Lec-24. 假设检验

主讲教师: 吴利苏 (**wulisu@sdust.edu.cn**)
主　　页: wulisu.cn

本次课内容

假设检验

- 假设检验的相关概念

假设检验

- 假设检验. 首先提出关于总体的假设, 然后根据样本对所提出的假设作出决策(接受 or 拒绝).
- 如何利用样本值对一个具体的假设进行推验?

例

某车间用一台包装机包装葡萄糖，包得的袋装糖重是一个随机变量，服从正态分布．当机器正常时，其均值为 0.5 kg，标准差为 0.015 kg．某日开工后检验包装机是否正常，随机地抽取它所包装地 9 袋糖，称得净重为

$$\begin{array}{ccccc} 0.497 & 0.506 & 0.518 & 0.524 & 0.498 \\ 0.511 & 0.520 & 0.515 & 0.512 & \end{array}$$

问机器是否正常？

分析: 用 μ 和 σ 分别表示这一天袋装糖重总体 X 得均值和方差. 由长期实践知, 标准差较稳定, 设为 $\sigma = 0.015$. 则

$$X \sim N(\mu, 0.015^2), \qquad \text{其中 } \mu \text{ 未知.}$$

分析: 用 μ 和 σ 分别表示这一天袋装糖重总体 X 得均值和方差. 由长期实践知, 标准差较稳定, 设为 $\sigma = 0.015$. 则

$$X \sim N(\mu, 0.015^2), \qquad \text{其中 } \mu \text{ 未知.}$$

- 如何根据样本值判断 $\mu = 0.5$ 还是 $\mu \neq 0.5$?

分析: 用 μ 和 σ 分别表示这一天袋装糖重总体 X 得均值和方差. 由长期实践知, 标准差较稳定, 设为 $\sigma = 0.015$. 则

$$X \sim N(\mu, 0.015^2), \qquad \text{其中 } \mu \text{ 未知.}$$

- 如何根据样本值判断 $\mu = 0.5$ 还是 $\mu \neq 0.5$?
  - 提出两个对立假设 $H_0: \mu = \mu_0 = 0.5$ 和 $H_1: \mu \neq \mu_0$;

分析: 用 μ 和 σ 分别表示这一天袋装糖重总体 X 得均值和方差. 由长期实践知, 标准差较稳定, 设为 $\sigma = 0.015$. 则

$$X \sim N(\mu, 0.015^2), \qquad \text{其中 } \mu \text{ 未知.}$$

- 如何根据样本值判断 $\mu = 0.5$ 还是 $\mu \neq 0.5$?
    - 提出两个对立假设 $H_0 : \mu = \mu_0 = 0.5$ 和 $H_1 : \mu \neq \mu_0$;
    - 根据一个合理的法则, 利用已知样本作出决策: 接受 H_0 or 拒绝 H_0.

分析: 用 μ 和 σ 分别表示这一天袋装糖重总体 X 得均值和方差. 由长期实践知, 标准差较稳定, 设为 $\sigma = 0.015$. 则

$$X \sim N(\mu, 0.015^2), \qquad \text{其中 } \mu \text{ 未知.}$$

- 如何根据样本值判断 $\mu = 0.5$ 还是 $\mu \neq 0.5$?
  - 提出两个对立假设 $H_0 : \mu = \mu_0 = 0.5$ 和 $H_1 : \mu \neq \mu_0$;
  - 根据一个合理的法则, 利用已知样本作出决策: 接受 H_0 or 拒绝 H_0.
  - 如果作出的判断是接受 H_0, 则 $\mu = \mu_0$, 则认为机器是正常的, 否则认为不正常的.

由于要检验的假设涉及总体均值, 故可借助于样本均值.

- $\bar{X}$ 是 μ 的无偏估计量, $\bar{X}$ 的观察值 $\bar{x}$ 一定程度上可以反映 μ 的大小. 因此若 H_0 为真, 则 $|\bar{x}-\mu_0|$ 不应该太大. 考虑 $\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$.
- 当 H_0 为真时,

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

选定一个适当的正数 k.

