

第九章 假设检验

在第八章, 我们讨论了未知参数 θ 的点估计和区间估计方法. 但这并不能解决完总体中有关未知参数 θ 的基本问题. 在例 7.1 和例 7.2 中, 我们分别提出了判断产品合格率 p 是否有 “ $p \geq 0.98$ ” 和彩电平均寿命 μ 是否有 “ $\mu > 5$ 万小时” 的问题. 解决这样的问题的方法是数理统计的另一个基本内容, 称为假设检验. 这一章, 我们就讨论假设检验的基本方法.

§9.1 假设检验的基本概念

§9.1.1 假设检验的基本思想

我们通过下面这个例子来说明假设检验的基本思想.

例 9.1 由于测量误差存在, 某测距仪测量值 X (单位: m) $\sim N(\mu, 10^2)$. 今对距离 500 m 的目标测量 9 次, 得到平均距离 $\bar{x} = 507$ m, 问该测距仪是否存在系统误差?

对这个问题可做如下分析:

- (1) 这不是一个参数估计问题;
- (2) 这是给定总体和样本下, 要求对问题 “该测距仪是否存在系统误差” 作出判断 “是” 还是 “否”. 这类问题称为统计假设检验问题.
- (3) 问题 “该测距仪是否存在系统误差” 的判断是对参数 μ 的判断. 这个判断涉及如下两个参数的集合:

$$\Theta_0 = \{\mu : \mu = 500\}, \quad \Theta_1 = \{\mu : \mu \neq 500\}$$

其中判断为不存在系统误差对应于 $\mu \in \Theta_0$, 判断存在系统误差对应于 $\mu \in \Theta_1$. 在统计学中这两个参数的非空集合都称作统计假设.

(4) 利用所给总体 $X \sim N(\mu, 10^2)$ 和样本均值 $\bar{x} = 507$ 去判断假设 “ $\mu \in \Theta_0$ ” 是否成立, 在统计学中被称为检验.

下面我们来分析这个问题的解决方法.

如果测距仪正常工作, 不存在系统误差, 则测量值 X 的期望 μ 应为 500 m. 而存在系统误差时, μ 就应不等于 500 m. 所以, 判断测距仪是否存在系统误差

等价于在测量值 $X \sim N(\mu, 10^2)$ 时判断是否有 “ $\mu = 500$ ”. 为此, 我们首先提出两个相互对立的假设:

$$H_0: \mu = \mu_0 = 500; \quad H_1: \mu \neq \mu_0. \quad (9.1.1)$$

H_0 称为原假设, H_1 称为备择假设. 这样, 判断测距仪是否存在系统误差又等价于判断 H_0 成立还是 H_1 成立.

这个问题的一般形式是: 当总体 $X \sim N(\mu, \sigma_0^2)$, 其中 σ_0^2 已知, 取得样本 X_1, X_2, \dots, X_n 及样本观测值 x_1, x_2, \dots, x_n 后, 要判断 $H_0: \mu = \mu_0, H_1: \mu \neq \mu_0$ 中哪一个成立.

这时, 我们掌握的样本数据为 $\bar{x} = 507$, 这与 $\mu_0 = 500$ 有差异. 我们的分析应从 $|\bar{x} - \mu_0|$ 入手.

注意样本均值 \bar{X} 是总体期望 μ 的无偏估计, 因此

(1) 若 H_0 成立, 即 $\mu = \mu_0$, 故 \bar{X} 与 μ_0 偏差的绝对值 $|\bar{X} - \mu_0|$ 取值应相对较小, 它的取值反映的是样本均值所产生的抽样误差.

(2) 若 H_1 成立, 即 $\mu \neq \mu_0$, 则 $|\bar{X} - \mu_0|$ 的取值应相对较大, 它的取值主要反映的是 $\mu \neq \mu_0$ 所产生的系统误差.

因此判断 H_0 成立还是 H_1 成立转化为判断 $|\bar{X} - \mu_0|$ 取值的相对大小. 而 $|\bar{X} - \mu_0|$ 取值的相对大小又与

$$|U| = \frac{|\bar{X} - \mu_0|}{\sigma_0/\sqrt{n}}$$

的相对大小一致, 从而就是判断 $|U|$ 取值的相对大小. 这就需要一个临界值, 以判断 $|U|$ 取值的相对大小. 下面来寻求这样一个临界值.

我们取一个很小的正数 α , 如 $\alpha = 0.05, \alpha = 0.01$. α 称为显著性水平. 并且假定 H_0 成立, 此时有

$$U = \frac{\bar{X} - \mu_0}{\sigma_0/\sqrt{n}} \sim N(0, 1), \quad (9.1.2)$$

从而有

$$P\left(|U| = \frac{|\bar{X} - \mu_0|}{\sigma_0/\sqrt{n}} > u_{1-\frac{\alpha}{2}}\right) = \alpha. \quad (9.1.3)$$

这样, 推出事件 $(|U| > u_{1-\frac{\alpha}{2}})$ 为一个小概率事件. 根据小概率事件在一次试验中是几乎不可能发生的小概率原理, 因此在 H_0 成立的假定下, 这个事件 $(|U| > u_{1-\frac{\alpha}{2}})$ 是几乎不可能发生的. 现在取显著性水平 $\alpha = 0.05$, 查表有 $u_{1-\frac{\alpha}{2}} = u_{0.975} = 1.96$, 抽样的结果是 $\bar{x} = 507$, 从而

$$|u| = \frac{|\bar{x} - \mu_0|}{\sigma_0/\sqrt{n}} = \frac{|507 - 500|}{10/\sqrt{9}} = 2.1 > 1.96 = u_{1-\frac{\alpha}{2}}.$$

这个小概率事件 ($|U| > u_{1-\frac{\alpha}{2}}$) 竟然发生了. 由小概率原理说明这个事件发生的概率并不小, 这与推导出该事件是小概率事件矛盾. 这个矛盾说明前面假定 H_0 成立是错误的, 应拒绝 H_0 成立, 从而接受备择假设 H_1 成立, 即认为这个测距仪测 500 m 处的物体平均测量值 $\mu \neq \mu_0 = 500$, 说明这个测距仪有系统误差.

还应该指出, 假设检验的结论与显著性水平 α 密切相关的. 在这个例子中, 若取 $\alpha = 0.01$, 则 $u_{1-\frac{\alpha}{2}} = u_{0.995} = 2.58$, 而

$$|u| = 2.1 < 2.58 = u_{1-\frac{\alpha}{2}}.$$

这样小概率事件 ($|U| > u_{1-\frac{\alpha}{2}}$) 没有发生. 所以没有理由拒绝原假设 H_0 成立, 则应该接受 H_0 成立, 从而可以认为该测距仪没有系统误差. 因此, 假设检验的结论必须说明在什么样的显著性水平 α 下得到的.

上面的推导方法反映了一般假设检验的思想方法. 假设检验又叫显著性检验, 其中 U 叫做检验统计量; $u_{1-\frac{\alpha}{2}}$ 叫做临界值; $W = \{|U| > u_{1-\frac{\alpha}{2}}\}$ 叫做检验的拒绝域. 一般地, 拒绝域在原假设 H_0 成立时是一个小概率事件. 当样本观测值代入检验统计量使得这个小概率事件发生了, 则应拒绝 H_0 成立, 从而接受 H_1 成立. 若样本观测值代入时这个小概率事件没有发生, 则不拒绝 H_0 成立, 即接受 H_0 成立.

上面的推导方法是一种概率的反证法, 是根据小概率事件在一次试验中几乎不可能发生的小概率原理引出矛盾. 与数学证明中逻辑的反证法是有区别的.

§9.1.2 双侧检验与单侧检验

在例 9.1 的检验中, 备择假设 H_1 为 $\mu \neq \mu_0$, μ 位于 μ_0 两端, 这样的检验叫做双侧检验. 注意此检验的拒绝域 $W = \{|U| > u_{1-\frac{\alpha}{2}}\}$ 也位于两端.

