概率论与数理统计

崔文泉 (wqcui@ust.edu.cn)

2021 Autumn

第五章　假设检验

目录

§ 5.1.1 原假设，备择假设

设$X_1, \cdots, X_n$为从总体X中抽取的样本，假设总体有概率函数$f(x;\theta)$。这里θ可以是多维的，是已知的或者未知的参数。我们的问题就是要基于样本$X_1, \cdots, X_n$检验假设“总体有概率函数$f(x;\theta)$”，表示为

$$H_0 : X_1, \cdots, X_n \sim f(x;\theta) \tag{1}$$

以及在承认这个分布的前提下,对有关参数θ的假设作出判断这两类问题。即

$$H_0 : \theta \in \Theta_0 \leftrightarrow H_1 : \theta \in \Theta_1 \tag{2}$$

§ 5.1.1 原假设，备择假设

在很多场合下，$\Theta_0 \bigcup \Theta_1 = \Theta$。在本章中，我们主要考虑如下三种形式的参数假设

$$
\begin{array}{lll}
1. & H_0 : \theta \geq \theta_0 \leftrightarrow H_1 : \theta < \theta_0 & (\text{左边检验}) \\
2. & H_0' : \theta \leq \theta_0 \leftrightarrow H_1' : \theta > \theta_0 & (\text{右边检验}) \\
3. & H_0'' : \theta = \theta_0 \leftrightarrow H_1'' : \theta \neq \theta_0 & (\text{双边检验})
\end{array}
$$

当概率函数$f(x;\theta)$完全指定时(即都已知)，称H_0为一个简单假设；否则，则称为复合假设。在具体的问题中，检验(1)可以表现为对参数进行某种约束的形式。

§5.1.1 原假设，备择假设

Example

假设某种产品的某个性能指标服从正态分布$N(\mu,\sigma^2)$，由以往的经验知道$\mu=10,\sigma^2=4$，现从某批产品中随机抽取部分产品测量该指标，得到$x_1,\cdots,x_n$，问这批产品在该指标上平均性能是否发生变化？

[分析] 我们用X表示总体，由题设知道，“该批产品在该指标性能上没有发生变化”就可以表示为

$$H_0:\mu=10$$

此假设可以等价的表示为

$$H_0:X\sim N(10,4)$$

§ 5.1.1 原假设，备择假设

显然，这是一个简单假设。若σ^2未知，则相当于要检验

$$H_0 : \mu = 10, \sigma^2 \text{任意}$$

或者等价的表示为

$$H_0 : X \sim N(10, \sigma^2), \sigma^2 > 0$$

此时，假设H_0就是一个复合假设。

§5.1.1 原假设，备择假设

在检验一个假设的过程中，我们经常用到下述术语:

- 原假设和对立假设:

  我们常把被检验的假设称为是“原假设”，也称为“零假设”，常用H_0表示；而与原假设相对立的一个备选的假设，就称为“对立假设”，常用H_1表示。例如在上例中，原假设就是H_0。我们没有明确指定H_0不成立时的分布，因此对立假设就可以表示为“$H_1: \mu \neq 10$”，而如果我们只是感兴趣$\mu = 10$与$\mu = 20$，那么对立假设就可以表示成“$H_1: \mu = 20$”。因此，一个检验问题也可以表示成(2)的形式。

§5.1.2 接受域与拒绝域，两类错误与功效函数

- 检验统计量，接受域和拒绝域:

  在检验一个假设中使用的统计量，称为检验统计量。使得原假设得以成立的那些样本所在的区域，就称为接受域；而使得原假设被否定的那些样本所在的区域，则称为拒绝域。设$T(X) = T(X_1, \cdots, X_n)$为检验假设(2)的检验统计量，则接受域可以表示成$T(x) = T(x_1, \cdots, x_n) \in \mathcal{X}_1$，其中$\mathcal{X}_1$为某个欧氏空间的某个区域。也可以表示成

$$\phi: \text{当} T(x) \in \mathcal{X}_1 \text{时，接受} H_0, \text{否则就否定} \tag{3}$$

  其中ϕ称为检验问题(2)的一个检验。

§5.1.2 接受域与拒绝域，两类错误与功效函数

- 功效函数和两类错误:

Definition

设总体分布包含若干个未知单数$\theta=(\theta_1,\cdots,\theta_k)$，$H_0$是关于这些参数的一个原假设。设有样本$X_1,\cdots,X_n$，$\phi$ 是基于这些样本构造的H_0的一个检验，则称

$$\beta_\phi(\theta)=\beta_\phi(\theta_1,\cdots,\theta_k)=P_{\theta_1,\cdots,\theta_k}(\text{在检验}\phi\text{之下，}H_0\text{被否定})$$

为检验ϕ的功效函数。

容易明白：当某一个特定的参数值$(\theta_1^0,\cdots,\theta_k^0)$使得$H_0$成立时，我们希望$\beta_\phi(\theta_1^0,\cdots,\theta_k^0)$ 尽量小(当H_0成立时，我们不希望否定它)；反之，若$(\theta_1^0,\cdots,\theta_k^0)$属于对立假设，则我们希望$\beta_\phi(\theta_1^0,\cdots,\theta_k^0)$ 尽量大(当H_0不成立时，我们希望否定它)。

§5.1.2 接受域与拒绝域，两类错误与功效函数

Definition

在检验一个假设过程中，称原假设H_0被否定而事实上H_0是正确的这种错误为“第一类错误”(Type-I error)，而称原假设被接受但是事实上H_0是错误的这种错误为“第二类错误”(Type-II error)。

- 我们也常常称第一类错误为“弃真”，而称第二类错误为“取伪”。这两个错误我们总是不能避免的。根据功效函数和两类错误的定义，易知

$$
\begin{aligned}
P(Type-I \quad error) &= \begin{cases} \beta_\phi(\theta), & \theta \in \Theta_0 \\ 0, & \theta \in \Theta_1 \end{cases} \\
P(Type-II \quad error) &= \begin{cases} 0, & \theta \in \Theta_0 \\ 1-\beta_\phi(\theta), & \theta \in \Theta_1 \end{cases}
\end{aligned}
$$

§5.1.3 显著性检验

这两类错误往往不可兼顾，我们也象区间估计理论中首要控制可靠性一样，这里首先控制第一类错误。因此，这种只限制第一类错误的原则，就称为“显著性检验”。

§5.1.3 显著性检验

Definition

设ϕ是假设H_0的一个检验，$\beta_\phi(\theta)$为它的功效函数，$0 < \alpha < 1$，若

$$\beta_\phi(\theta) \le \alpha, \quad \forall\ \theta \in \Theta_0$$

则称ϕ为一个水平α检验，或者说，检验ϕ有水平α.

在这些概念下，一个检验问题就可表述为

- 表述检验问题$H_0 : \theta \in \Theta_0 \leftrightarrow H_1 : \theta \in \Theta_1$;
- 给定检验水平α，抽样得到样本$X_1, \cdots, X_n$，构造$H_0 \leftrightarrow H_1$的一个显著性检验ϕ;
- 对检验问题作出推断。

§5.2.1 正态分布参数的检验

例5.2.1. 某餐厅每天营业额服从正态分布，以往老菜单其均值为8000元，标准差为640元。一个新的菜单挂出后，九天中平均每天营业额为8300元，经理想知道这个差别是否是由于新菜单而引起的。(这里假定标准差不变。)

§5.2.1 正态分布参数的检验

例5.2.1. 某餐厅每天营业额服从正态分布，以往老菜单其均值为8000元，标准差为640元。一个新的菜单挂出后，九天中平均每天营业额为8300元，经理想知道这个差别是否是由于新菜单而引起的。(这里假定标准差不变。)

解：确立原假设与备择假设：

$$H_0 : \mu = 8000 \leftrightarrow H_1 : \mu > 8000$$

现样本给出$\bar{x} = 8300$,这仅是μ的一个估计。

本例中可用样本均值作为检验统计量。当H_0为真时，新菜单挂出后，每天营业额仍服从正态分布$N(8000, 640^2)$.所以

$$\bar{X} \sim N(8000, \frac{640^2}{9})$$

当平均观察值$\bar{x} \geq c$时，拒绝H_0; 当平均观察值$\bar{x} < c$时，接受H_0。

§ 5.2.1 正态分布参数的检验

取$\alpha = 0.05$,这意味着

$$P(H_0\text{为真，但被拒}) = 0.05$$

这里，"H_0为真"表示样本实际是来自总体$N(8000, 640)$,"被拒绝"表示样本均值$\bar{X} \geq c$.

所以，$P_{\mu_0}(\bar{X} \geq c) = 0.05$.

