DOI :10.3969/J. ISSN. 1672-6693. 2011. 02. 008

变异系数的抽样分布及假设检验*

赵彦晖, 张水若, 邢瑞芳
(西安建筑科技大学 理学院, 西安 710055)

摘要 :变异系数是一项可靠性指标,它应用于既有结构的可靠性、医院统计及保险理论等方面,对变异系数进行假设检验具有现实意义。本文取自一般正态总体的子样,并利用样本均值 $\bar{X}$、样本标准差 S 和变异系数 $v=\frac{\sigma}{\mu}$ 的关系构造了一种含变异系数的抽样分布 $Z=v\frac{\bar{X}}{S}\sim z(v,n-1)$,同时给出了该分布的密度函数 $f_Z(z)=\frac{\left(\frac{n-1}{2}\right)^{\frac{n-1}{2}}}{\sqrt{\frac{\pi v^2}{2n}}\Gamma\left(\frac{n-1}{2}\right)}\int_0^{+\infty}y^{n-1}\mathrm{e}^{-\frac{1}{2}\left[(n-1)y^2+\frac{(yz-1)^2}{v^2/n}\right]}\mathrm{d}y$。本文给出一种对变异系数的小样本假设检验方法,根据该抽样分布及原假设成立的条件下确定检验统计量 $Z=v_0\frac{\bar{X}}{S}$,然后利用统计量构造一个使备择假设成立的小概率事件,由此得出拒绝域或拒绝条件。

关键词 :变异系数 ;抽样分布 ;假设检验
中图分类号 :O212 文献标识码 :A 文章编号 :1672-6693(2011)02-0040-03

变异系数是总体标准差与总体均值之比 $v=\frac{\sigma}{\mu}$,是衡量结构风险程度或稳定程度的可靠性指标,是一个无量纲量。在实际问题中有着广泛的应用,比如在既有结构的可靠性、医院统计以及保险理论等方面。本文建立了含变异系数的 z 分布,并利用该分布给出了一种变异系数的假设检验方法。

1 抽样分布的建立

引理 1 设总体 $X\sim N(\mu,\sigma^2)$,$X_1,X_2,\cdots,X_n$ 是来自总体的容量为 n 的样本,$\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ 为样本均值,$v=\frac{\sigma}{\mu}(\mu\neq 0)$,那么

$$U=\frac{\bar{X}}{\mu}\sim N(1,v^2/n)$$

概率密度为

$$f_U(x)=\frac{1}{\sqrt{2\pi}\sqrt{v^2/n}}\mathrm{e}^{-\frac{(x-1)^2}{2v^2/n}}$$ 证毕

证明 由文献[2-4]知

$$U=\frac{\bar{X}}{\mu}=\frac{1}{n\mu}\sum_{i=1}^{n}X_i$$

服从正态分布。又因 $EX_i=\mu$,$DX_i=\sigma^2(i=1,2,\cdots,n)$,所以

$$EU=\frac{1}{n\mu}\sum_{i=1}^{n}EX_i=\frac{1}{n\mu}\sum_{i=1}^{n}\mu=1$$

$$DU=\frac{1}{n^2\mu^2}\sum_{i=1}^{n}DX_i=\frac{1}{n^2\mu^2}\sum_{i=1}^{n}\sigma^2=\frac{\sigma^2}{n\mu^2}=\frac{v^2}{n}$$

故

$$U=\frac{\bar{X}}{\mu}\sim N(1,v^2/n)$$

U 的概率密度为

$$f_U(x)=\frac{1}{\sqrt{2\pi}\sqrt{v^2/n}}\mathrm{e}^{-\frac{(x-1)^2}{2v^2/n}}$$ 证毕

引理 2 设总体 $X\sim N(\mu,\sigma)^2$,$X_1,X_2,\cdots,X_n$ 是来自总体的容量为 n 的样本

$$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_I-\bar{X})^2$$

为样本方差,那么 $V=\frac{S}{\sigma}$ 的分布密度为

* 收稿日期 :2010-05-10 修回日期 :2010-07-14
资助项目 :国家自然科学基金(No. 50678143)
作者简介 :赵彦晖,男,教授,研究方向为概率统计计算与应用。

$$f_V(y)=\begin{cases}\dfrac{1}{2^{\frac{n-3}{2}}\Gamma\left(\dfrac{n-1}{2}\right)}(n-1)^{\frac{n-1}{2}}y^{n-2}e^{-\frac{(n-1)y^2}{2}}, & y>0\\ 0, & y\leqslant 0\end{cases}$$

