doi:10.16055/j.issn.1672-058X.2015.0009.015

单正态总体假设检验的两类错误概率关系分析

高发玲, 刘春霞, 孙建英, 郭　萍
(青岛理工大学 琴岛学院 ,山东 青岛 266100)

摘　要:针对参数假设检验中的两类错误概率关系问题,以单正态总体均值以及方差假设检验为例,通过检验所得的接受域和拒绝域,具体分析了假设检验的两类错误“弃真”和“取伪”概率的关系.该方法有效地补充了教材中在样本容量固定和控制第一类错误概率情形下两类错误概率变化关系的原理.

关键词:假设检验;两类错误;概率关系
中图分类号:O212.1　　**文献标志码**:A　　**文章编号**:1672-058X(2015)09-0057-04

参数假设检验的原则是小概率事件在一次试验中几乎不可能发生,利用此原则来接受原假设或者拒绝原假设.但小概率事件在一次试验几乎不可能发生但也有可能发生,故会导致犯两类错误,一类是“弃真”错误,即概率为 $\alpha=P\{拒绝\ H_0 \mid H_0\ 为真\}$[1];一类是“取伪“的错误,即概率为 $\beta=P\{接受\ H_0 \mid H_0\ 为假\}$.当样本容量 n 固定时,α 与 β 不能同时都小,通过单正态总体 $X \sim N(\mu,\sigma^2)$ 均值 μ 和方差 σ^2 的假设检验来分析 α 与 β 的关系,即 α 变小时,β 就变大;α 变大时,β 就变小[1,2];另外在控制住 α 时,当增大样本容量 n 时必导致 β 减小.

1　总体均值 μ 的假设检验的两类错误概率关系分析

(1) 方差 σ^2 已知的情形[3].检验假设 $H_0:\mu=\mu_0$,备择假设 $H_1:\mu=\mu_1\neq\mu_0$.

当 H_0为真时,有

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

由于 $P\left\{\left|\frac{\overline{X}-\mu_0}{\frac{\sigma}{\sqrt{n}}}\right|\geqslant u_{\frac{\alpha}{2}}\right\}=\alpha$,则犯第二类错误“取伪“的概率:

$$\beta=P\{接受\ H_0 \mid H_0\ 为假\}=P\{|U|\leqslant u_{\frac{\alpha}{2}} \mid \mu=\mu_1\neq\mu_0\}=$$

$$P\left\{-u_{\frac{\alpha}{2}}\leqslant\frac{\overline{X}-\mu_0}{\frac{\sigma}{\sqrt{n}}}\leqslant u_{\frac{\alpha}{2}} \mid \mu=\mu_1\neq\mu_0\right\}=P\left\{\mu_0-u_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\leqslant\overline{X}\leqslant\mu_0+u_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \mid \mu=\mu_1\neq\mu_0\right\} \tag{2}$$

收稿日期:2015-01-19;**修回日期**:2015-03-25.
作者简介:高发玲(1982-)女,讲师,硕士,从事概率论与随机过程研究.

H_0 为假,即实际 $\mu=\mu_1\neq\mu_0$,即$\overline{X}\sim N(\mu_1,\sigma^2)$,则

$$\frac{\overline{X}-\mu_1}{\frac{\sigma}{\sqrt{n}}}\sim N(0,1) \tag{3}$$

因此

$$\beta=P\left\{(\mu_0-\mu_1)\frac{\sqrt{n}}{\sigma}-u_{\frac{\alpha}{2}}\leqslant\frac{\overline{X}-\mu_1}{\frac{\sigma}{\sqrt{n}}}\leqslant(\mu_0-\mu_1)\frac{\sqrt{n}}{\sigma}+u_{\frac{\alpha}{2}}\right\}=$$

$$\Phi\left((\mu_0-\mu_1)\frac{\sqrt{n}}{\sigma}+u_{\frac{\alpha}{2}}\right)-\Phi\left((\mu_0-\mu_1)\frac{\sqrt{n}}{\sigma}-u_{\frac{\alpha}{2}}\right) \tag{4}$$

