

第八章 假设检验

第一节 概述

统计推断中的另一类重要问题是假设检验(Hypothesis testing).当总体的分布函数未知,或只知其形式而不知道它的参数的情况时,我们常需要判断总体是否具有我们所感兴趣的某些特性.这样,我们就提出某些关于总体分布或关于总体参数的假设,然后根据样本对所提出的假设作出判断:是接受还是拒绝.这就是本章所要讨论的假设检验问题.我们先从下面的例子来说明假设检验的一般提法.

例 8.1 某工厂用包装机包装奶粉,额定标准为每袋净重 0.5kg.设包装机称得奶粉重量 X 服从正态分布 $N(\mu, \sigma^2)$.根据长期的经验知其标准差 $\sigma=0.015(\text{kg})$.为检验某台包装机的工作是否正常;随机抽取包装的奶粉 9 袋,称得净重(单位: kg)为

0.499 0.515 0.508 0.512 0.498
0.515 0.516 0.513 0.524

问该包装机的工作是否正常?

由于长期实践表明标准差比较稳定,于是我们假设 $X \sim N(\mu, 0.015^2)$.如果奶粉重量 X 的均值 μ 等于 0.5kg,我们说包装机的工作是正常的.于是提出假设:

$$H_0: \mu = \mu_0 = 0.5;$$

$$H_1: \mu \neq \mu_0 = 0.5.$$

这样的假设叫统计假设.

1. 统计假设

关于总体 X 的分布(或随机事件的概率)的各种推断叫统计假设,简称假设,用“ H ”表示,例如:

(1) 对于检验某个总体 X 的分布,可以提出假设:

H_0 : X 服从正态分布, H_1 : X 不服从正态分布.

H_0 : X 服从泊松分布, H_1 : X 不服从泊松分布.

(2) 对于总体 X 的分布的参数,若检验均值,可以提出假设:

$$H_0: \mu = \mu_0; H_1: \mu \neq \mu_0.$$

$$H_0: \mu \leq \mu_0; H_1: \mu > \mu_0.$$

若检验标准差,可提出假设:

$$H_0: \sigma = \sigma_0; H_1: \sigma \neq \sigma_0.$$

$$H_0: \sigma \geq \sigma_0; H_1: \sigma < \sigma_0.$$

这里 μ_0, σ_0 是已知数,而 $\mu = E(X), \sigma^2 = D(X)$ 是未知参数.

上面对于总体 X 的每个论断,我们都提出了两个互相对立的(统计)假设: H_0 和 H_1 ,显然, H_0 与 H_1 只有一个成立,或 H_0 真 H_1 假,或 H_0 假 H_1 真,其中假设 H_0 ,称为原假设(Original hypothesis)(又叫零假设、基本假设),而 H_1 称为 H_0 的备择假设(又叫对立假设).

在处理实际问题时,通常把希望得到的陈述视为备择假设,而把这一陈述的否定作为原假设.例如在上例中, $H_0: \mu = \mu_0 = 0.5$ 为原假设,它的对立假设是 $H_1: \mu \neq \mu_0 = 0.5$.

统计假设提出之后,我们关心的是它的真伪.所谓对假设 H_0 的检验,就是根据来自总体的样本,按照一定的规则对 H_0 作出判断:是接受,还是拒绝,这个用来对假设作出判断的规则叫做检验准则,简称检验,如何对统计假设进行检验呢?下面我们结合上例来说明假设

检验的基本思想和做法.

2. 假设检验的基本思想

在例 8.1 中所提假设是

$$H_0: \mu = \mu_0 = 0.5 \text{ (备择假设 } H_1: \mu \neq \mu_0 \text{)}.$$

由于要检验的假设涉及总体均值 μ , 因此首先想到是否可借助样本均值 \bar{X} 这一统计量来进行判断. 从抽样的结果来看, 样本均值

$$\bar{X} = \frac{1}{9} (0.499 + 0.515 + 0.508 + 0.512 + 0.498 + 0.515 + 0.516 + 0.513 + 0.524) = 0.5110,$$

\bar{X} 与 $\mu = 0.5$ 之间有差异. 对于 \bar{X} 与 μ_0 之间的差异可以有两种不同的解释.

(1) 统计假设 H_0 是正确的, 即 $\mu = \mu_0 = 0.5$, 只是由于抽样的随机性造成了 \bar{X} 与 μ_0 之间的差异;

(2) 统计假设 H_0 是不正确的, 即 $\mu \neq \mu_0 = 0.5$, 由于系统误差, 也就是包装机工作不正常, 造成了 \bar{X} 与 μ_0 之间的差异.

对于这两种解释到底哪一种比较合理呢? 为了回答这个问题, 我们适当选择一个小正数 α ($\alpha = 0.1, 0.05$ 等), 叫做显著性水平 (Level of significance). 在假设 H_0 成立的条件下, 确定统计量 $\bar{X} - \mu_0$ 的临界值 λ_α , 使得事件 $\{|\bar{X} - \mu_0| > \lambda_\alpha\}$ 为小概率事件, 即

$$P\{|\bar{X} - \mu_0| > \lambda_\alpha\} = \alpha. \quad (8-1)$$

例如, 取定显著性水平 $\alpha = 0.05$. 现在来确定临界值 $\lambda_{0.05}$.

因为 $X \sim N(\mu, \sigma^2)$, 当 $H_0: \mu = \mu_0 = 0.5$ 为真时, 有 $X \sim N(\mu_0, \sigma^2)$, 于是

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \sim N\left(\mu_0, \frac{\sigma^2}{n}\right),$$

$$Z = \frac{\bar{X} - \mu_0}{\sqrt{\sigma^2/n}} = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0, 1),$$

所以

$$P\{|Z| > z_{\alpha/2}\} = \alpha.$$

由 (8-1) 式, 有

$$P\left\{|Z| > \frac{\lambda_\alpha}{\sigma/\sqrt{n}}\right\} = \alpha,$$

因此

$$\frac{\lambda_\alpha}{\sigma/\sqrt{n}} = z_{\frac{\alpha}{2}}, \quad \lambda_\alpha = z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}},$$

$$\lambda_{0.05} = z_{0.025} \times \frac{0.015}{\sqrt{9}} = 1.96 \times 0.015/3 = 0.0098.$$

故有

$$P\{|\bar{X} - \mu_0| > 0.0098\} = 0.05.$$

因为 $\alpha = 0.05$ 很小, 根据实际推断原理, 即“小概率事件在一次试验中几乎是不可能发生的”原理, 我们认为当 H_0 为真时, 事件 $\{|\bar{X} - \mu_0| > 0.0098\}$ 是小概率事件, 实际上是不可能发生的. 现在抽样的结果是

$$|\bar{x} - \mu_0| = |0.5110 - 0.5| = 0.0110 > 0.0098.$$

也就是说, 小概率事件 $\{|\bar{X} - \mu_0| > 0.0098\}$ 居然在一次抽样中发生了, 这说明抽样得到的结果与假设 H_0 不相符, 因而不能不使人怀疑假设 H_0 的正确性, 所以在显著性水平 $\alpha = 0.05$ 下, 我们拒绝 H_0 , 接受 H_1 , 即认为这一天包装机的工作是不正常的.

通过上例的分析, 我们知道假设检验的基本思想是小概率事件原理, 检验的基本步骤是:

- (1) 根据实际问题的要求, 提出原假设 H_0 及备择假设 H_1 ;
- (2) 选取适当的显著性水平 α (通常 $\alpha = 0.10, 0.05$ 等) 以及样本容量 n ;
- (3) 构造检验用的统计量 U , 当 H_0 为真时, U 的分布要已知, 找出临界值 λ_α 使 $P\{|\bar{X} - \mu_0| > \lambda_\alpha\} = \alpha$.

我们称 $|\bar{X} - \mu_0| > \lambda_\alpha$ 所确定的区域为 H_0 的拒绝域(Rejection region), 记作 W ;

- (4) 取样, 根据样本观察值, 计算统计量 U 的观察值 U_0 ;

(5) 作出判断, 将 U 的观察值 U_0 与临界值 λ_α 比较, 若 U_0 落入拒绝域 W 内, 则拒绝 H_0 接受 H_1 ; 否则就说 H_0 相容 (接受 H_0).

