第八章 假设检验

统计推断

两大基本问题

估计

假设检验

参数检验

分布拟合检验

第八章 假设检验

假设检验是统计推断的另一个重要内容，在数理统计中的理论研究与实际应用中占有重要地位。

所谓假设检验，就是由实际问题的需要，对总体的某个(些)我们关心的方面提出看法(一般称为假设)，再由样本所提供的信息，建立数学模型，根据一定的方法对所提出的假设作出接受还是拒绝的判断，这类统计推断问题在数理统计中称为统计假设检验问题，简称假设检验。

假设检验分类

参数假设检验

总体分布已知，检验关于未知参数的某个假设。

非参数假设检验

总体分布未知，针对未知分布的类型或它的某些特征(如独立性、相关性等)所作假设进行的检验。

本章主要介绍一些参数检验和关于总体分布的拟合检验法。

§1 假设检验

一、基本概念

1. 假设的数学模型

在假设检验问题中，通常根据实际问题，作出合理的假设，以此作为假设检验问题开始。

H_0: 关于总体的某个假设； H_1: 关于总体的另一个假设

其中 H_0 称为原假设/零假设（Null Hypothesis）

H_1 称为备择假设（Alternative Hypothesis）

如 检验总体均值 H_0： $\mu = \mu_0$； H_1： $\mu \neq \mu_0$

2. 两类错误

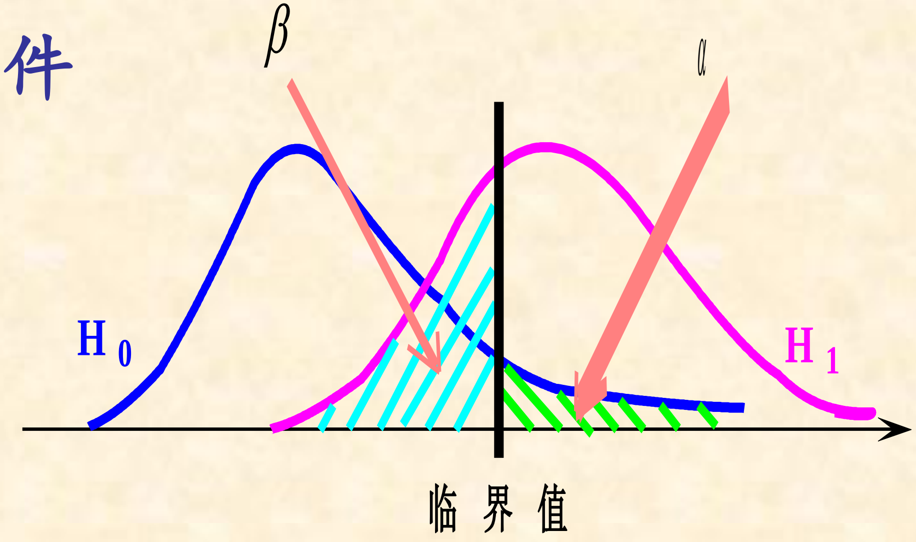
假设检验中，我们是依据样本值，按一定规则来判断原假设H_0的真伪，以决定对它的取舍。由于样本的随机性和样本提供信息的有限性，因此就存在犯两种错误的可能性。

所做判断 / 真实情况	接受H_0	拒绝H_0
H_0为真	正确	犯第Ⅰ类错误 弃真
H_0不真	犯第Ⅱ类错误 取伪	正确

所做判断 真实情况	接受H_0	拒绝H_0
H_0为真	正确	犯第Ⅰ类错误 弃真
H_0不真	犯第Ⅱ类错误 取伪	正确

说明

- 两类错误不是对立事件
- 两类错误概率之和 $\alpha+\beta$不一定等于1



所做判断 / 真实情况	接受H_0	拒绝H_0
H_0为真	正确	犯第 I 类错误 <u>弃真</u>
H_0不真	犯第 II 类错误 <u>取伪</u>	正确

理想的检验方法应该使犯两类错误的概率都尽可能的小，但在样本容量n固定时，<u>不可能同时</u>减小犯两类错误的概率，如果减小一个，就会增大犯另一类错误的概率。因此，要想同时减小犯两类错误的概率，只有增加样本容量n。

