第35卷第2期 海南师范大学学报(自然科学版) Vol.35 No.2
2022年6月 Journal of Hainan Normal University(Natural Science) Jun.2022

Doi:10.12051/j.issn.1674-4942.2022.02.012

两类独立随机变量和的概率估计

叶 柳,马 丽*
(海南师范大学 数学与统计学院/数据科学与智慧教育教育部重点实验室,海南 海口 571158)

摘 要:相互独立随机变量和的概率估计是概率统计中一个重要的研究方向。本文研究了两类独立随机变量和的概率估计:一类是n个相互独立的服从$\{-1,1\}$上的均匀分布的随机变量和S_n的最大值的概率估计,另一类是两个独立随机变量和的概率估计。首先,用全概率公式、递推的方法及随机变量的对称性给出了$\log P\left(\max\limits_{1\leqslant i\leqslant n}\left|S_i\right|\leqslant C\right)$的表达式,其中$C$为常数且$1\leqslant C<2$;对一般的$C\geqslant 1$,通过对偶数项和奇数项进行分类讨论,用全概率公式和递推的方法得到了该对数的下界。其次,对两个独立的随机变量,本文证明了如果其分布函数的对数大于等于幂函数,则这两个随机变量的和的分布函数的对数也大于等于某个幂函数。

关键词:简单随机游动;独立随机变量和的分布;概率估计。

中图分类号:O211.1 **文献标志码**:A **文章编号**:1674-4942(2022)02-0175-04

Probability Estimation of the Sum of Two Types of Independent Random Variables

YE Liu, MA Li*
(*School of Mathematics and Statistics / Key Laboratory of Data Science and Smart Education, Ministry of Education, Hainan Normal University, Haikou* 571158, *China*)

Abstract: The probability estimation of the sum of mutually independent random variables is an important research direction in probability and statistics. In this paper, the probability estimations of two types of independent random variables' sum are studied: the first one is the probability estimation of the maximum sum of n independent random variables that are uniformly distributed on $\{-1, 1\}$; the second one is the probability estimation of the sum of two independent random variables. Firstly, by using the total probability formula, recursive method and the symmetry of random variables, the expression of $\log P\left(\max\limits_{1\leqslant i\leqslant n}\left|S_i\right|\leqslant C\right)$ is given, where C is a constant greater than or equal to 1 and less than 2. For general $C\geqslant 1$, the low bound of $\log P\left(\max\limits_{1\leqslant i\leqslant n}\left|S_i\right|\leqslant C\right)$ is also got by considering whether n is even or not. Secondly, about the two independent random variables, this paper proves that if the logarithm of each random variable's distribution is greater than or equal to some power functions, then the logarithm of their sum's distribution is also greater than or equal to some power functions.

Keywords: simple random walk; the distribution of the sum of independent random variables; probability estimation

众所周知,中心极限定理在概率论与数理统计中起着非常重要的作用,特别是在大样本情形下,经常要

收稿日期:2021-12-25
基金项目:国家自然科学基金项目(11861029);海南省高层次人才项目(120RC589);海南省研究生创新科研课题(Qhys2021-301);海南省自然科学基金面上项目(122MS056)
第一作者:叶柳(1998—),广东河源人,硕士研究生,研究方向为随机微分方程。E-mail:1187621908@qq.com
*通信作者:马丽(1979—),河南南阳人,副教授,研究方向为随机微分方程。E-mail:malihnsd@163.com

借助于该定理得到统计量的渐近正态分布,从而进行统计推断。该定理描述了许多彼此没有什么相依关系,对随机现象均不能起到突出影响而均匀地起到微小作用的随机因素共同作用的结果呈现出正态分布的特点。正态分布由期望和方差确定,由大数定理可以确定一列随机变量和的期望。方差实际上是随机变量的和与其期望的差收敛到零的速度,该问题可以归为如下的小值概率问题。设$\{V_n\}$为一列非负随机变量,$h_n > 0$,当$n \to \infty$时,$P\left(V_n \leqslant h_n\right) \to 0$。特殊地,可以取$h_n = \varepsilon_n$或者取$h_n = C$或者$h_n = (1 - \delta)EV_n(0 < \delta \leqslant 1)$。