$1 - \Phi(\frac{c-8000}{640/3}) == 0.05$,

$c = 8000 + 1.645 \times 640/3 = 8350.9$

§ 5.2.1 正态分布参数的检验

一、单个正态总体$N(\mu,\sigma^2)$参数的检验

(1). 均值μ的检验 考虑如下三种形式的假设检验问题：

1. $H_0:\mu\geq\mu_0 \leftrightarrow H_1:\mu<\mu_0$
2. $H_0':\mu\leq\mu_0 \leftrightarrow H_1':\mu>\mu_0$
3. $H_0'':\mu=\mu_0 \leftrightarrow H_1'':\mu\neq\mu_0$

(a) 方差σ^2已知

我们知道$\bar{X}$是μ的一个无偏估计，因此一个看上去合理的检验为

$$\phi:\text{当}\bar{X}\geq c\text{时，接受}H_0,\text{不然就拒绝}$$

§5.2.1 正态分布参数的检验

为定出常数c，使检验ϕ有水平α，即令其功效函数在原假设H_0下

$$\beta_\phi(\mu) = P_{\mu\geq\mu_0}(\bar{X} < c) \leq \alpha$$

$$\Leftrightarrow P_{\mu\geq\mu_0}(\bar{X} - \mu < c - \mu) = \Phi(\frac{\sqrt{n}(c-\mu)}{\sigma}) \leq \alpha$$

显然，$\beta_\phi(\mu)$为μ的减函数，故欲使上述不等式对所有的$\mu \geq \mu_0$成立，只需$\beta_\phi(\mu_0) = \alpha$，则得到$c = \mu_0 - \sigma u_\alpha/\sqrt{n}$。从而得到假设$H_0 \leftrightarrow H_1$的一个检验为

$$\phi: 当\bar{X} \geq \mu_0 - \sigma u_\alpha/\sqrt{n}时，接受H_0, 不然就拒绝$$

类似的可以得到假设$H_0' \leftrightarrow H_1'$和$H_0'' \leftrightarrow H_1''$的一个检验分别为

$$\phi': 当\bar{X} \leq \mu_0 + \sigma u_\alpha/\sqrt{n}时，接受H_0', 不然就拒绝$$

和

$$\phi'': 当|\bar{X} - \mu_0| \leq \sigma u_{\alpha/2}/\sqrt{n}时，接受H_0'', 不然就拒绝$$

§ 5.2.1 正态分布参数的检验

例5.2.2. 微波炉在炉门关闭时的辐射量是一个重要得质量指标。某厂该指标服从正态分布$N(\mu,\sigma^2)$，常期以来$\sigma = 0.01$,且均值都符合要求不超过0.12.为检查近期产品的质量，抽查了25台，得其炉门关闭时辐射量的均值$\bar{x} = 0.1203$.试问在$\alpha = 0.05$水平上该厂炉门关闭时辐射量是否升高了？

§ 5.2.1 正态分布参数的检验

例5.2.2. 微波炉在炉门关闭时的辐射量是一个重要得质量指标。某厂该指标服从正态分布$N(\mu,\sigma^2)$，常期以来$\sigma = 0.01$,且均值都符合要求不超过0.12.为检查近期产品的质量，抽查了25台，得其炉门关闭时辐射量的均值$\bar{x} = 0.1203$.试问在$\alpha = 0.05$水平上该厂炉门关闭时辐射量是否升高了？

解：建立假设：

$$H_0 : \mu \le 0.12 \leftrightarrow H_1 : \mu > 0.12$$

在$\alpha = 0.05$时，$u_{0.95} = 1.645$,拒绝域应为$\{u \ge 1.645\}$.现由观测值求得

$$u = \frac{0.1203 - 0.12}{0.1/\sqrt{25}} = 0.015 < 1.645$$

因而在$\alpha = 0.05$水平下，不能拒绝H_0，认为当前生产的微波炉关门时的辐射量无明显升高。

§5.2.1 正态分布参数的检验

例5.2.3. 某厂生产需用玻璃纸作包装，按规定供应商供应的玻璃纸的横向延伸率不应低于65.已知该指标服从正态分布$N(\mu,\sigma^2)$,σ一直稳定于5.5.从近期来货中抽查了100个样品，得样本均值$\bar{x}=55.06$,试问在$\alpha=0.05$水平上能否接收这批玻璃纸？

§ 5.2.1 正态分布参数的检验

例5.2.3. 某厂生产需用玻璃纸作包装，按规定供应商供应的玻璃纸的横向延伸率不应低于65.已知该指标服从正态分布$N(\mu,\sigma^2)$,σ一直稳定于5.5.从近期来货中抽查了100个样品，得样本均值$\bar{x}=55.06$,试问在$\alpha=0.05$水平上能否接收这批玻璃纸？

解：建立假设：

$$H_0:\mu\geq 65 \leftrightarrow H_1:\mu<65$$

在$\alpha=0.05$时，$u_\alpha=-1.645$,拒绝域应取作$\{u\leq -1.645\}$.现由样本求得

$$u=\frac{55.06-65}{5.5/\sqrt{100}}=-18.07<-1.645$$

故应拒绝H_0，不能接收这批玻璃纸。

§5.2.1 正态分布参数的检验

例5.2.4. 某洗涤剂厂有一台瓶装洗洁精的灌装机，在生产正常时，每瓶洗洁精的净重服从正态分布，均值为454g，标准差为12g。为检查近期机器工作是否正常，从中抽出16瓶，称得其净重得平均值为$\bar{x} = 456.64$.试对机器工作正常与否作出判断。(取$\alpha = 0.01$,并假定σ不变。)

§5.2.1 正态分布参数的检验

例5.2.4. 某洗涤剂厂有一台瓶装洗洁精的灌装机，在生产正常时，每瓶洗洁精的净重服从正态分布，均值为454g，标准差为12g。为检查近期机器工作是否正常，从中抽出16瓶，称得其净重得平均值为$\bar{x} = 456.64$.试对机器工作正常与否作出判断。(取$\alpha = 0.01$,并假定σ不变。)

解：检验的假设为

$$H_0 : \mu = 454 \leftrightarrow H_1 : \mu \neq 454$$

在$\alpha = 0.01$时，$u_{1-\frac{\alpha}{2}} = u_{0.995} = 2.58$,从而拒绝域为$\{|u| \geq 2.58\}$.现由样本求得

$$u = \frac{456.64 - 454}{12/\sqrt{16}} = 0.88$$

由于$|u| < 2.58$,故不能拒绝H_0，认为机器正常。

§5.2.1 正态分布参数的检验

例5.2.5. 根据某地环境保护法规定，倾入河流的废水中某种有毒化学物质含量不得超过3ppm。该地区环保组织对沿河各厂进行检查，测定每日倾入河流得废水中该物质的含量。某厂连日的记录为

3.1	3.2	3.3	2.9	3.5	3.4	2.5	4.3
2.9	3.6	3.2	3.0	2.7	3.5	2.9	

试在显著性水平$\alpha = 0.05$上判断该厂是否符合环保规定(假设废水中有毒物质含量$X \sim N(\mu, \sigma^2)$.

§5.2.1 正态分布参数的检验

例5.2.5. 根据某地环境保护法规定，倾入河流的废水中某种有毒化学物质含量不得超过3ppm。该地区环保组织对沿河各厂进行检查，测定每日倾入河流得废水中该物质的含量。某厂连日的记录为

3.1	3.2	3.3	2.9	3.5	3.4	2.5	4.3
2.9	3.6	3.2	3.0	2.7	3.5	2.9	

试在显著性水平$\alpha=0.05$上判断该厂是否符合环保规定(假设废水中有毒物质含量$X\sim N(\mu,\sigma^2)$).

解：提出假设：

$$H_0:\mu\le 3 \leftrightarrow H_1:\mu>3$$

由于这里σ未知，故采用t检验，现
在$n=15$,在$\alpha=0.05$时$t_{0.95}(14)=1.7613$，故拒绝域为

$$\{t>1.7613\}$$

§ 5.2.1 正态分布参数的检验

现根据样本求得$\bar{x} = 3.2, s = 0.436$,从而有

$$u = \frac{3.2 - 3}{0.436/\sqrt{15}} = 1.7766 > 1.7613$$

样本落入拒绝域，因此在$\alpha = 0.05$水平上认为该厂废水中有毒物质含量超标，不符合环保规定。

§5.2.1 正态分布参数的检验

Example

对假设检验问题1，如果我们还要求“犯第二类错误的概率要小于指定的$\beta > 0$”该怎么办?

§5.2.1 正态分布参数的检验

Example

对假设检验问题1，如果我们还要求“犯第二类错误的概率要小于指定的$\beta > 0$”该怎么办?

解：根据功效函数和两类错误的定义，知道等价的要求

$$\beta_\phi(\mu) \geq 1 - \beta, \quad \mu < \mu_0 \tag{4}$$

但是，当$\mu < \mu_0$但μ接近μ_0时，$\beta_\phi(\mu) \approx \alpha$，而因为$\alpha, \beta$一般都很小，因此一般有$\alpha < 1 - \beta$，这就看出要求(1)无法达到。我们只能放松一些，要求对某个指定的$\mu_1 < \mu_0$，有

$$\beta_\phi(\mu) \geq 1 - \beta, \quad \mu < \mu_1 \tag{5}$$

§ 5.2.1 正态分布参数的检验

因为$\beta_\phi(\mu)$为μ的减函数，因此等价于要求

$$\beta_\phi(\mu_1) \geq 1-\beta$$

此即

$$\Phi\left(\frac{\sqrt{n}(\mu_0-\mu)}{\sigma} - u_\alpha\right) \geq 1-\beta$$

等价的得到

$$n \geq \sigma^2(u_\alpha+u_\beta)^2/(\mu_0-\mu)^2$$

也即要满足题目中的要求，样本大小至少要达到上式右边那么大。□

§5.2.1 正态分布参数的检验

(b) 方差σ^2未知

在σ^2已知时，我们曾求出假设$H_0 \leftrightarrow H_1$的检验为

$$\phi: 当\bar{X} \geq \mu_0 - \sigma u_\alpha/\sqrt{n}时，接受H_0,不然就拒绝$$

当方差σ^2未知时，我们用S^2估计σ^2。但是用S代替σ后，分布也发生了变化，由正态分布变成了自由度为$n-1$的t分布。类似的计算，可以得到$c = -t_\alpha(n-1)$. 从而得到假设$H_0 \leftrightarrow H_1$的一个检验为[1]

$$\phi: 当\bar{X} \geq \mu_0 - St_\alpha(n-1)/\sqrt{n}时，接受H_0,不然就拒绝$$

[1]我们也可以先提出如下形式的检验:

$$\phi: 当\sqrt{n}(\bar{X}-\mu_0)/S \geq -t_\alpha(n-1)时，接受H_0,不然就拒绝$$

然后验证此检验有水平α。事实上，对任意的$\mu \geq \mu_0$，都有

$$\beta_\phi(\mu) = P_\mu(\sqrt{n}(\bar{X}-\mu_0)/S < -t_\alpha(n-1)) \leq P_\mu(\sqrt{n}(\bar{X}-\mu)/S < -t_\alpha(n-1)) = \alpha$$

