

问题的提出:

盐厂用一台包装机包装食盐以利外运。规定每袋的标准重量为500g, 并假定包得的袋装盐重服从正态分布。据以往的经验知, 包装机工作时袋装盐重的标准差为10g。某天开工后随机抽测了16袋, 称得其平均重量为510g。试问: 该天包装机工作是否正常?

分析: 设该天包装机包装的袋装盐重为 X , 由题意知 $X \sim N(u, \sigma^2)$, 其中 $\sigma = 10\text{g}$, u 未知。我们关心的问题是: 包装机工作是否正常, 也即判断该天包装机包装的袋装盐重 X 的均值是否等于

标准重500g，也即是检验“假设： $\mu = 500\text{g}$ ”是否正确。这个问题是有关正态总体方差已知时，对总体均值的假设检验。

这是关于总体参数的假设检验问题。这类问题称为**参数检验**。

假设检验思想

先看一下我们刚才介绍的例子。由题意设 $X \sim N(\mu, \sigma^2)$ ，其中已知 $\sigma = \sigma_0 = 10\text{g}$ ，而 μ 未知。现在，问题的结论只能有两个，要么 $\mu = \mu_0 = 500\text{g}$ ，要么 $\mu \neq \mu_0 = 500\text{g}$ 。我们要判断哪

个命题成立？为此给出两个假设

$$H_0: u = u_0 = 500\text{g}, H_1: u \neq u_0 = 500\text{g}。$$

这是两个对立的假设，常称 H_0 为**原假设**， H_1 为**备择假设**。现在要利用样本观测值来检验 H_0 ，从而决定拒绝 H_0 或接受 H_0 。

由于这是对总体均值的假设检验，我们自然想到能否借助样本均值 \bar{X} ，因为样本均值是总体均值的无偏估计。因此，若原假设 H_0 正确，那么 \bar{X} 与 u_0 的偏差 $|\bar{X} - u_0|$ 不应太大。若 $|\bar{X} - u_0|$ 太大，我们有理由怀疑 H_0 的正确性而拒绝 H_0 ；反之，若 $|\bar{X} - u_0|$ 不是很大，即是说 \bar{X} 与 u_0 的差异是由随机因素引起的，故不能拒绝 H_0 。

当 H_0 成立时有

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

从而可将 $|\bar{X} - \mu_0|$ 大小的衡量归结为对 $|U|$ 大小的衡量。我们称 U 为**检验统计量**。其方法是给定一小正数 α 使得

$$P(|U| > u_{1-\alpha/2}) = \alpha.$$

再计算统计量 $|U|$ 的观测值 $|u|$, 若有 $|u| > u_{1-\alpha/2}$ 则拒绝 H_0 ; 否则, 就不能拒绝 H_0 。

为什么要作此判断呢? 这是根据**小概率事件的实际推断原理**而作出的。

因为 α 很小，通常取 $\alpha=0.05, 0.1$ 等，因而事件 $\{|U| > u_{1-\alpha/2}\}$ 是小概率事件。因此，若 H_0 正确，则在一次试验中，该事件实际上是不会发生的。例如在本例中，若取 $\alpha=0.05$ ，则 $u_{1-\alpha/2} = u_{0.975} = 1.96$ 。而今抽样结果为 $n = 16, \bar{x} = 510$ ，且知 $\sigma=10$ ，则统计量 $|U|$ 的观测值为

$$|u| = \left| \frac{510 - 500}{10/\sqrt{16}} \right| = 4 > 1.96,$$

也即有小概率事件 $\{|U| > u_{1-\alpha/2}\}$ 在一次试验中发生了。因此，我们认为原先的假设即原假设 H_0 是不成立的，因而拒绝 H_0 。从而认为包装机工作不正常。

这里包含了反证法思想，但与一般的反证法是不一样的。因为这里仅仅根据小概率事件的实际推断原理来论证的，因此，这里的反证法是带有概率性质的。给定的 α 称为**显著水平**。拒绝原假设成立的区域称为**拒绝区域**，记为 W ；拒绝域的边界值称为**临界值**。如在上例中，可求得拒绝域为 $W = \{|U| > u_{1-\alpha/2}\}$ ，临界值为 $u_{1-\alpha/2}$ 和 $u_{1-\alpha/2}$ 。

综上所述，假设检验的基本思想为：**对总体分布的未知参数作出某种假设，根据样本在原假设为真的前提下构造一个小概率事件，基于“小概率原理”（即：小概率事件在一次试验中几乎不可能发生）而对原假设作出拒绝或不拒绝。**

双侧检验与单侧检验

双侧检验: $H_0: u = u_0$, $H_1: u \neq u_0$

右侧检验: $H_0: u \leq u_0$, $H_1: u > u_0$ 或者

$H_0: u = u_0$, $H_1: u > u_0$;

左侧检验: $H_0: u \geq u_0$, $H_1: u < u_0$ 或者

$H_0: u = u_0$, $H_1: u < u_0$;

左侧检验和右侧检验统称为**单侧检验**。在同一显著水平 α 下, 两种形式下的右侧 (或左侧) 检验方法是一样的。故我们一般只考虑第一种。

上述检验名称来源于拒绝域的形式。

双侧检验拒绝域 $(-\infty, a) \cup (b, +\infty)$.