近年来,有很多学者关注了小值概率问题并在独立随机变量和的概率估计方面取得了重要进展。文献[1]利用鞅方法得到了独立随机变量泛函的集中不等式。文献[2]研究了独立随机变量部分和的大偏差概率估计以及它的渐近性质,文献[3]研究了一类独立的重尾随机变量随机和的大偏差概率,并且将$\left\{X_n, n \geqslant 1\right\}$为独立同分布情形推广到了独立不同分布情形。文献[4]得到相互独立且有界的随机变量和的概率不等式,并改进了其估计的界,但是文献[4]对随机变量有界的要求太高。文献[5]研究了Tauberian定理,并把不同形式的Tauberian定理统一到一种形式。基于文献[4]和文献[5],本文考虑研究如下两个问题:(1)相互独立的期望为0、方差为1的随机变量和的最大值的概率估计:设$\{X_i, i \geqslant 1\}$是独立同分布随机变量,且$EX_i = 0$和$EX_i^2 = 1$,对于任意的$t_0 > 0$,都有$E\exp\left(t_0\left|X_1\right|\right) < \infty$,$S_n = \sum_{i=1}^n X_i$,$\log P\left(\max\limits_{1 \leqslant i \leqslant n}\left|S_i\right| \leqslant C\right) \sim -\mathrm{A}n$(A是常数)中的常数A应为多少?(2)设$V_1 \geqslant 0$和$V_2 \geqslant 0$是两个独立的随机变量,且对任意小的$t > 0$有$\log P\left(V_i \leqslant t\right) \geqslant c_i t^{-\alpha_i}(i = 1, 2, \cdots)$,对任意小的$t > 0$,$\log P\left(V_1 + V_2 \leqslant t\right)$的下界为多少?

1 主要结果

本文主要结果如下:

定理1.1 令$\{X_i, i \geqslant 1\}$是独立同分布随机变量且$P\left(X_i = \pm 1\right) = \dfrac{1}{2}$,设$S_n = \sum_{i=1}^n X_i$,

(1)若$1 \leqslant C < 2$,则

$$\log P\left(\max_{1 \leqslant i \leqslant n}\left|S_i\right| \leqslant C\right) = \begin{cases} \log\dfrac{1}{2^{\frac{n-1}{2}}} = \dfrac{n-1}{2}\log\dfrac{1}{2} = -\dfrac{\log 2}{2}(n-1), & n\text{为奇数}, \\ \log\dfrac{1}{2^{\frac{n}{2}}} = \dfrac{n}{2}\log\dfrac{1}{2} = -\dfrac{\log 2}{2}n, & n\text{为偶数。} \end{cases} \tag{1}$$

(2)对于一般的$C \geqslant 1$,当n为偶数时,

$$\log P(\max_{1 \leqslant i \leqslant n}|S_i| \leqslant C) \geqslant \begin{cases} -\dfrac{\log 2}{2}n, & 1 \leqslant C < 2; \\ -\dfrac{\log 2}{2}(n-2), & C \geqslant 2\text{。} \end{cases}$$

当n为奇数时,

$$\log P\left(\max_{1 \leqslant i \leqslant n}\left|S_i\right| \leqslant C\right) \geqslant -\frac{\log 2}{2}(n-1)\text{。}$$

定理1.2 令$V_1 \geqslant 0$和$V_2 \geqslant 0$是两个独立的随机变量,且对任意小$t > 0$,

$$\log P\left(V_i \leqslant t\right) \geqslant c_i t^{-\alpha_i}, i = 1, 2, \cdots, \tag{2}$$

则对任意小的$t > 0$,有

$$\log P\left(V_1 + V_2 \leqslant t\right) \geqslant \left[c_1^{\frac{1}{1 + (\alpha_1 \vee \alpha_2)}} + c_2^{\frac{1}{1 + (\alpha_1 \vee \alpha_2)}}\right]^{1 + (\alpha_1 \vee \alpha_2)} t^{-(\alpha_1 \vee \alpha_2)}\text{。} \tag{3}$$

2 主要结果的证明

定理1.1的证明:

