

MAT220 第17讲：独立总体的检验

目录

1	配对数据回顾	2
2	独立总体	3
2.1	改变我们的假设	3
2.2	样本均值之差	3
2.3	构造 $\bar{X} - \bar{Y}$ 的抽样分布	4
3	独立均值的假设检验	5
3.1	归无假设与备择假设	5
3.2	σ 已知和未知时的检验统计量	6
4	检验两个二项概率	7
4.1	单样本比例回顾	7
4.2	两个独立样本比例之差	8
4.3	成功概率之差的假设检验	8
5	例：橙色彩虹糖	9
6	假设检验总结	11

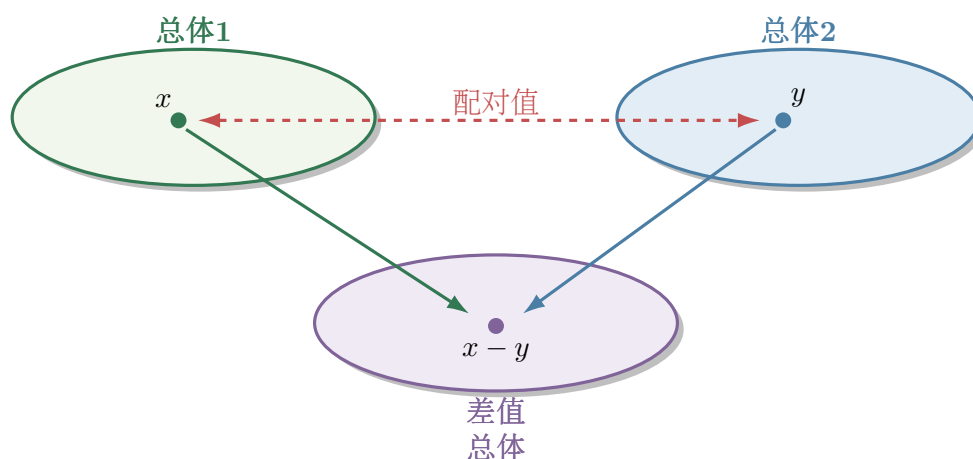
在第16讲中，我们比较了观测值自然配对的两个总体。通过这种配对，我们可以在每一对观测值内部作差，构造一个新的差值总体 $D = X - Y$ ，并将问题化为单样本 t 检验。在本讲中，我们将去掉这一假设，改为假设两个总体是独立的。正如我们将看到的，这一新假设仍然使我们能够同时对两个总体进行假设检验。不过，这一次我们不能再构造差值 $d = x - y$ 。因此，我们需要用另一种方法来弥补这种便利的丧失。最后，我们将使用这些新方法检验总体均值之差，并检验两个独立二项试验的成功概率之差。

1 配对数据回顾

在本讲中，我们将继续讨论比较两个不同总体参数的假设检验。给定两个总体，这里的核心问题是：

这两个总体平均而言是否相同？

原则上，我们可以比较两个总体的真实均值，并检验它们是否相等，从而回答这个问题。在进行这种检验的过程中，我们需要构造一个结合两个总体样本数据的抽样分布。在第16讲中，我们指出，一般来说，这是一项困难的工作。因此，我们通过假设两个总体自然配对来简化问题。由于一个总体中的每个观测值在另一个总体中都有唯一对应的观测值，这一假设使我们能够构造第三个总体，其中包含所有配对值之间的差：



在符号上，我们使用字母 D 表示这个“差值总体”。形式上， D 就是 $X - Y$ ，即用于定义两个总体的随机变量之差。

利用关于期望值和方差的一些标准结论，我们得到随机变量 D 的均值和方差为：

$$\mu_D = \mu_X - \mu_Y, \quad \sigma_D^2 = \sigma_X^2 + \sigma_Y^2 - 2 \text{Cov}(X, Y).$$

虽然均值具有简单的形式，但方差中出现了一个称为协方差的额外项。粗略地说，这个协方差项衡量配对值如何共同变化。由于 σ_D 通常未知，我们必须使用 s_D 估计差值的标准差，并使用学生 t 分布，从而得到以下检验统计量。

