高等工程數學（8）

厦門大學
XIAMEN UNIVERSITY

信息学院
（国家示范性软件学院）
School of Informatics

黄 煒
博士，副教授
Dr. Wei Huang

假设检验

数理统计（8）

厦門大學
XIAMEN UNIVERSITY
信息学院
（国家示范性软件学院）
School of Informatics
黄 炜
博士·副教授
Dr. Wei Huang

3 假设检验

统计推断的另一类重要问题是假设检验问题. 第二章参数估计是讨论在总体分布形式已知的前提下,借助于样本构造统计量以对总体的未知参数进行估计. 那么,当总体分布形式未知时,能否对总体的分布形式和参数进行估计或推断呢? 这就是本章所要讨论的假设检验.

在假设检验中,考虑问题的方式与估计理论不同,乃是先假设总体的分布形式或参数具有某种特征,然后利用样本所提供的信息来判断原先的假设是否合理,如果合理则认为假设可以接受,否则便否定原先的假设,从而对所研究的问题做出推断或分析. 例如,已知一个样本来自正态总体,问题是要证实这个样本是否来自均值是 μ_0 的正态总体. 又如,如果已知相互独立的两个样本,它们分别来自两个正态总体,那么能否说这两个正态总体的均值相同呢? 这类问题都需要对总体的未知参数作某种假设,并由样本构造适当的统计量,再根据样本数据对所作假设进行检验,从而对总体的性质作出推断. 这种处理问题的方法称为**假设检验**.

3.1 假设检验的基本概念

下面通过几个例子说明假设检验问题的提法与解决问题的方法，以及其所依据的原理，并介绍有关的一些概念和术语.

例 1 某车间用一台包装机包装葡萄糖，额定标准为每袋净重 0.5 kg. 包装机正常工作时，装包量服从正态分布，且根据长期的经验知其标准差 $\sigma=0.015$ kg. 为了检验包装机工作是否正常，随机抽取它所包装的葡萄糖 9 袋，称得净重(单位：kg)为

$$0.497,\ 0.506,\ 0.518,\ 0.524,\ 0.488,$$
$$0.511,\ 0.510,\ 0.515,\ 0.512$$

问包装机工作是否正常？

设装包量为 X，则 $X\sim N(\mu;\sigma^2)$，且 $\sigma=0.015$. 所谓工作正常，即是 $\mu=0.5$ kg，因而假设包装机工作正常，记为

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

现在需要由所得的 9 个数据来判断假设 H_0 是否合理？未知参数 μ 的一个较好估计值为

$$\bar{x}=\frac{1}{9}\sum_{i=1}^{9}x_i=0.509.$$

如果工作正常，那么估计值 $\bar{x}$ 与 $\mu_0=0.5$ 应该相差不大. 反之，如果样本观测值 $\bar{x}$ 与 μ_0 有显著差异，即 $|\bar{x}-\mu_0|$ 相当大时，则有理由认为工作不正常.

这个问题是在已知总体为正态分布且方差已知的前提下，检验均值是否为指定常数.

3.1 假设检验的基本概念

例 2 某厂生产两种型号的灯泡，根据以往的资料可以假设灯泡寿命服从正态分布，现随机地从这两种型号的灯泡中分别抽取容量为 40 的样本，测得其寿命与标准差的数据为

$$\text{甲型号：}\quad \overline{x}=1242\text{ 小时},\ s_1=90\text{ 小时},$$

$$\text{乙型号：}\quad \overline{y}=1258\text{ 小时},\ s_2=94\text{ 小时}.$$

问这两种型号的灯泡的寿命可以认为相同吗？

这个问题是在两个总体的分布类型都为正态分布的前提下，检验它们的方差和均值是否相同. 由于关心的问题是灯泡寿命，所以也可以将这个问题看成为在总体方差未知的情况下，检验这两个总体的总体均值是否相同.

我们把任何一个有关未确知的总体分布的命题称为**统计假设**，简称为**假设**. 从所举例子看出，统计假设常常要涉及两种情况的假设，如例 1 是检验总体均值 μ 是否为 $\mu_0=0.5$. 如果用 H_0 表示 $\mu=\mu_0$ 这个假设，那么另一个假设可用 H_1 表示，它就是 $\mu\neq\mu_0$. 我们称 H_0 为原假设(或零假设)，而称 H_1 为备选假设. 至于假设检验中的两个假设，把哪一个作为原假设，哪一个作为备选假设，要根据具体目的和要求而定，同时也要考虑数学处理的方便.

3.1 假设检验的基本概念

所谓**统计检验**，简称为**检验**，指的是根据样本观测值对假设作出判断的方法，判断的结论或者是认为原假设 H_0 成立，或者是认为原假设 H_0 不成立而是备选假设 H_1 成立. 称前者为接受 H_0，称后者为拒绝 H_0. 由于样本观测值 $(x_1, x_2, \cdots, x_n)$ 是样本空间中的一个点，因而统计检验的实质是将样本空间划分为两个不相交的集合 W_0 与 W_1，W_0 表示检验的接受域，W_1 表示检验的拒绝域. 也就是说，当 $(x_1, x_2, \cdots, x_n) \in W_0$ 时接受 H_0，而当 $(x_1, x_2, \cdots, x_n) \in W_1$ 时拒绝 H_0. 由于一个统计问题中的样本空间是可以事先知道的，因此确定了拒绝域 W_1，也就确定了 W_0. 一般是对每一个检验指明其拒绝域 W_1. 这样，所谓给出一个检验，实际上就是指明这个检验的拒绝域 W_1.

由于样本的随机性，在使用一个检验时，作出“接受 H_0”或“拒绝 H_0”的判断并不证明假设 H_0 一定正确或一定不正确，而只是表明检验使用者对假设 H_0 的一种倾向性意见. 因此，使用一个检验会面临犯两种错误的可能性：一是当 H_0 实际上成立，但由于样本观测值落在拒绝域 W_1 内而拒绝 H_0；二是当 H_0 实际上不成立，但由于样本观测值落在拒绝域外而接受 H_0. 我们称前者“弃真”的错误为**第一类错误**，而称后者“取伪”的错误为**第二类错误**. 为了清楚起见，我们将使用一个检验可能产生的各种后果列成表 3.1-1.

3.1 假设检验的基本概念

表 3.1-1

检验带来的后果		根据样本观测值所得的结论	
		当 $(x_1,\cdots,x_n)\notin W_1$ 接受 H_0	当 $(x_1,\cdots,x_n)\in W_1$ 拒绝 H_0
总体分布的实际情况(未知)	H_0 成立	判断正确	犯第一类错误
	H_0 不成立	犯第二类错误	判断正确

犯两类错误的大小可用概率来度量. 犯第一类错误的概率为

$$P\{\text{拒绝 } H_0 \mid H_0 \text{ 成立}\} = P\{(X_1,X_2,\cdots,X_n)\in W_1 \mid H_0 \text{ 成立}\}, \qquad (3.1\text{-}1)$$

称为检验的 I 类风险；而犯第二类错误的概率为

$$P\{\text{接受 } H_0 \mid H_0 \text{ 不成立}\} = P\{(X_1,X_2,\cdots,X_n)\notin W_1 \mid H_0 \text{ 不成立}\}, \qquad (3.1\text{-}2)$$

称为检验的 II 类风险. 当选用一个检验时，自然希望两类风险都尽可能小. 但是，两者往往不能兼顾. 当样本容量 n 固定时，一般情况下，I 类风险变小时，II 类风险会增大；反之，II 类风险变小时，I 类风险会增大. 通常是控制 I 类风险不超过某个预先指定的小正数 α 的前提下，使 II 类风险 β 尽可能小. 但 II 类风险 β 的计算比较复杂，所以一般只考虑 I 类风险至多为 α 的要求，这样得到的统计检验称为**显著性检验**，并称 α 为**显著性水平**.

对于一个显著性检验，它的拒绝域 W_1 由下式确定：

$$P\{(X_1,X_2,\cdots,X_n)\in W_1 \mid H_0 \text{ 成立}\} = \alpha. \qquad (3.1\text{-}3)$$

3.1 假设检验的基本概念

由于显著性检验主要是控制Ⅰ类风险不致太大，所以根据“小概率事件在一次试验中不应该发生”这一原理，最终判断“拒绝 H_0”时结论较为可靠，而最终判断“接受 H_0”的结论则不太可靠. 因为不知道Ⅱ类风险究竟有多大，所以通常用“不拒绝 H_0”来代替“接受 H_0”比较恰当.

设要检验的对象是某个未知参数 θ，那么解决一个统计假设检验问题的一般步骤如下.

1）建立原假设 H_0，必要时指明备选假设 H_1. 通常 H_0 与 H_1 是用某个关于 θ 的等式或不等式来表示的；

2）求出未知参数 θ 的一个较优的点估计 $\hat{\theta}(X_1,X_2,\cdots,X_n)$，一般用 θ 的极大似然估计；

3）以 $\hat{\theta}$ 为基础寻找一个检验统计量

$$T=t(X_1,X_2,\cdots,X_n),$$

要求 T 的分布当 H_0 成立时是已知的，从而可以通过查表或直接计算得到 T 的分位数；

4）对于给定的显著性水平 α，以检验统计量 T 为基础，由 H_0 与 H_1 的内容确定拒绝域 W_1，使得当 H_0 成立时，

$$P\{(X_1,X_2,\cdots,X_n)\in W_1\}=\alpha.$$

通常，W_1 是用关于 $t(x_1,x_2,\cdots,x_n)$ 的一个或几个不等式来给出的.

3.2 正态总体参数的假设检验

在假定总体分布为正态分布下，我们分别对一个总体或两个总体的情形讨论参数的假设检验问题. 由于正态总体下抽样分布有较好的结果，并且大多数非正态总体下检验统计量的分布可以用正态分布来近似，因此正态总体下参数的假设检验得到了广泛应用.