3. 显著性检验

只控制犯第Ⅰ类错误的概率，而不考虑第Ⅱ类错误的概率的检验，称为显著性检验（**Test of Statistical Significance**）。

即　控制 $P\{弃真\} \leqslant \alpha$

其中，α 称为显著性水平（**Significance Level**）。

4. 拒绝域与临界点

可以将样本空间划分为两个互不相交的区域：

$$\begin{cases} \text{拒绝域} & \text{拒绝} H_0 \text{的样本点所在的区域。} \\ \text{接受域} & \text{接受} H_0 \text{的样本点所在的区域。} \end{cases}$$

4. 拒绝域与临界点

拒绝域的边界点称为临界点。(也即接受域和拒绝域的分界点)

划分拒绝域和接受域的依据：

原假设H_0成立条件下选定的检验统计量的分布所具有的概率性质以及显著性水平α。

检验统计量　类似于参数估计问题，要借助样本的函数进行统计推断。用于假设检验问题的统计量称为检验统计量。

5. 假设检验原理

实际推断原理

小概率事件在一次试验中几乎是不发生的

反证法思想

首先提出假设，为检验其是否成立，用适当的统计方法来确定假设成立的可能性大小，如果可能性小，则认为假设不成立，拒绝它；反之，则间接肯定。

6. 假设检验的基本思想

设有某个假设H_0需要检验，先假设H_0正确，在H_0成立条件下，选取一个适当的检验统计量Z，由检验统计量Z的分布和给定显著性水平α 确定拒绝域和接受域。若H_0成立时，由样本值计算出检验统计量Z的观察值落在拒绝域内，依据实际推断原理此时原假设H_0是不正确的，应该拒绝H_0；否则，接受H_0。

二、假设检验的基本步骤

1. 根据实际问题提出原假设H_0 和备择假设H_1 ；

2. 确定检验统计量及其在H_0 成立条件下的概率分布；

3. 给定的显著性水平α，在H_0成立的条件下，确定拒绝域和临界点；

4. 由样本值计算检验统计量的观测值，依据该值所落入区域(拒绝域或接受域)作出判断：

接受H_0 或 拒绝H_0 。

例1（P178 例1）某车间用一台包装机包装葡萄糖，袋装糖的净重是一个RV，它服从正态分布。当机器正常时，其均值为0.5kg，标准差为0.015kg。某日开工后为检验包装机是否正常，随机地抽取它所包装的9袋糖，称得净重为（kg）：

0.497　0.506　0.518　0.524　0.498　0.511

0.520　0.515　0.512，

问机器是否正常？（$\alpha=0.05$）

分析

以μ和σ分别表示袋装糖的净重总体X的均值和标准差。长期实践表明：标准差比较稳定，于是σ=0.015，而这里μ未知，即 $X \sim N(\mu, 0.015^2)$。现在的问题是要根据样本来判断μ是否为0.5？

为此，我们进行如下的假设检验：

$$H_0: \ \mu = \mu_0 = 0.5; \quad H_1: \ \mu \neq \mu_0$$

解：检验 $H_0: \mu = \mu_0$； $H_1: \mu \neq \mu_0$

检验统计量 $Z = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \overset{H_0\text{真}}{\sim} N(0,1)$

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

控制 $P\{\text{拒绝}H_0 | H_0\text{为真}\} = P_{\mu_0}\left\{\left|\dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}\right| \geq k\right\} \overset{\text{令}}{=} \alpha$

拒绝域 $\left|\dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}\right| \geq z_{\frac{\alpha}{2}}$

$\bar{x} = 0.511, \sigma = 0.015$

$n = 9, \alpha = 0.05$

检验统计量的观测值 $\left|\dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}\right| = 2.2 > z_{\frac{0.05}{2}} = 1.96$