右侧检验拒绝域
 $(b, +\infty)$.

左侧检验拒绝域
 $(-\infty, a)$

两类错误

假设检验是依据小概率事件的推断原理来作出判断的，加之抽样的随机性，这就可能导致错误的判断。具体来说，有两类错误：

(1) 当原假设 H_0 实际上为真时，假设检验结果有可能拒绝 H_0 。这种“以真为假”，“弃真”的错误称为第一类错误。犯这种错误的概率为

$$P(\text{拒绝}H_0|H_0\text{为真})$$

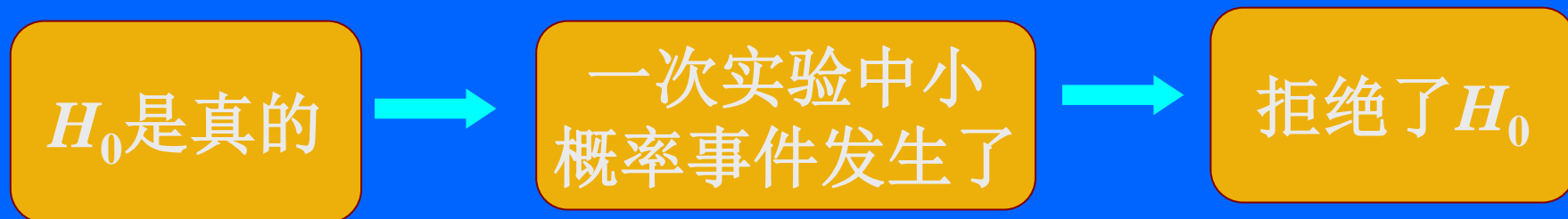
小于等于显著水平 α 。

(2) 当原假设 H_0 实际上不真时，假设检验结果将接受 H_0 。这种“以假为真”，“纳伪”的错误称为第二类错误。犯第二类错误的概率为 β ，其

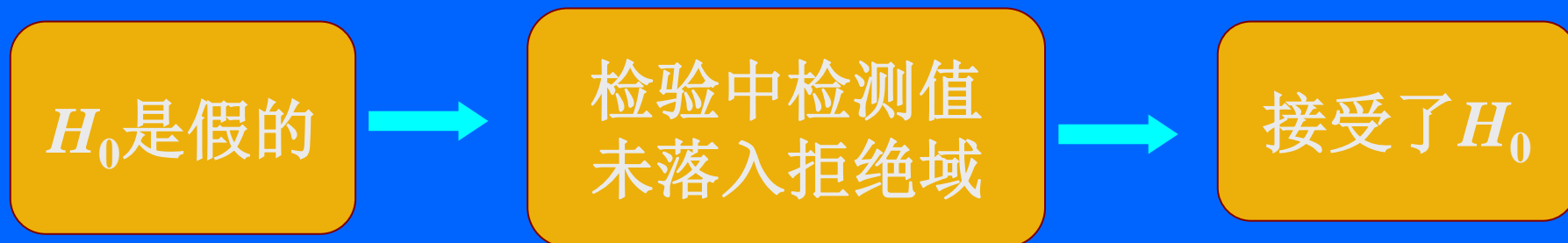
中

$$\beta = P(\text{接受}H_0|H_0\text{不真})$$

由于我们的假设检验思想，这两种错误是不可能完全避免的，因此，我们只希望犯这两类错误的概率越小越好。然而，当样本容量 n 确定之后，犯这两类错误的概率不能同时减小，减小其中一个，另一个往往会增大。目前，人们的通常做法是：先控制第一类错误的概率 α ，当 α 被选定后，再通过增大样本容量 n 使 β 减小。



由检验过程知： $P(\text{拒绝}H_0 | H_0\text{真}) \leq \alpha$



$$\beta = P(\text{接受}H_0 | H_0\text{假}) \neq 1 - \alpha$$

假设检验的基本步骤

根据前面的分析，假设检验的步骤如下：

- (1) 根据实际问题，提出原假设 H_0 与备择假设 H_1 ；
- (2) 选择适当的检验统计量，使其分布或极限分布在 H_0 成立的条件是已知的，确定该统计量的分布；
- (3) 选定显著性水平 α ，并根据统计量的分布查表确定临界值；
- (4) 根据样本观测值，计算出统计量的值，并将其与临界值作比较；
- (5) 下结论：若检验统计量的值落在拒绝区域内，则拒绝 H_0 ；否则，不拒绝 H_0 。

注：临界值是与 α （给定的显著性水平）相关的。因此，对同一假设检验，所给的 α 不同，可能使检验的结果大相径庭。所以 α 的选取是很重要的。