第 48 卷第 3 期 郑 州 大 学 学 报(理 学 版) Vol. 48 No. 3
2016 年 9 月 J. Zhengzhou Univ. (Nat. Sci. Ed.) Sep. 2016

负二项分布概率最大值的性质

丁 勇

(南京医科大学 康达学院 数学与计算机教研室 江苏 连云港 222000)

摘要: 负二项分布概率的最大值是每次试验成功的概率 p 和首次试验成功次数 r 的函数. 对确定的 r,该函数是 p 的单调上升的连续函数,仅当 $(r-1)/p$ 是整数时不可导;对确定的 p,该函数是 r 的单调下降函数.

关键词: 负二项分布; 概率最大值; 单调性

中图分类号: O211.1 **文献标志码:** A **文章编号:** 1671-6841(2016)03-0047-04

DOI: 10.13705/j.issn.1671-6841.2016050

0 引言

在伯努利实验的家族中,作为几何分布的一种延伸,负二项分布是重要的离散型分布之一,在排队论、可靠性以及群团型生态格局分布等方面有着重要的应用[1-3],有关该分布的理论和应用研究已取得很多成果[4-10]. 本文对负二项分布概率最大值的性质进行讨论,使我们对负二项分布有更深入的认识和理解.

1 负二项分布的最大值

在独立重复贝努里试验中,第 r 次成功的试验次数 X 是个随机变量,其一切可能值是 $r, r+1, \cdots\cdots, X$ 的概率分布称为负二项分布,且有[10]

$$P_r(X=k)=C_{k-1}^{r-1}p^r(1-p)^{k-r},\ 0<p<1, r\geqslant 1, k=r, r+1, r+2, \cdots, \tag{1}$$

这里 p 为每次试验成功的概率.

负二项分布也称为巴斯卡分布,当 $r=1$ 时成为几何分布. 对确定的 p,当 $r=1$ 时,负二项分布的图形是下降的. 当 $r>1$ 时,负二项分布的图形先上升,后下降. 图 1 为 $p=0.6$、$r=5$ 时负二项分布的图形. 已知 $k=\left[\frac{r-1}{p}\right]+1$ 时(这里[ ]表示向下取整)[5],$P_r(X=k)$ 有最大值,记为 $P^{\max}$,它是一个关于 r 和 p 的二元函数. 固定 r 或 p,则 $P^{\max}$ 是一个关于 p 或 r 的一元函数,分别记为 $P_r^{\max}(p)$ 和 $P_p^{\max}(r)$,本文对这两个函数的性质进行研究. 下面讨论中,$\alpha\sim\beta$ 表示 $\lim\frac{\alpha}{\beta}=1$.

2 $P_r^{\max}(p)$ 的性质

图 2 是 $r=10$ 时 $P_r^{\max}(p)$ 的图形,由图 2 可观察到,$P_r^{\max}(p)$ 是一条有尖点的单调上升的连续曲线,

$$\lim_{p\to 0}P_r^{\max}(p)=0, \lim_{p\to 1}P_r^{\max}(p)=1.$$

性质 1 $P_r^{\max}(p)$ 是单调上升的连续函数,且 $\lim\limits_{p\to 0}P_r^{\max}(p)=0, \lim\limits_{p\to 1}P_r^{\max}(p)=1$. 当 $\frac{r-1}{p}$ 不是整数时,$P_r^{\max}(p)$ 是可导函数,记 $k=\left[\frac{r-1}{p}\right]+1$,则 $\frac{\mathrm{d}P_r^{\max}(p)}{\mathrm{d}p}=C_{k-1}^{r-1}p^{r-1}(1-p)^{k-r-1}(r-kp)>0$,当 $\frac{r-1}{p}$ 是整数时,$P_r^{\max}(p)$ 是不可导函数.

收稿日期: 2016-03-06

作者简介: 丁勇(1956—),男,江苏淮安人,教授,主要从事生物统计研究,E-mail:yding@njmu.edu.cn.

引用本文: 丁勇. 负二项分布概率最大值的性质[J]. 郑州大学学报(理学版),2016,48(3):47-50.