但实际情况中也常常遇到单侧检验问题. 如在例 9.1 中, 若我们已经知道该测距仪的测量值不会偏低, 则我们关心的是该测距仪是否有测量值偏大的系统误差. 这样就应该检验

$$H_0: \mu = \mu_0; \quad H_1: \mu > \mu_0. \quad (9.1.4)$$

因备择假设中 μ 位于 μ_0 右端, 所以这个检验叫做右侧检验.

因为 H_0 成立时, 应有

$$U = \frac{\bar{X} - \mu_0}{\sigma_0/\sqrt{n}} \sim N(0, 1). \quad (9.1.5)$$

当 H_1 成立时, 有 $\mu > \mu_0$, 注意 \bar{X} 是 μ 的无偏估计量, 从而 U 取值有偏大的倾向, 故拒绝域应在临界值的右边. 于是显著性水平 α 取定后, 当 H_0 成立时, 有

$$P(U > u_{1-\alpha}) = \alpha. \quad (9.1.6)$$

基于小概率原理的概率反证法, 知此时拒绝域为 $W = \{U > u_{1-\alpha}\}$.

本例中, 若取 $\alpha = 0.05$, 则 $u_{1-\alpha} = u_{0.95} = 1.645$, 而

$$u = \frac{\bar{x} - \mu_0}{\sigma_0/\sqrt{n}} = \frac{507 - 500}{10/\sqrt{9}} = 2.1 > 1.645 = u_{1-\alpha},$$

则应当拒绝 H_0 , 认为该测距仪有测量值偏大的系统误差.

实际情况中, 我们还可能遇到这样的假设检验

$$H_0: \mu \leq \mu_0, \quad H_1: \mu > \mu_0. \quad (9.1.7)$$

我们假设例 9.1 中的模型, 即总体 $X \sim N(\mu, \sigma_0^2)$, 来导出检验 (9.1.7) 的拒绝域.

我们仍取检验统计量为

$$U = \frac{\bar{X} - \mu_0}{\sigma_0/\sqrt{n}},$$

注意若 H_1 成立时, U 取值有偏大的倾向, 故拒绝域应位于临界值右端.

但在此模型下,

$$\frac{\bar{X} - \mu}{\sigma_0/\sqrt{n}} \sim N(0, 1).$$

对显著性水平 α 取定, 有

$$P\left(\frac{\bar{X} - \mu}{\sigma_0/\sqrt{n}} > u_{1-\alpha}\right) = \alpha.$$

假定 H_0 成立时, 有

$$U = \frac{\bar{X} - \mu_0}{\sigma_0/\sqrt{n}} \leq \frac{\bar{X} - \mu}{\sigma_0/\sqrt{n}},$$

故以下事件有包含关系

$$\left(U = \frac{\bar{X} - \mu_0}{\sigma_0/\sqrt{n}} > u_{1-\alpha}\right) \subset \left(\frac{\bar{X} - \mu}{\sigma_0/\sqrt{n}} > u_{1-\alpha}\right),$$

于是有

$$P\left(U = \frac{\bar{X} - \mu_0}{\sigma_0/\sqrt{n}} > u_{1-\alpha}\right) \leq P\left(\frac{\bar{X} - \mu}{\sigma_0/\sqrt{n}} > u_{1-\alpha}\right) = \alpha,$$

故检验 (9.1.5) 的拒绝域为 $W = \{U > u_{1-\alpha}\}$.

假设为 (9.1.7) 的检验也叫做右侧检验. 注意右侧检验中原假设为 $\mu = \mu_0$ 或 $\mu \leq \mu_0$ 时, 其检验统计量与拒绝域都相同. 故这两种原假设在理论上没有区分, 但在实际意义上有区分.

同理, 假设为

$$H_0: \mu = \mu_0 \text{ (或 } \mu \geq \mu_0); \quad H_1: \mu < \mu_0 \quad (9.1.8)$$

的检验叫做左侧检验. 并可知在总体 $X \sim N(\mu, \sigma_0^2)$, 显著性水平为 α 时, 检验统计量仍为 $U = \frac{\bar{X} - \mu_0}{\sigma_0/\sqrt{n}}$, 而拒绝域为 $W = \{U < -u_{1-\alpha}\}$.

右侧检验与左侧检验合称单侧检验.

§9.1.3 两类错误

由前所述, 拒绝域是依据小概率原理推出的. 但几乎不可能发生不等于一定不发生. 因此假设检验的结论有可能是错误的. 这种错误有两类.

第一类错误叫做“弃真”, 即当 H_0 确实是正确的, 但检验统计量的值却落入了拒绝域, 我们就拒绝了 H_0 . 这样, 就犯下了第一类错误.

第二类错误叫做“取伪”, 即当 H_0 确实是错误的, 但检验统计量的值未落入拒绝域, 从而结论是不拒绝 H_0 . 这样, 犯了第二类错误.

犯第一类错误的概率是不大于显著性水平 α 的. 如以前面讨论的右侧检验为例, 当 H_0 为真时, $\frac{\bar{X} - \mu}{\sigma_0/\sqrt{n}} \sim N(0, 1)$, 且 $\mu \leq \mu_0$, 故有

$$\begin{aligned} P(\text{拒绝 } H_0) &= P\left(U = \frac{\bar{X} - \mu_0}{\sigma_0/\sqrt{n}} > u_{1-\alpha}\right) \\ &\leq P\left(\frac{\bar{X} - \mu}{\sigma_0/\sqrt{n}} > u_{1-\alpha}\right) = \alpha. \end{aligned}$$

犯第二类错误的概率记为 β , 即当 H_1 为真时, $\frac{\bar{X} - \mu}{\sigma_0/\sqrt{n}} \sim N(0, 1)$, 且 $\mu > \mu_0$, 故

$$\begin{aligned} \beta &= P(\text{不拒绝 } H_0) \\ &= P\left(U = \frac{\bar{X} - \mu_0}{\sigma_0/\sqrt{n}} \leq u_{1-\alpha}\right). \end{aligned}$$

β 的值一般不易求出. 但 $E(U) = \frac{\mu - \mu_0}{\sigma_0/\sqrt{n}}$. 由图 9.1 中 $\mu = \mu_0$ 和 $\mu = \mu_1 > \mu_0$ 时 U 的密度曲线, 可知 $\alpha + \beta \neq 1$. 且 α 越小 β 越大, 反之亦然.

这样, 我们不能同时使犯两类错误的概率都减小. 一般的做法有两种. 第一种是取定显著性水平 α , 再增大样本容量 n , 使 β 减小. 第二种是根据实际问题看哪一类错误后果严重, 再选取 α 的大小. 如第一类错误后果严重, 则 α 可取小一些, 如 $\alpha = 0.01, \alpha = 0.05$. 如第二类错误后果严重, 则 α 可取大一些, 如 $\alpha = 0.05, \alpha = 0.1$ 等.

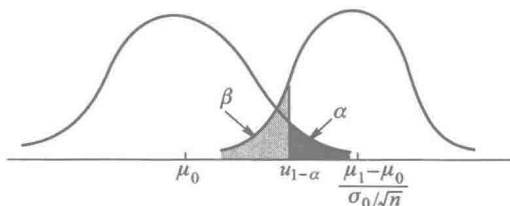


图 9.1

§9.1.4 假设检验的一般步骤

由前面的讨论, 我们以例 9.1 为例总结出假设检验的一般步骤:

(1) 根据实际问题, 提出原假设 H_0 和备择假设 H_1 . 在例 9.1 中, H_0 和 H_1 分别是

$$H_0: \mu = \mu_0 = 500, \quad H_1: \mu \neq \mu_0.$$

(2) 选取适当的检验统计量, 并在原假设 H_0 成立的条件下确定检验统计量的分布. 在例 9.1 中, 我们选取

$$U = \frac{\bar{X} - \mu_0}{\sigma_0 / \sqrt{n}}$$

为检验统计量. 在 H_0 成立时, 有

$$U = \frac{\bar{X} - \mu_0}{\sigma_0 / \sqrt{n}} \sim N(0, 1).$$

(3) 选取显著性水平 α , 并根据统计量的分布查表确定临界值, 从而得到检验的拒绝域. 在例 9.1 中, 取显著性水平 $\alpha = 0.05$, 查表求得临界值 $u_{1-\alpha/2} = u_{0.975} = 1.96$. 因此检验的拒绝域为 $W = \{|U| > u_{1-\alpha/2} = 1.96\}$.