1. 一个正态总体的情形

设$(X_1,X_2,\cdots,X_n)$是取自正态总体$N(\mu;\sigma^2)$的样本，分别讨论如何应用上述一般步骤解决μ与σ^2的各种假设检验问题.

(1) σ^2 已知时 μ 的检验

这时要检验假设

$$\mathrm{H}_0:\mu=\mu_0;\quad \mathrm{H}_1:\mu\neq\mu_0,$$

通常表述为“在显著性水平α下，针对H_1检验H_0”.

μ的极大似然估计量是$\overline{X}$，且$\overline{X}\sim N(\mu;\frac{\sigma^2}{n})$，从而$\sqrt{n}\frac{\overline{X}-\mu}{\sigma}\sim N(0;1)$. 当$\mathrm{H}_0$成立时，$\mu=\mu_0$，且记

$$U=u(X_1,X_2,\cdots,X_n)=\sqrt{n}\frac{\overline{X}-\mu_0}{\sigma},$$

3.2 正态总体参数的假设检验

则有 $u(X_1,X_2,\cdots,X_n)\sim N(0;1)$(注意，在 H_0 成立的条件下，$u(X_1,X_2,\cdots,X_n)$才是服从标准正态分布). 因此可用 U 作为检验统计量. 如果 H_0 成立，那么$|u(x_1,x_2,\cdots,x_n)|$应该比较小. 于是，当根据样本观测值$(x_1,x_2,\cdots,x_n)$发现$|u(x_1,x_2,\cdots,x_n)|$较大时，自然有理由认为 H_0 可能不成立，故拒绝域 W_1 可以取为

$$W_1=\{(x_1,x_2,\cdots,x_n)\ |\ |u(x_1,x_2,\cdots,x_n)|>k\}.$$

对给定的显著性水平 α，由于当 $\mu=\mu_0$ 时

$$P\{|U|>u_{1-\alpha/2}\}=\alpha,$$

所以临界值 $k=u_{1-\alpha/2}$. 这样便得到显著性水平 α 下检验的拒绝域.

$$W_1=\{(x_1,x_2,\cdots,x_n)\ |\ |u(x_1,x_2,\cdots,x_n)|>u_{1-\alpha/2}\}. \tag{3.2-1}$$

这个检验称为 **u 检验**.

有时，我们只关心总体均值是否增大(或减小). 譬如，经过工艺改革后，考虑产品的质量是否比以前有所提高. 若无明显提高，改革就失去了意义. 这时的问题是在新工艺下产品质量的均值 $\mu>\mu_0$ 是否成立. 假设形式成为

$$H_0:\mu\leqslant\mu_0,\quad H_1:\mu>\mu_0. \tag{3.2-2}$$

在 3.4 节中将说明上述假设形式与假设

$$H_0:\mu=\mu_0,\quad H_1:\mu>\mu_0 \tag{3.2-2$'$}$$

在同一显著性水平 α 下的检验法是一样的. 但后者便于在 H_0 成立下求出拒绝域，所以采用(3.2-2′)式表示的假设形式.

3.2 正态总体参数的假设检验

这种假设形式中的备选假设处于原假设的一侧，通常称为单侧假设，相应的检验称为**单侧检验**.

仍用样本均值 $\overline{X}$ 来估计 μ. 但不能当 $\overline{X}>\mu_0$ 时就认为 $\mu>\mu_0$，这是因为有随机波动，而只有 $\overline{X}$ 比 μ_0 大得适当多一些时才能认为 $\mu>\mu_0$ 真的成立. 因此，拒绝域可以取为

$$W_1=\{(x_1,x_2,\cdots,x_n)\mid u(x_1,x_2,\cdots,x_n)>k\}.$$

对给定的显著性水平 α，当 $\mu=\mu_0$ 时，

$$P\{U>u_{1-\alpha}\}=\alpha,$$

故临界值 $k=u_{1-\alpha}$. 于是，显著性水平 α 下单侧检验(3.2-2)的拒绝域为

$$W_1=\{(x_1,x_2,\cdots,x_n)\mid u(x_1,x_2,\cdots,x_n)>u_{1-\alpha}\}. \tag{3.2-3}$$

类似地，若考虑假设形式为

$$\mathrm{H}_0:\mu\geqslant\mu_0,\quad \mathrm{H}_1:\mu<\mu_0, \tag{3.2-4}$$

则可以得到在显著性水平 α 下的拒绝域

$$W_1=\{(x_1,x_2,\cdots,x_n)\mid u(x_1,x_2,\cdots,x_n)<-u_{1-\alpha}\}. \tag{3.2-5}$$

(2) σ^2 未知时 μ 的检验

我们要检验

$$\mathrm{H}_0:\mu=\mu_0,\quad \mathrm{H}_1:\mu\neq\mu_0.$$

μ 的极大似然估计量仍是 $\overline{X}$，但上述(1)中的 U 含有未知参数 σ，故不能取为检验统

3.2 正态总体参数的假设检验

计量. 考虑到用样本标准差 S 替代 σ 便得到统计量 $T=\sqrt{n}\frac{\overline{X}-\mu_0}{S}$，且 $T\sim t(n-1)$，因而取检验统计量为

$$T=t(X_1,X_2,\cdots,X_n)=\sqrt{n}\,\frac{\overline{X}-\mu_0}{S}.$$

当 H_0 成立时，$|t(x_1,x_2,\cdots,x_n)|$ 应该比较小，如果根据样本观测值发现 $|t|$ 较大，那么自然有理由认为 H_0 可能不成立. 于是，对于给定的显著性水平 α，检验的拒绝域为

$$W_1=\left\{(x_1,x_2,\cdots,x_n)\;\middle|\;\sqrt{n}\,\frac{|\overline{x}-\mu_0|}{s}>t_{1-\alpha/2}(n-1)\right\}. \qquad (3.2\text{-}6)$$

这个检验称为 t 检验.

对于单侧检验，将有关结论列于表 3.2-1 中.

3.2 正态总体参数的假设检验

表 3.2-1 一个正态总体下均值检验的拒绝域

假设		拒绝域 W_1 应满足的条件	
H_0	H_1	σ^2 已知	σ^2 未知
检验统计量及其分布		$U=\sqrt{n}\dfrac{\overline{X}-\mu_0}{\sigma}\sim N(0,1)$ $(\mu=\mu_0)$	$T=\sqrt{n}\dfrac{\overline{X}-\mu_0}{s}\sim t(n-1)$ $(\mu=\mu_0)$
$\mu=\mu_0$	$\mu\neq\mu_0$	$\sqrt{n}\dfrac{\lvert\bar{x}-\mu_0\rvert}{\sigma}>u_{1-\alpha/2}$	$\sqrt{n}\dfrac{\lvert\bar{x}-\mu_0\rvert}{s}>t_{1-\alpha/2}(n-1)$
$\mu\leqslant\mu_0$	$\mu>\mu_0$	$\sqrt{n}\dfrac{\bar{x}-\mu_0}{\sigma}>u_{1-\alpha}$	$\sqrt{n}\dfrac{\bar{x}-\mu_0}{s}>t_{1-\alpha}(n-1)$
$\mu\geqslant\mu_0$	$\mu<\mu_0$	$\sqrt{n}\dfrac{\bar{x}-\mu_0}{\sigma}<-u_{1-\alpha}$	$\sqrt{n}\dfrac{\bar{x}-\mu_0}{s}<-t_{1-\alpha}(n-1)$

3.2 正态总体参数的假设检验

(3) μ 已知时 σ^2 的检验

这时要检验

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

σ^2 的极大似然估计量是 $\widehat{\sigma^2}=\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2$，当 H_0 成立，即 $\sigma^2=\sigma_0^2$ 时，$\chi^2=\frac{n\widehat{\sigma^2}}{\sigma_0^2}=\frac{1}{\sigma_0^2}\sum_{i=1}^{n}(X_i-\mu)^2\sim\chi^2(n)$. 因此，取检验统计量为

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

当 H_0 成立时，由于 $\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2$ 应在 σ_0^2 周围波动，所以对于给定的显著性水平 α，检验的拒绝域为

$$W_1=\left\{(x_1,x_2,\cdots,x_n)\,\middle|\,\frac{1}{\sigma_0^2}\sum_{i=1}^{n}(x_i-\mu)^2<\chi^2_{\alpha/2}(n)\text{ 或 }\frac{1}{\sigma_0^2}\sum_{i=1}^{n}(x_i-\mu)^2>\chi^2_{1-\alpha/2}(n)\right\}. \tag{3.2-7}$$

这个检验称为 χ^2 检验.

3.2 正态总体参数的假设检验

(4) μ 未知时 σ^2 的检验

要检验

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

但由于 μ 未知,所以 $\frac{1}{\sigma_0^2}\sum_{i=1}^{n}(X_i-\mu)^2$ 不是统计量. 因而自然会想到用样本均值 $\overline{X}$ 替代 μ,这样便得到统计量 $\frac{1}{\sigma_0^2}\sum_{i=1}^{n}(X_i-\overline{X})^2$,且当 H_0 成立时,$\frac{1}{\sigma_0^2}\sum_{i=1}^{n}(X_i-\overline{X})^2\sim\chi^2(n-1)$. 因此,取检验统计量为

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

从而对于给定的显著性水平 α,检验的拒绝域为

$$W_1=\left\{(x_1,x_2,\cdots,x_n)\,\middle|\,\frac{1}{\sigma_0^2}\sum_{i=1}^{n}(x_i-\overline{x})^2<\chi^2_{\alpha/2}(n-1)\right.$$

或 $$\left.\frac{1}{\sigma_0^2}\sum_{i=1}^{n}(x_i-\overline{x})>\chi^2_{1-\alpha/2}(n-1)\right\}.\tag{3.2-8}$$

这个检验也称为 χ^2 检验.

