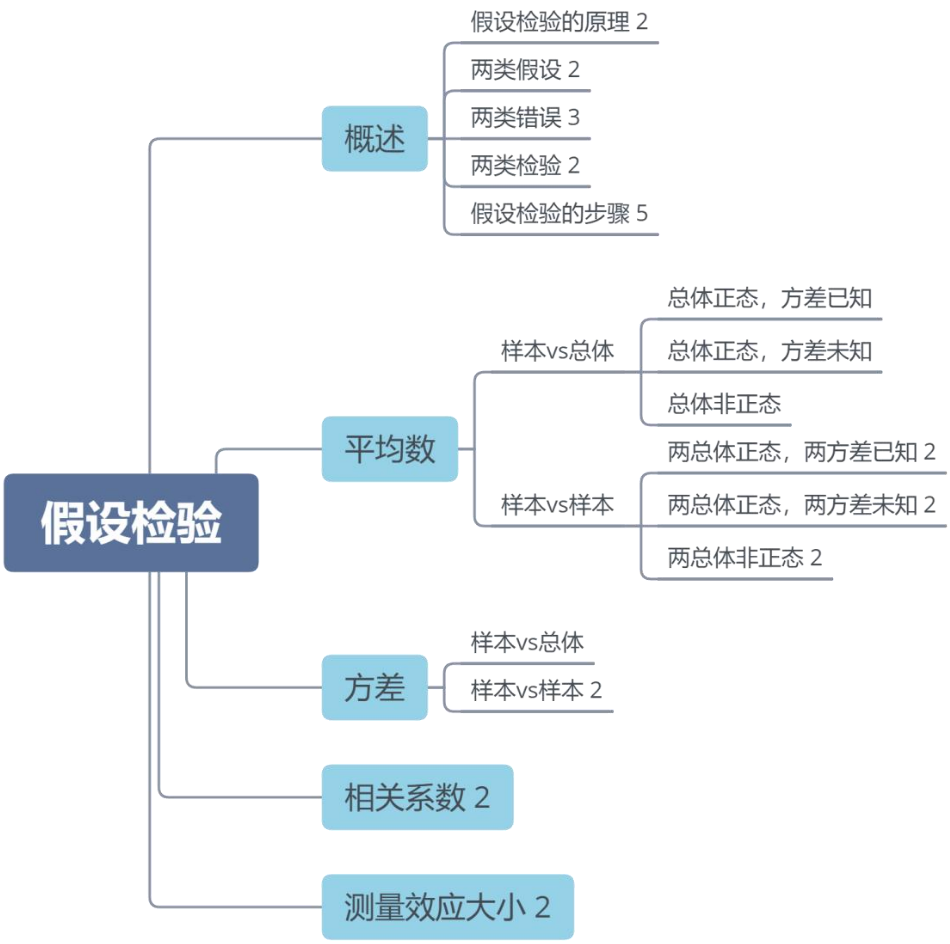
第八章　假设检验



一、概述

通过样本统计量得出的差异做出一般性结论，判断总体参数之间是否存在差异，其推论过程称作假设检验。

假设检验包括参数检验和非参数检验。若进行假设检验时，总体的分布形式已知，需要对总体的未知参数进行假设检验，称其为参数假设检验；若对总体分布形式所知甚少，需要对未知分布函数的形式及其他特征进行假设检验，称之为非参数假设检验。

（一）假设检验的原理

1.反证法

为了检验 H_0，首先需要假设 H_0 为真，若出现“不合理现象”，则不能接受 H_0，转而接受 H_1；若没有出现不合理现象，接受 H_0，拒绝 H_1。而“不合理现象”指小概率事件在一次试验中发生的情况。

2.小概率事件原理

小概率事件在一次试验中不可能发生，通常将概率不超过 0.05 或 0.01 的事件称为“小概率事件”。

总结：假设检验是利用了基于小概率事件的反证法。

（二）两类假设

1. H_1：希望得到证实的假设，又称备择假设、研究假设、科学假设或对立假设。

2. H_0：直接被检验的假设，又称虚无假设、无差假设、零假设或原假设。

在统计学中不能对 H_1 直接进行检验，所以需要建立与之对立的假设 H_0，两者有且只有一个正确，而 H_0 则是统计推论的出发点。

（三）两类错误

1. I型错误与II型错误

（1）I型错误：当 H_0 正确时，拒绝了 H_0 时所犯的错误，也叫α错误、弃真错误，其概率为α；指研究者得出了处理有效应的结论，而实际上并没有效果，即所谓的"无中生有"。

