

非参数核密度法厘定玉米区域产量保险费率研究 ——以河北安国市为例

王丽红 杨沛华 田志宏 闫仲勇
(中国农业大学 经济管理学院, 北京 100094)

摘要 本研究以非参数核密度法为核心,构建了农作物 GRP 保险费率厘定的方法体系,以此对河北省安国市 1980—2004 年的玉米产量风险损失进行估算和纯保险费率进行厘定,得出在 80% 的保障水平下,纯保险费率为 3.4%;在 70% 的保障水平下,纯保险费率为 2.9%。这一结果比较接近该区域传统的风险评估方法所得结果。该方法的运用旨在为区域农业保险费率的厘定提供了一种新的可供选择的方法途径。

关键词 玉米区域产量保险;非参数核密度模型;产量损失概率;费率厘定

中图分类号 F 840.66

文章编号 1007-4333(2007)01-0090-05

文献标识码 A

Maize GRP rate of premium deciding by nonparametric kernel density: ——A case study on Anguo city, Hebei province

Wang Lihong, Yang Ruihua, Tian Zhihong, Yan Zhongyong

(College of Economics and Management, China Agricultural University, Beijing 100094, China)

Abstract We used nonparametric kernel density technique as a core in developing an evaluation system for determining the GRP rate of premium based on our analysis of maize production in Anguo city, Hebei province during 1980—2004. The study shows that at 80% guarantee level, the pure rate of premium was 3.4%; and at 70% guarantee level, the pure rate of premium was 2.9%. These results were almost the same as those evaluated by the conventional method, which proves the feasibility of using nonparametric kernel density technique for risk evaluation, and provides a new practical way to determine crop GRP rate of premium. The result can serve as a guide to China agricultural insurance tests, and agricultural insurance subsidy policy.

Key words maize GRP; nonparametric kernel density model; production loss probability; premium rate deciding

农作物区域产量保险 (GRP) 是针对系统风险的保险,满足保险的风险一致性原则,可以有效地避免保险中出现的逆向选择和道德风险问题,美国和加拿大经过长期试验,均认同这种保险计划。我国地域辽阔,农业生产区域气候差异大,自然灾害的发生率和灾害损失率在时空上分布很不均衡,按照农业风险地域分异规律,实施区域农作物保险更切合实际。而在以往我国农作物保险的实践中,主要采取单一费率的形式,这一做法造成保户的保费负担与承受的风险特征不一致,诱发了逆向选择和道德风险的问题^[1],是造成我国农业保险市场失灵的主要因素之一。重新启动国内农业保险市场的一个现实

的迫切的问题,从技术层面上来看,就是实行农业保险区域化保费细分。从信息层面来看,就是利用现有的数据资料,科学地厘定或精算不同区域的保险费率。

对于纯保险费率的厘定,传统上运用的方法是参数估计法,主要为正态分布法(还有一些较为复杂但适用性较强的分布如广义 Beta 分布等),另外还有实际生产历史法(actual product history, APH)和经验费率法。这些方法的基本原理是通过确定区域农作物产量单产的分布模式及产量相应的保障水平,来确定农作物产量的期望损失,以期望损失占产量保障水平的百分比作为理论纯费率。同时对于农

收稿日期: 2006-05-18

作者简介: 王丽红,博士研究生, E-mail: wlhcau@163.com; 杨沛华,副教授,通讯作者,主要从事农业经济、经济计量研究, E-mail: cemyrh@163.com

作物产量的长期变动趋势等因素也加以考虑,利用计量经济模型进行相应的趋势调整,使得费率的厘定更切合实际,体现农业风险发生的系统性和区域风险的差异性的特点。但是,参数估计法要求有先验的分布函数、样本容量足够大并已知关于总体分布的有关信息,而实际应用中往往很难获得总体分布信息,选择先验的分布函数的过程也比较繁琐,样本较小时的估算结果也不稳定。相比之下,非参数估计法不限制样本母体的分布形式,是非参数密度估计中有关单样本模型较为典型的估计方法,估计结果在小样本中相对稳定,能够更好地体现区域风险的差异性特点^[2,3]。