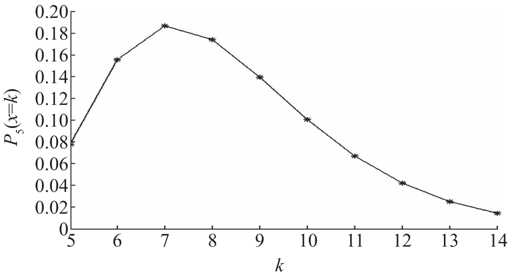

图1 负二项分布的图形($p=0.6,r=5$)

Fig. 1 Graphs of the negative binomial distribution($p=0.6,r=5$)

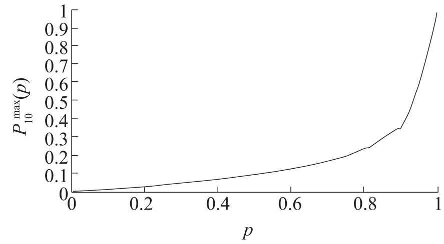

图2 负二项分布最大值 $P_{10}^{\max}(p)$ 的图形($p=0.001,0.002,\cdots,0.999$)

Fig. 2 Graphs of the maximumvalue $P_{10}^{\max}(p)$ of negative binomial distribution($p=0.001,0.002,\cdots,0.999$)

证明 记$\frac{r-1}{p}=\left[\frac{r-1}{p}\right]+t,0\leqslant t<1.$

先考虑 $t>0$ 的情况,取 Δp 充分小($\frac{r-1}{k}-p<\Delta p<\frac{r-1}{k-1}-p$),使$\left[\frac{r-1}{p+\Delta p}\right]=k-1=\left[\frac{r-1}{p}\right]$,即当 p 有微小改变时,k 不变,从而

$$\Delta P_r^{\max}=P_r^{\max}(p+\Delta p)-P_r^{\max}(p)=C_{k-1}^{r-1}(p+\Delta p)^r\left[1-(p+\Delta p)\right]^{k-r}-C_{k-1}^{r-1}p^r(1-p)^{k-r}=$$

$$C_{k-1}^{r-1}p^r(1-p)^{k-r}\left[(1+\frac{\Delta p}{p})^r(1-\frac{\Delta p}{1-p})^{k-r}-1\right]\sim C_{k-1}^{r-1}p^r(1-p)^{k-r}\left\{(1+\frac{r\Delta p}{p})\left[1-\frac{(k-r)\Delta p}{1-p}\right]-1\right\}\sim$$

$$C_{k-1}^{r-1}p^r(1-p)^{k-r}\left[\frac{r\Delta p}{p}-\frac{(k-r)\Delta p}{1-p}\right]=C_{k-1}^{r-1}p^{r-1}(1-p)^{k-r-1}\Delta p(r-kp),$$

故$\lim\limits_{\Delta p\to0}\Delta P_r^{\max}=0$. 由于 $t>0$,所以 $k=\left[\frac{r-1}{p}\right]+1<\frac{r-1}{p}+1$,从而 $kp<r-1+p<r$,即

$$\lim_{\Delta p\to0}\frac{\Delta P_r^{\max}}{\Delta p}=C_{k-1}^{r-1}p^{r-1}(1-p)^{k-r-1}(r-kp)>0.$$

再考虑 $t=0$ 的情况,此时 $k=\frac{r-1}{p}+1$,从而 $k-1>r-1$. 先取 Δp 充分小,($\frac{r-1}{k}-p<\Delta p<0$),使$\left[\frac{r-1}{p+\Delta p}\right]+1=k$,类似于上面证明可得$\lim\limits_{\Delta p\to0^-}\Delta P_r^{\max}=0$,且

$$\lim_{\Delta p\to0^-}\frac{\Delta P_r^{\max}}{\Delta p}=C_{k-1}^{r-1}p^{r-1}(1-p)^{k-r-1}(r-kp)=C_{k-1}^{r-1}p^{r-1}(1-p)^{k-r}>0.$$