配对数据的检验统计量

对于配对数据，首先计算差值 $d_i = x_i - y_i$ 。然后使用

$$t = \frac{\bar{d} - 0}{s_D / \sqrt{n}}, \quad df = n - 1.$$

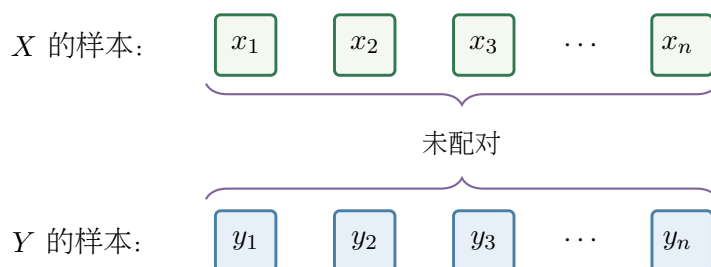
2 独立总体

2.1 改变我们的假设

在进行涉及两个总体的假设检验时，描述适当的抽样分布存在一个主要的技术障碍。为了简化问题，在上一讲中，我们假设两个总体自然配对。这使我们能够从总体 D 中抽取样本，并将其拆分为相应的 X 和 Y 样本，从而描述 $\bar{D}$ 的抽样分布。

在本讲中，我们不再假设总体是配对的。相反，我们将假设 X 和 Y 是独立的，这意味着二者之间不存在概率依赖关系。事实上，我们在本课程前面讨论概率时见过这个术语：如果一个事件的发生不会改变另一个事件的概率，则称这两个事件“独立”。类似地，如果知道一个随机变量的值不会改变另一个随机变量的概率分布，我们就称 X 和 Y 独立。在数学上，假设 X 和 Y 独立意味着二者之间没有相互作用，因此协方差 $\text{Cov}(X, Y)$ 为 0。

尽管这看起来很有希望，但我们也必须承认，在放宽配对数据假设的过程中，我们失去了一些东西。当 X 和 Y 的观测值配对时，我们能够构造一个由 $D = X - Y$ 控制的第三个总体，其中数据点具有 $d = x - y$ 的形式。这使我们能够从总体 D 中抽取样本，并将其拆分为 X 和 Y 的单独样本。如果去掉 X 和 Y 配对的假设，我们就不能再使用这种方法。换句话说，即使我们从 X 和 Y 中选取两个大小相同的简单随机样本，也没有自然的方法将它们配对成某种差值样本：



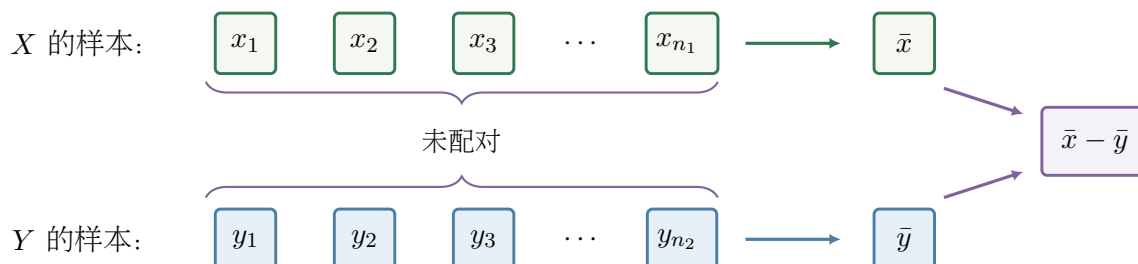
这意味着，没有自然的方法构造个体差值样本来检验差 $\mu_X - \mu_Y$ 。

2.2 样本均值之差

假设我们分别从 X 和 Y 中抽取两个独立随机样本。由于样本是独立的，因此没有必要要求两个样本具有相同的大小。令 n_1 表示从 X 中抽取的样本大小，令 n_2 表示从 Y 中抽取的样本大小。则样本均值 $\bar{x}$ 和 $\bar{y}$ 可以分别计算为：

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_{n_1}}{n_1} \quad \text{以及} \quad \bar{y} = \frac{y_1 + y_2 + \dots + y_{n_2}}{n_2}.$$

我们的最终目标是在不能直接使用差值总体 D 的情况下，描述一个以 $\mu_X - \mu_Y$ 为中心的抽样分布。根据可用的样本数据，我们可以考虑 $\bar{x} - \bar{y}$ ，从而构造均值之差：



