Statistics and Application 统计学与应用, **2026, 15(3), 92-101**
Published Online March 2026 in Hans. https://www.hanspub.org/journal/sa
https://doi.org/10.12677/sa.2026.153058

Hans汉斯

“指数分布抽样基本定理”之讨论

陈宣戈[1]，顾蓓青[1]，雷　平[2]，徐晓岭[1]

[1]上海对外经贸大学统计与数据科学学院，上海
[2]上海对外经贸大学国际经贸学院，上海

收稿日期：2026年2月13日；录用日期：2026年3月6日；发布日期：2026年3月17日

摘　要

论文针对李国安老师提出的“指数分布抽样基本定理”进行了讨论分析，简化了原文献中结论的相关证明，并修正了原文献中一些错误。论文最后通过两参数指数分布与三参数威布尔分布总体说明“指数分布抽样分布基本定理”的应用。

关键词

指数分布，次序统计量，抽样基本定理

Discussion on “Basic Theorem of Exponential Distribution Sampling”

Xuange Chen[1], Beiqing Gu[1], Ping Lei[2], Xiaoling Xu[1]

[1]School of Statistics and Data Science, Shanghai University of International Business and Economics, Shanghai
[2]School of International Business, Shanghai University of International Business and Economics, Shanghai

Received: February 13, 2026; accepted: March 6, 2026; published: March 17, 2026

Abstract

This paper discusses and analyzes the “Basic Theorem of Exponential Distribution Sampling” proposed by Professor Li Guoan. It simplifies the relevant proofs of the conclusions in original literature and corrects some errors in this literature. Finally, the paper illustrates the application of the theorem using the two-parameter exponential distribution and the three-parameter Weibull distribution populations.

文章引用: 陈宣戈, 顾蓓青, 雷平, 徐晓岭. “指数分布抽样基本定理”之讨论[J]. 统计学与应用, 2026, 15(3): 92-101.
DOI: 10.12677/sa.2026.153058

Keywords

Exponential Distribution, Order Statistic, Basic Theorem of Sampling

1. 引言

李国安老师在文献[1]-[3]中提出并证明了所谓的“指数分布抽样基本定理”，另外还在文献[4]中又增加了两个结论。事实上，文献[1]-[4]中所言的“指数分布抽样基本定理”，其实早在几十年前都已经被人发现，而且更为全面，具体可查阅文献[5]-[7]，这些结论也并已写入各类可靠性数学、可靠性统计的教科书中，例如文献[8]。国内最早写入概率论与数理统计教科书中的是魏宗舒先生，见文献[9]中的第 9 章第 4 节。文献[10]将更加全面的结论写入了概率论与数理统计教材。由此可以看到文献[1]-[4]仅仅是将原有的部分结果给予的一个名称而已(称其为“指数分布抽样基本定理”)，文献[1]-[3]中关于“指数分布抽样基本定理”的证明都是没必要给出的，再者，文献[4]中所谓增加的两个结果中关于 $X_{(n)}-\bar{X}$ 的密度函数是错误的，并由此其定理 3.3 中涉及 $X_{(n)}-\bar{X}$ 的数学期望与方差也都是错的。需要指出的是，其所谓的“指数分布抽样基本定理”是不全面的，最为核心的部分没有给出，并由此造成文献[4]中的许多推导过于复杂，影响论文的可读性。另外，指数分布有许多特别的性质，具体可查阅文献[11] [12]。

2. 指数分布总体次序统计量的分布

设总体 X 服从参数为 θ 的指数分布 $\mathrm{Exp}(1/\theta)$，其分布函数与密度函数分别为：

$$F(x)=1-\mathrm{e}^{-x/\theta},\quad f(x)=\frac{1}{\theta}\mathrm{e}^{-x/\theta},x>0,\theta>0$$

而 $X_1,X_2,\cdots,X_n$ 是来自总体 X 的容量为 n 的一个简单随机样本，其次序统计量记为 $X_{(1)}\le X_{(2)}\le\cdots\le X_{(n)}$。

