Test of Hypotheses 假设检验

1. 简介及定义

2. 单个正态总体 $X \sim N(\mu, \sigma^2)$ 中的假设检验

1. 简介

采样数据可以估计参数的值（点估计、区间估计或置信区间）

However，调查的目的并不是估计参数的值，而是对关于参数的两个相互矛盾的论断进行分析判断：哪一个是正确的

the objective of an investigation is not to estimate a parameter but to decide which of two contradictory claims about the parameter is correct.

实现上述任务所用的方法一般称为假设检验

Methods for accomplishing this comprise the part of statistical inference called hypothesis testing.

是一个关于某参数的论断、断言、主张、声称: 参数涉及总体的特征或者概率分布

- 例如在一个假设检验中，声称 μ =5 .75, 这里 μ 是某 PVC管道的直径的均值。
- 另一个例子就是给定一个结论 $h \geq 178cm$, 这里 h 是我们班男生的平均身高

在任何一个假设检验中，总是要考虑两个相互矛盾的假设

- 假设 $\mu = 5.75$
- 对立的假设就是 $\mu \neq 5.75$

- 声称 $h \geq 178cm$
- 对立的假设就是 $h < 178$

假设检验的目的就是基于样本信息，决定这两个相互矛盾的假设那个是对的，或者是可以接受的。

在检验假设的过程中，有一个假设通常在一开始就是被青睐的。
one of the claims is initially favored.

该初始时就被青睐（假定为合理正确）的假设或论断将不被拒绝，除非样本信息和该假设非常矛盾，并且能够提供非常强的证据支撑其对立假设

原假设（The null hypothesis），　记为 H_0 , 是这样一种论断：初始被认为是正确的 (the “prior belief” claim).

备择假设（The alternative hypothesis）, 记为 H_1, 与 H_0 相互矛盾的论断.

原假设被拒绝而接受备择假设的条件是：样本充分表明 H_0 是错误的。The null hypothesis will be rejected in favor of the alternative hypothesis only if sample evidence suggests that H_0 is false.

如果样本没有和 H_0 明显矛盾，我们继续认为 H_0 是对的，而不否决它。假设检验的两种可能结果是：拒绝 H_0，或者不能拒绝 H_0

定义: 假设检验就是这样一种方法：使用随机采样数据决定是否应该拒绝原假设。

如果我们检验 $H_0 : \mu = 5.75$ 相对备择假设 $H_1 : \mu \neq 5.75$. 仅当随机数据强烈表明 μ 是其它的一个数而不是 5.75 时，原假设 $H_0 : \mu = 5.75$ 应该被拒绝。

如果缺少这样的证据，H_0 就不应该被拒绝，因为它依然是非常可信的。In the absence of such evidence, should not be rejected, since it is still quite plausible.

检验一个论断是否正确，需要一个检验统计量 (A test statistic)：一个关于样本的函数。

被选择的检验统计量必须能非常有效的辨别区分两个假设：检验统计量的值在 H_0 是正确时候，必须和 H_0 是不正确时候差异非常明显

例 8.1 某工厂用包装机包装奶粉，额定标准为每袋净重0.5 kg. 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 $\sigma=0.015(\mathrm{kg})$. 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?

假设 $X\sim N(\mu,0.015^2)$. 如果奶粉重量 X 的均值 μ 等于 0.5 kg，包装机的工作是正常的.

于是提出假设：　$H_0:\mu=\mu_0=0.5$;　$H_1:\mu\neq\mu_0=0.5$.

从抽样的结果来看，样本均值

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

$\overline{X}$ 与 $\mu_0=0.5$ 之间有差异.　注，此处的 $\bar{X}$ 应该 $\bar{x}$，即样本均值的观察值

产生差异的原因是?

(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 $\overline{X}$ 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 $\overline{X}$ 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?

适当选择一个小正数 $\alpha(\alpha = 0.1, 0.05$ 等),叫做显著性水平

若H_0 为真,则 $Z = \dfrac{\overline{X} - \mu_0}{\sqrt{\sigma^2 / n}} = \dfrac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} \sim N(0,1)$,

取定显著性水平 $\alpha = 0.05$. $P\{|Z| > z_{\alpha/2}\} = \alpha$.