落入拒绝域，故应拒绝H_0。

说　明

1) 通常显著性水平α取值较小(一般取0.01, 0.05),

因而当H_0为真($\mu=\mu_0$)时,事件$\left\{\left|\dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\right|\geq z_{\frac{\alpha}{2}}\right\}$是一

个小概率事件,根据实际推断原理,若H_0为真,则

由一次试验得到的样本值$\bar{x}$满足不等式$\left|\dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\right|\geq z_{\frac{\alpha}{2}}$

几乎是不会发生的。

说　明

但现在,在一次观察中竟然出现了满足不等式

$\left|\dfrac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}\right| \geq z_{\frac{\alpha}{2}}$的$\overline{x}$,即小概率事件现在发生了,则

我们有理由怀疑假设H_0的正确性,因此拒绝H_0。

如果在一次观察中的样本值满足$\left|\dfrac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}\right| < z_{\frac{\alpha}{2}}$,

则我们没有理由拒绝H_0,故只能接受H_0。

此可以看出:显著性检验问题是实际推断原理和反证法思想的综合应用。

说　明

2) 假设检验中所用的反证法与确定数学中的反证法有区别。

一般反证法：在原假设成立的条件下得出明显矛盾的结果，从而下结论推翻原假设。

在假设检验中，我们在原假设成立的条件下得出的结论是一个具有很大概率的事件(即接受域)，它几乎必然成立，但仍有很小的概率不成立(即拒绝域)。那么为什么认为我们的反证法具有说服力呢？这主要是基于实际推断原理——小概率事件在一次观察中是不会发生的，如果发生了，就只能怀疑(否定)原假设。

依据备择假设的不同情况，将参数的假设检验作如下分类：

参数检验

- 双边备择假设　$H_0: \mu = \mu_0$；$H_1: \mu \neq \mu_0$
- 单边备择假设
  - 右边检验　$H_0: \mu \leqslant \mu_0$；$H_1: \mu > \mu_0$
  - 左边检验　$H_0: \mu \geqslant \mu_0$；$H_1: \mu < \mu_0$

单边检验的拒绝域

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

① 右边检验的拒绝域

$$H_0：\mu \leqslant \mu_0；H_1：\mu > \mu_0$$

检验统计量 $$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$$

拒绝域(形式) $$\frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} \geq k$$?

$H_0: \mu \leqslant \mu_0$; $H_1: \mu > \mu_0$

若$\mu \le \mu_0$,

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \ge \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \ge k$$

控制 $P\{拒绝H_0 | H_0为真\}$

$$= P_{\mu\le\mu_0}\left\{\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \ge k\right\}$$

事件$\left\{\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \ge k\right\} \supset \left\{\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \ge k\right\}$

$$\le P_{\mu\le\mu_0}\left\{\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \ge k\right\} \overset{令}{=} \alpha$$

$N(0, 1)$

可知$k = z_\alpha$　　　故拒绝域$\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \ge z_\alpha$

单边检验的拒绝域

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

② 左边检验的拒绝域

$$H_0: \mu \geq \mu_0;\quad H_1: \mu < \mu_0$$

检验统计量 $$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$$

拒绝域(形式) $$\frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} \leq k \ ?$$

H_0：$\mu \geqslant \mu_0$；H_1：$\mu < \mu_0$

若$\mu \geq \mu_0$,

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \leq \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \leq k$$

控制 $P\{拒绝H_0 | H_0为真\}$

$$= P_{\mu \geq \mu_0}\left\{\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \leq k\right\}$$

事件$\left\{\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \leq k\right\} \supset \left\{\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \leq k\right\}$

$$\leq P_{\mu \geq \mu_0}\left\{\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \leq k\right\} \overset{令}{=} \alpha$$

$N(0, 1)$

可知$k = -z_\alpha$ **故拒绝域** $\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} \leq -z_\alpha$

§2 正态总体均值的假设检验

一、单个正态总体均值的检验

设总体 $X \sim N(\mu, \sigma^2)$, $X_1, X_2, \cdots, X_n$是来自X的样本, 给定显著性水平α, 讨论以下各种检验的拒绝域。

