Vol. 37 (2017)
No. 4

数 学 杂 志
J. of Math. (PRC)

k-NORMAL DISTRIBUTION AND ITS APPLICATIONS

HAN Tian-yong[1], WEN Jia-jin[1], SONG An-chao[2], YE Jian-hua[1]

(*1.College of Information Science and Engineering, Chengdu University, Chengdu 610106, China*)

(*2.School of Statistics, Southwestern University of Finance and Economics, Chengdu 611130, China*)

Abstract: In this paper, we study the truncated variables and k-normal distribution. By using the theory of logarithmic concave function, we obtain the inequality chains involving variances of truncated variables and the function of truncated variables, which is the generalization of some classical results involving normal distribution and the hierarchical teaching model. Some simulation results and a real data analysis are shown.

Keywords: truncated random variables; k-normal distribution; hierarchical teaching model; logarithmic concave function; simulation

2010 MR Subject Classification: 62J10; 62P25; 60E05; 60E15; 26D15; 26E60

Document code: A **Article ID:** 0255-7797(2017)04-0737-14

1 Introduction

With the expansion of university enrollment, various work to improve students' ability all round was continued to be carried out. How to increasingly improve teaching quality in the courses with large number of students (such as advanced mathematics) are discussed repeatedly. Since the examination scores of the large number of students obey normal distribution, statistical theory is a natural research tool for study of a large scale teaching (see [1, 2]).

The math score of the students of some grades in a university is a random variable ξ_I, where $\xi_I \in I = [0, 100)$. Assume that the students are taught by divided into n classes according to their math scores, written as: Class$[a_1, a_2)$, Class$[a_2, a_3)$, $\cdots$, Class$[a_n, a_{n+1})$, where $n \geq 3, 0 = a_1 < a_2 < \cdots < a_{n+1} = 100$, and a_i, a_{i+1} are the lowest and the highest math scores of the students of the Class$[a_i, a_{i+1})$, respectively. This model of teaching is called hierarchical teaching model (see [1–4, 7]). This teaching model is often used in college English and college mathematics teaching. In teaching practice, the previously mentioned score maybe the math score of national college entrance examination or entrance exams which represent the mathematical basis of the students, or in mathematical language, the initial value of the teaching.

* **Received date:** 2016-02-25 **Accepted date:** 2016-09-28

Foundation item: Supported by the Natural Science Foundation of Sichuan Science and Technology Department (2014SZ0107).

Biography: Han Tianyong (1976–), male, born at Chengdu, Sichuan, associate professor, major in dynamical system, inequality and its application.

No doubt that this teaching model is better than traditional teaching model. However, the real reason for it's high efficiency and the further improvement are not found. As far as we know, not many papers were published to deal these since the difficulty of computing the indefinite integrals involving the normal distribution density function. In [3], by means of numerical simulation, the authors proved the variance of the hierarchical class is smaller. In [4], the authors established some general properties of the variance of the hierarchical teaching, and established a linear model of teaching efficiency of hierarchical teaching model. If the students are divided into Superior-Middle-Poor three classes, the authors believe that the three classes, especially the third one will benefit most from the hierarchical teaching.

In order to study the hierarchical teaching model, we need to give the definition of truncated variables.

Definition 1.1 Let $\xi_I \in I$ be a continuous random variable, and let its probability density function (p.d.f.) be $f : I \to (0, \infty)$. If $\xi_{I^*} \in I^* \subseteq I$ is also a continuous random variable and its probability density function is

$$f_{\xi_{I^*}} : I^* \to (0,\infty), f_{\xi_{I^*}}(t) \triangleq \frac{f(t)}{\displaystyle\int_{I^*} f},$$

then we call the random variable ξ_{I^*} a truncated variable of the random variable ξ_I, denoted by $\xi_{I^*} \subseteq \xi_I$; if $\xi_{I^*} \subseteq \xi_I$, and $I^* \subset I$, then we call the random variable ξ_{I^*} a proper Truncated Variable of the random variable ξ_I, denoted by $\xi_{I^*} \subset \xi_I$, here $I, I^* \subseteq (-\infty, \infty)$, I and I^* are intervals.

In the hierarchical teaching model, the math score of Class$[a_i, a_{i+1})$ is also a random variable $\xi_{[a_i,a_{i+1})} \in [a_i, a_{i+1})$. Since $[a_i, a_{i+1}) \subset I$, we say it is a proper truncated variables of the random variable ξ_I, written as $\xi_{[a_i,a_{i+1})} \subset \xi_I, i = 1, 2, \cdots, n$. Assume that Class$[a_i, a_{i+1})$ and Class$[a_{i+1}, a_{i+2})$ are merged into one, i.e.,

$$\text{Class}[a_i, a_{i+2}), 1 \le i \le n-1.$$

Since $[a_i, a_{i+1}) \subset [a_i, a_{i+2})$ and $[a_{i+1}, a_{i+2}) \subset [a_i, a_{i+2})$, we know that $\xi_{[a_i,a_{i+1})}$ and $\xi_{[a_{i+1},a_{i+2})}$ are the proper truncated variables of the random variable $\xi_{[a_i,a_{i+2})}$.