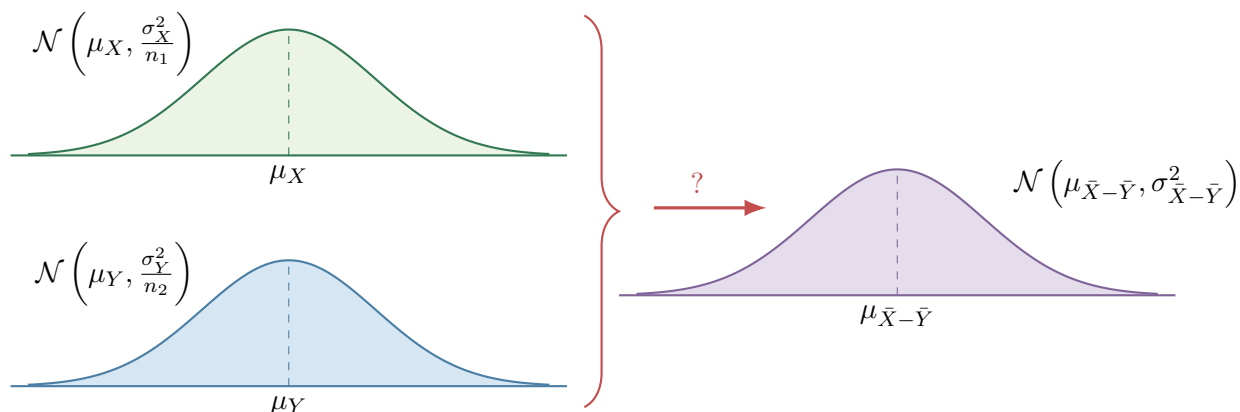
如果我们希望通过评估这个特定观测值 $\bar{x} - \bar{y}$ 出现的可能性来进行假设检验，那么就需要考虑一个包含所有可能随机获得的差 $\bar{x} - \bar{y}$ 的概率分布。因此，我们将研究随机变量 $\bar{X} - \bar{Y}$ 。

当样本足够大时，中心极限定理告诉我们， $\bar{X}$ 和 $\bar{Y}$ 的各自抽样分布近似为正态分布，其参数为：

$$\bar{X} \approx \mathcal{N}\left(\mu_X, \frac{\sigma_X^2}{n_1}\right) \quad \text{以及} \quad \bar{Y} \approx \mathcal{N}\left(\mu_Y, \frac{\sigma_Y^2}{n_2}\right).$$

我们强调，与配对数据的情况不同，两个样本不必具有相同的大小，即不要求 $n_1 = n_2$ 。

现在，我们将使用这两个正态分布来描述随机变量 $\bar{X} - \bar{Y}$ 的整体分布。示意性地，我们要寻找的是：



2.3 构造 $\bar{X} - \bar{Y}$ 的抽样分布

在上一讲中，我们看到，两个随机变量之差的期望值等于各个随机变量期望值之差。对于 $\bar{X} - \bar{Y}$ ，用符号表示为：

$$E(\bar{X} - \bar{Y}) = E(\bar{X}) - E(\bar{Y}).$$

由于 $E(\bar{X})$ 就是 $\bar{X}$ 的均值，即 μ_X ，而 $\bar{Y}$ 也同样如此，因此我们可以得出：

$$\mu_{\bar{X} - \bar{Y}} = E(\bar{X} - \bar{Y}) = E(\bar{X}) - E(\bar{Y}) = \mu_X - \mu_Y.$$

方便的是，这与配对数据中出现的中心相同：两个样本均值之差的期望值等于两个总体均值之差。

对于方差，我们实际上仍然可以使用第16讲中的同一结论：

$$\text{Var}(\bar{X} - \bar{Y}) = \text{Var}(\bar{X}) + \text{Var}(\bar{Y}) - 2 \text{Cov}(\bar{X}, \bar{Y}),$$