（2）II型错误：当 H_1 正确（H_0 错误）时，拒绝了 H_1（接受 H_0）时所犯的错误，也叫β错误、取伪错误，其概率为β；假设检验未能侦查到实际存在的处理效应，即所谓的"失之交臂"。

通常，将犯I型错误的概率α称为假设检验的显著性水平。经检验，若差异超过某一误差限度，则表明总体差异显著（差异具有统计学意义），但这并不一定意味着实际效果"显著"。

2.两类错误的关系

（1）因为α与β是在两个相互对立的前提下的概率，所以α+β不一定等于 1。

（2）在其他条件不变的情况下，α与β不可能同时减小或增大。

（3）在规定了α的情况下要同时尽量减少β，直接的方法就是增大样本容量。

3.统计检验力

（1）含义

指某个检验能够正确拒绝一个错误的虚无假设（H_0）的概率，它反映着正确辨认真实差异的能力，统计学中用（1-β）来表示。

（2）影响因素

①处理效应大小：处理效应越明显，越容易被检测到，统计检验力越大。

②显著性水平α：α增大，则β相应减小，所以（1-β）增大，即拒绝虚无假设的概率增大，统计检验力就越大。

③检验的方向性：单侧检验的统计检验力要高于双侧检验。

④样本容量：样本容量越大，标准误就越小，样本分布均值越集中，统计检验力就越大。

（四）两类检验

1.单侧检验

强调某方向的检验，如是否显著"大于""优于"等。如：

H_0：$\mu_1 \le \mu_0$　　　H1：$\mu_1 > \mu_0$

2.双侧检验

只强调差异不强调方向性的检验，如是否有显著差异。如：

H_0：$\mu_1 = \mu_0$　　　H_1：$\mu_1 \ne \mu_0$

（五）假设检验的步骤

1.列出已知条件和两类假设；

2.计算标准误；

3.根据样本分布计算相应的“Z/t/…”值；

4.查表获得临界值；

5.将算出来的值和查出来的临界值进行比较，若超过临界值，则说明差异显著。

二、平均数差异检验（样本 vs 总体）

对样本平均数与总体平均数之间的差异进行显著性检验。

（一）总体正态，方差已知

样本均值服从 Z 分布，使用 Z 检验：

$$SE_{\overline{x}}=\frac{\sigma}{\sqrt{n}}\text{，}Z=\frac{\overline{X}-\mu_0}{SE_{\overline{x}}}$$

（二）总体正态，方差未知

样本均值服从 t 分布，使用 t 检验：

$$SE_{\overline{x}}=\frac{s}{\sqrt{n-1}}=\frac{s_{n-1}}{\sqrt{n}}\quad t=\frac{\overline{X}-\mu_0}{SE_{\overline{x}}}\text{，}df=n-1$$

（三）总体非正态，但 n≥30 时

样本均值近似服从 Z 分布，使用 Z 检验：

$$SE_{\overline{x}}=\frac{\sigma}{\sqrt{n}}$$ 当 σ 未知时，可用样本标准差 s 代替

$$Z'=\frac{\overline{X}-\mu_0}{SE_{\overline{x}}}$$

三、平均数差异的显著性检验（样本 vs 样本）

对两个样本平均数之间的差异进行检验。目的在于通过样本平均数之间的差异（$\overline{X}_1-\overline{X}_2$）来检验两总体之间的差异（$\mu_1-\mu_2$）。

（一）两总体正态，两方差已知

样本均值之差服从 Z 分布，使用 Z 检验。

1.独立样本

$$SE_{D_{\overline{X}}}=\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$$

$$Z=\frac{\overline{X}_1-\overline{X}_2}{SE_{D_{\overline{X}}}}$$

2.相关样本

$$SE_{D_{\overline{X}}}=\sqrt{\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{n}-2r\times\frac{\sigma_1}{\sqrt{n}}\times\frac{\sigma_2}{\sqrt{n}}}$$

$$Z=\frac{\overline{X}_1-\overline{X}_2}{SE_{D_{\overline{X}}}}$$

（二）两总体正态，两方差未知

样本均值之差服从 t 分布，使用 t 检验。

1.独立样本（需进行方差齐性检验）

（1）方差齐性

$$SE_{D_{\overline{X}}}=\sqrt{\frac{n_1s_1^2+n_2s_2^2}{n_1+n_2-2}\times(\frac{n_1+n_2}{n_1n_2})}$$

$$t=\frac{\overline{X}_1-\overline{X}_2}{SE_{D_{\overline{X}}}}，\quad df=n_1+n_2-2$$

