概率论与数理统计第22讲

本讲义可在网址http://math.shekou.com下载

第七章　假设检验

统计推断的另一类重要问题是假设检验, 在总体分布未知或虽知其类型但含有未知参数的时候, 为推断总体的某些未知特性, 提出某些关于总体的假设. 我们需要根据样本所提供的信息以及运用适当的统计量, 对提出的假设作出接受或拒绝的决策, 假设检验是作出这一决策的过程.

假设检验 $\begin{cases} \text{参数假设检验} \\ \text{非参数假设检验} \end{cases}$

参数假设检验是针对总体分布函数中的未知参数而提出的假设进行检验, **非参数假设检验**是针对总体分布函数形式或类型的假设进行检验, 本章主要讨论单参数假设检验问题

§ 7.1 假设检验的基本概念

一, 引例

设一箱中有红白两种颜色的球共100个, 甲说这里有98个白球, 乙从箱中任取一个, 发现是红球, 问甲的说法是否正确?

先作假设H_0: 箱中确有98个白球.
如果假设H_0正确, 则从箱中任取一个球是红球的概率只有0.02, 是小概率事件. 通常认为在一次随机试验中, 概率小的事件不易发生, 因此, 若乙从箱中任取一个, 发现是白球, 则没有理由怀疑假设H_0的正确性. 今乙从箱中任取一个, 发现是红球, 即小概率事件竟然在一次试验中发生了, 故有理由拒绝假设H_0, 即认为甲的说法不正确.

二, 假设检验的基本思想

假设检验的基本思想实质上是带有某种概率性质的反证法. 为了检验一个假设H_0是否正确, 首先假定该假设H_0正确, 然后根据抽取到的样本对假设H_0作出接受或拒绝的决策. 如果样本观察值导致了不合理的现象发生, 就应拒绝假设H_0, 否则应接受假设H_0.

假设检验中所谓"不合理", 并非逻辑中的绝对矛盾, 而是基于人们在实践中广泛采用的原则, 即小概率事件在一次试验中是几乎不发生的. 但概率小到什么程度才能算作"小概率事件", 显然, 概率越小, 否定假设H_0越有说服力. 常记这个概率值为$\alpha(0<\alpha<1)$, 称为检验的显著性水平, 对不同的问题, 检验的显著性水平α不一定相同, 但一般应取为较小的值, 如0.1, 0.05, 或0.01等.

三、假设检验的两类错误

当假设H_0正确时, 小概率事件也有可能发生, 此时我们会拒绝假设H_0, 因而犯了"弃真"的错误, 称此为**第一类错误**. 犯第一类错误的概率恰好就是"小概率事件"发生的概率α, 即

$$P\{拒绝H_0|H_0为真\}=\alpha.$$

反之, 若假设H_0不正确, 但一次抽样检验结果, 未发生不合理的结果, 这时我们会接受H_0, 因而犯了"取伪"的错误, 称此为**第二类错误**. 记β为犯第二类错误的概率, 即

$$\mathrm{P}\{接受H_0|H_0不真\}=\beta.$$

理论上, 自然希望犯这两类错误的概率都很小. 当样本容量n固定时, α,β不能同时都小, 即α变小时, β就变大; 而β变小时, α就变大. 一般只有当样本容量n增大时, 才有可能使两者变小. 在实际应用中, 一般原则是: 控制犯第一类错误的概率, 即给定α, 然后通过增大样本容量n来减小β.

关于显著性水平α的选取: 若注重经济效益, α可取小些, 如α=0.01; 若注重社会效益, α可取大些, 如α=0.1; 若要兼顾经济效益和社会效益, 一般可取α=0.05.

四, 假设检验问题的一般提法

在假设检验问题中, 把要检测的假设H_0称为**原假设**(**零假设**或**基本假设**), 把原假设H_0的对立面称为**备择假设**或**对立假设**, 记为H_1.