其中 $\text{Cov}(\bar{X}, \bar{Y})$ 是 $\bar{X}$ 和 $\bar{Y}$ 的协方差。这里发生了关键的简化：由于假设两个样本独立，所以 $\bar{X}$ 和 $\bar{Y}$ 也独立，因此这个额外的协方差项为 0：

$$\text{Cov}(\bar{X}, \bar{Y}) = 0.$$

因此，上面的方差公式可以方便地简化为：

$$\text{Var}(\bar{X} - \bar{Y}) = \text{Var}(\bar{X}) + \text{Var}(\bar{Y}).$$

$\bar{X}$ 的方差就是其正态分布中的第二个参数，即 $\bar{X}$ 的标准误差项的平方，也就是 $\frac{\sigma_X^2}{n_1}$ 。类似地， $\bar{Y}$ 的方差为 $\frac{\sigma_Y^2}{n_2}$ 。因此，对于独立总体， $\bar{X} - \bar{Y}$ 分布的方差为：

$$\sigma_{\bar{X}-\bar{Y}}^2 = \text{Var}(\bar{X} - \bar{Y}) = \text{Var}(\bar{X}) + \text{Var}(\bar{Y}) = \frac{\sigma_X^2}{n_1} + \frac{\sigma_Y^2}{n_2}.$$

综合以上结果，我们现在可以完整描述近似表示 $\bar{X} - \bar{Y}$ 的正态分布。

独立均值之差的抽样分布

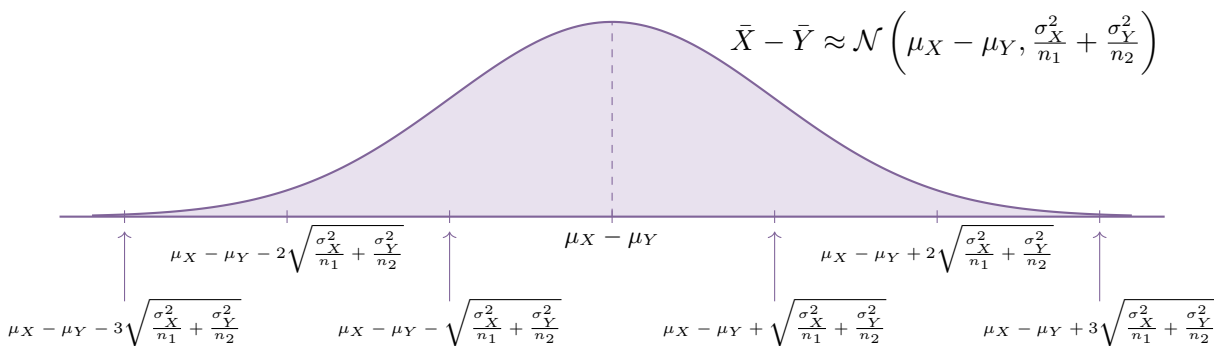
对于独立样本，

$$\bar{X} - \bar{Y} \approx \mathcal{N}\left(\mu_X - \mu_Y, \frac{\sigma_X^2}{n_1} + \frac{\sigma_Y^2}{n_2}\right).$$

我们的正态分布将以 $\mu_X - \mu_Y$ 为中心，并以标准误差为单位向两侧计数，其中标准误差是上述方差项的平方根：

$$\sqrt{\frac{\sigma_X^2}{n_1} + \frac{\sigma_Y^2}{n_2}}.$$

从几何上看，我们得到下图：



虽然这看起来很复杂，但重要的是，它完全是可处理的——我们可以使用一个公式来计算相关概率。

3 独立均值的假设检验

3.1 归无假设与备择假设

与上一讲中的讨论一样，当我们想要检验两个独立总体平均而言是否相等时，归无假设为

$$H_0 : \mu_X - \mu_Y = 0.$$

然后，我们从 X 和 Y 中抽取两个可能具有不同大小的独立样本。在归无假设成立时，抽样分布变为

$$\bar{X} - \bar{Y} \approx \mathcal{N}\left(0, \frac{\sigma_X^2}{n_1} + \frac{\sigma_Y^2}{n_2}\right).$$

来自 X 和 Y 的两个样本分别具有样本均值 $\bar{x}$ 和 $\bar{y}$ ，它们的差 $\bar{x} - \bar{y}$ 将落在上述正态分布上的某个位置。与我们之前见过的其他假设检验一样，观测值 $\bar{x} - \bar{y}$ 离中心越远，P值就越小。

与第16讲中的配对总体一样，如果我们希望检验 H_0 是否与某个相反的主张相冲突，就可以根据具体怀疑使用不同的备择假设：

$$H_1 : \begin{cases} \mu_X - \mu_Y < 0, & \text{表示 } X \text{ 总体的平均值较小,} \\ \mu_X - \mu_Y > 0, & \text{表示 } X \text{ 总体的平均值较大,} \\ \mu_X - \mu_Y \neq 0, & \text{仅表示两个平均值不同.} \end{cases}$$