1. σ^2已知，关于μ的检验

① H_0： $\mu = \mu_0$；　　H_1： $\mu \neq \mu_0$

② H_0： $\mu \leqslant \mu_0$；　　H_1： $\mu > \mu_0$

③ H_0： $\mu \geqslant \mu_0$；　　H_1： $\mu < \mu_0$

上述检验问题中，利用检验统计量 $Z = \dfrac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$ 确定拒绝域，这种检验法称为**Z检验法**。

2. σ^2未知，关于μ的检验

σ已知, $Z=\dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$

双边检验 H$_0$：$\mu=\mu_0$；H$_1$：$\mu\neq\mu_0$

检验统计量 $$t=\frac{\bar{X}-\mu_0}{S/\sqrt{n}}\overset{H_0\text{真}}{\sim} t(n-1)$$

拒绝域 $$\left|\frac{\bar{x}-\mu_0}{S/\sqrt{n}}\right|\geq t_{\frac{\alpha}{2}}(n-1)$$

此外，σ^2未知时，μ的单边检验的拒绝域如下，即**P189**表8.1.

假设	检验统计量	拒绝域
$H_0:\mu=\mu_0, H_1:\mu\neq\mu_0$		$\lvert t\rvert \geq t_{\frac{\alpha}{2}}(n-1)$
$H_0:\mu\geq\mu_0, H_1:\mu<\mu_0$	$t=\dfrac{\bar{X}-\mu_0}{S/\sqrt{n}}$	$t\leq -t_{\alpha}(n-1)$
$H_0:\mu\leq\mu_0, H_1:\mu>\mu_0$		$t\geq t_{\alpha}(n-1)$

上述检验利用 t 检验统计量得出拒绝域的检验法称为 t **检验法**。

二、两个正态总体均值差的检验

设$X_1, X_2, \cdots, X_{n_1}$是来自总体$X \sim N\left(\mu_1, \sigma^2\right)$的样本,

$Y_1, Y_2, \cdots, Y_{n_2}$是来自总体$Y \sim N\left(\mu_2, \sigma^2\right)$的样本,且$X$与

Y独立，它们的样本均值为$\bar{X}, \bar{Y}$,样本方差为S_1^2, S_2^2.

又设μ_1, μ_2, σ^2均未知,显著性水平为α，求检验

H_0: $\mu_1 - \mu_2 = \delta$(已知常数)；　H_1: $\mu_1 - \mu_2 \neq \delta$

的拒绝域？

H_0: $\mu_1-\mu_2=\delta$（已知常数）； H_1: $\mu_1-\mu_2\neq\delta$

检验统计量

$$t=\frac{\bar{X}-\bar{Y}-\delta}{\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\overset{H_0\text{真}}{\sim}t(n_1+n_2-2)$$

拒绝域

$$\frac{|\bar{x}-\bar{y}-\delta|}{\sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\geq t_{\frac{\alpha}{2}}(n_1+n_2-2)$$

此外，关于两个正态总体均值差的单边检验的拒绝域如下，即P189表8.1.

假设	检验统计量	拒绝域
$H_0:\mu_1-\mu_2=\delta, H_1:\mu_1-\mu_2\neq\delta$	$t=\dfrac{\bar{X}-\bar{Y}-\delta}{S_W\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}$,	$\lvert t\rvert\geq t_{\frac{\alpha}{2}}(n_1+n_2-2)$
$H_0:\mu_1-\mu_2\geq\delta, H_1:\mu_1-\mu_2<\delta$	$S_W=\sqrt{\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}$	$t\leq -t_{\alpha}(n_1+n_2-2)$
$H_0:\mu_1-\mu_2\leq\delta, H_1:\mu_1-\mu_2>\delta$		$t\geq t_{\alpha}(n_1+n_2-2)$

上述检验法称为 t 检验法。

此外，当两个正态总体的方差均已知时(不一定相等)，可以用**Z检验法**来检验两个正态总体均值差的假设。

检验统计量

$$Z=\frac{\bar{X}-\bar{Y}-\delta}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_1^2}{n_2}}}\overset{\mu_1-\mu_2=\delta}{\sim}N(0,1)$$

具体拒绝域见下表，即**P189**表8.1.