证明 由于 $V=\dfrac{S}{\sigma}\geqslant 0$,所以当 $y\leqslant 0$ 时,V 的分布密度为 $f_V(y)=0$。

又由文献[2] 知 $\chi^2=\dfrac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$,所以当 $y>0$ 时,V 的分布函数为

$$F_V(y)=P\{V\leqslant y\}=$$

$$P\left\{\chi^2=\frac{(n-1)S^2}{\sigma^2}\leqslant(n-1)y^2\right\}=\int_0^{(n-1)y^2}f_{\chi^2}(x)\,dx$$

因此 V 的分布密度为

$$f_V(y)=F'_V(y)=2(n-1)yf_{\chi^2}((n-1)y^2)=$$

$$2(n-1)y\frac{1}{2^{\frac{n-1}{2}}\Gamma\left(\frac{n-1}{2}\right)}[(n-1)y^2]^{\frac{n-1}{2}-1}e^{-\frac{(n-1)y^2}{2}}=$$

$$\frac{1}{2^{\frac{n-3}{2}}\Gamma\left(\frac{n-1}{2}\right)}(n-1)^{\frac{n-1}{2}}y^{n-2}e^{-\frac{(n-1)y^2}{2}}$$

从而 $V=\dfrac{S}{\sigma}$ 的分布密度为

$$f_V(y)=\begin{cases}\dfrac{1}{2^{\frac{n-3}{2}}\Gamma\left(\dfrac{n-1}{2}\right)}(n-1)^{\frac{n-1}{2}}y^{n-2}e^{-\frac{(n-1)y^2}{2}}, & y>0\\ 0, & y\leqslant 0\end{cases}$$

证毕

定理1 设总体 $X\sim N(\mu,\sigma^2)$,$X_1,X_2,\cdots,X_n$ 是来自总体的容量为 $n(n>1)$ 的样本。

$$\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i,\quad S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$$

分别为样本均值和样本方差,$S=\sqrt{S^2}$ 为样本标准差,$v=\dfrac{\sigma}{\mu}$ 为变异系数,那么随机变量

$$Z=v\frac{\bar{X}}{S}$$

仍为连续型分布,其密度为

$$f_Z(z)=\frac{\left(\frac{n-1}{2}\right)^{\frac{n-1}{2}}}{\sqrt{\frac{\pi v^2}{2n}}\Gamma\left(\frac{n-1}{2}\right)}\int_0^{+\infty}y^{n-1}e^{-\frac{1}{2}\left[(n-1)y^2+\frac{(yz-1)^2}{v^2/n}\right]}dy$$

证明 按引理1和引理2的记号,因为 $\bar{X}$ 与 S 相互独立[2-4],所以 $U=\dfrac{\bar{X}}{\mu}$ 与 $V=\dfrac{S}{\sigma}$ 也相互独立,又因为 $V\geqslant 0$ 且当 $y\leqslant 0$ 时 $f_V(y)=0$,所以随机变量

$$Z=v\frac{\bar{X}}{S}=\frac{\bar{X}/\mu}{S/\sigma}=\frac{U}{V}$$

其分布函数为

$$F_Z(z)=P\{Z\leqslant z\}=P\{U\leqslant Vz\}=$$

$$\int_0^{+\infty}\left[\int_{-\infty}^{yz}f_U(x)f_V(y)\,dx\right]dy$$

Z 的分布密度为

$$f_Z(z)=F'_Z(z)=\int_0^{+\infty}yf_U(yz)f_V(y)\,dy=$$

$$y\frac{1}{\sqrt{2\pi}\sqrt{v^2/n}}e^{-\frac{(yz-1)^2}{2v^2/n}}$$

$$\int_0^{+\infty}\frac{1}{2^{\frac{n-3}{2}}\Gamma\left(\frac{n-1}{2}\right)}(n-1)^{\frac{n-1}{2}}y^{n-2}e^{-\frac{(n-1)y^2}{2}}dy=$$

$$\frac{\left(\frac{n-1}{2}\right)^{\frac{n-1}{2}}}{\sqrt{\frac{\pi v^2}{2n}}\Gamma\left(\frac{n-1}{2}\right)}\int_0^{+\infty}y^{n-1}e^{-\frac{1}{2}\left[(n-1)y^2+\frac{(yz-1)^2}{v^2/n}\right]}dy$$

为了方便起见,笔者把定理1给出的分布称为 z 分布,记作 $Z=v\dfrac{\bar{X}}{S}\sim z(v,n-1)$,其中 $v=\dfrac{\sigma}{\mu}$ 为变异系数。 证毕

