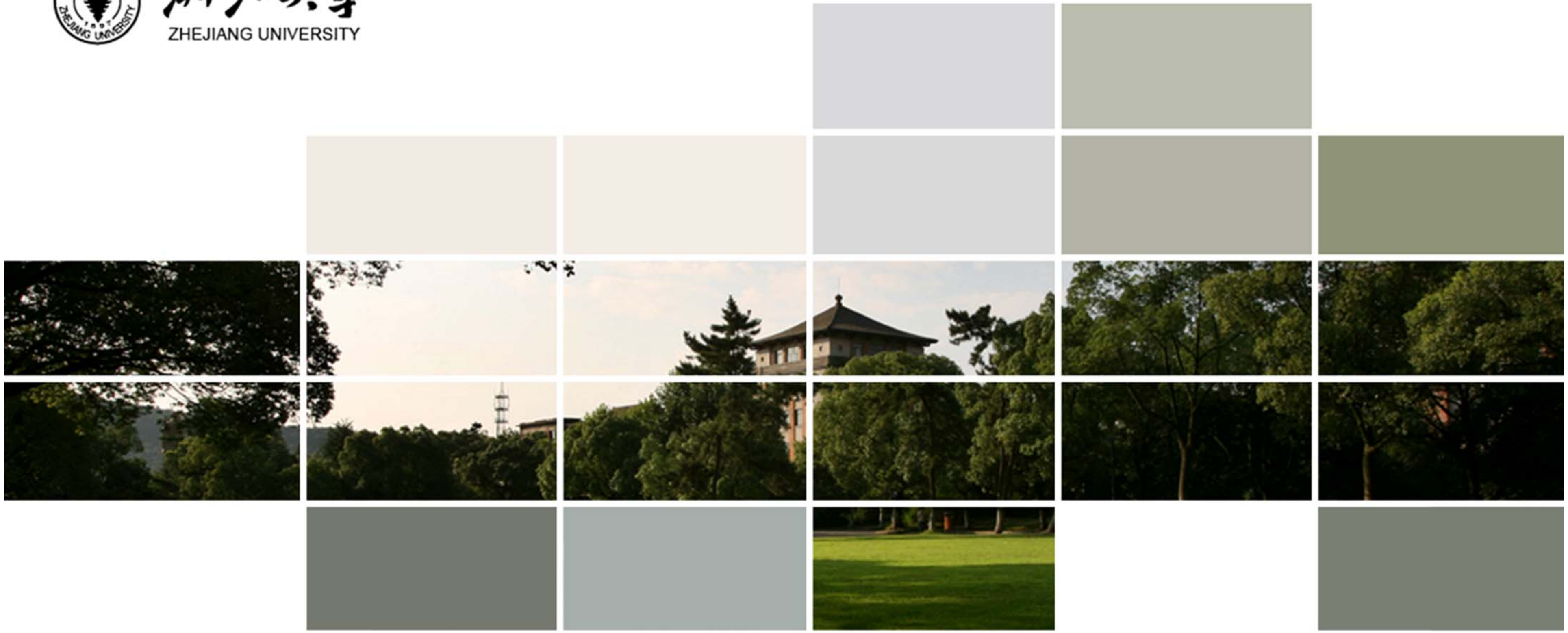
浙江大学
ZHEJIANG UNIVERSITY



第54讲　假设检验的基本思想

例1. 体重指数BMI是目前国际上常用的衡量人体胖瘦程度以及是否健康的一个标准. 专家指出, 健康成年人的BMI 取值应在 18.55- 24.99 之间.

某种减肥药广告宣称, 连续使用该种减肥药一个星期便可达到减肥的效果.

为了检验其说法是否可靠,随机抽取9位试验者(要求BMI 指数超过25、年龄在20-25岁女生),先让每位女生记录没有服用减肥药前的体重, 然后让每位女生服用该减肥药, 服药期间, 要求每位女生保持正常的饮食习惯, 连续服用该减肥药1周后, 再次记录各自的体重.

测得服减肥药前后的体重差值(服药前体重-服药后体重) (单位: kg):

1.5， 0.6， -0.3， 1.1， -0.8， 0， 2.2， -1.0， 1.4

问题：根据目前的样本资料能否认为该减肥药广告中的宣称是可靠的?

假设检验的目的是通过收集到的数据，来验证某个想要得到的结论。过程类似于法官的审判过程。

法官的立场基于“疑罪从无”：
法官宣告被告“有罪”是需要充分的证据来推翻被告是“无罪”的假设；
而宣判“无罪”，是由于没有充分的证据支持被告“有罪”，并不是有充分的证据支持被告“无罪”.

检验假设的过程是一个四步曲.

第一步，建立两个完全对立的假设：

原假设(零假设)H_0，备择假设 (对立假设) H_1。

原假设与备择假设是不对称的！

决定谁是原假设，依赖于立场、惯例、方便性.

