STAT1603 Lecture Notes
Created by He Entong

1 Convolution

We are interested in the *p.d.f.* of the sum of two independent variables X, Y. That is, the *p.d.f.* of $Z = X + Y$. We first consider the *c.d.f.* of Z.

$$F_Z(z) = F_Z(x+y) = \iint\limits_D f_{X,Y}(x,y)\ dxdy \tag{1}$$

where D is the domain of integration. Applying the variable substitution

$$\begin{cases} x = z - y \\ y = y \end{cases} \tag{2}$$

We obtain

$$\begin{aligned} F_Z(z) &= \iint\limits_{D'} f_{X,Y}(z-y,y)\ \mathbb{J}(x,y)dydz \\ &= \iint\limits_{D'} f_{X,Y}(z-y,y)\ \begin{vmatrix} \frac{\partial(z-y)}{\partial y} & \frac{\partial(z-y)}{\partial z} \\ \frac{\partial y}{\partial y} & \frac{\partial y}{\partial z} \end{vmatrix} dydz \\ &= \iint\limits_{D'} f_{X,Y}(z-y,y)\ dydz \\ &= \int_{D_Z} dz\ \int_{D_Y} f_{X,Y}(z-y,y)\ dy \end{aligned} \tag{3}$$

Then we have

$$\begin{aligned} f_Z(z) &= \frac{d}{dz}F_Z(z) \\ &= \frac{d}{dz}\int_{D_Z} dz\ \int_{D_Y} f_{X,Y}(z-y,y)\ dy \\ &= \int_{D_Y} f_{X,Y}(z-y,y)\ dy \end{aligned} \tag{4}$$

Now we may well generalize the *p.d.f.* of any random variable which is a definite function of two arbitrary random variables, that is, $Z = k(x,y)$

$$F_Z(z) = F_Z(k(x,y)) = \iint\limits_D f_{X,Y}(x,y)\ dxdy \tag{5}$$

The variable substitution is

$$\begin{cases} x = g(y,z), \quad g(y,z) \sim k(x,y) \\ y = y \end{cases} \tag{6}$$

Then we obtain

$$\begin{aligned} F_Z(z) &= \iint\limits_{D'} f_{X,Y}(g(y,z),y)\ \mathbb{J}_{g,y}(y,z)dydz \\ &= \iint\limits_{D'} f_{X,Y}(g(y,z),y)\ \begin{vmatrix} \frac{\partial g}{\partial y} & \frac{\partial g}{\partial z} \\ \frac{\partial y}{\partial y} & \frac{\partial y}{\partial z} \end{vmatrix} dydz \\ &= \iint\limits_{D'} f_{X,Y}(g(y,z),y)\left|\frac{\partial g}{\partial z}\right|\ dydz \end{aligned} \tag{7}$$

Hence,

$$f_Z(z) = \frac{d}{dz}F_Z(z) = \frac{d}{dz}\iint\limits_{D'} f_{X,Y}(g(y,z),y)\left|\frac{\partial g}{\partial z}\right| \, dydz$$
$$= \frac{d}{dz}\int_{D_Z} dz \int_{D_Y} f_{X,Y}(g(y,z),y)\left|\frac{\partial g}{\partial z}\right| \, dy \tag{8}$$
$$= \boxed{\int_{D_Y} f_{X,Y}(g(y,z),y)\left|\frac{\partial g}{\partial z}\right| \, dy}$$

2 Distribution of Summation of Normally Distributed Random Variables

Assume variables X_1, X_2 are both normally distributed, that gives

$$f_{X_1}(x) = \frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}, \; f_{X_2}(x) = \frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{(x-\mu_2)^2}{2\sigma_2^2}} \tag{9}$$

Define a new summation variable $W = X_1 + X_2$, applying the convolution we obtain

$$f_W(w) = \int_{-\infty}^{+\infty} f_{X_1,X_2}(w-x,x) \; dx$$
$$= \int_{-\infty}^{+\infty} f_{X_1}(w-x)f_{X_2}(x) \; dx$$
$$= \frac{1}{2\pi\sigma_1\sigma_2}\int_{-\infty}^{+\infty} e^{-\frac{(w-x-\mu_1)^2}{2\sigma_1^2}} e^{-\frac{(x-\mu_2)^2}{2\sigma_2^2}} \; dx$$
$$= \frac{1}{2\pi\sigma_1\sigma_2}exp\left\{-\frac{\sigma_1^2+\sigma_2^2}{2\sigma_1^2\sigma_2^2}\left[\frac{\sigma_1\sigma_2(w-(\mu_1+\mu_2))}{\sigma_1^2+\sigma_2^2}\right]^2\right\}\int_{-\infty}^{+\infty} exp\left\{-\frac{\sigma_1^2+\sigma_2^2}{2\sigma_1^2\sigma_2^2}\left[x+\frac{(\mu_1-w)\sigma_2^2-\mu_2\sigma_1^2}{\sigma_1^2+\sigma_2^2}\right]^2\right\} dx$$
$$= \frac{1}{\sqrt{2\pi(\sigma_1^2+\sigma_2^2)}}exp\left\{-\frac{1}{2}\frac{[w-(\mu_1+\mu_2)]^2}{\sigma_1^2+\sigma_2^2}\right\}$$

Hence we obtain that the sum of two normally distributed variables is also a normal variable, with the mean $\mu = \mu_1 + \mu_2$, and the variance $\sigma^2 = {\sigma_1}^2 + {\sigma_2}^2$

3 χ^2-Distribution, F Distribution and Student t-Distribution

3.1 χ^2-Distribution

If a random variable, denoted by X, follows a standard normal distribution, that is

$$X \sim N(0,1), \quad f_X(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2} \tag{10}$$

We are interested in the distribution of variable X^2. We have:

$$F(X^2 \le x) = F(X \le \sqrt{x})$$
$$= \int_{-\infty}^{\sqrt{x}} \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}$$
$$\text{Hence,} \quad f_{X^2}(x) = \frac{d}{dx}\int_{-\infty}^{\sqrt{x}} \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2} \tag{11}$$
$$= \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x}\frac{1}{2\sqrt{x}}$$
$$= \frac{e^{-\frac{1}{2}x}x^{-\frac{1}{2}}(\frac{1}{2})^{\frac{1}{2}}}{\Gamma(\frac{1}{2})}$$

Hence we conclude that the variable X^2 follows a gamma distribution, that is,

$$X^2 \sim \mathrm{Ga}(\frac{1}{2}, \frac{1}{2}) \tag{12}$$

Now we are interested in the summation of *i.i.d.* random variables $Z = X_1^2 + X_2^2 + \cdots + X_n^2$. We may first consider the summation of two variables with gamma distribution. That is

$$X_1^2 \sim \mathrm{Ga}(\beta, \alpha_1), \quad X_2^2 \sim \mathrm{Ga}(\beta, \alpha_2) \tag{13}$$

Then the *p.d.f.* of the summation variable $Z = X_1^2 + X_2^2$ is

$$\begin{aligned} f_Z(z) &= \int\limits_{x\in D} \frac{\beta^{\alpha_1}(z-x)^{\alpha_1-1}e^{-\beta(z-x)}}{\Gamma(\alpha_1)} \frac{\beta^{\alpha_2}x^{\alpha_2-1}e^{-\beta_2 x}}{\Gamma(\alpha_2)} \,\mathrm{d}x \\ &= \frac{\beta^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1)\Gamma(\alpha_2)} e^{-\beta z} \int\limits_{x\in D} (z-x)^{\alpha_1-1}x^{\alpha_2-1}\,\mathrm{d}x \\ &= z^{\alpha_1+\alpha_2-1}\frac{\beta^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1)\Gamma(\alpha_2)} e^{-\beta z}\ \mathrm{B}(\alpha_1, \alpha_2) \\ &= \frac{\beta^{\alpha_1+\alpha_2}z^{\alpha_1+\alpha_2-1}e^{-\beta z}}{\Gamma(\alpha_1+\alpha_2)} \\ \text{Hence,}\ \ & Z \sim \mathrm{Ga}(\beta, \alpha_1+\alpha_2) \end{aligned} \tag{14}$$

Using the feature of gamma distribution repeatedly, we obtain random variable Z, where $Z = X_1^2 + X_2^2 + \cdots X_n^2$ where $X_i^2 \sim Ga(\frac{1}{2}, \frac{1}{2})$ follows gamma distribution. That is,

$$Z \sim \mathrm{Ga}(\frac{1}{2}, \frac{n}{2}) \equiv \chi^2_{(n)}, \quad \boxed{f_Z(z) = \frac{z^{\frac{n}{2}}e^{-\frac{1}{2}z}}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})}} \tag{15}$$

We call the distribution a χ^2 distribution with n degrees of freedom.

3.2 F Distribution

F Distribution focuses on the distribution of a fraction of two random variables, both follow χ^2-distribution. We give

$$F = \frac{\frac{X_1}{m}}{\frac{X_2}{n}}, \quad X_1 \sim \chi^2_{(m)}, \quad X_2 \sim \chi^2_{(n)} \tag{16}$$

Then we are interested in the *p.d.f.* of random variable F, applying convolution we obtain

$$\begin{aligned} f_Z(x) &= \int_0^{+\infty} f_{X_1}(\frac{m}{n}xt) f_{X_2}(t) \frac{\partial}{\partial z}\left(\frac{m}{n}zt\right) \mathrm{d}t \\ &= \int_0^{+\infty} \frac{e^{-\frac{m}{2n}xt}(\frac{m}{n}xt)^{\frac{m}{2}-1}}{2^{\frac{m}{2}}\Gamma(\frac{m}{2})} \frac{e^{-\frac{1}{2}t}t^{\frac{n}{2}-1}}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} \frac{\partial}{\partial z}\left(\frac{m}{n}zt\right) \mathrm{d}t \\ &= \frac{m}{n} \frac{(\frac{m}{n}x)^{\frac{m}{2}-1}}{2^{\frac{m+n}{2}}\Gamma(\frac{m}{2})\Gamma(\frac{n}{2})} \int_0^{+\infty} e^{-\frac{m}{2n}xt}e^{-\frac{1}{2}t}t^{\frac{m+n}{2}-1}\,\mathrm{d}t \\ &= \boxed{\left(\frac{m}{n}\right)^{\frac{m}{2}} x^{\frac{m}{2}-1}\left(\frac{m}{n}x+1\right)^{-\frac{m+n}{2}} \frac{1}{\mathrm{B}(\frac{m}{2}, \frac{n}{2})}} \end{aligned} \tag{17}$$

3.3 Student t-Distribution

Given two random variable X_1, X_2, following standard normal distribution and χ^2 distribution respectively. Then we define the distribution a new random variable $T = \frac{X_1}{\sqrt{\frac{X_2}{n}}}$ is following as the t-distribution. The following will derive the *p.d.f.* of variable T.

$$
\begin{aligned}
X_1 \sim \mathrm{N}(0,1),\ \ f_{X_1}(x) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \\
X_2 \sim \chi^2(n),\ \ f_{X_2}(x) &= \frac{x^{\frac{n}{2}} e^{-\frac{1}{2}x}}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})}
\end{aligned}
\tag{18}
$$

$$
\begin{aligned}
f_T(t) &= \int_{x\in D} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x}{2n}t^2)} \frac{x^{\frac{n}{2}-1} e^{-\frac{1}{2}x}}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} \frac{\partial}{\partial t}\left(t\sqrt{\frac{x}{n}}\right) \mathrm{d}x \\
&= \int_{x\in D} \frac{1}{\sqrt{2\pi n}} e^{-\frac{1}{2}(\frac{x}{2n}t^2)} \frac{x^{\frac{n-1}{2}} e^{-\frac{1}{2}x}}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} \mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi n}} \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} \int_{-\infty}^{+\infty} x^{\frac{n-1}{2}} e^{-\frac{t^2+n}{2n}x} \mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi n}} \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} \Gamma\left(\frac{n+1}{2}\right)\left(\frac{2n}{t^2+n}\right)^{\frac{n+1}{2}} \\
&= \boxed{\frac{\Gamma(\frac{n+1}{2})}{\sqrt{\pi n}\,\Gamma(\frac{n}{2})}\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}}}
\end{aligned}
\tag{19}
$$

This is the *p.d.f.* of a t-distribution of n degrees of freedom. It is effective in t-Test while analyzing a small number of samples.