（2）方差不齐

柯克兰-柯克斯 t 检验

$$SE_{D\overline{X}}=\sqrt{\frac{s_1^2}{n_1-1}+\frac{s_2^2}{n_2-1}}，\quad t=\frac{\overline{X}_1-\overline{X}_2}{SE_{D_{\overline{X}}}}，\text{临界值}t_\alpha'=\frac{SE_{\overline{X}_1}^2t_{1(\alpha)}+SE_{\overline{X}_2}^2t_{2(\alpha)}}{SE_{\overline{X}_1}^2+SE_{\overline{X}_2}^2}$$

2.相关样本

（1）r 已知

$$SE_{D_{\overline{X}}}=\sqrt{\frac{s_1^2+s_2^2-2rs_1s_2}{n-1}}，\quad t=\frac{\overline{X}_1-\overline{X}_2}{SE_{D_{\overline{X}}}}，\quad df=n-1$$

（2）r 未知

$$SE_{D_{\overline{X}}}=\sqrt{\frac{s_d^2}{n-1}}=\sqrt{\frac{\sum d^2-\frac{(\sum d)^2}{n}}{n(n-1)}}，\quad t=\frac{\overline{X}_1-\overline{X}_2}{SE_{D_{\overline{X}}}}，\quad df=n-1，\quad d=X_{1i}-X_{2i}$$

（三）两总体非正态，但 n1 和 n2 均大于 30 时

样本均值之差近似服从 Z 分布，使用 Z 检验。

1.独立样本

$$SE_{D_{\bar{X}}}=\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}} \qquad Z'=\frac{\overline{X}_1-\overline{X}_2}{SE_{D_{\bar{X}}}}$$

当 σ 未知时，可用样本标准差 s 代替

2.相关样本

$$SE_{D_{\bar{X}}}=\sqrt{\frac{\sigma_1^2+\sigma_2^2-2r\sigma_1\sigma_2}{n}}，Z'=\frac{\overline{X}_1-\overline{X}_2}{SE_{D_{\bar{X}}}}$$

当 σ 未知时，可用样本标准差 s 代替

四、方差齐性检验

（一）样本方差与总体方差（样本 VS 总体）

当从正态分布的总体中随机抽取容量为 n 的样本时，其样本方差与总体方差的比值服从 χ^2 分布，使用 χ^2 检验，df=n-1：

$$\chi^2=\frac{ns^2}{\sigma_0^2}$$

（二）两个样本方差之间（样本 VS 样本）

通过样本方差之间的差异对其总体方差之间是否有差异进行判断。

1.独立样本

样本方差之比服从 F 分布，使用 F 检验：

$$F=\frac{s_{大}^2}{s_{小}^2}，df_1=n_1-1，df_2=n_2-1$$

2.相关样本

样本方差之比服从 t 分布，使用 t 检验：

$$SE=\sqrt{\frac{4s_1^2s_2^2(1-r^2)}{n-2}}，t=\frac{s_1^2-s_2^2}{SE}，df=n-2$$

五、相关系数的显著性检验

（一）积差相关系数的显著性检验

1.当ρ=0 时：

$$SE=\sqrt{\frac{1-r^2}{n-2}}$$

$$t=\frac{r-0}{SE}，\quad df=n-2$$

2.当ρ≠0 时，先通过查表将 r 和ρ转化为费舍 Z_r 和 Z_ρ，然后进行 Z 检验：

$$SE=\sqrt{\frac{1}{n-3}}，\quad Z=\frac{Z_r-Z_\rho}{SE}$$

（二）其他相关系数的显著性检验

原理与上述各类型的假设检验相同，这里不做过多介绍。具体知识可见相应参考书目。

六、测量效应大小

（一）科恩 d 值

1.含义

科恩 d 值测量了两个分布的分散程度。

2.计算方法

（1）总体值已知

科恩 d 值=平均数差/标准差=M－μ/s

（2）总体值未知

科恩 d 值=平均数差/样本标准差=M－μ/s（$\sqrt{\frac{SS}{df}}$）

3.评估标准

d的大小	评价效应大小
0<d<0.2	效应较小（平均数差异小于0.2个标准差）
0.2<d<0.8	差异中等（平均数差异约为0.5个标准差）
d>0.8	效应较大（平均数差异大于0.8个标准差）

（二）r^2

1.含义

测量由处理引起的方差百分率。

2.计算方法

$$r^2=\frac{t^2}{t^2+df}$$

3.评估标准

r^2的大小	评价效应大小
0.01<r^2<0.09	小效应
0.09<r^2<0.25	中效应
r^2>0.25	大效应