

假设检验（上）

一个事例引出假设检验的诸概念

事例：一条新建的南北交通干线全长 10 km。公路穿过一个隧道（长度忽略不计），隧道南面公路长 3.5 km，北面公路长 6.5 km。在刚刚通车的一个月中，隧道南发生了3起互不相关的交通事故，而隧道北没有发生交通事故，能否认为隧道南的路面更容易发生交通事故？

用 p 表示一起交通事故发生在隧道南的概率。如果每起交通事故在这 10 km 的路面上等可能地发生，则 $p = 0.35$ 。于是， $p > 0.35$ 表示隧道南的路面发生交通事故的概率比隧道北的路面发生交通事故的概率大。

为了判断 $p > 0.35$ 是否成立，先作假设

$$H_0 : p = 0.35$$

在统计学中，称 H_0 是**原假设**或**零假设**。再作备择假设 $H_1 : p > 0.35$ 。

在本问题中，如果拒绝 H_0 ，就应当接受 H_1 。

用 W 表示三起交通事故都发生在隧道南，如果 H_0 为真，则每一起事故发生在隧道南的概率都是 0.35。因为这三起交通事故的发生是独立的，所以

$$P(W) = 0.35^3 \approx 4.29\%$$

W 是一个小概率事件，一般不会发生。因为小概率事件 W 发生了，所以要拒绝 H_0 ，接受备择假设，即认为隧道南的路面比隧道北发生交通事故的概率大。

做出以上结论也有可能犯错误，犯错误的概率正是 $P(W) \approx 4.29\%$ 。这是因为当 H_0 成立，且 W 发生时，我们才犯错误。这类错误被称为**第一类错误**。因为小概率事件 W 的发生，我们才作出了拒绝 H_0 的判断，所以称 W 是**拒绝域**或**否定域**。

在正确计算的前提下，犯错误的原因是随机因素造成的。要有效减少犯错误的概率，可以通过**增加观测数据**解决。

在解决假设检验的问题时，无论作出拒绝还是接受原假设 H_0 的决定，都有可能犯错误。在统计学中，称拒绝 H_0 时犯的误差为第一类错误，称接受 H_0 时犯的误差为第二类错误。具体如下：

- (1) H_0 为真，统计推断的结果拒绝 H_0 ，犯第一类错误；
- (2) H_0 为假，统计推断的结果接受 H_0 ，犯第二类错误；
- (3) H_0 为真，统计推断的结果没有拒绝 H_0 ，不犯错误；
- (4) H_0 为假，统计推断的结果拒绝 H_0 ，不犯错误。

显著性检验

在应用问题中，常用到显著性检验。

显著性检验的任务是**依据观测数据或试验数据判断原假设 H_0 是否成立**。

如果观测数据或试验数据和原假设有显著的差异，就拒绝 H_0 ，并称检验显著。否则不能拒绝 H_0 ，并称检验不显著。注意，不能拒绝 H_0 并不表示一定要接受 H_0 。

显著水平 α 控制了拒绝 H_0 时犯第一类错误的概率：如果检验是显著的，拒绝 H_0 后，犯第一类错误的概率不超过 α 。如果检验不显著，就不能拒绝 H_0 ，但是马上作出接受 H_0 的决定还是会冒险的，因为这时犯第二类错误的概率可能较大。

显著性检验的原则是**控制犯第一类错误的概率不超过某个定值 α** ，比如说 $\alpha = 0.05$ 。

显著性检验对犯第二类错误的概率没有限制，因此，如何规定原假设和备择假设就有讲究了。

事例：亚硝酸盐广泛用于肉食品制作。鉴于亚硝酸盐对人体十分有害（引起中毒、致癌），所以各国对其作为食品添加剂的用量都有十分严格的限制。现在为了检验某食品厂灌肠的亚硝酸盐含量是否超标，预备随机抽查该厂10个灌肠样品，用显著性检验方法进行检验。如果控制犯第一类错误的概率不超过 0.01，应当对下面的哪个假设进行检验？

- (a) H_0 ：亚硝酸盐不超标 vs H_1 ：亚硝酸盐超标；
- (b) H_0 ：亚硝酸盐超标 vs H_1 ：亚硝酸盐不超标。

如果对于假设 (a) 进行检验，则依据数据拒绝 H_0 时，犯错误的概率不超过 0.01，即把合格灌肠判为不合格的概率不超过 0.01。因为 0.01 太小了，所以假设 (a) 保护了厂家的权益。当依据数据不能拒绝 H_0 时，若接受 H_0 ，则将冒很大的风险。因为这时犯（第二类）错误的概率是未知的，而不合格的灌肠流向市场会损害消费者的健康，造成严重后果。所以假设 (a) 不适用。