4 χ^2-Test

χ^2 test is used to determine the difference between the hypothesis model and the sample. It is usually applied on a $m \times n$ label, with each data denoted as f_{ij}. If under the hypothesis H_0, each data should be $f_{\text{hypo }ij}$, then we build up a variable

$$
\chi^2 = \sum_{i=1}^{m}\sum_{j=1}^{n} \frac{(X_{\text{data }ij} - X_{\text{hypo }ij})^2}{X_{\text{hypo }ij}} \sim \chi^2_{(m-1)(n-1)} \tag{20}
$$

This follows a χ^2 distribution with degree of freedom $(m-1)(n-1)$. Then we substitute the data in to get the test value of $\chi^2_{\text{T.S.}}$. And compare it with the correspond critical value given by $\chi^2_{(m-1)(n-1)}\alpha$.

第 7 章 假设检验

统计推断的另一个主要内容是统计假设检验. 在这一章里我们将讨论统计假设的设立及其检验问题.

7.1 假设检验的基本思想与概念

7.1.1 假设检验问题

我们从一个例子开始引出假设检验问题.

例 7.1.1: 某厂生产的合金强度服从正态分布 $N(\theta, 16)$,其中 θ 的设计值为不低于 110(Pa) . 为保证质量,该厂每天都要对生产情况做例行检查,以判断生产是否正常进行,即该合金的平均强度不低于 110(Pa) . 某天从生产中随机抽取 25 块合金,测得强度值为 $x_1, x_2, \cdots, x_{25}$ 其均值为 $\bar{x} = 108$(Pa) ,问当日生产是否正常?

对这个实际问题可作如下分析;

1. 这不是一个参数估计问题.
2. 这是在给定总体与样本下,要求对命题"合金平均强度不低于 110 Pa"作出回答:"是"还是"否"? 这类问题称为统计假设检验问题,简称假设检验问题.
3. 命题:"合金平均强度不低于 110 Pa"正确与否仅涉及参数 0,因此该命题是否正确将涉及如下两个参数集合:

$$\theta_0 = |\theta; \theta \geqslant 110|, \theta_1 = |\theta : \theta < 110|$$

   命题成立对应于"$\theta \in \theta_0$",命题不成立则对应"$\theta \in \theta_1$". 在统计学中这两个非空参数集合都称作统计假设,简称假设.
4. 我们的任务是利用所给总体 $N(\theta, 16)$ 和样本均值 $\bar{x} = 108$(Pa) 去判断假设 (命题) "$\theta \in \theta_0$". 是否成立,这里的"判断"在统计学中称为检验或检验法则.
   检验结果有两种:

   "假设不正确"——称为拒绝该假设;
   "假设正确"——称为接收该假设.

5. 若假设可用一个参数的集合表示,该假设检验问题称为参数假设检验问题,否则称为非参数假设检验问题,例 7.1.1 就是一个参数假设检验问题,而对假设"总体为正态分布"作出检验的问题就是一个非参数假设检验问题. 本章前三节讲述参数假设检验问题,最后一节 (7.4) 将讨论非参数假设检验问题.

7.1.2 假设检验的基本步骤

接下来我们来叙述假设检验的基本步骤.

一、建立假设

在假设检验中,常把一个被检验的假设称为原假设,用 H_0 表示,通常将不应轻易加以否定的假设作为原假设. 当 H_0 被拒绝时而接收的假设称为备择假设,用 H_1 表示,它们常常成对出现. 在

例 7.1.1 中,我们可建立如下两个假设:

$$H_0:\theta\in\Theta_0=\{\theta:\theta\geqslant 110| \quad \text{vs} \quad H_1;\theta\in\Theta_1=\{\theta:\theta<110\}$$

或简写为

$$H_0:\theta\geqslant 110 \quad \text{vs} \quad H_1:\theta<110$$

其中"vs"是 versus 的缩写,是"对"的意思,即表示 H_0 对 H_1 的假设检验问题.

二、选择检验统计量,给出拒绝域形式

由样本对原假设进行判断总是通过一个统计量完成的, 该统计量称为检验统计量. 比如, 在例 7.1.1中,样本均值 $\bar{x}$ 就是一个很好的检验统计量,因为要检验的假设是正态总体的均值,在方差已知场合,样本均值 x 是总体均值的充分统计量、使原假设被拒绝的样本观测值所在区域称为拒绝域,一般它是样本空间的一个子集,并用 W 表示,在例 7.1.1中,样本均值 $\bar{x}$ 愈大,意味着总体均值 θ 也大,样本均值 $\bar{x}$ 愈小,意味着总体均值 θ 也小,因此,在样本均值的取值中有一个临界值 c (待定),所以拒绝域为

$$\{W=|(x_1,\cdots,x_n);\bar{x}\leqslant c\}=\{\bar{x}\leqslant c\}$$

是合理的.

当拒绝域确定了,检验的判断准则跟着也确定了:

- 如果 $(x_1,\cdots,x_n)\in W$,则认为 H_0 不成立;
- 如果 $(x_1,\cdots,x_n)\in\overline{W}$,认为 H_0 成立;

一般将 $\overline{W}$ 称为接收域. 由此可见,一个拒绝域 W 唯一确定一个检验法则,反之,一个检验法则也唯一确定一个拒绝续. 在两个观测值 $n=2$ 场合,图 7.1.1 给出拒绝域的示意图.

通常我们将注意力放在拒绝域上. 正如在数学上我们不能用一个例子去证

明一个结论一样,用一个样本(例子)不能证明一个命题(假设)是成立的,但可以用一个例子(样本)推翻一个命题. 因此,从逻辑上看,注重拒绝域是适当的. 事实上,在"拒绝原假设"和"拒绝备择假设(从而接收原假设)"之间还有一个模糊域,如今我们把它并入接收域 (参见图 7.1.1),所以接收域是复杂的,将之称为保留域也许更恰当,但习惯上已把它称为接收域,没有必要再进行改变,只是应注意它的含义.

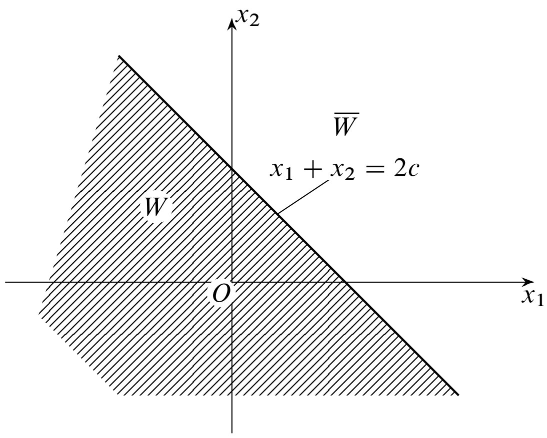


图 7.1.1: 拒绝域示意图

三、选择显着性水平

检验的结果与真实情况可能吻合也可能不吻合,因此,检验是可能犯错误的. 检验可能犯的错误有两类:其一是 H_0 为真但由于随机性使样本观测值落在拒绝域中,从而拒绝原假设 H_0,这种错

误称为第一类错误,其发生的概率称为犯第一类错误的概率,或称拒真概率,通常记为 α,即

$$\alpha = P\left(\text{拒绝}H_0 \middle| H_0\text{为真}\right) = P(X \in W), \theta \in \theta_0 \tag{7.1.1}$$

其中 $X = (x_1, \cdots, x_n)$ 表示样本. 另一种错误是 H_0 不真(即 H_1 为真)但由于随机性使样本观测值落在接受域中,从而接受原假设 H_0,这种错误称为第二类错误,其发生的概率称为犯第二类错误的概率,或称受伪概率,通常记为 β,即

$$\alpha = P\left(\text{接受}H_0 \middle| H_0\text{为真}\right) = P\left(X \in \overline{W}\right), \theta \in \Theta_0 \tag{7.1.2}$$

表7.1.1列出了检验的各种情况及两类错误.

表 7.1.1: 检验的两类错误

观察数据情况	总体情况	
	H_0 为真	H_1 为真
$(x_1, \cdots, x_n) \in W$	犯第一类错误	正确
$(x_1, \cdots, x_n) \in W^c$	正确	犯第二类错误

犯第一类错误的概率 α 和犯第二类错误的概率 β 可以用同一个函数表示,即所谓的势函数. 势函数是假设检验中最重要的概念之一,它的定义如下:

定义 7.1.1. 设检验问题

$$H_0 : \theta \in \Theta_0 \quad \text{vs} \quad H_1 : \theta \in \Theta_1$$

的拒绝域为 W,则样本观测值 X 落在拒绝域 W 内的概率称为该检验的**势函数**,记为

$$g(\theta) = P_\theta(X \in W), \quad \theta \in \boldsymbol{\Theta} = \Theta_0 \cup \Theta_1 \tag{7.1.3}$$

显然,势函数 $g(\theta)$ 是定义在参数空间 Θ 上的一个函数. 由 (7.1.1) 和 (7.1.2) 知,当 $\theta \in \Theta_0$ 时,$g(\theta) = \alpha = \alpha(\theta)$,当 $\theta \in \Theta_1$ 时,$g(\beta) = 1 - \beta = 1 - \beta(\theta)$. 由此可见,犯两类错误的概率都是参数 θ 的函数,并可由势函数得到,即:

$$g(\theta) = \begin{cases} \alpha(\theta), & \theta \in \Theta_0 \\ 1 - \beta(\theta), & \theta \in \Theta_1 \end{cases}$$

对例 7.1.1,其拒绝域为 $W = |\bar{x} \leqslant c|$,由 (7.1.3) 可以算出该检验的势函数

$$g(\theta) = P_\theta(\bar{x} \leqslant c) = P_\theta\left(\frac{\bar{x} - \theta}{4/5} \leqslant \frac{c - \theta}{4/5}\right) = \Phi\left(\frac{c - \theta}{4/5}\right)$$

这个势函数是 θ 的减函数(见图7.1.2)

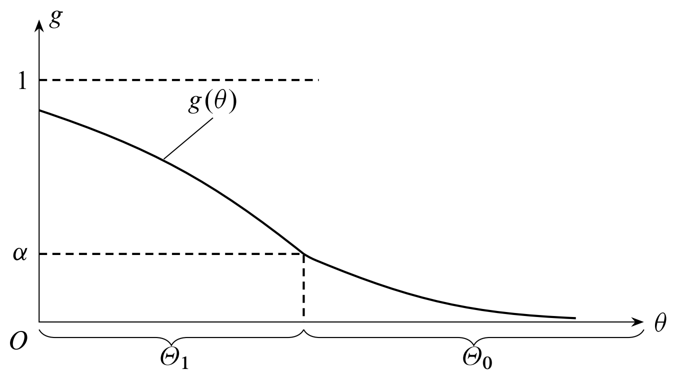


图 7.1.2: 例 7.1.1 的势函数 $g(\theta)$

利用这个势函数容易写出其犯两类错误的概率分别为

$$\alpha(\theta) = \Phi\left(\frac{c - \theta}{4/5}\right), \quad \theta \in \Theta_0 \tag{7.1.4}$$

$$\beta(\theta) = 1 - \Phi\left(\frac{c - \theta}{4/5}\right), \quad \theta \in \Theta_1 \tag{7.1.5}$$

由上述两个式子可以看出犯两类错误的概率 α, β 间的关系：

- 当 α 减小时，由 (7.1.4) 知，c 也随之减小，再由 (7.1.5) 知，c 的减小必导致 β 的增大；
- 当 β 减小时，由 (7.1.5) 知，c 会增大，再由 (7.1.4) 知，c 的增大必导致 α 的增大.