We remark here if $\xi_I \in I$ is a continuous random variable, and its p.d.f. is $f : I \to (0, \infty)$, then the integration $\displaystyle\int_I f$ converges , and it satisfies the following two conditions

$$\int_I f = 1; P\left(\xi_I \in \overline{I}\right) = \int_{\overline{I}} f,$$

where $P\left(\xi_I \in \overline{I}\right)$ is the probability of the random event $\xi_I \in \overline{I}$, and $\overline{I} \subseteq I$ is an interval.

According to the definitions of the mathematical expectation $E\xi_{I^*}$ and the variance $D\xi_{I^*}$ (see [8, 9]) with Definition 1.1, we are easy to get

$$E\xi_{I^*} \triangleq \int_{I^*} t f_{\xi_{I^*}} = \frac{\displaystyle\int_{I^*} tf}{\displaystyle\int_{I^*} f} \tag{1.1}$$

and

$$D\xi_{I^*} \triangleq E(\xi_{I^*} - E\xi_{I^*})^2 = \frac{\displaystyle\int_{I^*} t^2 f}{\displaystyle\int_{I^*} f} - \left(\frac{\displaystyle\int_{I^*} tf}{\displaystyle\int_{I^*} f} \right)^2, \tag{1.2}$$

where ξ_{I^*} is a truncated variable of the random variable ξ_I.

In the hierarchical teaching model, what we concerned about is the relationship between the variance of $\xi_{[a_i,a_{i+1})}$ and the variance of ξ_I, where $i = 1, 2, \cdots, n$. Its purpose is to determine the superiority and inferiority of the hierarchical teaching model and the traditional mode of teaching. If

$$D\xi_{[a_i,a_{i+1})} < D\xi_I, \forall i \in \{1, 2, \cdots, n\}, \tag{1.3}$$

then we believe that the hierarchical teaching model is better than the traditional mode of teaching. Otherwise, we believe that the hierarchical teaching model is not worth promoting.

2 k-Normal Distribution

The normal distribution (see [3, 4, 8, 9]) is considered as the most prominent probability distribution in statistics. Besides the important central limit theorem that says the mean of a large number of random variables drawn from a common distribution, under mild conditions, is distributed approximately normally, the normal distribution is also tractable in the sense that a large number of related results can be derived explicitly and that many qualitative properties may be stated in terms of various inequalities.

One of the main practical uses of the normal distribution is to model empirical distributions of many different random variables encountered in practice. For fit the actual data more accurately, many research for generalizing this distribution are carried out. Some representative examples are the following. In 2001, Armando and other authors extended the p.d.f. to the normal-exponential-gamma form which contains four parameters (see [5]). In 2005, Saralees generalized it into the form $K \exp\left\{-\left|\frac{x-\mu}{\sigma}\right|^s\right\}$ (see [6]). In 2014, Wen Jiajin rewrote the p.d.f as k-Normal Distribution as follows (see [7]).

Definition 2.1 If ξ is a continuous random variable and its p.d.f. is

$$f_k^{\mu,\sigma} : (-\infty, \infty) \to (0, \infty), f_k^{\mu,\sigma}(t) \triangleq \frac{k^{1-k^{-1}}}{2\Gamma(k^{-1})\sigma} \exp\left(-\frac{|t-\mu|^k}{k\sigma^k} \right),$$

then we call the random variable ξ follows the k-normal distribution, denoted by $\xi \sim N_k(\mu, \sigma)$, where $\mu \in (-\infty, \infty), \sigma \in (0, \infty), k \in (1, \infty)$, and $\Gamma(s) \triangleq \displaystyle\int_0^\infty x^{s-1}e^{-x}\mathrm{d}x$ is the gamma function.

For the p.d.f. $f_k^{\mu,\sigma}(t)$ of k-normal distribution, the graphs of the functions $f_{3/2}^{0,1}(t), f_2^{0,1}(t)$ and $f_{5/2}^{0,1}(t)$ are depicted in Figure 1 and $f_k^{0,1}(t)$ is depicted in Figure 2.

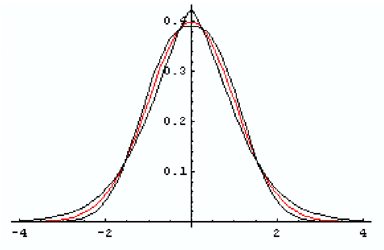

Figure 1: The graphs of the functions $f_{3/2}^{0,1}(t), f_2^{0,1}(t)$ and $f_{5/2}^{0,1}(t), -4 \leqslant t \leqslant 4$

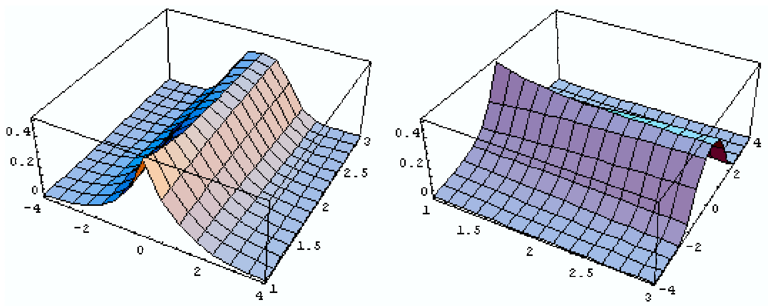

Figure 2: The graph of the function $f_k^{0,1}(t), -4 \leqslant t \leqslant 4, 1 < k \leqslant 3$

