Advances in Applied Mathematics 应用数学进展, 2020, 9(12), 2329-2332
Published Online December 2020 in Hans. http://www.hanspub.org/journal/aam
https://doi.org/10.12677/aam.2020.912272

基于幂函数与指数函数组合的洛伦兹曲线模型

丁善华，蒙　璐

贵州民族大学数据科学与信息工程学院，贵州　贵阳
Email: 867188984@qq.com, 2778651598@qq.com

收稿日期：2020年11月27日；录用日期：2020年12月21日；发布日期：2020年12月31日

摘　要

构建一类新的拟合效果更好的洛伦兹曲线模型，并针对1977年美国收入分配分组数据进行数值实验。实验结果表明，该类模型拟合效果较好。

关键词

洛伦兹曲线模型，幂函数，指数函数，数值实验

Lorenz Curve Model Based on Exponent Combination of Power Function and Exponential Function

Shanhua Ding, Lu Meng

School of Data Science and Information Engineering, Guizhou Minzu University, Guiyang Guizhou
Email: 867188984@qq.com, 2778651598@qq.com

Received: Nov. 27th, 2020; accepted: Dec. 21st, 2020; published: Dec. 31st, 2020

Abstract

A new kind of Lorenz curve model with better fitting effect is constructed, and numerical experiments are carried out on the grouped data of American income distribution in 1977. The experimental results show that this kind of model has better fitting effect.

Keywords

Lorentz Curve Model, Power Function, Exponential Function, Numerical Experiments

文章引用: 丁善华, 蒙璐. 基于幂函数与指数函数组合的洛伦兹曲线模型[J]. 应用数学进展, 2020, 9(12): 2329-2332.
DOI: 10.12677/aam.2020.912272

1. 引言

洛伦兹曲线作为刻画社会收入分配的不平等程度或贫富差距程度的一个有效工具，洛伦兹曲线得到了广泛而深入的研究。到目前为止，已经发现了几十种洛伦兹曲线模型用以拟合分组形式的收入分配数据。文献[1]列出了直到迄今为止的数理经济理论界发现的所有洛伦兹曲线模型，并进行了比较。随后不少人[2] [3] [4] [5]陆续提出了若干新模型。已有的这些模型存在诸多问题：例如拟合效果欠佳，不能提供整个[0, 1]区间上的良好逼近(例如见[6] [7])，有些模型甚至不满足洛伦兹曲线的条件(见[1] [7])。数理经济学界仍在积极寻找拟合效果良好的模型，任何新模型对收入分配分析都具有重要意义。

设收入分配的概率密度函数为 $f(x)$，对应的概率分布函数为 $F(x)$，则 $p=F(x)$ 表示收入低于或等于 x 的人口比例。记收入低于或等于 x 的人口群体拥有收入占总收入的比例为 $L(p)$，则 $L(p)=\frac{1}{\mu}\int_0^x tf(t)\mathrm{d}t p=F(x)$，其中 μ 为平均收入。记 $F(x)$ 的反函数为，则 $F^{-1}(p)$ $L(p)=\frac{1}{\mu}\int_0^p F^{-1}(q)\mathrm{d}q$。因此，只要知道收入分配的统计分布，就可以得到相应的洛伦兹曲线。可以通过入户调查来获取家庭收入和消费等数据，用 p_i 表示的是低收入群体的累计人口比例，L_i 表示的是该群体拥有的总收入比例，则其形式为 $\{(p_i,L_i)\}_{i=1}^n$。用最小二乘法对数据进行拟合，通过拟合来确定 $L(p,\tau)$ 参数向量 τ 的估值 $\hat{\tau}$。最后使用 $L(p,\hat{\tau})=\hat{L}(p)$ 来近似描述收入分配情况。$L(p,\tau)$ 是定义在区间 $[0,1]$ 上取值于区间 $[0,1]$ 的函数，并且满足

$$L(0,\tau)=0,\ L(1,\tau)=1,\ L'(p,\tau)\geq 0,\ L''(p,\tau)\geq 0. \tag{1}$$

即 $L(p,\tau)$ 是 $[0,1]$ 上关于 p 的凸增函数。

2. 构建洛伦兹曲线模型

文献[8]提出了形如 $G(p)=pe^{a(1-p)},a<0$ 的指数洛伦兹曲线模型。容易验证，当 $a>\frac{1}{4}$ 时，$G(p)$ 为 $[0,1]$ 上的凹函数，不满足洛伦兹曲线条件。最近，[9] [10] [11]提出了 $H(p)=f(p)g(p)$ 形式的凹凸组合的洛伦兹曲线模型，其中 $f(p)$ 为凹函数，$g(p)$ 为凸函数。因此，一个自然的想法能否将 $G(p)=pe^{a(1-p)},\left(a>\frac{1}{4}\right)$ 与某些凸函数进行组合。得到拟合效果更好的一类新的洛伦兹曲线模型。为此，我们考虑将其与幂函数组合，得到 $H(p)=p^\alpha e^{a(1-p)}$ 形式的洛伦兹曲线模型。

定理 1 当 $\frac{1}{4}\leq a\leq 1$，$\alpha\geq\frac{2a+1+\sqrt{4a^2+1}}{2}$ 时，$H(p)$ 满足洛伦兹曲线条件。

证明 为了证明 $H(p)$ 满足洛伦兹曲线条件，需验证 $H(p)$ 满足(1)式。显然，$H(0)=0$，$H(1)=1$。

$$\begin{aligned}H'(p)&=\alpha p^{\alpha-1}e^{a(1-p)}+p^\alpha(-a)e^{a(1-p)}\\&=\left(\alpha p^{\alpha-1}-ap^\alpha\right)e^{a(1-p)}\\&=e^{a(1-p)}p^{\alpha-1}(\alpha-ap).\end{aligned}$$