为了便于查阅,用表 3.2-1 和表 3.2-2 分别给出一个正态总体情形下总体均值和总体方差的各种检验的拒绝域,其显著性水平为 α.

3.2 正态总体参数的假设检验

表 3.2-2　一个正态总体下方差检验的拒绝域

假设		拒绝域 W_1 应满足的条件	
H_0	H_1	μ 已知	μ 未知
检验统计量及其分布		$\chi^2=\frac{1}{\sigma_0^2}\sum_{i=1}^{n}(X_i-\mu)^2\sim\chi^2(n)$ $(\sigma^2=\sigma_0^2)$	$\chi^2=\frac{1}{\sigma_0^2}\sum_{i=1}^{n}(X_i-\overline{X})^2\sim\chi^2(n-1)$ $(\sigma^2=\sigma_0^2)$
$\sigma^2=\sigma_0^2$	$\sigma^2\neq\sigma_0^2$	$\frac{1}{\sigma_0^2}\sum_{i=1}^{n}(x_i-\mu)^2<\chi_{\alpha/2}^2(n)$ 或 $\frac{1}{\sigma_0^2}\sum_{i=1}^{n}(x_i-\mu)^2>\chi_{1-\alpha/2}^2(n)$	$\frac{1}{\sigma_0^2}\sum_{i=1}^{n}(x_i-\overline{x})^2<\chi_{\alpha/2}^2(n-1)$ 或 $\frac{1}{\sigma_0^2}\sum_{i=1}^{n}(x_i-\overline{x})^2>\chi_{1-\alpha/2}^2(n-1)$
$\sigma^2\leqslant\sigma_0^2$	$\sigma^2>\sigma_0^2$	$\frac{1}{\sigma_0^2}\sum_{i=1}^{n}(x_i-\mu)^2>\chi_{1-\alpha}^2(n)$	$\frac{1}{\sigma_0^2}\sum_{i=1}^{n}(x_i-\overline{x})^2>\chi_{1-\alpha}^2(n-1)$
$\sigma^2\geqslant\sigma_0^2$	$\sigma^2<\sigma_0^2$	$\frac{1}{\sigma_0^2}\sum_{i=1}^{n}(x_i-\mu)^2<\chi_{\alpha}^2(n)$	$\frac{1}{\sigma_0^2}\sum_{i=1}^{n}(x_i-\overline{x})^2<\chi_{\alpha}^2(n-1)$

3.2 正态总体参数的假设检验

例 1 某种零件的尺寸服从正态分布，今从一批这种零件中随机抽取 6 件测得尺寸数据(单位：mm)如下：

$$32.50,\ 29.66,\ 31.64,\ 30.00,\ 31.87,\ 31.03$$

在显著性水平 $\alpha=0.05$ 与 $\alpha=0.01$ 下，能否认为这批零件尺寸的平均值为 32.5 mm?

解 由于检验

$$H_0:\mu=32.5,\quad H_1:\mu\neq 32.5,$$

且总体方差 σ^2 未知，所以应用 t 检验.

由测得数据算得 $\bar{x}=31.12$，$s=1.11$，故检验统计量的观测值为

$$|t|=\sqrt{n}\,\frac{|\bar{x}-32.5|}{s}=\sqrt{6}\times\frac{|31.12-32.5|}{1.11}=3.05.$$

在显著性水平 $\alpha=0.05$ 下，临界值 $t_{0.975}(5)=2.5703$. 由于$3.05>2.5703$，所以拒绝 H_0，即不能认为这批零件尺寸的平均值为 32.5 mm.

但在显著性水平 $\alpha=0.01$ 下，临界值 $t_{0.995}(5)=4.0322$. 由于 $3.05<4.0322$，所以不能拒绝 H_0，即可以认为这批零件尺寸的均值为 32.5 mm. ∎

3.2 正态总体参数的假设检验

例 2 从某厂生产的一批灯泡中随机地抽取 20 只进行寿命测试. 由测试结果算得 $\overline{x}=1960$ 小时，$s=200$ 小时. 假定灯泡寿命 $X\sim N(\mu;\sigma^2)$，其中 μ 与 σ^2 均未知. 在显著性水平 0.05 下能否认为这批灯泡的平均寿命达到国家标准 2000 小时？

解 因为 1960<2000，故考虑到厂方利益，为了控制符合国家标准的灯泡被误认为不合格，所以建立假设

$$H_0:\mu\geqslant 2000,\quad H_1:\mu<2000.$$

应用单侧的 t 检验. 由所给数据得到检验统计量的观测值为

$$t=\sqrt{n}\,\frac{\overline{x}-2000}{s}=\sqrt{20}\,\frac{1960-2000}{200}=-0.894,$$

而临界值 $t_{0.05}(20-1)=-t_{0.95}(19)=-1.729$. 显然，$-0.894>-1.729$，因此不能拒绝 H_0，即可以认为该厂生产的灯泡达到国家标准. ■

3.2 正态总体参数的假设检验

2. 两个正态总体的情形

设$(X_1,X_2,\cdots,X_m)$是取自正态总体$N(\mu_1;\sigma_1^2)$的样本，$(Y_1,Y_2,\cdots,Y_n)$是取自正态总体$N(\mu_2;\sigma_2^2)$的样本，并且这两个样本相互独立．讨论这两个正态总体的均值之间与方差之间差异的检验．

(1) σ_1^2 与 σ_2^2 均已知，$\mu_1-\mu_2$ 的检验

我们要检验假设

$$H_0:\mu_1-\mu_2=\delta,\quad H_1:\mu_1-\mu_2\neq\delta,$$

其中 $\delta\geqslant 0$ 是事先给定的常数，通常取 $\delta=0$．

$\mu_1-\mu_2$ 的极大似然估计量是 $\overline{X}-\overline{Y}$，且 $\overline{X}-\overline{Y}\sim N\left(\mu_1-\mu_2;\dfrac{\sigma_1^2}{m}+\dfrac{\sigma_2^2}{n}\right)$．

当 H_0 成立时，$\mu_1-\mu_2=\delta$，因而检验统计量取为

$$U=\frac{(\overline{X}-\overline{Y})-\delta}{\sqrt{\dfrac{\sigma_1^2}{m}+\dfrac{\sigma_2^2}{n}}},$$

且当 H_0 成立时，有 $U\sim N(0;1)$．于是应用 u 检验，对于给定的显著性水平 α，检验的拒绝域为

$$W_1=\left\{(x_1,x_2,\cdots,x_m;y_1,y_2,\cdots,y_n)\left|\frac{|(\overline{x}-\overline{y})-\delta|}{\sqrt{\dfrac{\sigma_1^2}{m}+\dfrac{\sigma_2^2}{n}}}>u_{1-\alpha/2}\right.\right\}.\qquad(3.2\text{-}9)$$

3.2 正态总体参数的假设检验

(2) $\sigma_1^2=\sigma_2^2=\sigma^2$ 但未知，$\mu_1-\mu_2$ 的检验

我们仍要检验

$$H_0:\mu_1-\mu_2=\delta,\quad H_1:\mu_1-\mu_2\neq\delta.$$

$\mu_1-\mu_2$ 的极大似然估计量仍是 $\overline{X}-\overline{Y}$，但由于 σ^2 未知，所以 U 不能作为检验统计量，故不能应用 u 检验. 考虑到

$$\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{S_w\sqrt{\dfrac{1}{m}+\dfrac{1}{n}}}\sim t(m+n-2),$$

且当 H_0 成立时，$\mu_1-\mu_2=\delta$，故可以取检验统计量为

$$T=\frac{(\overline{X}-\overline{Y})-\delta}{S_w\sqrt{\dfrac{1}{m}+\dfrac{1}{n}}},$$

其中

$$S_w^2=\frac{1}{m+n-2}\left[\sum_{i=1}^{m}(X_i-\overline{X})^2+\sum_{j=1}^{n}(Y_j-\overline{Y})^2\right].$$

于是，应用 t 检验可得，在显著性水平 α 下检验的拒绝域为

$$W_1=\left\{(x_1,x_2,\cdots,x_m;y_1,y_2,\cdots,y_n)\;\middle|\;\frac{|(\overline{x}-\overline{y})-\delta|}{S_w\sqrt{\dfrac{1}{m}+\dfrac{1}{n}}}>t_{1-\alpha/2}(m+n-2)\right\}.$$

(3.2-10)

3.2 正态总体参数的假设检验

(3) μ_1 与 μ_2 均未知，σ_1^2 与 σ_2^2 差异的检验

这时要检验假设

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

上述原假设 H_0 和备选假设 H_1 可以等价地表示为

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

考虑到

$$\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}\sim F(m-1,n-1),$$

且当 H_0 成立时，$\frac{\sigma_1^2}{\sigma_2^2}=1$，故取检验统计量为

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

当 H_0 成立时，$F=S_1^2/S_2^2\sim F(m-1,n-1)$. 于是，在显著性水平 α 下，检验的拒绝域为

$$W_1=\left\{(x_1,x_2,\cdots,x_m;y_1,y_2,\cdots,y_n)\,\middle|\,\frac{s_1^2}{s_2^2}<F_{\alpha/2}(m-1,n-1)\right.$$

$$\left.\text{或}\frac{s_1^2}{s_2^2}>F_{1-\alpha/2}(m-1,n-1)\right\}. \qquad (3.2\text{-}11)$$

这个检验称为 F 检验.

表 3.2-3 和表 3.2-4 分别给出了两个正态总体下它们的均值和方差之间差异的各种检验的拒绝域，其中显著性水平为 α.

