第八章　假设检验

一、假设检验的基本概念

二、正态总体参数的假设检验

在总体X的分布完全未知, 或只知其分布但不知其参数的情况下, 为了推断总体的某些未知特性, 我们需对X的分布或分布中的参数作出某种假设,

例如, 提出总体服从泊松分布的假设, 又如, 对正态总体提出数学期望等于μ_0的假设.

我们要建立一个检验法则，然后根据样本观测值, 利用统计分析的方法检验这一假设是否合理,从而作出接受或者拒绝这一假设的决策.

假设检验是作出这一决策的过程.

假设检验：根据试验目的提出假设，通过试验结果提供的信息（样本）对假设作出推断，作出拒绝或接受假设的决策。

第一节　　假设检验的基本概念

一、　假设检验的基本原理

二、　假设检验的相关概念

三、　假设检验的一般步骤

引例　某产品出厂检验规定: 次品率p不超过4%才能出厂. 现从一万件产品中任意抽查12件发现3件次品, 问该批产品能否出厂?

解 假设 $\boxed{p \le 0.04}$,则

$$P_{12}(3) = C_{12}^{3} p^{3}(1-p)^{9} \le 0.0097 < \boxed{0.01}$$

这是 小概率事件，一般在一次试验中是不会发生的,

现一次试验竟然发生, 故认为原假设不成立,

即该批产品次品率p>0.04, 该批产品不能出厂.

注　本检验方法是概率意义下的反证法.

一、假设检验的基本原理与方法

基本原理：

实际推断原理："一个小概率事件在一次试验中几乎是不可能发生的"．

基本方法： 概率意义下的反证法

(1) 提出假设；

(2) 根据试验结果（样本）计算概率；

(3) 推断小概率事件是否发生？

若发生，拒绝假设；
不发生，接受假设。

对总体X的分布或分布中的未知参数提出假设, 称为统计假设.

对立 → 原假设或零假设 H_0： $\boxed{p \le 0.04}$,

→ 备择假设或对立假设 H_1： $p > 0.04$,

假设检验的任务： 必须在原假设与备择假设之间作一选择.

在原假设下，利用样本观测值对实际问题进行分析，如果发生小概率事件，则拒绝接受原假设（接受备择假设）。
如果发生大概率事件，则接受原假设（拒绝备择假设）。
这种统计推断的问题 称为假设检验问题。

参数假设检验----只对总体中的参数提出假设进行检验。

非参数假设检验----对总体的分布提出假设进行检验。

例8.1.1 某车间生产一种零件，其长度X是一个随机变量, $X\sim N(\mu,0.5^2)$, 要求标准长度为100cm. 现从一批数量很大的零件中随机抽查9件，测得它们的长度(单位：cm)如下：

99.3 100.1 99.9 99.2 99.6 99.1 99.3 100.2 99.0,

问能否认为这批零件是合格的（取显著性水平α=0.05）？

已知$X \sim N(\mu, 0.5^2)$, 其中μ未知.

问题: 根据样本值判断 $\mu = 100$还是 $\mu \neq 100$.

提出两个互为对立的假设 $H_0: \mu = \mu_0 = 100$ 和 $H_1: \mu \neq \mu_0$.

再利用已知样本作出判断是接受假设 H_0(拒绝假设 H_1), 还是拒绝假设 H_0(接受假设 H_1).

如果作出的判断是接受 H_0, 则 $\mu = \mu_0$,

即认为这批零件是合格的, 否则, 认为是不合格的.

因为 $\overline{X}$ 是 μ 的无偏估计量，

所以若 H_0 为真，则 $|\overline{x}-\mu_0|$ 不应太大，

当 H_0 为真时，$\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$，

衡量 $|\overline{x}-\mu_0|$ 的大小可归结为衡量 $\dfrac{|\overline{x}-\mu_0|}{\sigma/\sqrt{n}}$ 的大小，

于是可以选定一个适当的正数 k，

当观察值 $\bar{x}$ 满足 $\dfrac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}} \geq k$时，拒绝假设$H_0$，

反之,当观察值 $\bar{x}$ 满足 $\dfrac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}} < k$时，接受假设$H_0$.