关于指数分布总体次序统计量的分布，最核心的结论是如下定理 1 与定理 2，证明可见文献[5]-[11]，将定理 1 与定理 2 称为“指数分布抽样基本定理”比较合适，因为这两个定理是指数分布统计推断的基础。而利用定理 1 与定理 2 可以深入研究涉及指数分布的诸多问题，具体可见文献[13]-[20]。

定理 1： 设 $X_{(1)}\le X_{(2)}\le\cdots\le X_{(n)}$ 是来自总体 $X\sim\mathrm{Exp}(1/\theta)$ 的一个容量为 n 的前 n 个次序统计量，记 $Y_i=\dfrac{(n-i+1)\left(X_{(i)}-X_{(i-1)}\right)}{\theta},i=1,2,\cdots,n$，其中 $X_{(0)}=0$，则 $Y_1,Y_2,\cdots,Y_n$ 相互独立，且同服从标准指数分布 $\mathrm{Exp}(1)$。

定理 2： 设总体 $X\sim\mathrm{Exp}(1/\theta)$，$X_{(1)}\le X_{(2)}\le\cdots\le X_{(n)}$ 是来自总体 X 的容量为 n 的前 n 个次序统计量，则 $X_{(k+1)}-X_{(k)}$，$X_{(k+2)}-X_{(k)}$，$\cdots$，$X_{(n)}-X_{(k)}$ 是来自指数分布总体 $\mathrm{Exp}(1/\theta)$ 样本容量为 $n-k$ 的前 $n-k$ 个次序统计量。

由文献[10] [11]可知如下定理 3：

定理 3： 设 $X_{(1)}\le X_{(2)}\le\cdots\le X_{(n)}$ 是来自总体 $X\sim\mathrm{Exp}(1/\theta)$ 的一个容量为 n 的前 n 个次序统计量，则(1) $X_{(k)}-X_{(k-1)},k=1,2,\cdots,n$ (其中 $X_{(0)}=0$)的数学期望与方差为：

$$E\left(X_{(k)}-X_{(k-1)}\right)=\frac{\theta}{n-k+1}，\quad D\left(X_{(k)}-X_{(k-1)}\right)=\frac{\theta^2}{\left(n-k+1\right)^2}$$

(2) $X_{(1)}$ 与 $X_{(k)}-X_{(1)},k=2,3,\cdots,n$ 独立，且

$$E\left(X_{(k)}\right)=\theta\sum_{i=1}^{k}\frac{1}{n-i+1}，\quad D\left(X_{(k)}\right)=\theta^2\sum_{i=1}^{k}\frac{1}{\left(n-i+1\right)^2}$$

对 $1\le j\le k\le n$ 有： $\operatorname{cov}\left(X_{(j)},X_{(k)}\right)=D\left(X_{(j)}\right)=\theta^2\sum_{i=1}^{j}\frac{1}{\left(n-i+1\right)^2}$

注：文献[4]中由于没有很好地利用定理 3 的结论，所以其论文中的定理 3.1 和定理 3.2 的计算比较复杂。

文献[1]-[3]给出所谓的“指数分布抽样基本定理”见如下定理 4：

定理 4：设 $X_{(1)}\le X_{(2)}\le\cdots\le X_{(n)}$ 是来自总体 $X\sim\operatorname{Exp}(1/\theta)$ 的一个容量为 n 的前 n 个次序统计量，则(1) $\frac{2nX_{(1)}}{\theta}\sim\chi^2(2)$；(2) $\frac{2n\left(\bar{X}-X_{(1)}\right)}{\theta}\sim\chi^2\left(2(n-1)\right)$；(3) $\bar{X}-X_{(1)}$ 与 $X_{(1)}$ 相互独立。