3.2 正态总体参数的假设检验

表 3.2-3　两个正态总体的均值差检验的拒绝域

假设		拒绝域 W_1 应满足的条件	
H_0	H_1	σ_1^2,σ_2^2 均已知	$\sigma_1^2=\sigma_2^2=\sigma^2$ 但未知
检验统计量及其分布		$U=\dfrac{(\overline{X}-\overline{Y})-\delta}{\sqrt{\dfrac{\sigma_1^2}{m}+\dfrac{\sigma_2^2}{n}}}$ $\sim N(0,1)(\mu_1-\mu_2=\delta)$	$T=\dfrac{(\overline{X}-\overline{Y})-\delta}{S_w\sqrt{\dfrac{1}{m}+\dfrac{1}{n}}}$ $\sim t(m+n-2)(\mu_1-\mu_2=\delta)$
$\mu_1-\mu_2=\delta$	$\mu_1-\mu_2\neq\delta$	$\dfrac{\lvert(\overline{x}-\overline{y})-\delta\rvert}{\sqrt{\dfrac{\sigma_1^2}{m}+\dfrac{\sigma_2^2}{n}}}>u_{1-\alpha/2}$	$\dfrac{\lvert(\overline{x}-\overline{y})-\delta\rvert}{s_w\sqrt{\dfrac{1}{m}+\dfrac{1}{n}}}>t_{1-\alpha/2}(m+n-2)$
$\mu_1-\mu_2\leqslant\delta$	$\mu_1-\mu_2>\delta$	$\dfrac{(\overline{x}-\overline{y})-\delta}{\sqrt{\dfrac{\sigma_1^2}{m}+\dfrac{\sigma_2^2}{n}}}>u_{1-\alpha}$	$\dfrac{(\overline{x}-\overline{y})-\delta}{s_w\sqrt{\dfrac{1}{m}+\dfrac{1}{n}}}>t_{1-\alpha}(m+n-2)$
$\mu_1-\mu_2\geqslant\delta$	$\mu_1-\mu_2<\delta$	$\dfrac{(\overline{x}-\overline{y})-\delta}{\sqrt{\dfrac{\sigma_1^2}{m}+\dfrac{\sigma_2^2}{n}}}<-u_{1-\alpha}$	$\dfrac{(\overline{x}-\overline{y})-\delta}{s_w\sqrt{\dfrac{1}{m}+\dfrac{1}{n}}}<-t_{1-\alpha}(m+n-2)$

3.2 正态总体参数的假设检验

表 3.2-4　两个正态总体的方差之间差异检验的拒绝域

假设		拒绝域 W_1 应满足的条件	
H_0	H_1	μ_1、μ_2 均已知	μ_1、μ_2 均未知
检验统计量及其分布		$F=\dfrac{n\sum_{i=1}^{m}(X_i-\mu_1)^2}{m\sum_{i=1}^{n}(Y_i-\mu_2)^2}\sim F(m,n)$ $(\sigma_1^2=\sigma_2^2)$	$F=\dfrac{S_1^2}{S_2^2}\sim F(m-1,n-1)$ $(\sigma_1^2=\sigma_2^2)$
$\sigma_1^2=\sigma_2^2$	$\sigma_1^2\neq\sigma_2^2$	$\dfrac{n\sum_{i=1}^{m}(x_i-\mu_1)^2}{m\sum_{i=1}^{n}(y_i-\mu_2)^2}<F_{\alpha/2}(m,n)$ 或 $\dfrac{n\sum_{i=1}^{m}(x_i-\mu_1)^2}{m\sum_{i=1}^{n}(y_i-\mu_2)^2}>F_{1-\alpha/2}(m,n)$	$\dfrac{s_1^2}{s_2^2}<F_{\alpha/2}(m-1,n-1)$ 或 $\dfrac{s_1^2}{s_2^2}>F_{1-\alpha/2}(m-1,n-1)$
$\sigma_1^2\leqslant\sigma_2^2$	$\sigma_1^2>\sigma_2^2$	$\dfrac{n\sum_{i=1}^{m}(x_i-\mu_1)^2}{m\sum_{i=1}^{n}(y_i-\mu_2)^2}>F_{1-\alpha}(m,n)$	$\dfrac{s_1^2}{s_2^2}>F_{1-\alpha}(m-1,n-1)$
$\sigma_1^2\geqslant\sigma_2^2$	$\sigma_1^2<\sigma_2^2$	$\dfrac{n\sum_{i=1}^{m}(x_i-\mu_1)^2}{m\sum_{i=1}^{n}(y_i-\mu_2)^2}<F_{\alpha}(m,n)$	$\dfrac{s_1^2}{s_2^2}<F_{\alpha}(m-1,n-1)$

3.2 正态总体参数的假设检验

例 3 设有甲、乙两种零件，但乙种零件比甲种零件制造简单造价低. 现从这两种零件中随机地各抽取 5 件，经过测试得到抗压强度值(单位:10^5 帕)如下：

甲种零件	88	87	92	90	91
乙种零件	89	89	90	84	88

假定这两种零件的抗压强度都服从正态分布，且它们的方差相等，问在显著性水平 $\alpha=0.05$下，能否用乙种零件替代甲种零件？

解 根据题意，如果乙种零件的平均抗压强度不低于甲种零件，那么由于乙种零件制造简单，造价低，故可用乙种零件替代甲种零件. 于是，本题要检验乙种零件的平均抗压强度是否不低于甲种零件.

从所给的数据可算得甲种零件样本均值 $\overline{x}=89.6$，样本方差 $s_1^2=4.30$；乙种零件样本均值 $\overline{y}=88.0$，样本方差 $s_2^2=5.50$；并且有 $\overline{x}>\overline{y}$. 为了控制实际上乙种零件的平均抗压强度不低于甲种零件而误认为低于甲种零件的错误，需要检验假设

$$H_0: \mu_1 \leqslant \mu_2, \quad H_1: \mu_1 > \mu_2,$$

即

$$H_0: \mu_1 - \mu_2 \leqslant 0, \quad H_1: \mu_1 - \mu_2 > 0.$$

3.2 正态总体参数的假设检验

应用 t 检验，由 $m=n=5$，$s_w^2=\frac{1}{8}(4s_1^2+4s_2^2)=4.90$，得到检验统计量的观测值为

$$t=\frac{\bar{x}-\bar{y}}{s_w\sqrt{\frac{1}{m}+\frac{1}{n}}}=\frac{89.6-88.0}{\sqrt{4.90}\times\sqrt{\frac{1}{5}+\frac{1}{5}}}=1.1429,$$

而在显著性水平 $\alpha=0.05$ 下，临界值 $t_{1-\alpha}(m+n-2)=t_{0.95}(8)=1.8595$. 显然，$1.1429<1.8595$，因而不能拒绝 H_0，也即是可以用乙种零件替代甲种零件. ■

例 4 在平炉上进行一项试验，以确定改变操作方法的建议能否增加钢的得率（其定义是$\frac{\text{可用钢材}}{\text{投入炉中金属总量}}\times 100\%$）. 试验是在同一平炉上进行的，且每炼一炉钢，除操作方法外其余条件都尽可能做到相同. 先用标准方法炼一炉，再用所建议的方法炼一炉，交替进行试验，各炼了 10 炉，其得率分别为

标准方法	78.1	72.4	76.2	74.3	77.4	78.4	76.0	75.5	76.7	77.3
新方法	79.1	81.0	77.3	79.1	80.0	79.1	79.1	77.3	80.2	82.1

假定这两个样本相互独立且都来自正态总体，问在显著性水平$\alpha=0.50$下，是否可以认为方差相等？若相等，则在 $\alpha=0.01$ 下再检验新方法是否提高了得率.

3.2 正态总体参数的假设检验

解 由所给数据算得 $\bar{x}=76.23, s_1^2=3.325; \bar{y}=79.43, s_2^2=2.225$.

为了应用 t 检验，先要检验方差齐性，即检验假设

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

应用 F 检验，得到检验统计量的观测值为

$$\frac{s_1^2}{s_2^2}=\frac{3.325}{2.225}=1.494,$$

而在显著性水平 $\alpha=0.50$ 下，临界值 $F_{0.75}(9,9)=1.59, F_{0.25}(9,9)=\frac{1}{F_{0.75}(9,9)}=\frac{1}{1.59}=0.629$. 显然

$$0.629<1.494<1.59,$$

所以不能拒绝 H_0，即可以认为它们的方差相等.

由于 $\bar{y}>\bar{x}$，所以为了控制新方法实际上没有提高得率而误认为提高了得率的错误，故在显著性水平 $\alpha=0.10$ 下检验假设

$$H_0': \mu_1 \geqslant \mu_2, \quad H_1': \mu_1<\mu_2.$$

应用 t 检验，得到检验统计量的观测值为

$$\frac{\bar{x}-\bar{y}}{s_w\sqrt{\frac{1}{m}+\frac{1}{n}}}=\frac{76.23-79.43}{\sqrt{\frac{1}{18}(9s_1^2+9s_2^2)}\sqrt{\frac{1}{10}+\frac{1}{10}}}=\frac{76.23-79.43}{\sqrt{2.775}\times\sqrt{0.2}}=-4.295,$$

而在 $\alpha=0.01$ 下，临界值 $t_{0.99}(18)=2.5524$. 显然，$-4.295<-2.5524$，因而拒绝 H_0，即可以认为所建议的方法提高了得率. ∎

3.2 正态总体参数的假设检验

例 5 甲、乙两台机床生产同一型号的滚珠. 今从它们生产的滚珠中分别抽取了 8 个与 9 个，测得直径（单位：mm）的数据如下：

甲机床	15.0	14.5	15.2	15.5	14.8	15.1	15.2	14.8	
乙机床	15.2	15.0	14.8	15.2	15.0	15.0	15.1	14.8	14.8

假定滚珠直径服从正态分布，问两台机床生产的滚珠直径是否可以认为具有同一分布 $(\alpha=0.05)$？

解 如果是服从同一分布，那么 $\sigma_1^2=\sigma_2^2$ 及 $\mu_1=\mu_2$，因此要先检验方差齐性，再检验均值是否相等.