1. 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!

例如：假设A:新药有某种毒副作用，假设B:新药无某种毒副作用. ——A选作原假设H_0!

“有毒副作用”错误地当成“无毒副作用”比“无毒副作用”错误地当成“有毒副作用”带来的后果更严重。

2. 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.

例1中原假设H_0:药物没有减肥效果.

备择假设 H_1: 药物有减肥效果.

3. 原假设取简单假设.只有一个参数(或分布)的假设称为简单假设.如果只有一个假设是简单假设，将其取为原假设.

参数假设的形式

设θ是反映总体指标某方面特征的量，是我们感兴趣的参数．一般参数θ的假设有三种情形：

H_0: $\theta=\theta_0$, H_1: $\theta<\theta_0$(左边检验)

H_0: $\theta=\theta_0$, H_1: $\theta>\theta_0$(右边检验)

H_0: $\theta=\theta_0$, H_1: $\theta\neq\theta_0$(双边检验)

在假设检验中

H_0: $\theta \geq \theta_0$, H_1: $\theta < \theta_0$（左边检验）与

H_0: $\theta = \theta_0$, H_1: $\theta < \theta_0$（左边检验）

的检验法则与检验效果是一致的.

同样的

H_0: $\theta \leq \theta_0$, H_1: $\theta > \theta_0$（右边检验）与

H_0: $\theta = \theta_0$, H_1: $\theta > \theta_0$（右边检验）

的检验法则与检验效果也是一致的.

如何检验假设？

根据收集的资料，针对假设，给出检验方法，然后对假设进行判断。

判断方法有二种：临界值法．P_值法.

以例1为例来说明减肥药有效？

还是无效？

设服用减肥药前后体重差值$X \sim N(\mu,\ \sigma^2)$，

并假定方差$\sigma^2 = 0.36.$

检验假设：$H_0 : \mu = 0, H_1 : \mu > 0,$

注意到：$\bar{X}$是μ的无偏估计，$\bar{X}$的取值大小反映了μ的取值大小，当原假设成立时，$\bar{X}$取值应偏小。因此

当$\bar{X} \geq C$时，拒绝原假设H_0,

当$\bar{X} < C$时，接受原假设H_0,

其中C是待定的常数.

如果统计量$T=T(X_1,\cdots,X_n)$的取值大小和原假设H_0是否成立有密切联系，可将其称为对应假设问题的**检验统计量**，而对应于拒绝原假设H_0时，样本值的范围称为**拒绝域**，记为W,其补集$\overline{W}$称为**接受域**.

第二步：给出检验统计量，并确定拒绝域的形式.

本例中的检验统计量为$\bar{X}$,拒绝域为

$$W=\left\{(X_1,\cdots,X_n):\bar{X}\geq C\right\}$$

C如何选择？——关键问题.

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).

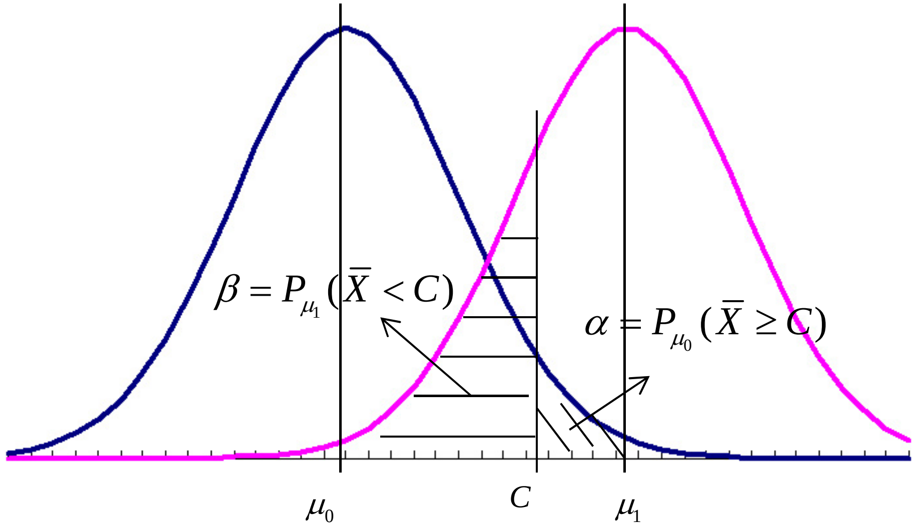
第II类错误：接受错误的原假设(取伪).

$$\alpha = P\{\text{第I类错误}\} = P\{\text{拒绝}H_0 \mid H_0\text{是真实的}\},$$

$$\beta = P\{\text{第II类错误}\} = P\{\text{接受}H_0 \mid H_0\text{是错误的}\}.$$

例如：设总体$X \sim N(\mu,1)$，则$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \sim N(\mu, \frac{1}{n})$，

$H_0: \mu = \mu_0, H_1: \mu = \mu_1 (> \mu_0)$，　拒绝域：$\bar{X} \geq C$.