(4) 根据样本观测值计算检验统计量的观测值, 并与临界值比较, 再对拒绝或接受原假设 H_0 做出判断.

在例 9.1 中, $\bar{x} = 507$, 从而

$$|u| = \frac{|\bar{x} - \mu_0|}{\sigma_0 / \sqrt{n}} = \frac{|507 - 500|}{10 / \sqrt{9}} = 2.1 > 1.96 = u_{1-\alpha/2},$$

结论是拒绝 H_0 成立.

§9.2 正态总体下参数的假设检验

在本节中, 我们讨论正态总体下参数的假设检验.

§9.2.1 一个正态总体下参数的假设检验

设总体 $X \sim N(\mu, \sigma^2)$, X_1, X_2, \dots, X_n 为来自总体 X 的样本. 分别记 \bar{X}, S^2 为样本均值和样本方差.

1. $\sigma^2 = \sigma_0^2$ 已知时, μ 的检验

当原假设为 $H_0: \mu = \mu_0$, 且 H_0 成立时, 由前面的讨论, 可知此时检验统计量

$$U = \frac{\bar{X} - \mu_0}{\sigma_0 / \sqrt{n}} \sim N(0, 1),$$

且对给定的显著性水平 α , 我们已经知道各种备择假设下检验的拒绝域 (见表 9.1). 这种检验因检验统计量为 U , 故叫做 u 检验.

2. σ^2 未知时, μ 的检验

当原假设为 $H_0: \mu = \mu_0$, 且 H_0 成立时, 由抽样分布定理 7.5, 知检验统计量

$$t = \frac{\bar{X} - \mu_0}{S / \sqrt{n}} \sim t(n-1), \quad (9.2.1)$$

当显著性水平 α 取定后, 注意到对各种备择假设下拒绝域的方向与备择假设方向一致, 与 σ^2 已知时推导的过程类似, 可得各种备择假设下检验的拒绝域 (表 9.1). 这种检验因检验统计量为 t , 故叫做 t 检验.

表 9.1 一个正态总体均值的假设检验表

H_0	H_1	σ^2 已知	σ^2 未知
		在显著性水平 α 下关于 H_0 的拒绝域	
$\mu = \mu_0$	$\mu \neq \mu_0$	$ U > u_{1-\frac{\alpha}{2}}$	$ t > t_{1-\frac{\alpha}{2}}(n-1)$
$\mu = \mu_0$	$\mu > \mu_0$	$U > u_{1-\alpha}$	$t > t_{1-\alpha}(n-1)$
$\mu = \mu_0$	$\mu < \mu_0$	$U < -u_{1-\alpha}$	$t < -t_{1-\alpha}(n-1)$

例 9.2 已知某种元件的使用寿命 (单位: h) 服从标准差为 $\sigma = 120$ h 的正态分布. 按要求, 该种元件的使用寿命不得低于 1800 h 才合格. 今从一批这种元件中随机抽取 36 件, 测得其平均寿命为 1750 h. 判断这批元件是否合格 ($\alpha = 0.05$)?

解 一件元件寿命不低于 1800 h 合格, 则一批元件应其平均寿命不低于 1800 h 才算合格, 则应检验假设

$$H_0: \mu \geq \mu_0 = 1800, \quad H_1: \mu < 1800.$$

这是一个左侧检验. 因元件寿命 $X \sim N(\mu, \sigma^2)$, 且 $\sigma = 120$ 已知, 故应用 u 检验.

$\alpha = 0.05$, 则查表知 $u_{1-\alpha} = u_{0.95} = 1.645$. 又已知 $\sigma = 120, n = 36, \bar{x} = 1750, \mu_0 = 1800$, 于是得检验统计量 U 的值

$$u = \frac{\bar{x} - \mu_0}{\sigma_0/\sqrt{n}} = \frac{1750 - 1800}{120/\sqrt{36}} = -2.5 < -u_{1-\alpha} = -1.645.$$

因 $u = -2.5$ 落入拒绝域, 从而拒绝 H_0 , 认为这批元件不合格.

例 9.3 某厂生产的一种型号的电阻元件其电阻值 X (单位: Ω) $\sim N(2.64, \sigma^2)$ 改变生产工艺后, 从生产线上随机取 10 个电阻测得其电阻值为

2.13, 2.42, 2.65, 2.74, 2.82, 2.62, 2.39, 2.76, 2.88, 2.54.

问新工艺对该电阻元件的电阻值有无显著影响 ($\alpha = 0.05$)?

解 若新工艺对电阻值无显著影响, 则工艺改革前后其电阻平均值应相等. 故考虑以下假设检验

$$H_0: \mu = \mu_0 = 2.64, \quad H_1: \mu \neq \mu_0.$$

因是正态总体下总体方差 σ^2 未知, 应选用 t 检验. 由样本观测值, 算得 $\bar{x} = 2.59, s = 0.23$.

又有 $\alpha = 0.05, n = 10$, 查表, 得 $t_{1-\frac{\alpha}{2}}(n-1) = t_{0.975}(9) = 2.2622$, 算得

$$|t| = \frac{|\bar{x} - \mu_0|}{s/\sqrt{n}} = \frac{|2.59 - 2.64|}{0.23/\sqrt{10}} = 0.6875.$$

因为 $|t| = 0.6875 < t_{1-\frac{\alpha}{2}}(n-1) = 2.2622$, 故不拒绝 H_0 , 认为新工艺对该电阻元件的电阻值无显著影响.

3. μ 未知时, σ^2 的检验

当原假设为 $H_0: \sigma^2 = \sigma_0^2$, 且 H_0 成立时, 由抽样分布定理 7.4, 知检验统计量

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2(n-1). \quad (9.2.2)$$

当显著性水平 α 取定后, 表 9.2 给出了各种备择假设下检验的拒绝域. 这种检验检验统计量为 χ^2 , 故叫做 χ^2 检验.

表 9.2 一个正态总体方差的 χ^2 检验

H_0	H_1	显著性水平 α 下关于 H_0 的拒绝域
$\sigma^2 = \sigma_0^2$	$\sigma^2 \neq \sigma_0^2$	$\chi^2 < \chi_{\frac{\alpha}{2}}^2(n-1)$ 或 $\chi^2 > \chi_{1-\frac{\alpha}{2}}^2(n-1)$
$\sigma^2 = \sigma_0^2$	$\sigma^2 > \sigma_0^2$	$\chi^2 > \chi_{1-\alpha}^2(n-1)$
$\sigma^2 = \sigma_0^2$	$\sigma^2 < \sigma_0^2$	$\chi^2 < \chi_{\alpha}^2(n-1)$

例 9.4 电池厂生产的某型号电池, 其使用寿命长期以来服从方差 $\sigma^2 = 100$ 的正态分布. 现对一批电池随机抽取 9 个, 测得寿命 (单位: h) 如下:

678, 670, 650, 680, 672, 612, 601, 605, 674.

问能否认为这批电池寿命的波动性较过去有显著的增大 ($\alpha = 0.05$)?

解 因方差是刻画随机变量取值的波动性. 因此这是正态总体下 μ 未知 σ^2 的检验. 对应假设为

$$H_0: \sigma^2 = \sigma_0^2 = 100, \quad H_1: \sigma^2 > \sigma_0^2.$$

由样本观测值算得 $s^2 = 1125.9$. 又 $\alpha = 0.05, n = 9$, 查表, 得 $\chi_{1-\alpha}^2(n-1) = \chi_{0.95}^2(8) = 15.5$, 又算得 χ^2 统计量值

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{8 \times 1125.9}{100} = 90.07 > 15.5 = \chi_{1-\alpha}^2(n-1),$$

故拒绝 H_0 , 认为这批电池寿命的波动性较过去有显著增大.

