467, p = .021$ ), 两组之间存在统计学显著差异



where  $df$  = degrees of freedom

Note: 不要把结果报告为“显著差异”，而应将其报告为“**统计**显著差异”。这是因为你对结果是否重要的决定，不应该仅仅基于你的统计检验，还需要看效应量和功效。

因此，需要向读者表明这个“意义”是一个统计意义。



# One-way ANOVA 后续

我的方差分析的 p 值大于 0.05，现在该怎么办？

报告单因素方差分析的结果（例如，“根据单因素方差分析 ( $F(2,27) = 1.397$ ,  $p = 0.15$ ) 确定的各组平均数之间没有统计学显著差异”）。未达到统计显著性结果并不意味着不应报告组平均值  $\pm$  标准差。然而，通常不会进行 post hoc test 事后测试，也不应进行事后测试。

我的 p 值小于 0.05，现在该怎么办？

首先，你需要先报告你的结果。再进行 post hoc test。

什么是 post hoc test（两两样本之间的多重比较 或 事后分析）？

方差分析测试会告诉组之间是否存在总体差异，但它不会告诉你哪些特定组存在差异——这需要“两两样本之间的多重比较”。只有在组平均值显示出总体统计显著差异（即，具有统计显著性的单向方差分析结果）时，才应进行两两样本之间的多重比较。两两样本之间的多重比较需要控制实验误差率（通常  $\alpha = 0.05$ ），方法与使用多个 t 检验相同。也被称为后验测试；也就是说，在事件之后进行。





# post hoc test (两两样本之间的多重比较)

假设在总体的单因素方差分析结果是显著的，因此我们得出结论，并非所有的总体均值都是相等的。我们现在以事后（事后）比较的形式一次比较两个。我们进行以下六项测试

Test 1:  $H_0: \mu_1 = \mu_2$  vs.  $H_1: \mu_1 \neq \mu_2$

Test 3:  $H_0: \mu_1 = \mu_4$  vs.  $H_1: \mu_1 \neq \mu_4$

Test 5:  $H_0: \mu_2 = \mu_4$  vs.  $H_1: \mu_2 \neq \mu_4$

Test 2:  $H_0: \mu_1 = \mu_3$  vs.  $H_1: \mu_1 \neq \mu_3$

Test 4:  $H_0: \mu_2 = \mu_3$  vs.  $H_1: \mu_2 \neq \mu_3$

Test 6:  $H_0: \mu_3 = \mu_4$  vs.  $H_1: \mu_3 \neq \mu_4$

Consider testing 3 *true* null hypothesis. In using  $\alpha = 0.05$  for each test, the probability of making a correct retention is 0.95. The probability of making three consecutive correct retentions =  $0.95 \times 0.95 \times 0.95 \approx 0.86$ . Therefore, the probability of making at least one incorrect decision =  $1 - 0.86 = 0.14$ . This is the **family-wise type I error rate**.

The simplest of these adjustments is called the ***Bonferroni correction*** (Dunn 1961)

$$p' = m * p$$

“multiply all your raw p-values by m tests”

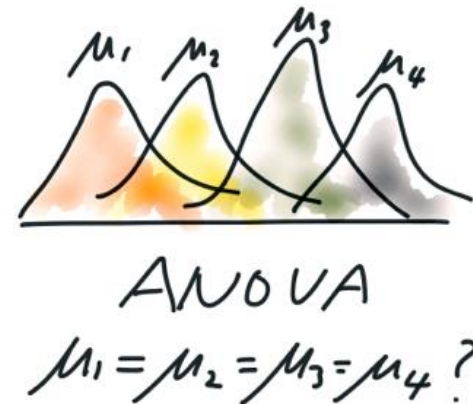


# 双因素方差分析



# 方差分析 (Analysis of Variance, ANOVA)

- 研究分类型自变量对数值型应变量的影响
  - 1个或多个分类型自变量
  - 1个数值型因变量
- 通过检验多个总体均值是否相等来判断是否有显著影响
  - 通过分析数据的误差 -> 判断各总体均值是否相等
- 方差分析 Vs 假设检验
  - 假设检验：一次只能研究2个样本
    - 需要比较的次数随因素的数量增多而增加
    - Type I error 发生的可能性增大
  - 方差分析：同时分析多个样本
    - 提高检验效率
    - 将所有信息结合在一起，增加了分析的可靠性
- 有单因素方差分析和双因素方差分析
  - 单因素方差分析 (One-way ANOVA)
    - 涉及1个分类型自变量对数值因变量影响
  - 双因素方差分析 (Two-way ANOVA)
    - 涉及2个自变量对数值因变量影响
    - 它分为只考虑主效应 (main effect) 的双因素方差分析和考虑交互效应 (interaction) 的双因素方差分析



