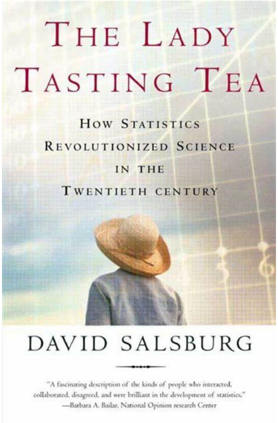
生物统计学

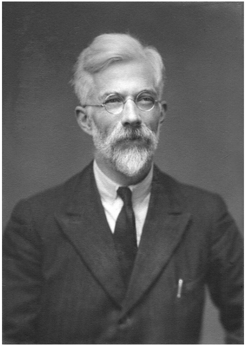
第六章 假设检验的理论基础

云南大学 生命科学学院

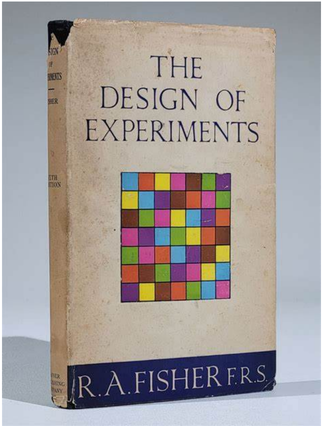
会泽百家 至公天下

统计学家 David Salsburg 所著的《女士品茶：统计学如何变革了科学和生活》一书，开篇讲述了这么一个故事…





Ronald Fisher (1890-1962)



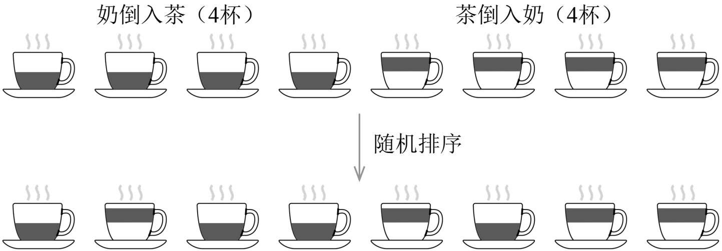


图 **6.1** 品茶试验

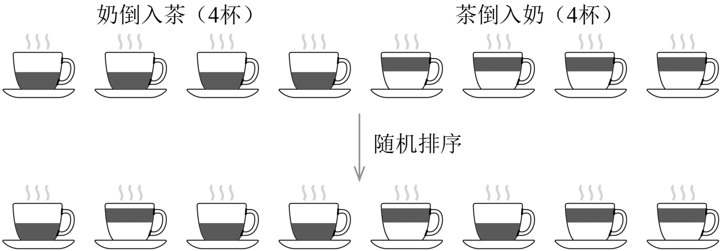


图 **6.1** 品茶试验

Fisher 设计的试验要求 8 杯茶的测试顺序完全随机!

制茶顺序一共有多少种可能?

制茶顺序一共有多少种可能?

$$\mathrm{C}_n^m = \frac{n!}{m!(n-m)!} = \frac{8!}{4!(8-4)!} = 70$$

制茶顺序一共有多少种可能?

$$\mathrm{C}_n^m = \frac{n!}{m!(n-m)!} = \frac{8!}{4!(8-4)!} = 70$$

假如女士没有品鉴能力，靠猜测完全分辨正确的概率?

制茶顺序一共有多少种可能？

$$\mathrm{C}_n^m = \frac{n!}{m!(n-m)!} = \frac{8!}{4!(8-4)!} = 70$$

假如女士没有品鉴能力，靠猜测完全分辨正确的概率？

$$\frac{1}{70} \approx 0.014$$

概率论中的小概率原理，认为一个事件发生的概率如果很小，那么它在一次试验中几乎不可能发生（但在多次重复试验中又几乎必然发生）。

概率论中的小概率原理，认为一个事件发生的概率如果很小，那么它在一次试验中几乎不可能发生（但在多次重复试验中又几乎必然发生）。

Fisher 用5%的标准来要求显著性！

概率论中的小概率原理，认为一个事件发生的概率如果很小，那么它在一次试验中几乎不可能发生（但在多次重复试验中又几乎必然发生）。

Fisher 用5%的标准来要求显著性！

$$\frac{1}{70} < 0.05$$

概率论中的小概率原理，认为一个事件发生的概率如果很小，那么它在一次试验中几乎不可能发生（但在多次重复试验中又几乎必然发生）。

Fisher 用5%的标准来要求显著性！

$$\frac{1}{70} < 0.05$$

按照 Fisher 的思想，“女士完全分辨正确 **8** 杯茶”是统计上显著的！

假如 Fisher 的试验设置了 6 杯茶，两种制茶方式各 3 杯，结果会如何?