§9.2.2 两个正态总体下参数的假设检验

设总体 $X \sim N(\mu_1, \sigma_1^2)$, 总体 $Y \sim N(\mu_2, \sigma_2^2)$. 从两个总体中分别抽取两个独立样本 X_1, X_2, \dots, X_{n_1} 和 Y_1, Y_2, \dots, Y_{n_2} , 以 $\bar{X}, \bar{Y}, S_1^2, S_2^2$ 分别记两样本的样本均值和样本方差.

1. σ_1^2, σ_2^2 已知时, $\mu_1 = \mu_2$ 的检验

当原假设 $H_0: \mu_1 = \mu_2$ 成立时, 由抽样分布定理 7.7, 知此时检验统计量

$$U = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0, 1), \quad (9.2.3)$$

从而对各种备择假设有此时 u 检验的拒绝域 (表 9.3).

2. $\sigma_1^2 = \sigma_2^2 = \sigma^2$, 且 σ^2 未知时, $\mu_1 = \mu_2$ 的检验

当原假设 $H_0: \mu_1 = \mu_2$ 成立时, 由抽样分布定理 7.8, 知此时检验统计量

$$T = \frac{\bar{X} - \bar{Y}}{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2), \quad (9.2.4)$$

其中

$$S_w = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}},$$

从而对各种备择假设有此时 t 检验的拒绝域 (表 9.3).

表 9.3 两个正态总体均值的假设检验表

H_0	H_1	σ_1^2, σ_2^2 已知	σ_1^2, σ_2^2 未知, $\sigma_1^2 = \sigma_2^2$
		在显著性水平 α 下关于 H_0 的拒绝域	
$\mu_1 = \mu_2$	$\mu_1 \neq \mu_2$	$ U > u_{1-\frac{\alpha}{2}}$	$ T > t_{1-\frac{\alpha}{2}}(n_1 + n_2 - 2)$
$\mu_1 = \mu_2$	$\mu_1 > \mu_2$	$U > u_{1-\alpha}$	$T > t_{1-\alpha}(n_1 + n_2 - 2)$
$\mu_1 = \mu_2$	$\mu_1 < \mu_2$	$U < -u_{1-\alpha}$	$T < -t_{1-\alpha}(n_1 + n_2 - 2)$

例 9.5 在例 8.16 中, 请判断甲、乙两厂生产的蓄电池其电容量有无显著差别 ($\alpha = 0.05$).

解 现两个总体服从正态分布, 且 $\sigma_1^2 = \sigma_2^2 = \sigma^2$, 但 σ^2 未知. 其检验假设应为

$$H_0: \mu_1 = \mu_2; \quad H_1: \mu_1 \neq \mu_2.$$

由例 8.16, 知 $\bar{x} = 140.5, \bar{y} = 139.9, n_1 = 8, n_2 = 10, s_w = 2.36$. 且 $t_{1-\frac{\alpha}{2}}(n_1 + n_2 - 2) = t_{0.975}(16) = 2.1199$. 故

$$t = \frac{\bar{x} - \bar{y}}{s_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{140.5 - 139.9}{2.36 \times \sqrt{\frac{1}{8} + \frac{1}{10}}} = 0.5360.$$

因 $|t| = 0.5360 < t_{0.975}(16) = 2.1199$, 故不拒绝 H_0 , 认为甲、乙两厂生产的蓄电池其电容量无显著差别.

在例 8.16 和例 9.5 中, 其计算比较麻烦, 我们可采用统计软件进行计算, 如 SAS, SPSS 等统计软件. 本书第十一章介绍了 SPSS 软件的应用方法. 用 SPSS 软件解决例 8.16 和例 9.5 的问题, 输入数据后, 软件分析的结果有很多. 但关于 t 检验的分析结果为表 9.4:

表 9.4 t 检验分析结果

	Levene 方差齐性检验		两样本均值差的 t 检验						
	F	p 值	t	df	双边 p 值	均值差的 平均值	均值差的 标准差	均值差的 95% 置信区间	
								下界	上界
z	0.503	0.488	0.537	16	0.599	0.600 00	1.118 10	-1.770 27	2.970 27
			0.527	13.853	0.607	0.600 00	1.139 34	-1.846 08	3.046 08

由此分析结果, 知 $\bar{x} - \bar{y} = 0.6; s_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = 1.1181$; 统计量值 $t = 0.537$; 自由度 $n_1 + n_2 - 2 = 16$, 概率 $P(|t| > 0.537) = 0.599$. 由这些分析结果, 可得到检验的结论. 另外, Levene 方差齐性检验是检验 σ_1^2 是否等于 σ_2^2 . 其结论是相等.

此表检验结果第二行是当 σ_1^2 不等于 σ_2^2 的检验结果. 同时, t 检验的分析结果还给出了置信区间的结论, 其中置信上限是 2.970 27, 置信下限为 -1.770 27; 置信度为 95%. 因为软件分析结果更为精确, 例 8.16 和例 9.5 的计算结果与软件分析结果的细微差别为计算误差.

3. μ_1, μ_2 未知时, 方差 $\sigma_1^2 = \sigma_2^2$ 的检验

当原假设 $H_0: \sigma_1^2 = \sigma_2^2$ 成立时, 由抽样分布定理 7.9, 知此时检验统计量

$$F = \frac{S_1^2}{S_2^2} \sim F(n_1 - 1, n_2 - 1). \quad (9.2.5)$$

从而对各种备择假设可得到检验的拒绝域 (表 9.5). 这个检验由于检验统计量为 F , 称为 F 检验.

表 9.5 两正态总体方差的假设检表

H_0	H_1	显著性水平 α 下关于 H_0 的拒绝域
$\sigma_1^2 = \sigma_2^2$	$\sigma_1^2 \neq \sigma_2^2$	$F < F_{\frac{\alpha}{2}}(n_1 - 1, n_2 - 1)$ 或 $F > F_{1-\frac{\alpha}{2}}(n_1 - 1, n_2 - 1)$
$\sigma_1^2 = \sigma_2^2$	$\sigma_1^2 > \sigma_2^2$	$F > F_{1-\alpha}(n_1 - 1, n_2 - 1)$
$\sigma_1^2 = \sigma_2^2$	$\sigma_1^2 < \sigma_2^2$	$F < F_{\alpha}(n_1 - 1, n_2 - 1)$

例 9.6 在例 8.16 中, 若去掉条件 $\sigma_1^2 = \sigma_2^2$. 请问是否有 σ_1^2 显著大于 σ_2^2 ($\alpha = 0.05$)?

解 由题意, 检验假设为

$$H_0: \sigma_1^2 \leq \sigma_2^2, \quad H_1: \sigma_1^2 > \sigma_2^2.$$

由例 8.16, 知 $s_1^2 = 6.57, s_2^2 = 4.77, n_1 = 8, n_2 = 10$, 又 $\alpha = 0.05, F_{1-\alpha}(n_1 - 1, n_2 - 1) = F_{0.95}(7, 9) = 3.29$, 且

$$F = \frac{S_1^2}{S_2^2} = \frac{6.57}{4.77} = 1.38.$$

因为 $F = 1.38 < F_{0.95}(7, 9) = 3.29$, 故不拒绝 H_0 , 即不能认为 σ_1^2 显著大于 σ_2^2 .

§9.3 自然指数分布族均值参数的检验

对于非正态总体 X , 其分布函数 $F(x, \theta)$ 中含有未知参数 θ , 且 $E(X) = m(\theta), D(X) = \sigma^2(\theta)$. 当取得样本 X_1, X_2, \dots, X_n 且样本容量 n 较大时 (通常要求 $n \geq 50$), 由独立同分布中心极限定理, 在原假设 $H_0: \theta = \theta_0$ 成立时, 近似地有

$$U = \frac{\bar{X} - m(\theta_0)}{\sigma(\theta_0)/\sqrt{n}} \sim N(0, 1). \quad (9.3.1)$$

当显著性水平 α 给定后, 对双侧检验有拒绝域

$$W = \{|U| > u_{1-\frac{\alpha}{2}}\}. \quad (9.3.2)$$

对单侧检验, 拒绝域的导出则十分复杂.