这一现象说明：在样本量给定的条件下，α 与 β 中一个减小必导致另一个增大，这不是偶然的，而具有一般性. 这进一步说明：在样本量一定的条件下不可能找到一个使 α, β 都小的检验. 在此背景下，只能采取折中方案. 英国统计学家 Neyman 和 Pearson 提出水平为 α 的显著性检验的概念.

定义 7.1.2. 对检验问题 $H_0 : \theta \in \Theta_0 \quad \text{vs} \quad H_1 : \theta \in \Theta_1$，如果一个检验满足对任意的 $\theta \in \Theta_0$，都有

$$g(\theta) \leqslant \alpha$$

则称该检验是**显着性水平**为 α 的**显着性检验**，简称水平为 α 的检验.

提出显著性检验的概念就是要控制犯第一类错误的概率 α，但也不能使得 α 过小（α 过小会导致 β 过大），在适当控制 α 中制约 β. 最常用的选择是 $\alpha = 0.05$，有时也选择 $\alpha = 0.10$ 或 $\alpha = 0.01$.

四、给出拒绝域

在确定显著性水平后，我们可以定出检验的拒绝域 W. 在例 7.1.1中，若取 $\alpha = 0.05$，则要求对任意的 $\theta \geqslant 100$ 有 $g(\theta) = \Phi(5(c-\theta)/4) \leqslant 0.05$，由于 $g(\theta)$ 是关于 θ 的单调减函数 (见图 7.1.2)，只需要

$$g(110) = \Phi(5(c-110)/4) = 0.05$$

成立即可. 由此可先确定标准正态分布的 0.05 分位数 $u_{0.05} = -u_{0.95}$，它使得

$$\frac{5(c-110)}{4} = u_{0.05}$$

从而 c 的值为

$$c = 110 + 0.8u_{0.05} = 110 - 0.8 \times 1.645 = 108.684$$

所以，检验的拒绝域为

$$W = \{\bar{x} \leqslant 108.684\}$$

若令 $u = \frac{\bar{x}-110}{4/5}$，则拒绝域有另一种表示，即

$$W = \{u \leqslant u_{0.05}\} = \{u \leqslant -1.645\}$$

五、做出判断

在有了明确的拒绝域 W 后，根据样本观测值我们可以做出判断：

- 当 $\bar{x} \leqslant 108.684$ 或 $u \leqslant -1.645$ 时，则拒绝 H_0，即接收 H_1；
- 当 $\bar{x} > 108.684$ 或 $u > -1.645$ 时，则接收 H_0.

在例 7.1.1 中，由于

$$\bar{x} = 108 < 108.684$$

因此拒绝原假设，即认为该日生产不正常.

✍ 习 题 7.1

1. 设 $x_1, \cdots, x_n$ 是来自 $N(\mu, 1)$ 的样本，考虑如下假设检验问题

$$H_0 : \mu = 2 \quad \text{vs} \quad H_1, \mu = 3$$

若检验由拒绝域为 $w = \{\bar{x} \geqslant 2.6\}$ 确定.

(1) 当 $n=20$ 时求检验犯两类错误的概率;

(2) 如果要使得检验犯第二类错误的概率 $\beta \leqslant 0.01$,n 最小应取多少?

(3) 证明:当 $n \to +\infty$ 时,$\alpha \to 0, \beta \to 0$.

2. 设 $x_1, \cdots, x_{10}$ 是来自 $0-1$ 总体 $b(1,p)$ 的样本,考虑如下检验问题

$$H_0: p=0.2 \quad \text{vs} \quad H_1, p=0.4$$

取拒绝域为 $W=\{\bar{x} \geqslant 0.5\}$,求该检验犯两类错误的概率.

3. 设 $x_1, \cdots, x_{16}$ 是来自正态总体 $N(\mu, 4)$ 的样本,考虑检验问题

$$H_0: \mu=6 \quad \text{vs} \quad H_1: \mu \neq 6$$

拒绝域取为 $W=\{|\bar{x}-6| \geqslant c\}$,试求 c 使得检验的显著性水平为 0.05,并求该检验在 $\mu=6.5$ 处犯第二类错误的概率.

4. 设总体为均匀分布 $U(0,\theta)$, $x_1, \cdots, x_n$ 是样本,考虑检验问题

$$H_0: \theta \geqslant 3 \quad \text{vs} \quad H_1: \theta<3$$

拒绝域取为 $W=\{x_{(n)} \leqslant 2.5\}$,求检验犯第一类错误的最大值 α,若要使得该最大值 α 不超过 0.05,n 至少应取多大?

5. 在假设检验问题中,若检验结果是接受原假设,则检验可能犯哪一类错误? 若检验结构是拒绝原假设,则又有可能犯哪一类错误?

7.2 正态总体参数假设检验

参数假设检验常见的有三种基本形式

$$\begin{aligned}
&(1)\ H_0: \theta \leqslant \theta_0 \quad \text{vs } H_1: \theta>\theta_0 \\
&(2)\ H_0: \theta \geqslant \theta_0 \quad \text{vs } H_1; \theta<\theta_0 \\
&(3)\ H_0: \theta=\theta_0 \quad \text{vs } H_1: \theta \neq \theta_0
\end{aligned}$$

一般说来,对这三种假设所采用的检验统计量是相同的,差别在拒绝域上. 当备择假设 H_1 在原假设 H_0 一侧时的检验称为**单侧检验**,当备择假设 H_1 分散在原假设 H_0 两侧时的检验称为**双侧检验**. 以上 (1),(2) 是单侧检验,(3) 是双侧检验. 识别单侧与双侧检验有益于以后构造其拒绝域.

本节对正态总体参数检验分别进行讨论.

7.2.1 单个正态总体均值的检验

设 $x_1, \cdots, x_n$ 是来自 $N(\mu, \sigma^2)$ 的样本,考虑如下三种关于 μ 的检验问题

$$H_0: \mu \leqslant \mu_0 \quad \text{vs} \quad H_1: \mu>\mu_0 \tag{7.2.1}$$

$$H_0: \mu \geqslant \mu_0 \quad \text{vs} \quad H_1: \mu<\mu_0 \tag{7.2.2}$$

$$H_0: \mu=\mu_0 \quad \text{vs} \quad H_1: \mu \neq \mu_0 \tag{7.2.3}$$

由于正态总体含两个参数,总体方差 σ^2 已知与否对检验有影响. 下面我们分 σ 已知和未知两种情况叙述.

一、σ 已知时的 μ 检验

对于单侧检验问题(7.2.1),由于 μ 的点估计是,且 $\bar{x} \sim N\left(\mu, \sigma^2/n\right)$,故选用检验统计量

$$u=\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} \tag{7.2.4}$$

是恰当的. 直觉告诉我们:当样本均值 $\bar{x}$ 不超过设定均值 μ_0 时,应接收原假设;当样本均值 $\bar{x}$ 超过 μ_0 时,应拒绝原假设. 可是,在有随机性存在的场合,如果 $\bar{x}$ 比 μ_0 大一点就拒绝原假设似乎不当,只有当 $\bar{x}$ 比 μ_0 大到一定程度时拒绝原假设才是恰当的,这就存在一个临界值 c,拒绝域为

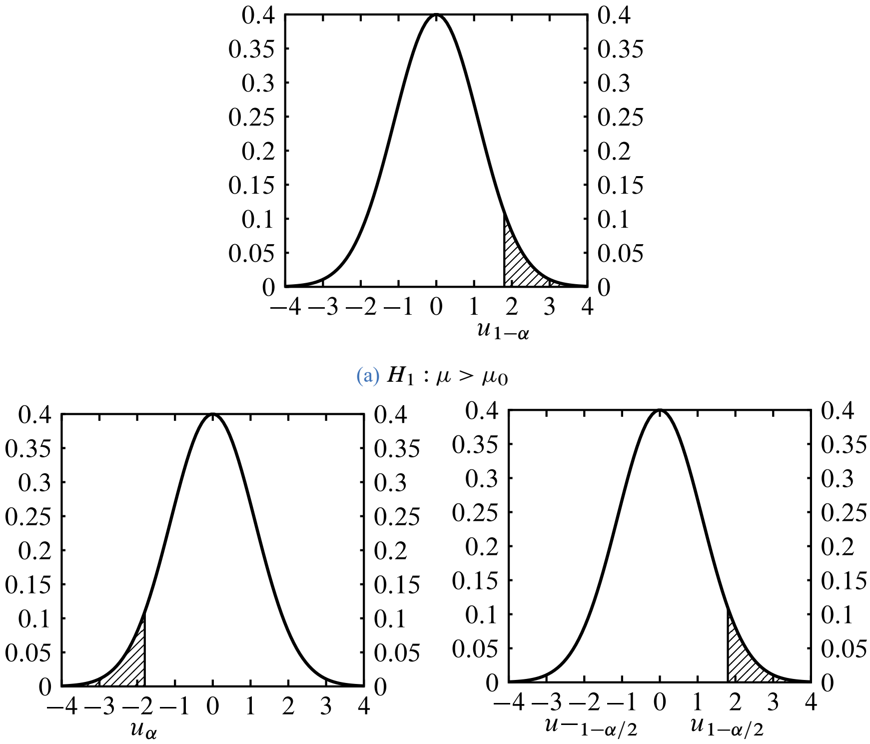
$$W = \{(x_1, \cdots, x_n) : u \geqslant c\} \tag{7.2.5}$$

常简记为 $\mu \geqslant c$,若要求检验的显著性水平为 α,则 c 满足

$$P_{\mu_0}(u \geqslant c) = \alpha$$

由于在 $\mu = \mu_0$ 时 $\mu \sim N(0, 1)$,故 $c = \mu_{1-\alpha}$(见图 7.2.1(a)),最后的拒绝域为

$$W = \{u \geqslant u_{1-\alpha}\} \tag{7.2.6}$$



(a) $H_1 : \mu > \mu_0$

(b) $H_1 : \mu < \mu_0$　　(c) $H_1 : \mu \neq \mu_0$

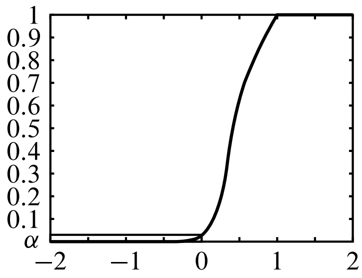
图 7.2.1: μ 检验的拒绝域

该检验用的检验统计量是 μ 统计量,故一般称为 μ 检验. 该检验的势函数是 μ 的函数,它可用正态分布写出,具体如下:对 $\mu \in (\infty, +\infty)$,

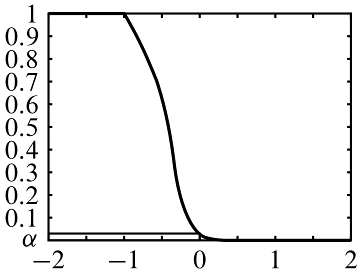
$$\begin{aligned} g(\mu) &= P_\mu(X \in W) = P_\mu(u \geqslant u_{1-\alpha}) \\ &= P_\mu\left(\frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \geqslant u_{1-a}\right) \\ &= P_\mu\left(\frac{\bar{x} - \mu + \mu - \mu_0}{\sigma/\sqrt{n}} \geqslant u_{1-\alpha}\right) \\ &= P_\mu\left(\frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \geqslant \frac{\mu_0 - \mu}{\sigma/\sqrt{n}} + u_{1-\alpha}\right) \\ &= 1 - \Phi\left(\sqrt{n}(\mu_0 - \mu)/\sigma + u_{1-a}\right) \end{aligned}$$

由此可见,势函数是 μ 的增函数,其图形见图 7.2.2(a),由增函数性质知,只要 $g(\mu_0 = \alpha)$ 就可保证在 $\mu \leqslant \mu_0$ 时有 $g(\mu) \leqslant \mu_0$,所以上述求出的检验是水平为 α 的检验.