Integrate the $f_k^{\mu,\sigma}(t)$ on $(-\infty,+\infty)$ by substitutions, we obtain that for all $\mu \in (-\infty,\infty)$, $\sigma \in (0,\infty)$ and $k \in (0,\infty)$, we have

$$f_k^{\mu,\sigma}(t) > 0, \forall t \in (-\infty,\infty) \text{ and } \int_{-\infty}^{\infty} f_k^{\mu,\sigma}(t)\,\mathrm{d}t = 1. \tag{2.1}$$

If $\xi \sim N_k(\mu,\sigma)$, then we have

$$E\xi \triangleq \int_{-\infty}^{\infty} t f_k^{\mu,\sigma}(t)\,\mathrm{d}t = \mu. \tag{2.2}$$

Formula (2.2) still holds for all $k > 0$.

Lemma 2.1 If $\xi \sim N_k(\mu,\sigma)$, then we have

$$D\xi \triangleq E(\xi - E\xi)^2 = \frac{k^{2k^{-1}}\Gamma(3k^{-1})}{\Gamma(k^{-1})}\sigma^2 \begin{cases} > \sigma^2, & 1 < k < 2, \\ = \sigma^2, & k = 2, \\ < \sigma^2, & 2 < k < \infty. \end{cases} \tag{2.3}$$

Proof It's easy to obtain that

$$D\xi = \frac{k^{2k^{-1}}\Gamma(3k^{-1})}{\Gamma(k^{-1})}\sigma^2. \tag{2.4}$$

By the graph of the function $\omega(k)$ (depicted in Figure 3), we know that the function $\omega(k) = \frac{k^{-2k}\Gamma(3k)}{\Gamma(k)}$ is monotonically increasing. Hence the function$\omega_*(k) = \omega\left(\frac{1}{k}\right) = \frac{k^{2k^{-1}}\Gamma\left(3k^{-1}\right)}{\Gamma(k^{-1})}$ is monotonically decreasing. Note that $\omega_*(2) = 1$, we get

$$\frac{k^{2k^{-1}}\Gamma\left(3k^{-1}\right)}{\Gamma\left(k^{-1}\right)}\begin{cases} > 1, & 1 < k < 2, \\ = 1, & k = 2, \\ < 1, & 2 < k < \infty. \end{cases} \tag{2.5}$$

Using (2.4) and (2.5), we get our desired result (2.3).

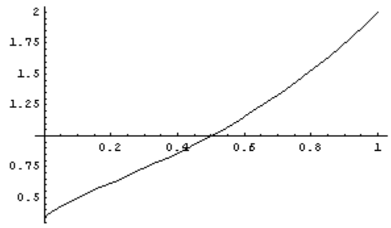

Figure 3: The graph of the function $\omega(k), 0 < k < 1$

According to the previous results, we find that k-normal distribution is a new distribution similar to but different from the normal distribution and the generalized normal distribution (see [5, 6]), it is also a natural generalization of the normal distribution, and it can be used to fit a number of empirical distributions with different skewness and kurtosis as well.

We remark here that k-normal distribution has similar but distinct form to the generalized normal distribution in [6]. By Definition 2.1, we know that $f_2^{\mu,\sigma}(t)$ is the p.d.f. of normal distribution $N(\mu,\sigma)$. But the p.d.f. for $s = 2$ (in [6]) is

$$\frac{1}{\sqrt{\pi}\sigma}\exp\left\{-\left(\frac{x-\mu}{\sigma}\right)^2\right\},$$

which does not match with normal distribution. So, to a certain extent, k-normal distribution is a better form of the generalized normal distribution.

3 Main Results

In this section, we will study the relationship among the variances of truncated variables. The main result of the paper is as follows.

Theorem 3.1 Let the p.d.f. $f : I \to (0,\infty)$ of the random variable ξ_I be differentiable, and let $D\xi_{I_*}, D\xi_{I^*}, D\xi_I$ be the variances of the truncated variables $\xi_{I_*}, \xi_{I^*}, \xi_I$, respectively. If

(i) $f : I \to (0,\infty)$ is a logarithmic concave function;

(ii) $\xi_{I_*} \subset \xi_I, \xi_{I^*} \subset \xi, I_* \subset I^*$, then we have the inequalities

$$D\xi_{I_*} < D\xi_{I^*} < D\xi_I. \tag{3.1}$$

Before prove Theorem 3.1, we first establish the following three lemmas.