选择方向

相减的顺序必须与备择假设一致。如果使用样本差 $\bar{x} - \bar{y}$ ，那么：

- $\mu_X < \mu_Y$ 写为 $H_1 : \mu_X - \mu_Y < 0$;
- $\mu_X > \mu_Y$ 写为 $H_1 : \mu_X - \mu_Y > 0$ 。

3.2 σ 已知和未知时的检验统计量

假设两个总体标准差 σ_X 和 σ_Y 都已知。在这种情况下， $\bar{X} - \bar{Y}$ 分布的标准误差为

$$\sqrt{\frac{\sigma_X^2}{n_1} + \frac{\sigma_Y^2}{n_2}}.$$

当我们将观测到的差转换为 z 分数时，示意性地取以下商：

$$z = \frac{\text{观测值} - \text{均值}}{\text{标准误差}}.$$

在本例中，我们假设归无假设成立，即 $\mu_X - \mu_Y = 0$ 。标准误差已在上面给出，并且可以由现有信息直接计算。因此，观测到的差 $\bar{x} - \bar{y}$ 通过以下公式转换为 z 分数：

$$z = \frac{(\bar{x} - \bar{y}) - 0}{\sqrt{\frac{\sigma_X^2}{n_1} + \frac{\sigma_Y^2}{n_2}}} = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_X^2}{n_1} + \frac{\sigma_Y^2}{n_2}}}.$$

标准差已知时的独立均值

当 σ_X 和 σ_Y 已知时，为检验 $H_0 : \mu_X - \mu_Y = 0$ ，使用公式：

$$z = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_X^2}{n_1} + \frac{\sigma_Y^2}{n_2}}}.$$

现在假设 σ_X 或 σ_Y 中至少有一个未知。在这种情况下，我们没有可以代入前一个公式的数值。因此，次优的方法是用各自的样本估计值替换两个总体标准差，并改用近似的学生 t 分布：

$$z = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_X^2}{n_1} + \frac{\sigma_Y^2}{n_2}}} \quad \xrightarrow{\text{将 } \sigma_X \text{ 和 } \sigma_Y \text{ 替换为 } s_X \text{ 和 } s_Y} \quad t = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_X^2}{n_1} + \frac{s_Y^2}{n_2}}}.$$

σ_X 或 σ_Y 未知 学生 t 分布,
 $df = \min(n_1 - 1, n_2 - 1)$

在本课程中，按照约定，我们使用两个可用自由度中较小的一个。

标准差未知时的独立均值

当任一总体标准差未知时，为检验 $H_0: \mu_X - \mu_Y = 0$ ，使用公式

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_X^2}{n_1} + \frac{s_Y^2}{n_2}}},$$

其中

$$df = \min(n_1 - 1, n_2 - 1).$$

之后，可以使用相应的 z 表或 t 表计算适当的P值。

4 检验两个二项概率

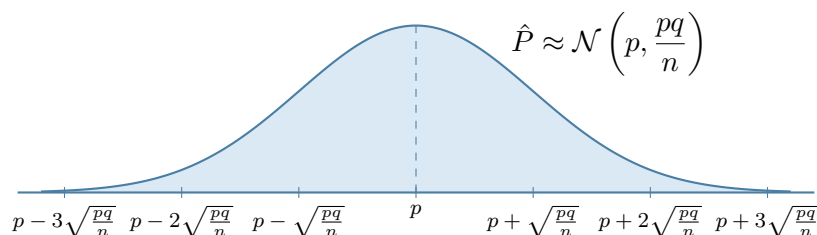
4.1 单样本比例回顾

假设我们有一个成功概率为 p 的单个二项试验。如果进行 n 次试验并观察到 r 次成功，那么可以根据样本数据估计 p 。这个估计值称为样本比例，可以简单地计算为：

$$\hat{p} = \frac{r}{n}.$$

正如我们在第13讲和第15讲中看到的，在这种情况下，中心极限定理给出正态近似：

$$\hat{P} \approx \mathcal{N}\left(p, \frac{pq}{n}\right), \quad q = 1 - p.$$



在这种情况下，我们必须始终检查样本是否足够大，以确保正态近似合理。在本课程中，我们使用 $np > 5$ 和 $nq > 5$ 这一标准。

4.2 两个独立样本比例之差

现在假设我们有两个二项试验:

- 试验1的真实成功概率为 p_1 , 样本大小为 n_1 , 样本比例为 $\hat{p}_1$;
- 试验2的真实成功概率为 p_2 , 样本大小为 n_2 , 样本比例为 $\hat{p}_2$ 。

分别有

$$\hat{P}_1 \approx \mathcal{N}\left(p_1, \frac{p_1 q_1}{n_1}\right) \quad \text{以及} \quad \hat{P}_2 \approx \mathcal{N}\left(p_2, \frac{p_2 q_2}{n_2}\right),$$

其中 $q_1 = 1 - p_1$, $q_2 = 1 - p_2$ 。样本比例与样本均值具有相同的作用, 我们可以取它们的差, 得到一个新变量 $\hat{P}_1 - \hat{P}_2$ 。如果假设这两个试验独立, 那么协方差项再次为 0, 从而得到以下形式的正态分布:

$$\hat{P}_1 - \hat{P}_2 \approx \mathcal{N}\left(p_1 - p_2, \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}\right).$$

以此为基础, 我们现在将描述一种检验两个试验是否具有相同成功概率的方法。

4.3 成功概率之差的假设检验

为了检验两个成功概率是否相等, 我们从归无假设开始:

$$H_0 : p_1 - p_2 = 0.$$

在这一假设下, 上述正态分布将以 0 为中心:

$$\mathcal{N}\left(0, \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}\right).$$

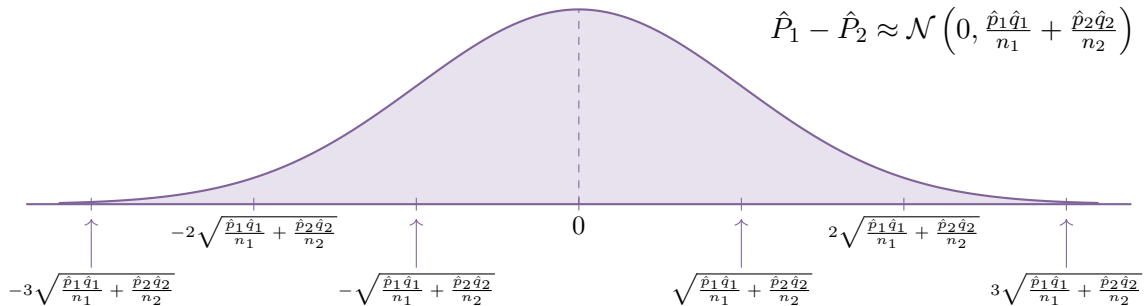
但是, 我们不能直接使用这个分布进行检验, 因为实际上并不知道 p_i 和 q_i 的值。为什么? 因为归无假设只对差 $p_1 - p_2$ 作出了陈述, 而没有给出它们的具体数值。因此, 我们必须将 p_1 和 p_2 视为未知, 并使用样本比例估计它们的值。这意味着我们的正态分布将变为:

$$\hat{P}_1 - \hat{P}_2 \approx \mathcal{N}\left(0, \frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}\right),$$

其中 $\hat{q}_1 = 1 - \hat{p}_1$, $\hat{q}_2 = 1 - \hat{p}_2$ 。从几何上看, 这仍然是一条以 0 为中心的钟形曲线, 但这一次我们按照标准误差

$$\sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$$

来计数:



相应的 z 分数以中心 0 为起点，按照上述标准误差项进行计数，因此公式为：

$$z = \frac{\text{观测值} - \text{均值}}{\text{标准误差}} = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}}.$$