① 当样本 n 固定,α 减小时,由于$\int_{u_{\frac{\alpha}{2}}}^{+\infty}\varphi(x)\mathrm{d}x=\frac{\alpha}{2}$,故$\frac{\alpha}{2}$减小,必然有 $u_{\frac{\alpha}{2}}$ 增大,从而 $\Phi\left((\mu_0-\mu_1)\frac{\sqrt{n}}{\sigma}+u_{\frac{\alpha}{2}}\right)$ 增大,而 $\Phi\left((\mu_0-\mu_1)\frac{\sqrt{n}}{\sigma}-u_{\frac{\alpha}{2}}\right)$ 减小,即 $\beta=\Phi\left((\mu_0-\mu_1)\frac{\sqrt{n}}{\sigma}+u_{\frac{\alpha}{2}}\right)-\Phi\left((\mu_0-\mu_1)\frac{\sqrt{n}}{\sigma}-u_{\frac{\alpha}{2}}\right)$ 增大.

② 当控制住 α 时,由于$\beta=\int_{(\mu_0-\mu_1)\frac{\sqrt{n}}{\sigma}-u_{\frac{\alpha}{2}}}^{(\mu_0-\mu_1)\frac{\sqrt{n}}{\sigma}+u_{\frac{\alpha}{2}}}\varphi(x)\mathrm{d}x$(其中$\varphi(x)$ 为标准正态分布的密度函数),积分区间是以$(\mu_0-\mu_1)\frac{\sqrt{n}}{\sigma}$为中心,以 $u_{\frac{\alpha}{2}}$ 为半径的邻域,当增大样本容量 n,α 固定不变时,半径 $u_{\frac{\alpha}{2}}$ 不变,而积分区间的中心$(\mu_0-\mu_1)\frac{\sqrt{n}}{\sigma}$向右移($\mu_0-\mu_1>0$时)或者左移($\mu_0-\mu_1<0$时).由于标准正态分布密度函数关于 y 轴对称,右侧单调递减,左侧单调递增,由此不管积分区间的中心$(\mu_0-\mu_1)\frac{\sqrt{n}}{\sigma}$右移还是左移,都会导致$\beta=\int_{(\mu_0-\mu_1)\frac{\sqrt{n}}{\sigma}-u_{\frac{\alpha}{2}}}^{(\mu_0-\mu_1)\frac{\sqrt{n}}{\sigma}+u_{\frac{\alpha}{2}}}\varphi(x)\mathrm{d}x$ 减小.

(2) 方差 σ^2 未知的情形.检验假设 $H_0:\mu=\mu_0$,备择假设 $H_1:\mu=\mu_1\neq\mu_0$.

当 H_0为真时,有

$$T=\frac{\overline{X}-\mu_0}{\frac{S}{\sqrt{n}}}\sim t(n-1) \tag{5}$$

由于 $P\left\{\left|\frac{\overline{X}-\mu_0}{\frac{S}{\sqrt{n}}}\right|\geqslant t_{\frac{\alpha}{2}}(n-1)\right\}=\alpha$,则犯第二类错误“取伪”的概率:

$$\beta=P\{\text{接受 }H_0\mid H_0\text{ 为假}\}=P\{|T|\leqslant t_{\frac{\alpha}{2}}(n-1)\mid\mu=\mu_1\neq\mu_0\}=$$

$$P\left\{-t_{\frac{\alpha}{2}}(n-1)\leqslant\frac{\overline{X}-\mu_0}{\frac{S}{\sqrt{n}}}\leqslant t_{\frac{\alpha}{2}}(n-1)\mid\mu=\mu_1\neq\mu_0\right\}=$$

$$P\left\{\mu_0-t_{\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}}\leqslant\overline{X}\leqslant\mu_0+t_{\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}}\mid\mu=\mu_1\neq\mu_0\right\} \tag{6}$$