假设	检验统计量	拒绝域
$H_0:\mu_1-\mu_2=\delta, H_1:\mu_1-\mu_2\neq\delta$	$Z=\dfrac{\bar{X}-\bar{Y}-\delta}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}$	$\lvert Z\rvert\geq z_{\frac{\alpha}{2}}$
$H_0:\mu_1-\mu_2\geq\delta, H_1:\mu_1-\mu_2<\delta$		$Z\leq -z_{\alpha}$
$H_0:\mu_1-\mu_2\leq\delta, H_1:\mu_1-\mu_2>\delta$		$Z\geq z_{\alpha}$

三、基于成对数据的检验

有时为了比较两种产品、两种仪器、两种方法等的差异，我们常在相同条件下做对比试验，得到一批成对的观察值，然后分析观察数据作出推断，这种方法常称为逐对比较法。

X	x_1	x_2	…	x_n
Y	y_1	y_2	…	y_n
$D = X - Y$	$x_1 - y_1$	$x_2 - y_2$	…	$x_n - y_n$

由于各种因素影响，不能将X或Y的观察值(上述表格中的第一行或第二行)看成是一个样本的样本值，如：**P186** 例3，为了测试两台光谱仪的测量结果有无显著差异，现利用这两台仪器分别对9个试块进行测量，分别得到9对数据。但由于9个试块之间存在差异(成分含量、金属含量、均匀性等各不同)，故不能将每台仪器测得的9个数据看成是来自同一总体的样本。

X	x_1	x_2	…	x_n
Y	y_1	y_2	…	y_n
$D=X-Y$	x_1-y_1	x_2-y_2	…	x_n-y_n

但是每一对数据之间的差值，即

$$d_i = x_i - y_i \ (i = 1,2,\ldots,n)$$

却是由同一因素引起的。如例3中，都是由于两台仪器之间的差异所引起的。

因此，可以将d_i看成是来自同一总体的样本值，即

$$D_i = X_i - Y_i \ (i = 1,2,\ldots,n)$$

$D_1, D_2, \ldots, D_n$ 是 *i.i.d.* 样本。

X	x_1	x_2	…	x_n
Y	y_1	y_2	…	y_n
$D = X - Y$	$x_1 - y_1$	$x_2 - y_2$	…	$x_n - y_n$

一般，若$D_i \sim N\left(\mu_D, \sigma_D^2\right), (i = 1, 2, \ldots, n)$，则可认为$D_1, D_2, \ldots, D_n$是来自正态总体$N\left(\mu_D, \sigma_D^2\right)$的一个样本，其中，$\mu_D, \sigma_D^2$未知。我们可基于这一样本检验假设

① H_0：$\mu_D = 0$； H_1：$\mu_D \neq 0$

② H_0：$\mu_D \leqslant 0$； H_1：$\mu_D > 0$

② H_0：$\mu_D \geqslant 0$； H_1：$\mu_D < 0$

对于上述检验问题，选用 t 检验，即

检验统计量 $$t=\frac{\bar{D}}{S_D/\sqrt{n}} \overset{\mu_D=0}{\sim} t(n-1)$$

其中：$\bar{D}$是样本均值，S_D是样本标准差。

拒绝域如下（P186表8.1）

双边检验 $$\left|\frac{\bar{d}}{s_D/\sqrt{n}}\right| \geq t_{\frac{\alpha}{2}}(n-1)$$

右边检验 $$\frac{\bar{d}}{s_D/\sqrt{n}} \geq t_{\alpha}(n-1)$$

左边检验 $$\frac{\bar{d}}{s_D/\sqrt{n}} \leq -t_{\alpha}(n-1)$$

小结

- 本节介绍了参数假设检验问题，以正态总体为例，讨论了有关总体均值的一些假设检验问题。

作业

Pages 218, 219：

第2, 3, 6, 7, 8题