农业保险费率厘定研究中关于非参数方法的运用主要集中于核密度估计法。Turvey 等^[1]在 1993 曾采用非参数估计量对农作物产量分布的保险费率进行了估计,由于采用的样本过小,限制了核密度估计的效果。Barry K. G. 等在 1998 年采用非参数核密度估计农作物产量密度并用以确定了农作物产量的保险费率^[2],随后对该方法进行了重新的探讨,在 2000 年提出了适应性核密度算法^[4],优化了估计的效果。刘长标^[5]将非参数密度估计的信息扩散法尝试用于农作物区域产量风险的评估。谭英平^[6]则探讨了非参数核密度估计的方法中带宽(组间参数)的确定方法。钟甫宁等^[7]采用正态分布函数作为核函数对各地区农作物受灾率进行了估算。这些研究不同侧重地探讨了非参数核密度方法在农业保险中的运用。总的来看,由于人们对非参数核密度法还没有足够的认识,该方法在农业区域保险费率厘定方面的应用和探索性研究还比较少。基于此,本文针对推行区域农业保险的需要,整合了目前非参数核密度法的研究成果,结合自身数据处理的经验,系统地提出了一个非参数核密度估计基础上的风险评估和保险费率厘定方法体系,并且通过河北安国市玉米保险费率的厘定过程和结果,说明在现有的数据信息条件下该方法的适用性,旨在为农作物保险费率的厘定提供一个可供选择的方法途径。

1 区域产量损失风险评估及费率厘定的非参数方法

1.1 产量损失风险估算的核密度模型

准确估算农作物产量损失发生的概率关系到保

险概率厘定的科学性,因此估算农作物产量损失发生概率方法的选择非常重要。核函数是一个权重函数,它描述了数据的分布情况,本文采用非参数核密度的信息扩散模型进行估算。

设农作物产量损失率为 l ,显然 $l \in [0, 1]$,某地区第 t 年的生产损失率的样本观测数据为 $x_t, t = 1, 2, \dots, n$ 。设的信息按正态分布规律扩散给样本空间 $[0, 1]$ 中的每个样本点 l ,信息扩散模型的形式为

$$g_{xt}(l) = \frac{1}{h\sqrt{2\pi}} \exp\left[-\frac{(x_t - l)^2}{2h^2}\right] \quad (1)$$

式中: h 为信息扩散系数,即组间宽度参数(bandwidth),这个参数决定了密度函数的平滑程度, h 越大密度越平滑,选择一个恰当的组间宽度参数是非参数核密度估计的重要步骤。

选择组间宽度参数的方法有总量确定法(cross-validation)、Silverman 的“经验法则”(rule-of-thumb)和插入法(plug-in)^[2]。本研究拟采用 Silverman 的“经验法则”,最优的 h 参数为

$$h = \left[\int_0^1 k(t)^2 dt \right]^{1/5} \left[\int_0^1 \phi(y)^2 dy \right]^{-1/5} n^{-1/5} \quad (2)$$

式中: $k(t) = \int_0^1 k(t)^2 dt$, ϕ 为待估的密度, $\phi(y) = \frac{\partial^2 \phi(y)}{\partial y^2}$,如果未知密度呈方差为 σ^2 的正态分布,则以正态分布函数作为核函数,根据 Barry G K 和 Alan P K 的工作^[4],式(2)简化为

$$h_{opt} = 1.06 \cdot \sigma \cdot n^{-1/5} \quad (3)$$

在实际工作中,由于数据往往偏离正态分布,将范围因子从 1.06 降到 0.9 效果较好,这已得到证明^[4]。

1.2 纯费率的厘定过程

本研究确定区域农作物产量保险费率的估算过程为:1)相对波动产量的分解;2)农作物产量损失概率的估算;3)保险费率的厘定。其中一些环节有较多可替代方法供选择,总结见表 1。