例如, 有一封装罐装可乐的生产流水线, 每罐的标准容量规定为350毫升. 质检员每天都要检验可乐的容量是否合格, 已知每罐的容量服从正态分布, 且生产比较稳定时, 其标准差σ=5毫升. 某日上班后, 质检员每隔半小时从生产线上取一罐, 共抽取了6罐, 测得容量(单位为毫升)如下:

353, 345, 357, 339, 355, 360.

试问生产线工作是否正常?

本例的假设检验问题可简记为

$$H_0:\mu=\mu_0,\ H_1:\mu\neq\mu_0,\ (\mu_0=350)\quad(1.1)$$

形如(1.1)式的备择假设H_1, 表示μ可能大于μ_0也可能小于μ_0, 称为**双侧(边)备择假设**. 形如(1.1)式的假设检验**称为双侧(边)假设检验**.

在实际问题中, 有时还需要检验下列形式的假设

$$H_0:\mu\le\mu_0,\ H_1:\mu>\mu_0, \qquad (1.2)$$

$$H_0:\mu\ge\mu_0,\ H_1:\mu<\mu_0, \qquad (1.3)$$

形如(1.2)式的假设检验称为**右侧(边)检验**;
形如(1.3)式的假设检验称为**左侧(边)检验**;
右侧(边)检验和左侧(边)检验统称为**单侧(边)检验**.

为检验提出的假设, 通常需构造检验统计量, 并取总体的一个样本值, 根据该样本提供的信息来判断假设是否成立. 当检验统计量取某个区域W中的值时, 我们拒绝原假设H_0, 则称W为**拒绝域**, 拒绝域的边界点称为**临界点**.

五, 假设检验的一般步骤

(1) 根据实际问题的要求, 充分考虑和利用已知的背景知识, 提出原假设H_0及备择假设H_1;

(2) 给定显著性水平α以及样本容量n;

(3) 确定检验统计量U, 并在原假设H_0成立的前提下导出U的概率分布, 要求U的分布不依赖于任何未知参数;

(4) 确定拒绝域, 即依据直观分析先确定拒绝域的形式, 然后根据给定的显著性水平α和U的分布, 由

$$P\{\text{拒绝}H_0|H_0\text{为真}\}=\alpha$$

确定拒绝域的临界值, 从而确定拒绝域;

(5) 作一次具体的抽样, 根据得到的样本观察值和所得的拒绝域, 对假设H_0作出拒绝或接受的判断.

例1 某化学日用品有限责任公司用包装机包装洗衣粉, 洗衣粉包装机在正常工作时, 装包量$X \sim N(500,2^2)$(单位:g), 每天开工后, 需先检验包装机工作是否正常. 某天开工后, 在装好的洗衣粉中任取9袋, 其重量如下:

505, 499, 502, 506, 498, 498, 497, 510, 503

假设总体标准差σ不变, 即$\sigma=2$, 试问这天包装机工作是否正常?($\alpha=0.05$)

解 (1) 提出假设检验:

$$H_0:\mu=500, \quad H_1:\mu\neq 500$$

(2) 以H_0成立为前提, 确定检验H_0的统计量及其分布.

$$U=\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}=\frac{\bar{X}-500}{2/3}\sim N(0,1).$$

(3) 对给定的显著性水平α=0.05, 确定H_0的接受域 $\bar{W}$或拒绝域W. 取临界点为$u_{\alpha/2}$=1.96, 使$P\{|U|>u_{\alpha/2}\}=\alpha$.

故 H_0 被接受与拒绝的区域分别为

$$\overline{W}=[-1.96,1.96]$$

$$W=(-\infty,-1.96)\bigcup(1.96,+\infty).$$

(4) 由样本计算统计量 U 的值

$$u=\frac{502-500}{2/3}=3.$$

(5) 对假设 H_0 作出推断

因为 $u\in W$(拒绝域)，故认为这天洗衣粉包装机工作不正常.