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

对于给定的小概率值α (0<α<1), 令

$$P\{\frac{|\bar{X}-\mu_0|}{\sigma/\sqrt{n}} \geq k\} = \alpha,$$

由标准正态分布分位点的定义得 $k=u_{\alpha/2}$，

当$\dfrac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}} \geq u_{\alpha/2}$时,拒绝$H_0$, $\dfrac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}} < u_{\alpha/2}$时，接受$H_0$.

假设检验过程如下:

在实例中若取定 $\alpha = 0.05$,

则 $k = u_{\alpha/2} = u_{0.025} = 1.96$,

又已知 $n = 9$, $\sigma = 0.5$,

由样本算得 $\bar{x} = 99.52$, 即有 $\dfrac{|\bar{x} - \mu_0|}{\sigma / \sqrt{n}} = \dfrac{|99.52 - 100|}{0.5 / \sqrt{9}} = 2.88 > 1.96$,

于是拒绝假设H_0, 认为这批零件不合格。

二、假设检验的相关概念

显著性水平

在例8.1.1中，为使 $P\{\frac{|\overline{X}-\mu_0|}{\sigma/\sqrt{n}} \geq k\} = \alpha$,

当样本容量固定时,选定α后,数 k 就可以确定为$u_{\alpha/2}$,然后按照

$u = \dfrac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}$ 的观察值的绝对值大于等于 k 还是小于 k来作决定.

如果$|u| = \left|\dfrac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}\right| \geq u_{\alpha/2}$,则称 $\overline{x}$ 与μ_0的差异是显著的,

则我们拒绝 H_0,

反之,如果$|u| = \left|\dfrac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}\right| < u_{\alpha/2}$,则称 $\overline{x}$ 与μ_0的差异是不显著的,

则我们接受 H_0,

上述关于 $\overline{x}$ 与 μ_0 有无显著差异的判断是 在显著性水平 α 之下作出的.

显著性水平α

α=P{小概率事件|假设成立},

α越小（大），小概率事件越不容易（容易）发生，越容易拒绝（接受）原假设。

α 通常取为0.01，0.05或0.1.

假设检验就是根据样本, 构造一个适当的统计量, 按照某种规则, 决定是接受H_0(拒绝H_1)还是拒绝H_0 (接受H_1), 所使用的统计量称为检验统计量.

$$u=\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$$ —— 检验统计量。

当检验统计量取某个区域W中的值时, 我们就拒绝原假设H_0, 则称区域W为拒绝域. 拒绝域的边界点称为临界点.

当检验统计量的值不属于拒绝域W（即属于W^c）时, 我们就接受原假设H_0, 则称此区域为接受域.

如在前面实例中,

拒绝域为 $|u|\geq u_{\alpha/2}$,

临界点为 $u=-u_{\alpha/2}$, $u=u_{\alpha/2}$.

两类错误及记号

假设检验作决策的依据是样本， 样本取值有随机性，因而假设检验所作出的决定有可能是错误的。 这种错误有两类：

(1) 当原假设H_0为真, 观察值却落入拒绝域, 而作出了拒绝H_0的判断, 此类错误称做**第一类错误**, 又叫**弃真错误**, 这类错误是“以真为假”。 犯第一类错误的概率P{拒绝H_0|H_0为真}

不超过显著性水平α.

(2) 当原假设 H_0 不真, 而观察值却落入接受域, 而作出了接受 H_0 的判断, 此类错误称做**第二类错误**, 又叫**取伪错误**, 这类错误是“以假为真”。

犯第二类错误的概率记为β. $\beta = P\{$接受 H_0 | 当 H_0 不真$\}$.

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I类错误
H_0 不真	犯第II类错误	正确

当样本容量 n 一定时, 若减少犯某一类错误的概率, 则犯另一类错误的概率往往增大.

若要使犯两类错误的概率都减小, 除非增加样本容量.