即此检验有水平α，因此它是检验问题$H_0 \leftrightarrow H_1$的一个显著性水平为α的检验。

§ 5.2.1 正态分布参数的检验

类似的可以得到假设$H_0' \leftrightarrow H_1'$和$H_0'' \leftrightarrow H_1''$的一个检验分别为

ϕ' : 当$\bar{X} \leq \mu_0 + St_\alpha(n-1)/\sqrt{n}$时，接受$H_0'$, 不然就拒绝

和

ϕ'' : 当$|\bar{X} - \mu_0| \leq St_{\alpha/2}(n-1)/\sqrt{n}$时，接受$H_0''$, 不然就拒绝

§ 5.2.1 正态分布参数的检验

(2). 方差σ^2的检验 考虑如下三种形式的假设检验问题：

1. $H_0 : \sigma^2 \geq \sigma_0^2 \leftrightarrow H_1 : \sigma^2 < \sigma_0^2$

图

2. $H_0' : \sigma^2 \leq \sigma_0^2 \leftrightarrow H_1' : \sigma^2 > \sigma_0^2$
3. $H_0'' : \sigma^2 = \sigma_0^2 \leftrightarrow H_1'' : \sigma^2 \neq \sigma_0^2$

假定总体均值μ未知，由于S^2是σ^2的一个无偏估计，对假设$H_0 \leftrightarrow H_1$，故一个看上去合理的检验为

$$\phi : \text{当}S^2 \geq c\text{时，接受}H_0,\text{不然就拒绝}$$

其功效函数

$$\beta_\phi(\sigma^2) = P(S^2 < c) = P((n-1)S^2/\sigma^2 < (n-1)c/\sigma^2)$$

为σ^2的减函数，

§ 5.2.1 正态分布参数的检验

其功效函数

$$\beta_\phi(\sigma^2) = P(S^2 < c) = P((n-1)S^2/\sigma^2 < (n-1)c/\sigma^2)$$

为σ^2的减函数，因此欲使$\beta_\phi(\sigma^2) \le \alpha$，$\forall \sigma^2 \ge \sigma_0^2$，只需$\beta_\phi(\sigma_0^2) = \alpha$，从而得到$c = \frac{\sigma_0^2}{n-1}\chi^2_{1-\alpha}(n-1)$。故检验$\phi$为

$$\phi: 当S^2 \ge \frac{\sigma_0^2}{n-1}\chi^2_{1-\alpha}(n-1)时，接受H_0, 不然就拒绝$$

类似可以得到检验问题2，3的一个检验分别为

$$\phi': 当S^2 \le \frac{\sigma_0^2}{n-1}\chi^2_{\alpha}(n-1)时，接受H_0', 不然就拒绝$$

和

$$\phi'': 当S^2 \le \frac{\sigma_0^2}{n-1}\chi^2_{\alpha/2}(n-1)或者S^2 \ge \frac{\sigma_0^2}{n-1}\chi^2_{1-\alpha/2}(n-1)时，$$

接受H_0'',不然就拒绝.检验ϕ''是一个习惯上取法。

§5.2.1 正态分布参数的检验

详细的推导如下：

对假设问题3，一个看上去合理的检验为

$$\phi'' : \text{当} c_1 \leq \frac{S^2}{\sigma^2} \leq c_2 \text{时，接受} H_0'', \text{不然就拒绝}$$

为定出c_1, c_2，令其功效函数

$$P(\frac{S^2}{\sigma_0^2} > c_2 \text{或者} \frac{S^2}{\sigma_0^2} < c_1) = P(\frac{S^2}{\sigma_0^2} > c_2) + P(\frac{S^2}{\sigma_0^2} < c_1) \leq \alpha$$

类似于区间估计里的处理方法，我们方便上取

$$P\left(\frac{S^2}{\sigma_0^2} > c_2\right) \leq \alpha/2$$

$$P\left(\frac{S^2}{\sigma_0^2} < c_1\right) \leq \alpha/2$$

从而定出c_1和c_2。

§ 5.2.1 正态分布参数的检验

另外，**需要注意的是**上述推导我们都是假设总体的均值μ未知。若总体的均值μ已知，则需要修改S^2为

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$$

相应的分布也从自由度为$n-1$的χ^2变成了自由度为n的χ^2。

§5.2.1 正态分布参数的检验

例5.2.6. 某种导线的电阻服从$N(\mu,\sigma^2)$，μ未知，其中一个质量指标是电阻标准差不得大于0.005Ω.现从中抽取了九根导线测其电阻，测得样本标准差$s=0.0066$，试问在$\alpha=0.05$水平上能否认为这批导线的电阻波动合格？

§ 5.2.1 正态分布参数的检验

例5.2.6. 某种导线的电阻服从$N(\mu,\sigma^2)$，μ未知，其中一个质量指标是电阻标准差不得大于0.005Ω.现从中抽取了九根导线测其电阻，测得样本标准差$s=0.0066$，试问在$\alpha=0.05$水平上能否认为这批导线的电阻波动合格？

解：建立假设：

$$H_0:\sigma\le 0.005 \leftrightarrow H_1:\sigma>0.005$$

这是一个单边检验，在$n=9,\alpha=0.05$时，$\chi^2_{0.95}(8)=15.507$，拒绝域为$W=\{\chi^2\ge 15.507\}$,现由样本求得

$$\chi^2=\frac{8\dot{0}.0066^2}{0.005^2}=13.94<15.507$$

故可接受原假设，在$\alpha=0.05$水平上认为这批导线的电阻波动合格。

§5.2.2 两样本正态总体的情形

为了检验某肥料是否能显著提高玉米产量，可以设计一个随机试验：选择两块条件一样的试验区，把两试验区各分成若干小块，一个试验区的各小块施肥，另一个试验区的各小块不施肥，最后统计收成，可以采用如下的检验方法来检验玉米产量差别，从而知道肥料是否有效。

设总体$X \sim N(\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2), -\infty < \mu_1, \mu_2 < \infty, \sigma_1^2, \sigma_2^2 > 0; X_1, \cdots X_n$是从总体$X$中抽取的一个样本，$Y_1, \cdots, Y_n$是从总体$Y$中抽取的一个样本。设来自不同总体的样本相互独立。下面设考虑有关均值差$\mu_1 - \mu_2$和方差比$\sigma_1^2 \sigma_2^2$的检验。取显著性水平为α.下面举例说明。

§5.2.2 两样本正态总体的情形

例5.2.7. 甲乙两个农业试验区种植玉米，除了甲区磷肥外，其他试验条件都相同，把两个试验区分别均分成10个和9个小区统计产量(单位：千克)，得数据如下：

甲区	62	57	65	63	58	57	60	60	58
乙区	50	59	56	57	56	55	57		

假定甲乙两区中每小块得玉米产量分别服从$N(\mu_1,\sigma^2),N(\mu_2,\sigma^2)$,其中$\mu_1,\mu_2,\sigma^2$未知。试问在显著性水平$\alpha=0.1$下磷肥对玉米得产量是否有效？

§5.2.2 两样本正态总体的情形

例5.2.7. 甲乙两个农业试验区种植玉米，除了甲区磷肥外，其他试验条件都相同，把两个试验区分别均分成10个和9个小区统计产量(单位：千克)，得数据如下：

甲区	62	57	65	63	58	57	60	60	58
乙区	50	59	56	57	56	55	57		

假定甲乙两区中每小块得玉米产量分别服从$N(\mu_1, \sigma^2), N(\mu_2, \sigma^2)$,其中$\mu_1, \mu_2, \sigma^2$未知。试问在显著性水平$\alpha = 0.1$下磷肥对玉米得产量是否有效？

解：磷肥对玉米产量有效果等价于$\mu_1 > \mu_2$,故将其设为对立假设，假设检验问题是构造基于$\mu_1 - \mu_2$的极大似然估计$\bar{X} - \bar{Y}$的检验统计量

$$T = \frac{\bar{X} - \bar{Y}}{S_\omega \sqrt{\frac{1}{m} + \frac{1}{n}}}$$

§5.2.2 两样本正态总体的情形

当H_0成立时，$T \sim t_{m+n-2}$,于是拒绝域为$\{T > t_{m+n-2(\alpha)}\}$，由所得数据算出检验统计量得观测值为:

$$t = \frac{\bar{x} - \bar{y}}{s_\omega \sqrt{\frac{1}{m} + \frac{1}{n}}} = 3.23.$$