1.1 节中已经详细地讨论了产量损失概率的估算,下面具体分析相对波动产量分解和保险费率的厘定。

1)相对波动产量的分解。在此将影响作物最终产量形成的自然和非自然因素划分为农业技术措施、气象条件和随机误差 3 部分。具体表达为

$$y_t = \hat{y}_t + y_{wt} + \epsilon_t \quad (4)$$

表1 纯费率的厘定过程

Table 1 Process of deciding pure agricultural insurance rate

步骤	模型	关键问题	备选方法
相对波动产量分解	$y_t = \hat{y}_t + y_{wt} + \varepsilon_t$ $y_{wt} = y_t - \hat{y}_t; y_{rt} = y_{wt} / \hat{y}_t$	\hat{y}_t 的估算	移动平均法 指数平滑法 时间变量回归
产量损失概率估算	$g_{xt}(l) = \frac{1}{h\sqrt{2}} \exp\left[-\frac{(x_t - l)^2}{2h^2}\right]$	h 的确定	总量确定法 插入法 经验法则法 正态分布法 非参数估计法
保险费率厘定	$r = E(y_t) / \hat{y}_t$	$E(y_t)$	

Shows the calculating system by which we try to decide the rate of GRP agricultural insurance.

其中, y_t 为第 t 年的作物单产, \hat{y}_t 为由技术水平决定的趋势产量; y_{wt} 为以气象因素为主的短周期变化因子影响的产量; ε_t 为随机误差项, 主要包括受病虫害、社会因素等随机因素影响的产量, ε_t 一般比较小, 如果忽略 ε_t 的影响, 那么

$$y_{wt} = y_t - \hat{y}_t \quad (5)$$

得到 y_{wt} 后, 为了消除地区间生产水平差异的影响, 采用波动产量除以趋势产量的方法得到相对波动产量 y_{rt} 。本文采用时间变量回归法对 \hat{y}_t 进行估计。

2) 费率的厘定。采用 Alan P K 和 Barry K G^[2] 提出的非参数保险费率厘定方法。具体形式如下

$$E(y_t) = p(y_t < \hat{y}_t) [\hat{y}_t - E(y_t | y_t < \hat{y}_t)] \quad (6)$$

$$r = E(y_t) / \hat{y}_t \quad (7)$$

其中, $E(\cdot)$ 为实际产量的期望; $p(\cdot)$ 为第 t 年的实际产量小于保障水平下的产量的概率; \hat{y}_t 为保障水平的百分率; r 为区域产量保险费率。

2 河北省安国市玉米区域产量保险费率厘定分析

基于河北省安国市 1980—2004 年的玉米播种面积和玉米总产量资料, 对该市玉米生产的纯保险费率进行测算。河北省是农业生产大省, 具有大陆性季风气候, 降水量时空分布不均匀, 自然灾害如旱、涝、风、雹等分布的区域差异较大, 为实施区域农作物保险提供了良好的基础。

2.1 安国市玉米相对波动产量的求取

利用安国市 1980—2004 年玉米总产量和玉米播种面积计算出各年单位面积产量, 形成玉米单产时间序列, 对玉米单产序列与时间, 得到玉米的趋势产量

$$\hat{y}_t = 2112.7971 + 441.3796t - 9.4382t^2$$

该模型 F 统计量为 59.65, t 、 F 检验均通过了 5% 的显著性水平检验。表 2 给出了 1980—2004 年安国市玉米相对波动产量 y_{rt} 。

表2 1980—2004年安国市玉米相对波动产量

Table 2 Variations in relative maize yield in Anguo city between 1980—2004

年份	1980	1981	1982	1983	1984	1985	1986	1987	1988
相对波动产量	42.22	2.17	6.02	-4.02	-10.87	-9.79	-1.99	-5.09	-6.45
年份	1989	1990	1991	1992	1993	1994	1995	1996	1997
相对波动产量	-10.25	-13.86	-3.34	-3.78	-4.86	5.16	11.67	12.4	10.67
年份	1998	1999	2000	2001	2002	2003	2004		
相对波动产量	10.94	9.80	5.94	2.35	-26.52	-2.83	-2.24		

Gives the estimated yield wave data in order to evaluate the maize production risk.