Tests of Between-Subjects Effects

Dependent Variable: Score on Beck's Depression Inventory

| Source            | Type III Sum of Squares | df | Mean Square | F        | Sig. | Partial Eta Squared |
|-------------------|-------------------------|----|-------------|----------|------|---------------------|
| Corrected Model   | 3687.775 <sup>a</sup>   |    |             |          |      |                     |
| Intercept         | 117629.949              | 1  | 117629.949  | 3199.098 | .000 | .972                |
| gender            | 332.202                 | 1  | 332.202     | 9.035    | .003 | .089                |
| medicine          | 2519.647                | 3  | 839.882     | 22.842   | .000 | .427                |
| gender * medicine | 567.784                 | 3  | 189.261     | 5.147    | .002 | .144                |
| Error             | 3382.815                | 92 | 36.770      |          |      |                     |
| Total             | 128105.000              |    |             |          |      |                     |
| Corrected Total   | 7070.590                | 99 |             |          |      |                     |

a. R Squared = .522 (Adjusted R Squared = .485)

© 2018 www.spss-tutorials.com

# 双因素方差分析 (Two-way ANOVA)

前面介绍完了单因子方差分析，但是当我们的因子大于一个的时候，我们又该怎么分析呢？同样抛个样例问题出来。假设现在我们想了解北京城市人口空间分布是否受不同环路（一环、二环、三环乃至四、五、六环）或新老城区的显著影响。所以该问题是一个典型的双因子问题，可以拆分为如下的情况：

| 因子 | 新城区 | 老城区 |
|----|-----|-----|
| 一环 | 人口  | 人口  |
| 二环 | 人口  | 人口  |
| 三环 | 人口  | 人口  |

对于该问题我们可以考虑用单因子方差分析来解决——即通过考虑两个因子间所有的组合来分析是否有显著影响。（二环+新城区，二环+老城区，三环+新城区，……，六环+老城区）通过这样组合来得到最后的单因子水平。但是这样处理的问题是，我们无法了解到底是新老城区的因素影响了人口的空间分布，或者是不同的环路影响了人口的空间分布，亦或是二者共同影响。所以我们需要新的方法来分析。这就是题目所述的双因子方差分析。