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

$$P\{|Z|>z_{\alpha/2}\}=\alpha.$$ 这是一个小概率事件！！！

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”

下面我们验证一下，在本次抽样试验中，上述事件是否发生。

已知 $\Phi(1.96)=0.975,$ 故 $z_{0.025}=1.96,$

$$|z|=\left|\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\right|=\left|\frac{0.5110-0.5}{0.015/\sqrt{9}}\right|=2.2>z_{0.025},$$

说明：一个小概率事件竟然在一次随机抽样中发生了

故假设不合理，拒绝 H_0，接受 H_1

假设检验中的两类错误

第I类错误： H_0 是为真，但是被拒绝。犯该错误的概率为 $P(H_0$ 被拒绝$|H_0$ 为真$) = \alpha$， α 称为显著性水平

第II类错误： H_0 是假的时候，被接受了。犯该错误的概率为 $P(H_0$被接受$|H_0$ 为假$) = \beta$

（1）z-检验法：对服从正态分布的总体均值 μ的检验，总体的方差 σ^2 已知----双边检验 (Two-tailed test)

原假设, $H_0: \mu = \mu_0$

备择假设, $H_1: \mu \neq \mu_0$,

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

拒绝域 $|Z| = \left|\frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}\right| > z_{\alpha/2}$

如果 $|z_0| = \left|\frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}\right| > z_{\alpha/2}$, 拒绝 H_0; 否则，接受 H_0

（2）Z-检验法：对服从正态分布的总体均值 μ的 单侧检验

总体的方差 σ^2 已知 ---- upper or lower tailed test

原假设, H_0: $\mu=\mu_0$ or H_0: $\mu \le \mu_0$

备择假设, H_1: $\mu > \mu_0$,

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

拒绝域 $Z=\dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} > z_\alpha$

如果 $z_0=\dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} > z_\alpha$,

则拒绝 H_0; 否则接受 H_0

原假设, H_0: $\mu=\mu_0$ or H_0: $\mu \ge \mu_0$

备择假设, H_1: $\mu < \mu_0$,

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

拒绝域 $Z=\dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} < -z_\alpha$

如果 $z_0=\dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} < -z_\alpha$,

如果 H_0; 否则，接受 H_0

Attention:这里用的是单侧 α 分位点

（3）t-检验法：对服从正态分布的总体均值 μ的双侧检验

总体的方差 σ^2 未知 --- Two-tailed test

原假设, H_0: $\mu = \mu_0$

备择假设, H_1: $\mu \neq \mu_0$,

检验统计量 $t = \dfrac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim t(n-1)$

拒绝域 $|t| = \left|\dfrac{\bar{X} - \mu_0}{S/\sqrt{n}}\right| > t_{\alpha/2}(n-1)$

如果 $|t_0| = \left|\dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}\right| > t_{\alpha/2}(n-1)$, 则拒绝 H_0; 否则接受 H_0

(4) t-检验法：对服从正态分布的总体均值 μ的单侧检验

总体的方差 σ^2 未知 -

原假设, $H_0: \mu = \mu_0$ 或 $H_0: \mu \leq \mu_0$ 备择假设, $H_1: \mu > \mu_0$, 检验统计量 $t = \dfrac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim t(n-1)$ 拒绝域 $t = \dfrac{\bar{X} - \mu_0}{S/\sqrt{n}} > t_\alpha(n-1)$ 如果 $t = \dfrac{\bar{X} - \mu_0}{S/\sqrt{n}} > t_\alpha(n-1)$, 则拒绝 H_0; 否则接受 H_0	原假设, $H_0: \mu = \mu_0$ or $H_0: \mu \geq \mu_0$ 备择假设, $H_1: \mu < \mu_0$, 检验统计量 $t = \dfrac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim t(n-1)$ 拒绝域 $t = \dfrac{\bar{X} - \mu_0}{S/\sqrt{n}} < -t_\alpha(n-1)$ 如果 $t = \dfrac{\bar{X} - \mu_0}{S/\sqrt{n}} < -t_\alpha(n-1)$, 则拒绝 H_0; 否则接受 H_0

Attention:这里用的是单侧 α 分位点

（5）χ^2检验：对服从正态分布的方差 σ^2 的双侧检验，总体的均值 μ 未知

原假设, $H_0: \sigma^2 = \sigma_0^2$
备择假设, $H_1: \sigma^2 \neq \sigma_0^2$,

检验统计量 $\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2(n-1)$
拒绝域 $\chi^2 > \chi^2_{\alpha/2}(n-1)$ 或 $\chi^2 < \chi^2_{1-\alpha/2}(n-1)$