2.2 玉米生产风险的测算

采用玉米产量灾害损失样本, 估算安国市玉米

生产损失风险。为了计算的可行, 笔者将连续的信息扩散模型(1)通过如下的方式进行离散化。

1) 连续模型的离散化

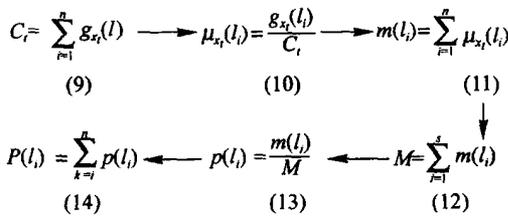
设已知灾害损失样本 $\{x_t, t = 1, 2, \dots, n\}$, 将生产损失 l 的样本空间 $[0, 1]$ 分成 s 等份。则原连续型风险分布概率密度函数变为:

$$g_{x_i}(l_i) = \frac{1}{h \sqrt{2}} \exp \left[- \frac{(x_t - l_i)^2}{2h^2} \right] \quad (8)$$

其中, $t = 1, 2, \dots, n; i = 1/s, 2/s, \dots, 1$ 。

用式(8)、(9)和(10), 计算出 n 年产量损失率的归一化信息分布, 对归一化的信息分布进行处理, 得到关于农作物区域产量风险损失率 $m(l_i)$, 它代表所考察的区域产量灾害损失样本经过信息扩散后

取值为 l 的样本个数。令 $M = \sum_{i=1}^s m(l_i)$, M 为 n 年内作物区域产量损失的样本空间上所有样本个数的总和。显然, 在理论上应当有 $M = n$, 用式(12)和(13)得到式(14), 其中 $P(l_i)$ 为损失样本 $\{x_t\}$ 经过信息扩散后得到的损失率超过 l 的概率估计值, 即为所求的风险估计值。具体过程见图 1。



The figure shows how to estimate step by step the probability of cost rate which is larger than l .

图 1 $P(l_i)$ 的求解流程

Fig. 1 Calculation flow of $P(l_i)$

2) 玉米生产风险估算结果

利用 1980—2004 年安国市玉米产量灾害损失样本数据(表 2), 对相对波动产量序列进行平稳性检验, 结果见表 3。有(AC)和偏自相关系数(PAC)很小且趋于 0, 并且 Q 统计量(Q-Stat)的伴随概率(Prob)均大于 0.05; 单位根检验的结果是, 在 5% 的

表 3 相对波动产量数列水平自相关检验

Table 3 Auto-regression tests of changes in relative yield series

滞后期	自相关系数	偏自相关系数	Q 统计量	伴随概率
1	0.32	0.32	2.80	0.09
2	0.23	0.15	4.38	0.11
3	0.02	-0.10	4.39	0.22
4	-0.12	-0.15	4.83	0.31

Shows that the relative wave yield series is stable.

显著性水平下, ADF 检验的值为 -3.0, 因此序列是平稳的。

取 $s = 20$, 运用 Silverman 的“经验法则”做出玉米产量相对波动分布形态的核密度拟合图(图 2), 得到区域产量风险的最优带宽。

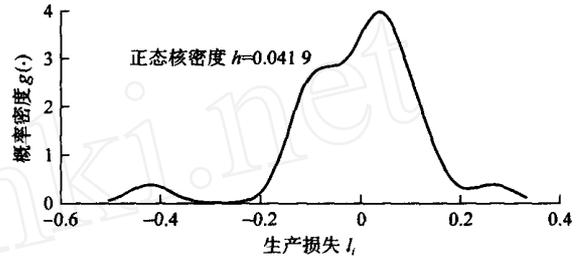


图 2 非参数核密度函数法拟合安国市玉米波动

Fig. 2 Variations in maize yield as estimated by nonparametric kernel density technique