# Two-way ANOVA 双因素方差分析

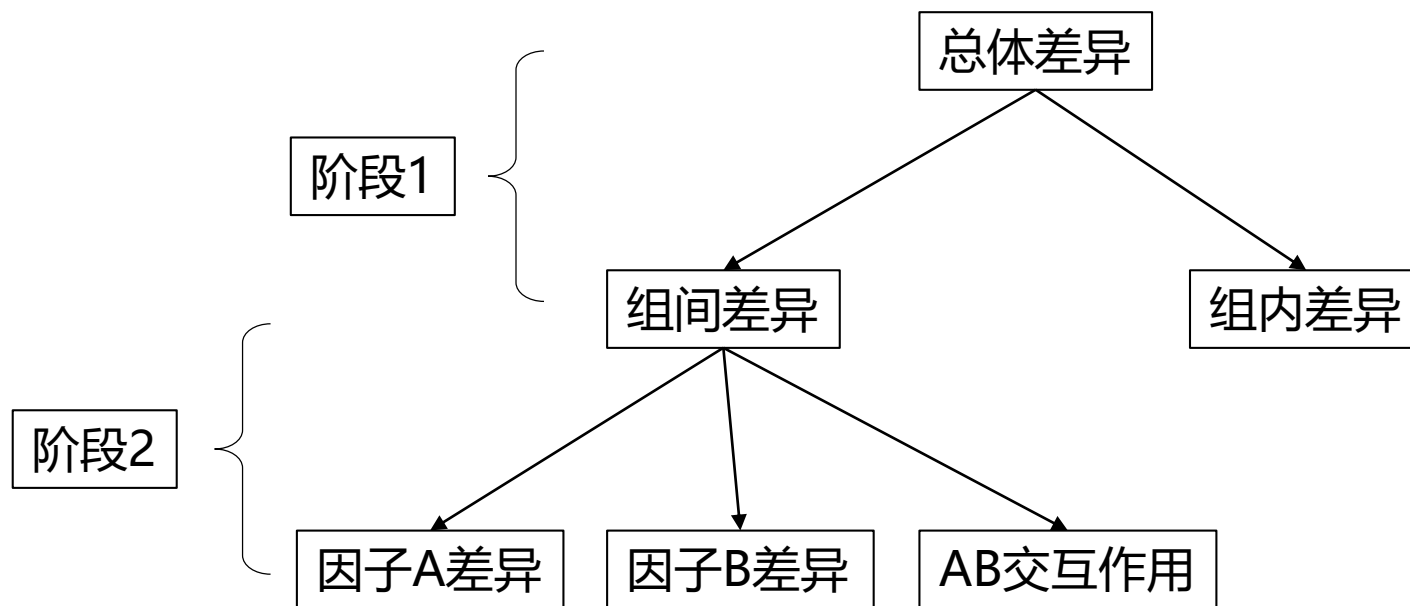
## 双因素方差分析法的类型

双因素方差分析有两种类型：

一个是无交互作用的双因素方差分析，它假定因素A 和因素B的效应之间是相互独立的，不存在相互关系；

另一个是有交互作用的双因素方差分析，它假定因素A和因素B的结合会产生出一种新的效应。

例如，若假定不同地区的消费者对某种颜色有与其他地区消费者不同的特殊偏爱，这就是两个因素结合后产生的新效应，属于有交互作用的背景；否则，就是无交互作用的背景。有交互作用的双因素方差分析已超出本书的范围，这里介绍无交互作用的双因素方差分析。



双因素方差分析的结构



## 一、适用条件

两因素方差分析，需要满足6个条件：

条件1：观察变量唯一，且为连续变量。

条件2：有两个分组变量，且都为分类变量。

条件3：观测值相互独立。

条件4：观察变量不存在显著的异常值。

条件5：各组、各水平观察变量为正态(或近似正态)分布。

条件6：相互比较的各处理水平(组别)的总体方差相等，即通过方差齐性检验。

## 二、统计量计算

本部分以两因素两水平的析因设计为例，介绍两因素方差分析的统计量计算。

| A 因素<br>(两个水平) | B 因素(两个水平)    |               |
|----------------|---------------|---------------|
|                | 水平 1( $b_1$ ) | 水平 2( $b_2$ ) |
| 水平 1( $a_1$ )  | $(a_1b_1)$    | $(a_1b_2)$    |
| 水平 2( $a_2$ )  | $(a_2b_1)$    | $(a_2b_2)$    |

图1 两因素两水平的析因设计示意图

两因素方差分析可以首先按照单因素方差分析的方法把变异分为总变异 $SS_{总}$ 、组间变异 $SS_{组间}$ 、组内变异 $SS_{组内}$

- **单独效应**是指其他因素的水平固定时，同一因素不同水平间的差别。如图1中，当A因素固定在1水平时，B因素的单独效应为 $(a_1b_2)-(a_1b_1)$ 。
- **主效应**是指某一因素各水平间的平均差别。如图1中，B因素水平为1时，A因素的单独效应为 $(a_1b_1)-(a_2b_1)$ ；B因素固定为2时，A因素的单独效应为 $(a_1b_2)-(a_2b_2)$ ，平均后的A因素的主效应为 $[(a_1b_1)-(a_2b_1)+(a_1b_2)-(a_2b_2)]/2$ 。
- **交互作用**是指当因素的各个单独效应随另一因素变化而变化。如图1中，A与B的交互作用表示为 $AB=[(a_2b_2-a_1b_2)-(a_2b_1-a_1b_1)]/2$ 。若存在交互作用，在统计分析时须逐一分析各因素的单独效应。反之，若不存在交互作用，说明两因素的作用效果相互独立，逐一分析各因素的主效应即可。



# ANOVA 类型-2: two-way ANOVA

## (不考虑双因子交互作用的方差分析)

| 被试   | 剂量1        | 剂量2        | 剂量3        | 剂量4        |
|------|------------|------------|------------|------------|
| 年龄段1 | Average 1  | Average 2  | Average 3  | Average 4  |
| 年龄段2 | Average 5  | Average 6  | Average 7  | Average 8  |
| 年龄段3 | Average 9  | Average 10 | Average 11 | Average 12 |
| 年龄段4 | Average 13 | Average 14 | Average 15 | Average 16 |
| 年龄段5 | Average 17 | Average 18 | Average 19 | Average 20 |
| 年龄段6 | Average 21 | Average 22 | Average 23 | Average 24 |