Lemma 3.1 Let $\xi_I \in I$ be a continuous random variable, and let its p.d.f. be $f : I \to (0, \infty)$. If $\xi_{I_*} \subseteq \xi_I, \xi_{I^*} \subseteq \xi_I, I_* \subseteq I^*$, then we have

$$\xi_{I_*} \subseteq \xi_{I^*}; \tag{3.2}$$

if $\xi_{I_*} \subseteq \xi_I, \xi_{I^*} \subseteq \xi_I, I_* \subset I^*$, then we have

$$\xi_{I_*} \subset \xi_{I^*}. \tag{3.3}$$

Proof By virtue of the hypotheses, we get

$$f_{\xi_{I_*}} : I_* \to (0, \infty), f_{\xi_{I_*}}(t) \triangleq \frac{f(t)}{\displaystyle\int_{I_*} f} \quad \text{and} \quad f_{\xi_{I^*}} : I^* \to (0, \infty), f_{\xi_{I^*}}(t) \triangleq \frac{f(t)}{\displaystyle\int_{I^*} f},$$

thus

$$f_{\xi_{I_*}}(t) = \frac{f(t)\Big/\displaystyle\int_{I^*} f}{\displaystyle\int_{I_*}\left(f\Big/\displaystyle\int_{I^*} f\right)} = \frac{f_{\xi_{I^*}}(t)}{\displaystyle\int_{I_*} f_{\xi_{I^*}}}.$$

It follows therefore from the above facts and Definition 1.1 that we have

$$I_* \subseteq I^* \Rightarrow \xi_{I_*} \subseteq \xi_{I^*} \quad \text{and} \quad I_* \subset I^* \Rightarrow \xi_{I_*} \subset \xi_{I^*}.$$

Lemma 3.2 Let the function $f : I \to (0, \infty)$ be differentiable. If f is a logarithmic concave function, then we have

$$f(v) - \frac{f'(v)}{f(v)} \int_u^v f(t)\,\mathrm{d}t > 0, \forall (u, v) \in I^2. \tag{3.4}$$

Proof We define an auxiliary function F of the variables u and v as

$$F : I^2 \to (-\infty, \infty) \text{ by } F(u, v) \triangleq f(v) - [\log f(v)]' \int_u^v f(t)\,\mathrm{d}t.$$

If $v = u$, then we have $F(u, v) = f(u) - [\log f(u)]' \displaystyle\int_u^u f(t)\,\mathrm{d}t = f(u) > 0$.

By Cauchy mean value theorem, there exists a real number $\theta \in (0, 1)$ for $u \neq v$ such that

$$\frac{f(v) - f(u)}{\displaystyle\int_u^v f(t)\,\mathrm{d}t} = \frac{f'[u + \theta(v - u)]}{f[u + \theta(v - u)]} = \{\log f[u + \theta(v - u)]\}'. \tag{3.5}$$

If $u < v$, then we have

$$u < u + \theta(v - u) < v. \tag{3.6}$$

Combining (3.5) and (3.6), we obtain

$$\frac{F(u,v)-f(u)}{\displaystyle\int_u^v f(t)\,\mathrm{d}t}=\{\log f[u+\theta(v-u)]\}'-[\log f(v)]'\geqslant 0.$$

So $F(u,v)\geqslant f(u)>0$. This proves inequality (3.4) for $u<v$.

If $u>v$, then we have

$$v<u+\theta(v-u)<u. \tag{3.7}$$

Combining (3.5) and (3.7), we obtain

$$\frac{F(u,v)-f(u)}{\displaystyle\int_u^v f(t)\,\mathrm{d}t}=\{\log f[u+\theta(v-u)]\}'-[\log f(v)]'\leqslant 0.$$

Since $\displaystyle\int_u^v f(t)\,\mathrm{d}t<0$, we have $F(u,v)\geqslant f(u)>0$. So inequality (3.4) is also holds for the last case.

Lemma 3.3 Let the function $f:I\to(0,\infty)$ be differentiable. If f is a logarithmic concave function, then the function

$$G:I^2\to[0,\infty),G(u,v)\triangleq\begin{cases}\dfrac{\displaystyle\int_u^v t^2 f}{\displaystyle\int_u^v f}-\left(\dfrac{\displaystyle\int_u^v tf}{\displaystyle\int_u^v f}\right)^2, & u\neq v,\\ 0, & u=v\end{cases}$$

satisfies the following inequalities

$$\frac{\partial G(u,v)}{\partial v}\begin{cases}>0, & \forall(u,v)\in I^2,u<v,\\ <0, & \forall(u,v)\in I^2,u>v.\end{cases} \tag{3.8}$$

Proof For the convenience of notation, two real numbers with same sign α and β will be written as $\alpha\sim\beta$.

By the definition, we know that

$$c>0,\alpha\in(-\infty,\infty)\Rightarrow c\alpha\sim\alpha. \tag{3.9}$$

The power mean inequality asserts (see [10]) that

$$\left(\frac{\displaystyle\int_u^v |g|^\gamma f}{\displaystyle\int_u^v f}\right)^{\frac{1}{\gamma}}\geqslant\frac{\displaystyle\int_u^v |g|f}{\displaystyle\int_u^v f}\geqslant\left|\frac{\displaystyle\int_u^v gf}{\displaystyle\int_u^v f}\right|,\ \forall\gamma>1,$$

then we are easy to get

$$G(u,v)=G(v,u)\geqslant 0,\forall(u,v)\in I^2. \tag{3.10}$$

We first prove the case of $u < v$,

$$\begin{aligned}\frac{\partial G}{\partial v} &= \frac{f(v)}{\left(\int_u^v f\right)^2}\left[v^2\int_u^v f-\int_u^v t^2 f-2\int_u^v tf\left(v-\frac{\int_u^v tf}{\int_u^v f}\right)\right]\\ &\sim v^2\int_u^v f-\int_u^v t^2 f-2\int_u^v tf\left(v-\frac{\int_u^v tf}{\int_u^v f}\right),\end{aligned}$$

i.e.,

$$\frac{\partial G}{\partial v}\sim H(u,v), \tag{3.11}$$

where

$$H(u,v)\triangleq v^2\int_u^v f-\int_u^v t^2 f-2\int_u^v tf\left(v-\frac{\int_u^v tf}{\int_u^v f}\right). \tag{3.12}$$