再取 Δp 充分小($0<\Delta p<\frac{r-1}{k-2}-p$),使$\left[\frac{r-1}{p+\Delta p}\right]+1=k-1$,

$$\Delta P_r^{\max}=P_r^{\max}(p+\Delta p)-P_r^{\max}(p)=C_{k-2}^{r-1}(p+\Delta p)^r\left[1-(p+\Delta p)\right]^{k-r-1}-C_{k-1}^{r-1}p^r(1-p)^{k-r}=$$

$$\frac{(k-2)!}{(r-1)!\ (k-r)!}p^r(1-p)^{k-1-r}\left[(k-r)(1+\frac{\Delta p}{p})^r(1-\frac{\Delta p}{1-p})^{k-1-r}-(k-1)(1-p)\right]\sim$$

$$\frac{(k-2)!}{(r-1)!\ (k-r)!}p^r(1-p)^{k-1-r}\left[(\frac{r-1}{p}+1-r)(1+\frac{r\Delta p}{p})(1-\frac{(k-1-r)\Delta p}{1-p})-\frac{r-1}{p}(1-p)\right]=$$

$$\frac{(k-2)!}{(r-2)!\ (k-r)!}p^{r-1}(1-p)^{k-r}\left[(1+\frac{r\Delta p}{p})(1-\frac{(k-1-r)\Delta p}{1-p})-1\right]\sim C_{k-2}^{r-2}p^{r-1}(1-p)^{k-r}\cdot$$

$$\left[\frac{r\Delta p}{p}-\frac{(k-1-r)\Delta p}{1-p})\right]=C_{k-2}^{r-2}p^{r-2}(1-p)^{k-1-r}\Delta p\left[r(1-p)-p(\frac{r-1}{p}-r)\right]=C_{k-2}^{r-2}p^{r-2}(1-p)^{k-1-r}\Delta p,$$

所以$\lim\limits_{\Delta p\to0^+}\Delta P_r^{\max}=0$ 且$\lim\limits_{\Delta p\to0^+}\frac{\Delta P_r^{\max}}{\Delta p}=C_{k-2}^{r-2}p^{r-2}(1-p)^{k-r-1}>0$,从而当$\frac{r-1}{p}$是整数时,$\lim\limits_{\Delta p\to0^-}\frac{\Delta P_r^{\max}}{\Delta p}\neq\lim\limits_{\Delta p\to0^+}\frac{\Delta P_r^{\max}}{\Delta p}$,所以此时 $P_r^{\max}(p)$是不可导函数.

由于无论哪种情况都有 $\Delta P_r^{\max}$ 的极限为0 和$\frac{\Delta P_r^{\max}}{\Delta p}$的极限大于0,所以 $P_r^{\max}(p)$是单调上升的连续函数.

当 $r=1$ 时,$P_r^{\max}(p)=p$,显然有 $\lim\limits_{p\to0}P_r^{\max}(p)=0$,$\lim\limits_{p\to1}P_r^{\max}(p)=1$. 当 $r>1$ 时,因为 $\frac{r-1}{p}\leqslant k=\left[\frac{r-1}{p}\right]+1\leqslant\frac{r-1}{p}+1$,故当 $p\to0$ 时,$k\sim\frac{r-1}{p}\to\infty$. 当 $n\to\infty$ 时,根据斯特林公式 $n!\sim\sqrt{2\pi n}\left(\frac{n}{\mathrm{e}}\right)^n$ 可得

$$P_r^{\max}(p)=\frac{(k-1)!}{(r-1)!\,(k-r)!}p^r(1-p)^{k-r}\sim\frac{\sqrt{2\pi(k-1)}\left(\frac{k-1}{\mathrm{e}}\right)^{k-1}p^r(1-p)^{k-r}}{(r-1)!\,\sqrt{2\pi(k-r)}\left(\frac{k-r}{\mathrm{e}}\right)^{k-r}}=$$

$$\frac{p\mathrm{e}^{1-r}}{(r-1)!}\sqrt{\frac{k-1}{k-r}}[(k-1)p]^{r-1}(k-1)^{k-r}\left[\frac{(k-1)(1-p)}{k-r}\right]^{k-r},$$

由于 $\lim\limits_{p\to0}\frac{p\mathrm{e}^{1-r}}{(r-1)!}=0$,$\lim\limits_{p\to0}\sqrt{\frac{k-1}{k-r}}=\lim\limits_{p\to0}\sqrt{\frac{r-1-p}{r-1-pr}}=1$,$\lim\limits_{p\to0}[(k-1)p]^{r-1}=\lim\limits_{p\to0}(r-1-p)^{r-1}=(r-1)^{r-1}$,根据罗必塔法则可得:

$$\lim_{p\to0}(k-1)^{k-r}=\lim_{p\to0}\left(\frac{r-1-p}{p}\right)^{\frac{r-1-pr}{p}}=\lim_{p\to0}\mathrm{e}^{\frac{r-1-pr}{p}[\ln(r-1-p)-\ln p]}=\lim_{p\to0}\mathrm{e}^{-\frac{(r-1-pr)2}{r-1}\left(\frac{1}{r-1-p}+\frac{1}{p}\right)}=0;$$