3. 两类错误

由于我们是根据样本作出接受 H_0 或拒绝 H_0 的决定, 而样本具有随机性, 因此在进行判断时, 我们可能会犯两个方面的错误: 一类错误是, 当 H_0 为真时, 而样本的观察值 U_0 落入拒绝域 W 中, 按给定的法则, 我们拒绝了 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弃真概率, 通常记为 α , 即

$$P\{\text{拒绝 } H_0 \mid H_0 \text{ 为真}\} = \alpha;$$

另一种错误是, 当 H_0 不真时, 而样本的观察值落入拒绝域 W 之外, 按给定的检验法则, 我们却接受了 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取伪概率, 通常记为 β , 即

$$P\{\text{接受 } H_0 \mid H_0 \text{ 不真}\} = \beta.$$

显然这里的 α 就是检验的显著性水平. 总体与样本各种情况的搭配如表 8-1.

表 8-1

H_0	判断结论		犯错误的概率
真	接受	正确	0
	拒绝	犯第一类错误	α
假	接受	犯第二类错误	β
	拒绝	正确	0

对给定的一对 H_0 和 H_1 , 总可以找到许多拒绝域 W . 当然我们希望寻找这样的拒绝域 W , 使得犯两类错误的概率 α 与 β 都很小. 但是在样本容量 n 固定时, 要使 α 与 β 都很小是不可能的, 一般情形下, 减小犯其中一类错误的概率, 会增加犯另一类错误的概率, 它们之间的关系犹如区间估计问题中置信水平与置信区间的长度的关系那样. 通常的做法是控制犯第一类错误的概率不超过某个事先指定的显著性水平 α ($0 < \alpha < 1$), 而使犯第二类错误的概率

也尽可能地小.具体实行这个原则会有许多困难,因而有时把这个原则简化成只要求犯第一类错误的概率等于 α , 称这类假设检验问题为显著性检验问题, 相应的检验为显著性检验. 在一般情况下, 显著性检验法则是较容易找到的, 我们将在以下各节中详细讨论.

在实际问题中, 要确定一个检验问题的原假设, 一方面要根据问题要求检验的是什么, 另一方面要使原假设尽量简单, 这是因为在下面将讲到的检验法中, 必须要了解某统计量在原假设成立时的精确分布或渐近分布.

在接下来的讨论中, 我们先介绍正态总体下参数的几种显著性检验, 再介绍总体分布函数的假设检验.

第二节 单个正态总体的假设检验

1. 单个正态总体数学期望的假设检验

(1) σ^2 已知, 关于 μ 的假设检验 (Z 检验法(Z -test)). 设总体 $X \sim N(\mu, \sigma^2)$, 方差 σ^2 已知, 检验假设:

$$H_0: \mu = \mu_0; H_1: \mu \neq \mu_0 \quad (\mu_0 \text{ 为已知常数})$$

由

$$\bar{X} \sim N\left(\mu, \frac{\sigma}{n}\right), \quad \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim N(0, 1),$$

我们选取

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \quad (8-2)$$

作为此假设检验的统计量, 显然当假设 H_0 为真 (即 $\mu = \mu_0$ 正确) 时, $Z \sim N(0, 1)$, 所以对于给定的显著性水平 α , 可求 $z_{\alpha/2}$ 使

$$P\{|Z| > z_{\alpha/2}\} = \alpha,$$

见图 8-1, 即

$$P\{Z < -z_{\alpha/2}\} + P\{Z > z_{\alpha/2}\} = \alpha.$$

从而有

$$\begin{aligned} P\{Z > z_{\alpha/2}\} &= \alpha/2, \\ P\{Z \leq z_{\alpha/2}\} &= 1 - \alpha/2. \end{aligned}$$

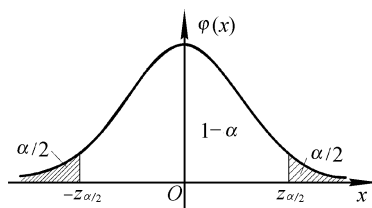


图 8-1

利用概率 $1 - \alpha/2$, 反查标准正态分布函数表, 得临界值 $z_{\alpha/2}$, 也称双侧 α 分位点.

另一方面, 利用样本观察值 x_1, x_2, \dots, x_n 计算统计量 Z 的观察值

$$z_0 = \frac{\bar{x} - \mu_0}{S / \sqrt{n}}. \quad (8-3)$$

如果: (a) $|z_0| > z_{\alpha/2}$, 则在显著性水平 α 下, 拒绝原假设 H_0 (接受备择假设 H_1), 所以 $|z_0| > z_{\alpha/2}$ 便是 H_0 的拒绝域.

(b) $|z_0| \leq z_{\alpha/2}$, 则在显著性水平 α 下, 接受原假设 H_0 , 认为 H_0 正确.

这里我们是利用 H_0 为真时服从 $N(0, 1)$ 分布的统计量 Z 来确定拒绝域的, 这种检验法称为 Z 检验法 (或称 U 检验法). 例 8.1 中所用的方法就是 Z 检验法. 为了熟悉这类假设检

验的具体作法，现在再看下面的例子。

例 8.2 根据长期经验和资料的分析，某砖厂生产的砖的“抗断强度” X 服从正态分布，方差 $\sigma^2=1.21$ 。从该厂产品中随机抽取 6 块，测得抗断强度如下（单位： $\text{kg} \cdot \text{cm}^{-2}$ ）：如下

32.56 29.66 31.64 30.00 31.87 31.03

检验这批砖的平均抗断强度为 $32.50 \text{ kg} \cdot \text{cm}^{-2}$ 是否成立（取 $\alpha=0.05$ ，并假设砖的抗断强度的方差不会有什么变化）？

解 提出假设 $H_0: \mu = \mu_0 = 32.50$; $H_1: \mu \neq \mu_0$ 。

选取统计量

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}},$$

若 H_0 为真，则 $Z \sim N(0, 1)$ 。

对给定的显著性水平 $\alpha=0.05$ ，求 $z_{\alpha/2}$ 使

$$P\{|Z| > z_{\alpha/2}\} = \alpha,$$

这里 $z_{\alpha/2} = z_{0.025} = 1.96$ 。

计算统计量 Z 的观察值：

$$|z_0| = \left| \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} \right| = \left| \frac{31.13 - 32.50}{1.1\sqrt{6}} \right| \approx 3.05.$$

判断：由于 $|z_0| = 3.05 > z_{0.025} = 1.96$ ，所以在显著性水平 $\alpha=0.05$ 下否定 H_0 ，即不能认为这批产品的平均抗断强度是 $32.50 \text{ kg} \cdot \text{cm}^{-2}$ 。

把上面的检验过程加以概括，得到了关于方差已知的正态总体期望值 μ 的检验步骤：

(a) 提出待检验的假设 $H_0: \mu = \mu_0$; $H_1: \mu \neq \mu_0$ 。

(b) 构造统计量 Z ，并计算其观察值 z_0 ：

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}, \quad z_0 = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}.$$

(c) 对给定的显著性水平 α ，根据

$$P\{|Z| > z_{\alpha/2}\} = \alpha, \quad P\{Z > z_{\alpha/2}\} = \alpha/2, \quad P\{Z \leq z_{\alpha/2}\} = 1 - \alpha/2$$

查标准正态分布表，得双侧 α 分位点 $z_{\alpha/2}$ 。

(d) 作出判断：（根据 H_0 的拒绝域）：

若 $|z_0| > z_{\alpha/2}$ ，则拒绝 H_0 ，接受 H_1 ；

若 $|z_0| \leq z_{\alpha/2}$ ，则接受 H_0 。

(2) 方差 σ^2 未知，关于 μ 的假设检验（ t 检验法(t-test)）。设总体 $X \sim N(\mu, \sigma^2)$ ，方差 σ^2 未知，检验假设：

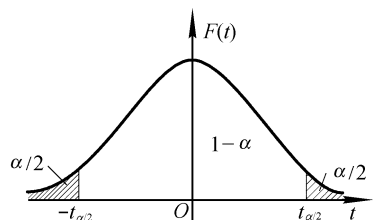
$$H_0: \mu = \mu_0; \quad H_1: \mu \neq \mu_0.$$

由于 σ^2 未知， $\frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$ 便不是统计量，这时我们自然想到用 σ^2 的无偏估计量——样本

方差 S^2 代替 σ^2 ，由于

$$\frac{\bar{X} - \mu}{S / \sqrt{n}} \sim t_{(n-1)},$$

因此选取样本的函数



$$t = \frac{\bar{X} - \mu_0}{S / \sqrt{n}} \quad (8-4)$$

图 8-2

作为统计量, 当 H_0 为真 ($\mu = \mu_0$) 时 $t \sim t(n-1)$, 对给定的检验显著性水平 α , 由

$$P\{|t| > t_{\alpha/2}(n-1)\} = \alpha,$$

$$P\{t > t_{\alpha/2}(n-1)\} = \alpha/2,$$

得 t 分布分位点 $t_{\alpha/2}(n-1)$, 也称双侧 α 分点, 见图 8-2, 直接查 t 分布表 (见附表 4)。

利用样本观察值, 计算统计量 t 的观察值

$$t_0 = \frac{\bar{x} - \mu_0}{s / \sqrt{n}},$$

因而原假设 H_0 的拒绝域为

$$|t_0| = \left| \frac{\bar{x} - \mu_0}{s / \sqrt{n}} \right| > t_{\alpha/2}(n-1). \quad (8-5)$$

所以, 若 $|t_0| > t_{\alpha/2}(n-1)$, 则拒绝 H_0 , 接受 H_1 ; 若 $|t_0| \leq t_{\alpha/2}(n-1)$, 则接受原假设 H_0 .