如果 $\chi_0^2 > \chi^2_{\alpha/2}(n-1)$ or $\chi_0^2 < \chi^2_{1-\alpha/2}(n-1)$, 则拒绝 H_0;

否则 接受 H_0

（6）χ^2-检验：对服从正态分布的方差 σ^2 的单侧检验，总体的均值 μ 未知

原假设，H_0: $\sigma^2 \le \sigma_0^2$
备择假设，H_1: $\sigma^2 > \sigma_0^2$，

原假设，H_0: $\sigma^2 \ge \sigma_0^2$
备择假设，H_1: $\sigma^2 < \sigma_0^2$，

检验统计量 $\chi^2=\frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2(n-1)$

拒绝域 $\chi^2 > \chi_\alpha^2(n-1)$

拒绝域 $\chi^2 < \chi_{1-\alpha}^2(n-1)$

如果 $\chi^2 > \chi_\alpha^2(n-1)$，
则拒绝 H_0; 否则 接受 H_0

如果 $\chi^2 < \chi_{1-\alpha}^2(n-1)$，
则拒绝 H_0; 否则 接受 H_0

Attention:这里用的是单侧 α 分位点

（7）χ^2-检验：对服从正态分布的方差 σ^2 的单侧检验，总体的均值 μ 已知

原假设, $H_0: \sigma^2 = \sigma_0^2$
备择假设, $H_1: \sigma^2 \neq \sigma_0^2$,

检验统计量 $\chi^2 = \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma_0^2} \sim \chi^2(n)$
拒绝域 $\chi^2 > \chi^2_{\alpha/2}(n)$ 或 $\chi^2 < \chi^2_{1-\alpha/2}(n)$

如果 $\chi_0^2 > \chi^2_{\alpha/2}(n)$ or $\chi_0^2 < \chi^2_{1-\alpha/2}(n)$, 则拒绝 H_0; 否则 接受 H_0

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(单位：kg·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 kg·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 $Z=\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$,

若H_0为真，则$Z\sim N(0,1)$.

对给定的显著性水平$\alpha=0.05$， $z_{\alpha/2}=z_{0.025}=1.96$. 拒绝域 $|Z|=\left|\dfrac{\bar{X}-\mu_0}{\frac{\sigma}{\sqrt{n}}}\right|>z_{\alpha/2}$

计算统计量Z的观察值：

$$|z_0|=\left|\frac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}\right|=\left|\frac{31.13-32.50}{1.1/\sqrt{6}}\right|\approx 3.05 > z_{0.025}=1.96$$

因此在显著性水平$\alpha=0.05$下否定H_0，不能认为这批产品的平均抗断强度是32.50 kg·cm^{-2}.

把上面的检验过程加以概括，就可得到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：

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

(c) 对给定的显著性水平 α，查标准正态分布表，

得双侧 α 分位点 $z_{\alpha/2}$.

拒绝域 $|Z|=\left|\frac{\bar{X}-\mu_0}{\frac{\sigma}{\sqrt{n}}}\right|>z_{\alpha/2}$

(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 $|z_0|\leqslant z_{\alpha/2}$，则接受 H_0.

例 8.3 用某仪器间接测量温度，重复 5 次，所得的数据是 1 250 ℃，1 265 ℃，1 245 ℃，1 260 ℃，1 275 ℃，而用别的精确办法测得温度为 1 277 ℃（可看做温度的真值），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 问题是要检验假设： $H_0:\mu=\mu_0=1\,277$； $H_1:\mu\neq\mu_0$.

由于 σ^2 未知（即仪器的精度不知道），我们选取统计量 $t=\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}}$.

拒绝域 $|t|=\left|\dfrac{\bar{X}-\mu_0}{S/\sqrt{n}}\right|>t_{\alpha/2}(n-1)$

查 t 分布表得双侧 α 分位点 $t_{\alpha/2}(n-1)=t_{0.025}(4)=2.776$.

t 的观察值为 $|t_0|=\left|\dfrac{\bar{x}-\mu_0}{s/\sqrt{n}}\right|=\left|\dfrac{1\,259-1\,277}{\sqrt{(570)/(4\times5)}}\right|=\left|\dfrac{-18}{5.339}\right|>3>t_{0.025}(4)$

所以应拒绝 H_0，认为该仪器间接测量有系统偏差.