特别地, 当总体 X 服从自然指数分布族分布, 对原假设 $H_0: m = m_0, H_1: m \neq m_0$ 的检验, 检验统计量为

$$U = \frac{\bar{X} - m_0}{\sqrt{V(m_0)/n}}, \quad (9.3.3)$$

双侧检验的拒绝域为 (9.3.2) 式, 其中 $V(m)$ 为 X 的方差函数.

下面讨论自然指数分布族总体均值参数 m 的单侧检验. 若检验假设为 $H_0: m \leq m_0, H_1: m > m_0$, 且方差函数 $V(m)$ 在 $m \leq m_0$ 时是 m 的单调递增函数, 则近似地有

$$\frac{\bar{X} - m}{\sqrt{\frac{V(m)}{n}}} \sim N(0, 1),$$

从而对给定的显著性水平 α , 近似地有

$$P\left(\frac{\bar{X} - m}{\sqrt{V(m)/n}} > u_{1-\alpha}\right) = \alpha.$$

当 H_0 成立时, 注意 $V(m)$ 在 $m \leq m_0$ 时是单调递增函数, 故有

$$U = \frac{\bar{X} - m_0}{\sqrt{V(m_0)/n}} \leq \frac{\bar{X} - m}{\sqrt{V(m_0)/n}} \leq \frac{\bar{X} - m}{\sqrt{V(m)/n}}.$$

从而以下事件有包含关系

$$\left(U = \frac{\bar{X} - m_0}{\sqrt{V(m_0)/n}} > u_{1-\alpha}\right) \subset \left(\frac{\bar{X} - m}{\sqrt{V(m)/n}} > u_{1-\alpha}\right),$$

有

$$P(U > u_{1-\alpha}) \leq P\left(\frac{\bar{X} - m}{\sqrt{V(m)/n}} > u_{1-\alpha}\right) = \alpha,$$

故此时右侧检验的拒绝域为

$$W = \{U > u_{1-\alpha}\}. \quad (9.3.4)$$

显然对右侧检验若原假设为 $H_0: m = m_0$, 其拒绝域也是 (9.3.4).

同理, 对左侧检验, $H_0: m \geq m_0$ (或 $m = m_0$), $H_1: m < m_0$, 当方差函数 $V(m)$ 在 $m \geq m_0$ 时是 m 的单调递增函数, 可得拒绝域为

$$W = \{U < -u_{1-\alpha}\}. \quad (9.3.5)$$

例 9.7 在某批电子设备中任取 50 台, 测得平均失效时间为 290 h. 已知此电子设备寿命 X 服从指数分布, 试问该电子设备平均寿命是否低于 320 h ($\alpha = 0.05$)?

解 因为指数分布为自然指数分布族分布, 且 $m = E(X)$, $V(m) = D(X) = m^2$ 是 m 的单调递增函数. 故检验假设

$$H_0: m \geq m_0 = 320, \quad H_1: m < 320.$$

又因为 $\alpha = 0.05$, 查表, 有 $-u_{1-\alpha} = -u_{0.95} = -1.645$, 且

$$u = \frac{\bar{x} - m_0}{\sqrt{\frac{V(m_0)}{n}}} = \frac{290 - 320}{320/\sqrt{50}} = -0.66.$$

因 $u = -0.66 > -u_{0.95} = -1.645$, 从而不拒绝 H_0 , 即不能认为该电子设备平均寿命低于 320 h.

例 9.8 某厂生产的某型号齿轮若直径误差超过 0.3 mm 则为次品. 该厂规定出厂产品次品率不得超过 2%. 某天从生产线上抽取 100 个该型齿轮检验, 发现其中有 3 个直径误差超过 0.3 mm. 问该天该厂生产的该型齿轮能否出厂 ($\alpha = 0.05$)?

解 注意该天齿轮能否出厂相当于该天次品率是否超过 2%. 设随机变量 X 为

$$X = \begin{cases} 1, & \text{抽一件产品为次品,} \\ 0, & \text{抽一件产品为正品.} \end{cases}$$

则 X 服从 0-1 分布 $B(1, p)$. 其中 p 为该厂产品次品率, 则抽取 100 个样品是从总体 X 中抽取容量为 100 的样本. 注意 X 属于自然指数分布族分布, 且 $m = E(X) = p$, $V(m) = D(X) = p(1-p)$. 易见方差函数 $V(m) = m(1-m)$ 在 $(0, \frac{1}{2})$ 内是单调递增函数. 现检验假设

$$H_0: p \leq p_0 = 0.02, \quad H_1: p > 0.02.$$

又 $\alpha = 0.05$, 查表得 $u_{1-\alpha} = u_{0.95} = 1.645$, 又有 $\bar{x} = \frac{1}{100} \sum_{i=1}^{100} x_i = \frac{3}{100} = 0.03$, 故

$$u = \frac{\bar{x} - m_0}{\sqrt{V(m_0)/n}} = \frac{0.03 - 0.02}{\sqrt{\frac{0.02 \times (1 - 0.02)}{200}}} = 1.01.$$

因 $u = 1.01 < 1.645$, 故不拒绝 H_0 , 可认为该厂产品次品率不大于 2%. 故可以出厂.

§9.4 总体分布的 χ^2 拟合优度检验

前面讨论的都是总体分布已知时参数的假设检验问题, 叫做参数检验. 但在实际应用中常遇到总体分布是未知的, 我们却要检验总体分布是否服从某一种分布, 这样的检验叫做非参数检验. 这一节我们介绍检验总体分布的 χ^2 拟合优度检验. 这种检验方法是皮尔逊 1900 年提出的.

设 $F_0(x)$ 为已知分布函数, 从总体 X 中取得样本观测值 x_1, x_2, \dots, x_n 后, 我们要检验总体 X 是否有分布函数 $F_0(x)$. 为此, 提出检验假设:

$$\begin{aligned} H_0: & \text{总体 } X \text{ 有分布函数 } F_0(x); \\ H_1: & \text{总体 } X \text{ 的分布函数不是 } F_0(x). \end{aligned} \quad (9.4.1)$$

拟合优度检验中, 备择假设 H_1 可省略不写.

为了检验这个假设, 需要寻求一个适当的检验统计量. 皮尔逊提出了构造这个统计量的思想和方法:

(1) 将数轴 $(-\infty, +\infty)$ 做划分:

$$-\infty = a_0 < a_1 < a_2 < \dots < a_k = +\infty,$$

令

$$I_i = (a_{i-1}, a_i], \quad i = 1, 2, \dots, k.$$

一般可取 $5 \leq k \leq 16$.

(2) 当原假设 H_0 成立时, 求总体 X 取值落入 I_i 的概率, 即

$$p_i = P(a_{i-1} < X \leq a_i) = F_0(a_i) - F_0(a_{i-1}), \quad i = 1, 2, \dots, k.$$

并计算样本观测值落入 I_i 的频率 $\frac{n_i}{n}$, 其中 n_i 为样本观测值 x_1, x_2, \dots, x_n 中落入区间 I_i 的个数, $i = 1, 2, \dots, k$.

(3) 由伯努利大数律可知, 当 H_0 成立且 n 较大时, $\frac{n_i}{n}$ 应与 p_i “充分接近”, 即分布函数 $F_0(x)$ 对总体 X 的分布函数拟合得相当好, 则

$$\left| \frac{n_i}{n} - p_i \right|, \quad i = 1, 2, \dots, k$$

应充分小, 从而 $\left(\frac{n_i}{n} - p_i \right)^2$ 也充分小. 但这个充分小的程度是应与 p_i 相对而言的. 于是 $F_0(x)$ 对总体分布函数的拟合优度 (即总体分布函数是 $F_0(x)$) 表现在

$\left(\frac{n_i}{n} - p_i\right)^2$ 的加权平方和 (这个“权”应与 p_i 有关) 应较小. 皮尔逊给出了下面的加权平方和 χ^2 作为拟合优度的检验统计量.

$$\chi^2 = \sum_{i=1}^k \frac{n}{p_i} \left(\frac{n_i}{n} - p_i\right)^2 = \sum_{i=1}^k \frac{(n_i - np_i)^2}{np_i}. \quad (9.4.2)$$

注意第二步计算 p_i 时, 若 $F_0(x)$ 中含有未知参数或未知参数向量 θ , 则应首先得到 θ 的极大似然估计 $\hat{\theta}$, 再计算 p_i , 此时, p_i 的值应是其估计值 \hat{p}_i .