上述利用 t 统计量得出的检验法称为 t 检验法。

在实际中, 正态总体的方差常为未知, 所以我们常用 t 检验法来检验关于正态总体均值的问题。

例 8.3 用某仪器间接测量温度, 重复 5 次, 所得的数据是 1250° , 1265° , 1245° , 1260° , 1275° , 而用别的精确办法测得温度为 1277° (可看作温度的真值), 试问此仪器间接测量有无系统偏差?

这里假设测量值 X 服从 $N(\mu, \sigma^2)$ 分布。

解 问题是要检验:

$$H_0: \mu = \mu_0 = 1277; H_1: \mu \neq \mu_0.$$

由于 σ^2 未知 (即仪器的精度不知道), 我们选取统计量

$$t = \frac{\bar{X} - \mu_0}{S / \sqrt{n}}.$$

当 H_0 为真时, $t \sim t(n-1)$, t 的观察值为

$$|t_0| = \left| \frac{\bar{x} - \mu_0}{s / \sqrt{n}} \right| = \left| \frac{1259 - 1277}{\sqrt{(570)/(4 \times 5)}} \right| = \left| \frac{-18}{5.339} \right| > 3.$$

对于给定的检验水平 $\alpha = 0.05$, 由

$$P\{|t| > t_{\alpha/2}(n-1)\} = \alpha,$$

$$P\{t > t_{\alpha/2}(n-1)\} = \alpha/2,$$

$$P\{t > t_{0.025}(4)\} = 0.025,$$

查 t 分布表得双侧 α 分位点

$$t_{\alpha/2}(n-1) = t_{0.025}(4) = 2.776.$$

因为 $|t_0| > 3 > t_{0.025}(4) = 2.776$, 故应拒绝 H_0 , 认为该仪器间接测量有系统偏差。

(3) 双边检验与单边检验. 在上面讨论的假设检验中, H_0 为 $\mu = \mu_0$, 而备择假设 $H_1: \mu \neq \mu_0$ 意思是 μ 可能大于 μ_0 , 也可能小于 μ_0 , 称为双边备择假设, 而称形如 $H_0: \mu = \mu_0$, $H_1: \mu \neq \mu_0$ 的假设检验为双边检验. 有时我们只关心总体均值是否增大, 例如, 检验某种新

工艺是否可以提高材料的强度, 这时所考虑的总体的均值应该越大越好, 如果我们能判断在新工艺下总体均值较以往正常生产得大, 则可考虑采用新工艺. 此时, 我们需要检验假设

$$H_0: \mu = \mu_0; H_1: \mu > \mu_0. \quad (8-6)$$

(我们在这里作了不言而喻的假定, 即新工艺不可能比旧的更差), 形如 (8-6) 的假设检验, 称为右边检验, 类似地, 有时我们需要检验假设

$$H_0: \mu = \mu_0; H_1: \mu < \mu_0. \quad (8-7)$$

形如 (8-7) 的假设检验, 称为左边检验, 右边检验与左边检验统称为单边检验.

下面来讨论单边检验的拒绝域.

设总体 $X \sim N(\mu, \sigma^2)$, σ^2 为已知, x_1, x_2, \dots, x_n 是来自 X 的样本观察值. 给定显著性水平 α , 我们先求检验问题

$$H_0: \mu = \mu_0; H_1: \mu > \mu_0.$$

的拒绝域.

取检验统计量 $Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$, 当 H_0 为真时, Z 不应太大, 而在 H_1 为真时, 由于 \bar{X} 是 μ

的无偏估计, 当 μ 偏大时, \bar{X} 也偏大, 从而 Z 往往偏大, 因此拒绝域的形式为

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \geq k, \quad k \text{ 待定}.$$

因为当 H_0 为真时, $\frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \sim N(0, 1)$, 由

$$P\{\text{拒绝 } H_0 \mid H_0 \text{ 为真}\} = P\left\{\frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \geq k\right\} = \alpha$$

得 $k = z_\alpha$, 故拒绝域为

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \geq z_\alpha. \quad (8.8)$$

类似地, 左边检验问题

$$H_0: \mu = \mu_0; H_1: \mu < \mu_0.$$

的拒绝域为

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \leq -z_\alpha. \quad (8-9)$$

例 8.4 从甲地发送一个信号到乙地, 设发送的信号值为 μ , 由于信号传送时有噪声迭加到信号上, 这个噪声是随机的, 它服从正态分布 $N(0, 2^2)$, 从而乙地接到的信号值是一个服从正态分布 $N(\mu, 2^2)$ 的随机变量. 设甲地发送某信号 5 次, 乙地收到的信号值为:

8.4 10.5 9.1 9.6 9.9

由以往经验, 信号值为 8, 于是乙方猜测甲地发送的信号值为 8, 能否接受这种猜测? 取 $\alpha = 0.05$.

解 按题意需检验假设

$$H_0: \mu = 8; H_1: \mu > 8.$$

这是右边检验问题，其拒绝域如（8.8）式所示，

$$\text{即 } Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \geq z_{0.05} = 1.645.$$

而现在

$$z_0 = \frac{9.5 - 8}{2 / \sqrt{5}} = 1.68 > 1.645,$$

所以拒绝 H_0 ，认为发出的信号值 $\mu > 8$ 。

2. 单个正态总体方差的假设检验 (χ^2 检验法 (χ^2 -test))

(1) 双边检验. 设总体 $X \sim N(\mu, \sigma^2)$, μ 未知, 检验假设

$$H_0: \sigma^2 = \sigma_0^2; H_1: \sigma^2 \neq \sigma_0^2.$$

其中 σ_0^2 为已知常数.

由于样本方差 S^2 是 σ^2 的无偏估计, 当 H_0 为真时, 一般

来说, 比值 $\frac{S^2}{\sigma_0^2}$ 一般来说应在 1 附近摆动, 而不应过分大于 1

或过分小于 1, 由第六章知, 当 H_0 为真时,

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2(n-1). \quad (8-10)$$

所以对于给定的显著性水平 α 有 (见图 8-3)

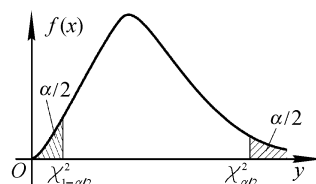


图 8-3

$$P\{\chi^2_{1-\alpha/2}(n-1) \leq \chi^2 \leq \chi^2_{\alpha/2}(n-1)\} = 1 - \alpha. \quad (8-11)$$

对于给定的 α , 查 χ^2 分布表可求得 χ^2 分布分位点 $\chi^2_{1-\alpha/2}(n-1)$ 与 $\chi^2_{\alpha/2}(n-1)$.

由 (8-11) 知, H_0 的接受域是

$$\chi^2_{1-\alpha/2}(n-1) \leq \chi^2 \leq \chi^2_{\alpha/2}(n-1); \quad (8-12)$$

H_0 的拒绝域为

$$\chi^2 < \chi^2_{1-\alpha/2}(n-1) \text{ 或 } \chi^2 > \chi^2_{\alpha/2}(n-1). \quad (8-13)$$

这种用服从 χ^2 分布的统计量对单个正态总体方差进行假设检验的方法, 称为 χ^2 检验法.