$$\lim_{p\to0}\left[\frac{(k-1)(1-p)}{k-r}\right]^{k-r}=\lim_{p\to0}\left[\frac{(r-1-p)(1-p)}{r-1-pr}\right]^{\frac{r-1-pr}{p}}=$$

$$\lim_{p\to0}\mathrm{e}^{\left(\frac{r-1-pr}{p}\right)[\ln(r-1-p)+\ln(1-p)-\ln(r-1-pr)]}=\lim_{p\to0}\mathrm{e}^{\frac{(r-1-pr)2}{r-1}\left(\frac{r}{r-1-p}-\frac{1}{1-p}-\frac{1}{r-1-pr}\right)}=1.$$

所以 $\lim\limits_{p\to0}P_r^{\max}(p)=0$. 令 $\frac{1}{p}=1+\varepsilon$,当 $p\to1$ 时,$\varepsilon\to0$,取 $\varepsilon<\frac{1}{r-1}$,由于 $k=\left[\frac{r-1}{p}\right]+1=[r-1+\varepsilon(r-1)]+1=r$,所以当 $p\to1$ 时,$k=r$,故有 $\lim\limits_{p\to1}P_r^{\max}(p)=C_{k-1}^{r-1}p^r(1-p)^{k-r}=1$.

3　$P_p^{\max}(r)$ 的性质

性质 2　记 $k_r=\left[\frac{r-1}{p}\right]+1$,$m=k_{r+1}-k_r$,则 $m\geqslant1$,当 $\frac{1}{2}\leqslant p\leqslant1$ 时,$1\leqslant m\leqslant2$.

证明　$k_r+1=\left[\frac{r-1}{p}\right]+2=\left[\frac{r-1}{p}+1\right]+1=\left[\frac{r}{p}+1-\frac{1}{p}\right]+1\leqslant\left[\frac{r}{p}\right]+1=k_{r+1}$,当 $\frac{1}{2}\leqslant p\leqslant1$ 时,$k_{r+1}=\left[\frac{r}{p}\right]+1=\left[\frac{r-1}{p}+\frac{1}{p}\right]+1\leqslant\left[\frac{r-1}{p}+2\right]+1\leqslant\left[\frac{r-1}{p}\right]+3=k_r+2$.

性质 3　对确定的 p,$\lim\limits_{r\to\infty}P_p^{\max}(r)=0$,$P_p^{\max}(r+1)\leqslant P_p^{\max}(r)\leqslant P_p^{\max}(1)=p$,即负二项分布的概率最大值是 r 的单调下降函数,且不超过 p.

证明　由于 $\frac{r-1}{p}\leqslant k=\left[\frac{r-1}{p}\right]+1\leqslant\frac{r-1}{p}+1$,故当 $r\to\infty$ 时,$k\sim\frac{r-1}{p}\to\infty$,根据斯特林公式

$$P_p^{\max}(r)=\frac{(k-1)!}{(r-1)!\,(k-r)!}p^r(1-p)^{k-r}\sim\frac{\sqrt{2\pi(k-1)}\left(\frac{k-1}{\mathrm{e}}\right)^{k-1}p^r(1-p)^{k-r}}{\sqrt{2\pi(r-1)}\sqrt{2\pi(k-r)}\left(\frac{r-1}{\mathrm{e}}\right)^{r-1}\left(\frac{k-r}{\mathrm{e}}\right)^{k-r}}=$$

$$\sqrt{\frac{k-1}{r-1}}\frac{p}{\sqrt{2\pi(k-r)}}\left[\frac{p(k-1)}{r-1}\right]^{r-1}\left[\frac{(k-1)(1-p)}{k-r}\right]^{k-r}.$$