皮尔逊证明了当 $n \rightarrow \infty$ 时, χ^2 统计量的极限分布为 $\chi^2(k-r-1)$ 分布, 其中 r 是 θ 中未知参数的个数. 由前面的分析知道若 H_0 不成立时, 即总体 X 的分布函数不是 $F_0(x)$ 时, χ^2 统计量的值应偏大. 故当显著性水平 α 给定后, 此检验的拒绝域为

$$W = \{\chi^2 > \chi^2_{1-\alpha}(k-r-1)\}. \quad (9.4.3)$$

还需要注意 χ^2 拟合优度检验要求样本容量 n 较大, 一般要求 $n \geq 50$, 同时还要求每个 $n_i \geq 5, i = 1, 2, \dots, k$. 若某些区间 I_i 内 n_i 不到 5, 可合并相邻的区间 I_i 使其相应的 $n_i \geq 5$. 当然这样一合并, 相应的 k 也就减小了.

例 9.9 经统计, 某机场 100 天中每天因故误了班机的人数如下

一天误班机人数 x_i	0	1	2	3	4	5	6
频数 n_i	14	27	26	20	7	3	3

检验该机场一天中误班机人数是否服从泊松分布 ($\alpha = 0.05$).

解 令 X 表示该机场一天中误班机人数, 则原假设为

$$H_0: X \sim P(\lambda).$$

因 λ 未知, 故可得 λ 的极大似然估计值为

$$\hat{\lambda} = \bar{x} = \frac{1}{100}(14 \times 0 + 27 \times 1 + 26 \times 2 + 20 \times 3 + 7 \times 4 + 3 \times 5 + 3 \times 6) = 2,$$

从而 H_0 为 $X \sim P(2)$. 此时, p_i 的估计值为

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

故有计算表 9.6:

表 9.6

i	n_i	\hat{p}_i	$n\hat{p}_i$	$(n_i - n\hat{p}_i)^2/n\hat{p}_i$
0	14	0.135 3	13.53	0.016 3
1	27	0.270 7	27.07	0.000 2
2	26	0.270 7	27.07	0.042 3
3	20	0.180 4	18.04	0.212 9
4	7	0.090 2	9.02	0.452 4
5	3	0.052 7	5.27	0.101 1
6	3			
合计	100			0.825 2

故得 $\chi^2 = 0.825\ 2$. 因合并区间后 $k = 6$, 估计的未知参数个数 $r = 1$. 又 $\alpha = 0.05$, 查表 $\chi^2_{1-\alpha}(k-r-1) = \chi^2_{0.95}(4) = 9.488 > \chi^2 = 0.825\ 2$. 从而不拒绝 H_0 , 可以认为该机场一天中误班机人数服从 $\lambda = 2$ 的泊松分布.

例 9.10 某厂对噪声污染进行分析时, 经测试 100 次得噪声数据如下 (单位: dB)

81.0	81.5	82.5	83.0	79.0	79.5	80.0	79.5	80.0	79.5
84.0	80.5	79.0	77.0	78.0	77.5	77.0	78.5	75.0	75.5
76.0	76.0	75.0	73.0	73.5	74.0	73.0	73.5	73.0	74.5
74.0	75.5	75.0	75.0	75.5	76.0	77.5	77.5	72.0	77.5
78.0	78.5	78.0	78.5	77.5	78.0	78.5	78.0	73.5	73.5
73.0	74.5	76.5	76.0	75.5	75.0	71.0	71.5	71.5	72.0
72.0	71.5	74.0	74.0	73.5	72.5	71.5	74.5	73.0	74.5
65.0	65.5	67.0	67.5	63.0	68.0	68.5	69.0	69.5	70.0
76.0	76.0	76.5	66.0	70.5	70.0	76.5	76.5	76.0	75.0
76.0	84.5	81.0	82.0	82.0	88.0	85.0	86.0	86.5	83.5

检验该厂噪声 X 是否服从正态分布 ($\alpha = 0.05$).

解 原假设为

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

但 μ 和 σ 为未知参数, 从而要求 μ 和 σ^2 的极大似然估计值. 由计算器计算得 $\hat{\mu} = \bar{x} = 75.76, \hat{\sigma}^2 = b_2 = 4.76^2$. 故原假设此时为

$$H_0: X \sim N(75.56, 4.76^2).$$

注意测试数据的最小值为 62.5, 最大值为 88.5. 即所有数据都落在区间 $[62.5, 88.5]$ 上. 将区间 $[62.5, 88.5]$ 等分为 13 个 a 区间, 每个小区间长度为 2.

故得计算表 9.7. 其中

$$\hat{p}_i = \Phi\left(\frac{b_i - 75.76}{4.76}\right) - \Phi\left(\frac{a_i - 75.76}{4.76}\right), \quad i = 1, 2, \dots, 13.$$

算得 $\chi^2 = 6.6326$. 由于合并区间后, $k = 9$, 估计的未知参数个数 $r = 2$. 又 $\alpha = 0.05$. 查表, 可得临界值 $\chi^2_{1-\alpha}(k-r-1) = \chi^2_{0.95}(6) = 12.592$.

因为 $\chi^2 = 6.6326 < 12.592 = \chi^2_{0.95}(6)$. 从而接受 H_0 , 认为噪声服从正态分布.

表 9.7

区间 $(a_i, b_i]$	n_i	$n\hat{p}_i$	$(n_i - n\hat{p}_i)^2 / n\hat{p}_i$
$(-\infty, 64.5]$	1	6.37	0.417 1
$(64.5, 66.5]$	3		
$(66.5, 68.5]$	4		
$(68.5, 70.5]$	5	7.13	0.636 3
$(70.5, 72.5]$	8	11.18	0.904 5
$(72.5, 74.5]$	18	14.92	0.635 8
$(74.5, 76.5]$	22	16.56	1.787 1
$(76.5, 78.5]$	17	15.53	0.139 1
$(78.5, 80.5]$	8	12.38	1.549 6
$(80.5, 82.5]$	6	8.13	0.558 0
$(82.5, 84.5]$	4	7.70	0.005 1
$(84.5, 86.5]$	3		
$(86.5, +\infty)$	1		
合计	100		6.632 6

χ^2 拟合检验的计算量较大, 可采用 SPSS 统计软件进行计算.

§9.5 复习分析题

例 9.11 设总体 $X \sim N(\mu, \sigma_0^2)$, X_1, X_2, \dots, X_n 为样本. 检验假设为

$$H_0: \mu = \mu_0, \quad H_1: \mu > \mu_0.$$

(1) 求检验犯第二类错误的概率 $\beta = G(\mu)$.

(2) 证明: 当 n 固定时, α 与 β 不能同时减少; 且 α 固定时, 增大 n 能使 β 减少.

(3) 设 $\sigma_0 = 1$, $\mu_0 = 10$, 当 $\mu = 11$, $\alpha = 0.05$ 时, 为了使 $\beta \leq 0.05$, 应取多大容量的样本.

分析 当原假设 H_0 成立时, 有

$$U = \frac{\bar{X} - \mu_0}{\sigma_0/\sqrt{n}} \sim N(0, 1).$$

对此右侧检验, 拒绝域为 $W = \{U > u_{1-\alpha}\}$, 从而不难计算犯第二类错误的概率 $\beta = G(\mu)$. 而要证明 n 固定时, α 与 β 不能同时减少, 则可找出 $u_{1-\alpha}$ 与 $u_{1-\beta}$ 的关系式再分析.