It follows from (3.9) and (3.12) that

$$\frac{\partial H}{\partial v} \sim v\int_u^v f-v^2 f(v)-\left[\int_u^v f-2vf(v)\right]\frac{\int_u^v tf}{\int_u^v f}-f(v)\left(\frac{\int_u^v tf}{\int_u^v f}\right)^2,$$

i.e.,

$$\frac{\partial H}{\partial v}\sim H_*(u,v,w)\triangleq v\int_u^v f-v^2 f(v)-\left[\int_u^v f-2vf(v)\right]w-f(v)w^2, \tag{3.13}$$

where

$$u=\frac{\int_u^v uf}{\int_u^v f}<w=\frac{\int_u^v tf}{\int_u^v f}<\frac{\int_u^v vf}{\int_u^v f}=v. \tag{3.14}$$

Since

$$H_*(u,v,v)=v\int_u^v f-v^2 f(v)-\left[\int_u^v f-2vf(v)\right]v-f(v)v^2=0, \tag{3.15}$$

so by (3.9) and (3.15), we have

$$H_*(u,v,w) \sim \int_u^v f-2vf(v)+f(v)(v+w).$$

Hence

$$\frac{\partial H}{\partial v}\sim H_*(u,v,w)\sim H^*(u,v), \tag{3.16}$$

where

$$H^*(u,v) \triangleq \int_u^v f - vf(v) + f(v)\frac{\int_u^v tf}{\int_u^v f}. \tag{3.17}$$

Combining (3.9), (3.14), (3.17), $v > u$ with Lemma 3.2, we can do the straight calculation as follows

$$\begin{aligned}\frac{\partial H^*}{\partial v} &= \frac{f(v)}{\int_u^v f}\left(v - \frac{\int_u^v tf}{\int_u^v f}\right)\left[f(v) - \frac{f'(v)}{f(v)}\int_u^v f\right]\\ &\sim f(v) - \frac{f'(v)}{f(v)}\int_u^v f > 0.\end{aligned}$$

By (3.17) and $v > u$, we get

$$H^*(u,v) > H^*(u,u) = \lim_{v\to u}\left[\int_u^v f - vf(v) + f(v)\frac{\int_u^v tf}{\int_u^v f}\right] = 0. \tag{3.18}$$

By (3.16) and (3.18), we get

$$\frac{\partial H}{\partial v} > 0. \tag{3.19}$$

By (3.19) and $v > u$, we get

$$H(u,v) = v^2\int_u^v f - \int_u^v t^2 f - 2\int_u^v tf\left(v - \frac{\int_u^v tf}{\int_u^v f}\right) > H(u,u) = 0. \tag{3.20}$$

From (3.11) and (3.20), for the case of $v > u$, result (3.8) of Lemma 3.3 follows immediately.

Next, we prove the case of $u > v$. Based on the above analysis, we obtain the following relations

$$\begin{aligned}&\frac{\partial G}{\partial v}\sim H(u,v);\\ &v < w < u \Rightarrow \frac{\partial H}{\partial v} \sim H_*(u,v,w) \sim -H^*(u,v);\\ &v < w < u \Rightarrow \frac{\partial H^*}{\partial v} \sim f(v) - \frac{f'(v)}{f(v)}\int_u^v f > 0 \Rightarrow H^*(u,v) < 0;\\ &H^*(u,v) < 0 \Rightarrow \frac{\partial H}{\partial v} > 0 \Rightarrow H(u,v) < 0;\\ &\frac{\partial G}{\partial v}\sim H(u,v) < 0 \Rightarrow \frac{\partial G}{\partial v} < 0.\end{aligned}$$

Thus inequalities (3.8) still hold for $u > v$. This completes our proof.

Now we turn our attention to the proof of Theorem 3.1.

Proof Without loss of generality, we can assume that

$$I^* = [a, b], I = [\alpha, \beta], -\infty \leqslant \alpha \leqslant a < b \leqslant \beta \leqslant \infty.$$

Note that

$$I^* \subset I \Rightarrow \alpha \leqslant a < b < \beta \text{ or } \alpha < a < b \leqslant \beta.$$

If $\alpha \leqslant a < b < \beta$, so according to (1.2), (3.10) and Lemma 3.3, we get

$$D\xi_{I^*} = G(a, b) < G(a, \beta) = G(\beta, a) \leqslant G(\beta, \alpha) = G(\alpha, \beta) = D\xi_I,$$

hence

$$D\xi_{I^*} < D\xi_I. \tag{3.21}$$

If $\alpha < a < b \leqslant \beta$, so, according to (1.2), (3.10) and Lemma 3.3, we get

$$D\xi_{I^*} = G(a, b) \leqslant G(a, \beta) = G(\beta, a) < G(\beta, \alpha) = G(\alpha, \beta) = D\xi_I.$$

That is to say, inequality (3.21) still holds.