- 因素 (factors), 独立变量: 年龄, 剂量

Main effect

**NHST:**

H0-1: 不同性别组间没有差异

H0-2: 不同剂量组间没有差异

H1-1: 不同性别组间有差异

H1-2: 不同剂量组间有差异



# ANOVA 类型-2: two-way ANOVA

## (考虑双因子交互作用的方差分析)

(除了每个main effect外, 还可以研究因子间 interaction)

| 被试 | 剂量1         | 剂量2         | 剂量3         | 剂量4         |
|----|-------------|-------------|-------------|-------------|
| 饭后 | A11,A12,... | B11,B12,... | C11,C12,... | D11,D12,... |
| 空腹 | A21,A22,... | B21,B22,... | C21,C22,... | D21,D22,... |

- 因素 (factors), 2 独立变量: 时间效应 , 剂量效应
- 时间和剂量的结合 可能产生新的效应

Main effect

Interaction effect

**NHST:**

H0-1: 不同时间组间没有差异

H0-2: 不同剂量组间没有差异

H0-3: 时间和剂量因素间没相互作用

H1-1: 不同时间组间有差异

H1-2: 不同剂量组间有差异

H1-3: 时间和剂量因素间有相互作用



### 例如双因素方差分析的基本步骤

#### 1、提出原假设：

|       | 因素A (r) | 因素B (s) | AB交互作用 (rs) |
|-------|---------|---------|-------------|
| 无显著差异 | H01     | H02     | H03         |
| 有显著差异 | H11     | H12     | H13         |

#### 2、F值检验

|       |                                      |             |
|-------|--------------------------------------|-------------|
| 拒绝H01 | $F_{0.05}(d_{fa}, d_{fe}) < F_A$     | 因素A不同存在显著差异 |
| 拒绝H02 | $F_{0.05}(d_{fb}, d_{fe}) < F_B$     | 因素B不同存在显著差异 |
| 拒绝H03 | $F_{0.05}(d_{fab}, d_{fe}) < F_{AB}$ | 因素A与B存在交互效应 |



| 变异来源           | 自由度 | SS   |
|----------------|-----|--|
| 组间变异           | 3   | $SS_{\text{处理}} = \frac{1}{n} (T_1^2 + T_2^2 + T_3^2 + T_4^2) - C$ |
| <b>A</b> 因素主效应 | 1   | $SS_A = \frac{1}{2n} (A_1^2 + A_2^2) - C$                          |
| <b>B</b> 因素主效应 | 1   | $SS_B = \frac{1}{2n} (B_1^2 + B_2^2) - C$                          |
| <b>AB</b> 交互作用 | 1   | $S_{AB} = SS_{\text{组间}} - SS_A - SS_B$                            |

图2 两因素两水平的设计析因设计 $SS_{\text{组间}}$ 的变异分解



# ANOVA 类型-3: One-way repeated-measures ANOVA (单因素重复测量方差分析)

| 被试       | Day 1     | Day 2     | Day 3     | Day 4     |
|----------|-----------|-----------|-----------|-----------|
| 被试参加相同实验 | 所有被试      | 所有被试      | 所有被试      | 所有被试      |
|          | s1,s2,... | s1,s2,... | s1,s2,... | s1,s2,... |

- 因素 (factor), 独立变量: Day
- 因变量: 检测指标 (一个)
- Between Group (组间) : Day
- Within Group (组内): 不同被试

**NHST:**

H0:组间没有差异

H1:组间有差异

## 重复测量 repeated measures:

where an observation has been made on the same individual more than once. This usually involves measurements taken at different time points



# ANOVA 类型-4: Two-way repeated-measures ANOVA (双因素重复测量方差分析)

| 被试   | 剂量1       | 剂量2        | 剂量3        | 剂量4        |
|------|-----------|------------|------------|------------|
| 年龄段1 | Average 1 | Average 2  | Average 3  | Average 4  |
| 年龄段2 | Average 5 | Average 6  | Average 7  | Average 8  |
| 年龄段3 | Average 9 | Average 10 | Average 11 | Average 12 |