由$\alpha = 0.1$得临界值为$t_{m+n-2}(\alpha/2) = t_{17}(0.1) \approx 1.33 < 3.23$,因此拒绝$H_0$,即可在显著性水平0.1下认为磷肥对玉米得产量有显著性影响。

§ 5.2.2 两样本正态总体的情形

假定了两个正态总体的方差是相等的，即$\sigma_1^2 = \sigma_2^2 = \sigma$.现在我们根据样本来检验这个方差齐性的假设，即要检验

$$H_0 : \frac{\sigma_1^2}{\sigma_2^2} = 1 \leftrightarrow H_1 : \frac{\sigma_1^2}{\sigma_2^2} \neq 1.$$

§ 5.2.2 两样本正态总体的情形

假定了两个正态总体的方差是相等的，即$\sigma_1^2 = \sigma_2^2 = \sigma$.现在我们根据样本来检验这个方差齐性的假设，即要检验

$$H_0 : \frac{\sigma_1^2}{\sigma_2^2} = 1 \leftrightarrow H_1 : \frac{\sigma_1^2}{\sigma_2^2} \neq 1.$$

解：因为σ_1^2和σ_2^2的极大既然估计分别是：

$$\hat{\sigma}_1^2 = \frac{1}{m}\sum_{i=1}^{m}(X_i - \bar{X})^2, \hat{\sigma_2^2} = \frac{1}{n}\sum_{i=1}^{m}(Y_i - \bar{Y})^2.$$

在$\theta = \frac{\sigma_1^2}{\sigma_2^2}$的极大似然估计$\hat{\theta} = \frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2}$的基础上可以构造检验统计量

$$F = \frac{S_1^2}{S_2^2} = \frac{(m-1)\hat{\sigma}_1^2/m}{(n-1)\hat{\sigma}_2^2/n}$$

§5.2.2 两样本正态总体的情形

注意到F中的分子和分母分别是X和Y的样本方差。当零假设成立时，$F \sim F_{m-1,n-1}$,于是拒绝域为

$$\{F < F_{m-1,n-1}(\alpha/2)\text{或}F > F_{m-1,n-1}(1-\alpha/2)\}$$

由数据算得检验统计量F的观测值$f = 1.19$,如果取显著性水平$\alpha = 0.2$,那么临界值为

$$F_{9,8}(0.1) = 2.44, F_{9,8}(0.9) = 1/F_{8,9}(0.1) = 0.41$$

(如果$X \sim F_{m,n}$,则$1/X \sim F_{n,m}$).

易见$0.41 < 1.19 < 2.44$,因此不能拒绝H_0,即在显著性水平0.2下可以认为上例中所作的方差齐性假定是合理的。

§5.2.2 两样本正态总体的情形

例5.2.8. 甲、乙两台机床分别加工某种轴，轴的直径分别服从正态分布$N(\mu_1,\sigma_1^2)$与$N(\mu_1,\sigma_2^2)$,为比较两台机床的加工精度有无显著差异。从各自加工的轴中分别抽取若干根轴测其直径，结果如下(取$\alpha=0.05$)：

总体	样本容量	直径							
X(甲)	8	20.5	19.8	19.7	20.4	20.1	20.0	19.0	19.9
Y(乙)	7	20.7	19.8	19.5	20.8	20.4	19.6	20.2	

§ 5.2.2 两样本正态总体的情形

例5.2.8. 甲、乙两台机床分别加工某种轴，轴的直径分别服从正态分布$N(\mu_1,\sigma_1^2)$与$N(\mu_1,\sigma_2^2)$,为比较两台机床的加工精度有无显著差异。从各自加工的轴中分别抽取若干根轴测其直径，结果如下(取$\alpha=0.05$):

总体	样本容量	直径							
X(甲)	8	20.5	19.8	19.7	20.4	20.1	20.0	19.0	19.9
Y(乙)	7	20.7	19.8	19.5	20.8	20.4	19.6	20.2	

解：建立假设：

$$H_0:\sigma_1^2=\sigma_2^2 \leftrightarrow H_1:\sigma_1^2\neq\sigma_2^2$$

在$n=8,m=7,\alpha=0.05$时，$F_{0.025}(7,6)=\frac{1}{F_{0.975}(6,7)}=\frac{1}{5.12}=0.195$，拒绝域为$W=\{F\leq 0.195 \quad or \quad F\geq 5.70\}$, 现由样本求得$s_X^2=0.2164, s_Y^2=0.2729$，从而$F=0.793$，未落入拒绝域，因而在$\alpha=0.05$水平上认为两台机床加工精度一致.

§ 5.2.2 两样本正态总体的情形

例5.2.9. 设在例5.2.8中，两台机床加工的轴是同类的，其直径分别服从正态分布，经例5.2.8的F检验已确认其方差相等，即认为两台机床加工的直径分别服从$N(\mu_1, \sigma^2)$ 与$N(\mu_2, \sigma^2)$，现进一步检验两台机床加工的轴的平均直径是否一致(取$\alpha = 0.05$)。

§5.2.2 两样本正态总体的情形

例5.2.9. 设在例5.2.8中，两台机床加工的轴是同类的，其直径分别服从正态分布，经例5.2.8的F检验已确认其方差相等，即认为两台机床加工的直径分别服从$N(\mu_1,\sigma^2)$ 与$N(\mu_2,\sigma^2)$，现进一步检验两台机床加工的轴的平均直径是否一致(取$\alpha=0.05$)。

解：问题相当于检验

$$H_0:\mu_1=\mu_2 \leftrightarrow H_1:\mu_1\neq\mu_2$$

由于两总体方差一致但未知，故统计量为

$$T=\frac{\bar{X}-\bar{Y}}{S_W\sqrt{\frac{1}{n}+\frac{1}{m}}},\quad 其中S_W^2=\frac{(n-1)S_X^2+(m-1)S_X^2}{n+m-2}$$

§ 5.2.2 两样本正态总体的情形

在$n = 8, m = 7, \alpha = 0.05$时，$t_{0.975}(13) = 2.1604$，从而拒绝域为$\{|t| > 2.1604\}$,现由样本求得$\bar{x} = 19.295, \bar{y} = 20.143, s_W^2 = 0.2425, s_W = 0.4924$，从而$T = -0.8554$，未落入拒绝域，因而在$\alpha = 0.05$水平上认为两台机床加工的轴的平均直径一致.

§5.2.3 基于成对数据的检验

在上述两样本正态总体的假设检验中，要求两个样本是独立的，但是没有要求样本量相等。

有一类数据叫做成对数据$\{(X_1, X_2), \cdots, (X_n, Y_n)\}$, 比如一个病人在用药前后测得的指标分别为$X$和$Y$,则$X$与$Y$总是一起出现的，且由于它们是同一个个体的指标，故具有很大的相关性而绝对不是独立的，这与两样本正态总体有本质区别。另外，两样本检验问题要求样本$X_1, \cdots, X_m$是同分布的($Y_1, \cdots, Y_m$ 亦然)，而成对数据则无此要求，而要求$X_1 - Y_1, \cdots, X_n - Y_n$是同分布。比如病人可以来自两个不同性别、种族、年龄层的人。

要检验用药前后的指标有无显著差别，可以构造一个新的总体$Z = Y - X$及样本$Z_1 = X_1 - Y_1, \cdots, Z_n = X_n - Y_n$,相应的假设检验是一样本的！

在实际问题中，如果发现有两个样本且其样本量是相等的，则要求检查独立性和同分布性，否则可能是成对数据。

§ 5.2.4 0-1分布中未知参数p的检验

产品验收时，需要检验不合格率是否小于某给定的一个数。

设$(X_1, \cdots, X_n)$是取自总体X的一个样本，该总体服从0-1分布，取1的概率为p。常见的假设有三种：

(1) $H_0 : p = p_0 \leftrightarrow H_1 : p \neq p_0$;

(2) $H_0 : p = p_0 \leftrightarrow H_1 : p > p_0$或$H_0 : p \leq p_0 \leftrightarrow H_1 : p > p_0$;

(3) $H_0 : p = p_0 \leftrightarrow H_1 : p < p_0$或$H_0 : p \geq p_0 \leftrightarrow H_1 : p < p_0$;

假定样本量n较大，取显著性水平为α.由于p的极大似然估计为$\bar{X}$,取"标准化"过的检验统计量

$$T = \sqrt{n}\frac{\bar{X} - p_0}{\sqrt{p_0(1 - p_0)}}$$

其中p_0和$p_0(1 - p_0)/n$分别为$\bar{X}$在零假设$p = p_0$下的期望和方差，

§ 5.2.4 0-1分布中未知参数p的检验

从而当H_0成立时，由中心极限定理近似地有$T \sim N(0,1)$.于是上述三种检验的拒绝域分别为

$$\{|T| > u_{\alpha/2}\}, \{T > u_\alpha\}, \{T < -u_\alpha\}$$

§ 5.2.4 0-1分布中未知参数p的检验

例5.2.11. 某厂产品不合格率通常为0.5.厂方希望知道原料产地的改变是否对产品的质量发生显著的影响。现在随机地从原料产地改变后的产品中抽取了80个样本进行检验，发现有5个是不合格品。试问，在显著性水平0.1下，厂方由此可以得出什么结论?

§5.2.4 0-1分布中未知参数p的检验

例5.2.11. 某厂产品不合格率通常为0.5.厂方希望知道原料产地的改变是否对产品的质量发生显著的影响。现在随机地从原料产地改变后的产品中抽取了80个样本进行检验，发现有5个是不合格品。试问，在显著性水平0.1下，厂方由此可以得出什么结论?