By Lemma 3.1, we have $\xi_{I_*} \subset \xi_I, \xi_{I^*} \subset \xi_I, I_* \subset I^* \Rightarrow \xi_{I_*} \subset \xi_{I^*}$. Using inequality (3.21) for ξ_{I_*}, ξ_{I^*}, we can obtain

$$D\xi_{I_*} < D\xi_{I^*}. \tag{3.22}$$

Combining inequalities (3.21) and (3.22), we get inequalities (3.1).

This completes the proof of Theorem 3.1.

From Theorem 3.1 we know that if the probability density function of the random variable ξ_I is differentiable and log concave, and ξ_{I_*} is the proper truncated variables of the random variable ξ_{I^*}, the variance of ξ_{I_*} is less than the variance of ξ_{I^*}. This result is of great significance in the hierarchical teaching model, see the next theorem.

For the convenience of use, Theorem 3.1 can be slightly generalized as follows.

Theorem 3.2 Let $\varphi : I \to (-\infty, \infty)$ and $f : I \to (0, \infty)$ be differentiable functions, where f be the p.d.f. of the random variable ξ_I, and let $D\varphi(\xi_{I_*}), D\varphi(\xi_{I^*})$ with $D\varphi(\xi_I)$ be the variances of the truncated variables $\varphi(\xi_{I_*}), \varphi(\xi_{I_*})$ with $\varphi(\xi_I)$, respectively. If

(i) $\varphi'(t) > 0, \forall t \in I$;

(ii) the function $(f \circ \varphi^{-1})(\varphi^{-1})' : \varphi(I) \to (0, \infty)$ is log concave;

(iii) $\xi_{I_*} \subset \xi_I, \xi_{I^*} \subset \xi_I, I_* \subset I^*$,

then we have the following inequalities

$$D\varphi(\xi_{I_*}) < D\varphi(\xi_{I^*}) < D\varphi(\xi_I). \tag{3.23}$$

Proof Set $\overline{\xi} = \varphi(\xi), \overline{f} = (f \circ \varphi^{-1})(\varphi^{-1})'$. By condition (i), we can see that $\xi = \varphi^{-1}(\overline{\xi}), \overline{f} = (f \circ \varphi^{-1})(\varphi^{-1})' > 0$ and

$$\int_{\varphi(I)} \overline{f} = \int_{\varphi(I)} \left(f \circ \varphi^{-1}\right)(t) \left[\varphi^{-1}(t)\right]' \mathrm{d}t = \int_{\varphi(I)} f\left[\varphi^{-1}(t)\right] \mathrm{d}\varphi^{-1}(t) = \int_I f = 1.$$

Thus $\overline{f}:\varphi(I)\to(0,\infty)$ is a p.d.f. of the random variable $\overline{\xi}$.

By condition (ii), we can see that $\overline{f}$ is a logarithmic concave function. Combining conditions (i) and (iii) with Lemma 3.1, we have

$$I_*\subset I^*\subset I\Rightarrow\varphi(I_*)\subset\varphi(I^*)\subset\varphi(I)\Rightarrow\overline{\xi}_{\varphi(I_*)}\subset\overline{\xi}_{\varphi(I^*)}\subset\overline{\xi}_{\varphi(I)}.$$

We can deduce from Theorem 3.1 that the following is true

$$D\overline{\xi}_{\varphi(I_*)}=D\varphi(\xi_{I_*})<D\overline{\xi}_{\varphi(I^*)}=D\varphi(\xi_{I^*})<D\overline{\xi}_{\varphi(I)}=D\varphi(\xi_I).$$

Thus inequalities (3.23) is valid.

4 Applications

In the hierarchical teaching model, the math score of the students of some grade in a university is a random variable ξ_I, where $I=[0,100),\xi_I\subset\xi,\ \xi\in(-\infty,\infty)$. By using the central limit theorem (see [8]), we know that ξ follows a normal distribution, that is, $\xi\sim N_2(\mu,\sigma)$. If, in the grade, the top students and poor students are few, that is to say, the variance $D\xi$ of the random variable ξ is small, according to Figure 1 and Figure 2 with Lemma 2.1, we believe that there is a real number $k\in[2,\infty)$ such that $\xi\sim N_k(\mu,\sigma)$. Otherwise, there is a real number $k\in(1,2)$ such that $\xi\sim N_k(\mu,\sigma)$. Then the k, σ of $N_k(\mu,\sigma)$ can be determined according to [5].

We have collected three real data sets $X1$, $X2$ and $X3$, which are all math test score of the students from the unhierarchical, the first level (superior) and the second level (poor) classes, containing 263, 149 and 145 records, respectively. For further analyzing the data, we first estimate parameters k, μ, σ of $N_k(\mu,\sigma)$, then draw probability density function of $N_k(\mu,\sigma)$ and frequency histogram of the corresponding data set in the same coordinate system, which also contains the probability density function curve graph of normal distribution. After that, we obtain three graphs for $X1$, $X2$ and $X3$, respectively (see Figure 4, Figure 5 and Figure 6 in Appendix B). These three figures show that k-normal distribution is superior to normal distribution since kurtosis is bigger and variance is smaller.