假如 Fisher 的试验设置了 6 杯茶，两种制茶方式各 3 杯，结果会如何？

假如女士没有品鉴能力，靠猜测完全分辨正确的概率为

$$\frac{1}{\mathrm{C}_6^3} = \frac{1}{20} = 0.05$$

即使 6 杯全对，统计上的显著性也达不到小于**0.05**的水平！

定义 (3.4)

在一个有 n 个等可能结果的随机试验中，事件 A 包含其中 m 个结果，则事件 A 的发生概率可定义为

$$P(A) = \frac{m}{n} \tag{3.3}$$

表 **6.1** 女士品茶试验的概率计算

试验可能的结果	排序方式的数量	概率	累积概率
8 杯全对	1	$\frac{1}{70}$ (0.014)	0.014
6 杯对、2 杯错	16	$\frac{16}{70}$ (0.229)	0.243
4 杯对、4 杯错	36	$\frac{36}{70}$ (0.514)	0.757
2 杯对、6 杯错	16	$\frac{16}{70}$ (0.229)	0.986
8 杯全错	1	$\frac{1}{70}$ (0.014)	1.000

总结起来，Fisher 的显著性检验思想可归纳成以下几点:

总结起来，Fisher 的显著性检验思想可归纳成以下几点:

1. 有一个明确的零假设;

总结起来，Fisher 的显著性检验思想可归纳成以下几点：

1. 有一个明确的零假设；
2. 设计一组试验，观察随机变量 x，且当零假设成立时，x 有已知的概率分布；

总结起来，Fisher 的显著性检验思想可归纳成以下几点：

1. 有一个明确的零假设；
2. 设计一组试验，观察随机变量 x，且当零假设成立时，x 有已知的概率分布；
3. 将 x 的取值根据对零假设的不利程度排序；

总结起来，Fisher 的显著性检验思想可归纳成以下几点：

1. 有一个明确的零假设；
2. 设计一组试验，观察随机变量 x，且当零假设成立时，x 有已知的概率分布；
3. 将 x 的取值根据对零假设的不利程度排序；
4. 根据试验的当前观测值 x_c，计算 x_c 和比 x_c 更不利于零假设的可能取值的概率，并得到和事件概率 $P(x \geq x_c|H_0)$ ($x > x_c$ 表示比 x_c 更不利于零假设 H_0 的值)；

总结起来，Fisher 的显著性检验思想可归纳成以下几点:

1. 有一个明确的零假设;
2. 设计一组试验，观察随机变量 x，且当零假设成立时，x 有已知的概率分布;
3. 将 x 的取值根据对零假设的不利程度排序;
4. 根据试验的当前观测值 x_c，计算 x_c 和比 x_c 更不利于零假设的可能取值的概率，并得到和事件概率 $P(x \geq x_c|H_0)$ ($x > x_c$ 表示比 x_c 更不利于零假设 H_0 的值);
5. 选择一个显著性水平 α，当 $P(x \geq x_c|H_0) < \alpha$ 时拒绝零假设，反之，接受零假设。

定义 (6.1)

样本所来自的总体与已知总体之间，或两个样本所来自的总体之间，存在零差异的假设，称为零假设 (null hypothesis)，又称无效假设，记作 H_0。

定义 (6.2)

与零假设对立的假设，称为备择假设 (alternative hypothesis)，记作 H_1。

例 (6.1)

新生儿出生体重可能与地区经济状况有关，为检验这一假说，在经济欠发达地区某医院产科搜集足期分娩的 100 例健康婴儿的出生体重，平均值 3.37 kg。假设基于普查的结果显示全国新生儿出生体重平均为 3.4 kg，标准差 0.68 kg。进行假设检验该如何设定零假设和备择假设?