例 8.5 某厂生产的某种型号的电池, 其寿命长期以来服从方差 $\sigma^2 = 5000$ (h^2) 的正态分布, 现有一批这种电池, 从它的生产情况来看, 寿命的波动性有所改变, 现随机抽取 26 只电池, 测得其寿命的样本方差 $s^2 = 9200$ (h^2). 问根据这一数据能否推断这批电池的寿命的波动性较以往有显著的变化 (取 $\alpha = 0.02$)?

解 本题要求在 $\alpha = 0.02$ 下检验假设

$$H_0: \sigma^2 = 5000; H_1: \sigma^2 \neq 5000.$$

现在 $n = 26$,

$$\chi^2_{\alpha/2}(n-1) = \chi^2_{0.01}(25) = 44.314,$$

$$\chi_{1-\alpha/2}^2 (n-1) = \chi_{0.99}^2(25) = 11.524,$$

$$\sigma_0^2 = 5000.$$

由 (8-13) 拒绝域为

$$\frac{(n-1)s^2}{\sigma_0^2} > 44.314$$

或

$$\frac{(n-1)s^2}{\sigma_0^2} < 11.524$$

由观察值 $s^2=9200$ 得 $\frac{(n-1)s^2}{\sigma_0^2} = 46 > 44.314$, 所以拒绝 H_0 , 认为这批电池寿命的波动性较以往

有显著的变化.

(2) 单边检验 (右检验或左检验). 设总体 $X \sim N(\mu, \sigma^2)$, μ 未知, 检验假设

$$H_0: \sigma^2 \leq \sigma_0^2; H_1: \sigma^2 > \sigma_0^2. \text{ (右检验)}$$

由于 $X \sim N(\mu, \sigma^2)$, 因此随机变量

$$\chi^{*2} = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2 (n-1).$$

当 H_0 为真时, 统计量

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} \leq \chi^{*2}.$$

对于显著性水平 α , 有 (见图 8-4).

$$P\{\chi^{*2} > \chi_{\alpha}^2 (n-1)\} = \alpha$$

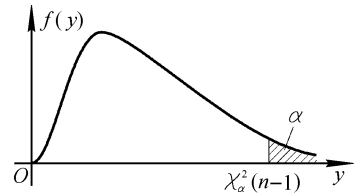


图 8-4

于是有

$$P\{\chi^2 > \chi_{\alpha}^2 (n-1)\} \leq P\{\chi^{*2} > \chi_{\alpha}^2 (n-1)\} = \alpha.$$

可见, 当 α 很小时, $\{\chi^2 > \chi_{\alpha}^2 (n-1)\}$ 是小概率事件, 在一次的抽样中认为不可能发生,

所以 H_0 的拒绝域为

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} > \chi_{\alpha}^2 (n-1) \text{ (右检验)}. \quad (8.14)$$

类似地, 可得左检验假设 $H_0: \sigma^2 \geq \sigma_0^2, H_1: \sigma^2 < \sigma_0^2$ 的拒绝域为

$$\chi^2 < \chi_{1-\alpha}^2 (n-1) \text{ (左检验)}. \quad (8-15)$$

例 8.6 今进行某项工艺革新, 从革新后的产品中抽取 25 个零件, 测量其直径, 计算得样本方差为 $s^2=0.00066$, 已知革新前零件直径的方差 $\sigma^2=0.0012$, 设零件直径服从正态分布, 问革新后生产的零件直径的方差是否显著减小 ($\alpha=0.05$) ?

解 提出假设: $H_0: \sigma^2 \geq \sigma_0^2 = 0.0012$; $H_1: \sigma^2 < \sigma_0^2$.
选取统计量

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}.$$

$$\chi^{*2} = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1), \text{ 且当 } H_0 \text{ 为真时, } \chi^{*2} \leq \chi^2$$

对于显著性水平 $\alpha=0.05$, 查 χ^2 分布表得

$$\chi_{1-\alpha}^2(n-1) = \chi_{0.95}^2(24) = 13.848,$$

当 H_0 为真时,

$$P\{\chi^2 < \chi_{1-\alpha}^2(n-1)\} \leq P\left\{\frac{(n-1)S^2}{\sigma^2} < \chi_{1-\alpha}^2(n-1)\right\} = \alpha.$$

故拒绝域为

$$\chi^2 < \chi_{1-\alpha}^2(n-1) = 13.848.$$

根据样本观察值计算 χ^2 的观察值

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{24 \times 0.00066}{0.0012} = 13.2.$$

作判断: 由于 $\chi^2 = 13.2 < \chi_{1-\alpha}^2(n-1) = 13.848$, 即 χ^2 落入拒绝域中, 所以拒绝 $H_0: \sigma^2 \geq \sigma_0^2$, 即认为革新后生产的零件直径的方差小于革新前生产的零件直径的方差.

最后我们指出, 以上讨论的是在均值未知的情况下, 对方差的假设检验, 这种情况在实际问题中较多. 至于均值已知的情况下, 对方差的假设检验, 其方法类似, 只是所选的统计量为

$$\chi^2 = \frac{\sum_{i=1}^n (X_i - \mu)^2}{\sigma_0^2}.$$

当 $\sigma^2 = \sigma_0^2$ 为真时, $\chi^2 \sim \chi^2(n)$.

关于单个正态总体的假设检验可列表 8-2.

表 8-2

检验参数	条件	H_0	H_1	H_0 的拒绝域	检验用的统计量	自由度	分位点
数学期望	σ^2 已知	$\mu = \mu_0$ $\mu \leq \mu_0$ $\mu \geq \mu_0$	$\mu \neq \mu_0$ $\mu > \mu_0$ $\mu < \mu_0$	$ Z > z_{\alpha/2}$ $Z > z_{\alpha}$ $Z < -z_{\alpha}$	$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$		$\pm z_{\alpha/2}$ z_{α} $-z_{\alpha}$
	σ^2 未知	$\mu = \mu_0$ $\mu \leq \mu_0$ $\mu \geq \mu_0$	$\mu \neq \mu_0$ $\mu > \mu_0$ $\mu < \mu_0$	$ t > t_{\alpha/2}$ $t > t_{\alpha}$ $t < -t_{\alpha}$	$t = \frac{\bar{X} - \mu_0}{S / \sqrt{n}}$	$n-1$	$\pm t_{\alpha/2}$ t_{α} $-t_{\alpha}$
方差	μ 未知	$\sigma^2 = \sigma_0^2$ $\sigma^2 \leq \sigma_0^2$ $\sigma^2 \geq \sigma_0^2$	$\sigma^2 \neq \sigma_0^2$ $\sigma^2 > \sigma_0^2$ $\sigma^2 < \sigma_0^2$	$\chi^2 > \chi_{\alpha/2}^2$ $\chi^2 < \chi_{1-\alpha/2}^2$ $\chi^2 > \chi_{\alpha}^2$ $\chi^2 < \chi_{1-\alpha}^2$	$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$	$n-1$	$\begin{cases} \chi_{\alpha/2}^2 \\ \chi_{1-\alpha/2}^2 \end{cases}$ χ_{α}^2 $\chi_{1-\alpha}^2$
	μ 已知	$\sigma^2 = \sigma_0^2$ $\sigma^2 \leq \sigma_0^2$ $\sigma^2 \geq \sigma_0^2$	$\sigma^2 \neq \sigma_0^2$ $\sigma^2 > \sigma_0^2$ $\sigma^2 < \sigma_0^2$	$\chi^2 > \chi_{\alpha/2}^2$ $\chi^2 < \chi_{1-\alpha}^2$ $\chi^2 > \chi_{\alpha}^2$ $\chi^2 < \chi_{1-\alpha}^2$	$\chi^2 = \frac{\sum_{i=1}^n (X_i - \mu)^2}{\sigma_0^2}$	n	$\begin{cases} \chi_{\alpha/2}^2 \\ \chi_{1-\alpha/2}^2 \end{cases}$ χ_{α}^2 $\chi_{1-\alpha}^2$

注：上表中 H_0 中的不等号改成等号，所得的拒绝域不变。

第三节 两个正态总体的假设检验

上一节介绍了单个正态总体的数学期望与方差的检验问题，在实际工作中还常碰到两个正态总体的参数假设检验。与单个正态总体的参数假设检验不同的是，我们所关注的不是某个单一参数的值的假设检验，而是着重考虑两个总体之间的差异，即两个正态总体参数比较的假设检验问题。