证明：文献[1]-[3]都给出了定理 3 的证明，下面利用定理 1 给出证明。事实上，

$$\sum_{i=1}^{n}X_i=\sum_{i=1}^{n}X_{(i)}=\sum_{i=1}^{n}(n-i+1)\left(X_{(i)}-X_{(i-1)}\right)$$

$$\frac{2n\left(\bar{X}-X_{(1)}\right)}{\theta}=\frac{2}{\theta}\left(\sum_{i=1}^{n}X_{(i)}-nX_{(1)}\right)=\frac{2}{\theta}\sum_{i=2}^{n}(n-i+1)\left(X_{(i)}-X_{(i-1)}\right)=\sum_{i=2}^{n}Y_i\sim\chi^2\left(2(n-1)\right)$$

又 $2Y_1=\frac{2nX_{(1)}}{\theta}\sim\chi^2(2)$，于是 $X_{(1)}$ 与 $\bar{X}-X_{(1)}$ 相互独立。

文献[4]又将上述的“指数分布抽样基本定理”补充了两个结论，即如下定理 5：

定理 5：设 $X_{(1)}\le X_{(2)}\le\cdots\le X_{(n)}$ 是来自总体 $X\sim\operatorname{Exp}(1/\theta)$ 的一个容量为 n 的前 n 个次序统计量，则(1) $\frac{2nX_{(1)}}{\theta}\sim\chi^2(2)$， $\frac{2n\left(\bar{X}-X_{(1)}\right)}{\theta}\sim\chi^2\left(2(n-1)\right)$， $\bar{X}-X_{(1)}$ 与 $X_{(1)}$ 相互独立；

(2) $-2(n-1)\ln\left[1-\exp\left(-\frac{X_{(n)}-X_{(1)}}{\theta}\right)\right]\sim\chi^2(2)$，且 $X_{(1)}$ 与 $X_{(n)}-X_{(1)}$ 相互独立；(3) $X_{(1)}$ 与 $X_{(n)}-\bar{X}$ 相互独立，且 $X_{(n)}-\bar{X}$ 的密度函数为：

$$f_{X_{(n)}-\bar{X}}(z)=\sum_{k=0}^{n-2}(-1)^k C_{n-1}^{k}\frac{(n-1-k)^{n-2}(n-1)}{n^{n-2}\theta}\exp\left[-\frac{(n-1)(k+1)z}{(n-1-k)\theta}\right],z>0$$

注：值得指出的是，关于定理 5 文献[4]给出的证明比较复杂，而且所给出的 $X_{(n)}-\bar{X}$ 的密度函数是错误的，又由于文献[4]的定理 3.3 中利用了 $X_{(n)}-\bar{X}$ 的密度函数求取 $X_{(n)}-\bar{X}$ 的数学期望与方差，所以其得到的 $X_{(n)}-\bar{X}$ 的数学期望与方差也是错误的。

下面首先给出定理 5 (2)及 $X_{(1)}$ 与 $X_{(n)}-\bar{X}$ 相互独立的证明。事实上，

由于 $X_{(n)}-X_{(1)}=\sum_{i=2}^{n}\left(X_{(i)}-X_{(i-1)}\right)=\theta\sum_{i=2}^{n}\frac{1}{n-i+1}\frac{(n-i+1)\left(X_{(i)}-X_{(i-1)}\right)}{\theta}=\theta\sum_{i=2}^{n}\frac{Y_i}{n-i+1}$

则 $X_{(1)}$ 与 $X_{(n)}-X_{(1)}$ 相互独立。

记 $D_n=X_{(n)}-X_{(1)}$， 由文献[10]可知 D_n 的密度函数为：

$$f_{D_n}(y)=\frac{n-1}{\theta}\left(1-\mathrm{e}^{-y/\theta}\right)^{n-2}\mathrm{e}^{-y/\theta},y>0$$

进而 D_n 的分布函数为：$F_{D_n}(y)=\left(1-\mathrm{e}^{-y/\theta}\right)^{n-1},y>0$

由此 $-2\ln F_{D_n}(D_n)\sim\chi^2(2)$，即 $-2(n-1)\ln\left[1-\exp\left(-\frac{X_{(n)}-X_{(1)}}{\theta}\right)\right]\sim\chi^2(2)$

