假设检验

定义

- 不求取参数的具体数值，而是给未知参数赋予一个特定意义的范围 (Θ^*) ，根据观测数据推断其是否落在该指定范围内。

- 假设检验包括假设和检验两个阶段：
  - 假设阶段：对于总体的某个待推断的未知方面，根据解决问题的需要，作出一个假设，假设可能正确也可能错误。
  - 检验阶段：为判断做出的假设是否正确，从总体中抽取样本，并根据样本提供的信息，按一定原则（落在指定区间的概率大小）对所作假设加以检验，进而以具体检验情况或者实际检验结果做出决定：接受或者拒绝所作假设。

基本概念

- 原假设、备择假设：
  设 θ 为总体的一个未知参数，其一切可能取值的集合即为 Θ 。则关于 θ 的任一假设可用 $\theta \in \Theta$ 来表示，其中 Θ^* 为 Θ 的一个真子集。
  在假设检验中，将在假设阶段所作的假设称为原假设，为了使得问题表述更明确，通常还提出一个与之相对的假设，称为备择假设（一般就是原假设的反面）。

- 双边假设与单边假设
  1. 双边假设：$H_0:\ \theta = \theta_0,\quad H_1:\ \theta \neq \theta_0$
  2. 左边假设：$H_0:\ \theta = \theta_0(\theta \geq \theta_0),\quad H_1:\ \theta < \theta_0$
  3. 右边假设：$H_0:\ \theta = \theta_0(\theta \leq \theta_0),\quad H_1:\ \theta > \theta_0$

  NOTE: 左（右）边假设可以理解为取接受域的左（右）边界为判定条件。左右边假设统称为单边假设。

- 检验统计量：
  在检验过程中，通常通过构造统计量，并对其观测值进行检验。用于检验的统计量 V 为检验统计量。

- 接受域、拒绝域、临界点
  统计假设检验的依据是样本，样本的取值可能对原假设 H_0 不利或有利。可以根据某合理准则（一般由检验量的选择确定）将样本空间分成两个部分，分别表现为对原假设有利或者不利。
  在检验过程中，当样本落入拒绝域时，拒绝原假设；当样本落入接受域时，接受原假设。拒绝域的边界称为临界点。

- 显著性水平：设 $(X_1, X_2, \cdots, X_n)$ 为样本，W 为样本空间的一个子集，对于给定的小概率 $\alpha \in (0, 1)$ ，对任一的待推断 $\theta \in \Theta_0$ ，若 W 满足

  $$P_\theta((X_1, X_2, cdots, X_n) \in W) \leq \alpha$$

  则 W 构成了原假设的一个拒绝域。称 α 为显著性水平，并称此由 W 构成拒绝与的检验方式为显著性水平为 α 的检验。
  NOTE: 显著性水平 α 的一个直观理解就是当所取样本事件发生的概率低于 α 时，拒绝做出的假设。

假设检验的步骤

1. 根据实际问题，建立原假设 H_0 和备择假设 H_1 。
2. 在 H_0 为真的前提下，选择合适的检验统计量 V 。
3. 由 H_1 确定出拒绝域的形式。
   对于给定的显著性水平 α ，其对应的拒绝域为：
   - 双边假设：$(V < V_{\frac{\alpha}{2}}) \cup (V > V_{1-\frac{\alpha}{2}})$
   - 左边假设：$V < V_\alpha$
   - 右边假设：$V > V_{1-\alpha}$
4. 根据样本值计算检验量的观测值，若落入拒绝域，则拒绝 H_0，接受 H_1 ；否则接受 H_0。

两类错误及检验的评价标准

在给定的 α 下，接受还是拒绝原假设完全取决于样本的观测值，因此所作的检验可能导致以下两类问题的产生：

真实情况\所作判断	接受 H_0	拒绝 H_0
H_0 为真	正确	第一类错误（弃真）
H_0 为假	第二类错误（纳伪）	正确

- 重要性质：**当样本容量一定时，犯两类错误的概率不会同时减小**

假设检验的评价准则

- 功效函数：设 θ 为总体的待推断参数，对于一个具有拒绝域 W 的检验 τ ，定义

  $$\beta(\theta)=P_\theta(W)=P_\theta((X_1,X_2,\cdots,X_n)\in W)$$

  为该检验的功效函数。

- 功效函数的意义：功效函数描述的是样本落在拒绝域中的概率。若设犯两类错误的概率为 α 和 β ，接受域为 θ_0 ，拒绝域为 θ_1，则

  $$\begin{aligned}\alpha&=\beta(\theta) \qquad &(\theta\in\Theta_0)\\ \beta&=1-\beta(\theta) \qquad &(\theta\in\Theta_1)\end{aligned}$$

- 一致最优检验
  对给定的 $\alpha\in(0,1)$ ，设 τ^* 为一个水平 α 的检验，若对于任意一个水平 α 的检验 τ ，都有

  $$\beta_{\tau^*}(\theta)>\beta_\tau(\theta),\ \ \forall\theta\in\Theta_1$$

  则称 τ^* 为一致最优检验（UMP）。
  - 一致最优检验，在第一类错误控制在 α 以内的情况下，总是使得第二类错误达到最小。
  - 双边检验一般不存在 UMP 检验。