由所给数据算得 $\overline{x}=15.0125$，$\overline{y}=14.9889$，$s_1^2=0.0955$，$s_2^2=0.0261$，且 $m=8$，$n=9$，$s_w^2=0.0585$.

在显著性水平 $\alpha=0.05$ 下，检验方差齐性假设

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

应用 F 检验，得到检验统计量的观测值

$$\frac{s_1^2}{s_2^2}=\frac{0.0955}{0.0261}=3.6590,$$

而临界值 $F_{0.975}(7,8)=4.53$，$F_{0.025}(7,8)=\dfrac{1}{F_{0.975}(8,7)}=\dfrac{1}{4.90}=0.2041$. 显然，$0.2041<3.6590<4.53$，因此不能拒绝 H_0，即认为它们的方差相等.

3.2 正态总体参数的假设检验

再检验假设

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

应用 t 检验，得到检验统计量的观测值

$$\frac{\bar{x}-\bar{y}}{s_w\sqrt{\frac{1}{m}+\frac{1}{n}}} = \frac{15.0125-14.9889}{\sqrt{0.0585}\times\sqrt{\frac{1}{8}+\frac{1}{9}}} = 0.2008.$$

而在显著性水平 $\alpha=0.05$ 下，临界值 $t_{0.975}(15)=2.1315$，显然有 $|0.2008|<2.1315$，因此不能拒绝 H_0，即认为均值没有差异.

综合上述两个结果，可以认为两台机床生产的滚珠直径具有同一分布. ∎

3.3 非正态总体大样本参数检验

在正态总体参数检验中，由于检验统计量都有精确分布，因而对样本大小没有限制。对于非正态总体，一般难以找到具有精确分布的检验统计量。当样本容量 n 较大时，往往用检验统计量的渐近分布代替它的精确分布，从而得到近似的拒绝域。这里，点估计的渐近正态性起着重要的作用。

例 1 某县早稻收割面积为 100 万亩（1 亩＝666.7 m^2），现随机抽取 150 亩，得到平均亩产量 $\bar{x}=320$ kg，且样本标准差 $s=100$ kg，问在显著性水平 $\alpha=0.05$ 下，能否预计这 100 万亩早稻的平均亩产量为 340 kg？

解 要检验假设

$$H_0:E(X)=340,\quad H_1:E(X)\neq 340,$$

其中 X 表示亩产量。由于 $n=150$ 较大，所以近似地有

$$\sqrt{n}\frac{\overline{X}-E(X)}{S}\sim N(0;1).$$

当 H_0 成立时，$E(X)=340$。这时上式为

$$U=\sqrt{n}\frac{\overline{X}-340}{S}\sim N(0;1).$$

因此，由

$$P\{|U|>u_{1-\alpha/2}\}\approx\alpha$$

3.3 非正态总体大样本参数检验

得到检验的拒绝域为

$$W_1 = \left\{ (x_1, x_2, \cdots, x_n) \,\middle|\, \sqrt{150}\,\frac{|\overline{x} - 340|}{s} > u_{1-\alpha/2} \right\}.$$

现在，$\overline{x}=320$，$s=100$，从而算得检验统计量的观测值是

$$\sqrt{150}\,\frac{|320 - 340|}{100} = 2.4495,$$

而由 $\alpha=0.05$ 得 $1-\frac{\alpha}{2}=0.975$，查表可得分位数 $u_{0.975}=1.96$. 显然 $2.4495>1.96$，因而拒绝 H_0，即不能预计这 100 万亩的平均亩产量为 340 kg. ■

例 2 某厂生产一批产品，质量检查规定，其次品率不大于 1% 才能出厂销售. 现从这批产品中随机地抽取 100 件进行检查，发现有 3 件次品，问这批产品能否出厂？($\alpha=0.05$)

解 由于样本的次品率为 3%，大于 1%，所以考虑厂方利益要检验假设

$$H_0: p \leqslant 0.01, \quad H_1: p > 0.01,$$

其中 p 为该批产品的次品率. 总体 X 服从 0-1 分布 $B(1,p)$，p 的极大似然估计量是 $\overline{X}$，且近似地有

$$\sqrt{100}\,\frac{\overline{X} - p}{\sqrt{p(1-p)}} \sim N(0;1).$$

当 H_0 成立时 $p\leqslant 0.01$，从而取检验统计量为

3.3 非正态总体大样本参数检验

$$U=\sqrt{100}\,\frac{\overline{X}-0.01}{\sqrt{0.01\times(1-0.01)}}.$$

于是，在显著性水平 α 下，检验的拒绝域为

$$W_1\approx\left\{(x_1,x_2,\cdots,x_n)\;\middle|\;\sqrt{100}\,\frac{\overline{x}-0.01}{\sqrt{0.01\times(1-0.01)}}>u_{1-\alpha}\right\}.$$

现 $\overline{x}=0.03$，$\alpha=0.05$，故检验统计量的观测值为 $\sqrt{100}\times\dfrac{0.03-0.01}{\sqrt{0.0099}}=2.01$，而 $u_{1-\alpha}=u_{0.95}=1.645$. 显然，$2.01>1.645$，因而拒绝 H_0，即这批产品不能出厂销售.

3.3 非正态总体大样本参数检验

例 3 在两种工艺条件下纺出一批细纱，现随机地各抽取 100 个样品测试其强力(单位：牛)经计算得到在这两种工艺条件下样本均值和样本标准差为

甲工艺：$\bar{x}=2.80, s_1=0.280$,

乙工艺：$\bar{y}=2.86, s_2=0.285$,

问在 $\alpha=0.05$ 下，两种工艺条件纺出的细纱的平均强力有无显著差异？

解 根据题意，这是检验两个总体均值是否相等的问题，即检验假设

$$H_0: E(X)=E(Y), \quad H_1: E(X)\neq E(Y),$$

其中 X、Y 分别表示甲、乙两种工艺条件纺出的细纱总体. 取检验统计量

$$U=\frac{\bar{X}-\bar{Y}}{\sqrt{\frac{S_1^2}{m}+\frac{S_2^2}{n}}},$$

则当 H_0 成立时，U 的渐近分布为 $N(0;1)$. 于是在显著性水平 α 下，检验的拒绝域为

$$W_1\approx\left\{(x_1,x_2,\cdots,x_m;y_1,y_2,\cdots,y_n)\,\middle|\,\frac{|\bar{x}-\bar{y}|}{\sqrt{\frac{s_1^2}{m}+\frac{s_2^2}{n}}}>u_{1-\alpha/2}\right\}.$$

现在，$m=n=100, \bar{x}=2.80, \bar{y}=2.86, s_1^2=(0.280)^2, s_2^2=(0.285)^2$，由此得到检验统计量的观测值为

$$\frac{|2.80-2.86|}{\sqrt{\frac{(0.280)^2+(0.285)^2}{100}}}=1.5018,$$

而临界值 $u_{0.975}=1.96$. 显然，$1.5018<1.96$，因而不能拒绝 H_0，即在这两种工艺条件下纺出的细纱的平均强力无显著差异. ■

3.3 非正态总体大样本参数检验

例 4 为确定某种除虫剂对植物的影响，取 1000 株植物做试验. 在没有施这种除虫剂的 100 株植物中，有 53 株长势良好；而在施除虫剂的 900 株中，有 689 株长势良好. 问在显著性水平$\alpha=0.01$下这种除虫剂的效果是否显著.

解 根据题意要检验假设

$$H_0: p_1 \geqslant p_2, \quad H_1: p_1 < p_2,$$

其中 p_1 是没有施除虫剂的植物中长势良好的比例，而 p_2 是施除虫剂的植物中长势良好的比例. 取检验统计量

$$T=\frac{\overline{X}_m-\overline{Y}_n}{\sqrt{\frac{\overline{X}_m(1-\overline{X}_m)}{m}+\frac{\overline{Y}_n(1-\overline{Y}_n)}{n}}},$$

当 H_0 成立时，T 的分布近似为 $N(0;1)$. 由所给数据知，$m=100$，$n=900$，$\overline{x}_m=0.530$，$\overline{y}_n=0.766$，从而得到检验统计量的观测值

$$\frac{\overline{x}_m-\overline{y}_n}{\sqrt{\frac{\overline{x}_m(1-\overline{x}_m)}{m}+\frac{\overline{y}_n(1-\overline{y}_n)}{n}}}=\frac{0.530-0.766}{\sqrt{\frac{0.53\times 0.47}{100}+\frac{0.766\times 0.234}{900}}}=-4.550,$$

而临界值$-u_{1-\alpha}=-u_{0.99}=-2.326$. 显然，$-4.550<-2.326$，因此拒绝 H_0，即这种除虫剂的效果是显著的. ∎

3.4 检验的优劣

与估计问题相类似，对于一个假设检验问题可以构造出许多不同的检验，它们的Ⅰ类风险都至多为α. 例如，设总体$X\sim N(\mu;\sigma^2)$，其中μ与σ^2均未知，我们要在显著性水平α下检验假设

$$\mathrm{H}_0:\mu=\mu_0,\quad \mathrm{H}_1:\mu\neq\mu_0,$$

其中μ_0是常数，在3.2节中取检验统计量为

$$T=\sqrt{n}\,\frac{\overline{X}-\mu_0}{S},$$

从而得到的拒绝域为(3.2-6)式，即拒绝域是

$$W_1=\left\{(x_1,x_2,\cdots,x_n)\,\middle|\,\sqrt{n}\,\frac{|\overline{x}-\mu_0|}{s}>t_{1-\alpha/2}(n-1)\right\}.$$

不过，由于 $$P\{T<t_{2\alpha/3}(n-1)\quad 或\quad T>t_{1-\alpha/3}(n-1)\}=\alpha,$$

所以拒绝域也可以取为

$$\widetilde{W}_1=\left\{(x_1,x_2,\cdots,x_n)\,\middle|\,\sqrt{n}\,\frac{\overline{x}-\mu_0}{s}<t_{2\alpha/3}(n-1)\quad 或\quad \sqrt{n}\,\frac{\overline{x}-\mu_0}{s}>t_{1-\alpha/3}(n-1)\right\}.$$

这个检验当然与上一个检验不同. 那么，为什么在3.2节中检验的拒绝域取为W_1呢？这是在保证Ⅰ类风险至多为α的前提下，要使Ⅱ类风险尽可能小的原则所决定的. 本节对这一原则作一些初步的讨论，以便对检验的本质有进一步的理解.