显著性检验

在给定样本容量的情况下，只对犯第一类错误的概率加以控制（使其不超过给定的显著性水平α）, 而不考虑犯第二类错误的概率的检验, 称为**显著性检验**.

若做出接受H_1的结论，则结论较可靠。

若做出接受H_0的结论，结论未必可靠。

双边备择假设与双边假设检验

在 $H_0: \mu = \mu_0$ 和 $H_1: \mu \neq \mu_0$ 中, 备择假设 H_1表示 μ可能大于μ_0, 也可能小于 μ_0, 称为双边备择假设,

形如 $H_0: \mu = \mu_0$, $H_1: \mu \neq \mu_0$ 的假设检验称为双边假设检验.

右边检验与左边检验

形如 $H_0: \mu = \mu_0 (\mu \leq \mu_0)$, $H_1: \mu > \mu_0$ 的假设检验称为右边检验.

形如 $H_0: \mu = \mu_0 (\mu \geq \mu_0)$, $H_1: \mu < \mu_0$ 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.

在假设检验中

$H_0: \mu \le \mu_0, H_1: \mu > \mu_0$（右边检验）与

$H_0: \mu = \mu_0, H_1: \mu > \mu_0$（右边检验）

的检验法则与检验效果是一致的。

同样地，

$H_0: \mu \ge \mu_0, H_1: \mu < \mu_0$（左边检验）与

$H_0: \mu = \mu_0, H_1: \mu < \mu_0$（左边检验）

的检验法则与检验效果也是一致的。

单边检验的拒绝域

设总体 $X \sim N(\mu, \sigma^2)$, σ为已知, $X_1, X_2, \cdots, X_n$ 是来自总体 X 的样本,给定显著性水平 α,

则 右边检验的拒绝域为 $u = \dfrac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} \geq u_\alpha$,

左边检验的拒绝域为 $u = \dfrac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} \leq -u_\alpha$.

三、假设检验的一般步骤

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. 给定显著性水平 α 以及样本容量 n；

3. 确定检验统计量Z（其选取与原假设有关）以及拒绝域形式;

4. 按 $P\{拒绝 H_0 | H_0 为真\} \le \alpha$ 求出拒绝域 W；

5. 取样，根据样本观察值计算检 验统计量Z的观察值z，$z \in W$时，拒绝H_0；否则接受H_0.

注： 原假设与备择假设是不对称的。

决定谁是原假设，依赖于立场、惯例、方便性。

原则 1. 保护原假设。

如果错误地拒绝假设A比错误地拒绝假设B带来更严重的后果---A选作原假设。

例如：假设A：新药有某种毒副作用。
假设B：新药无某种毒副作用。

假设A选作原假设H_0.

原则 2. 原假设为维持现状。

为解释某些现象或效果的存在性，原假设常取为“无效果”、“无改进”、“无差异”，拒绝原假设表示有有利证据支持备择假设。

例8.1.2 某工厂生产的固体燃料推进器的燃烧率服从正态分布 $N(\mu,\sigma^2), \mu = 40\text{cm/s}, \sigma = 2\text{cm/s}$,现在用新方法生产了一批推进器，从中抽取 ***n*=25**只，测得燃烧率的样本均值为 $\bar{x} = 41.25\text{cm/s}$. 设在新方法下总体的标准差仍为$\sigma = 2\text{cm}/s$，问这批新推进器的燃烧率是否较以往生产的推进器的燃烧率有显著提高？取显著性水平$\alpha = 0.05$.

解 $\mu_0 = 40$，依题意检验假设为

H_0： $\mu = \mu_0$（即新方法未提高燃烧率）

H_1： $\mu > \mu_0$（即新方法提高了燃烧率）

这是一个右边检验问题，其检验统计量为

$$u = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} \sim N(0,1)$$

拒绝域为 $u \geq u_\alpha = u_{0.05} = 1.645$

而 $u = \dfrac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \dfrac{41.25 - 40}{2 / \sqrt{25}} = 3.125 > 1.645$

所以, 认为这批新推进器较以往提高了燃烧率.