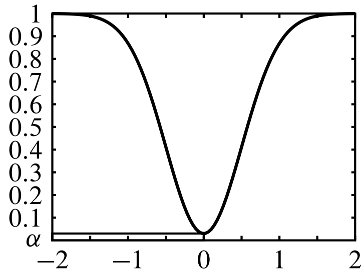
对单侧检验问题 (7.2.1),可以类似进行讨论. 仍选用 μ 作为检验统计量,考虑到 (7.2.1) 的备择



(a) $H_1: \mu > \mu_0$



(b) $H_1: \mu < \mu_0$



(c) $H_1: \mu \neq \mu_0$

图 7.2.2: $g(\mu)$ 的图形

假设 H_1 在左侧,故其拒绝域应有如下形式

$$W = \{\mu \leqslant c\}$$

对给定的显著性水平 $\alpha(0 < \alpha < 1)$,由 $P_{\mu_0}(u \leqslant c) = \alpha$ 可定出 $c = \mu_\alpha$(图 7.2.1(b)),最后的拒绝域为

$$W = \{\mu \leqslant \mu_\alpha\} \tag{7.2.7}$$

判断准则是类似的:当 $\mu \leqslant \mu_\alpha$ 时拒绝原假设 H_0(接收 H_1),否则接收 H_0. 该检验的势函数也是 μ 的减函数(图 7.2.2(b)),具体如下:对 $\mu \in (-\infty, +\infty)$,

$$\begin{aligned} g(\mu) &= P_\mu(u \leqslant u_a) \\ &= P_\mu\left(\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} \leqslant u_{\mathrm{a}}\right) \\ &= P_\mu\left(\frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \leqslant \frac{\mu_0-\mu}{\sigma/\sqrt{n}} + u_\alpha\right) \\ &= \Phi\left(\sqrt{n}(\mu_0-\mu)/\sigma + u_a\right) \end{aligned}$$

对双侧检验问题 (7.2.3),也可类似进行讨论. 仍选用 μ 作为检验统计量,考虑到 (7.2.3) 的备择假设 H_1 分散在二侧,故其拒绝域亦应在二侧,即拒绝域应有如下形式

$$W = \{|\mu| \geqslant c\}$$

对给定的显著性水平 $\alpha(0 < \alpha < 1)$,由 $P_{\mu_0}(|\mu| \geqslant c) = \alpha$ 可定出 $c = \mu_{1-\alpha/2}$(见图 7.2.1(c)),最后的拒绝域为

$$W = \{|\mu| \geqslant_{1-\alpha/2}\} \tag{7.2.8}$$

判断准则仍类似:当 $\mu \geqslant_{1-\alpha/2}$ 时拒绝原假设 H_0(接收 H_1),否则接收 H_0. 该检验的势函数仍是 μ 的函数(图7.2.2(c)),具体如下:对 $\mu \in (-\infty, +\infty)$,

$$\begin{aligned} g(\mu) = P_\mu(|u| \geqslant u_{1-\alpha/2}) &= P_\mu\left(\frac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}} \geqslant u_{1-a/2}\right) \\ &= 1 - P_\mu\left(\frac{\mu_0-\mu}{\sigma/\sqrt{n}} - u_{1-a/2} \leqslant \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \leqslant \frac{\mu_0-\mu}{\sigma/\sqrt{n}} + u_{1-a/2}\right) \end{aligned}$$

$$= 1 - \Phi\left(\sqrt{n}(\mu_0 - \mu)/\sigma + u_{1-a/2}\right) + \Phi\left(\sqrt{n}(\mu_0 - \mu)/\sigma - u_{1-a/2}\right)$$

由图 7.2.2 可以看出,对三种假设检验问题,只要在 $\mu = \mu_0$ 处控制检验犯第一类错误的概率为 α,则检验就是水平为 α 的显著性检验. 这不是偶然的现象,具有一般性,这为求显著性检验提供了很大的方便.

例 7.2.1: 从甲地发送一个讯号到乙地. 设乙地接受到的讯号值是一个服从正态分布 $N(\mu, 0.2^2)$ 的随机变量,其中 μ 为甲地发送的真实讯号值. 现甲地重复发送同一讯号 5 次,乙地接收到的讯号值为

$$8.05 \quad 8.15 \quad 8.2 \quad 8.1 \quad 8.25$$

设接受方有理由猜测甲地发送的讯号值为 8,问能否接受这猜测?

解: 这是一个假设检验的问题,总体 $X \sim N\left(\mu, 0.2^2\right)$,待检验的原假设 H_0 与备择假设 H_1 分别为

$$H_0 : \mu = 8 \quad \text{vs} \quad H_1 : \mu \neq 8$$

这是一个双侧检验问题,检验的拒绝域为 $\left\{|u| \geqslant u_{1-\alpha/2}\right\}$. 取置信水平 $\alpha = 0.05$,则查表知 $u_{0.975} = 1.96$. 现该例中观测值可计算得出 $\bar{x} = 8.15, u = \sqrt{5}(8.15-8)/0.2 = 1.68$, μ 值未落入拒绝域内,故不能拒绝原假设,即接受原假设,可认为猜测成立.

二、σ 未知时的 t 检验

对检验问题 (7.2.1),由于 σ 未知,(7.2.4) 给出的 μ 含未知参数 σ 而无法计算,需要做修改. 一个自然的想法是将 (7.2.4) 中未知的 σ 替换成样本标准差 s,这就形成 t 检验统计量

$$t = \frac{\sqrt{n}(\bar{x} - \mu_0)}{s} \tag{7.2.9}$$

由定理 5.4.1 知,在 $\mu = \mu_0$ 时 $t \sim t(n-1)$,从而检验问题 (7.2.1) 的拒绝域为

$$W = \{t \geqslant t_{1\text{-}a}(n-1)\} \tag{7.2.10}$$

对检验问题 (7.2.2),拒绝域为

$$W = \{t \leqslant t_a(n-1)\} \tag{7.2.11}$$

对检验问题 (7.2.3),拒绝域为

$$W = \left\{|t| \geqslant t_{1-a/2}(n-1)\right\} \tag{7.2.12}$$

同样可证明这三个检验都是水平为 α 的检验.

例 7.2.2: 某厂生产的某种铝材的长度服从正态分布,其均值设定为 240cm. 现从该厂抽取 5 件产品,测得其长度为(单位:cm)

$$239.7 \quad 239.6 \quad 239 \quad 240 \quad 239.2$$

试判断该厂此类铝材的长度是否满足设定要求?

这是一个关于正态均值的双侧假设检验问题. 原假设是 $H_0 : \mu = 240$, 备择假设是 $H_1 : \mu = 240$. 由于 σ 未知, 故采用 t 检验, 其拒绝域为 $\left\{|t| \geqslant t_{1-a/2}(n-1)\right\}$, 若取 $\alpha = 0.05$, 则查表得 $t_{0.975}(4) = 2.776$. 现由样本计算得到 $\bar{x} = 239.5, s = 0.4$,故

$$t = \sqrt{5} \cdot |239.5 - 240|/0.4 = 2.795$$

由于 $2.795 > 2.776$,故拒绝原假设,认为该厂生产的铝材的长度不满足设定要求.

综上,关于单个正态总体的均值的检验问题可汇总成表 7.2.1.

表 7.2.1: 单个正态总体均值的假设检验

检验法	条件	原假设 H_0	备择假设 H_1	检验统计量	拒绝域
		$\mu \leqslant \mu_0$	$\mu > \mu_0$		$\{u \geqslant u_1 - \alpha\}$
μ 检验	σ 已知	$\mu \geqslant \mu_0$	$\mu < \mu_0$	$\mu = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$	$\{u \leqslant u_\alpha\}$
		$\mu = \mu_0$	$\mu \neq \mu_0$		$\{\lvert u\rvert \geqslant u_{1-\alpha/2}\}$
		$\mu \leqslant \mu_0$	$\mu > \mu_0$		$\{t \geqslant \mu_{1-\alpha}(n-1)\}$
t 检验	σ 已知	$\mu \geqslant \mu_0$	$\mu < \mu_0$	$t = \frac{\bar{x}-\mu_0}{s/\sqrt{n}}$	$\{t \leqslant t_\alpha(n-1)\}$
		$\mu = \mu_0$	$\mu \neq \mu_0$		$\{\lvert t\rvert \geqslant t_{1-\alpha/2}(n-1)\}$

三、假设检验与置信区间的关系

细心的读者可能会发现，这里用的检验统计量与 6.5.5 节中置信区间所用的枢轴量是相同的，这不是偶然的，两者之间存在非常密切的关系，现叙述如下.

设 $x_1, \cdots, x_n$ 是来自正态总体 $N(\mu, \sigma^2)$ 的样本，现讨论在 σ 未知场合关于均值 μ 的检验问题. 分三种情况：

考虑双侧检验问题

$$H_0 : \mu = \mu_0 \quad \text{vs} \quad H_1 : \mu \neq \mu_0$$

则水平为 α 的检验的接收域为

$$\overline{W} = \left\{|\bar{x} - \mu_0| \leqslant \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1)\right\}$$

它可以改写为

$$\overline{W} = \left\{\bar{x} - \frac{s}{\sqrt{n}} t_{1-a/2}(n-1) \leqslant \mu_0 \leqslant \bar{x} + \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1)\right\}$$

并且有 $P_{\mu_0}(\overline{W} = 1-\alpha)$，这里 μ_0 并无限制，若让 μ_0 在 $(-\infty, +\infty)$ 内取值，就可得到 μ 的 $1-\alpha$ 置信区间：$\bar{x} \pm \frac{s}{\sqrt{n}} t_{1-a/2}(n-1)$. 反之，若有一个如上的 $1-\alpha$ 置信区间，也可获得关于 $H_0 : \mu = \mu_0$ 的水平为 α 的显著性检验. 所以，"正态均值 μ 的 $1-\alpha$ 置信区间"与"关于 $H_0 : \mu = \mu_0$ 的双侧检验问题的水平为 α 的检验"是一一对应.

考虑单侧检验问题

$$H_0 : \mu \leqslant \mu_0 \quad \text{vs} \quad H_1 : \mu > \mu_0$$

其水平为 α 的检验的接受域为

$$\begin{aligned}\overline{W} &= \left\{\bar{x} - \mu_0 \geqslant \frac{s}{\sqrt{n}} t_{1-a}(n-1)\right\} \\ &= \left\{\mu_0 \leqslant \bar{x} - \frac{s}{\sqrt{n}} t_{1-\alpha}(n-1)\right\}\end{aligned}$$

这就给出了参数 μ 的 $1-\alpha$ 置信上限. 反之，对上述给定的 p 的 $1-\alpha$ 置信上限，我们也可以得到关于 $H_0 : u \leqslant u_0$ 的单侧检验问题的水平为 α 的检验，它们之间也是一一对应的. 类似地，对另一个单侧检验问题，其水平为 α 的检验与参数 μ 的 $1-\alpha$ 置信下限也是一一对应的.

7.2.2 两个正态总体均值差的检验

设 $x_1, \cdots, x_m$ 是来自正态总体 $N(\mu_1, \sigma_1^2)$ 的样本，$y_1, \cdots, y_n$ 是来自另一个正态总体 $N(\mu_2, \sigma_2^2)$ 的样本，两个样本相互独立. 考虑如下三类检验问题：

$$H_0 : \mu_1 - \mu_2 \leqslant 0 \quad \text{vs} \quad H_1 : \mu_1 - \mu_2 > 0 \tag{7.2.13}$$

$$H_0 : \mu_1 - \mu_2 \geqslant 0 \quad \text{vs} \quad H_1 : \mu_1 - \mu_2 < 0 \tag{7.2.14}$$

$$H_0 : \mu_1 - \mu_2 = 0 \quad \text{vs} \quad H_1 : \mu_1 - \mu_2 \neq 0 \tag{7.2.15}$$

这里主要分两种情形进行讨论.