例 (6.1)

新生儿出生体重可能与地区经济状况有关，为检验这一假说，在经济欠发达地区某医院产科搜集足期分娩的 100 例健康婴儿的出生体重，平均值 3.37 kg。假设基于普查的结果显示全国新生儿出生体重平均为 3.4 kg，标准差 0.68 kg。进行假设检验该如何设定零假设和备择假设?

解

零假设 $H_0 : \mu = 3.4$；备择假设 $H_1 : \mu < 3.4$。

例 (6.1)

新生儿出生体重可能与地区经济状况有关，为检验这一假说，在经济欠发达地区某医院产科搜集足期分娩的 100 例健康婴儿的出生体重，平均值 3.37 kg。假设基于普查的结果显示全国新生儿出生体重平均为 3.4 kg，标准差 0.68 kg。进行假设检验该如何设定零假设和备择假设?

解

零假设 $H_0 : \mu = 3.4$；备择假设 $H_1 : \mu < 3.4$。

零假设 H_0：无分辨能力；备择假设 H_1：有分辨能力。

定义 (6.3)

利用样本构造的用于检验零假设是否成立，且具有已知概率分布的统计量，称为检验统计量 (test statistic)。

定义 (6.4)

检验统计量的一个阈值，沿着不利于零假设的方向检验统计量超过该阈值时拒绝零假设，则该阈值称为检验临界值 (critical value)。

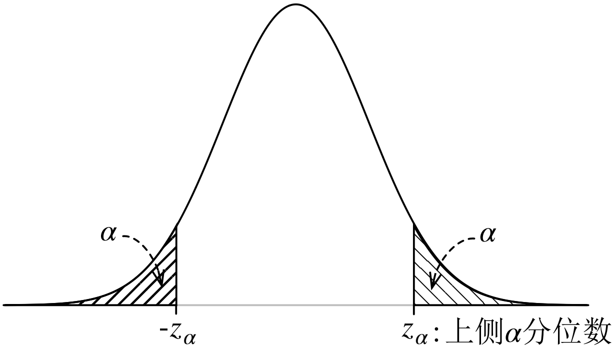


图 **6.2** 正态分布上侧与下侧分位数

例 (6.2)

...100 例健康婴儿的出生体重，平均值 3.37 kg。假设基于普查的结果显示全国新生儿出生体重平均为 3.4 kg，标准差 0.68 kg。如何通过检验临界值来作出统计推断?

例 (6.2)

...100 例健康婴儿的出生体重，平均值 3.37 kg。假设基于普查的结果显示全国新生儿出生体重平均为 3.4 kg，标准差 0.68 kg。如何通过检验临界值来作出统计推断?

$$z_c = \frac{3.37 - 3.4}{0.68/\sqrt{100}} \approx -0.441$$

$$z_{0.95} = -1.644854$$

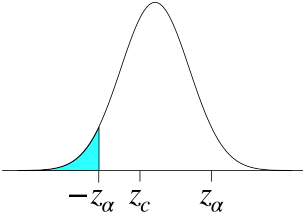
```
> qnorm(0.95, mean = 0, sd = 1, lower.tail = FALSE)
```

例 (6.2)

...100 例健康婴儿的出生体重，平均值 3.37kg。假设基于普查的结果显示全国新生儿出生体重平均为 3.4kg，标准差 0.68kg。如何通过检验临界值来作出统计推断?

解

因为 z_c 大于 $z_{0.95} \approx -1.645$，处于检验临界值 $z_{0.95}$ 的右侧，而不利于零假设的方向在检验统计量取值域的左侧，所以检验结论为：接受零假设。



定义 (6.5)

当零假设成立时，得到样本观测值以及比样本观测值更不利于零假设的值的概率，称为相伴概率 (companion probability)，记作 P，通常直接称为P值。

定义 (6.6)

假设检验中事先确定的一个作为判断界限的小概率标准，称为显著性水平，记作 α。

例 (6.3)

...100 例健康婴儿的出生体重，平均值 3.37 kg。假设基于普查的结果显示全国新生儿出生体重平均为 3.4 kg，标准差 0.68 kg。选择显著性水平 $\alpha = 0.05$，该如何通过相伴概率来作出统计推断?