解：总体$X \sim B(1,p)$,其中p未知。在显著性水平$\alpha = 0.1$下，产品质量无变化等价于$p = 0.05$,故我们要检验

$$H_0 : p = 0.05 \leftrightarrow H_1 : p \neq 0.05$$

由于$\bar{x} = 5/80 = 0.0625$,因此检验统计量T的观测值

$$t = \sqrt{n}\frac{\bar{x} - p_0}{\sqrt{p_0(1-p_0)}} = 0.513$$

由$\alpha = 0.10$得临界值$u_0.05 = 1.645$.易见，$|t| < 1.645$,因此不能拒绝H_0,即在近似显著性水平0.01下可以认为原料产地得改变对该厂产品的质量没有发生显著的影响。

§5.2.4 0-1分布中未知参数p的检验

例5.2.12. 某厂产品的优质品率一直保持在40%，近期技监部门来厂抽查，共抽查了12件产品，其中优质品为5件，在$\alpha = 0.05$水平上能否认为其优质品率仍保持在40%?

§ 5.2.4 0-1分布中未知参数p的检验

例5.2.12. 某厂产品的优质品率一直保持在40%，近期技监部门来厂抽查，共抽查了12件产品，其中优质品为5件，在$\alpha = 0.05$水平上能否认为其优质品率仍保持在40%?

解：以X记检查一个产品时优质品的个数，则$X \sim b(1, p)$.检验问题为：

$$H_0 : p = 0.4 \leftrightarrow H_1 : p \neq 0.4$$

拒绝域为$\{T \leq c_1 \ or \ T \geq c_2\}, \quad c_1 < c_2$

其中$T = \sum_{i=1}^{12} X_i$. c_1与c_2应满足：当$p_0 = 0.4$时

$$使\ p(T \leq c_1) \leq \frac{\alpha}{2}\ 成立的最大整数;$$

$$使\ p(T \geq c_2) \leq \frac{\alpha}{2}\ 成立的最小整数.$$

§ 5.2.4 0-1分布中未知参数p的检验

现用二项分布来求临界值，这里$T \sim b(12, 0.4)$.在$n = 12, p = 0.4$时，T取不同值的概率如下：

k	0	1	2	...	8	9
$p(T = k)$	0.0022	0.0174	0.0639	...	0.0420	0.012457

现取$\alpha = 0.05$,由于

$p(T \leq 1) = 0.0196 < 0.025, p(T \leq 2) = 0.0835 > 0.025$,故取$c_1 = 1$，又由于

$p(T \geq 8) = 0.0573 > 0.025, p(T \geq 9) = 0.0153 < 0.025$,故取$c_2 = 9$.从而拒绝域为

$$\{T \leq 1\text{或}T \geq 9\}$$

现$T = 5$,未落入拒绝域，因而在$\alpha = 0.05$水平上认为该厂优质品率无明显变化。

§5.2.4 0-1分布中未知参数p的检验

例5.2.13. 如果在上一例中抽检50件产品，那么该检验问题的拒绝域是什么？

§5.2.4 0-1分布中未知参数p的检验

例5.2.13. 如果在上一例中抽检50件产品，那么该检验问题的拒绝域是什么？

解：这里样本容量较大，可用正态近似方法寻求临界值c_1与c_2，由$n = 50, p_0 = 0.4, \alpha = 0.05$时，$u_{0.975} = 1.96$可知

$$np_0 - 0.5 - u_0.975\sqrt{np_0(1-p_0)} = 12.71$$
$$np_0 + 0.5 + u_0.975\sqrt{np_0(1-p_0)} = 27.29$$

故取$c_1 = 12, c_2 = 28$,则水平未0.05的检验的拒绝域为

$$\{T \leq 12\text{或}T \geq 28\}$$

§ 5.3 拟合优度检验

前面的假设检验基本上是在假定总体是正态的条件下做的，但是这个假设不一定成立，需要收集样本$(X_1, \cdots, X_n)$来检验它。一般地，检验

$$H_0: X\text{服从某种分布}$$

可以采用$Karl\ Pearson$提出的χ^2拟合优度检验。

§5.3.1 离散总体情形

(1)理论分布不含未知参数的情形

设某总体X服从一个离散分布，且根据经验得知总体落在类别$a_1, \cdots, a_k$的理论频率分别为$p_1, \cdots, p_k$,现从该总体抽得一个样本量为n得样本，其落在类别$a_1, \cdots, a_k$的观测数分别为$n_1, \cdots, n_k$.感兴趣的问题是检验理论频率是否正确，即下面假设是否正确：

$$H_0 : P(X \in a_1) = p_1, \cdots, X \in a_k) = p_k$$

这类问题只提零假设而不提对立假设，相应的检验方法称为拟合优度检验。显然，在零假设下，各类别的理论频数分别为$np_1, \cdots, np_k$,将理论频数和观测频数列于下表：

类别	a_1	a_2	$\cdots$	a_k
理论频数	np_1	np_2	$\cdots$	np_k
观测频数	n_1	n_2	$\cdots$	n_k

§ 5.3.1 离散总体情形

由大数定律知，在零假设成立时，n_i/n依概率收敛于p_i，故理论频数np_i与观测频数n_i接近。而检验统计量取为：

$$\chi^2 = \sum_{i=1}^{k} \frac{(n_i - np_i)^2}{np_i}.$$

简单地，就是

$$\chi^2 = \sum \frac{(O - E)^2}{E},$$

其中O为观测频数，E为期望频数。

这个统计量中每项的分母的选取有点讲究，我们可以这样粗略地解释：假设n_i服从$Poisson$分布，则n_i的均值和方差均为np_i,从而$(n_i - np_i)/\sqrt{np_i}$的极限分布为标准正态分布，

§ 5.3.1 离散总体情形

因此χ^2近似为k个服从自由度为1的χ^2分布的随机变量之和，由于$\sum_{i=1}^{k}(n_i - np_i) = 0$,故这$k$个随机变量满足一个约束，从而$\chi^2$的自由度为$k-1$。事实上，可以严格证明，在一定的条件下，$\chi^2$的极限分布就是自由度为$k-1$的$\chi^2$分布，但其证明超出本课程要求范围。

§ 5.3.1 离散总体情形

例5.3.1. 有人制造一个含有6个面的骰子，并声称是均匀的。现设计一个实验来检验此命题：连续投掷600次，发现出现六面的频数分别为97,104,82,110,93,114.问能否再显著性水平0.2下认为骰子是均匀的?

§ 5.3.1 离散总体情形

例5.3.1. 有人制造一个含有6个面的骰子，并声称是均匀的。现设计一个实验来检验此命题：连续投掷600次，发现出现六面的频数分别为97,104,82,110,93,114.问能否再显著性水平0.2下认为骰子是均匀的?

解：该问题设计的总体是一个有6个类别的离散总体，记出现六个面的概率分别为$p_1, \cdots, p_6$,则零假设可表示为

$$H_0: p_i = 1/6, i = 1, \cdots, 6.$$

在零假设下，理论频数都是100，故检验统计量χ^2的取值为

$$\frac{(97-100)^2}{100} + \frac{(104-100)^2}{100} + \cdots + \frac{(114-100)^2}{100} = 6.94,$$

跟自由度为$6-1=5$的χ^2分布的上0.05分位数$\chi^2_5(0.2) \approx 7.29$比较，不能拒绝零假设，即可在显著性水平0.2下认为骰子是均匀的。

§5.3.1 离散总体情形

例5.3.2. 一支香烟中的尼古丁含量X服从正态分布$N(\mu,1)$,合格标准规定μ不能超过1.5mg。为对一批香烟的尼古丁含量是否合格作判断，则可建立如下假设：

$$H_0:\mu\le 1.5 \leftrightarrow H_1:\mu>1.5$$

这是在方差已知情况下对正态分布的均值作单边检验，所用的检验统计量为

$$U=\frac{\bar{X}-1.5}{1/\sqrt{n}}$$

拒绝域是

$$W=u\ge u_{1-\alpha}$$

现随机抽取一盒(20支)香烟，测得平均每支香烟的尼古丁含量为$\bar{x}=1.97mg$,则可求得检验统计量得值为$u=2.10$.下表对四个不同得显著性水平α分别列出相应得拒绝域和所下得结论：

§5.3.1 离散总体情形

显著性水平(α)	拒绝域	$u=2.10$时得结论
0.05	$\{u \geq 1.645\}$	拒绝H_0
0.025	$\{u \geq 1.96\}$	拒绝H_0
0.01	$\{u \geq 2.33\}$	接受H_0
0.005	$\{u \geq 2.58\}$	接受H_0

从上表可看出，当α相对大一些时，U得临界值就小，从而2.10超过了临界值，故应拒绝H_0;而当α减小时，临界值便增大，2.10就可能不超过临界值，这时便接受H_0.

现在换一个角度来看这一问题。用$\mu=\mu_0=1.5$时检验统计量U得分布$N(0,1)$可求得

$$P_{\mu_0}(2.10)=0.0179$$

这里0.0179是这个问题中拒绝H_0的最小的显著性水平，比它稍大一点便会导致接受H_0，这种"拒绝H_0的最小的显著性水平"就称为p值。

§5.3.1 离散总体情形

例5.3.3. 某大公司得人事部门希望了解公司职工得病假是否均匀分布在周一到周五，以便合理安排工作。如今抽取了100名病假职工，其病假日分布如下：

工作日	周一	周二	周三	周四	周五
频数	17	27	10	28	18

试问该公司职工病假是否均匀分布在一周五个工作日中($\alpha = 0.05$)?