- 若 $\frac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}} \geq k$, 则拒绝假设 H_0;
- 若 $\frac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}} < k$, 则接受假设 H_0.

这里的 k 的取值应该保证:

H_0 为真, 做出拒绝 H_0 的决策

是一个小概率事件.

给定一个较小的数 $\alpha(0 < \alpha < 1)$, 考虑

$$P\{\text{ 拒绝}H_0 \mid H_0\text{为真}\} \le \alpha,$$

即

$$P\left\{\frac{|\bar{X} - \mu_0|}{\sigma/\sqrt{n}} \ge k \mid H_0\text{为真}\right\} \le \alpha.$$

为确定 k, 取等号

$$P\left\{\frac{|\bar{X} - \mu_0|}{\sigma/\sqrt{n}} \ge k \mid H_0\text{为真}\right\} = \alpha.$$

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

$$k = z_{\alpha/2}.$$

因此，

- 若 $\frac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}} \geq z_{\alpha/2}$，则拒绝假设 H_0；
- 若 $\frac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}} < z_{\alpha/2}$，则接受假设 H_0.

包装机假设检验得过程如下：

解：取 $\alpha = 0.05$，则 $k = z_{\alpha/2} = z_{0.025} = 1.96$.

$n = 9$, $\sigma = 0.015$, $\bar{x} = 0.511$.

所以

$$\frac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}} = 2.2 > 1.96.$$

所以拒绝 H_0，认为包装机工作不正常. $\square$

以上所采取得检验法是符合实际推断原理的.
由于 α 通常取得很小, 一般取 $\alpha = 0.01, 0.05$.
因此当 H_0 为真 (即 $\mu = \mu_0$) 时,
$\left\{\frac{|\bar{X}-\mu_0|}{\sigma/\sqrt{n}} \geq z_{\alpha/2}\right\}$ 是一个小概率事件.

根据实际推断原理, 就可以认为:

如果 H_0 为真, 由一次试验得到不等式

$$\frac{|\bar{X}-\mu_0|}{\sigma/\sqrt{n}} \geq z_{\alpha/2}$$

观察值 $\bar{x}$, 几乎是不会发生的.

在一次试验中,

- 若
$$\frac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}} \geq z_{\alpha/2},$$
则有理由怀疑原来的假设 H_0 的正确性, 因而拒绝 H_0.

- 若
$$\frac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}} < z_{\alpha/2},$$
则没有理由拒绝 H_0, 因而接受 H_0.

显著性水平

在上例中，当样本容量 n 固定，选定 α 后，就可以确定阈值 $k = z_{\alpha/2}$.

- 若 $|z| = \frac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}} \geq k$，则称 $\bar{x}$ 与 μ_0 的差异是显著的，此时拒绝 H_0.
- 若 $|z| = \frac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}} < k$，则称 $\bar{x}$ 与 μ_0 的差异是不显著的，此时接受 H_0.
- α 称为显著性水平. $\bar{x}$ 与 μ 的有无显著差异的判断是在显著性水平 α 之下作出的.
- 检验统计量 $Z = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$.

原假设与备择假设

假设检验问题常叙述为:

在显著性水平 α 下, 检验假设

$$H_0 : \mu = \mu_0, \qquad H_1 : \mu \neq \mu_0.$$

或称为" 在显著水平 α 下, 针对 H_1 检验 H_0".

- H_0 称为原假设或零假设;
- H_1 称为备择假设.
  (意指在原假设被拒绝后可供选择的假设)

拒绝域与临界点

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

如前面的实例中，拒绝域为

$$\frac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}} \geq z_{\alpha/2},$$

临界点为 $z = z_{\alpha/2}$.

两类错误

由于样本的随机性，任一检验规则在应用时，都有可能发生错误的判断——两类错误.