1. 两正态总体数学期望假设检验

(1) 方差已知, 关于数学期望的假设检验 (Z 检验法). 设 $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$, 且 X, Y 相互独立, σ_1^2 与 σ_2^2 已知, 要检验的是

$$H_0: \mu_1 = \mu_2; H_1: \mu_1 \neq \mu_2. \text{ (双边检验)}$$

怎样寻找检验用的统计量呢? 从总体 X 与 Y 中分别抽取容量为 n_1, n_2 的样本 X_1, X_2, \dots, X_{n_1} 及 Y_1, Y_2, \dots, Y_{n_2} , 由于

$$\bar{X} \sim N\left(\mu_1, \frac{\sigma_1^2}{n_1}\right), \quad \bar{Y} \sim N\left(\mu_2, \frac{\sigma_2^2}{n_2}\right),$$

$$E(\bar{X} - \bar{Y}) = E(\bar{X}) - E(\bar{Y}) = \mu_1 - \mu_2,$$

$$D(\bar{X} - \bar{Y}) = D(\bar{X}) + D(\bar{Y}) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2},$$

因此随机变量 $\bar{X} - \bar{Y}$ 也服从正态分布, 即

$$\bar{X} - \bar{Y} \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right).$$

从而

$$\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{(\sigma_1^2/n_1) + (\sigma_2^2/n_2)}} \sim N(0, 1).$$

于是我们按如下步骤判断.

$$(a) \text{ 选取统计量 } Z = \frac{\bar{X} - \bar{Y}}{\sqrt{(\sigma_1^2/n_1) + (\sigma_2^2/n_2)}}, \quad (8-16)$$

当 H_0 为真时, $Z \sim N(0, 1)$.

(b) 对于给定的显著性水平 α , 查标准正态分布表求 $z_{\alpha/2}$ 使

$$P\{|Z| > z_{\alpha/2}\} = \alpha, \text{ 或 } P\{Z \leq z_{\alpha/2}\} = 1 - \alpha/2. \quad (8.17)$$

(c) 由两个样本观察值计算 Z 的观察值 z_0 :

$$z_0 = \frac{\bar{x} - \bar{y}}{\sqrt{(\sigma_1^2/n_1) + (\sigma_2^2/n_2)}}.$$

(d) 作出判断:

若 $|z_0| > z_{\alpha/2}$, 则拒绝假设 H_0 , 接受 H_1 ;

若 $|z_0| \leq z_{\alpha/2}$, 则与 H_0 相容, 可以接受 H_0 .

例 8.7 A, B 两台车床加工同一种轴, 现在要测量轴的椭圆度. 设 A 车床加工的轴的椭圆度 $X \sim N(\mu_1, \sigma_1^2)$, B 车床加工的轴的椭圆度 $Y \sim N(\mu_2, \sigma_2^2)$, 且 $\sigma_1^2 = 0.0006(\text{mm}^2)$, $\sigma_2^2 = 0.0038(\text{mm}^2)$, 现从 A, B 两台车床加工的轴中分别测量了 $n_1 = 200$, $n_2 = 150$ 根轴的椭圆度, 并计算得样本均值分别为 $\bar{x} = 0.081(\text{mm})$, $\bar{y} = 0.060(\text{mm})$. 试问这两台车床加工的轴的椭圆度是否有显著性差异 (给定 $\alpha = 0.05$)?

解 提出假设 $H_0: \mu_1 = \mu_2; H_1: \mu_1 \neq \mu_2$.

选取统计量

$$Z = \frac{\bar{X} - \bar{Y}}{\sqrt{(\sigma_1^2/n_1) + (\sigma_2^2/n_2)}},$$

在 H_0 为真时, $Z \sim N(0, 1)$.

给定 $\alpha=0.05$, 因为是双边检验, $\alpha/2=0.025$.

$$P\{|Z| > z_{\alpha/2}\} = 0.05, \quad P\{Z > z_{\alpha/2}\} = 0.025,$$

$$P\{Z \leq z_{\alpha/2}\} = 1 - 0.025 = 0.975.$$

查标准正态分布表, 得

$$z_{\alpha/2} = z_{0.025} = 1.96.$$

计算统计量 Z 的观察值 z

$$z_0 = \frac{\bar{x} - \bar{y}}{\sqrt{(\sigma_1^2/n_1) + (\sigma_2^2/n_2)}} = \frac{0.081 - 0.060}{\sqrt{(0.0006/200) + (0.0038/150)}} = 3.95.$$

作判断: 由于 $|z_0| = 3.95 > 1.96 = z_{\alpha/2}$, 因此拒绝 H_0 , 即在显著性水平 $\alpha=0.05$ 下, 认为两台车床加工的轴的椭圆度有显著差异.

用 Z 检验法对两正态总体的均值作假设检验时, 必须知道总体的方差, 但在许多实际问题中总体方差 σ_1^2 与 σ_2^2 往往是未知的, 这时只能用如下的 t 检验法.

(2) 方差 σ_1^2 , σ_2^2 未知, 关于均值的假设检验 (t 检验法). 设两正态总体 X 与 Y 相互独立, $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$, σ_1^2 , σ_2^2 未知, 但知 $\sigma_1^2 = \sigma_2^2$, 检验假设

$$H_0: \mu_1 = \mu_2; H_1: \mu_1 \neq \mu_2. \text{ (双边检验)}$$

从总体 X, Y 中分别抽取样本 X_1, X_2, \dots, X_{n_1} 与 Y_1, Y_2, \dots, Y_{n_2} , 则随机变量

$$t = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_w \sqrt{(1/n_1) + (1/n_2)}} \sim t_{(n_1+n_2-2)},$$

式中 $S_w^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}$, S_1^2 , S_2^2 分别是 X 与 Y 的样本方差.

当假设 H_0 为真时, 统计量

$$t = \frac{\bar{X} - \bar{Y}}{S_w \sqrt{(1/n_1) + (1/n_2)}} \sim t_{(n_1+n_2-2)}. \quad (8-18)$$

对给定的显著性水平 α , 查 t 分布得 $t_{\alpha/2}(n_1+n_2-2)$, 使得

$$P\{|t| > t_{\alpha/2}(n_1+n_2-2)\} = \alpha. \quad (8-19)$$

再由样本观察值计算 t 的观察值

$$t_0 = \frac{\bar{x} - \bar{y}}{S_w \sqrt{(1/n_1) + (1/n_2)}}, \quad (8-20)$$

最后作出判断:

若 $|t_0| > t_{\alpha/2}(n_1+n_2-2)$, 则拒绝 H_0 ;

若 $|t_0| \leq t_{\alpha/2}(n_1+n_2-2)$, 则接受 H_0 .

例 8.8 在一台自动车床上加工直径为 2.050mm 的轴, 现在每相隔两小时, 各取容量都为 10 的样本, 所得数据列表如表 8-3 所示.

表 8-3

零件加工编号	1	2	3	4	5	6	7	8	9	10
第一个样本	2.066	2.063	2.068	2.060	2.067	2.063	2.059	2.062	2.062	2.060
第二个样本	2.063	2.060	2.057	2.056	2.059	2.058	2.062	2.059	2.059	2.057

假设直径的分布是正态的, 由于样本是取自同一台车床, 可以认为 $\sigma_1^2 = \sigma_2^2 = \sigma^2$, 而 σ^2 是未知常数. 问这台自动车床的工作是否稳定 (取 $\alpha=0.01$)?

解 这里实际上是已知 $\sigma_1^2 = \sigma_2^2 = \sigma^2$, 但 σ^2 未知的情况下检验假设 $H_0: \mu_1 = \mu_2$; $H_1: \mu_1 \neq \mu_2$. 我们用 t 检验法, 由样本观察值算得:

$$\begin{aligned}\bar{x} &= 2.063, & \bar{y} &= 2.059, \\ s_1^2 &= 0.00000956, & s_2^2 &= 0.00000489, \\ s_w^2 &= \frac{9 \times s_1^2 + 9 \times s_2^2}{10 + 10 - 2} = \frac{0.0000086 + 0.0000044}{18} = 0.0000072.\end{aligned}$$

由 (8-20) 式计算得

$$t_0 = \frac{2.063 - 2.059}{\sqrt{0.0000072 \times (2/10)}} = 3.3.$$

对于 $\alpha=0.01$, 查自由度为 18 的 t 分布表得 $t_{0.005}(18)=2.878$. 由于 $|t_0| = 3.3 > t_{0.005}(18)=2.878$, 于是拒绝原假设 $H_0: \mu_1 = \mu_2$. 这说明两个样本在生产上是有差异的, 可能这台自动车床受时间的影响而生产不稳定.