这就是用于计算适当P值的公式。

检验两个独立二项概率

为了检验 $H_0 : p_1 - p_2 = 0$ ，使用

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}}.$$

P值使用 $Z \sim \mathcal{N}(0, 1)$ 计算。

与第4.1节中的正态分布一样，我们必须首先确认两个二项试验可以合理地用正态分布近似。在本课程中，在进行近似之前，必须确认以下四个量

$$n_1 \hat{p}_1, \quad n_1 \hat{q}_1, \quad n_2 \hat{p}_2, \quad n_2 \hat{q}_2$$

均大于 5。

5 例：橙色彩虹糖

假设一位教授向两个班级发放彩虹糖。他想检验班级1中橙色彩虹糖的比例是否小于班级2中橙色彩虹糖的比例。

记录的数据如下：

- 班级1共数了 740 颗彩虹糖，其中 139 颗是橙色的；
- 班级2共数了 741 颗彩虹糖，其中 150 颗是橙色的。

我们使用显著性水平 $\alpha = 0.05$ 进行检验。

步骤1：整理信息

重要的数值为

$$n_1 = 740, \quad r_1 = 139, \quad n_2 = 741, \quad r_2 = 150, \quad \alpha = 0.05.$$

令 p_1 表示班级1中橙色彩虹糖的真实比例，令 p_2 表示班级2中的真实比例。归无假设表示两个比例相等，而我们的怀疑是第一个比例更小：

$$H_0 : p_1 - p_2 = 0, \quad H_1 : p_1 - p_2 < 0.$$

这是一个左尾检验。

步骤2：计算P值

首先计算两个样本比例：

$$\hat{p}_1 = \frac{r_1}{n_1} = \frac{139}{740} \approx 0.188, \quad \hat{p}_2 = \frac{r_2}{n_2} = \frac{150}{741} \approx 0.202.$$

因此，

$$\hat{q}_1 = 1 - \hat{p}_1 \approx 0.812, \quad \hat{q}_2 = 1 - \hat{p}_2 \approx 0.798.$$

接下来检查正态近似条件：

$$n_1 \hat{p}_1 \approx 740(0.188) = 139 > 5,$$

$$n_1 \hat{q}_1 \approx 740(0.812) = 601 > 5,$$

$$n_2 \hat{p}_2 \approx 741(0.202) = 150 > 5,$$

$$n_2 \hat{q}_2 \approx 741(0.798) = 591 > 5.$$

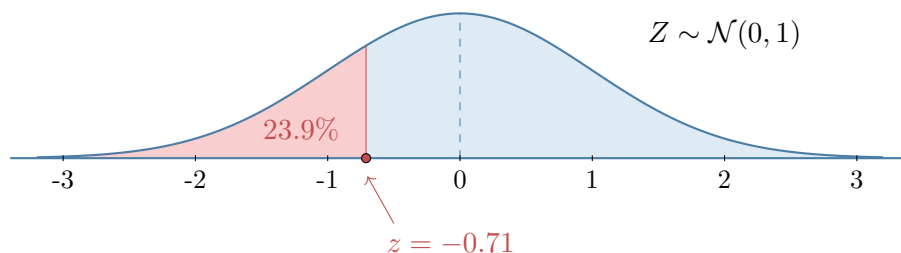
四个值都大于 5，因此正态近似是合理的。现在可以利用这些信息计算 z 分数：

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}} \approx \frac{0.188 - 0.202}{\sqrt{\frac{(0.188)(0.812)}{740} + \frac{(0.202)(0.798)}{741}}} \approx -0.71.$$

由于这是左尾检验，P值可以直接从 z 表中读取：

$$P\text{值} = P(\hat{P}_1 - \hat{P}_2 \leq -0.0146) \approx P(Z \leq -0.71) \approx 0.2389.$$

从几何上看，我们刚刚计算了左尾区域，其面积约占整个图形的 23.9%：



按照构造，这个面积与根据归无假设建立的抽样分布中，样本数据 $\hat{p}_1 - \hat{p}_2$ 左侧的面积相同。换句话说，如果两个班级中橙色彩虹糖的真实比例确实相同，那么得到至少像本次结果一样，在 $\hat{p}_1 < \hat{p}_2$ 方向上偏离这么远的差值，其概率约为 23.9%。