一、σ_1, σ_2 已知时的两样本 μ 检验

此时 $\mu_1 - \mu_2$ 的点估计 $\bar{x} - \bar{y}$ 的分布完全已知,

$$\bar{x} - \bar{y} \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}\right)$$

由此可采用 μ 检验方法,检验统计量为

$$u = (\bar{x} - \bar{y}) \Big/ \sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}$$

在 $\mu_1 = \mu_2$ 时,$\mu \sim N(0,1)$. 检验的拒绝域取决于备择假设的具体内容. 对检验问题 (7.2.13),检验的拒绝域为

$$W = \{u \geqslant u_{1-\alpha}\}, \tag{7.2.16}$$

对检验问题 (7.2.14),检验的拒绝域为

$$W = \{\mu \leqslant \mu_\alpha\}, \tag{7.2.17}$$

对检验问题 (7.2.15),检验的拒绝域为

$$W = \left\{|\mu| \geqslant \mu_{1-\alpha/2}\right\}. \tag{7.2.18}$$

二、$\sigma_1 = \sigma_2 = 0$ 但未知时的两样本 t 检验

在 $\sigma_1^2 = \sigma_2^2 = \sigma^2$ 但未知时,首先 $\bar{x} - \bar{y} \sim N\left(\mu_1 - \mu_2, (\frac{1}{m} + \frac{1}{n})\sigma^2\right)$,其次,由于

$$\frac{1}{\sigma^2}\sum_{i=1}^{\infty}(x_i - \bar{x})^2 \sim \chi^2(m-1), \quad \frac{1}{\sigma^2}\sum_{i=1}^{n}(y_i - \bar{y})^2 \sim \chi^2(n-1)$$

故 $\frac{1}{\sigma^2}\left(\sum(x_i - \bar{x})^2 + \sum(y_i - \bar{y})^2\right) \sim \chi^2(m+n-2)$,记

$$s_w^2 = \frac{1}{m+n-2}\left[\sum_{i=1}^{m}(x_i - \bar{x})^2 + \sum_{i=1}^{n}(y_i - \bar{y})^2\right]$$

于是有

$$t = \frac{(\bar{x} - \bar{y}) - (\mu_1 - \mu_2)}{s_w\sqrt{\frac{1}{m} + \frac{1}{n}}} \sim t(m+n-2)$$

这就给出了检验统计量为

$$t = \frac{(\bar{x} - \bar{y})}{s_w\sqrt{\frac{1}{m} + \frac{1}{n}}}$$

对检验问题 (7.2.13),检验的拒绝域为

$$W = \{t \geqslant t_{1-\alpha}(m+n-2)\}, \tag{7.2.19}$$

对检验问题 (7.2.14),检验的拒绝域为

$$W = \{t \leqslant t_\alpha(m+n-2)\}, \tag{7.2.20}$$

对检验问题 (7.2.15),检验的拒绝域为

$$W = \left\{|t| \geqslant t_{1-\alpha/2}(m+n-2)\right\}. \tag{7.2.21}$$

例 7.2.3: 某厂铸造车间为提高铸件的耐磨性而试制了一种镍合金铸件以取代铜合金铸件,为此,从两种铸件中各抽取一个容量分别为 8 和 9 的样本,测得其硬度(一种耐磨性指标)为

镍合金:76.43 76.21 73.58 69.69 65.29 70.83 82.75 72.34

铜合金：73.66　64.27　69.34　71.37　69.77　68.12　67.27　68.07　62.61

根据专业经验，硬度服从正态分布，且方差保持不变，试在显著性水平 $\alpha = 0.05$ 下判断镍合金的硬度是否有明显提高.

解： 用 X 表示镍合金的硬度，Y 表示铜合金的硬度，则由假定，$X \sim N(\mu_1, \sigma^2)$，$Y \sim N(\mu_2, \sigma^2)$，要检验的假设是：$H_0 : \mu_1 = \mu_2 \quad \text{vs} \quad H_1 : \mu_1 > \mu_2$. 由于两者方差未知但相等，故采用两样本 t 检验，经计算，

$$\bar{x} = 73.39, \quad \bar{y} = 68.2756, \quad \sum_{i=1}^{8}(x_i - \bar{x})^2 = 205.7958$$

$$\sum_{i=1}^{9}(y_i - \bar{y})^2 = 91.1552$$

从而 $s_w = \sqrt{\frac{1}{8+9-2}(205.7958 + 91.1552)} = 4.4494$，

$$t = \frac{73.39 - 68.2756}{4.4494 \cdot \sqrt{\frac{1}{7} + \frac{1}{8}}} = 2.2210$$

查表知 $t_{0.95}(15) = 1.753$，由于 $t > t_{0.95}(15)$，故拒绝原假设，可判断镍合金硬度有显著提高.

利用假设检验与置信区间的关系对其他情况下的检验问题可仿 6.5.5 节中两正态总体均值差的置信区间类似进行，我们下面以表格形式列出，而不作推导（l 的表达式见 6.5.5 节）.

表 7.2.2: 两个正态总体均值的假设检验

检验法	条件	原假设 H_0	备择假设 H_1	检验统计量	拒绝域
		$\mu_1 \leqslant \mu_2$	$\mu_1 > \mu_2$		$\{u \geqslant u_{1-\alpha}\}$
u 检验	σ_1, σ_2 已知	$\mu_1 \geqslant \mu_2$	$\mu_1 < \mu_2$	$u = \frac{(\bar{x}-\bar{y})}{\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}}$	$\{u \leqslant u_\alpha\}$
		$\mu_1 = \mu_2$	$\mu_1 \neq \mu_2$		$\{\lvert u\rvert \geqslant u_{1-\alpha/2}\}$
	σ_1, σ_2 未知	$\mu_1 \leqslant \mu_2$	$\mu_1 > \mu_2$		$\{t \geqslant t_{1-\alpha}(m+n-2)\}$
t 检验		$\mu_2 \geqslant \mu_2$	$\mu_1 < \mu_2$	$t = \frac{(\bar{x}-\bar{y})}{s_w\sqrt{\frac{1}{m}+\frac{1}{n}}}$	$\{t \leqslant t_\alpha(m+n-2)\}$
	$\sigma_1 = \sigma_2$	$\mu_1 = \mu_2$	$\mu_1 \neq \mu_2$		$\{\lvert t\rvert \geqslant t_{1-\alpha/2}(m+n-2)\}$
大样本	σ_1, σ_2 未知	$\mu_1 \leqslant \mu_2$	$\mu_1 > \mu_2$		$\{u \geqslant u_{1-\alpha}\}$
检验	m, n	$\mu_1 \geqslant \mu_2$	$\mu_1 < \mu_2$	$u = \frac{(\bar{x}-\bar{y})}{\sqrt{\frac{s_x^2}{n}+\frac{s_y^2}{m}}}$	$\{u \leqslant u_\alpha\}$
	充分大	$\mu_1 = \mu_2$	$\mu_1 \neq \mu_2$		$\{\lvert u\rvert \geqslant u_{1-\alpha/2}\}$
近似	σ_1, σ_2 未知	$\mu_1 \leqslant \mu_2$	$\mu_1 > \mu_2$		$\{t \geqslant t_{1-\alpha}(l-1)\}$
t 检验	m, n	$\mu_1 \geqslant \mu_2$	$\mu_1 < \mu_2$	$t = \frac{(\bar{x}-\bar{y})}{\sqrt{\frac{s_x^2}{n}+\frac{s_y^2}{m}}}$	$\{t \leqslant t_\alpha(l-1)\}$
	不很大	$\mu_1 = \mu_2$	$\mu_1 \neq \mu_2$		$\{\lvert t\rvert \geqslant t_{1-\alpha/2}(l-1)\}$

7.2.3 正态总体方差的检验

一、单个正态总体方差的 χ^2 检验

设 $x_1, \cdots, x_n$ 是来自 $N(\mu, \sigma^2)$ 的样本，对方差亦可考虑如下三个检验问题：

$$H_0 : \sigma^2 \leqslant \sigma_0^2 \quad \text{vs} \quad H_1 : \sigma^2 > \sigma_0^2 \tag{7.2.22}$$

$$H_0 : \sigma^2 \geqslant \sigma_0^2 \quad \text{vs} \quad H_1 : \sigma^2 < \sigma_0^2 \tag{7.2.23}$$

$$H_0 : \sigma^2 = \sigma_0^2 \quad \text{vs} \quad H_1 : \sigma^2 \neq \sigma_0^2 \tag{7.2.24}$$

此处通常假定 μ 未知，它们采用的检验统计量是相同的，均为

$$\chi^2 = (n-1)s^2/\sigma_0^2 \tag{7.2.25}$$

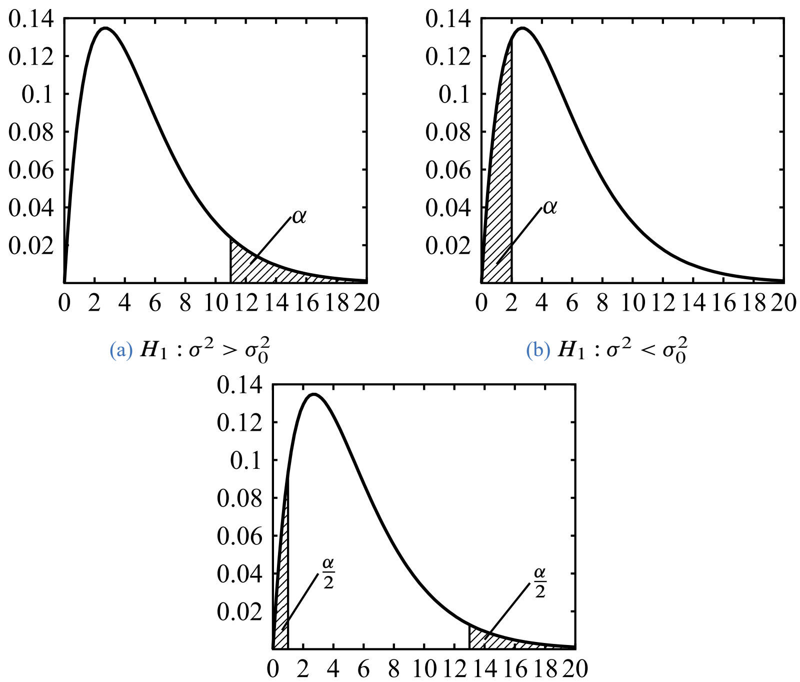
在 $\sigma^2=\sigma_0^2$ 时，$\chi^2\sim\chi^2(n-1)$，于是，若取显著性水平为 α，则对应三个检验问题的拒绝域依次分别为

$$W=\{x^2\geqslant\chi^2_{1-\alpha}(n-1)\}$$

$$W=\{\chi^2\leqslant\chi^2_{\alpha}(n-1)!\}$$

$$W=\left\{\chi^2\leqslant\chi^2_{\alpha/2}(n-1)\text{或}\chi^2\geqslant\chi^2_{1-\alpha/2}(n-1)\right\}$$

χ^2 分布是偏态分布，三种拒绝域形式见图 7.2.3



(a) $H_1:\sigma^2>\sigma_0^2$ (b) $H_1:\sigma^2<\sigma_0^2$

(c) $H_1:\sigma^2\neq\sigma_0^2$

图 7.2.3: χ^2 检验的拒绝域 $(n-1=6,\alpha=0.05)$

例 7.2.4: 某类钢板每块的重量 X 服从正态分布，其一项质量指标是钢板重量的方差不得超过 $0.016kg^2$. 现从某天生产的钢板中随机抽取 25 块，得其样本方差 $s^2=0.025kg^2$，问该天生产的钢板重量的方差是否满足要求.

解： 这是一个关于正态总体方差的单侧检验问题. 原假设为 $H_0:\sigma^2\leqslant 0.016$，备择假设为 $H_1:\sigma^2>0.016$，此处 $n=25$，若取 $\alpha=0.05$，则查表知 $\chi^2_{0.95}(24)=36.415$，现计算可得

$$\chi^2=\frac{(n-1)s^2}{\sigma_0^2}=\frac{24\times 0.025}{0.016}=37.5>36.415$$

由此，在显著性水平 0.05 下，我们拒绝原假设，认为该天生产的钢板重量不符合要求.