六, 多参数与非参数假设检验问题

原则上, 以上介绍的所有单参数假设检验的内容也适用于多参数与非参数假设检验问题, 只需在某些细节上作适当调整即可, 这里仅说明下列两点:

(1) 对多参数假设检验问题, 要寻求一个包含所有待检参数的检验统计量, 使之服从一个已知的确定分布;

(2) 非参数假设检验问题可近似地化为一个多参数假设检验问题.

§7.2 单正态总体的假设检验

一, 总体均值的假设检验

当检验关于总体均值μ(数学期望)的假设时, 该总体中的另一个参数, 即方差σ^2是否已知, 会影响到对于检验统计量的选择, 故下面分两种情形进行讨论.

1. 方差σ^2已知情形

设总体$X\sim N(\mu,\sigma^2)$, 其中总体方差σ^2已知, $X_1,X_2,\ldots,X_n$是取自总体X的一个样本, $\bar{X}$为样本均值.

(1) 检验假设H_0:$\mu=\mu_0$, H_1:$\mu\neq\mu_0$. 其中μ_0为已知常数. 则当H_0为真时,

$$U=\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\sim N(0,1), \qquad (2.1)$$

故选取U作为检验统计量, 记其观察值为u, 相应的检验法称为***u*检验法**.

因为 $\overline{X}$是μ的无偏估计量, 当H_0成立时, $|u|$不应太大, 当H_1成立时, $|u|$有偏大的趋势, 故拒绝域形式为

$$|u| = \left| \frac{\overline{x} - \mu_0}{\sigma / \sqrt{n}} \right| > k \quad (k\text{待定})$$

对于给定的显著性水平α, 查标准正态分布表得$k=u_{\alpha/2}$, 使

$$P\{|U| > u_{\alpha/2}\} = \alpha$$

由此即得拒绝域为

$$|u| = \left|\frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}\right| > u_{\alpha/2}. \qquad (2.2)$$

即 $W = (-\infty, -u_{\alpha/2}) \cup (u_{\alpha/2}, +\infty)$

根据一次抽样后得到的样本观察值 $x_1, x_2, \ldots, x_n$ 计算出 U 的观察值 u, 若 $|u| > u_{\alpha/2}$, 则拒绝原假设 H_0, 即认为总体均值与 μ_0 有显著差异; 若 $|u| \le u_{\alpha/2}$, 则接受原假设 H_0, 即认为总体均值与 μ_0 无显著差异.

类似地, 对单侧检验有:

(2)右侧检验: 检验假设H_0:$\mu\le\mu_0$, H_1:$\mu>\mu_0$, 其中μ_0为已知常数. 可得拒绝域为

$$u=\frac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}>u_\alpha \qquad (2.3)$$

(3)左侧检验: 检验假设H_0:$\mu \geq \mu_0$, H_1:$\mu < \mu_0$, 其中μ_0为已知常数. 可得拒绝域为

$$u = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} < -u_\alpha \qquad (2.4)$$

例1 某车间生产钢丝, 用X表示钢丝的折断力, 由经验判断$X \sim N(\mu,\sigma^2)$, 其中$\mu=570$, $\sigma^2=8^2$; 今换了一批材料, 从性能上看估计折断力的方差不会有什么变化(即仍有$\sigma^2=8^2$), 但不知折断力的均值μ和原先有无差别. 现抽得样本, 测得其折断力为:

578 572 570 568 572 570 570 572 596 584

取$\alpha=0.05$, 试检验折断力均值有无变化?