Further more, as shown in the histograms, the variance of $X1$, $X2$ and $X3$ is decreasing. By observing the proportion of scores less than 60 of $X1$, $X2$ and $X3$, we find that the hierarchical teaching model bring better results, and that the second category (represented by $X3$) classes receive more significant benefits from this teaching model.

According to Theorem 3.1 and Lemma 2.1, we have

Theorem 4.1 In the hierarchical teaching model, if $\xi\sim N_k(\mu,\sigma)$, where $k>1$, then for all $i,n:1\leqslant i\leqslant n-1,n\geqslant3$, we have

$$D\xi_{I_*}<D\xi_{I^*}<D\xi_I<D\xi=\frac{k^{2k^{-1}}\Gamma(3k^{-1})}{\Gamma(k^{-1})}\sigma^2,\tag{4.1}$$

where

$$I_*=[a_i,a_{i+1})\ \text{or}\ [a_{i+1},a_{i+2}),I^*=[a_i,a_{i+2}),I=[0,100).$$

We accomplish simulation analysis about Theorem 3.1. The procedure of simulation design is shown in Appendix A. The results of the simulation are listed in the tables (see Tables 1–4 in Appendix A). By comparing the data in these tables, we find that, no matter how to change the parameters k, μ or σ, the variance of truncated variable is strictly less than that of untruncated variable. For example, for any k, μ or σ as shown in Tables 1–4,

$$D\xi_{[0,60)} < D\xi_{[0,80)} < D\xi_{(-\infty,\infty)},\ \ D\xi_{[60,80)} < D\xi_{[60,100)} < D\xi_{(-\infty,\infty)},$$

this does verify the truth of Theorem 3.

From Tables 1 and 3, we see that for each σ and $I \subset (-\infty,\infty)$, if

$$\xi_1 \sim N_3(\mu,\sigma), \xi_2 \sim N_2(\mu,\sigma), \xi_3 \sim N_{1.5}(\mu,\sigma),$$

then $D\xi_{1I} < D\xi_{2I} < D\xi_{3I}$. From Tables 2 and 4, for each μ and $I \subset (-\infty,\infty)$, if

$$\eta_1 \sim N_3(\mu,\sigma), \eta_2 \sim N_2(\mu,\sigma), \eta_3 \sim N_{1.5}(\mu,\sigma),$$

then $D\eta_{1I} < D\eta_{2I} < D\eta_{3I}$. The truth of Theorem 3.1 is verified.

Actually in appendix, the data set X1 is the math test score of unhierarchical students, X2 and X3 are math test score of hierarchical students. We have figured out their variances

$$D(X1) = 254.2813,\ \ D(X2) = 172.8042,\ \ D(X3) = 161.0640.$$

The facts $D(X3) < D(X1)$ and $D(X2) < D(X1)$, just show that the hierarchical teaching is more efficiency than unhierarchical teaching.

References

[1] Yao Hui, Dai Yong, Xie Lin. Pareto-geometric distribution[J]. J. Math., 2012, 32(2): 339–351.

[2] Deng Yuhui. Probablity distribution of sample spacing[J]. J. Math., 2004, 24(6): 685–689.

[3] Yang Chaofeng, Pu Yingjuan. Bayes analysis of hierarchical teaching[J]. Math. Prac. The. (in Chinese), 2004, 34(9): 107–113.

[4] Han Tianyong, Wen Jiajin. Normal distribution and associated teaching efficiency[J]. Math. Prac. The. (in Chinese). 2014, 44(6): 183–193.

[5] Armando D, Graciela G, Ramon M. A practical procedure to estimate the shape parameter in the generalized Gaussian distribution, technique report[OL]. Available: http://www.cimat.mx/reportes/enlinea/I-01-18_eng.pdf, 2001.

[6] Saralees N. A generalized normal distribution[J]. J. Appl. Stat. 2008, 32(7): 685–694.

[7] Wen Jianjin, Han Tianyong, Cheng S S. Quasi-log concavity conjecture and its applications in statistics[J]. J. Inequal. Appl., 2014, DOI: 10.1186/1029-242X-2014-339.

[8] Johnson O. Information theory and the central limit theorem[M]. London: Imperial College Press, 2004.

[9] Wlodzimierz B. The normal distribution: characterizations with applications[M]. New York: Springer-Verlag, 1995.

[10] Wang Wanlan. Approaches to prove inequalities (in Chinese)[M] Harbin: Harbin Institute of Technology Press, 2011.

[11] Tong T L. An adaptive solution to ranking and selection problems[J]. Ann. Stat., 1978, 6(3): 658–672.

[12] Bagnoli M, Bergstrom T. Log-concave probability and its applications[J]. Econ. The., 2005, 26(2): 445–469.

k-正态分布及其应用

韩天勇[1], 文家金[1], 宋安超[2], 叶建华[1]

(1.成都大学信息科学与工程学院, 四川 成都 610106)

(2.西南财经大学统计学院, 四川 成都 611130)

摘要: 近本文研究了截断随机变量和k-正态分布. 利用对数凹函数理论, 获得了涉及截断随机变量和截断随机变量的函数的方差的不等式链, 推广了涉及正态分布和分层教学模型的一些经典结论. 同时在附录部分给出了仿真结果.