- 无偏检验：
  设对于检验 τ ，其功效为 $\beta(\theta)$ ，若对任意 $\theta_0\in\Theta_0$ 以及 $\theta_1\in\Theta_1$ ，都有：

  $$\beta(\theta_0)\le\beta(\theta_1)$$

  则称 τ 为一个无偏检验。
  - 无偏检验要求一个检验翻地一类错误的概率总不超过不犯第二类错误的概率。

正态总体的假设检验

单正态总体的情形

σ^2 已知的情形下，检验 μ

构造检验统计量为

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

则拒绝域为

原假设 H_0	备择假设 H_1	拒绝域
$\mu=\mu_0$	$\mu\neq\mu_0$	$\|U\|>u_{1-\frac{\alpha}{2}}$
$\mu\ge\mu_0$	$\mu<\mu_0$	$U<-u_{1-\alpha}$
$\mu\le\mu_0$	$\mu>\mu_0$	$U>u_{1-\alpha}$

σ^2 未知的情形下，检验 μ

构造检验统计量为

$$T=\frac{\bar{X}-\mu_0}{\frac{S}{\sqrt{n}}}\sim t(n-1)$$

则拒绝域为

原假设 H_0	备择假设 H_1	拒绝域
$\mu = \mu_0$	$\mu \neq \mu_0$	$\|T\| > t_{1-\frac{\alpha}{2}}$
$\mu \geq \mu_0$	$\mu < \mu_0$	$T < -t_{1-\alpha}$
$\mu \leq \mu_0$	$\mu > \mu_0$	$T > t_{1-\alpha}$

μ 已知的情形下，检验 σ^2

构造检验统计量为

$$\chi^2 = \frac{\sum_{i=0}^{n}(X_i - \mu)^2}{\sigma_0^2} \sim \chi^2(n)$$

则拒绝域为

原假设 H_0	备择假设 H_1	拒绝域
$\sigma^2 = \sigma_0^2$	$\sigma^2 \neq \sigma_0^2$	$\chi^2 < \chi^2_{\frac{\alpha}{2}}(n)\ or\ \chi^2 > \chi^2_{1-\frac{\alpha}{2}}(n)$
$\sigma^2 \geq \sigma_0^2$	$\sigma^2 < \sigma_0^2$	$\chi^2 < \chi^2_{\alpha}(n)$
$\sigma^2 \leq \sigma_0^2$	$\sigma^2 > \sigma_0^2$	$\chi^2 > \chi^2_{1-\alpha}(n)$

μ 未知的情形下，检验 σ^2

构造检验统计量为

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

则拒绝域为

原假设 H_0	备择假设 H_1	拒绝域
$\sigma^2 = \sigma_0^2$	$\sigma^2 \neq \sigma_0^2$	$\chi^2 < \chi^2_{\frac{\alpha}{2}}(n-1)\ or\ \chi^2 > \chi^2_{1-\frac{\alpha}{2}}(n-1)$
$\sigma^2 \geq \sigma_0^2$	$\sigma^2 < \sigma_0^2$	$\chi^2 < \chi^2_{\alpha}(n-1)$
$\sigma^2 \leq \sigma_0^2$	$\sigma^2 > \sigma_0^2$	$\chi^2 > \chi^2_{1-\alpha}(n-1)$

两个正态总体的情形

σ_1^2 和 σ_2^2 均已知的情形下，检验 $\mu_1 - \mu_2$

构造检验统计量为

$$U = \frac{\bar{X} - \bar{Y} - \delta}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}} \sim N(0,1)$$

则拒绝域为

原假设 H_0	备择假设 H_1	拒绝域
$\mu_1 - \mu_2 = \delta$	$\mu_1 - \mu_2 \neq \delta$	$\|U\| > u_{1-\frac{\alpha}{2}}$
$\mu_1 - \mu_2 \geq \delta$	$\mu_1 - \mu_2 < \delta$	$U < -u_{1-\alpha}$
$\mu_1 - \mu_2 \leq \delta$	$\mu_1 - \mu_2 > \delta$	$U > u_{1-\alpha}$

σ_1^2 和 σ_2^2 均未知但 $\sigma_1^2 = \sigma_2^2$ 的情形下，检验 $\mu_1 - \mu_2$

构造检验统计量为

$$T = \frac{\bar{X} - \bar{Y} - \delta}{\sqrt{\frac{1}{n} + \frac{1}{m}}S_w} \sim T(n+m-2)$$

则拒绝域为

原假设 H_0	备择假设 H_1	拒绝域
$\mu_1-\mu_2=\delta$	$\mu_1-\mu_2\neq\delta$	$\|\|T\|\|>t_{1-\frac{\alpha}{2}}$
$\mu_1-\mu_2\geq\delta$	$\mu_1-\mu_2<\delta$	$T<-t_{1-\alpha}$
$\mu_1-\mu_2\leq\delta$	$\mu_1-\mu_2>\delta$	$T>t_{1-\alpha}$