又 $X_{(n)}-\bar{X}=\left(X_{(n)}-X_{(1)}\right)-\left(\bar{X}-X_{(1)}\right)=\sum_{i=2}^{n}\left(X_{(i)}-X_{(i-1)}\right)-\frac{1}{n}\sum_{i=2}^{n}(n-i+1)\left(X_{(i)}-X_{(i-1)}\right)$

$$=\sum_{i=2}^{n}\left(1-\frac{n-i+1}{n}\right)\left(X_{(i)}-X_{(i-1)}\right)=\theta\sum_{i=2}^{n}\left(1-\frac{n-i+1}{n}\right)\frac{1}{n-i+1}\frac{(n-i+1)\left(X_{(i)}-X_{(i-1)}\right)}{\theta}$$

$$=\theta\sum_{i=2}^{n}\left(1-\frac{n-i+1}{n}\right)\frac{1}{n-i+1}Y_i$$

则 $X_{(1)}$ 与 $X_{(n)}-\bar{X}$ 相互独立。

下面说明定理 5 中关于 $X_{(n)}-\bar{X}$ 的密度函数表达式是不正确的。

不失一般性，假设 $\theta=1$，记 $a_i=\left(1-\frac{n-i+1}{n}\right)\frac{1}{n-i+1},i=2,3,\cdots,n$，并记 $Z_i=a_iY_i$，$i=2,3,\cdots,n$，$Z=X_{(3)}-\bar{X}=\sum_{i=2}^{3}Z_i$，由于 $Y_2,Y_3,\cdots,Y_n$ 相互独立，且均服从标准指数分布 $\mathrm{Exp}(1)$，进而 $Z_2,Z_3,\cdots,Z_n$ 也相互独立，且 $Z_i\sim\mathrm{Exp}(1/a_i),i=2,3,\cdots,n$

当 $n=3$ 时，$a_2=\frac{1}{2}-\frac{1}{3}=\frac{1}{6},a_3=1-\frac{1}{3}=\frac{2}{3}$，此时，对 $z>0$

$$\begin{aligned}F_Z(z)&=P(Z_2+Z_3\le z)=\int_0^z\frac{1}{a_2}\mathrm{e}^{-z_2/a_2}\mathrm{d}z_2\int_0^{z-z_2}\frac{1}{a_3}\mathrm{e}^{-z_3/a_3}\mathrm{d}z_3\\&=\int_0^z\frac{1}{a_2}\mathrm{e}^{-z_2/a_2}\left[1-\mathrm{e}^{-(z-z_2)/a_3}\right]\mathrm{d}z_2\\&=\int_0^z\frac{1}{a_2}\mathrm{e}^{-z_2/a_2}\mathrm{d}z_2-\mathrm{e}^{-z/a_3}\int_0^z\frac{1}{a_2}\mathrm{e}^{-(1/a_2-1/a_3)z_2}\mathrm{d}z_2\\&=1-\mathrm{e}^{-z/a_2}-\frac{1}{a_2}\mathrm{e}^{-z/a_3}\frac{1}{1/a_2-1/a_3}\int_0^z\left(\frac{1}{a_2}-\frac{1}{a_3}\right)\mathrm{e}^{-(1/a_2-1/a_3)z_2}\mathrm{d}z_2\\&=1-\mathrm{e}^{-z/a_2}-\frac{1}{a_2}\mathrm{e}^{-z/a_3}\frac{a_2a_3}{a_3-a_2}\left[1-\mathrm{e}^{-(1/a_2-1/a_3)z}\right]\\&=1-\mathrm{e}^{-z/a_2}-\frac{a_3}{a_3-a_2}\mathrm{e}^{-z/a_3}+\frac{a_3}{a_3-a_2}\mathrm{e}^{-z/a_2}\\&=1+\frac{a_2}{a_3-a_2}\mathrm{e}^{-z/a_2}-\frac{a_3}{a_3-a_2}\mathrm{e}^{-z/a_3}\end{aligned}$$

$$f_Z(z)=-\frac{1}{a_3-a_2}\mathrm{e}^{-z/a_2}+\frac{1}{a_3-a_2}\mathrm{e}^{-z/a_3}=2\mathrm{e}^{-3z/2}-2\mathrm{e}^{-6z}$$