§5.3.1 离散总体情形

例5.3.3. 某大公司得人事部门希望了解公司职工得病假是否均匀分布在周一到周五，以便合理安排工作。如今抽取了100名病假职工，其病假日分布如下：

工作日	周一	周二	周三	周四	周五
频数	17	27	10	28	18

试问该公司职工病假是否均匀分布在一周五个工作日中($\alpha = 0.05$)?

解：若病假是均匀分布在五个工作日内，则应有$p_i = \frac{1}{5}, i = 1, 2, \cdots, 5$,以$A_i$表示"病假在周$i$"，则要检验假设

$$H_0 : P(A_i) = \frac{1}{5}, \quad i = 1, 2, \cdots, 5$$

采用统计量$\chi^2 = \sum_{i=1}^{r} \frac{(n_i - np_i)^2}{np_i}$,由于$r = 5$,在$\alpha = 0.05$时，

§ 5.3.1 离散总体情形

$\chi^2_{0.95}(4) = 9.49$, 因而拒绝域为

$$W = \{\chi^2 \geq 9.49\}$$

为计算统计量χ^2得值，可列成如下计算表格：

Table: χ^2值计算表

工作日	n_i	np_i	$(n_i - np_i)^2/(np_i)$
周一	17	20	0.45
周二	27	20	2.45
周三	10	20	5.00
周四	28	20	3.20
周五	18	20	0.20
合计	100		11.30

§ 5.3.1 离散总体情形

由上表知

$$\chi^2 = \sum_{i=1}^{5} \frac{(n_i - np_i)^2}{np_i} = 11.30 > 9.49$$

这样表明样本落在拒绝域中，因而在$\alpha = 0.05$水平上拒绝原假设H_0,认为该公司职工病假在五个工作日中不是均匀分布的。

§ 5.3.1 离散总体情形

(2)理论分布包含若干未知参数的情形

当理论总体总含有未知的参数时，理论频数np_i一般也与这些参数有关，此时应该用适当的估计如极大似然估计代替这些参数以得到p_i的估计$\hat{p}_i$，得到的统计量记为

$$\chi^2 = \sum_{i=1}^{k} \frac{(n_i - n\hat{p}_i)^2}{n\hat{p}_i}.$$

在零假设下，检验统计量的χ^2的极限分布是自由度为$k-1-r$的χ^2 分布.

§ 5.3.1 离散总体情形

例5.3.4. 从某人群中随机抽取100个人的血液，并测定他们在某基因位点处的基因型。假设该位点只有两个等位基因A 和a, 这100个基因型中AA,Aa 和aa 的个数分别为30，40，30，则能否再0.05水平下认为该群体在此位点处达到Hardly-Weinberg 平衡态？

§5.3.1 离散总体情形

例5.3.4. 从某人群中随机抽取100个人的血液，并测定他们在某基因位点处的基因型。假设该位点只有两个等位基因A 和a，这100个基因型中AA,Aa 和aa 的个数分别为30，40，30，则能否再0.05水平下认为该群体在此位点处达到Hardly-Weinberg 平衡态？

解：取零假设为

$$H_0 : Hardly - Weinberg\text{平衡态成立}$$

设人群中等位基因A的频率为p,则该人群在此位点处达到$Hardly - Weinberg$平衡态指的是在人群中3个基因型的频率分别为$P(AA) = p^2, P(Aa) = 2p(1-p), P(aa) = (1-p)^2$, 即零假设可等价地写成

$$H_0 : P(AA) = p^2, P(Aa) = 2p(1-p), P(aa) = (1-p)^2.$$

§5.3.1 离散总体情形

在H_0下，3个基因型的理论频数为

$$100 \times \hat{p}^2, 100 \times 2 \times \hat{p}(1-\hat{p}), 100 \times (1-\hat{p})^2,$$

其中$\hat{p}$等于估计的等位基因频率0.5，代入χ^2统计量表达式，得统计量的值等于4.该统计量的值大于自由度为3-1-1=1(恰好一个自由参数被估计)的χ^2 分布上0.05分为数3.84，故可在0.05的水平下认为未达到Hardly-Weinberg平衡态。

§ 5.3.1 离散总体情形

例5.3.5. 在一高速路的收费站记录了106分钟内在每一分钟内到达收费站的车辆个数。数据如下表，若用X表示每一分钟内到达收费站的车辆个数，试问X是否服从某个Poisson分布？

Table: 一分钟内达到的车辆个数

车辆个数(x_i)/每分钟	出现的次数	车辆个数(x_i)/每分钟	出现的次数
0	0	10	13
1	0	11	10
2	1	12	5
3	3	13	6
4	5	14	4
5	7	15	5
6	13	16	4
7	12	17	0
8	8	18	1
9	9		n=106

§ 5.3.1 离散总体情形

解：依题设，即要检验假设

$$H_0 : P(X = x) = \frac{\lambda^x}{x!}e^{-\lambda}, \quad x = 0, 1, \cdots$$

由极大似然估计方法，知$\hat{\lambda} = \bar{x} = 9.09$。代入到上述分布里，并计算表中取每一个$x_i$的概率，如

$$P(X = 5) = 0.058$$
$$P(X < 4) = 0.052$$

得到如下

§ 5.3.1 离散总体情形

Table: χ^2计算表

区间	ν_i	$\hat{p}_i$	$n\hat{p}_i$	$\nu_i^2/n\hat{p}_i$
$0 \le x < 5$	9	0.052	5.51	14.70
$5 \le x < 5$	7	0.058	6.5	7.97
$6 \le x < 7$	13	0.088	9.33	18.11
$7 \le x < 8$	12	0.115	12.19	11.81
$8 \le x < 9$	8	0.131	13.89	4.61
$9 \le x < 10$	9	0.132	13.99	5.79
$10 \le x < 11$	13	0.120	12.72	13.29
$11 \le x < 12$	10	0.099	10.49	9.53
$12 \le x < 13$	5	0.075	7.95	3.14
$13 \le x < 14$	6	0.054	5.72	6.29
$14 \le x$	14	0.076	8.06	24.32
	106	1.0	106	119.56

§ 5.3.1 离散总体情形

因此

$$z = \sum_{i=1}^{k} \frac{\nu_i^2}{n\hat{p}_i} - n = 119.56 - 106 = 13.56$$

由$\alpha = 0.05$和$k - r - 1 = 9$，所以$\chi^2_{0.05}(9) = 16.92$。因此$z < \chi^2_{0.05}(9) = 16.92$ 所以接受原假设H_0，即可以认为每分钟到达收费站的车辆个数服从某个Poisson分布。□

§ 5.3.1 离散总体情形

例5.3.6. 在某交叉路口记录每15秒钟内通过的汽车数量，共观察了25分钟，得100个记录，经常整理得：

通过的汽车数量	0	1	2	3	4	5	6	7	8	9	10	11
频数	1	5	15	17	26	11	9	8	3	2	2	1

在$\alpha = 0.05$水平上检验如下假设：通过该交叉路口的汽车数量服从泊松分布$P(\lambda)$.

本例中只观察到$0, 1, \cdots, 11$等12个值。这相当于把总体分成12类，每一类出现的概率分别为：

$$p_i = \frac{\lambda^i}{i!}e^{-\lambda}, \quad i = 0, 1, \cdots, 10$$

$$p_{11} = \sum_{i=11}^{\infty} \frac{\lambda^i}{i!}e^{-\lambda}$$

§ 5.3.1 离散总体情形

从而把所要检验的原假设记

$$H_0 : P(A_i) = p_i, \quad i = 0, 1, \cdots, 11$$

其中A_i表示15秒钟内通过交叉路口的汽车为i辆, $i = 0, 1, \cdots, 10$, A_{11} 表示15秒钟内通过交叉路口的汽车超过10辆。

总体分布中含有未知参数λ,当然这个λ可以用样本均值$\bar{x} = 4.28$去估计。

§ 5.3.1 离散总体情形

从而,用极大似然估计$\bar{x} = 4.28$去代替λ，我们有

$$p_i = \frac{4.28^i}{i!}e^{-4.28}, \quad i = 0, 1, \cdots, 10$$

$$p_{11} = \sum_{i=11}^{\infty} \frac{4.28^i}{i!}e^{-4.28}$$

其次，由于我们要采用检验统计量的近似分布来确定拒绝域，因而要求各n_i不能过少，通常要求$n_i \geq 5$,当某些频数小于5时，通常的做法时将临近若干组合并。在本例中，$n_0 = 1 < 5$,因而可将$i = 0, i = 1$的两组合并，同样由于$i \geq 8$时各组频数小于5。

§ 5.3.1 离散总体情形

采用Pearson-χ^2 统计量
在$\alpha = 0.05$时，$\chi^2_{0.95}(8-1-1) = \chi^2_0.95(6) = 12.592$，拒绝域为

$$W = \chi^2 \geq 12.592$$

由样本数据计算$\chi^2 = 5.7897 < 12.592$, 故在$\alpha = 0.05$ 水平上，可保留H_0, 认为15秒钟内通过交叉路口的汽车数量服从参数$\lambda = 4.28$ 的泊松分布。

§ 5.3.2 列联表的独立性检验和齐一性检验

(1)独立性检验

下面考虑很常见的列联表。列联表是一种按两个属性作双向分类的表。例如肝癌病人可以按所在医院(属性A)和是否最终死亡(属性B)分类。目的是看不同医院的疗效是否不同。又如婴儿可按喂养方式(属性A，分两个水平：母乳喂养与人工喂养)和小儿牙齿发育状况(属性B,分两个水平:正常与异常)来分类。这两个例子中两个属性都只有两个水平，相应的列联表称为"四格表"，一般地，如果第一个属性有a个水平，第二个属性有b个水平，称为$a \times b$表.