	原假设为真	原假设不真
根据样本拒绝原假设	第 I 类错误	正确
根据样本接受原假设	正确	第 II 类错误

- 第 I 类错误：拒绝真实的原假设 (弃真).
- 第 II 类错误：接受错误的原假设 (取伪).

- $$P_1 = P\{\text{第 I 类错误}\} = P\{\text{拒绝 } H_0 \mid H_0 \text{ 为真}\} = \alpha$$

- $$P_2 = P\{\text{第 II 类错误}\} = P\{\text{接受 } H_0 \mid H_0 \text{ 不真}\}$$

在确定检验法则时, 我们应尽可能使 P_1, P_2 都较小. 当样本容量一定时, P_1, P_2 往往相互制约. 若减少犯第一类错误的概率, 则犯第二类错误的概率往往增大.

所以要使犯两类错误的概率都减小, 只能增加样本容量.

一个记号

$$
\begin{aligned}
P_1 &= P\{\text{第 I 类错误}\} \\
&= P\{H_0 \text{ 为真时拒绝 } H_0\} \\
&= P\{\text{拒绝 } H_0 \mid H_0 \text{ 为真}\} \\
&= P_{\mu_0}\{\text{拒绝 } H_0\} \\
&= P_{\mu \in H_0}\{\text{拒绝 } H_0\}
\end{aligned}
$$

* $P_{\mu_0}\{\bullet\}$ 表示参数 $\mu = \mu_0$ 时，事件 $\{\bullet\}$ 的概率.
* $P_{\mu \in H_0}\{\bullet\}$ 表示参数 μ 取 H_0 规定的值时，事件 $\{\bullet\}$ 的概率.

显著性检验

- 只对犯第一类错误的概率加以控制，而不考虑犯第二类错误的概率的检验，称为**显著性检验**.

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.
- 这样的假设检验称为双边假设检验.

单边检验

但有时，我们只关心总体均值是否增大或减少，

- 形如 $H_0 : \mu \leq \mu_0$, $H_1 : \mu > \mu_0$ 的假设检验称为右边检验.
- 形如 $H_0 : \mu \geq \mu_0$, $H_1 : \mu < \mu_0$ 的假设检验称为左边检验.
- 右边与左边检验统称为单边检验.

单边检验的拒绝域

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

- 右边检验 $H_0 : \mu \leq \mu_0$, $H_1 : \mu > \mu_0$ 的拒绝域

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \geq z_\alpha.$$

- 左边检验 $H_0 : \mu \geq \mu_0$, $H_1 : \mu < \mu_0$ 的拒绝域

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \leq -z_\alpha.$$

证: 右边检验的情况.

H_0 中的全部 μ 都比 H_1 中的 μ 要小, 当 H_1 为真时, 观察值往往偏大. 拒绝域

$$\bar{x} \geq k.$$

下面确定常数 k.

$$\begin{aligned} P\{\text{拒绝} H_0 \mid H_0 \text{为真}\} &= P_{\mu \leq \mu_0}\{\bar{x} \geq k\} \\ &= P_{\mu \leq \mu_0}\left\{\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} \geq \frac{k-\mu_0}{\sigma/\sqrt{n}}\right\} \\ &\leq P_{\mu \leq \mu_0}\left\{\frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \geq \frac{k-\mu_0}{\sigma/\sqrt{n}}\right\}. \end{aligned}$$

要控制 $P\{拒绝H_0 \mid H_0为真\} \le \alpha$, 只需令

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

则 $\frac{k-\mu_0}{\sigma/\sqrt{n}} = z_\alpha \Rightarrow k = \mu_0 + \frac{\sigma}{\sqrt{n}} z_\alpha$. 拒绝域为

$$\bar{x} \ge \mu_0 + \frac{\sigma}{\sqrt{n}} z_\alpha, \qquad 即\, z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} \ge z_\alpha.$$

类似, 左边检验 $H_0 : \mu \ge \mu_0$, $H_1 : \mu < \mu_0$ 的拒绝域为

$$z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} \le -z_\alpha.$$

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 与备择假设 H_1;
2. 给定显著性水平 α 以及样本容量 n;
3. 确定检验统计量以及拒绝域形式;
4. 按 $P\{当H_0为真时拒绝H_0\} = \alpha$ 求出拒绝域;
5. 取样, 根据样本观察值确定接受还是拒绝 H_0.