3.4.1 功效函数

前面所提出的检验仅仅保证Ⅰ类风险至多为 α，至于Ⅱ类风险究竟有多大是不清楚的. 这对应用来说是不够的. 例如，厂方出售一批产品给买主，厂方声称其产品的不合格率不超过1%，如果用前述的检验假设

$$H_0: p \leqslant 0.01, \quad H_1: p > 0.01$$

的方法所规定的拒绝域来确定验收规则，那么厂方自然是满意的，因为这个检验保证当产品的不合格率实际上不超过1%而被买主拒收的概率不超过 α. 但买方是不放心的，他更关心的是产品不合格率实际超过1%而被他买下的可能性有多大，即他更希望知道使用这个检验将会面临多大的Ⅱ类风险.

给定一个参数型统计问题，设总体参数 θ 的取值范围为 Θ，即 $\theta \in \Theta$，$(X_1, X_2, \cdots, X_n)$ 是取自该总体的一个容量为 n 的样本. 未知参数 θ 的假设检验问题可以更清楚地表述如下：对于备选假设 $H_1: \theta \in \Theta_1$，检验原假设 $H_0: \theta \in \Theta_0$，这里的 Θ_0 与 Θ_1 是参数空间 Θ 的两个互不相交的真子集，记为

$$H_0: \theta \in \Theta_0, \quad H_1: \theta \in \Theta_1.$$

当集合 Θ_0（或 Θ_1）仅含一个元素时，则称假设 H_0（或假设 H_1）为简单假设，否则称为复合假设. 例如，对于正态总体 $N(\mu; \sigma^2)$，要检验假设

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

则当 σ^2 已知时，Θ_0 仅含一个元素 μ_0，所以 H_0 是简单假设，但 $\Theta_1 = (-\infty, \mu_0) \cup (\mu_0, +\infty)$，故 H_1 是复合假设.

3.4.1 功效函数

定义 3.4-1 对于检验问题

$$H_0:\theta\in\Theta_0,\quad H_1:\theta\in\Theta_1$$

来说，一个检验 ϕ 的拒绝域若为 W_1，则对任一 $\theta\in\Theta_1$，称概率

$$P_\theta\{(X_1,X_2,\cdots,X_n)\in W_1\}$$

为检验 ϕ 对于备选假设 H_1 在 θ 处的**功效**.

定义 3.4-1 中的 P_θ 表示按总体参数取 θ 值时的分布来计算随机事件的概率.

一个检验 ϕ 的Ⅱ类风险与其功效之间有着简单关系，即当 $\theta\in\Theta_1$ 时，Ⅱ类风险为

$$P_\theta\{(X_1,X_2,\cdots,X_n)\notin W_1\}=1-P_\theta\{(X_1,X_2,\cdots,X_n)\in W_1\}.\qquad(3.4\text{-}1)$$

当 H_1 为复合假设时，对于每一个 $\theta\in\Theta_1$，检验 ϕ 都有一个功效，因而可以把功效看成为 $\theta\in\Theta_1$ 的一个函数.

另一方面，如果 $\theta\in\Theta_0$，那么概率 $P_\theta\{(X_1,X_2,\cdots,X_n)\in W_1\}$ 恰是检验 ϕ 的Ⅰ类风险. 当 H_0 是复合假设时，对于每一个 $\theta\in\Theta_0$，检验 ϕ 都有一个Ⅰ类风险值.

综合上述两方面，可以定义检验 ϕ 的功效函数如下.

定义 3.4-2 设 W_1 是检验 ϕ

$$H_0:\theta\in\Theta_0,\quad H_1:\theta\in\Theta_1$$

的拒绝域，则函数

$$\beta(\theta)=P_\theta\{(X_1,X_2,\cdots,X_n)\in W_1\},\theta\in\Theta\qquad(3.4\text{-}2)$$

称为检验 ϕ 的**功效函数**，必要时记为 $\beta_\phi(\theta)$.

3.4.1 功效函数

按功效函数的定义可知，当 $\theta\in\Theta_0$ 时，$\beta_\phi(\theta)$ 是检验 ϕ 的 Ⅰ 类风险；而当 $\theta\in\Theta_1$ 时，$\beta_\phi(\theta)$ 是检验 ϕ 对于备选假设 H_1 在 θ 处的功效，且 $1-\beta_\phi(\theta)$ 是检验 ϕ 的 Ⅱ 类风险. 因此，功效函数通过拒绝域 W_1 把两类风险包含在一个统一的形式之中.

例 1 设 $(X_1,X_2,\cdots,X_n)$ 是取自正态总体 $N(\mu;1)$ 的样本，其中 μ 未知，但 $\mu\in\Theta=\{\mu_0,\mu_1\}$，且 $\mu_0<\mu_1$. 对于检验假设

$$H_0:\mu=\mu_0,\quad H_1:\mu=\mu_1,$$

H_0 与 H_1 显然都是简单假设，应用 u 检验得到拒绝域为

$$W_1=\{(x_1,x_2,\cdots,x_n)\mid\sqrt{n}(\overline{x}-\mu_0)>u_{1-\alpha}\},$$

其中 α 为显著性水平. 这个检验的功效函数为

$$\begin{aligned}\beta(\mu)&=P_\mu\{(X_1,X_2,\cdots,X_n)\in W_1\}=P_\mu\{\sqrt{n}(\overline{X}-\mu_0)>u_{1-\alpha}\}\\&=P_\mu\{\sqrt{n}(\overline{X}-\mu)>u_{1-\alpha}-\sqrt{n}(\mu-\mu_0)\}\\&=1-P_\mu\{\sqrt{n}(\overline{X}-\mu)\leqslant u_{1-\alpha}-\sqrt{n}(\mu-\mu_0)\}.\end{aligned}$$

由于 $\overline{X}\sim N(\mu;\frac{1}{n})$，故 $\sqrt{n}(\overline{X}-\mu)\sim N(0;1)$，所以记 $\Phi(x)$ 为标准正态分布 $N(0;1)$ 的分布函数，则有

$$\beta(\mu)=1-\Phi(u_{1-\alpha}-\sqrt{n}(\mu-\mu_0)),\mu\in\Theta=\{\mu_0;\mu_1\}.$$

3.4.1 功效函数

由此看出，当 $\mu>\mu_0$，即 $\mu-\mu_0>0$ 时，$\beta(\mu)$ 是随 μ 的增大而单调增大的，且 $\beta(\mu_0)=1-\Phi(u_{1-\alpha})=1-(1-\alpha)=\alpha$，因而 $\beta(\mu_1)>\alpha$. 这个检验的Ⅱ类风险为

$$1-\beta(\mu_1)=\Phi(u_{1-\alpha}-\sqrt{n}(\mu_1-\mu_0)). \tag{3.4-1}$$

由(3.4-1)式可知，对于固定的样本容量 n，当 α 减小(即Ⅰ类风险减小)时，由于 $u_{1-\alpha}$ 增大而使 $1-\beta(\mu_1)$ 增大，即Ⅱ类风险增大；反之，若 α 增大，则由于 $u_{1-\alpha}$ 减小而使 $1-\beta(\mu_1)$ 减小，即Ⅱ类风险减小. 这就说明了，当样本容量固定时，Ⅰ类风险变小会使Ⅱ类风险增大，而Ⅱ类风险变小会使Ⅰ类风险变大. 在给定显著性水平 α 下，要使Ⅱ类风险不大于 ε，则只有确定适当的样本容量 n，使下式成立：

$$\Phi(u_{1-\alpha}-\sqrt{n}(\mu_1-\mu_0))\leqslant\varepsilon. \tag{3.4-2}$$

解不等式(3.4-2)得

$$\sqrt{n}\geqslant\frac{u_{1-\alpha}-u_\varepsilon}{\mu_1-\mu_0},\quad 即\quad n\geqslant\left(\frac{u_{1-\alpha}-u_\varepsilon}{\mu_1-\mu_0}\right)^2. \qquad ▌$$

3.4.1 功效函数

例 2 设$(X_1,X_2,\cdots,X_n)$是取自正态总体 $N(\mu;1)$的样本，其中 μ 未知，但 $\mu\in\Theta=(-\infty,+\infty)$. 对于检验假设

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

应用 u 检验得到拒绝域为

$$W_1=\{(x_1,x_2,\cdots,x_n)\mid\sqrt{n}\mid\overline{x}-\mu_0\mid>u_{1-\alpha/2}\}.$$

这个 u 检验的功效函数为

$$\begin{aligned}\beta(\mu)&=P_\mu\{(X_1,X_2,\cdots,X_n)\in W_1\}=P_\mu\{\sqrt{n}\mid\overline{X}-\mu_0\mid>u_{1-\alpha/2}\}\\&=1-P_\mu\{\sqrt{n}\mid\overline{X}-\mu_0\mid\leqslant u_{1-\alpha/2}\}=1-P_\mu\{-u_{1-\alpha/2}-\sqrt{n}(\mu-\mu_0)\\&\quad\leqslant\sqrt{n}(\overline{X}-\mu)\leqslant u_{1-\alpha/2}-\sqrt{n}(\mu-\mu_0)\}\\&=1-[\Phi(u_{1-\alpha/2}-\sqrt{n}(\mu-\mu_0))-\Phi(-u_{1-\alpha/2}\\&\quad-\sqrt{n}(\mu-\mu_0))],\ -\infty<\mu<+\infty.\end{aligned}$$

由于 $1-\Phi(u_{1-\alpha/2}-\sqrt{n}(\mu-\mu_0))=\Phi(-u_{1-\alpha/2}+\sqrt{n}(\mu-\mu_0))$，所以有

$$\beta(\mu)=\Phi(-u_{1-\alpha/2}+\sqrt{n}(\mu-\mu_0))+\Phi(-u_{1-\alpha/2}-\sqrt{n}(\mu-\mu_0)),\qquad(3.4\text{-}3)$$

故函数 $\beta(\mu)$的图形关于直线 $\mu=\mu_0$ 是左右对称的，如图 3.4-1 所示.

$$\beta(\mu_0)=\Phi(-u_{1-\alpha/2})+\Phi(-u_{1-\alpha/2})=\Phi(u_{\alpha/2})+\Phi(u_{\alpha/2})=\alpha,$$

当 $\mu\neq\mu_0$ 时，$\beta(\mu)>\alpha$.