← 无重复检测值的双因素方差分析称为**双因素测量方差分析**(Two-way ANOVA)

| 被试 | 剂量1         | 剂量2             | 剂量3         | 剂量4         |
|----|-------------|-----------------|-------------|-------------|
| 饭后 | A11,A12,... | B11,B12,<br>... | C11,C12,... | D11,D12,... |
| 空腹 | A21,A22,... | B21,B22,<br>... | C21,C22,... | D21,D22,... |

← 有重复检测的因素，这时的双因素方差分析称为**双因素重复测量方差分析**(Two-way repeated-measure ANOVA)



# Two-way ANOVA vs Two-way repeated ANOVA

## Two-way ANOVA

|         | Wheat | Corn | Soy | Rice |
|---------|-------|------|-----|------|
| Blend X | 123   | 138  | 110 | 151  |
| Blend Y | 145   | 165  | 140 | 167  |
| Blend Z | 156   | 176  | 185 | 175  |

**2个因素:**

自变量1: 粮食种类 (wheat, corn, soy, rice)

自变量2: 肥料种类 (blend X, Y, X)

**1个因变量: 1种测量值, 测量值只量1次**

## Two-way repeated measures ANOVA

|            | Crop  |      |     |      |
|------------|-------|------|-----|------|
| Fertilizer | Wheat | Corn | Soy | Rice |
| Blend X    | 123   | 128  | 166 | 151  |
|            | 156   | 150  | 178 | 125  |
|            | 112   | 174  | 187 | 117  |
|            | 100   | 116  | 153 | 155  |
|            | 168   | 109  | 195 | 158  |
| Blend Y    | 135   | 175  | 140 | 167  |
|            | 130   | 132  | 145 | 183  |
|            | 176   | 120  | 159 | 142  |
|            | 120   | 187  | 131 | 167  |
|            | 155   | 184  | 126 | 168  |
| Blend Z    | 156   | 186  | 185 | 175  |
|            | 180   | 138  | 206 | 173  |
|            | 147   | 178  | 188 | 154  |
|            | 146   | 176  | 165 | 191  |
|            | 193   | 190  | 188 | 169  |

**2个因素:**

自变量1: 粮食种类 (wheat, corn, soy, rice)

自变量2: 肥料种类 (blend X, Y, X)

**1个因变量: 1种测量值, 测量值量>1次 (多次重复测量 repeated measure)**



# 小结

- “Two-way” 双向 是指有2个因素（如不同时间，不同条件）
- “Repeated-measures” 重复测量 是指用一个受试者接受了不止一种治疗或是处于不止一种的环境下
- 双因素重复测量方差分析也称为**受试者内部 (within-subjects)** 方差分析
- 双因素重复测量方差分析通常用于在两个或多个时间点测量因变量的研究中，例如当受试者经历两个或多个条件时（即，这两个因素是“时间”和“条件”）的测量值。双向重复测量方差分析除了可以检验受试者在“不同时间”或是“不同条件”下测量是否存在差异外。**双因素重复测量方差分析的一个主要目的是加强验证两个因素是否存在相互作用。**
- 双因素重复测量方差分析的检验可以是两个重复测量的因素，也可以是一个重复测量因素和一个非重复测量因素。但只要存在任何重复因素，则应使用重复测量方差分析



# 双因素重复测量方差分析（例子）

研究A、B两种饲料对家兔的增重效果，选择20只家兔，随机分成两组，第一组用饲料A饲养，第二组用饲料B饲养，并于试验开始后第1个月(time1)、第2个月(time2)、第三个月(time3)分别测量2组家兔体重，试比较A、B两种饲料对家兔的增重效果有无差别

| group | time1 | time2 | time3 |
|-------|-------|-------|-------|
| A     | 2     | 5.3   | 7.4   |
| A     | 1.9   | 3.9   | 6.3   |
| A     | 1.1   | 3     | 5     |
| A     | 1.8   | 5.2   | 6.2   |
| A     | 1.2   | 3.1   | 4.3   |
| A     | 1.1   | 2.5   | 4.7   |
| A     | 1.6   | 4.2   | 5.6   |
| A     | 1.5   | 4     | 5.8   |
| A     | 1.4   | 3.6   | 5.2   |
| A     | 1.8   | 4.3   | 5.5   |
| B     | 1.7   | 3.8   | 5.4   |
| B     | 1.2   | 3.2   | 4.6   |
| B     | 1     | 2.8   | 5.8   |
| B     | 1.9   | 4.3   | 7.6   |
| B     | 1.8   | 4.9   | 6.8   |
| B     | 1.4   | 4.1   | 5.8   |
| B     | 1.5   | 3.8   | 5.2   |
| B     | 1.6   | 3.7   | 5.1   |
| B     | 1.5   | 3.8   | 4.9   |
| B     | 1.5   | 4.2   | 5.5   |