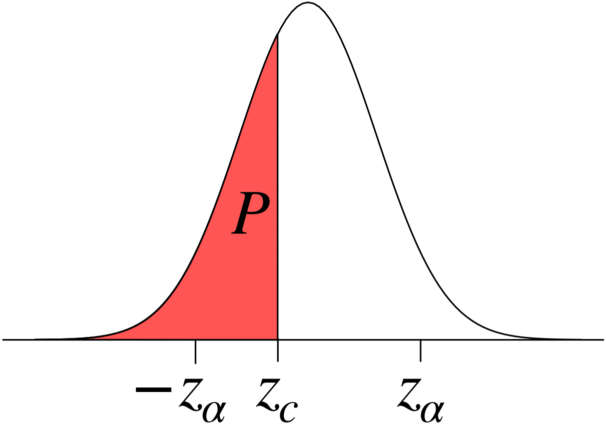
例 (6.3)

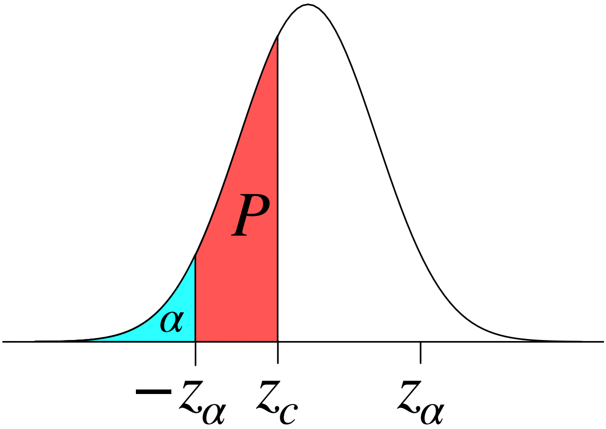
...100 例健康婴儿的出生体重，平均值 3.37 kg。假设基于普查的结果显示全国新生儿出生体重平均为 3.4 kg，标准差 0.68 kg。选择显著性水平 $\alpha = 0.05$，该如何通过相伴概率来作出统计推断?

解

$$P(z \leq z_c|H_0) = P(z \leq -0.441|H_0) \approx 0.330$$

```
> pnorm(q = -0.441, lower.tail = TRUE)
[1] 0.3296065
```


P
$-z_\alpha$
z_c
z_α


P
α
−zα
zc
zα

定义 (6.7)

检验统计量的取值范围中，对应拒绝零假设结论的区域，称为拒绝域 (rejection region)。

定义 (6.8)

检验统计量的取值范围中，对应接受零假设结论的区域，称为接受域 (acceptance region)。

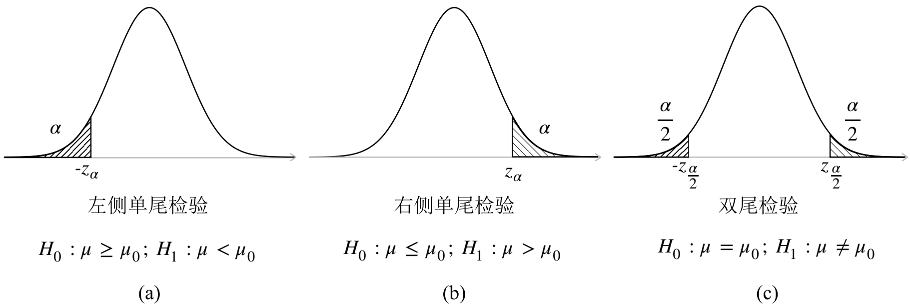


图 **6.3** 单尾检验与双尾检验

定义 (6.9)

拒绝域出现在检验统计量抽样分布单侧尾部的检验方法，称为单尾检验(one-tailed test)，或单侧检验 (one-sided test)。

定义 (6.10)