当 $\frac{1}{4}\leq a\leq 1$，$p\in[0,1]$，$\alpha\geq\frac{2a+1+\sqrt{4a^2+1}}{2}$ 时，易知 $H'(p)\geq 0.$

$$\begin{aligned}H''(p)&=(-a)\mathrm{e}^{a(1-p)}p^{\alpha-1}(\alpha-ap)+\mathrm{e}^{a(1-p)}(\alpha-1)p^{(\alpha-2)}(\alpha-ap)+\mathrm{e}^{a(1-p)}p^{(\alpha-1)}(\alpha-a)\\&=\mathrm{e}^{a(1-p)}p^{\alpha-2}\left(-ap(\alpha-ap)+(\alpha-1)(\alpha-ap)+p(\alpha-a)\right)\\&=\mathrm{e}^{a(1-p)}p^{\alpha-2}\left(-ap\alpha+a^2p^2+\alpha^2-\alpha ap-\alpha+ap+\alpha p-ap\right)\\&=\mathrm{e}^{a(1-p)}p^{\alpha-2}\left(-2ap\alpha+a^2p^2+\alpha^2-\alpha+\alpha p\right)\\&=\mathrm{e}^{a(1-p)}p^{\alpha-2}\left(a^2-(2ap+1)\alpha+a^2p^2+\alpha p\right).\end{aligned}$$

因为 $p\in[0,1]$，$\frac{1}{4}\le a\le 1$时，所以当$\alpha\ge\frac{2a+1+\sqrt{4a^2+1}}{2}$时，$H''(p)\ge 0$，定理证毕。

3. 数值实验

利用 Basmann (1990)实例分析中所给出的 1977 年美国 100 个分位点的详细收入分配分组数据对周递芝(2018)中 $J(p)=p^{\alpha}\sin\left(\frac{\pi}{2}p\right),\left(\alpha>1+\frac{\pi}{2}\right)$ 以及我们的 $H(p)=p^{\alpha}\mathrm{e}^{a(1-p)},\left(\alpha>\frac{2a+1+\sqrt{4a^2+1}}{2},\frac{1}{4}\le a\le 1\right)$ 进行数值实验，并比较。定义残差平方和为 $SSR=\sum_{i=1}^{n}\left[L(p,\hat{\tau})-L_i\right]^2$，最大绝对误差 $MAS=\max_{1\le i\le n}\left|L(p,\hat{\tau})-L_i\right|$，基尼系数 $G=1-2\int_0^1 L(p)\mathrm{d}p$。借助 Lingo11 和 matlab 得出的实验结果见表 1。

Table 1. Test results of mean square error of income group data of the United States in 1977
表 1. 1977 年美国收入分组数据均方误差检验结果

	SSR	MAS	G
$J(p)$	0.1191427E-02	0.8428122	0.339733
$H(P)$	0.8449939E-03	0.0530572	0.378377

通过数值实验结果的比对，$H(p)$的残差平方和和最大绝对误差都是最小，这表明了我们构建的新模型拟合效果有了很大的改善，该模型的拟合精度更高，适应性更强。因此，我们构造的洛伦兹曲线模型是对既有洛伦兹曲线模型的有效拓展。

致　谢

作者对同行评阅人的意见和建议表示深深感谢。

参考文献

[1] Schader, M. and Schmid, F. (1994) Fitting Parametric Lorenz Curves to Grouped Income Distributions a Critical Note. *Empirical Economics*, **19**, 361-370. https://doi.org/10.1007/BF01205943

[2] Ogwang, T. and Rao, U.L.G. (1996) A New Functional Form for Approximating the Lorenz Curve. *Economics Letters*, **52**, 21-29. https://doi.org/10.1016/0165-1765(96)00833-6

[3] Ryu, H.K. and Slottje, D.J. (1996) Two Flexible Functional Form Approaches for Approximating the Lorenz Curve. *Journal of Econometrics*, **72**, 251-274. https://doi.org/10.1016/0304-4076(94)01722-0

[4] Sarabia, J.M. (1997) A Hierarchy of Lorenz Curves Based on the Generalized Tukey's Lambda Distribution. *Econometric Reviews*, **16**, 305-320. https://doi.org/10.1080/07474939708800389

[5] Sarabia, J.M., Gastillo, E. and Pascual, M. (2005) Mixture Lorenz Curves. *Economics Letters*, **89**, 89-94. https://doi.org/10.1016/j.econlet.2005.05.015

[6] Ogwang, T. and Rao, U.L.G. (2000) Hybrid models of the Lorenz Curve. *Economics Letters*, **69**, 39-44. https://doi.org/10.1016/S0165-1765(00)00274-3

[7] Sarabiaa, J.M., Castillob, E. and Slottjec, D.J. (1999) An Ordered Family of Lorenz Curves. *Journal of Econometrics*, **91**, 43-60. https://doi.org/10.1016/S0304-4076(98)00048-7

[8] Basmann, R.L., Hayes, K.J. and Slottje, D.J. (1990) A General Functional Form for Approximating the Lorenz Curve. *Journal of Econometrics*, **43**, 77-90. https://doi.org/10.1016/0304-4076(90)90108-6

[9] 周递芝. 基于三角函数组合的洛伦兹曲线模型[J]. 经济研究导刊, 2018, 355(5): 13-14.

[10] 邢丽, 储昌木, 孙娇娇. 基于对数组合的一类洛伦兹曲线模型[J]. 数学的实践与认识, 2016(5): 204-209.

[11] 储昌木, 邢丽, 田应富. 基于凹凸组合的一类洛伦兹曲线模型[J]. 经济数学, 2015(2): 97-100.