关键词: 截断随机变量; k-正态分布; 分层教学模型; 对数凹函数; 仿真

MR(2010)**主题分类号:** 62J10; 62P25; 60E05; 60E15; 26D15; 26E60

中图分类号: O174.13; O211.3; O211.5

Appendix

A The Simulation and Comparison of Variances of Truncated k-Normal Variable

The procedure of simulation design is as follows

Step 1 Choose the appropriate parameter k, μ and σ in the distribution $N_k(\mu,\sigma)$;

Step 2 Generate 200 random numbers obeying the distribution $\xi \sim N_k(\mu,\sigma)$;

Step 3 Use the 200 numbers to calculate the variance for six truncated k-normal variables $\xi_{(-\infty,\infty)}$, $\xi_{[0,60)}$, $\xi_{[60,80)}$, $\xi_{[80,100)}$, $\xi_{[0,80)}$ and $\xi_{[60,100)}$;

Step 4 Repeat Step 1 and Step 2 for 50 times;

Step 5 Calculate the mean of 50 variances for each truncated k-normal variable, denoted by $D\xi_{(-\infty,\infty)}$, $D\xi_{[0,60)}$, $D\xi_{[60,80)}$, $D\xi_{[80,100)}$, $D\xi_{[0,80)}$ and $D\xi_{[60,100)}$ respectively ;

Step 6 Change the value of k, μ and σ, and repeat Step 1, Step 2, Step 3, Step 4. All the results are listed in Tables 1–4 (NaN indicates there is no random number for corresponding truncated variable).

Table 1: $k=3$, σ=10

μ \ var	$D\xi_{(-\infty,+\infty)}$	$D\xi_{[0,60)}$	$D\xi_{[60,80)}$	$D\xi_{[80,100)}$	$D\xi_{[0,80)}$	$D\xi_{[60,100)}$
70	76.5217	9.5024	27.1810	8.6149	50.9796	50.5629
75	76.9029	5.0242	26.6868	16.2052	38.6029	63.2846
80	77.2020	NaN	23.6349	24.2553	26.4003	71.7841

Table 2: $k=3$, μ=75

var / σ	$D\xi_{(-\infty,+\infty)}$	$D\xi_{[0,60)}$	$D\xi_{[60,80)}$	$D\xi_{[80,100)}$	$D\xi_{[0,80)}$	$D\xi_{[60,100)}$
6	27.2043	NaN	16.9083	4.1962	17.0017	27.1108
8	49.64	NaN	24.20	9.14	27.05	46.5981
13	130.55	13.16	29.18	27.78	60.37	87.9405

Table 3: $k=1.5$, σ=10

var / μ	$D\xi_{(-\infty,+\infty)}$	$D\xi_{[0,60)}$	$D\xi_{[60,80)}$	$D\xi_{[80,100)}$	$D\xi_{[0,80)}$	$D\xi_{[60,100)}$
70	127.6834	38.651	27.5998	23.7547	75.811	68.07633
75	126.641	32.5665	27.0468	24.73	63.3965	74.69793
80	125.2832	24.7078	26.1625	25.9007	50.8872	77.04843

Table 4: $k=1.5$, μ=75

var / σ	$D\xi_{(-\infty,+\infty)}$	$D\xi_{[0,60)}$	$D\xi_{[60,80)}$	$D\xi_{[80,100)}$	$D\xi_{[0,80)}$	$D\xi_{[60,100)}$
6	44.7223	NaN	20.4474	12.8553	25.2661	39.75123
8	83.019	20.067	24.9876	20.8246	42.5282	59.61493
13	209.1055	61.4701	28.6449	27.5559	101.9277	89.09083

B Curve Fitting for Three Real Data Sets $X1$, $X2$ and $X3$

The results of curve fitting for three real data sets are as follows (see Figure 4–6)

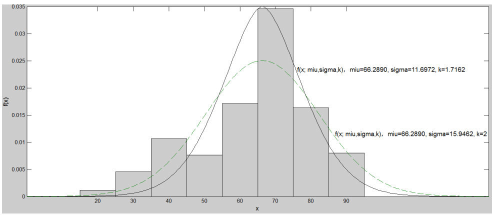

Figure 4: Fitting $X1$

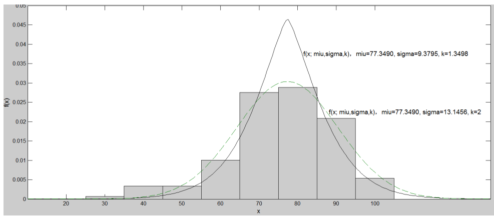

Figure 5: Fitting $X2$

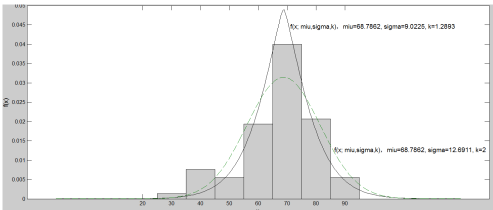

Figure 6: Fitting $X3$