H_0 为假,即实际

$$\mu=\mu_1\neq\mu_0,\frac{\overline{X}-\mu_1}{\frac{S}{\sqrt{n}}}\sim t(n-1) \tag{7}$$

所以

$$\beta=P\left\{(\mu_0-\mu_1)\frac{\sqrt{n}}{S}-t_{\frac{\alpha}{2}}(n-1)\leqslant\frac{\overline{X}-\mu_1}{\frac{S}{\sqrt{n}}}\leqslant(\mu_0-\mu_1)\frac{\sqrt{n}}{S}+t_{\frac{\alpha}{2}}(n-1)\right\} \tag{8}$$

① 当样本 n 固定,α 减小时,由于$\int_{t_{\frac{\alpha}{2}}(n-1)}^{+\infty}f(x)\mathrm{d}x=\frac{\alpha}{2}$,故$\frac{\alpha}{2}$减小,必然有 $t_{\frac{\alpha}{2}}(n-1)$ 增大,则 β 增大.

② 当控制住 α 时,由于$\beta=\int_{(\mu_0-\mu_1)\frac{\sqrt{n}}{S}-t_{\frac{\alpha}{2}}(n-1)}^{(\mu_0-\mu_1)\frac{\sqrt{n}}{S}+t_{\frac{\alpha}{2}}(n-1)}f(x)\mathrm{d}x$(其中 $f(x)$ 为 $t(n-1)$ 分布的密度函数),积分区间是以$(\mu_0-\mu_1)\frac{\sqrt{n}}{S}$为中心,以 $t_{\frac{\alpha}{2}}(n-1)$ 为半径的邻域,当增大样本容量 n,α 固定不变时,半径 $t_{\frac{\alpha}{2}}(n-1)$ 将变小,且积分区间的中心$(\mu_0-\mu_1)\frac{\sqrt{n}}{S}$向右移($\mu_0-\mu_1>0$时)或者左移($\mu_0-\mu_1<0$时).由于 t 分布密度函数关于 y 轴对称,右侧单调递减,左侧单调递增,由此不管积分区间的中心$(\mu_0-\mu_1)\frac{\sqrt{n}}{\sigma}$右移还是左移,都会导致$\beta=\int_{(\mu_0-\mu_1)\frac{\sqrt{n}}{S}-t_{\frac{\alpha}{2}}(n-1)}^{(\mu_0-\mu_1)\frac{\sqrt{n}}{S}+t_{\frac{\alpha}{2}}(n-1)}f(x)\mathrm{d}x$ 减小.

2　总体方差 σ^2 的假设检验两类错误概率关系分析

检验假设 $H_0:\sigma^2=\sigma_0^2$,备择假设 $H_1:\sigma^2={\sigma_1}^2\neq{\sigma_0}^2$.当 H_0为真时,$\chi^2=\frac{(n-1)S^2}{\sigma_0^2}\sim\chi^2(n-1)$[4],有

$$P\left\{\frac{(n-1)S^2}{\sigma_0^2}\geqslant\chi^2_{\frac{\alpha}{2}}(n-1)\right\}=\frac{\alpha}{2};P\left\{\frac{(n-1)S^2}{\sigma_0^2}\leqslant\chi^2_{1-\frac{\alpha}{2}}(n-1)\right\}=\frac{\alpha}{2} \tag{9}$$

$$P\{(n-1)S^2\geqslant\sigma_0^2\chi^2_{\frac{\alpha}{2}}(n-1)\}=\frac{\alpha}{2};P\{(n-1)S^2\leqslant\sigma_0^2\chi^2_{1-\frac{\alpha}{2}}(n-1)\}=\frac{\alpha}{2}$$