其中

$$S_w=\sqrt{\frac{(n-1)S_1^2+(m-1)S_2^2}{n+m-2}}$$

μ_1 和 μ_2 均未知的情形下，检验 σ_1^2/σ_2^2

构造检验统计量为

$$F=\frac{S_1^2}{S_2^2}\sim F(n-1,m-1)$$

则拒绝域为

原假设 H_0	备择假设 H_1	拒绝域
$\sigma^2=\sigma_0^2$	$\sigma^2\neq\sigma_0^2$	$F<F_{\frac{\alpha}{2}}(n-1,m-1)\ or\ F>F_{1-\frac{\alpha}{2}}(n-1,m-1)$
$\sigma^2\geq\sigma_0^2$	$\sigma^2<\sigma_0^2$	$F<F_{\alpha}(n-1,m-1)$
$\sigma^2\leq\sigma_0^2$	$\sigma^2>\sigma_0^2$	$F>F_{1-\alpha}(n-1,m-1)$

样本容量的选取

假设检验在样本容量固定时，不能同时控制犯第一类错误和犯第二类错误的概率。但是当样本容量不受限制的时候，可以通过选取合适的 n 的值，使犯取伪错误的概率 β 控制在希望的限度，具体地，当 n 满足下式时，犯第二类错误的概率不超过给定的 β：

- 对于单边检验：

$$\sqrt{n}\geq(u_{1-\alpha}+u_{1-\beta})\sigma/\delta$$

- 对于双边检验：

$$\sqrt{n}\geq(u_{1-\frac{\alpha}{2}}+u_{1-\beta})\sigma/\delta$$

总体分布的假设检验

拟合优度检验

- 定义：总体的分布不再是已知属于某种类型的分布族，问题转变为对总体是否属于某分布族进行统计推断：基于样本信息，验证总体数据与假定之可能分布的拟合程度。

具体方式

Pearson χ^2 检验

设完备事件组 A_1，A_2，$\cdots$，A_k，原假设

$$H_0:P(A_i)=p_i,\quad i=1,2,\cdots,k$$

进行 n 次独立重复试验，事件出现的频数分别为

$$v_1,\ v_2,\ \cdots,\ v_k$$

- 定理（Pearson）：若假设 H_0 成立，则当 $n\to\infty$ 时，

$$V=\sum_{i=1}^{k}\frac{(v_i-np_i)^2}{np_i}$$

  趋向于 $\chi^2(k-r-1)$ 分布，其中 r 是用最大似然估计法估计的未知参数的个数。

- 拒绝域：$V > \chi^2_{1-\alpha}(k-r-1)$

- 应用范围：Pearson χ^2 检验是针对离散多项分布的检验，对于连续分布，可在如下推广后，再来应用：
  - 首先将实数轴划分为 k 个区间
  - 计算拟合分布 F_0 在各个区间的概率
  - 同时计算样本点落在各个区间的频数
  - 再利用 Pearson χ^2 检验法检验

偏度、峰度检验法

对于标准正态分布，偏度系数 $v_1 = 0$；峰度系数 $v_2 = 3$
总体 X 的偏度和峰度的矩估计量分别为：

$$G_1 = B_3/B_2^{\frac{3}{2}}$$

称为样本偏度；

$$G_2 = B_4/B_2^2$$

称为样本峰度，其中 B_k 为样本 k 阶中心矩。

若总体 X 为正态变量，则可证得：当 n 充分大时，近似地有：

$$G_1 \sim N(0, \sigma_1^2)$$
$$G_2 \sim N(\mu_2, \sigma_2^2)$$

其中

$$\sigma_1^2 = \frac{6(n-2)}{(n+1)(n+3)}$$
$$\mu_2 = 3 - \frac{6}{(n+1)}$$
$$\sigma_2^2 = \frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}$$

- 检验方式：
  由大数定律可得，样本偏度和样本峰度依概率收敛于总体偏度和总体峰度。故当 $H_0 \sim N(\mu, \sigma^2)$ 为真且 n 足够大时，G_1 与 v_1 的偏离不应该太大，G_2 与 v_2 的偏离不应该太大。于是：

  $$U_1 = G_1/\sigma_1 \sim N(0,1)$$
  $$U_2 = (G_2 - \mu_2)/\sigma_2 \sim N(0,1)$$

  的绝对值的观测值不应该过大。
  - 拒绝域：设 $k1$，$k2$ 满足

    $$P_{H_0}(|U_1| \geq k_1) = \frac{\alpha}{2}$$
    $$P_{H_0}(|U_2| \geq k_2) = \frac{\alpha}{2}$$

    即 $k_1 = u_{1-\frac{\alpha}{4}}$，$k_2 = u_{1-\frac{\alpha}{4}}$ 则拒绝域为：

    $$|U_1| \geq k_1 \ or \ |U_2| \geq k_2$$