由于 $\lim\limits_{r\to\infty}p\sqrt{\frac{k-1}{r-1}}=p\lim\limits_{r\to\infty}\sqrt{\frac{1}{p}-\frac{1}{r-1}}=p\sqrt{\frac{1}{p}}=\sqrt{p}$,$\lim\limits_{r\to\infty}\frac{1}{\sqrt{2\pi(k-r)}}=\lim\limits_{r\to\infty}\sqrt{\frac{p}{2\pi[r(1-p)-1]}}=0$,再根据第二个重要极限,$\lim\limits_{r\to\infty}\left[\frac{p(k-1)}{r-1}\right]^{r-1}=\lim\limits_{r\to\infty}\left(\frac{r-1-p}{r-1}\right)^{r-1}=\lim\limits_{r\to\infty}\left[\left(1-\frac{p}{r-1}\right)^{-\frac{r-1}{p}}\right]^{-p}=\mathrm{e}^{-p}\lim\limits_{r\to\infty}\left[\frac{(k-1)(1-p)}{k-r}\right]^{k-r}=$

$$\lim_{r\to\infty}\left[\frac{\left(\frac{r-1}{p}-1\right)(1-p)}{\frac{r-1}{p}-r}\right]^{\frac{r-1}{p}-r}=\lim_{r\to\infty}\left[\frac{(r-1-p)(1-p)}{r-1-pr}\right]^{\frac{r-1-pr}{p}}=\lim_{r\to\infty}\left(1+\frac{p^2}{r-1-pr}\right)^{\frac{r-1-pr}{p^2}p}=\mathrm{e}^p,$$

所以$\lim\limits_{r\to\infty} P_p^{\max}(r)=0$.

再证 $P_p^{\max}(r+1)=C_{k_{r+1}-1}^{r}p^{r+1}(1-p)^{k_{r+1}-r-1}\leqslant C_{k_r-1}^{r-1}p^{r}(1-p)^{k_r-r}=P_p^{\max}(r)$,记 $k_{r+1}-k_r=m$ 时,上述不等式为$\frac{(k_r+m-1)!}{r!\ (k_r+m-1-r)!}p^{r+1}(1-p)^{k_r+m-r-1}\leqslant\frac{(k_r-1)!}{(r-1)!\ (k_r-r)!}p^{r}(1-p)^{k_r-r}$,化简得

$$\frac{(k_r+m-1)!}{r(k_r+m-1-r)!}p(1-p)^{m-1}\leqslant\frac{(k_r-1)!}{(k_r-r)!}. \tag{2}$$

下面用归纳法对(2)式进行证明. 当 $m=1$ 时,(2)式为$\frac{k_r}{r}p\leqslant 1$,由于 $k_r\frac{p}{r}=\left\{\left[\frac{r-1}{p}\right]+1\right\}\frac{p}{r}\leqslant(\frac{r-1}{p}+1)\frac{p}{r}=1-\frac{1-p}{r}<1$,故当 $m=1$ 时,不等式成立.

归纳假设不等式(2)成立,将 m 换作 $m+1$,要证$\frac{(k_r+m)(k_r+m-1)!}{r(k_r+m-r)(k_r+m-1-r)!}p(1-p)^{m}\leqslant\frac{(k_r-1)!}{(k_r-r)!}$,根据归纳假设,即要证$\frac{(k_r+m)}{(k_r+m-r)}(1-p)\leqslant 1$,即$(k_r+m)p\geqslant r$,由于 $k_{r+1}=k_r+m=\left[\frac{r}{p}\right]+1\geqslant\frac{r}{p}$,故不等式成立.

4 讨论

本文就负二项分布概率最大值的问题进行了探讨,发现了负二项分布概率最大值的二元分布图(图3),相邻两个最大值之差的图形,呈现出震荡的、振幅不断变大的趋势(图4). 负二项分布概率最大值是 p 的连续函数,且在可数个点不可导,其值随着 p 的增加而增加(图3)、随着 r 的增加而减少(图3),当 r 增加1个单位时,极大值点右移,至少右移一个单位,当$\frac{1}{2}\leqslant p\leqslant 1$ 时,极大值点最多右移2个单位.

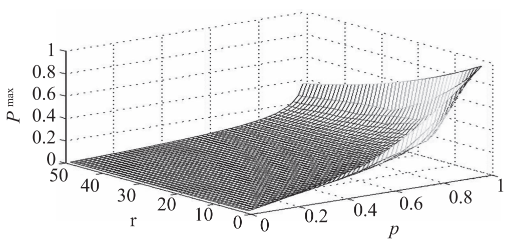

图3 负二项分布概率最大值 $P^{\max}$ 分布图 ($p=0.02,0.04,\cdots,0.98;r=1,2,\cdots,49$)