解 (1) $\beta = P\{\text{接受 } H_0 | H_0 \text{ 不真}\}$. 但 H_0 不真时, 有 $\frac{\bar{X} - \mu}{\sigma_0/\sqrt{n}} \sim N(0, 1)$, 故此时由 $E(U) = \frac{\mu - \mu_0}{\sigma_0/\sqrt{n}}, D(U) = 1$, 得

$$U = \frac{\bar{X} - \mu_0}{\sigma_0/\sqrt{n}} \sim N\left(\frac{\mu - \mu_0}{\sigma_0/\sqrt{n}}, 1\right).$$

于是有

$$\begin{aligned}\beta &= G(\mu) = P(U < u_{1-\alpha}) \\ &= \Phi\left(u_{1-\alpha} - \frac{\mu - \mu_0}{\sigma_0/\sqrt{n}}\right).\end{aligned}$$

(2) 由 u 分位点定义, 知 $\beta = \Phi(-u_{1-\beta})$, 故有

$$u_{1-\alpha} - \frac{\mu - \mu_0}{\sigma_0/\sqrt{n}} = -u_{1-\beta},$$

从而有

$$u_{1-\alpha} + u_{1-\beta} = \frac{\mu - \mu_0}{\sigma_0/\sqrt{n}}.$$

当 n 固定时, 取 $\mu > \mu_0$ 固定. 注意到 $\Phi(x)$ 的单调性及 $\Phi(u_{1-\alpha}) = 1 - \alpha$. 当 α 减少时, 从而有 $u_{1-\alpha}$ 增大, 故有 $u_{1-\beta}$ 减小 β 增大. 因此 n 固定时, 不能同时减少 α 和 β .

另外, 当 α 固定时, 则 $u_{1-\alpha}$ 固定, 增大 n 从而使 $u_{1-\beta}$ 增大, 即是使 β 减小.

(3) 当 $\sigma_0 = 1, \mu_0 = 10, \mu = 11, \alpha = 0.05$, 此时, 有

$$\begin{aligned}\beta &= G(11) = \Phi\left(u_{0.95} - \frac{11 - 10}{1/\sqrt{n}}\right) \\ &= \Phi(1.645 - \sqrt{n}).\end{aligned}$$

要使 $\beta \leq 0.05$, 注意 $\Phi(-1.645) = 0.05$, 故有

$$-1.645 \geq 1.645 - \sqrt{n}.$$

解得 $n \geq (2 \times 1.645)^2 = 10.824$, 即样本容量 n 至少取 11.

例 9.12 农业研究所为了研究某种化肥对农作物的效力, 在 13 个单位面积地块进行试验, 得到农作物的单位面积产量 (单位: kg) 如下:

未施肥地块: 29, 27, 32, 31, 28, 32, 31;

施化肥地块: 34, 35, 32, 33, 34, 30.

设农作物的单位面积产量服从正态分布, 检验施用该种化肥能否显著提高农作物的单位面积产量 ($\alpha = 0.05$).

分析 设 X 表示未施肥地块农作物的单位面积产量, $X \sim N(\mu_1, \sigma_1^2)$; Y 表示施化肥地块农作物的单位面积产量, $Y \sim N(\mu_2, \sigma_2^2)$, 其中参数 $\mu_1, \sigma_1^2, \mu_2, \sigma_2^2$ 都是未知的. 因此应首先检验是否 $\sigma_1^2 = \sigma_2^2$, 若相等, 再检验是否有 $\mu_1 = \mu_2$.

解 由数据, 可算得

$$n_1 = 7, \bar{x} = 30, s_1^2 = 4.0;$$

$$n_2 = 6, \bar{y} = 33, s_2^2 = 3.2.$$

首先检验假设

$$H_0: \sigma_1^2 = \sigma_2^2, \quad H_1: \sigma_1^2 \neq \sigma_2^2,$$

$\alpha = 0.05$. 查表得 $F_{1-\frac{\alpha}{2}}(n_1 - 1, n_2 - 1) = F_{0.975}(6, 5) = 6.98$, 且统计量值为

$$F = \frac{s_1^2}{s_2^2} = \frac{4}{3.2} = 1.25.$$

因为 $F = 1.25 < F_{0.975}(6, 5) = 6.98$, 故应接受 H_0 , 即可认为 $\sigma_1^2 = \sigma_2^2$.

现在检验假设

$$H_0: \mu_1 = \mu_2; \quad H_1: \mu_1 < \mu_2,$$

$\alpha = 0.05$. 查表得 $t_{1-\alpha}(n_1 + n_2 - 2) = t_{0.95}(11) = 1.796$, 且统计量值为

$$T = \frac{\bar{x} - \bar{y}}{s_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{30 - 33}{\sqrt{\frac{6 \times 4 + 5 \times 3.2}{11}} \times \sqrt{\frac{1}{7} + \frac{1}{6}}} = -2.83.$$

因为 $T = -2.83 < -t_{0.95}(1) = -1.796$, 故应拒绝 H_0 , 而接受 H_1 , 即认为施化肥后能显著提高农作物的单位面积产量.

例 9.13 医生对于慢走是否能降低血压这一问题的研究感兴趣, 随机地选取 8 个病人慢走一个月, 得到以下的血压数据 (单位: mmHg):

病人编号	1	2	3	4	5	6	7	8
慢走前 (x_i)	134	122	118	130	144	125	127	133
慢走后 (y_i)	130	120	123	127	138	121	132	135

问慢走是否可以使血压降低 ($\alpha = 0.05$)?

分析 设慢走前后的血压值都服从正态分布且方差相等, 则可由检验两总体均值是否相等的 t 检验法检验. 但本例这种数据叫做配对数据. 是在相同的条件下由慢走前后血压的改变得到的一对数据. 设 $D_i = X_i - Y_i$, 若 $D_i \sim N(\mu_D, \sigma_D^2)$, 则可导出关于 μ_D 的检验方法, 叫做配对数据的检验.

解 设 $D_i = X_i - Y_i \sim N(\mu_D, \sigma_D^2)$, 则问题应检验

$$H_0: \mu_D = 0; \quad H_1: \mu_D > 0.$$

由抽样分布定理知

$$\frac{\bar{D} - \mu_D}{S_D/\sqrt{n}} \sim t(n-1),$$

其中 \bar{D} 为样本 D_1, \dots, D_n 的样本均值, S_D^2 为样本方差.

当 H_0 成立时,

$$t = \frac{\bar{D}}{S_D/\sqrt{n}} \sim t(n-1),$$

故当 $\alpha = 0.05$ 时, 检验的拒绝域为

$$W = \left\{ t = \frac{\bar{D}}{S_D/\sqrt{n}} > t_{1-\alpha}(n-1) \right\}.$$

由假设检验的理论可知对配对数据而言, 采用配对数据 t 检验比采用两总体均值的 t 检验效率更高.

本例中, $d_i = x_i - y_i$ 的数据为 4, 2, -5, 3, 6, 4, -5, -2, 算得 $\bar{d} = 1.7143$, $s_D = 3.8607$, $n = 8$, 故

$$t = \frac{1.7143}{3.8607/\sqrt{8}} = 1.2529.$$

又 $\alpha = 0.05$, $t_{1-\alpha}(n-1) = t_{0.95}(7) = 1.8946$, 因 $t = 1.2529 < t_{1-\alpha}(n-1) = 1.8946$, 故不拒绝 H_0 , 从而不能认为慢走会显著降低血压.

习 题 九

一、选择题

1. 在假设检验问题中, 由于样本的随机性, 造成以下 () 的错误称为第二类错误.
 - (A) H_0 本来真实, 检验后作出拒绝 H_0 的判断
 - (B) H_0 本来不真实, 检验后作出接受 H_1 的判断
 - (C) H_1 本来不真实, 检验后作出接受 H_1 的判断
 - (D) H_1 本来真实, 检验后作出拒绝 H_1 的判断

2. 总体 $X \sim N(\mu, \sigma^2)$, σ^2 未知, x_1, x_2, \dots, x_n 为样本值, 在显著性水平 $\alpha = 0.05$ 时, 检验 $H_0: \mu = \mu_0; H_1: \mu \neq \mu_0$, 且拒绝 H_0 . 则当显著性水平改为 $\alpha = 0.01$ 时, 下列结论正确的是 ().

- (A) 必定拒绝 H_0 (B) 必定接受 H_0
(C) 第二类错误变小 (D) 可能接受, 也可能拒绝 H_0

3. 在一个确定的检验问题中, 与判断结果有关的因素有 ().