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叠加到信号上，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，问能否接受这种猜测？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，

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

拒绝域 $Z=\dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}>z_\alpha=z_{0.05}=1.645$

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

所以拒绝 H_0，认为发出的信号值 $\mu>8$.

例 8.5 某厂生产的某种型号的电池，其寿命服从方差 $\sigma^2=5\,000(\mathrm{h}^2)$ 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 $S^2=9\,200(\mathrm{h}^2)$. 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 $\alpha=0.02$)? $n=26$,

解 本题要求在 $\alpha=0.02$ 下检验假设: $H_0:\sigma^2=\sigma_0^2=5000, H_1:\sigma^2\neq\sigma_0^2$

检验统计量 $\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}\sim\chi^2(n-1)$

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

或 $\chi^2<\chi^2_{1-\alpha/2}(n-1)=\chi^2_{0.99}(25)=11.524$,

由观察值 $S^2=9\,200$, 得 $\dfrac{(n-1)S^2}{\sigma_0^2}=46>44.314$,

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.

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：$H_0:\sigma^2\geqslant\sigma_0^2=0.0012$；$H_1:\sigma^2<\sigma_0^2$.

选取统计量 $\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}$.

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

根据样本观察值计算 χ^2 的观察值

$$\chi^2=\frac{(n-1)S^2}{\sigma_0^2}=\frac{24\times0.00066}{0.0012}=13.2<\chi_{1-\alpha}^2(n-1)=13.848$$

即 χ^2 落入拒绝域中，因此拒绝 H_0：$\sigma^2\geqslant\sigma_0^2$，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.

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	$\mu=\mu_0$	$\mu\neq\mu_0$	$\lvert Z\rvert>z_{\alpha/2}$	$Z=\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$		$\pm z_{\alpha/2}$
		$\mu\leqslant\mu_0$	$\mu>\mu_0$	$Z>z_\alpha$			z_α
		$\mu\geqslant\mu_0$	$\mu<\mu_0$	$Z<-z_\alpha$			$-z_\alpha$
	σ^2 未知	$\mu=\mu_0$	$\mu\neq\mu_0$	$\lvert t\rvert>t_{\alpha/2}$	$t=\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}}$	$n-1$	$\pm t_{\alpha/2}$
		$\mu\leqslant\mu_0$	$\mu>\mu_0$	$t>t_\alpha$			t_α
		$\mu\geqslant\mu_0$	$\mu<\mu_0$	$t<-t_\alpha$			$-t_\alpha$
方差	μ 未知	$\sigma^2=\sigma_0^2$	$\sigma^2\neq\sigma_0^2$	$\begin{cases}\chi^2>\chi^2_{\alpha/2}\\ \chi^2<\chi^2_{1-\alpha/2}\end{cases}$	$\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}$	$n-1$	$\begin{cases}\chi^2_{\alpha/2}\\ \chi^2_{1-\alpha/2}\end{cases}$
		$\sigma^2\leqslant\sigma_0^2$	$\sigma^2>\sigma_0^2$	$\chi^2>\chi^2_\alpha$			χ^2_α
		$\sigma^2\geqslant\sigma_0^2$	$\sigma^2<\sigma_0^2$	$\chi^2<\chi^2_{1-\alpha}$			$\chi^2_{1-\alpha}$
	μ 已知	$\sigma^2=\sigma_0^2$	$\sigma^2\neq\sigma_0^2$	$\begin{cases}\chi^2>\chi^2_{\alpha/2}\\ \chi^2<\chi^2_{1-\alpha}\end{cases}$	$\chi^2=\dfrac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\sigma_0^2}$	n	$\begin{cases}\chi^2_{\alpha/2}\\ \chi^2_{1-\alpha}\end{cases}$
		$\sigma^2\leqslant\sigma_0^2$	$\sigma^2>\sigma_0^2$	$\chi^2>\chi^2_\alpha$			χ^2_α
		$\sigma^2\geqslant\sigma_0^2$	$\sigma^2<\sigma_0^2$	$\chi^2<\chi^2_{1-\alpha}$			$\chi^2_{1-\alpha}$

注：上表中 H_0 中的不等号改成等号，所得的拒绝域不变.