2. 两正态总体方差的假设检验 (F 检验法 (F -test))

(1) 双边检验. 设两正态总体 $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$, X 与 Y 独立, X_1, X_2, \dots, X_{n_1} 与 Y_1, Y_2, \dots, Y_{n_2} 分别是来自这两个总体的样本, 且 μ_1 与 μ_2 未知. 现在要检验假设 $H_0: \sigma_1^2 = \sigma_2^2$; $H_1: \sigma_1^2 \neq \sigma_2^2$.

在原假设 H_0 成立下, 两个样本方差的比应该在 1 附近随机地摆动, 所以这个比不能太大又不能太小. 于是我们选取统计量

$$F = \frac{S_1^2}{S_2^2}. \quad (8-21)$$

显然, 只有当 F 接近 1 时, 才认为有 $\sigma_1^2 = \sigma_2^2$.

由于随机变量 $F^* = \frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2} \sim F(n_1-1, n_2-1)$, 所以当假设 $H_0: \sigma_1^2 = \sigma_2^2$ 成立时, 统计量

$$F = \frac{S_1^2}{S_2^2} \sim F(n_1-1, n_2-1).$$

对于给定的显著性水平 α , 可以由 F 分布表求得临界值

$$F_{1-\frac{\alpha}{2}}(n_1-1, n_2-1) \text{ 与 } F_{\alpha/2}(n_1-1, n_2-1)$$

使得

$$P\{F_{1-\frac{\alpha}{2}}(n_1-1, n_2-1) \leq F \leq F_{\alpha/2}(n_1-1, n_2-1)\} = 1 - \alpha$$

(图 8-5), 由此可知 H_0 的接受区域为

$$F_{1-\frac{\alpha}{2}}(n_1-1, n_2-1) \leq F \leq F_{\alpha/2}(n_1-1, n_2-1);$$

而 H_0 的拒绝域为

$$F < F_{1-\frac{\alpha}{2}}(n_1-1, n_2-1),$$

或 $F > F_{\alpha/2}(n_1-1, n_2-1).$

然后, 根据样本观察值计算统计量 F 的观察值, 若 F 的观察值落在拒绝域中, 则拒绝 H_0 , 接受 H_1 ; 若 F 的观察值落在接受域中, 则接受 H_0 .

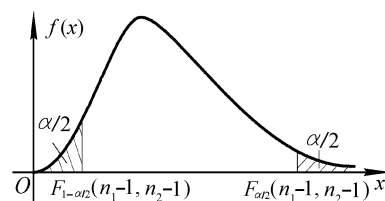


图 8-5

例 8.9 在例 8.8 中我们认为两个总体的方差 $\sigma_1^2 = \sigma_2^2$, 它们是否真的相等呢? 为此我们来检验假设 $H_0: \sigma_1^2 = \sigma_2^2$ (给定 $\alpha=0.1$).

解 这里 $n_1=n_2=10$, $s_1^2=0.00000956$, $s_2^2=0.00000489$, 于是统计量 F 的观察值为

$$F=0.00000956/0.00000489=1.95.$$

查 F 分布表得

$$F_{\alpha/2}(n_1-1, n_2-1) = F_{0.05}(9,9)=3.18,$$

$$F_{1-\alpha/2}(n_1-1, n_2-1) = F_{0.95}(9,9)=1/F_{0.05}(9,9)=1/3.18.$$

由样本观察值算出的 F 满足

$$F_{0.95}(9,9)=1/3.18 < F=1.95 < 3.18 = F_{0.05}(9,9).$$

可见它没有落入拒绝域, 因此不能拒绝原假设 $H_0: \sigma_1^2 = \sigma_2^2$, 从而认为两个总体的方差无显著差异.

注意: 在 μ_1 与 μ_2 已知时, 要检验假设 $H_0: \sigma_1^2 = \sigma_2^2$, 其检验方法类似于均值未知的情况, 只是所采用的检验统计量是

$$F = \frac{\frac{1}{n_1} \sum_{i=1}^{n_1} (X_i - \mu_1)^2}{\frac{1}{n_2} \sum_{i=1}^{n_2} (Y_i - \mu_2)^2} \sim F(n_1, n_2).$$

关于两个正态总体的假设检验了列表 8-4.

表 8-4

检 条	H_0	H_1	H_0 的拒	检验用的	自由	分位点
-----	-------	-------	----------	------	----	-----

验 参 数	件			绝域	统计量	度	
均 值	σ_1^2, σ_2^2 已知	$\mu_1 = \mu_2$ $\mu_1 \leq \mu_2$ $\mu_1 \geq \mu_2$	$\mu_1 \neq \mu_2$ $\mu_1 > \mu_2$ $\mu_1 < \mu_2$	$ Z > z_{\alpha/2}$ $Z > z_{\alpha}$ $Z < -z_{\alpha}$	$Z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$		$\pm z_{\alpha/2}$ z_{α} $-z_{\alpha}$
	σ_1^2, σ_2^2 未知 $\sigma_1^2 = \sigma_2^2$	$\mu_1 = \mu_2$ $\mu_1 \leq \mu_2$ $\mu_1 \geq \mu_2$	$\mu_1 \neq \mu_2$ $\mu_1 > \mu_2$ $\mu_1 < \mu_2$	$ t > t_{\alpha}$ $t > t_{\alpha}$ $t < -t_{\alpha}$	$t = \frac{\bar{X} - \bar{Y}}{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$	$n_1 + n_2 - 2$	$\pm t_{\alpha/2}$ t_{α} $-t_{\alpha}$
方 差	μ_1, μ_2 未知	$\sigma_1^2 = \sigma_2^2$ $\sigma_1^2 \leq \sigma_2^2$ $\sigma_1^2 \geq \sigma_2^2$	$\sigma_1^2 \neq \sigma_2^2$ $\sigma_1^2 > \sigma_2^2$ $\sigma_1^2 < \sigma_2^2$	$F > F_{\alpha/2}$ 或 $F < F_{1-\alpha/2}$ $F > F_{\alpha}$ $F < F_{1-\alpha}$	$F = \frac{S_1^2}{S_2^2}$	$(n_1 - 1, n_2 - 1)$	$F_{\alpha/2}$ 或 $F_{1-\alpha/2}$ F_{α} $F < F_{1-\alpha}$
	μ_1, μ_2 已知	$\sigma_1^2 = \sigma_2^2$ $\sigma_1^2 \leq \sigma_2^2$ $\sigma_1^2 \geq \sigma_2^2$	$\sigma_1^2 \neq \sigma_2^2$ $\sigma_1^2 > \sigma_2^2$ $\sigma_1^2 < \sigma_2^2$	$F > F_{\alpha/2}$ 或 $F < F_{1-\alpha/2}$ $F > F_{\alpha}$ $F < F_{1-\alpha}$	$F = \frac{\frac{1}{n_1} \sum_{i=1}^{n_1} (X_i - \mu_1)^2}{\frac{1}{n_2} \sum_{i=1}^{n_2} (X_i - \mu_2)^2}$	(n_1, n_2)	$F_{\alpha/2}$ 或 $F_{1-\alpha/2}$ F_{α} $F < F_{1-\alpha}$

(2) 单边检验.可作类似的讨论,限于篇幅,这里不作介绍了.

第四节 总体分布函数的假设检验

上两节中,我们在总体分布形式为已知的前提下,讨论了参数的检验问题.然而在实际问题中,有时不能确知总体服从什么类型的分布,此时就要根据样本来检验关于总体分布的假设.例如检验假设:“总体服从正态分布”等.本节仅介绍 χ^2 检验法.

所谓 χ^2 检验法是在总体的分布为未知时,根据样本值 x_1, x_2, \dots, x_n 来检验关于总体分布的假设

H_0 : 总体 X 的分布函数为 $F(x)$;

H_1 : 总体 X 的分布函数不是 $F(x)$ (8-22)

的一种方法(这里的备择假设 H_1 可不必写出).