解 建立假设$H_0:\mu=\mu_0=570$, $H_1:\mu\neq 570$

查正态分布表得$u_{\alpha/2}=u_{0.025}=1.96$, 从而拒绝域为

$$|u|>1.96$$

计算出

$$\overline{x}=\frac{1}{10}\sum_{i=1}^{10}x_i=575.20,\ \sigma^2=64$$

所以,

$$|u|=\left|\frac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}\right|=2.06>1.96$$

故应拒绝H_0, 即认为折断力的均值发生了变化

例2 有一工厂生产一种灯管, 已知灯管的寿命X服从正态分布$N(\mu,40000)$, 根据以往的生产经验, 知道灯管的平均寿命不会超过1500小时. 为了提高灯管的平均寿命, 工厂采用了新的工艺. 为了弄清楚新工艺是否真的能提高灯管的平均寿命, 他们测试了采用新工艺生产的25只灯管的寿命, 其平均值是1575小时. 尽管样本的平均值大于1500小时, 试问: 可否由此判定这恰是新工艺的效应而非偶然的原因使得抽出的这25只灯管的平均寿命较长呢?

解 可把问题转述为假设检验问题:

H_0:μ≤1500, H_1:μ>1500

从而可用右侧检验法来检验, 相应于 μ_0=1500, σ=200, n=25取显著性水平为 α=0.05, 查附表得u_α=1.645, 因已测出 $\bar{x}$=1575, 从而

$$u=\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}=\frac{1575-1500}{200}\sqrt{25}=1.875$$

$$u=\frac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}=\frac{1575-1500}{200}\sqrt{25}=1.875$$

由于$u=1.875>u_{\alpha}=1.645$, 从而否定原假设H_0, 接受备择假设H_1, 即认为新工艺事实上提高了灯管的平均寿命.

2. 方差σ^2未知情形

设总体$X\sim N(\mu,\sigma^2)$, 其中总体方差σ^2未知, $X_1,X_2,\ldots,X_n$是取自X的一个样本, $\bar{X}$与S^2分别为样本均值与样本方差.

(1) 检验假设$H_0:\mu=\mu_0$, $H_1:\mu\neq\mu_0$. 其中μ_0为已知常数. 则当H_0为真时,

$$T=\frac{\bar{X}-\mu_0}{S/\sqrt{n}}\sim t(n-1), \qquad (2.5)$$

$$T=\frac{\bar{X}-\mu_0}{S/\sqrt{n}}\sim t(n-1), \qquad (2.5)$$

故选取T为检验统计量, 记其观察值为t, 相应的检验法称为t检验法. 由于 $\bar{X}$是μ的无偏估计量, S^2是σ^2的无偏估计量, 当H_0成立时, $|t|$不应太大, 当H_1成立时, $|t|$有偏大的趋势. 对于给定的显著性水平α, 查分布表得$t_{\alpha/2}(n-1)$, 使

$$P\{|T|>t_{\alpha/2}(n-1)\}=\alpha$$

由此得拒绝域为

$$|t| = \left|\frac{\bar{x} - \mu_0}{s/\sqrt{n}}\right| > t_{\alpha/2}(n-1).$$

即 $W = (-\infty, -t_{\alpha/2}(n-1)) \cup (t_{\alpha/2}(n-1), +\infty)$ (2.6)

根据一次抽样后得到的样本观察值 $x_1, x_2, \ldots, x_n$ 计算出 T 的观察值 t, 若 $|t| > t_{\alpha/2}(n-1)$, 则拒绝原假设 H_0, 即认为总体均值与 μ_0 有显著差异; 若 $|t| \le t_{\alpha/2}(n-1)$, 则接受原假设 H_0, 即认为总体均值与 μ_0 无显著差异.

类似地, 对单侧检验有:

(2)右侧检验: 检验假设$H_0:\mu\le\mu_0$, $H_1:\mu>\mu_0$, 其中μ_0为已知常数. 可得拒绝域为

$$t=\frac{\overline{x}-\mu_0}{s/\sqrt{n}}>t_\alpha(n-1) \qquad (2.7)$$

(3)左侧检验: 检验假设H_0:$\mu \geq \mu_0$, H_1:$\mu < \mu_0$, 其中μ_0为已知常数. 可得拒绝域为

$$t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}} < -t_{\alpha}(n-1) \qquad (2.8)$$

例3 水泥厂用自动包装机包装水泥, 每袋额定重量是50kg, 某日开工后随机抽查了9袋, 称得重量如下:

49.6 49.3 50.1 50.0 49.2 49.9 49.8 51.0 51.2

设每袋重量服从正态分布, 问包装机工作是否正常(α=0.05)?