犯两类错误的概率相互制约

Neyman-Pearson原则：

首先控制犯第I类错误的概率不超过某个常数$\alpha\in(0,1)$，再寻找检验，使得犯第II类错误的概率尽可能小. α称为显著水平.

常取$\alpha = 0.01, \quad 0.05, \quad 0.1$等.

第三步，根据显著水平和统计量的分布确定临界值——临界值法

在例1中，取显著水平$\alpha = 0.05$,

当$H_0:\mu=0$成立时,$\Rightarrow \dfrac{\bar{X}}{0.6/\sqrt{9}} \sim N(0,1)$,（统计量的分布）

犯第I类错误的概率可如下计算：

$$P\{\bar{X} \geq C \mid \mu=0\} = P\left\{\frac{\bar{X}}{\sigma/\sqrt{n}} \geq \frac{C}{\sigma/\sqrt{n}} \,\middle|\, \mu=0\right\}$$

$$=1-\Phi\left(\frac{C}{\sigma/\sqrt{n}}\right) \leq \alpha = 0.05. \quad \left(0.05=\Phi\left(-z_{0.05}\right)\right)$$

$$\Rightarrow \frac{C}{0.6/\sqrt{9}} \geq z_{0.05} = 1.645. \Rightarrow C \geq 0.329.$$

根据Neyman $-$ *Pearson*原则，为使犯第**II**类错误的概率尽可能小，应取$C = 0.329$.因此, 拒绝域$W = \{\bar{X} \geq 0.329\}$.

第四步：根据样本得出结论.

根据实际样本资料，得$\bar{x} = 0.522 > 0.329$.

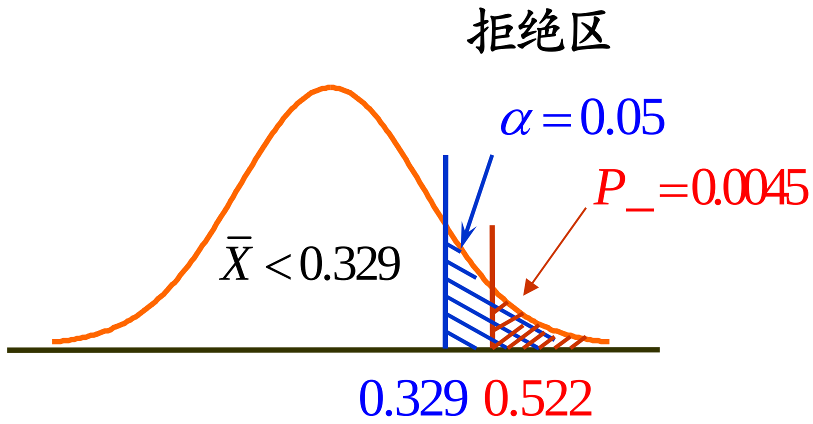
当原假设H_0成立时，样本落在拒绝域的概率不超过0.05，是小概率事件。

根据实际推断原理，有充分的理由拒绝原假设，认为厂家的宣传是可靠的.

同理，若$\alpha = 0.01$, 拒绝域$W = \{\bar{X} \geq 0.465\}$，拒绝原假设.

第三’步：计算最小显著水平——P_值法

P_-值：当原假设H_0成立时，检验统计量取比观察到的结果更为极端的数值的概率.



$$P_-=P\{\bar{X}\geq\bar{x}=0.522\mid\mu=0\}$$

$$=1-\Phi\left(\frac{0.522}{0.6/\sqrt{9}}\right)=0.0045$$

$$<\alpha=0.05$$

概率这么小的事件！
竟然发生了！！
拒绝原假设！！！

第四’步：比较P_值与显著水平，得出结论.

$P_$值与显著水平α的关系：

(1) 若$P_ \leq \alpha$，等价于样本落在拒绝域内，因此，拒绝原假设，称检验结果在水平α下是统计显著的.

(2) 若$P_ > \alpha$，等价于样本不落在拒绝域内，因此，不拒绝（接受）原假设，称检验结果在水平α下是统计不显著.

临界值法处理假设检验问题的基本步骤

（1）根据实际问题提出原假设和备择假设；

（2）提出检验统计量和拒绝域的形式；

（3）在给定的显著水平α下，根据Neyman-Pearson原则求出拒绝域的临界值;

（4）根据实际样本观测值作出判断.

P_值法处理假设检验问题的基本步骤

（1）根据实际问题提出原假设和备择假设；

（2）提出检验统计量和拒绝域的形式；

（$3'$）计算检验统计量的观测值与$P_$值；

（$4'$）根据给定的显著水平α，作出判断.