再分析假设 (b)。当依据数据拒绝 H_0 时，犯错误的概率不超过 0.01，这就保护了消费者的利益。因为超标灌肠流向市场的概率不超过 1%。当依据数据不能拒绝 H_0 ，即便认为灌肠的亚硝酸盐超标而犯错误，也不会造成严重后果，因为厂家进一步减少添加剂的用量就是了。所以，本例应当对假设 (b) 进行检验。

显著性检验对原假设 H_0 有利。

对于参数的假设检验问题，当数据在手，又不需要考虑其他因素的情况下，**如果选备择假设 H_1 和已有的事实或数据相符，选原假设 H_0 和已有的事实或数据相悖，则有利于得出拒绝 H_0 的结论**，否则不能从假设检验中得到更多信息。

正态分布参数的显著性检验

- 1、先做假设。
- 2、根据假设和样本，确定一个枢轴量，得到其分布。
- 3、根据假设和样本，计算枢轴量的观测值是否在拒绝域中。
- 4、选择接受或者拒绝原假设。

双边检验

如果 X_1, X_2, \cdots, X_n 是总体 $N(\mu, \sigma^2)$ 的样本观测值, σ **已知**时, $H_0: \mu = \mu_0$ vs $H_1: \mu \neq \mu_0$ 的显著水平为 α 的拒绝域是

$$\{|Z| \geq z_{\alpha/2}\}, \text{ 其中 } Z = \frac{\bar{X}_n - \mu_0}{\sigma/\sqrt{n}}$$

如果 $|Z| \geq z_{\alpha/2}$ 发生, 则称检验是显著的, 表示**结论和假设有显著差异**。这时, 拒绝 H_0 犯错误的概率不超过 α 。由于这种检验方法是基于正态分布的方法, 所以又称为正态检验法或 Z 检验法。

如果 X_1, X_2, \cdots, X_n 是总体 $N(\mu, \sigma^2)$ 的样本, σ **未知**时, 假设 $H_0: \mu = \mu_0$ vs $H_1: \mu \neq \mu_0$ 的显著水平为 α 的拒绝域是

$$\{|T| \geq t_{\alpha/2}(n-1)\}, \text{ 其中 } T = \frac{\bar{X}_n - \mu_0}{S/\sqrt{n}}$$

如果 $|T| \geq t_{\alpha/2}(n-1)$ 发生, 则称检验是显著的, 这时拒绝 H_0 犯错误的概率不超过 α 。由于这种检验方法是基于 t 分布的方法, 所以又称为 t 检验法。

上面两例都是检验 $H_0: \mu = \mu_0$ vs $H_1: \mu \neq \mu_0$ 。因为当 \bar{X}_n 比 μ_0 大许多或小许多时, 都倾向于拒绝原假设 H_0 , 所以这种检验又被称为**双边检验**。

单边检验

在许多问题中, 还常常需要检验总体均值 μ 是否大于某 个定值 μ_0 , 或检验总体方差 σ^2 是否大于某个定值 σ_0^2 。这时需要作**单边假设检验**。另外, 人们还可以利用数据提供的信息设计原假设和**对立假设**, 这时也需要作单边假设检验。

设 X_1, X_2, \cdots, X_n 是总体 $N(\mu, \sigma^2)$ 的样本, 如果标准差 σ 未知, 则假设 $H_0: \mu \geq \mu_0$ vs $H_1: \mu < \mu_0$ 的显著水平为 α 的拒绝域是

$$W_\alpha = \{T \leq -t_\alpha(n-1)\}, \text{ 其中 } T = \frac{\bar{X}_n - \mu_0}{S/\sqrt{n}}.$$

如果 W_α 发生, 则检验显著。检验显著时, 拒绝 H_0 犯错误的概率不超过 α 。

同理, 对于总体 $N(\mu, \sigma^2)$ 的样本 $X_1, X_2, \cdots, X_n, \sigma$ 未知时, 在显著水平 α 下, 假设 $H_0: \mu \leq \mu_0$ vs $H_1: \mu > \mu_0$ 的拒绝域是

$$W_\alpha = \{T \geq t_\alpha(n-1)\}, \quad \text{其中 } T = \frac{\bar{X}_n - \mu_0}{S/\sqrt{n}}.$$

如果 W_α 发生, 则检验显著。检验显著时, 拒绝 H_0 犯错误的概率不超过 α 。

实际上, 对于相同的显著水平和适当的单边原假设, 因为临界值 $t_\alpha(m) < t_{\alpha/2}(m)$, 所以双边检验显著时, 单边检验也一定显著。反之, 单边检验显著时, 双边检验不一定显著。造成这一结果的原因还在于, **单边检验的原假设和对立假设的设计往往已经利用了数据提供的信息, 而双边检验的设计往往没有利用数据的信息**。这是因为双边检验通常是得到试验数据之前所设计的检验, 所以该设计无法利用试验数据提供的信息。当试验数据在手时, 应当根据数据提供的信息作单边检验。

要深入, 还是要多做点题, 在实际问题中理解。

成对数据的假设检验（PPT例题）

大样本检验（PPT例题）