注意,若总体 X 为离散型随机变量,则假设(8-22)相当于

H_0 : 总体 X 的分布律为 $P\{X=x_i\}=p_i, i=1, 2, \dots$; (8-23)

若总体 X 为连续型随机变量,则假设(8-22)相当于

$$H_0: \text{总体 } X \text{ 的概率密度为 } f(x). \quad (8-24)$$

在用 χ^2 检验法检验假设 H_0 时, 若在假设 H_0 下 $F(x)$ 的形式已知, 而其参数值未知, 此时需先用最大似然估计法估计参数, 然后再作检验.

χ^2 检验法的基本思想与方法如下:

(1) 将随机试验可能取值的全体 Ω 分为 k 个互不相容的子集 A_1, A_2, \dots, A_k ($\bigcup_{i=1}^k A_i = \Omega$,

$A_i A_j = \emptyset, i \neq j; i, j = 1, 2, \dots, k$), 我们仍然用 A_i ($i = 1, 2, \dots, k$) 表示事件 $\{X \in A\}$, 即随机变量 X 的取值落在子集 A 中的事件, 于是在 H_0 为真时, 可以计算概率

$$\hat{p}_i = P(A_i) \quad (i = 1, 2, \dots, k).$$

(2) 寻找用于检验的统计量及相应的分布, 在 n 次试验中, 用 f_i ($i = 1, 2, \dots, k$) 表示随机变量 X 的取值落在子集 A_i 中的次数. 事件 A_i 出现的频率 $\frac{f_i}{n}$ 与概率 \hat{p}_i 往往有差异, 但由大数定律可以知道, 如果样本容量 n 较大 (一般要求 n 至少为 50, 最好在 100 以上), 在 H_0 成立条件下 $\left(\frac{f_i}{n} - \hat{p}_i\right)^2$ 的值应该比较小, 基于这种想法, 皮尔逊使用

$$\chi^2 = \sum_{i=1}^k \frac{(f_i - n\hat{p}_i)^2}{n\hat{p}_i} \quad (8-25)$$

作为检验 H_0 的统计量, 并证明了如下的定理.

定理 8.1 若 n 充分大 ($n \geq 50$), 则当 H_0 为真时 (不论 H_0 中的分布属什么分布), 统计量 (8-25) 总是近似地服从自由度为 $k-r-1$ 的 χ^2 分布, 其中 r 是被估计的参数的个数.

(3) 对于给定的检验水平 α , 查表确定临界值 $\chi_{\alpha}^2(k-r-1)$ 使

$$P\{\chi^2 > \chi_{\alpha}^2(k-r-1)\} = \alpha,$$

从而得到 H_0 的拒绝域为

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

(4) 由样本值 x_1, x_2, \dots, x_n 计算 χ^2 的值, 并与 $\chi_{\alpha}^2(k-r-1)$ 比较.

(5) 作结论: 若 $\chi^2 > \chi_{\alpha}^2(k-r-1)$, 则拒绝 H_0 , 即不能认为总体分布函数为 $F(x)$; 否则接受 H_0 .

例 8.10 一本书的一页中印刷错误的个数 X 是一个随机变量, 现检查了一本书的 100 页, 记录每页中印刷错误的个数, 其结果如表 8-5 所示. 其中 f_i 是观察到有 i 个错误的页数. 问能否认为一页书中的错误个数 X 服从泊松分布 (取 $\alpha = 0.05$)?

表 8-5

错误个数 i	0	1	2	3	4	5	6	≥ 7
页数 f_i	36	40	19	2	0	2	1	0
A_i	A_0	A_1	A_2	A_3	A_4	A_5	A_6	A_7

解 由题意首先提出假设:

H_0 : 总体 X 服从泊松分布.

$$P\{X=i\} = \frac{e^{-\lambda} \lambda^i}{i!}, \quad i=0, 1, 2, \dots,$$

这里 H_0 中参数 λ 为未知, 所以需先来估计参数. 由最大似然估计法, 得

$$\hat{\lambda} = \bar{x} = \frac{0 \times 36 + 1 \times 40 + \dots + 6 \times 1 + 7 \times 0}{100} = 1.$$

将试验结果的全体分为 A_0, A_1, \dots, A_7 两两不相容的事件. 若 H_0 为真, 则 $P\{X=i\}$ 有估计

$$\hat{p} = \hat{P}\{X=i\} = \frac{e^{-1} 1^i}{i!} = \frac{e^{-1}}{i!}, \quad i=0, 1, 2, \dots.$$

例如

$$\hat{p}_0 = \hat{P}\{X=0\} = e^{-1},$$

$$\hat{p}_1 = \hat{P}\{X=1\} = e^{-1},$$

$$\hat{p}_2 = \hat{P}\{X=2\} = \frac{e^{-1}}{2},$$

.....

$$\hat{p}_7 = \hat{P}\{X \geq 7\} = 1 - \sum_{i=0}^6 \hat{p}_i = 1 - \sum_{i=1}^6 \frac{e^{-1}}{i!}.$$

计算结果如表 8-6 所示. 将其中有些 $np_i < 5$ 的组予以适当合并, 使新的每一组内有 $np_i \geq 5$, 如表 8-6 所示, 此处并组后 $k=4$, 但因在计算概率时, 估计了一个未知参数 λ , 故

$$\chi^2 = \sum_{i=1}^4 \frac{(f_i - n\hat{p}_i)^2}{n\hat{p}_i} \sim \chi^2(4-1-1).$$

计算结果为 $\chi^2 = 1.460$ (表 8-6). 因为 $\chi_{\alpha}^2(4-1-1) = \chi_{0.05}^2(2) = 5.991 > 1.46$, 所以在显著性水平为 0.05 下接受 H_0 , 即认为总体服从泊松分布.

表 8-6

A_i	f_i	\hat{p}_i	$n\hat{p}_i$	$f_i - n\hat{p}_i$	$(f_i - n\hat{p}_i)^2 / n\hat{p}_i$
A_0	36	e^{-1}	36.788	-0.788	0.017
A_1	40	e^{-1}	36.788	3.212	0.280
A_2	19	$e^{-1}/2$	18.394	0.606	0.020
A_3	2	$e^{-1}/6$	6.131		
A_4	0	$e^{-1}/24$	1.533		

A_5	2	$e^{-1}/120$	0.307	-3.03	1.143
A_6	1	$e^{-1}/720$	0.051		
A_7	0	$1 - \sum_{i=1}^6 \hat{p}_i$	0.008		
Σ	1.460				

例 8.11 研究混凝土抗压强度的分布.200 件混凝土制件的抗压强度以分组形式列出(见

表 8-7) . $n = \sum_{i=1}^6 f_i = 200$.要求在给定的显著性水平 $\alpha = 0.05$ 下检验假设

$$H_0: \text{抗压强度 } X \sim N(\mu, \sigma^2).$$

表 8-7

压强区间($\times 98\text{kPa}$)	频数 f_i
190~200	10
200~210	26
210~220	56
220~230	64
230~240	30
240~250	14

解 原假设所定的正态分布的参数是未知的, 我们需先求 μ 与 σ^2 的最大似然估计值. 由第七章知, μ 与 σ^2 的最大似然估计值为

$$\hat{\mu} = \bar{x},$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2.$$

设 x_i^* 为第 i 组的组中值, 我们有

$$\bar{x} = \frac{1}{n} \sum_i x_i^* f_i = \frac{195 \times 10 + 205 \times 26 + \cdots + 245 \times 14}{200} = 221,$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_i (x_i^* - \bar{x})^2 f_i = \frac{1}{200} \{(-26)^2 \times 10 + (-16)^2 \times 26 + \cdots + 24^2 \times 14\} = 152,$$

$$\hat{\sigma} = 12.33.$$

原假设 H_0 改写成 X 是正态 $N(221, 12.33^2)$ 分布, 计算每个区间的理论概率值

$$\hat{p}_i = P\{a_{i-1} \leq X < a_i\} = \Phi(\mu_i) - \Phi(\mu_{i-1}), \quad i=1, 2, \cdots, 6,$$

其中

$$\mu_i = \frac{a_i - \bar{x}}{\hat{\sigma}},$$

$$\Phi(\mu_i) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\mu_i} e^{-\frac{t^2}{2}} dt.$$