若取 $\alpha=0.01$，则 $\chi^2_{0.99}(24)=42.98$，则不能拒绝原假设. 所以，显著性水平的大小对检验结果是有影响的. 关于这方面的讨论我们在 7.4 节进行.

二、两个正态总体方差比的 F 检验

设 $x_1,\cdots,x_m$ 是来自 $N(\mu_1,\sigma_1^2)$ 的样本，$y_1,\cdots,y_n$ 是来自 $N(\mu_2,\sigma_2^2)$ 的样本，考虑如下三个假设检验问题

$$H_0:\sigma_1^2\leqslant\sigma_2^2\quad\text{vs}\quad H_1:\sigma_1^2>\sigma_2^2 \tag{7.2.26}$$

$$H_0:\sigma_1^2\geqslant\sigma_2^2\quad\text{vs}\quad H_1:\sigma_1^2<\sigma_2^2 \tag{7.2.27}$$

$$H_0:\sigma_1^2=\sigma_2^2 \quad \text{vs} \quad H_1:\sigma_1^2\neq\sigma_2^2 \tag{7.2.28}$$

此处 μ_1,μ_2 均未知，记 s_x^2,s_y^2 分别是由 $x_1,\cdots,x_m$ 算得的 σ_1^2 的无偏估计和由 $y_1,\cdots,y_n$ 算得的 σ_2^2 的无偏估计 (两个都是样本方差)，则可建立如下的检验统计量

$$F=\frac{s_x^2}{s_y^2} \tag{7.2.29}$$

当 $\sigma_1^2=\sigma_2^2$ 时，(7.2.29) 的 $F\sim F(m-1,n-1)$，由此给出三个检验问题对应的拒绝域依次分别为

$$W=\{F\geqslant F_{1-\alpha}(m-1,n-1)\}$$

$$W=\{F\leqslant F_{1-\alpha}(m-1,n-1)\}$$

$$W=\left\{F\leqslant F_{\alpha/2}(m-1,n-1)\right\}\text{或}W=\left\{F\geqslant F_{\alpha/2}(m-1,n-1)\right\}.$$

例 7.2.5: 甲、乙两台机床加工某种零件，零件的直径服从正态分布，总体方差反映了加工精度，为比较两台机床的加工精度有无差别，现从各自加工的零件中分别抽取 7 件产品和 8 件产品，测得其直径为

X机床甲	16.2	16.4	15.8	15.5	16.7	15.6	15.8	
Y机床乙	15.9	16.0	16.4	16.1	16.5	15.8	15.7	15.0

这就形成了一个双侧假设检验问题，原假设是 $H_0:\sigma_1^2=\sigma_2^2$，备择假设为 $H_1:\sigma_1^2\neq\sigma_2^2$. 此处 $m=7,n=8$，经计算，$s_x^2=0.2729,s_y^2=0.2164$，于是 $F=\frac{0.2729}{0.2164}=1.261$. 若取 $\alpha=0.05$，查表知 $F_{0.975}(6,7)=5.12$, $F_{0.025}=\frac{1}{F_{0.025}(7,6)}=\frac{1}{5.70}=0.175$. 其拒绝域为

$$W=\{F\leqslant 0.175\text{或}F\geqslant 5.12\}$$

由此可见，样本未落人拒绝域，即在 0.05 水平下可以认为二台机床的加工精度一致.

关于正态总体方差的假设检验汇总列于表 7.2.3中.

表 7.2.3: 正态总体方差的假设检验

检验法	H_0	H_1	检验统计量	拒绝域
χ^2 检验	$\sigma^2\leqslant\sigma_0^2$ $\sigma^2\geqslant\sigma_0^2$ $\sigma^2=\sigma_0^2$	$\sigma^2>\sigma_0^2$ $\sigma^2<\sigma_0^2$ $\sigma^2\neq\sigma_0^2$	$\chi^2=\frac{(n-1)s^2}{\sigma_0^2}$	$\chi^2\geqslant\chi_{1-\alpha}^2(n-1)$ $\chi^2\leqslant\chi_{\alpha}^2(n-1)$ $\chi^2\leqslant\chi_{\alpha/2}^2(n-1)$ 或 $\chi^2\leqslant\chi_{1-\alpha}^2(n-1)$
F 检验	$\sigma^2\leqslant\sigma_0^2$ $\sigma^2\geqslant\sigma_0^2$ $\sigma^2=\sigma_0^2$	$\sigma^2>\sigma_0^2$ $\sigma^2<\sigma_0^2$ $\sigma^2\neq\sigma_0^2$	$F=\frac{s_x^2}{s_y^2}$	$F\geqslant F_{1-\alpha}(m-1,n-1)$ $F\leqslant F_{\alpha}(m-1,n-1)$ $F\leqslant F_{\alpha/2}(m-1,n-1)$或 $F\geqslant F_{1-\alpha/2}(m-1,n-1)$

✍ 习 题 7.2

1. 有一批枪弹，出厂时，其初速 $v\sim N(950,100)$ (单位：m/s) . 经过较长时间储存，取 9 发进行测试，得样本值 (单位：m/s) 如下：

   914 920 910 934 953 945 912 924 940

   据经验，枪弹经储存后其初速仍服从正态分布，且标准差保持不变，问是否可认为这批枪弹的初速有显著降低 ($\alpha=0.05$)?
2. 已知某炼铁厂铁水含碳量服从正态分布 $N(4.55,0.108^2)$. 现在测定了 9 炉铁水，其平均含碳量为 4.484，如果铁水含碳量的方差没有变化，可否认为现在生产的铁水平均含碳量仍为 4.55 ($\alpha=0.05$)?

3. 由经验知某零件质量 $X \sim N\left(15, 0.05^2\right)$（单位:g）,技术革新后,抽出 6 个零件,测得质量为:

$$14.7 \quad 15.1 \quad 14.8 \quad 15.0 \quad 15.2 \quad 14.6$$

已知方差不变,问平均质量是否仍为 15g?（取 α=0.05）

4. 化肥厂用自动包装机包装化肥，每包的质量服从正态分布，其平均质量为 100kg，标准差为 1.2kg. 某日开工后,为了确定这天包装机工作是否正常,随机抽取 9 袋化肥,称得质量如下:

$$99.3 \quad 98.7 \quad 100.5 \quad 101.2 \quad 98.3 \quad 99.7 \quad 99.5 \quad 102.1 \quad 100.5$$

设方差稳定不变,问这一天包装机的工作是否正常?（取 α=0.05）

5. 设需要对某正态总体的均值进行假设检验

$$H_0: \mu = 15, \quad H_1: \mu < 15$$

已知 $\sigma^2 = 2.5$,取 $\alpha = 0.05$,若要求当 H_1 中的以 $\mu \leqslant 13$ 时犯第二类错误的概率不超过 0.05,求所需的样本容量.

6. 从一批钢管抽取 10 根,测得其内径（单位:mm）为:

100.36	100.31	99.99	100.11	100.64
100.85	99.42	99.91	99.35	100.10

设这批钢管内直径服从正态分布 $N(\mu, \sigma^2)$,试分别在下列条件下检验假设 (α=0.05).

$$H_0: \mu = 100 \quad \text{vs} \quad H_1: \mu > 100$$

(1) 已知 $\sigma = 0.5$.

(2) σ 未知.

7. 假定考生成绩服从正态分布,在某地一次数学统考中,随机抽取了 36 位考生的成绩,算得平均成绩为 66.5 分,标准差为 15 分,问在显著性水平 0.05 下,是否可以认为这次考试全体考生的平均成绩为 70 分?

8. 一个小学校长在报纸上看到这样的报道:“这一城市的初中学生平均每周看 8h 电视.”她认为她所在学校的学生看电视的时间明显小于该数字为此她在该校随机调查了 100 个学生,得知平均每周看电视的时间 $barx = 6.5$ h,样本标准差为 $s = 2$ h. 问是否可以认为这位校长的看法是对的? 取 α=0.05.

9. 设在木材中抽出 100 根，测其小头直径，得到样本平均数为 $\bar{x} = 11.2cm$，样本标准差 $s = 2.6cm$,问该批木材小头的平均直径能否认为不低于 12cm($\alpha = 0.05$)?

10. 考察一鱼塘中鱼的含汞量,随机地取 10 条鱼测得各条鱼的含汞量（单位:mg）为:

$$0.8 \quad 1.6 \quad 0.9 \quad 0.8 \quad 1.2 \quad 0.4 \quad 0.7 \quad 1.0 \quad 1.2 \quad 1.1$$

设鱼的含汞量服从正态分布 $N(\mu, \sigma^2)$，试检验假设 $H_0 : \mu \leqslant 1.2$ vs $\mathrm{H}_1; \mu > 1.2$（取 $\alpha = 0.10$）.

11. 如果一个矩形的宽度 w 与长度 l 的比 $\frac{w}{l} = \frac{1}{2}(\sqrt{5} - 1) \approx 0.618$,这样的矩形称为黄金矩形. 下面列出某工艺品工厂随机取的 20 个矩形宽度与长度的比值.

0.693	0.749	0.654	0.670	0.662	0.672	0.615	0.606	0.690	0.628
0.668	0.611	0.606	0.609	0.553	0.570	0.844	0.576	0.933	0.630

设这一工厂生产的矩形的宽度与长度的比值总体服从正态分布,其均值为 μ,试检验假设（取 α=0.05）

$$H_0: \mu = 0.618 \quad \text{vs} \quad H_1: \mu \neq 0.618$$

12. 下面给出两种型号的计算器充电以后所能使用的时间 (h) 的观测值

型号 A	5.5	5.6	6.3	4.6	5.3	5.0	6.2	5.8	5.1	5.2	5.9

型号 B 3.8 4.3 4.2 4.0 4.9 4.5 5.2 4.8 4.5 3.9 3.7 4.6

设两样本独立且数据所属的两总体的密度函数至多差一个平移量. 试问能否认为型号 A 的计算器平均使用时间比型号 B 来得长（$(\alpha=0.01)$?

13. 从某锌矿的东、西两支矿脉中，各抽取样本容量分别为 9 与 8 的样本进行测试，得样本含锌平均数及样本方差如下：

$$东支:\bar{x}_1=0.230, s_1^2=0.1337$$

$$西支:\bar{x}_2=0.269, s_2^2=0.1736$$

若东、西两支矿脉的含锌量都服从正态分布且方差相同，问东、西两支矿脉含锌量的平均值是否可以看作一样 $(\alpha=0.05)$?

14. 在针织品漂白工艺过程中，要考察温度对针织品断裂强力（主要质量指标）的影响. 为了比较 70°C 与 80°C 的影响有无差别，在这两个温度下，分别重复做了 8 次试验，得数据如下（单位：N）：

70°C 时的强力：18.5, 18.8, 19.8, 20.9, 21.5, 19.5, 21.0, 21.2

80°C 时的强力：17.7, 20.3, 20.0, 18.8, 19.0, 20.1 20.0, 19.1

根据经验，温度对针织品断裂强度的波动没有影响. 问在 70°C 时的平均断裂强力与 80°C 时的平均断裂强力间是否有显著差别?（假定断裂强力服从正态分布，α=0.05.）

15. 一药厂生产一种新的止痛片，厂方希望验证服用新药片后至开始起作用的时间间隔较原有止痛片至少缩短一半，因此厂方提出需检验假设

$$H_0:\mu_1=2\mu_2 \quad \text{vs} \quad H_1;\dot{\mu}_1>2\mu_2$$

此处 μ_1,μ_2 分别是服用原有止痛片和服用新止痛片后至开始起作用的时间间隔的总体的均值. 设两总体均为正态分布且方差分别为已知值 σ_1^2,σ_2^2，现分别在两总体中取一样本 $x_1,\cdots,x_n$ 和 $y_1,\cdots,y_m$，设两个样本独立. 试给出上述假设检验问题的检验统计景及拒绝域.