将 h 等参数值代入式(8), 根据以上讨论的 $P(l_i)$ 求解流程, 得到安国市 1980—2004 年玉米产量损失在大于 20% 的概率为 0.85、大于 30% 的概率为 0.75, 即玉米产量损失小于 20% 的概率为 0.15, 小于 30% 的概率为 0.25。根据郭迎春等^[3]对河北气候风险分区的研究成果, 安国市的玉米生产风险介于 2 级(1%~2.5%)和 3 级(2.6%~5%)之间。这说明非参数核密度法厘定区域农作物产量纯保险费率是可行的。

2.3 安国市玉米区域产量纯保险费率的厘定

运用式(6)和(7), 对安国市玉米区域产量保险费率(r)进行厘定。结果见表 4。

表 4 1980—2004 年安国市玉米区域产量纯保险费率

Table 4 Maize GRP agricultural insurance rate in Anguo city 1980—2004

风险水平指标	80%保障	70%保障
	水平下	水平下
损失概率/ %	0.15	0.25
保障水平下的期望产量/ (kg/hm ²)	4 611.90	4 035.45
纯保险费率/ %	3.40	2.90

The table shows that at 80% and 70% guarantee level, yield loss probability, the expectation yield of maize in Anguo city and the rate of agricultural insurance.

设定安全系数为 15%, 营业费用 20%, 预定节余率 5%, 代入毛保险费率 = 纯保险费率 \times (1 + 安全系数) \times (1 + 营业费用) \times (1 + 预定节余率) 的计算式^[8], 则 2 种保障水平下玉米保险的毛保险费率分别为 4.9% 和 4.2%。

3 结 论

1) 本文构建了以非参数核密度估计法为核心的农作物 GRP 保险费率厘定的方法体系。在实际应用过程中,这种方法可以近似地分离出农作物生产的系统风险,通过信息扩散模型的离散化使这一模型可程序化运算,显示了本方法较之参数估计法的优越性;该费率的厘定过程也说明,在我国现有的统计资料条件下,运用非参数核密度法来评估农作物区域产量损失风险在操作上是可行的。

2) 对安国市玉米保险具体实例的估算结果显示,在 80%和 70%的保障水平下,纯保险费率分别为 3.4%和 2.9%(表 4),与郭迎春、闫宜玲通过区域风险评估得出的研究结果比较接近。这是运用玉米多年的产量波动数据,以数理统计理论估算的具有平均意义的结果。应该说明,在纯费率厘定时,不同的费率厘定方法得出的纯费率是有差别的。如果参考不同方法得出的纯费率估算值,就会得出较为理想的纯费率厘定结果。本文估算的结果可以作为区域纯保险费率厘定的参照之一。

3) 推行区域农业保险,实行差别保险费率是我国推行农业保险的必然方向。由于非参数方法本身存在的优越性,非参数核密度估计法为区域农作物

产量损失分布和费率厘定提供了一种可供选择的途径,有助于农业风险决策的科学化。

参 考 文 献

- [1] Thrvey C G, Zhao C. Parametric and nonparametric crop yield distributions and their effects on all-risk crop insurance premiums[D]. Guelph: University of Guelph, 1993
- [2] Barry K G, Alan P K. Nonparametric estimation of crop yield distributions: implications for rating group-risk crop insurance contracts[J]. Amer J Agr Econ, 1998, 80: 139-153
- [3] 郭迎春, 闫宜玲. 农业自然风险评估及区域农业保险费率的确定方法[J]. 应用气象学报, 1998(5): 232-237
- [4] Alan P K, Barry K G. Nonparametric estimation of crop insurance rates revisited[J]. Amer J Agr Econ, 2000, 83: 463-478
- [5] 刘长标. 农作物区域产量保险的精算研究[D]. 北京: 中国人民大学, 2000
- [6] 谭英平. 非参数密度估计在个体损失分布中的应用[J]. 统计研究, 2003(8): 40-44
- [7] 钟甫宁, 邢鹂. 粮食单产波动的地区性差异及对策研究[J]. 中国农业资源与区划, 2004(3): 16-19
- [8] 虞国柱, 李军. 农业保险[M]. 北京: 中国人民大学出版社, 2005: 400-401