故犯第二类错误"取伪"的概率:

$$\beta=P\{\sigma_0^2\cdot\chi^2_{1-\frac{\alpha}{2}}(n-1)\leqslant(n-1)S^2\leqslant\sigma_0^2\cdot\chi^2_{\frac{\alpha}{2}}(n-1)\mid\sigma^2=\sigma_1^2\neq\sigma_0^2\}=$$

$$P\left\{\frac{\sigma_0^2}{\sigma_1^2}\cdot\chi^2_{1-\frac{\alpha}{2}}(n-1)\leqslant\frac{(n-1)S^2}{\sigma_1^2}\leqslant\frac{\sigma_0^2}{\sigma_1^2}\cdot\chi^2_{\frac{\alpha}{2}}(n-1)\right\} \tag{10}$$

① 当样本 n 固定,α 减小时,由于$\int_{\chi^2_{\frac{\alpha}{2}}(n-1)}^{+\infty}f(x)\mathrm{d}x=\frac{\alpha}{2}$,$\int_{\chi^2_{1-\frac{\alpha}{2}}(n-1)}^{+\infty}f(x)\mathrm{d}x=1-\frac{\alpha}{2}$,故$\frac{\alpha}{2}$减小,必然有$\chi^2_{\frac{\alpha}{2}}(n-1)$ 增大,$\chi^2_{1-\frac{\alpha}{2}}(n-1)$ 减小,则 β 增大.

② 当控制住α时,由于$\beta = \int_{\frac{\sigma_0^2}{\sigma_1^2}\cdot\chi^2_{1-\frac{\alpha}{2}}(n-1)}^{\frac{\sigma_0^2}{\sigma_1^2}\cdot\chi^2_{\frac{\alpha}{2}}(n-1)} f(x)\mathrm{d}x$(其中$f(x)$为$\chi^2(n-1)$分布的密度函数),积分区间$\left[\frac{\sigma_0^2}{\sigma_1^2}\cdot\chi^2_{1-\frac{\alpha}{2}}(n-1),\frac{\sigma_0^2}{\sigma_1^2}\cdot\chi^2_{\frac{\alpha}{2}}(n-1)\right]$,当增大样本容量$n$,$\alpha$不变时,$\frac{\sigma_0^2}{\sigma_1^2}$为正数,而随着$n$的增大,$\chi^2(n-1)$的密度函数图形越来越对称,从而积分区间的长度减小,而密度函数$f(x)$大于0,从而引起$\beta = \int_{\frac{\sigma_0^2}{\sigma_1^2}\cdot\chi^2_{1-\frac{\alpha}{2}}(n-1)}^{\frac{\sigma_0^2}{\sigma_1^2}\cdot\chi^2_{\frac{\alpha}{2}}(n-1)} f(x)\mathrm{d}x$减小.

3 结 论

对于参数假设检验的两类错误概率关系,在单正态总体均值和方差假设检验中来分析两类错误概率的具体变化关系.通过固定样本容量,犯第一类错误“弃真”的概率变大,犯第二类错误“取伪”的概率将会变小;当控制“弃真”概率时,增大样本容量会导致“取伪”概率减小.

参考文献:

[1] 吴赣昌.概率论与数理统计[M].4版.北京:中国人民大学出版社,2011

[2] 魏宗舒.概率论与数理统计教程[M].北京:高等教育出版社,2004

[3] 韩非,于云霞.检验假设的两类错误概率的具体形式[J].新乡师范高等专科学校学报,2007,21(2):18-19

[4] 盛骤,谢式千,潘承毅. 概率论与数理统计及其应用[M].北京:高等教育出版社,2004

Probability Relation Analysis of Two Types of Errors in Single Normal Population Hypothesis Test

GAO Fa-ling[1] ,LIU Chun-xia[2] ,SUN Jianying[3] ,GUO Ping[4]

(Qindao College,Qingdao Technological University,Qingdao 266100,China

Abstract: In view of two types of errors probability relationship in parameter hypothesis test, taking single normal population mean and variance hypothesis test for example, through the accepted domain and the rejection region by the test, probability relation of two types of errors in hypothesis test “abandoned the true” and “accept the bad” is analyzed. The method is an effective supplement for the principle of probability relationship of two types of errors under the condition of fixed sample capacity and controlling first type of error probability.

Keywords: hypothesis test; two types of errors; probability relation.