实际应用中，常见的一个问题是考察两个属性是否独立。即零假设是

$$H_0 : \text{属性}A\text{与属性}B\text{独立}.$$

这是列联表的独立性检验问题。

§ 5.3.2 列联表的独立性检验和齐一性检验

假设样本量为n，第(i,j)格的频数为n_{ij}.记$p_{ij}=P$(属性A,B分布处于水平i,j),$u_i=P$(属性A有水平i),$v_j=P$(属性B有水平j).则零假设就是$p_{ij}=u_iv_j$.将u_i和v_j看成参数，则总的独立参数有$a-1+b-1=a+b-2$个。它们的极大似然估计为

$$\hat{u}_i=\frac{n_{i\cdot}}{n},\hat{v}_i=\frac{n_{\cdot j}}{n}.$$

正好是它们的频率。其中

$$n_{i\cdot}=\sum_{j=1}^{b}n_{ij},n_{\cdot j}=\sum_{i=1}^{a}n_{ij}.$$

§ 5.3.2 列联表的独立性检验和齐一性检验

在H_0下，第(i,j)格的理论频数为$n\hat{p}_{ij}=n_{i}.n._{j}/n$,因此在$H_0$下，$\sum_{i=1}^{a}\sum_{j=1}^{b}(n_{ij}-n\hat{p}_{ij})$应该较小。故取检验统计量为

$$\chi^2=\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{(n_{ij}-n_{i}.n._{j}/n)^2}{n_{i}.n._{j}/n}.$$

在零假设下χ^2的极限分布是有自由度为$k-1-r=ab-1-(a+b-2)=(a-1)(b-1)$的$\chi^2$分布。对于四格表，自由度为1.

§5.3.2 列联表的独立性检验和齐一性检验

例5.3.7. 设随机从某指定一大群人中，调查了201名人的收入情况，结果如下表。其中A表示收入，$1, 2, 3$分别表示低、中、高收入；B表示文化支出，$1, 2$分别表示“少”和“多”。试问：文化支出多少和收入高低有关系吗?

B \ A	1	2	3	和
1	63	37	60	160
2	16	17	8	41
和	79	54	68	201

§5.3.2 列联表的独立性检验和齐一性检验

例5.3.7. 设随机从某指定一大群人中，调查了201名人的收入情况，结果如下表。其中A表示收入，$1,2,3$分别表示低、中、高收入；B表示文化支出，$1,2$分别表示“少”和“多”。试问：文化支出多少和收入高低有关系吗？

B \ A	1	2	3	和
1	63	37	60	160
2	16	17	8	41
和	79	54	68	201

解：由题设我们提出假设

$$H_0 : \text{文化支出多少和收入高低没有关系}$$

则相当于要检验属性A和B是相互独立的. 对χ^2检验统计量代入数据得到$Z_0 = 7.2078$ 查表的$\chi^2_{0.05}(2) = 5.991 < Z_0$，因此我们否定$H_0$，即认为文化支出的多少和收入的高低是有关系的。$\square$

§5.3.2 列联表的独立性检验和齐一性检验

例5.3.8. 某地调查了3000名失业人员，按性别文化程度分类如下：

性别 \ 文化程度	大专以上	中专技校	高中	初中及以下	合计
男	40	138	620	1043	1841
女	20	72	442	625	1159
合计	60	210	1062	1668	3000

试问在$\alpha = 0.05$水平上检验失业人员的性别与文化程度是否有关。

§ 5.3.2 列联表的独立性检验和齐一性检验

例5.3.8. 某地调查了3000名失业人员，按性别文化程度分类如下：

性别 \ 文化程度	大专以上	中专技校	高中	初中及以下	合计
男	40	138	620	1043	1841
女	20	72	442	625	1159
合计	60	210	1062	1668	3000

试问在$\alpha = 0.05$水平上检验失业人员的性别与文化程度是否有关。

解：这是列联表的独立性检验问题。在本例中$a = 2, b = 4$,在$\alpha = 0.05$水平下，$\chi^2_{0.95}((a-1)(b-1)) = \chi^2_{0.95}(3) = 7.815$,因而拒绝域为

$$W = \chi^2 \geq 7.815$$

§5.3.2 列联表的独立性检验和齐一性检验

统计量

$$\chi^2 = \sum_{i=1}^{a}\sum_{j=1}^{b}\frac{(n_{ij}-n\hat{p}_{i\cdot}\hat{p}_{\cdot j})^2}{n\hat{p}_{i\cdot}\hat{p}_{\cdot j}}$$

从而由样本计算

$$\begin{aligned}\chi^2 &= \frac{(40-36.8)^2}{36.8}+\frac{(20-23.2)^2}{23.2}+\cdots\\ &\quad+\frac{(1043-1023.6)^2}{1023.6}+\frac{(625-644.4)^2}{644.4}=7.236\end{aligned}$$

由于$\chi^2 = 7.326 < 7.815$,样本落入接受域，从而在$\alpha = 0.05$水平上可认为失业人员性别与文化程度无关。

§5.3.2 列联表的独立性检验和齐一性检验

例5.3.9. 目前有的零售商店开展上门服务的业务，有的不开展此项业务。为了解这项业务的开展与否与其月销售额是否有关，调查了363个商店，结果如下，结果如下：

月销售额 / 服务方式	≤ 1000	(1000,5000]	(5000,10000]	(10000,20000]	> 2000	合计
上门服务	32	111	104	40	14	301
不上门服务	29	24	6	2	1	62
合计	61	135	110	42	15	363

试问在$\alpha = 0.01$水平上检验服务方式与月销售额是否有关。

§ 5.3.2 列联表的独立性检验和齐一性检验

例5.3.9. 目前有的零售商店开展上门服务的业务，有的不开展此项业务。为了解这项业务的开展与否与其月销售额是否有关，调查了363个商店，结果如下，结果如下：

月销售额 / 服务方式	≤ 1000	(1000,5000]	(5000,10000]	(10000,20000]	> 2000	合计
上门服务	32	111	104	40	14	301
不上门服务	29	24	6	2	1	62
合计	61	135	110	42	15	363

试问在$\alpha = 0.01$水平上检验服务方式与月销售额是否有关。

解：这也是列联表的独立性检验问题，在本例中
中$a = 2, b = 5$,在$\alpha = 0.01$水平
下，$\chi^2_{0.99}((a-1)(b-1)) = \chi^2_{0.95}(4) = 13.277$,因而拒绝域为

$$W = \chi^2 \geq 13.277$$

§ 5.3.2 列联表的独立性检验和齐一性检验

为计算统计量，先在独立性假设下求出各$n\hat{p}_{i\cdot}\hat{p}_{\cdot j} = \frac{n_{i\cdot}n_{\cdot i}}{n}$如下：

$n\hat{p}_{i\cdot}\hat{p}_{\cdot j}$	≤ 1000	(1000,5000]	(5000,10000]	(10000,20000]	> 2000	合计
上门服务	50.6	111.9	91.2	34.8	12.4	301
不上门服务	10.4	23.1	18.8	7.2	2.6	62
合计	61	135	110	42	15	363

由此计算出$\chi^2 = 56.38 > 13.277$,样本落在拒绝域中，这说明是否开展上门服务这项业务与月销售额有关。

§5.3.2 列联表的独立性检验和齐一性检验

(2) 齐一性检验

跟列联表有关的另一类重要的检验是齐一性检验，即检验某一个属性A的各个水平对应的另一个属性B的分布全部相同。这种检验跟独立性检验有着本质的区别。独立性问题中两属性都是随机的；而齐一性问题中属性A是非随机的，这样涉及到的分布实际上是条件分布。虽然如此，所采用的检验方法跟独立性检验完全一样。

§ 5.3.2 列联表的独立性检验和齐一性检验

例5.3.10. 下表是甲乙两医院肝癌病人生存情况。需要根据这些数据判断两医院的治疗效果是否一样。(下表是甲乙两院肝癌的近期疗效)

	生存	死亡	合计
甲院	150(n_{11})	88(n_{12})	238($n_{1\cdot}$)
乙院	36(n_{21})	18(n_{22})	54($n_{2\cdot}$)
合计	186($n_{\cdot 1}$)	106($n_{\cdot 2}$)	292(n)

§5.3.2 列联表的独立性检验和齐一性检验

例5.3.10. 下表是甲乙两医院肝癌病人生存情况。需要根据这些数据判断两医院的治疗效果是否一样。(下表是甲乙两院肝癌的近期疗效)

	生存	死亡	合计
甲院	150(n_{11})	88(n_{12})	238($n_{1\cdot}$)
乙院	36(n_{21})	18(n_{22})	54($n_{2\cdot}$)
合计	186($n_{\cdot 1}$)	106($n_{\cdot 2}$)	292(n)