## 一、适用条件

两因素重复测量资料方差分析，需要满足6个条件：

条件1：观察变量唯一，且为连续变量。

条件2：有两个分析因素。

条件3：观察变量为重复测量数据，即不满足独立性。

条件4：观察变量不存在显著的异常值。

条件5：各组、各水平(时间点)观察变量为正态(或近似正态)分布。

条件6：相互比较的各处理水平(组别)的总体方差相等，即方差齐性。

## 二、统计量计算

两因素重复测量资料的方差分析需考虑两个因素的影响，一个因素是处理分组，可通过施加干预和随机分组来实现；另一个因素是测量时间，由研究者根据专业知识和要求确定。因此，两因素重复测量资料的变异可分解为处理因素、时间因素、处理因素和时间因素的交互作用、受试对象间的随机误差和重复测量的随机误差。

| 变异来源   | $SS$   | $\nu$                             | $MS$   | $F$                                 |
|--------|--|-----------------------------------|--|-------------------------------------|
| 总变异    | $\sum X^2 - C$   | $N - 1$                           |  |                                     |
| 处理     | $\sum (\sum X_i)^2 / pn - C$                               | $g - 1$                           | $\frac{SS_{处理}}{\nu_{处理}}$                     | $\frac{MS_{处理}}{MS_{个体}}$           |
| 时间     | $\sum (\sum X_j)^2 / gn - C$                               | $p - 1$                           | $\frac{SS_{时间}}{\nu_{时间}}$                     | $\frac{MS_{时间}}{MS_{重复}}$           |
| 处理×时间  | $\sum_i (\sum_j X_{ij})^2 / n - C - SS_{处理} - SS_{时间}$     | $(g - 1)(p - 1)$                  | $\frac{SS_{处理 \times 时间}}{\nu_{处理 \times 时间}}$ | $\frac{MS_{处理 \times 时间}}{MS_{重复}}$ |
| 个体误差   | $\sum_i \sum_k (\sum X_{ik})^2 / p - C - SS_{处理}$          | $g \times (n - 1)$                | $\frac{SS_{个体}}{\nu_{个体}}$                     |                                     |
| 重复测量误差 | $SS_{总} - SS_{处理} - SS_{时间} - SS_{处理 \times 时间} - SS_{个体}$ | $g \times (p - 1) \times (n - 1)$ | $\frac{SS_{重复}}{\nu_{重复}}$                     |                                     |

其中： $i (i=1, 2, \dots, g)$ 为处理因素的水平数， $j (j=1, 2, \dots, p)$ 为测量时间点数， $k (k=1, 2, \dots, n)$ 为每个处理组内受试对象数。 $N$ 为全部受试对象总数， $N=p \times n \times k$ 。 $C=(\sum X)^2 / N$ 。



(四) 交互作用判断

1. 建立检验假设，确定检验水准

$H_0$ : 不同饲料的增重效果和时间因素之间的交互作用等于0，即干预因素和时间因素不存在交互作用

$H_1$ : 不同饲料的增重效果和时间因素之间的交互作用不等于0，即干预因素和时间因素存在交互作用

$\alpha = 0.05$

2. 计算检验统计量

由图1 公式和图2 数据可得计算结果如图6所示。

| 变异来源   | <i>SS</i> | <i>v</i> | <i>MS</i> | <i>F</i>  |
|--------|-----------|----------|-----------|-----------|
| 总变异    | 196.890   | 59       | 3.337     |           |
| 处理     | 1.667E-04 | 1        | 1.667E-04 | 1.524E-04 |
| 时间     | 170.161   | 2        | 85.081    | 437.308   |
| 处理×时间  | 0.041     | 2        | 0.021     | 0.106     |
| 个体误差   | 19.683    | 18       | 1.094     |           |
| 重复测量误差 | 7.004     | 36       | 0.195     |           |

图6 两因素重复测量方差分析表

3. 确定*P*值，作出推断结论

根据图6中的统计量查阅(*F*界值表)，处理因素与时间的交互项 $F_{\text{处理} \times \text{时间}} = 0.106 < F_{0.05(2, 36)} = 3.26$ ， $P > 0.05$ ，可认为两种饲料对家兔的增重效果在不同时间点上的变化没有统计学意义。因此，可直接分析主效应。