解 建立假设H_0:μ=50, H_1:$\mu\neq$50

样本容量n=9, 因此t分布的自由度为8, 对于给定的显著性水平α=0.05, 查t分布表得$t_{\alpha/2}(n-1)=t_{0.025}(8)=2.306$, 从而拒绝域为$|t|>2.306$. 由数据计算出 $\bar{x}=49.9$, $s^2=0.29$,

$$|t|=\left|\frac{\overline{x}-50}{s/\sqrt{n}}\right|=0.56<2.306$$

故应接受H_0, 即认为包装机工作正常.

例4 一公司声称某种类型的电池的平均寿命至少为21.5小时. 有一实验室检验了该公司制造的6套电池, 得到如下的寿命小时数:

19, 18, 22, 20, 16, 25

试问: 这些结果是否表明, 这种类型的电池低于该公司所声称的寿命? (显著性水平α=0.05).

解 可把上述问题归纳为下述假设检验问题: $H_0:\mu\geq 21.5$, $H_1:\mu<21.5$.

$\mu_0=21.5$, $n=6$, 对于给定的显著性水平 $\alpha=0.05$, 查表得$t_\alpha(n-1)=t_{0.05}(5)=2.015$.

再据测得的6个寿命小时数算得

$$\bar{x}=20, \quad s^2=10.$$

$$t=\frac{\bar{x}-\mu_0}{s/\sqrt{n}}=\frac{20-21.5}{\sqrt{10}}\sqrt{6}=-1.162$$

因$t=-1.162>-2.015=-t_{0.05}(5)$, 所以不能否定原假设$H_0$, 认为这类电池寿命不比公司宣称的寿命差.

三，总体方差的假设检验

设 $X \sim N(\mu,\sigma^2)$, $X_1,X_2,\ldots,X_n$ 是取自 X 的一个样本，$\bar{X}$ 与 S^2 为样本均值和样本方差.

(1) 检验假设 $H_0:\sigma^2=\sigma_0^2, H_1=\sigma^2\neq\sigma_0^2$.其中 σ_0^2 为已知常数. 则当 H_0 为真时,

$$\chi^2=\frac{n-1}{\sigma_0^2}S^2\sim\chi^2(n-1); \qquad (2.9)$$

故选取 χ^2 作为检验统计量. 相应的检验法称为 χ^2 **检验法**.

由于 S^2 是σ^2 的无偏估计量，当 H_0 成立时, S^2 应当在σ_0^2附近，当 H_1 成立时, χ^2 有偏小或偏大的趋势，故拒绝域形式为

$$\chi^2<k_1 \quad 或 \quad \chi^2>k_2 \quad (k_1,k_2 待定).$$

对于给定的显著性水平α，查分布表得

$$k_1=\chi^2_{1-\alpha/2}(n-1), k_2=\chi^2_{\alpha/2}(n-1)$$

使

$$P\{\chi^2<\chi^2_{1-\alpha/2}(n-1)\}=\frac{\alpha}{2}, P\{\chi^2>\chi^2_{\alpha/2}(n-1)\}=\frac{\alpha}{2}$$

由此即得拒绝域为

$$\chi^2=\frac{n-1}{\sigma_0^2}s^2<\chi^2_{1-\alpha/2}(n-1)\text{或}\chi^2=\frac{n-1}{\sigma_0^2}s^2>\chi^2_{\alpha/}$$

即$W=[0,\chi^2_{1-\alpha/2}(n-1))\cup(\chi^2_{\alpha/2}(n-1),+\infty)$ (2.10)

根据一次抽样后得到的样本观察值 $x_1,x_2,\ldots,x_n$ 计算出χ^2的观察值，若 $\chi^2<\chi^2_{1-\alpha/2}(n-1)$或$\chi^2>\chi^2_{\alpha/2}(n-1)$，则拒绝原假设 H_0，若$\chi^2_{1-\alpha/2}(n-1)\le\chi^2\le\chi^2_{\alpha/2}(n-1)$，则接受假设 H_0.