为了计算出统计量 χ^2 的值，我们把需要进行的计算列表如下（见表 8-8）。

表 8-8

压强区间 X	频数 f_i	标准化区间 [μ_i, μ_{i+1}]	$\hat{p} = \Phi(\mu_{i+1}) - \Phi(\mu_i)$	$n\hat{p}_i$	$(f_i - n\hat{p}_i)^2$	$\frac{(f_i - n\hat{p}_i)^2}{n\hat{p}_i}$
190~200	10	$(-\infty, -1.70)$	0.045	9	1	0.11
200~210	26	$[-1.70, -0.89)$	0.142	28.4	5.76	0.20
210~220	56	$[-0.89, -0.08)$	0.281	56.2	0.04	0.00
220~230	64	$[-0.08, 0.73)$	0.299	59.8	17.64	0.29
230~240	30	$[0.73, 1.54)$	0.171	34.2	17.64	0.52
240~250	14	$[1.54, +\infty)$	0.062	12.4	2.56	0.23
Σ			1.000	200		1.35

从上面计算得出 χ^2 的观察值为 1.33. 在检验水平 $\alpha=0.05$ 下，查自由度 $m=6-2-1=3$ 的 χ^2 分布表，得到临界值 $\chi_{0.05}^2(3)=7.815$. 因为 $\chi^2=1.33 < 7.815 = \chi_{0.05}^2(3)$ ，不能拒绝原假设，所以认为混凝土制件的抗压强度的分布是正态分布 $N(221, 152)$ 。

小 结

有关总体分布的未知参数或未知分布形式的种种推断叫做统计假设. 一般统计假设分为原假设 H_0 （在实际问题中至关重要的假设）及与原假设 H_0 对立假设即是备择假设 H_1 . 假设检验就是人们根据样本提供的信息作出“接受 H_0 、拒绝 H_1 ”或“拒绝 H_0 、接受 H_1 ”的判断。

假设检验的思想是小概率原理，即小概率事件在一次试验中几乎不会发生. 这种原理是人们处理实际问题中公认的原则。

由于样本的随机性，当 H_0 为真时，我们可能会作出拒绝 H_0 、接受 H_1 的错误判断（弃真错误）或当 H_0 不真时，我们可能会作出接受 H_0 、拒绝 H_1 的错误判断（取伪错误），见表 8-9。

表 8-9

假设检验的两类错误		
真实情况（未知）	所作决策	
	接受 H_0	拒绝 H_0
H_0 为真	正确	犯第一类错误
H_0 不真	犯第二类错误	正确

当样本容量 n 固定时，我们无法同时控制犯二类错误，即减小犯第一类错误的概率，就会增大犯第二类错误的概率，反之亦然. 在假设检验中我们主要控制（减小）犯第一类错误的概率. 使 $P\{\text{拒绝 } H_0 | H_0 \text{ 为真}\} \leq \alpha$ ，其中 α 很小. ($0 < \alpha < 1$)， α 称为检验的显著性水平，这种只对犯第一类错误的概率加以控制而不考虑犯第二类错误的概率的检验称为显著性假设检验。

单个、两个正态总体的均值、方差的假设检验是本章重点问题，读者需掌握 Z 检验法、

χ^2 检验法、 t 检验法等.这些检验法中原假设 H_0 备择假设 H_1 及 H_0 的拒绝域分别见表 8-2、表 8-4.

重要术语及主题

原假设	备择假设	检验统计量	单边检验	双边检验
显著性水平	拒绝域	显著性检验	一个正态总体的参数的检验	
两个正态总体均值差、方差比的检验			总体分布函数的假设检验	

习 题 八

1. 已知某炼铁厂的铁水含碳量在正常情况下服从正态分布 $N(4.55, 0.108^2)$. 现在测了 5 炉铁水, 其含碳量 (%) 分别为

4.28 4.40 4.42 4.35 4.37

问若标准差不改变, 总体平均值有无显著性变化 ($\alpha=0.05$)?

2. 某种矿砂的 5 个样品中的含镍量 (%) 经测定为:

3.24 3.26 3.24 3.27 3.25

设含镍量服从正态分布, 问在 $\alpha=0.01$ 下能否接收假设: 这批矿砂的含镍量为 3.25?

3. 在正常状态下, 某种牌子的香烟一支平均 1.1 克, 若从这种香烟堆中任取 36 支作为样本; 测得样本均值为 1.008 (g), 样本方差 $s^2=0.1(\text{g}^2)$. 问这堆香烟是否处于正常状态? 已知香烟 (支) 的重量 (单位: g) 近似服从正态分布 (取 $\alpha=0.05$).

4. 某公司宣称由他们生产的某种型号的电池其平均寿命为 21.5 小时, 标准差为 2.9 小时. 在实验室测试了该公司生产的 6 只电池, 得到它们的寿命 (以 h 计) 为 19, 18, 20, 22, 16, 25, 问这些结果是否表明这种电池的平均寿命比该公司宣称的平均寿命要短? 设电池寿命近似地服从正态分布 (取 $\alpha=0.05$).

5. 测量某种溶液中的水分, 从它的 10 个测定值得出 $\bar{x}=0.452(\%)$, $s=0.037(\%)$. 设测定值总体为正态, μ 为总体均值, σ 为总体标准差, 试在水平 $\alpha=0.05$ 下检验.

(1) $H_0: \mu=0.5(\%); H_1: \mu < 0.5(\%)$.

(2) $H'_0: \sigma=0.04(\%); H'_1: \sigma < 0.04(\%)$.

6. 某种导线的电阻服从正态分布 $N(\mu, 0.005^2)$. 今从新生产的一批导线中抽取 9 根, 测其电阻, 得 $s=0.008 \Omega$. 对于 $\alpha=0.05$, 能否认为这批导线电阻的标准差仍为 0.005?

7. 有两批棉纱, 为比较其断裂强度 (单位: kg), 从中各取一个样本, 测试得到:

第一批棉纱样本: $n_1=200$, $\bar{x}=0.532\text{kg}$, $s_1=0.218\text{kg}$;

第二批棉纱样本: $n_2=200$, $\bar{y}=0.57\text{kg}$, $s_2=0.176\text{kg}$.

设两强度总体服从正态分布, 方差未知但相等, 两批强度均值有无显著差异 (取 $\alpha=0.05$)?

8. 两位化验员 A, B 对一种矿砂的含铁量各自独立地用同一方法做了 5 次分析, 得到样本方差分别为 0.4322($\%^2$)与 0.5006($\%^2$). 若 A, B 所得的测定值的总体都是正态分布, 其方差分别为 σ_A^2 , σ_B^2 , 试在水平 $\alpha=0.05$ 下检验方差齐性的假设:

$$H_0: \sigma_A^2 = \sigma_B^2; \quad H_1: \sigma_A^2 \neq \sigma_B^2.$$

9. 在 π 的前 800 位小数的数字中, 0, 1, ..., 9 相应的出现了 74, 92, 83, 79, 80, 73, 77,

75, 76, 91 次. 试用 χ^2 检验法检验假设

$$H_0: P(X=0) = P(X=1) = P(X=2) = \cdots = P(X=9) = 1/10,$$

其中 X 为 π 的小数中所出现的数字, $\alpha=0.10$.

10.在一副扑克牌(52张)中任意抽3张,记录3张牌中含红桃的张数,放回,然后再任抽3张,如此重复64次,得到如表8-10所示结果,试在水平 $\alpha=0.01$ 下检验假设:

表 8-10

含红桃张数 Y	0	1	2	3
出现次数	21	31	12	0

H_0 : Y 服从二项分布,

$$P\{Y=i\} = C_3^i \left(\frac{1}{4}\right)^i \left(\frac{3}{4}\right)^{3-i}, \quad i=0, 1, 2, 3.$$

11.在某公路上,50min 之间,观察每15秒内过路的汽车的辆数,得到频数分布如表8-11所示:

表 8-11

过路的车辆数 X	0	1	2	3	4	5
次数 f_i	92	68	28	11	1	0

问这个分布能否认为是泊松分布 ($\alpha=0.10$)?

12.测得300只电子管的寿命(以小时计)如表8-12所示,试取水平 $\alpha=0.05$ 下的检验假设:

H_0 : 寿命 X 服从指数分布,其密度为

$$f(t) = \begin{cases} \frac{1}{200} e^{-\frac{t}{200}}, & t > 0, \\ 0, & \text{其他.} \end{cases}$$

表 8-12

寿命	只数
$0 < t \leq 100$	121
$100 < t \leq 200$	78
$200 < t \leq 300$	43
$t > 300$	58