- (A) 样本值及样本容量 (B) 显著性水平 α
(C) 所选的检验统计量 t (D) (A), (B) 同时成立

4. 设总体 $X \sim N(\mu, 16)$, X_1, X_2, \dots, X_{16} 为来自总体 X 的样本, 当 σ^2 已知时, 检验 $H_0: \mu = 50, H_1: \mu < 50$, 取 $\alpha = 0.05$, 再取 $U = \bar{X} - 50$ 作为检验统计量, 则可得检验的拒绝域 W 为 ().

- (A) $W = \{|U| < 1.96\}$ (B) $W = \{U < -1.645\}$
(C) $W = \{|U| < 1.645\}$ (D) $W = \{U < -1.96\}$

二、填空题

1. X_1, X_2, \dots, X_n 是来自正态总体 $N(\mu, 4)$ 的样本. 在显著性水平 α 下检验 $H_0: \mu = 0; H_1: \mu \neq 0$. 现取拒绝域 $W = \left\{ \sqrt{n} \frac{|\bar{X}|}{2} > u_{1-\frac{\alpha}{2}} \right\}$, 则当实际情况为 $\mu = 1$ 时, 犯第二类错误的概率为 ____.

2. 设总体 $X \sim N(\mu, \sigma^2)$, 且 σ^2 未知, 当检验 $H_0: \mu = \mu_0$ 时, 应采用统计量 ____, 当 H_0 真实时, 统计量应服从 ____ 分布.

3. 设总体 $X \sim N(\mu, \sigma^2)$, X_1, X_2, \dots, X_{25} 为样本. 检验 $H_0: \sigma^2 = 100$ 时, 应取统计量 ____, 当 H_0 真实时, 统计量应服从 ____ 分布.

三、解答题和证明题

1. 设总体 $X \sim N(\mu, \sigma^2)$, 且 σ 已知. 今抽取容量为 n 的样本, 得样本均值 \bar{x} , 做如下检验

$$H_0: \mu = \mu_0, \quad H_1: \mu = \mu_1 \quad (\mu_1 > \mu_0).$$

若已求得临界值 $k (0 < k < \mu_1 - \mu_0)$, 按如下原则判断假设的真伪: 当 $\bar{x} < \mu_0 + k$ 时, 接受 H_0 ; 当 $\bar{x} \geq \mu_0 + k$ 时, 拒绝 H_0 . 证明: 检验犯第一类错误及第二类错误的概率分别为

$$\alpha_n = 1 - \Phi\left(\frac{k}{\sigma\sqrt{n}}\right), \quad \beta_n = \Phi\left(\frac{\mu_0 + k - \mu_1}{\sigma/\sqrt{n}}\right).$$

2. 设总体 $X \sim N(\mu, 9)$, μ 未知, X_1, X_2, \dots, X_{25} 为样本. \bar{X} 为样本均值. 检验 $H_0: \mu = \mu_0; H_1: \mu \neq \mu_0$, 取检验的拒绝域 $W = \{|\bar{X} - \mu_0| \geq C\}$. 试确定常数 C , 使得检验的显著性水平 $\alpha = 0.05$.

3. 某种化工原料的含脂率 $X \sim N(0.27, 0.16^2)$. 对经处理后的这种原料取样测试 10 个样品. 测得含脂率如下:

0.19, 0.24, 0.04, 0.08, 0.20, 0.12, 0.31, 0.29, 0.13, 0.07.

(1) 若处理前后方差不变, 问处理前后的平均含脂率有无显著变化 ($\alpha = 0.05$)?

(2) 若不知方差处理前后方差是否变化, 问处理后的平均含脂率有无显著降低 ($\alpha = 0.05$)?

4. 某型弹壳直径 X (单位: mm) $\sim N(\mu, \sigma^2)$. 规定标准为 $\mu = 8, \sigma = 0.09$. 某车间新生产一批这种弹壳. 已知这批弹壳直径的标准差为标准值. 现抽测 9 枚弹壳, 得平均值为 7.97 试问这批弹壳是否合格 ($\alpha = 0.05$).

5. 某种导线的电阻服从正态分布, 要求电阻的标准差不得超过 0.004Ω . 今从某厂生产的一批导线中任意抽取 10 根, 测得 $s = 0.006 \Omega$, 能否认为这批导线电阻的标准差显著偏大 ($\alpha = 0.05$)?

6. 甲、乙两种枪弹的速度 (单位: m/s) 分别服从正态总体 $N(\mu_1, \sigma_1^2)$ 和 $N(\mu_2, \sigma_2^2)$. 在相同条件下进行速度测定, 得到如下结果:

枪弹甲: $n_1 = 30, \bar{x} = 2805, s_1 = 120.41$;

枪弹乙: $n_2 = 25, \bar{y} = 2680, s_2 = 105.00$.

(1) 由以往统计, 知 $\sigma_1 = 118, \sigma_2 = 115$, 问两种枪弹速度有无显著差异 ($\alpha = 0.05$).

(2) 若不知 σ_1, σ_2 , 但知 $\sigma_1 = \sigma_2 = \sigma$, 问枪弹甲的速度是否比枪弹乙快 ($\alpha = 0.05$).

7. 下表分别给出两个文学家马克·吐温的 8 篇小品文以及思诺特格拉斯的 10 篇小品文中由 3 个字母组成的词的比例:

马克·吐温	0.225	0.262	0.217	0.240	0.230	0.229	0.235	0.217		
思诺特格拉斯	0.209	0.205	0.196	0.210	0.202	0.207	0.224	0.223	0.220	0.201

设两组数据分别来自两个方差相等的正态总体, 问两个作家所写的小品文中由 3 个字母组成的词的比例是否有显著差异 ($\alpha = 0.05$)?

8. 有 A, B 两个化验室, 每天同时对某厂的冷却水中抽样检验水中含氧量 (单位: ppm), 下面是 7 天的检验记录:

A 室: 1.15, 1.86, 0.75, 1.82, 1.14, 1.65, 1.90;

B 室: 1.00, 1.90, 0.90, 1.80, 1.20, 1.70, 1.95.

设两个化验室测定值都服从正态分布. 问

(1) 两个化验室测定值的方差是否相同 ($\alpha = 0.05$)?

(2) 两个化验室的测定值有无显著差异 ($\alpha = 0.01$)?

9. 已知某种电子元件的使用寿命服从指数分布 $e(\lambda)$, 抽查 100 个样品, 测得样本均值 $\bar{x} = 1950$ h, 能否认为这种电子元件的平均寿命为 2000 h ($\alpha = 0.05$)?

10. 某种产品的次品率原为 10%, 改进工艺试制了一批新产品, 抽查 200 件样品中发现了 13 件次品, 是否可以认为改进了工艺后显著降低了产品的次品率 ($\alpha = 0.05$)?

11. 以下是某地区 100 个月中各月发生的较大地震的次数:

一个月的较大地震的次数	0	1	2	3	≥ 4
月数	57	31	8	3	1

检验数据是否来自泊松总体 ($\alpha = 0.05$).

12. 一供货商声称他们厂生产的电子元件的寿命 (单位: h) 服从均值为 200 的指数分布. 现随机取 1000 只这种元件, 测得数据如下:

寿命 x	$x \leq 150$	$150 < x \leq 300$	$300 < x \leq 450$	$450 < x \leq 600$
只数	543	258	120	48
寿命 x	$600 < x \leq 750$	$x > 750$		
只数	20	11		

取 $\alpha = 0.05$, 检验供货商说法是否属实.

13. 对某汽车零件制造厂所生产的汽缸螺栓直径进行抽样检验, 测得 100 个数据, 分组统计如下:

分组 $(a_i, b_i]$	频数 n_i	分组 $(a_i, b_i]$	频数 n_i
(10.93,10.95]	5	(11.01,11.03]	17
(10.95,10.97]	8	(11.03,11.05]	6
(10.97,10.99]	20	(11.05,11.07]	6
(10.99,11.01)	34	(11.09,11.11)	4

检验螺栓直径是否服从正态分布 ($\alpha = 0.05$).