这个 u 检验的Ⅱ类风险为

$$1-\beta(\mu)=\Phi(u_{1-\alpha/2}-\sqrt{n}(\mu-\mu_0))$$

3.4.1 功效函数

$$-\Phi(-u_{1-\alpha/2}-\sqrt{n}(\mu-\mu_0)). \qquad (3.4\text{-}4)$$

对于固定的样本容量 n，当 $\mu\to\pm\infty$ 时，$1-\beta(\mu)\to 0$，故当实际的总体均值 μ 偏离原假设较大时，这个 u 检验的效果较好，但当实际的总体均值 μ 接近原假设时，由于

$$\lim_{\mu\to\mu_0}(1-\beta(\mu))=1-\alpha,$$

而 α 较小，故 $1-\alpha$ 较大，所以这个 u 检验的效果不好. 这表明，无论样本容量 n 多大，要想对所有的 $\mu\in H_1$，Ⅱ类风险都很小是不可能的. 但是，可以使用(3.4-4)式确定样本容量 n，使得当 $|\mu-\mu_0|\geqslant\delta(\delta>0)$ 时，Ⅱ类风险不大于 ε. 事实上，由图 3.4-1 和(3.4-4)式知，n 应满足下式：

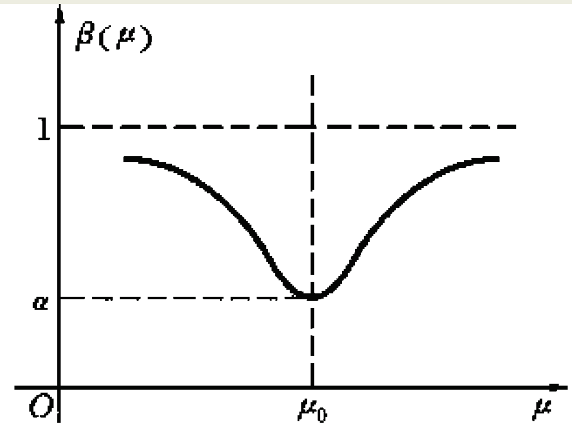

图 3.4-1

$$\Phi(u_{1-\alpha/2}-\sqrt{n}\delta)-\Phi(-u_{1-\alpha/2}-\sqrt{n}\delta)\leqslant\varepsilon.$$

通常，因 n 较大，故有 $\Phi(-u_{1-\alpha/2}-\sqrt{n}\delta)\approx 0$，从而近似地有

$$\Phi(u_{1-\alpha/2}-\sqrt{n}\delta)\leqslant\varepsilon.$$

由此可知，只要样本容量满足

$$u_{1-\alpha/2}-\sqrt{n}\delta\leqslant u_\varepsilon\text{，即 } n\geqslant\left(\frac{u_{1-\alpha/2}-u_\varepsilon}{\delta}\right)^2$$

就能使 $\mu\in H_1$ 且 $|\mu-\mu_0|\geqslant\delta$ 时，犯第二类错误的概率不超过 ε. ∎

3.4.1 功效函数

例 3 设$(X_1,X_2,\cdots,X_n)$是取自正态总体 $N(\mu;1)$的样本，其中 μ 未知，但 $\mu\in\Theta=(-\infty,+\infty)$. 我们要检验复合假设

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

应用 u 检验得到拒绝域为

$$W_1=\{(x_1,x_2,\cdots,x_n)\mid\sqrt{n}(\overline{x}-\mu_0)>u_{1-\alpha}\}.$$

这个 u 检验的功效函数为

$$\beta(\mu)=P_\mu\{\sqrt{n}(\overline{X}-\mu_0)>u_{1-\alpha}\}=P_\mu\{\sqrt{n}(\overline{X}-\mu)>u_{1-\alpha}-\sqrt{n}(\mu-\mu_\alpha)\}$$
$$=1-\Phi(u_{1-\alpha}-\sqrt{n}(\mu-\mu_0)),\mu\in(-\infty,+\infty).$$

其图形如图 3.4-2 所示.

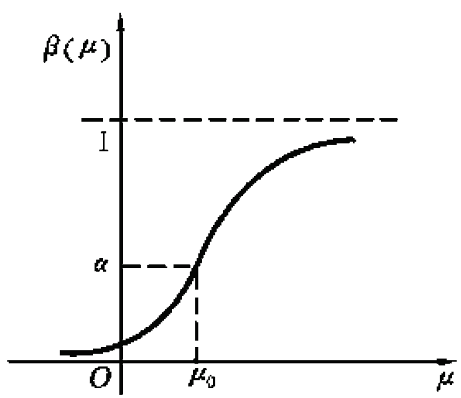

图 3.4-2

由图 3.4-2 看出，当 H_0 成立，即 $\mu\leqslant\mu_0$ 时，$\beta(\mu)\leqslant\alpha$，从而这个 u 检验符合显著性水平为 α 的要求，即Ⅰ类风险至多为 α. 这个 u 检验的Ⅱ类风险为

$$1-\beta(\mu)=\Phi(u_{1-\alpha}-\sqrt{n}(\mu-\mu_0)).\qquad(3.4\text{-}5)$$

当 $\mu>\mu_0$ 时，$1-\beta(\mu)<1-\alpha$，且 $\mu\to+\infty$时，$1-\beta(\mu)\to0$. 而当 $\mu\leqslant\mu_0$ 时，$1-\beta(\mu)\geqslant1-\alpha$.

容易看出，当标志Ⅰ类风险的显著性水平的 α 减小(增大)时，Ⅱ类风险会相应地增大(减小)；并且不论 n 多么大，只要 n 给定，总存在 μ_0 附近的点 $\mu(\mu>\mu_0)$使 $1-\beta(\mu)$几乎等于 $1-\alpha$. 但与例 2 一样，只要 $\mu-\mu_0\geqslant\delta(\delta>0$ 是给定的常数)，则可以利用

3.4.1 功效函数

(3.4-5)式确定样本容量 n，使Ⅱ类风险不大于 ε. 例如，设 $\mu_0=0$，$\alpha=0.05$，现希望当 $\mu\geqslant 1$ 时，这个 u 检验的Ⅱ类风险不大于 0.10，那么 n 必须满足

$$\Phi(u_{0.95}-\sqrt{n}(1-0))\leqslant 0.10,$$

即

$$u_{0.95}-\sqrt{n}\leqslant u_{0.10}=-u_{0.90}.$$

于是

$$\sqrt{n}\geqslant u_{0.95}+u_{0.90}=1.645+1.282=2.927,$$

即 $n\geqslant 8.57$.

因此，样本容量至少要 9 才能使这个 u 检验的Ⅱ类风险不大于 0.10. ∎

3.4.2 最大功效检验

如前所述，对于同一个假设检验问题，在相同的显著性水平 α 下可以给出不同的检验，这些检验的功效函数是不同的. 于是，产生一个如何选择最优检验的问题. Neyman-Pearson 提出一个评选标准：在Ⅰ类风险满足显著性水平 α 下使Ⅱ类风险尽可能小，即要求这个检验的功效函数 $\beta(\theta)$ 满足下述条件：

$$\begin{cases} \beta(\theta) \leqslant \alpha, & \theta \in \Theta_0, \\ \beta(\theta)\ \text{尽可能大}, & \theta \in \Theta_1. \end{cases} \tag{3.4-6}$$

定义 3.4-3 给定一个参数型统计问题，其总体参数 $\theta \in \Theta$，要检验假设

$$H_0: \theta \in \Theta_0, \quad H_1: \theta \in \Theta_1.$$

如果存在一个显著性水平 α 的检验 ϕ^*，使得对于任意一个显著性水平 α 的检验 ϕ，均有

$$\beta_{\phi^*}(\theta) \geqslant \beta_\phi(\theta), \quad \forall \theta \in \Theta_1,$$

则称 ϕ^* 为这个假设检验问题在显著性水平 α 下的**一致最大功效检验**. 当 H_1 为简单假设时，则称 ϕ^* 为**最大功效检验**.

如何寻找一致最大功效检验呢？Neyman-Pearson 就 H_0 和 H_1 都为简单假设的情形给出了构造最大功效检验的方法.

3.4.2 最大功效检验

定理 3.4-1（Neyman-Pearson 定理） 设总体 X 的概率密度（或概率函数）为 $f(x;\theta)$，$\theta\in\Theta=\{\theta_0,\theta_1\}$，$(X_1,X_2,\cdots,X_n)$是取自这个总体的一个样本，要检验简单假设

$$\mathrm{H}_0:\theta=\theta_0,\quad \mathrm{H}_1:\theta=\theta_1.$$

给定显著性水平 α，如果存在临界值 k，使

$$P_{\theta_0}\left\{\frac{\prod_{i=1}^{n} f(X_i;\theta_1)}{\prod_{i=1}^{n} f(X_i;\theta_0)}>k\right\}=\alpha, \tag{3.4-7}$$

那么，以

$$W_1^*=\left\{(x_1,x_2,\cdots,x_n)\,\middle|\,\frac{\prod_{i=1}^{n} f(x_i;\theta_1)}{\prod_{i=1}^{n} f(x_i;\theta_0)}>k\right\} \tag{3.4-8}$$

为拒绝域的检验 ϕ^* 是该假设检验问题在显著性水平 α 下的最大功效检验.