步骤3：将P值与 α 比较

比较得到：

$$0.2389 > 0.05.$$

因此，我们不拒绝 H_0 。

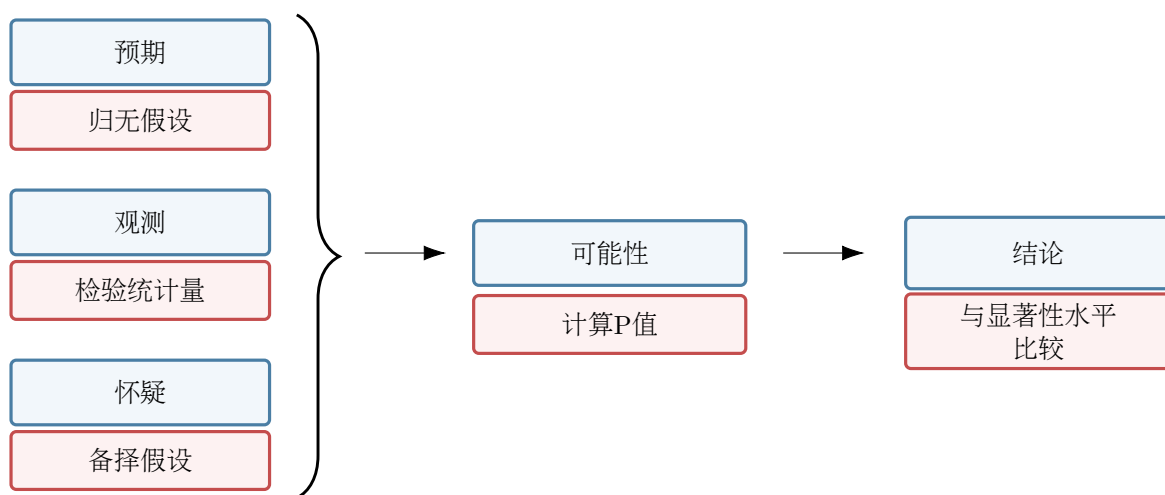
步骤4：给出解释

解释

在 5% 的显著性水平下，样本没有提供有说服力的证据，表明班级1中橙色彩虹糖的比例低于班级2中的比例。

6 假设检验总结

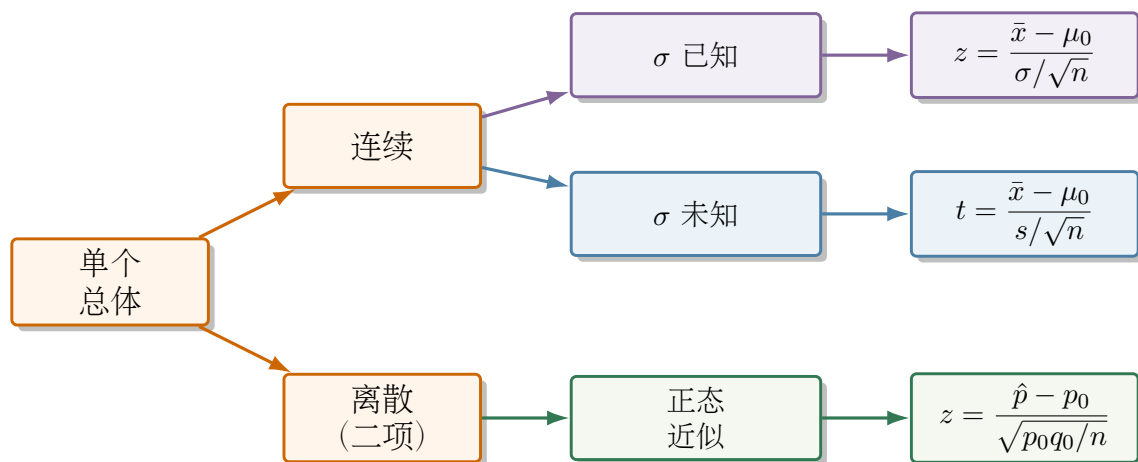
我们见过的所有假设检验都遵循相同的整体模式：



到目前为止，我们已经学习了七种主要的假设检验：

1. 在 σ 已知时，检验单个总体均值 μ ；
2. 在 σ 未知时，检验单个总体均值 μ ；
3. 检验单个二项成功概率 p ；
4. 对配对数据检验 $\mu_X - \mu_Y$ ；
5. 在总体标准差已知时，对独立数据检验 $\mu_X - \mu_Y$ ；
6. 在总体标准差未知时，对独立数据检验 $\mu_X - \mu_Y$ ；
7. 对两个独立二项试验检验差 $p_1 - p_2$ 。

这些检验可以分为两大类：单个总体的检验，以及同时检验两个总体。在单个总体的情况下，下列流程图整理了不同的检验。



当同时检验两个总体时，我们得到以下流程图：