16. 已知维尼纶纤度在正常条件下服从正态分布，且标准差为 0.048. 从某天产品中抽取 5 根纤维，测得其纤度为

1.32, 1.55, 1.36, 1.40, 1.44,

问这一天纤度的总体标准差是否正常?（取 $\alpha=0.05$）

17. 某电工器材厂生产一种保险丝. 测量其熔化时间，依通常情况方差为 400，今从某天产品中抽取容量为 25 的样本，测量其熔化时间并计算得 $\bar{x}=62.24, s^2=52=404.77$，问这天保险熔化时间分散度与通常有无显著差异?（取 α=0.05，假定熔化时间跟从正态分布.）

18. 某种导线的质量标准要求其电阻的标准差不得超过 $0.005(\Omega)$. 今在一批导线中随机抽取样品 9 根，测得样本标准差为 $s=0.007(\Omega)$，设总体为正态分布. 问在水平 $\alpha=0.05$ 下能否认为这批导线的标准差显著地偏大?

19. 两台车床生产同一种滚珠，滚珠直径服从正态分布. 从中分别抽取 8 个和 9 个产品，测得其直径为

甲车床：15.0, 14.5, 15.2, 15.5, 14.8, 15.1, 15.2, 14.8

乙车床：15.2, 15.0, 14.8, 15.2, 15.0, 15.0, 14.8, 15.1, 14.8

比较两台车床生产的滚珠直径的方差是否有明显差异 $(\alpha=0.05)$.

20. 有两台机器生产金属部件，分别在两台机器所生产的部件中各取一容量为 $m=14$ 和 $n=12$ 的样本，测得部件重量的样本方差分别为 $s_1^2=15.46, s_2^2=9.66$，设两样本相互独立，试在水平

α=0.05 下检验假设

$$H_0 : \sigma_1^2 = \sigma_2^2 \quad \text{vs} \quad H_1; \sigma_1^2 > \sigma_2^2$$

21. 测得两批电子器件的样品的电阻 (Ω) 为

A 批（x）	0.140	0.138	0.143	0.142	0.144	0.137
B 批（y）	0.135	0.140	0.142	0.136	0.138	0.140

设这两批器材的电阻值分别服从分布 $N\left(\mu_1, \sigma_1^2\right), N\left(\mu_2, \sigma_2^2\right)$，且两样本独立.
(1) 试检验两个总体的方差是否相等?（α=0.05）
(2) 试检验两个总体的均值是否相等?（α=0.05）

22. 某厂使用两种不同的原料生产同一类型产品，随机选取使用原料 A 生产的样品 22 件，测得平均质量为 2.36（kg），样本标准差为 0.57（kg）. 取使用原料 B 生产的样品 24 件，测得平均质量为 2.55（kg），样本标准差为 0.48（kg）. 设产品质量服从正态分布，两个样本独立. 问能否认为使用原料 B 生产的产品平均质量较使用原料 A 显著大?（取 α=0.05.）

7.3 其他分布参数的假设检验

7.3.1 指数分布参数的假设检验

指数分布是一类重要的分布，有广泛的应用. 设 $x_1, \cdots, x_n$，是来自指数分布 $\text{Exp}(1/\theta)$ 的样本，θ 为其均值，现考虑关于 θ 的如下检验问题：

$$H_0 : \theta \leqslant \theta_0 \quad \text{vs} \quad H_1 : \theta > \theta_0 \tag{7.3.1}$$

拒绝域的自然形式是 $W = \{\bar{x} \geqslant c\}$，下面讨论 $\bar{x}$ 的分布.

为寻找检验统计量，我们考察参数 θ 的充分统计量 $\bar{x}$. 在 $\theta = \theta_0$ 时，$n\bar{x} = \sum_{i=1}^n x_1 \sim Ga(n, 1/\theta_0)$，由伽玛分布性质可知

$$\chi^2 = \frac{2n\bar{x}}{\theta_0} \sim \chi^2(2n) \tag{7.3.2}$$

于是可用 χ^2 作为检验统计量并利用 $\chi^2(2n)$ 的分位数建立检验的拒绝域，对检验问题 7.3.1，拒绝域为

$$W = \left\{\chi^2 \geqslant \chi^2_{1-a}(2n)\right\} \tag{7.3.3}$$

关于 θ 的另两种检验问题的处理方法是类似的. 对检验问题

$$H_0 : \theta \geqslant \theta_0 \quad \text{vs} \quad H_1 : \theta < \theta_0 \tag{7.3.4}$$

$$H_0 : \theta = \theta_0 \quad \text{vs} \quad H_1 : \theta \neq \theta_0 \tag{7.3.5}$$

检验统计量仍然是 7.3.2的 χ^2，拒绝域分别为

$$W = \left\{\chi^2 \leqslant \chi^2_a(2n)\right\} \tag{7.3.6}$$

$$W = \left\{\chi^2 \leqslant \chi^2_{\alpha/2}(2n) \text{ 或 } \chi^2 \geqslant \chi^2_{\alpha/2}(2n)\right\} \tag{7.3.7}$$

例 7.3.1: 设我们要检验某种元件的平均寿命不小于 6000h，假定元件寿命为指敬分布，现取 5 个元件投人试验，观测到如下 5 个失效时间（h）

395　4094　119　11572　6133

这是一个假设检验问题,检验的假设为

$$H_0:\theta \geqslant 6000 \quad \text{vs} \quad H_1:\theta < 6000$$

经计算,$\bar{x}=4462.6$,故检验统计量为

$$\chi^2=\frac{10\bar{x}}{\theta_0}=\frac{10\times 4462.6}{6000}=7.4377$$

若取 $\alpha=0.05$,则查表知 $\chi^2_{0.05}(10)=3.94$,由于 $\chi^2>\chi^2_{0.05}(10)$,故接受原假设可以认为平均寿命不低于 6000 h.

7.3.2 比例 p 检验

比例 p 可看作某事件发生的概率,即可看作二点分布 $b(1,p)$ 中的参数. 作 n 次独立试验,以 x 记该事件发生的次数,则 $x\sim b(n,p)$. 我们可以根据 x 检验关于 p 的一些假设. 先考虑如下单边假设检验问题.

$$H_0:p\leqslant p_0 \quad \text{vs} \quad H_1:p>p_0 \tag{7.3.8}$$

直观上看,一个显然的检验方法是取如下的拒绝域 $W=\{x\geqslant c\}$,由于 x 只取整数值,故 c 可限制在非负整数中. 然而,一般情况下对给定的 α,不一定能正好取到一个 c 使

$$P(x\geqslant c;p_0)=\sum_{i=c}^{n}\binom{n}{i}p_0^i(1-p_0)^{n-i}=\alpha \tag{7.3.9}$$

能恰巧使得7.3.9成立的 c 值是罕见的. 这是在对离散总体作假设检验中普遍会遇到的问题,在这种情况下,较常见的是找一个 c_0,使得

$$\sum_{i=c_0}^{n}\binom{n}{i}p_0^i(1-p_0)^{n-i}>a>\sum_{i=c_0+1}^{n}\binom{n}{i}p_0^i(1-p_0)^{i-1}$$

于是,若取 $c=c_0$,这相当于把检验的显著性水平提高了一些,由 α 提高到到 $\sum_{i=c_0}^{n}\binom{n}{i}p_0^i(1-p_0)^{n-i}$ 若取 $c=c_0+1$,此时相当于把显著性水平由 α 降低到 $\sum_{i=c_0+1}^{n}\binom{n}{i}p_0^i(1-p_0)^{i-1}$,因为后者可保证 (7.3.9) 的左侧不大于 α,故取 $c=c_0+1$ 可得水平为 α 的检验,

对检验问题

$$H_0:p\geqslant p_0 \quad \text{vs} \quad H_1:p<p_0 \tag{7.3.10}$$

处理方法是类似的,检验的拒绝域为 $W=\{x\geqslant c\}$,c 为满足

$$\sum_{i=0}^{c}\binom{n}{i}p_0^i(1-p_0)^{n-i}\leqslant a$$

的最大正整数. 对检验问题

$$H_0:p=p_0 \quad \text{vs} \quad H_1;p\neq p_0, \tag{7.3.11}$$

检验的拒绝域 $W=\{x\leqslant c_1\}$ 或 $\{x\geqslant c_2\}$,其中 c_1 为满足

$$\sum_{i=0}^{c_1}\binom{n}{i}p_0^i(1-p_0)^{n-i}\leqslant\frac{\alpha}{2}$$

的最大整数,c_2 为满足

$$\sum_{i=c_2}^{n_1}\binom{n}{i}p_0^i(1-p_0)^{n-i}\leqslant\frac{\alpha}{2}$$

的最小整数.

例 7.3.2: 某厂生产的产品优质品率一直保持在 40%,近期对该厂生产的该类产品抽检 20 件,其中优质品 7 件,在 $\alpha=0.05$ 下能否认为优质品率仍保持在 40%?

这是一个假设检验问题, 以 p 表示优质品率, x 表示 20 件产品中的优质品件数, 则 $x\sim$

$b\square(20,p)$,待检验的原假设为

$$H_0: p=0.4 \quad \text{vs} \quad H_1: p\neq 0.4$$

拒绝域为 $W=\{x\leqslant c_1\}$ 或 $\{x\geqslant c_2\}$,下求 c_1 与 c_2.

由于

$$P(x\leqslant 3)=0.0160<0.025<P(x\leqslant 4)=0.0510$$

故取 $c_1=3$,又因

$$P(x\geqslant 11)=0.0565>0.025>P(x\geqslant 12)=0.0210$$

从而 $c_2=12$,拒绝域为 $W=\{x\leqslant 3\}$ 或 $\{x\geqslant 12\}$. 附带指出,该拒绝域的显著水平实际上不是 0.05,而是 $0.0160+0.021=0.0370$,本例中,由于观测值没有落入拒绝域,故接受原假设.

7.3.3 大样本检验

前一小节我们介绍了对二点分布参数 p 的检验问题, 我们看到临界值的确定比较繁琐, 使用不太方便. 如果样本量较大, 我们可用近似的检验方法——大样本检验. 其一般思路如下: 设 $x_1,\cdots,x_n$,是来自某总体的样本,又设该总体均值为 θ,方差为 θ 的函数,记为 $\sigma^2(\theta)$,譬如,对二点分布 $b(1,\theta)$,其方差 $\theta(1-\theta)$ 是均值 θ 的函数,则对下列三类假设检验问题:

(1) $H_0:\theta\leqslant\theta_0 \quad \text{vs} \quad H_0:\theta>\theta_0$;

(2) $H_0:\theta\geqslant\theta_0 \quad \text{vs} \quad H_0:\theta<\theta_0$;

(3) $H_0:\theta=\theta_0 \quad \text{vs} \quad H_0:\theta\neq\theta_0$

在样本容量 n 充分大时,利用中心极限定理,$\bar{x}\dot{\sim}N\left(\theta,\sigma^2(\theta)/n\right)$,故在 $\theta=\theta_0$ 时,可采用如下检验统计量

$$u=\frac{\sqrt{n}(\bar{x}-\theta_0)}{\sqrt{\sigma^2(\theta_0)}}\dot{\sim}N(0,1) \tag{7.3.12}$$

近似地确定拒绝域. 对应上述三类检验问题的拒绝域依次分别为

$$W=\{u\geqslant u_{1-\alpha}\}$$
$$W=\{u\leqslant u_a\}$$
$$W=\left\{|u|\geqslant u_{1-a/2}\right\}$$

例 7.3.3: 某厂生产的产品不合格率为 10%,在一次例行检查中,随机抽取 80 件,发现有 11 件不合格品,在 $\alpha=0.05$ 下能否认为不合格品率仍为 10%?

解: 这是关于不合格品率 θ 的检验,假设为

$$H_0:\theta=0.1 \quad \text{vs} \quad H_1:\theta\neq 0.1$$

我们可以仿例 7.3.2 的方法求拒绝域,但要把此拒绝域找出来是困难的,如今 $n=80$ 比较大,因此可采用大样本检验方法. 由 (7.3.12),$\theta_0=0.1,\sigma^2(\theta_0=0.1\times 0.9)$,检验统计量为

$$u=\frac{\sqrt{80}\left(\frac{11}{80}-0.1\right)}{\sqrt{0.1\times 0.9}}=1.118$$

若取 $\alpha=0.05$,则 $u_{0.975}=1.96$,故拒绝域为 $W=\{|u|\geqslant 1.96\}$. 如今 $u=1.118$ 未落入拒绝域,故不能拒绝原假设,可以认为不合格品率仍为 10%.