拒绝域出现在检验统计量抽样分布两侧尾部的检验方法，称为双尾检验(two-tailed test)，或双侧检验 (two-sided test)。

利用样本统计量的抽样分布，计算在零假设成立时，

利用样本统计量的抽样分布，计算在零假设成立时，

得到样本观测值以及比样本观测值更不利于零假设的值的概率——相伴概率，即 P 值（可理解为样本观测值的极端程度）。

利用样本统计量的抽样分布，计算在零假设成立时，

得到样本观测值以及比样本观测值更不利于零假设的值的概率——相伴概率，即P 值（可理解为样本观测值的极端程度）。

然后，根据小概率原理做出对零假设接受或拒绝的判断。

假设检验的一般操作流程可归纳如下:

1. 根据数据的具体情况，设定零假设和备择假设;
2. 选择显著性水平;
3. 计算检验统计量和相伴概率;
4. 根据检验统计量（与检验临界值比较）或相伴概率（与显著性水平比较）的情况，做出推断，得出检验结论。

	H_0 成立	H_0 不成立
拒绝 H_0	第一类错误	正确判断
接受 H_0	正确判断	第二类错误

图 **6.4** 两类错误

定义 (6.11)

当零假设成立时，检验方法拒绝零假设，此类判断错误称为第一类错误 (type I error)，又称 I 型错误、α 错误、“弃真”错误。

定义 (6.12)

当零假设不成立时，检验方法接受零假设，此类判断错误称为第二类错误 (type II error)、又称 II 型错误、β 错误、“纳伪”错误。

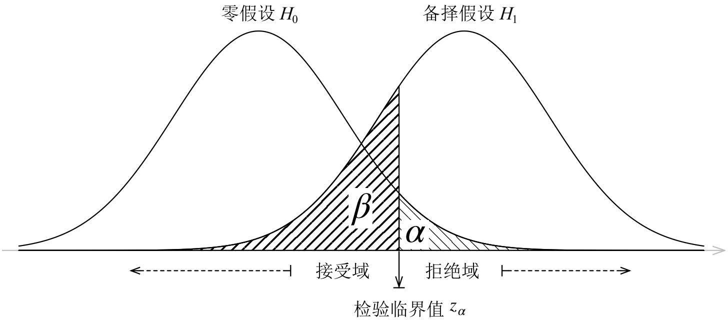


图 **6.5** 接受域与拒绝域

将假设检验视为根据样本信息做出的决策，结果应包含四种可能性：

- 零假设成立时，接受零假设；
- 零假设成立时，拒绝零假设（第一类错误，弃真概率 α）；
- 零假设不成立时，接受零假设（第二类错误，纳伪概率 β）；
- 零假设不成立时，拒绝零假设。

假定对某总体平均数 μ_0 进行区间估计，置信区间可表示为

$$P(\overline{x} - z_{\frac{\alpha}{2}}\sigma_{\overline{x}} \le \mu_0 \le \overline{x} + z_{\frac{\alpha}{2}}\sigma_{\overline{x}}) = 1 - \alpha \tag{6.31}$$

假定对某总体平均数 μ_0 进行区间估计，置信区间可表示为

$$P(\overline{x} - z_{\frac{\alpha}{2}}\sigma_{\overline{x}} \leq \mu_0 \leq \overline{x} + z_{\frac{\alpha}{2}}\sigma_{\overline{x}}) = 1 - \alpha \tag{6.31}$$

将连续不等式重新整理成包含检验统计量的形式，有

$$P(-z_{\frac{\alpha}{2}} \leq \frac{\overline{x} - \mu_0}{\sigma_{\overline{x}}} \leq z_{\frac{\alpha}{2}}) = 1 - \alpha \tag{6.32}$$

假设检验与区间估计是统计推断的两种方式，形式上二者可以相互转换。

置信区间对应接受域，置信度 $1-\alpha$ 对应显著性水平 α。

6.5 关于假设检验的几点说明

本章小结

1. 假设检验的基本原理
   - 女士品茶
   - Fisher 的显著性检验思想
   - Neyman-Pearson 假设检验理论
2. 假设检验的两类错误
   - 两类错误的定义
   - 两类错误的控制
3. 假设检验的功效
   - 检验的功效
4. 假设检验和区间估计的对偶关系
5. 关于假设检验的几点说明