2 变异系数 v 的假设检验

1)为了对变异系数 v 进行假设检验,笔者提出检验假设

$$H_0: v=v_0,\ H_1: v\neq v_0$$

由定理1知 $Z=v_0\dfrac{\bar{X}}{S}$ 为统计量,$Z\sim z(v,n-1)$,故对显著性水平 $\alpha(\alpha=0.1,0.05,0.01$ 等),可构造小概率事件

$$P\{Z\leqslant z_{1-\alpha/2}(v_0,n-1)\ 或\ Z\geqslant z_{\alpha/2}(v_0,n-1)\}=\alpha$$

其中 $z_\alpha(v_0,n-1)$ 为 $z(v_0,n-1)$ 分布的上侧分位数。于是得拒绝域(或拒绝条件)

$$Z\leqslant z_{1-\alpha/2}(v_0,n-1)\ 或\ Z\geqslant z_{\alpha/2}(v_0,n-1)$$

2)如果笔者提出检验假设

$$H_0: v=v_0,\ H_1: v>v_0$$

则由 $Z=v_0\dfrac{\bar{X}}{S}\sim z(v,n-1)$,可构造小概率事件

$$P\{Z\leqslant z_{1-\alpha}(v_0,n-1)\}=\alpha$$

于是得拒绝域(或拒绝条件)

$$Z\leqslant z_{1-\alpha}(v_0,n-1)$$

3)如果笔者提出检验假设

$$H_0: v=v_0,\ H_1: v<v_0$$

则由 $Z=v_0\dfrac{\bar{X}}{S}\sim z(v,n-1)$,可构造小概率事件

$$P\{Z \geqslant z_{\alpha}(v_0, n-1)\} = \alpha$$

于是得拒绝域(或拒绝条件)

$$Z \geqslant z_{\alpha}(v_0, n-1)$$

总之,根据样本观测值计算检验统计量观测值,由上述拒绝域(或拒绝条件)来确定是否拒绝原假设以此来检验结论。

参考文献:

[1] 周昌隆. 变异系数的抽样分布[J]. 赣南师范学院学报,1996,6(1):19-21.

[2] 赵彦晖,杨金林. 概率统计[M]. 北京:科学出版社,2006:108.

[3] 王文森. 变异系数—一个衡量离散程度简单而有用的统计指标[J]. 中国统计,2007,6:41-42.

[4] E-勒克斯(美). 概率论与数理统计(引论)[M]. 北京:人民教育出版社,1999:43.

[5] 孙祝岭. 正态分布变差系数的抽样检验[J]. 数理统计与应用概率,1997,9(12):271-273.

[6] 杨玉清,张宏. 变异系数差异的显著性检验[J]. 东北农业大学学报,1994,25(1):27-31.

[7] 谢峰,周飞跃. 变异系数的假设检验[J]. 工业技术经济,1999,18(4):82,87.

[8] 付靖. 变异系数在既有结构可靠性检验中的应用研究初探[D]. 西安:西安建筑科技大学,2009.

The Sample Distribution and Hypothesis Test of Coefficient of Variation

ZHAO Yan-hui, ZHANG Shui-ruo, XING Rui-fang

(College of Science, Xi'an University of Architecture & Technology, Xi'an 710055, China)

Abstract: Coefficient of variation is a reliability index. It is widely used in evaluation of reliability of existing structure, hospital statistics, insurance theory and so on. Hypothesis testing of the coefficient of variation is realistic significance. The sample of normal distribution is extracted, according to the relationship between sample mean $\bar{X}$, sample standard deviation S and coefficient of variation $v = \frac{\sigma}{\mu}$, the sample distribution of containing coefficient of variation $Z = v\frac{\bar{X}}{S} \sim z(v, n-1)$ was given in this paper, at the same time its density function $f_Z(z) = \dfrac{\left(\frac{n-1}{2}\right)^{\frac{n-1}{2}}}{\sqrt{\frac{\pi v^2}{2n}}\Gamma\left(\frac{n-1}{2}\right)}\int_0^{+\infty} y^{n-1} e^{-\frac{1}{2}\left[(n-1)y^2 + \frac{(yz-1)^2}{v^2/n}\right]} dy$ is given. A way of hypothesis test of the coefficient of variation is established in small sample. Through using the sample distribution and null hypothesis, test statistics $Z = v_0\frac{\bar{X}}{S}$ is determined. Then a small probability events what makes alternative hypothes right is constructed, using this statistics. Thus rejection regions or rejection conditions are given.

Key words: coefficient of variation; sample distribution; hypothesis test

(责任编辑 游中胜)