例 7.3.4: 某建筑公司宣称其磨下建筑工地平均每天发生事故数不超过 0.6 起,现记录了该公司磨下建筑工地 200 天的安全生产情况,事故数记录如下:

一天发生的事故数	0	1	2	3	4	5	$\geqslant 6$	合计
天数	102	59	30	8	0	1	0	200

试检验该建筑公司的宣称是否成立.(取 $\alpha=0.05$.)

解: 以 X 记该建筑公司麿下建筑工地一天发生的事故数,可认为 $X\sim P(\lambda)$(见习题 7.4.4),现要

检验的假设是：

$$H_0 : \lambda \leqslant 0.6 \quad \text{vs} \quad H_1 : \lambda > 0.6$$

由于 $n = 200$ 很大，故可以采用大样本检验，泊松分布的均值和方差都是 λ，而

$$\bar{x} = \frac{1}{200}(0 \times 102 + 1 \times 59 + 2 \times 30 + 3 \times 8 + 4 \times 0 + 5 \times 1) = 0.74$$

由 (7.3.12)，检验统计量为

$$u = \frac{\sqrt{n}(\bar{x} - \lambda)}{\lambda} = \frac{\sqrt{200}(0.74 - 0.6)}{\sqrt{0.6}} = 2.556$$

若取 $\alpha = 0.05$，则 $u_{0.95} = 1.645$，拒绝域为 $W = \{u \geqslant 1.645\}$. 如今 $u = 2.556$，已落入拒绝域，故拒绝原假设，认为该建筑公司的宣称明显不成立.

大样本检验是近似的. 近似的含义是指检验的实际显著性水平与原先设定的显著性水平有差距，这是由于诸如（7.3.12）中 u 的分布与 $N(0,1)$ 有距离. 如果 n 很大，则这种差异就很小. 实用中我们一般并不清楚对一定的 n，u 的分布与 $N(0,1)$ 的差异有多大，因而也就不能确定检验的实际水平与设定水平究竟差多少. 在区间估计中也有类似问题. 因此，大样本方法是一个"不得已而为之"的方法. 只要有基于精确分布的方法一般总是首先要加以考虑的.

7.3.4 检验的 p 值

假设检验的结论通常是简单的. 在给定的显著水平下，不是拒绝原假设就是保留原假设. 然而有时也会出现这样的情况：在一个较大的显著水平（比如 $\alpha = 0.05$）下得到拒绝原假设的结论，而在一个较小的显著水平（比如 $\alpha = 0.01$）下却会得到相反的结论. 这种情况在理论上很容易解释：因为显著水平变小后会导致检验的拒绝域变小，于是原来落在拒绝域中的观测值就可能落入接受域，但这种情况在应用中会带来一些麻烦. 假如这时一个人主张选择显著水平 $\alpha = 0.05$，而另一个人主张选 $\alpha = 0.01$，则第一个人的结论是拒绝 H_0，而后一个人的结论是接受 H0，我们该如何处理这一问题呢? 下面从一个例子谈起.

例 7.3.5: 一支香烟中的尼古丁含量 X 服从正态分布 $N(\mu, 1)$，质量标准规定 μ 不能超过 1.5 mg. 现从某厂生产的香烟中随机抽取 20 支测得其中平均每支香烟的尼古丁含量为 $\bar{x} = 1.97$ mg，试问该厂生产的香烟尼古丁含量是否符合质量标准的规定. 这是一个单侧假设检验问题，

$$\text{原假设} H_0 : \mu \leqslant 1.5, \qquad \text{备择假设} H_1 : \mu > 1.5$$

由于总体的标准差已知，故采用 μ 检验，由数据，

$$u = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{1.97 - 1.5}{1/\sqrt{20}} = 2.10$$

对一些的显著性水平，表 7.3.1 列出了相应的拒绝域和检验结论. 我们看到，不同的 α 有不同的结论.

表 7.3.1: 例 7.3.5 的拒绝域

显著性水平	拒绝域	$\mu = 2.10$ 对应的结论
$\alpha = 0.05$	$\mu \geqslant 1.645$	拒绝 H_0
$\alpha = 0.025$	$\mu \geqslant 1.96$	拒绝 H_0
$\alpha = 0.01$	$\mu \geqslant 2.33$	接受 H_0
$\alpha = 0.005$	$\mu \geqslant 2.58$	接受 H_0

现在换一个角度来看，在 $\mu = 1.5$ 时，μ 的分布是 $N(0,1)$. 此时可算得，$P(u \geqslant 2.10) = 0.0179$，若以 0.0179 为基准来看上述检验问题，可得

- 当 $\alpha < 0.0179$ 时，$u_{1-a} > 2.10$. 于是 2.10 就不在 $\{u > u_{1-a}\}$ 中，此时应接受原假设；
- 当 $\alpha \geqslant 0.0179$ 时，$u_{1-a} \leqslant 2.10$. 于是 2.10 就落在 $\{u \geqslant u_{1-a}\}$ 中，此时应拒绝 H_0.

由此可以看出,0.0179 是能用观测值 2.10 做出“拒绝 H_0”的最小的显著性水平,这就是 p 值,直观图形见图 7.3.1.

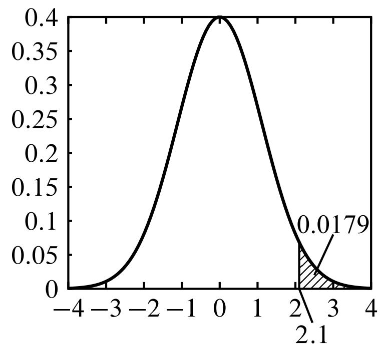


图 7.3.1: 观测值 $\mu = 2.10$ 对应的 p 值

定义 7.3.1. 在一个假设检验问题中,利用观测值能够做出拒绝原假设的最小显著性水平称为**检验的 p 值**.

引进检验的 p 值的概念有明显的好处. 第一,它比较客观,避免了事先确定显著水平;其次,由检验的 p 值与人们心目中的显著性水平 α 进行比较可以很容易做出检验的结论;

- 如果 $\alpha \geqslant p$,则在显著性水平 α 下拒绝 H_0;
- 如果 $\alpha < p$,则在显著性水平 α 下应保留 H_0.

p 值在应用中很有用,如今的统计软件中对检验问题一般都会给出检验的 p 值.

例 7.3.6: 设 $x_1, \cdots, x_n$ 是来自 $b(1,\theta)$ 的样本,要检验如下假设:

$$H_0 : \theta \leqslant \theta_0 \quad \text{vs} \quad H_1 ; \theta > \theta_0$$

若取检验的显著性水平为 α,则我们可给出检验的拒绝域的形式为:$W = |\sum x_i \geqslant c|$ 这里我们很难对一般的 n 和 α 定出 c 的表达式,我们只能说 c 是满足 $P_{\theta_0}|\sum x_i \geqslant c| \leqslant \alpha$ 的最小正整数. 这样的叙说总觉得不是很自然,事实上,我们并不需要定出 c,在得到观测值 $\sum x_i = t_0$ 后,我们只需要计算如下的概率即可

$$p = P_{\theta_0}\left|\sum x_i \geqslant c\right|$$

这就是检验的 p 值. 譬如,$n = 40, \theta_0 = 0.1, t_0 = 8$,则

$$p = 1 - 0.9^{40} - \binom{40}{1} 0.1 \times 0.9^{39} - \cdots - \binom{40}{7} 0.1^7 \times 0.9^{33} = 0.0419$$

于是,若取 $\alpha = 0.05$,由于 $p < \alpha$,则应拒绝原假设.

对双边假设检验,同样可以计算检验的 p 值,下面以一个例子加以说明.

例 7.3.7: 某工厂两位化验员甲、乙分别独立地用相同方法对某种聚合物的含氯量进行测定. 甲测 9 次,样本方差为 0.7292;乙测 11 次,样本方差为 0.2114. 假定测量数据服从正态分布,试对两总体方差作一致性检验.

解: 这是 7.2.3 节中关于二个正态总体的方差相等的检验,待检验的假设是

$$H_0 : \sigma_{\text{甲}}^2 = \sigma_{\text{乙}}^2, \quad \text{vs} \quad H_1 : \sigma_F^2 \neq \sigma_Z^2$$

检验统计量为 $F = s_{\text{甲}}^2 / s_{\text{乙}}^2$,在原假设成立下,$F \sim F(8,10)$,拒绝域为

$$W = \left\{F \geqslant F_{1-\alpha/2}(8,10) \text{或} F \leqslant F_{\alpha/2}(8,10)\right\}$$

如今我们不是把拒绝域具体化,而是由观测值算得 $F = 0.7292/0.2114 = 3.4494$,再去计算该检验的 p 值.

在这种双侧检验情况下,如何由观测值 $F = 3.4494$ 算得 p 值呢?

首先,我们用 F 分布算得

$$P(F \geqslant 3.4494) = 0.0354.$$

其次考虑到双侧检验的拒绝域 W 分散在两端,且两端尾部概率相等(见图 7.3.2),据此可定出 p 值为

$$p = 2P(F \geqslant 3.4494) = 0.0708$$

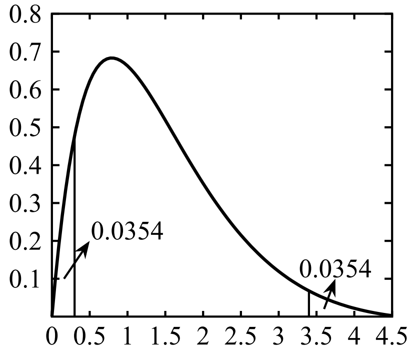


图 7.3.2: 观测值 $F = 3.4494$ 对应的 p 值由两端尾部概率之和确定

此 p 值不算很小,若心目中的 $\alpha = 0.05$,则接受两方差相等的假设.

✍习 题 7.3

1. 从一批服从指数分布的产品中抽取 10 个进行寿命试验,观测值如下(单位:h):

   1643 1629 426 132 1522 432 1759 1074 528 283

   根据这批数据能否认为其平均寿命不低于 1100h?(取 $(\alpha = 0.05)$
2. 某厂一种元件平均使用寿命为 1200h,偏低,现厂里进行技术革新,革新后任选 8 个元件进行寿命试验,测得寿命数据如下:

   2686 2001 2082 792 1660 4105 1416 2089

   假定元件寿命服从指数分布,取 $(\alpha = 0.05)$,问革新后元件的平均寿命是否有明显提高?
3. 有人称某地成年人中大学毕业生比例不低于 30%,为检验之,随机调查该地 15 名成年人,发现有 3 名大学毕业生,取 $\alpha = 0.05$,问该人看法是否成立?并给出检验的 p 值.
4. 某大学随机调查 120 名男同学,发现有 50 人非常喜欢看武侠小说,而随机调查的 85 名女同学中有 23 人喜欢,用大样本检验方法在 $\alpha = 0.05$ 下确认:男女同学在喜爱武侠小说方面有无显著差异?并给出检验的 p 值.
5. 假定电话总机在单位时间内接到的呼叫次数服从泊松分布,现观测了 40 个单位时间,接到的呼叫次数如下:

   0 2 3 2 3 2 1 0 2 2 1 2 2 1 3 1 1 4 1 1

   5 1 2 2 3 3 1 3 1 3 4 0 6 1 1 1 4 0 1 3

   在显著性水平 0.05 下能否认为单位时间内平均呼叫次数不低于 2.5 次?并给出检验的 p 值.
6. 通常每平方米某种布上的疵点数服从泊松分布,现观测该种布 $100m^2$,发现有 126 个疵点,在显著性水平为 0.05 下能否认为该种布每平方分米土平均疵点数不超过 1 个?并给出检验的 p 值.