类似地，对单侧检验有:

(2) 右侧检验：检验假设:

$H_0:\sigma^2\le\sigma_0^2, H_1:\sigma^2>\sigma_0^2$. 其中$\sigma_0$为已知常数.

可得拒绝域为

$$\chi^2=\frac{n-1}{\sigma_0^2}s^2>\chi_\alpha^2(n-1) \qquad (2.11)$$

(3) 左侧检验：检验假设:

$H_0:\sigma^2\geq\sigma_0^2, H_1:\sigma^2<\sigma_0^2$. 其中$\sigma_0$为已知常数.

可得拒绝域为

$$\chi^2=\frac{n-1}{\sigma_0^2}s^2<\chi^2_{1-\alpha}(n-1) \qquad (2.12)$$

例5 厂生产的某种型号的电池, 其寿命(以小时计)长期以来服从方差$\sigma^2=5000$的正态分布, 现有一批这种电池, 从它的生产情况来看, 寿命的波动性有所改变. 现在随机取26只电池, 测出其寿命的样本方差$s^2=9200$. 问根据这一数据能否推断这批电池的寿命的波动性较以往的有显著的变化(取$\alpha=0.02$)?

解 检验假设$H_0:\sigma^2=5000, H_1:\sigma^2\neq 5000.$

$$\alpha=0.02, n=26, \sigma_0^2=5000, s^2=9200,$$

$$\chi^2_{\alpha/2}(n-1)=\chi^2_{0.01}(25)=44.314,$$

$$\chi^2_{1-\alpha/2}(n-1)=\chi^2_{0.99}(25)=11.524$$

算出

$$\chi^2=\frac{(n-1)s^2}{\sigma_0^2}=46>44.314,$$

故拒绝H_0, 认为这批电池寿命的波动性较以往有显著变化.

例6 某工厂生产金属丝, 产品指标为折断力, 折断力的方差被用作工厂生产精度的表征. 方差越小, 表明精度越高. 以往工厂一直把方差保持在64(kg^2)与64以下. 最近从一批产品中抽取10根作折断力试验, 测得的结果(单位为千克)如下:

578,572,570,568,572,570,572,596,584,570

由上述样本数据得: $\bar{x}=575.2$, $s^2=76.74$.

判断折断力的方差是否变大了.

解 作假设检验 $H_0:\sigma^2\le 64$, $H_1:\sigma^2>64$

$$\bar{x}=575.2, s^2=75.74, n=10,$$

$$\alpha=0.05, \sigma_0^2=64,$$

$$\chi_\alpha^2(n-1)=\chi_{0.05}^2(9)=16.919.$$

$$\chi^2=\frac{n-1}{\sigma_0^2}s^2=\frac{9\times 75.74}{64}=10.65$$

$$\le 16.919=\chi_{0.05}^2(9)$$

不能拒绝原假设H_0. 生产流程正常无须调整

课堂练习

1. 某饲养厂规定, 屠宰的肉用鸡体重不得少于3kg, 现从该饲养厂的鸡群中, 随机抓16只, 且计算平均体重 $\bar{x}$=2.8kg, 标准差 s=0.2kg, 设肉用鸡重量X服从正态分布, 试以α=0.05的显著性水平作出该批鸡可否屠宰的判断.

2. 某厂生产一种保险丝. 规定保险丝熔化时间的方差不超过400, 现从一批产品中抽取25个, 测得其熔化时间的方差为388.58, 试根据所给数据, 检验这批产品的方差是否符合要求(α=0.05).

作业 习题7-2 第234页开始
第1,4,5,9题