## (五) 两因素重复测量方差分析(主效应分析)

### 1. 建立检验假设，确定检验水准

处理因素

$H_0$ : 不同饲料对家兔增重效果有影响

$H_1$ : 不同饲料对家兔增重效果没有影响

时间因素

$H_0$ : 不同时间对家兔增重效果有影响

$H_1$ : 不同时间对家兔增重效果没有影响

$\alpha = 0.05$

### 2. 计算检验统计量

尽管处理因素与时间的交互项无统计学意义，但计算主效应时，在模型中仍需要保留交互项。因此计算结果和图6一致。

### 3. 确定 $P$ 值，作出推断结论

处理因素 $F_{\text{处理}} = 1.524\text{E}-04 < F_{0.05(1, 18)} = 4.41$ ,  $P > 0.05$ , 表示两种饲料对家兔的增重效果差异无统计学意义。时间因素 $F_{\text{时间}} = 437.308 > F_{0.05(2, 36)} = 3.26$ ,  $P < 0.05$ , 表示时间对家兔的增重效果差异有统计学意义。





# 课程内容安排

| 上课日期 | 章节 | 教学内容            | 教学要点  | 作业           | 随堂测          | 学时 |
|------|----|-----------------|---|--------------|--------------|----|
| 2.16 | 1  | 数据可视化, 描述性统计    | 1. 课程介绍 & 数据类型                                      | 作业1<br>(8%)  | 测试1<br>(8%)  | 2  |
| 2.23 |    |                 | 2. 描述性统计Descriptive Statistics & 数据常用可视化            |              |              | 2  |
| 3.2  |    |                 | 3. 大数定理 & 中心极限定理                                    |              |              | 2  |
| 3.9  |    |                 | 4. 常用概率分布   |              |              | 2  |
| 3.16 | 2  | 推断性统计, 均值差异检验   | 5. 统计推断基础-1: 置信区间 Confidence Interval *             | 作业2<br>(10%) | 测试2<br>(10%) | 2  |
| 3.23 |    |                 | 6. 统计推断基础-2: 假设检验, I及II类错误, 统计量, p-值                |              |              | 2  |
| 3.30 |    |                 | 7. 数值数据的均值比较-1: 单样本及双样本t-检验, 效应量, 功效                |              |              | 2  |
| 4.6  |    |                 | 8. 数值数据的均值比较-2: One-Way ANOVA, 正态性检验                |              |              | 2  |
| 4.13 |    |                 | 9. 数值数据的均值比较-3: Two-Way ANOVA                       |              |              | 2  |
| 4.20 | 3  | 比例差异检验          | 10. 样本和置信区间预估 *                                     | 作业3<br>(6%)  | 测试3<br>(6%)  | 2  |
| 4.27 |    |                 | 11. 类别数据的比例比较-1: 单样本比例推断                            |              |              | 2  |
| 5.7  |    |                 | 12. 类别数据的比例比较-2: 联立表的卡方检验                           |              |              | 2  |
| 5.11 | 4  | 协方差, 相关分析, 回归分析 | 13. 相关分析 (Pearson r, Spearman rho, Kendal' s tau) * | 作业4<br>(6%)  | 测试4<br>(6%)  | 2  |
| 5.18 |    |                 | 14. 简单回归分析  |              |              | 2  |
| 5.25 |    |                 | 15. 多元回归Multiple Regression                         |              |              | 2  |
| 6.1  | 5  | Course Summary  | 16. 课程总结 *  |              |              | 2  |
|      |    |                 | Total   | 30%          | 30%          | 32 |

\* 随堂测试

# 第2章总结（推论性统计 – 均值估算 & 均值差异检验）

- **置信区间**

- 区间估算
- 置信水平，样本量，两种关系
- Z-dist, Student-t, Welch-t

- **假设检验**

- 一类，二类错误
- 统计量，p-值，效应量，功效，几者间的关系
- 单样本检验，双样本检验
- 使用条件
- 正态性检验

- **ANOVA**

- 单样本，双样本方差分析
- 正态性检验，方差齐性检验
- 多重比较问题，后分析（两两样本间的多重比较）
- 主效应，交互作用，重复测量
- 使用条件



谢谢，下周见！



让开，  
我要去学习了