Fig. 3 Graphs of the maximum value $P^{\max}$ of the negative binomial distribution($p=0.02,0.04,\cdots,0.98;r=1,2,\cdots,49$)

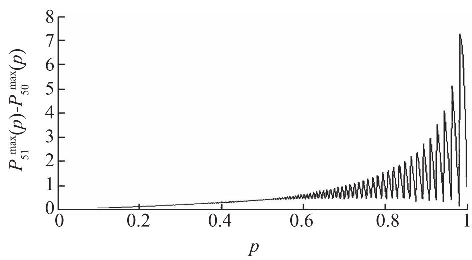

图4 相邻两个最大值之差 $P_{r+1}^{\max}(p)-P_r^{\max}(p)$ 的图形($r=50$)

Fig.4 The difference between the adjacent two maximum $P_{r+1}^{\max}(p)-P_r^{\max}(p)$ of the negative binomial distributionmaximum($r=50$)

参考文献:

[1] 蒋仁言,左明健著. 可靠性模型与应用[M]. 北京:机械工业出版社,1999.
[2] 田乃硕. 休假随机服务系统[M]. 北京:北京大学出版社,2001.
[3] 陈峰,杨树勤. 论负二项分布的应用条件[J]. 中国卫生统计, 1995,4:21-22.
[4] 周源泉,李宝盛. 预定成功数的负二项分布预测[J]. 质量与可靠性,2013 (1):1-8.
[5] 何春. 负二项分布概率的最大值点[J]. 生物数学学报,2011,26(1):160-162
[6] GUPTA R C,ONG S H. A new generalization of the negative binomial distribution[J]. Computational statistics and data analysis,2003, 45(2):287-300.
[7] 牛燕影,王增富,田乃硕. 负二项分布类的条件概率封闭性[J]. 数学理论与应用,2005,25(3):101-103.
[8] 熊加兵,陈光曙. 负二项分布随机变量的分解定理[J]. 大学数学, 2008,24(1):108-110.
[9] 王新利,陈光曙. 几何分布和负二项分布高阶矩的递推公式[J]. 高等数学研究,2011,14(2):15-16.
[10] 康殿统. 负二项分布的两个不同定义[J]. 河西学院学报. 2014,30(5):22-30.

(下转第56页)

2001,20(1):5 - 22.

[12] TRIPATHY A, AGRAWAL A, RATH S K. Classification of sentiment reviews using n-gram machine learning approach[J]. Expert systems with applications, 2016, 57(15):117 - 126.

[13] LIU Q, HE Q, SHI Z. Extreme support vector machine classifier [C] //Procceedings of the 12th Pacific-Asia conference on advances in knowledge discovery and data mining. Berlin, 2008.

A Smooth Extreme Learning Machine for Classification

YANG Liming, ZHANG Siyun, REN Zhuo

(*College of Science*, *China Agricultural University*, *Beijing* 100083, *China*)

Abstract: Extreme learning machine (ELM) had a high learning speed and a good generalization ablity. Smoothing strategy was an important technology for non-smooth problems. By combining a smoothing technique with ELM, a smooth ELM (SELM) framework was proposed. Moreover, the Newton-Armijo algorithm was used to solve the SELM, and resulting algorithm converged globally and quadratically. The proposed SELM had less decision variables and better abitities to deal with nonlinear problems than the existing smooth support vector machine. Numerical experiments demonstrated that the speed of SELM was much faster than that of the existing ELM algorithms based on optimization theory. Compared with other popular support vector machines, the proposed SELM achieved better or similar generalization. The results demonstrated the feasibility and effectiveness of the proposed algorithm.

Key words: extreme learning machine(ELM); smooth approach; Newton-Armijo algorithm; neural networks

(责任编辑:王浩毅)

(上接第50页)

The Characters of the Probability Maximum Value for Negative Binomial Distribution

DING Yong

(*Department of Mathematics and Computer Science*, *KangDa College of Nanjing Medical University*, *Lianyungang* 222000, *China*)

Abstract: The character of probability maximum value for negative binomial distribution was explored. The probability maximum value for negative binomial distribution was a function of p and r, where p was the probability of success for each test, and r was the number of the first successful test. It was a monotonically increasing continuous function of p when r was given, only $(r-1)/p$ was a integer, its derivative did not exist, and a monotone decreasing function of r when p was given.

Key words: negative binomial distribution; maximum probability; monotonicity

(责任编辑:方惠敏)