(1)由全概率公式及X_i的独立性可得

$$
\begin{aligned}
P\left(\max_{1\leqslant i\leqslant n}\left|S_i\right|=1\right)&=P\left(\max_{1\leqslant i\leqslant n}\left|S_i\right|=1\middle|X_1=1\right)P\left(X_1=1\right)+P\left(\max_{1\leqslant i\leqslant n}\left|S_i\right|=1\middle|X_1=-1\right)P\left(X_1=-1\right)\\
&=\frac{1}{2}P\left(\max_{1\leqslant i\leqslant n}\left|S_i\right|=1\middle|X_1=1\right)+\frac{1}{2}P\left(\max_{1\leqslant i\leqslant n}\left|S_i\right|=1\middle|X_1=-1\right)\\
&=\frac{1}{2}P\left(\max_{1\leqslant i\leqslant n}\left|S_i\right|=1\middle|X_1=1,X_2=1\right)P\left(X_2=-1\middle|X_1=1\right)\\
&+\frac{1}{2}P\left(\max_{1\leqslant i\leqslant n}\left|S_i\right|=1\middle|X_1=1,X_2=1\right)P\left(X_2=1\middle|X_1=-1\right)\\
&=\frac{1}{4}P\left(\max_{1\leqslant i\leqslant n-2}\left|S_i\right|=1\right)+\frac{1}{4}P\left(\max_{1\leqslant i\leqslant n-2}\left|S_i\right|=1\right)\\
&=\frac{1}{2}P\left(\max_{1\leqslant i\leqslant n-2}\left|S_i\right|=1\right)\\
&=\frac{1}{2^2}P\left(\max_{1\leqslant i\leqslant n-4}\left|S_i\right|=1\right)\\
&=\begin{cases}\dfrac{1}{2^{\frac{n-2}{2}}}P\left(\max\limits_{1\leqslant i\leqslant 2}\left|S_i\right|=1\right),n\text{为偶数}\\[2ex]\dfrac{1}{2^{\frac{n-1}{2}}}P\left(\max\limits_{1\leqslant i\leqslant 1}\left|S_i\right|=1\right),n\text{为奇数}\end{cases}\\
&=\begin{cases}\dfrac{1}{2^{\frac{n-2}{2}}}\left(\dfrac{1}{4}+\dfrac{1}{4}\right)=\dfrac{1}{2^{\frac{n}{2}}},n\text{为偶数};\\[2ex]\dfrac{1}{2^{\frac{n-1}{2}}}\left(\dfrac{1}{2}+\dfrac{1}{2}\right)=\dfrac{1}{2^{\frac{n-1}{2}}},n\text{为奇数。}\end{cases}
\end{aligned}
$$

注意到$|S_n|$只能取整数值,因此,当$1\leqslant C<2$时,

$$
\log P\left(\max_{1\leqslant i\leqslant n}\left|S_i\right|\leqslant C\right)=\begin{cases}\log\dfrac{1}{2^{\frac{n-1}{2}}}=\dfrac{n-1}{2}\log\dfrac{1}{2}=-\dfrac{\log 2}{2}(n-1),\quad n\text{为奇数};\\[2ex]\log\dfrac{1}{2^{\frac{n}{2}}}=\dfrac{n}{2}\log\dfrac{1}{2}=-\dfrac{\log 2}{2}n,\quad n\text{为偶数。}\end{cases}
$$

(2)当$C\geqslant 1$时,设P_n为从零点出发的随机游动前n步没跑出区间$[-C,C]$的概率,$P_{k,n}$为从k出发的随机游动前n步没跑出区间$[-C,C]$的概率。则由全概率公式可得

$$
\begin{aligned}
P\left(\max_{1\leqslant i\leqslant n}\left|S_i\right|\leqslant C\right)&=\frac{1}{4}P\left(\max_{1\leqslant i\leqslant n}\left|S_i\right|\leqslant C\middle|X_1=1,X_2=1\right)+\frac{1}{4}P\left(\max_{1\leqslant i\leqslant n}\left|S_i\right|\leqslant C\middle|X_1=1,X_2=-1\right)\\
&+\frac{1}{4}P\left(\max_{1\leqslant i\leqslant n}\left|S_i\right|\leqslant C\middle|X_1=-1,X_2=1\right)+\frac{1}{4}P\left(\max_{1\leqslant i\leqslant n}\left|S_i\right|\leqslant C\middle|X_1=-1,X_2=-1\right)\\
&=\frac{1}{2}P_{n-2}+\frac{1}{4}P_{2,n-2}+\frac{1}{4}P_{-2,n-2}\\
&\geqslant\frac{1}{2}P_{n-2}\geqslant\frac{1}{2^2}P_{n-4}\geqslant\cdots\\
&\geqslant\begin{cases}\dfrac{1}{2^{\frac{n-2}{2}}}P_2,\qquad n\text{为偶数};\\[2ex]\dfrac{1}{2^{\frac{n-1}{2}}}P_1,\qquad n\text{为奇数。}\end{cases}
\end{aligned}
$$

又由于$P_2=\begin{cases}\dfrac{1}{2},\qquad 1\leqslant C<2;\\[1ex]1,\qquad C\geqslant 2\text{。}\end{cases}$