解：这是一个齐一性检验问题。检验统计量χ^2的观测值为0.1195，远远小于自由度为1的χ^2分布的上0.05分位数，故可以接受零假设，即在水平0.05下可以认为两个医院的疗效无差别。

当有某个格子的频数较小时，如果允许的话可以合并格子使每个格子的频数足够大。

§5.3.3 连续总体情形

设$(X_1,\cdots,X_n)$是取自总体X的一个样本，记X的分布函数为$F(x)$,需要检验的那种分布中含有r个总体参数$\theta_1,\cdots,\theta_r$.我们要在显著性水平α下检验

$$H_0: F(x)=F_0(x;\theta_1,\cdots,\theta_r),$$

其中$F_0(x;\theta_1,\cdots,\theta_r)$表示需要检验的那种分布的分布函数。例如，当我们要检验

$$H_0: X\sim(\mu,\sigma^2)$$

时，$r=2,\theta_1=\mu,\theta_2=\sigma^2$.

$$F_0(x;\mu,\sigma^2)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi\sigma^2}}exp\{-\frac{1}{2\sigma^2}(t-\mu)^2\}dt$$

上述假设可以通过适当的离散化总体分布，采用拟合优度法来作检验。首先把实数轴分成k个子区间$(a_{j-1},a_j],j=1,\cdots,k,$其中$a_0$可以取$-\infty$,$a_k$可以取$\infty$.这样构造了一个离散总体，其取值就是这k个区间。

§5.3.3 连续总体情形

记

$$
\begin{aligned}
p_j &= P_{H_0}(a_{j-1} < X \leq a_j) \\
&= F_0(a_j;\theta_1,\cdots,\theta_r) - F_0(a_{j-1};\theta_1,\cdots,\theta_r), j = 1,\cdots,k
\end{aligned}
$$

如果H_0成立，则概率p_j应该与数据落在区间$(a_{j-1}, a_j]$的频率$f_j = n_j/n$接近，其中n_j表示相应的频数。当p_i的取值不含未知参数时，取检验统计量

$$
\chi^2 = \sum_{j=1}^{k} \frac{(n_j - np_j)^2}{np_j},
$$

否则取

$$
\chi^2 = \sum_{j=1}^{k} \frac{(n_j - n\hat{p}_j)^2}{n\hat{p}_j},
$$

其中$\hat{p}_i$是p_i中的未知参数换成适当的估计后得到的p_i的估计.

§ 5.3.3 连续总体情形

拒绝域取为

$$\{\chi^2 > \chi^2_{k-r-1}(\alpha)\}$$

如果p_i中不含未知参数，则r=0.

使用χ^2进行拟合优度检验时一般要求

$$n > 50, n\hat{p}_j \geq 5, j = 1, \cdots, k,$$

如果不满足这个条件，最好把某些组作适当合并。

§5.3.3 连续总体情形

例5.3.11. 为研究混凝土抗压强度的分布，抽取了200件混凝土制件测定其抗压强度，经整理得频数分布表如下：

抗压强度区间$(a_{i-1}, a_i]$	频数n_i
$(190, 200]$	10
$(200, 210]$	26
$(210, 220]$	56
$(220, 230]$	64
$(230, 240]$	30
$(240, 250]$	14
合计	200

试问在$\alpha = 0.05$水平上检验抗压强度的分布是否为正态分布。

§5.3.3 连续总体情形

解：若用$F(x)$表示$N(\mu,\sigma^2)$的分布函数，则本例便要检验假设：

$$H_0: \text{抗压强度的分布为} F(x)$$

由于$F(x)$中有两个未知参数μ,σ^2,因而需要用它们的极大似然估计去替代。这里仅给出了样本的分组数据，因此只能用组中值(即区间中点)去代替替原始数据，然后求μ,σ^2得MLE。现在6个组中值分别为

$x_1=195, x_2=205, x_3=215, x_4=225, x_5=235, x_6=245$,于是

$$\hat{\mu}=\bar{x}=\frac{1}{200}\sum_{i=1}^{6}n_i x_i=221$$

$$\hat{\sigma^2}=s_n^2=\frac{1}{200}\sum_{i=1}^{6}n_i(x_i-\bar{x})^2=152, \quad \hat{\sigma}=s_n=12.33$$

§5.3.3 连续总体情形

在$N(221,152)$分布下，求出落在区间$(a_{i-1},a_i]$内的概率的估计值：

$$\hat{p}_i=\Phi\left(\frac{a_i-221}{\sqrt{152}}\right)-\Phi\left(\frac{a_{i-1}-221}{\sqrt{152}}\right),\quad i=1,2,\cdots,r$$

不过常将a_0定为$-\infty$，将a_r定为$+\infty$.本例中$r=6$.采用检验统计量

$$\chi^2=\sum_{i=1}^{r}\frac{(n_i-n\hat{p}_i)^2}{n\hat{p}_i}$$

在$\alpha=0.05$时，$\chi^2_{0.95}(6-2-1)=7.815$,因而拒绝域为

$$W=\chi^2\geq 7.815$$

由样本计算$\chi^2=1.332<7.815$,这表明样本落入接受域，可接受抗压强度服从正态分布的假定。

总结：

1.理解假设检验的一些基本概念：零假设、对立假设、两类错误、拒绝域、显著性水平、功效.

2.学会将实际问题转化为假设检验问题来处理.

3.一样本和两样本正态总体均值和方差的假设检验.

4.$0-1$分布参数的假设检验.

5.拟合优度检验、列联表的独立性和齐一性检验.

表 6.2.1 一样本正态总体$N(\mu, \sigma^2)$.

检验对象	检验统计量	分布	拒绝域†
μ (σ^2已知)	$U = \sqrt{n}(\bar{X} - \mu_0)/\sigma$	$N(0,1)$	$\begin{cases} \lvert U\rvert > u_{\alpha/2} \\ U > u_\alpha \\ U < -u_\alpha \end{cases}$
μ (σ^2 未知)	$T = \sqrt{n}(\bar{X} - \mu_0)/S$	t_{n-1}	$\begin{cases} \lvert T\rvert > t_{n-1}(\alpha/2) \\ T > t_{n-1}(\alpha) \\ T < -t_{n-1}(\alpha) \end{cases}$
σ^2 (μ 已知)	$\chi^2 = \frac{1}{\sigma_0^2}\sum_{i=1}^n (X_i - \mu)^2$	χ_n^2	$\begin{cases} \chi^2 > \chi_n^2(\alpha/2)\text{或者}\chi^2 < \chi_n^2(1-\alpha/2) \\ \chi^2 > \chi_n^2(\alpha) \\ \chi^2 < \chi_n^2(1-\alpha) \end{cases}$
σ^2 (μ 未知)	$\chi^2 = \frac{1}{\sigma_0^2}\sum_{i=1}^n (X_i - \bar{X})^2$	χ_{n-1}^2	$\begin{cases} \chi^2 > \chi_{n-1}^2(\alpha/2)\text{或者}\chi^2 < \chi_{n-1}^2(1-\alpha/2) \\ \chi^2 > \chi_{n-1}^2(\alpha) \\ \chi^2 < \chi_{n-1}^2(1-\alpha) \end{cases}$

有关均值的检验：对立假设分别为$\mu \neq \mu_0, \mu > \mu_0$和$\mu < \mu_0$.
有关方差的检验：对立假设分别为$\sigma^2 \neq \sigma_0^2$,$\sigma^2 > \sigma_0^2$和$\sigma^2 < \sigma_0^2$.

表 6.2.2 两样本正态总体的假设检验

检验对象	检验统计量	分布	拒绝域[†]
均值(方差已知)	$U=\frac{\bar{X}-\bar{Y}}{\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}}$	$N(0,1)$	$\begin{cases}\lvert U\rvert>u(\alpha/2)\\ U>u(\alpha)\\ U<-u(\alpha)\end{cases}$
均值(方差未知)[‡]	$T=\frac{\bar{X}-\bar{Y}}{S_w\sqrt{\frac{1}{m}+\frac{1}{n}}}$	t_{m+n-2}	$\begin{cases}\lvert T\rvert>t_{m+n-2}(\alpha/2)\\ T>t_{m+n-2}(\alpha)\\ T<-t_{m+n-2}(\alpha)\end{cases}$
方差(均值已知)	$F=\frac{\sum_{i=1}^m(X_i-\mu_1)^2/m}{\sum_{i=1}^n(X_i-\mu_2)^2/n}$	$F_{m,n}$	$\begin{cases}F>F_{m,n}(\alpha/2)\text{或}F<\frac{1}{F_{n,m}(\alpha/2)}\\ F>F_{m,n}(\alpha)\\ F<\frac{1}{F_{n,m}(\alpha)}\end{cases}$
方差(均值未知)	$F=\frac{S_1^2}{S_2^2}$	$F_{m-1,n-1}$	$\begin{cases}F>F_{m-1,n-1}(\alpha/2)\text{或}F<\frac{1}{F_{n-1,m-1}(\alpha/2)}\\ F>F_{m-1,n-1}(\alpha)\\ F<\frac{1}{F_{n-1,m-1}(\alpha)}\end{cases}$

有关均值的检验：对立假设分别为$\mu_1\neq\mu_2,\mu_1>\mu_2$和$\mu_1<\mu_2$.
有关方差的检验：对立假设分别为$\sigma_1^2\neq\sigma_2^2$,$\sigma_1^2>\sigma_2^2$和$\sigma_1^2<\sigma_2^2$.