这个定理的证明可参阅有关文献，例如[7]. $\blacksquare$

这里说明一下 Neyman-Pearson 定理的含义. 它本质上是利用检验统计量

3.4.2 最大功效检验

$$l(X_1,X_2,\cdots,X_n)=\frac{\prod_{i=1}^{n}f(X_i;\theta_1)}{\prod_{i=1}^{n}f(X_i;\theta_0)},$$

而 $l(x_1,x_2,\cdots,x_n)$是似然函数 $L(\theta)$在 θ_1 与 θ_0 处的值之比. 由于似然函数 $L(\theta)$可以看成为当参数取 θ 值时，样本$(X_1,X_2,\cdots,X_n)$落在$(x_1,x_2,\cdots,x_n)$的邻域内(或取$(x_1,x_2,\cdots,x_n)$)的可能性大小的一个度量，因此当 H_0 成立即 $\theta=\theta_0$ 时，$l(x_1,x_2,\cdots,x_n)$的值应该较小；反之，如果根据样本观测值$(x_1,x_2,\cdots,x_n)$发现 $l(x_1,x_2,\cdots,x_n)$较大时，自然有理由认为 H_0 可能根本不成立.

3.4.2 最大功效检验

例 4 设总体分布为离散型，其概率函数 $f(x;\theta)$ 由下表给出，其中 $\theta \in \{0,1\}$.

θ \ x ($f(x;\theta)$)	0	1	2
0	0.8	0.1	0.1
1	0.2	0.6	0.2

X_1 是取自这个总体的大小为 1 的样本，要检验

$$H_0:\theta = 0, \quad H_1:\theta = 1.$$

容易算得

$$l(x_1) = \begin{cases} 0.25, & x_1 = 0, \\ 6, & x_1 = 1, \\ 2, & x_1 = 2. \end{cases}$$

在显著性水平 $\alpha = 0.10$ 下，最大功效检验 ϕ^* 给出的拒绝域为

$$W_1^* = \{x_1 \mid x_1 = 1\}.$$

事实上，这个检验在 $\theta = 1$ 处的功效为

$$\beta_{\phi^*}(1) = P_{\theta=1}\{X_1 \in W_1^*\} = P_{\theta=1}\{X_1 = 1\} = 0.6.$$

从而这个检验的 Ⅰ 类风险为 0.1，Ⅱ 类风险为 $1-0.6=0.4$. 如果考虑另一个检验 ϕ，它的拒绝域为

$$W_1 = \{x_1 \mid x_1 = 2\}.$$

显然，它的 Ⅰ 类风险仍为 0.1，但 Ⅱ 类风险为

$$1-\beta_\phi(1) = 1 - P_{\theta=1}\{X_1 = 2\} = 1 - 0.2 = 0.8. \qquad \blacksquare$$

3.4.2 最大功效检验

例 5 设$(X_1, X_2, \cdots, X_n)$是取自正态总体$N(\mu;1)$的一个样本，其中μ未知，要检验

$$\mathrm{H}_0: \mu = \mu_0, \quad \mathrm{H}_1: \mu = \mu_1.$$

其中$\mu_0 < \mu_1$. 在显著性水平α下，求最大功效检验的拒绝域.

解 由于

$$\prod_{i=1}^{n} f(x_i;\mu_1) = (2\pi)^{-\frac{n}{2}} \exp\left[-\frac{1}{2}\sum_{i=1}^{n}(x_i - \mu_1)^2\right],$$

$$\prod_{i=1}^{n} f(x_i;\mu_0) = (2\pi)^{-\frac{n}{2}} \exp\left[-\frac{1}{2}\sum_{i=1}^{n}(x_i - \mu_0)^2\right],$$

所以

$$l(x_1, x_2, \cdots, x_n) = \exp\left[n(\mu_1 - \mu_0)\bar{x} - \frac{n}{2}(\mu_1^2 - \mu_0^2)\right].$$

由 Neyman-Pearson 定理知，最大功效检验的拒绝域为

3.4.2 最大功效检验

$$
\begin{aligned}
W_1^* &= \{(x_1, x_2, \cdots, x_n) \mid l(x_1, x_2, \cdots, x_n) > k\} \\
&= \{(x_1, x_2, \cdots, x_n) \mid \bar{x} > k_1\} \\
&= \{(x_1, x_2, \cdots, x_n) \mid \sqrt{n}(\bar{x} - \mu_0) > k_0\},
\end{aligned}
$$

其中 $k_1 = [\ln k + \frac{n}{2}(\mu_1^2 - \mu_0^2)]/n(\mu_1 - \mu_0)$， $k_0 = \sqrt{n}(k_1 - \mu_0)$.

当 H_0 成立，即 $\mu = \mu_0$ 时，$\sqrt{n}(\bar{X} - \mu_0) \sim N(0;1)$. 因此，由(3.4-7)式可知，取临界值 $k_0 = u_{1-\alpha}$ 便有

$$P_{\mu_0}\{\sqrt{n}(\bar{X} - \mu_0) > u_{1-\alpha}\} = \alpha.$$

从而，最大功效检验的拒绝域

$$W_1^* = \{(x_1, x_2, \cdots, x_n) \mid \sqrt{n}(\bar{x} - \mu_0) > u_{1-\alpha}\}. \quad \blacksquare$$

3.4.2 最大功效检验

例 6 设$(X_1,X_2,\cdots,X_n)$是取自 $X\sim N(\mu;1)$的一个样本，其中 μ 未知. 证明对于单侧假设检验问题

$$H_0:\mu\leqslant\mu_0,\quad H_1:\mu>\mu_0,$$

例 5 给出的拒绝域为

$$W_1^*=\{(x_1,x_2,\cdots,x_n)\mid\sqrt{n}(\bar{x}-\mu_0)>u_{1-\alpha}\}$$

的 u 检验 ϕ^* 是显著性水平 α 下的一致最大功效检验.

证 对任一固定的 $\mu_1>\mu_0$，对于检验

$$H_0:\mu=\mu_0,\quad H_1:\mu=\mu_1,$$

由例 5 知，ϕ^* 是显著性水平 α 下的最大功效检验. 由于 ϕ^* 的拒绝域 W_1^* 与 μ_1 的取值无关，所以 ϕ^* 也是对于检验

$$H_0:\mu=\mu_0,\quad H_1:\mu>\mu_0 \qquad (*)$$

的一致最大功效检验.

设 ϕ 是原假设问题

$$H_0:\mu\leqslant\mu_0,\quad H_1:\mu>\mu_0$$

的任意一个显著性水平 α 下的检验，它的拒绝域为 W_1. 由显著性水平 α 的要求知，当 $\mu\leqslant\mu_0$ 时，

$$P_\mu\{(X_1,X_2,\cdots,X_n)\in W_1\}\leqslant\alpha,$$

3.4.2 最大功效检验

因而 ϕ 也是假设检验问题(*)的一个显著性水平 α 下的检验.但因为 ϕ^* 是检验问题(*)的一致最大功效检验,所以对任一 $\mu>\mu_0$,有

$$\beta_{\phi^*}(\mu) \geqslant \beta_{\phi}(\mu).$$

再则,由例 3 知,本例所给出的检验 ϕ^* 是原假设问题的一个显著性水平 α 下的检验.因此,ϕ^* 是这个检验问题的一致最大功效检验. ■

可以类似地证明,表 3.2-1 和表 3.2-2 中仅含一个总体参数的单侧检验都是一致最大功效检验.

求一致最大功效检验,一般来说是很困难的,更为严重的是,一致最大功效检验不一定存在.有兴趣的读者可以参阅理论性较强的数理统计教材,例如文献[7].

作业

习题 3

6. 从两批电子元件中随机抽取一些样品，测得它们的电阻(单位:欧姆)如下:

甲批	0.140	0.138	0.143	0.142	0.144	0.137
乙批	0.135	0.140	0.142	0.136	0.138	0.140

假定这两批电子元件的电阻都服从正态分布，问在显著性水平 0.05 下，能否认为这两个正态总体的方差相等?

8. 电话交换台每分钟接到呼唤的次数服从 Poisson 分布 $P(\lambda)$，今观测了 100 个时段，每个时段一分钟，总共有 585 次呼唤. 问在显著性水平 0.10 下，能否认为该电话交换台每分钟接到呼唤的次数服从 $\lambda=6$的 Poisson 分布?

10. 设$(X_1,X_2,\cdots,X_n)$是取自正态总体 $N(\mu;1)$的样本，其中 μ 未知，要检验假设

$$H_0:\mu\geqslant 0,\quad H_1:\mu<0.$$

在显著性水平 α 下，采用拒绝域为

$$W_1=\{(x_1,x_2,\cdots,x_n)\,|\,\sqrt{n}\bar{x}<-u_{1-\alpha}\}$$

的 u 检验.

(1) 求这个 u 检验的功效函数 $\beta(\mu)$;

(2) 当 $\alpha=0.05$ 时，如果要求 $\mu\leqslant -0.1$ 时这个 u 检验的 Ⅱ 类风险不大于 0.05，那么样本容量 n 至少应取多大?

謝謝觀看！

厦門大學
XIAMEN UNIVERSITY

信息学院
（国家示范性软件学院）
School of Informatics

黄 煒
博士，副教授
Dr. Wei Huang