故n为偶数时，

$$\log P(\max_{1\leqslant i\leqslant n}|S_i|\leqslant C)\geqslant\begin{cases}-\dfrac{\log 2}{2}n, & 1\leqslant C<2;\\ -\dfrac{\log 2}{2}(n-2), & C\geqslant 2。\end{cases}$$

n为奇数时，$\log P\left(\max\limits_{1\leqslant i\leqslant n}\left|S_i\right|\leqslant C\right)\geqslant-\dfrac{\log 2}{2}(n-1)$。

定理1.2的证明：分三种情况考虑。

(a)当$\alpha_1=\alpha_2=\alpha$时，对于任意的$0<\lambda<1$，由$V_1\geqslant 0$和$V_2\geqslant 0$是两个独立的随机变量可得

$$\log P\left(V_1+V_2\leqslant t\right)\geqslant\log P\left(V_1\leqslant\lambda t\right)+\log P\left(V_2\leqslant(1-\lambda)t\right)\geqslant-\left(c_1\lambda^{-\alpha}+c_2(1-\lambda)^{-\alpha}\right)t^{-\alpha}。\tag{4}$$

式(4)对任意的$0<\lambda<1$都成立，因此，式(4)左侧大于等于右侧的最大值，下面求右侧的最值点。设$f(\lambda)=-\left(c_1\lambda^{-\alpha}+c_2(1-\lambda)^{-\alpha}\right)t^{-\alpha}$，令$f'(\lambda)=0$即

$$\alpha t^{-\alpha}\left(c_1\lambda^{-(\alpha+1)}-c_2(1-\lambda)^{-(\alpha+1)}\right)=0。$$

解方程可得$\lambda_0=\dfrac{c_2^{-\frac{1}{1+\alpha}}}{c_1^{-\frac{1}{1+\alpha}}+c_2^{-\frac{1}{1+\alpha}}}$，代入可得$f\left(\lambda_0\right)=[c_1^{\frac{1}{1+\alpha}}+c_2^{\frac{1}{1+\alpha}}]^{1+\alpha}t^{-\alpha}$。因此，

$$\log P\left(V_1+V_2\leqslant t\right)\geqslant\left[c_1^{\frac{1}{1+\alpha}}+c_2^{\frac{1}{1+\alpha}}\right]^{1+\alpha}t^{-\alpha}。$$

(b)当$\alpha_1>\alpha_2$时，与(a)同理得到

$$\begin{aligned}\log P\left(V_1+V_2\leqslant t\right)&\geqslant\log P\left(V_1\leqslant\lambda t\right)+\log P\left(V_2\leqslant(1-\lambda)t\right)\\&\geqslant-c_1(\lambda t)^{-\alpha_1}-c_2\left[(1-\lambda)t\right]^{-\alpha_2},\end{aligned}\tag{5}$$

由于$\alpha_1>\alpha_2$且$0<t<1$，则$-t^{-\alpha_1}<-t^{-\alpha_2}$，

$$\log P\left(V_1+V_2\leqslant t\right)\geqslant-c_1(\lambda t)^{-\alpha_1}-c_2(1-\lambda)^{-\alpha_2}t^{-\alpha_1},$$

又由于$0<1-\lambda<1$，由函数的单调性可得

$$\log P\left(V_1+V_2\leqslant t\right)\geqslant-c_1(\lambda t)^{-\alpha_1}-c_2\left[(1-\lambda)t\right]^{-\alpha_1},$$

此时与情况(a)相同，故

$$\log P\left(V_1+V_2\leqslant t\right)\geqslant\left[c_1^{\frac{1}{1+\alpha_1}}+c_2^{\frac{1}{1+\alpha_1}}\right]^{1+\alpha_1}t^{-\alpha_1}。\tag{6}$$

(c)当$\alpha_1<\alpha_2$时，与(b)情况类似，故可得

$$\log P\left(V_1+V_2\leqslant t\right)\geqslant\left[c_1^{\frac{1}{1+\alpha_2}}+c_2^{\frac{1}{1+\alpha_2}}\right]^{1+\alpha_2}t^{-\alpha_2}。\tag{7}$$

综上，由式(5)、(6)和(7)可得

$$\log P\left(V_1+V_2\leqslant t\right)\geqslant\left[c_1^{\frac{1}{1+(\alpha_1\vee\alpha_2)}}+c_2^{\frac{1}{1+(\alpha_1\vee\alpha_2)}}\right]^{1+(\alpha_1\vee\alpha_2)}t^{-\left(\alpha_1\vee\alpha_2\right)}。$$

(下转第199页)