原假设的提出一般参考以下几个方面

- 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!
- 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”等, 拒绝原假设表示有较强的理由支持备择假设.
- 原假设取简单假设. 把只有一个参数 (或分布) 的假设取为原假设.

例

公司从生产商购买牛奶，公司怀疑生产商在牛奶中掺水以牟利，通过测定牛奶的冰点，可以检验出牛奶是否掺水。天然牛奶的冰点温度近似服从正态分布，均值 $\mu_0 = -0.545$ ℃，标准差 $\sigma = 0.008$ ℃，牛奶掺水可使冰点温度升高而接近水的冰点温度 (0 ℃)，测得生产商提交得 5 批牛奶得冰点温度，其均值为 $\bar{x} = -0.535$ ℃，问是否可以认为生厂商在牛奶中掺水？取 $\alpha = 0.05$.

解: 假设检验

$$H_0 : \mu \leq \mu_0 = -0.545(\text{即牛奶未掺水}),$$

$$H_1 : \mu \geq \mu_0(\text{即牛奶掺水}),$$

这是右边检验问题, 其拒绝域为

$$z = \frac{\bar{x} - \mu_0}{\sigma\sqrt{n}} \geq z_{0.05} = 1.645$$

$z = 2.7951 > 1.645$. z 的值落在拒绝域中, 所以在显著性水平 $\alpha = 0.05$ 下拒绝 H_0, 即认为生产商中掺了水. $\square$

例

设 $(X_1, ..., X_n)$ 是来自正态总体 $N(\mu, 9)$ 的一个样本，其中 μ 为未知参数，检验

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

拒绝域

$$w_1 = \{(x_1, ..., x_n) | |\bar{x} - \mu_0| \geq C\}.$$

(1) 确定常数 C，使得显著性水平为 $\alpha = 0.05$.
(2) 在固定样本容量 $n = 25$ 的情况下，分析犯两类错误的概率 α 和 β 之间的关系.

解: (1) 若 H_0 成立, 则 $\frac{\bar{X}-\mu_0}{3/\sqrt{n}} \sim N(0,1)$.

$$
\begin{aligned}
P\{(x_1, \ldots, x_n) \in w_1\} &= P\{|\bar{X} - \mu_0| \geq C\} \\
&= P\left\{\frac{|\bar{X} - \mu_0|}{3/\sqrt{n}} \geq \frac{C}{3/\sqrt{n}}\right\} \\
&= 0.05
\end{aligned}
$$

$\frac{C}{3/\sqrt{n}} = z_{0.025} = 1.96$, 则 $C = \frac{5.88}{\sqrt{n}}$.

(2) $n = 25$, 若 H_0 成立, 则

$$P\{(x_1, ..., x_n) \in w_1\} = 2(1 - \Phi(\frac{5C}{3})) = \alpha.$$

若 H_0 不成立, 不妨假设 $\mu = \mu_1 \neq \mu_0$.

$$\begin{aligned}
\beta &= P\{(x_1, ..., x_n) \notin w_1\} \\
&= P\{|\bar{x} - \mu_0| < C\} \\
&= P\{-C + \mu_0 < \bar{x} < C + \mu_0\} \\
&= P\left\{\frac{5}{3}(-C + \mu_0 - \mu) < \frac{5}{3}(\bar{X} - \mu) < \frac{5}{3}(C + \mu_0 - \mu)\right\} \\
&= \Phi\left(\frac{5}{3}(C + \mu_0 - \mu_1)\right) - \Phi\left(\frac{5}{3}(-C + \mu_0 - \mu_1)\right).
\end{aligned}$$

当 C 较小时, α 较大, β 较小.
当 C 较大时, α 较小, β 较大. $\square$