而文献[4]中给出了 $n=3$ 时 $X_{(3)}-\bar{X}$ 的密度函数为 $f_Z(z)=\frac{4}{3}\mathrm{e}^{-z}-\frac{4}{3}\mathrm{e}^{-4z}$，这显然是不正确的。下面首先给出参数不等的相互独立的指数分布之和的密度函数，其次可得到 $X_{(n)}-\bar{X}$ 的密度函数。

定理 6 [11]：设随机变量 $Y_1,Y_2,\cdots,Y_n$ 相互独立，且 $Y_i\sim\mathrm{Exp}(1/\theta_i),i=1,2,\cdots,n$，$\theta_1,\theta_2,\cdots,\theta_n$ 都不相等，

记 $Z=\sum_{i=1}^{n}Y_i$， $C_{i,n}(\theta_1,\cdots,\theta_n)=\prod_{k\neq i}^{n}\left(1-\frac{\theta_k}{\theta_i}\right)^{-1}$，则 Z 的分布函数为：

$$F_Z(z)=1-\sum_{i=1}^{n}C_{i,n}(\theta_1,\cdots,\theta_n)\mathrm{e}^{-z/\theta_i},z>0$$

定理 7：设 $X_{(1)}\le X_{(2)}\le\cdots\le X_{(n)}$ 是来自总体 $X\sim\mathrm{Exp}(1/\theta)$ 的一个容量为 n 的前 n 个次序统计量，则 $X_{(n)}-\bar{X}$ 的密度函数为：

$$f_{X_{(n)}-\bar{X}}(z)=\frac{1}{\theta}\sum_{i=1}^{n-1}\frac{1}{\theta_i}C_{i,n-1}(\theta_1,\cdots,\theta_{n-1})\mathrm{e}^{-z/(\theta\theta_i)},z>0$$

其中 $\theta_i=b_{i+1}^{-1},i=1,2,\cdots,n-1$， $b_j^{-1}=\left(1-\frac{n-j+1}{n}\right)\frac{1}{n-j+1},j=2,3,\cdots,n$

$$C_{i,n-1}(\theta_1,\cdots,\theta_{n-1})=\prod_{k\neq i}^{n-1}\left(1-\frac{\theta_k}{\theta_i}\right)^{-1},i=1,2,\cdots,n-1$$

证明：记 $Y_i=\frac{(n-i+1)\left(X_{(i)}-X_{(i-1)}\right)}{\theta},b_i^{-1}=\left(1-\frac{n-i+1}{n}\right)\frac{1}{n-i+1},i=2,3,\cdots,n$

则
$$\frac{X_{(n)}-\bar{X}}{\theta}=\sum_{i=2}^{n}\left(1-\frac{n-i+1}{n}\right)\frac{1}{n-i+1}Y_i=\sum_{i=2}^{n}\frac{Y_i}{b_i}$$

记 $Z_i=\frac{Y_{i+1}}{b_{i+1}},\theta_i=b_{i+1}^{-1},i=1,2,\cdots,n-1$，则 $Z_1,Z_2,\cdots,Z_{n-1}$ 相互独立，且

$$Z_i\sim\mathrm{Exp}(1/\theta_i),i=1,2,\cdots,n$$

令 $Z=\frac{X_{(n)}-\bar{X}}{\theta}=\sum_{i=1}^{n-1}Z_i$，则由定理 6 可知 Z 的密度函数为：

$$f_Z(z)=\sum_{i=1}^{n-1}\frac{1}{\theta_i}C_{i,n-1}(\theta_1,\cdots,\theta_{n-1})\mathrm{e}^{-z/\theta_i},z>0$$

进而 $X_{(n)}-\bar{X}$ 的密度函数为： $f_{X_{(n)}-\bar{X}}(z)=\frac{1}{\theta}\sum_{i=1}^{n-1}\frac{1}{\theta_i}C_{i,n-1}(\theta_1,\cdots,\theta_{n-1})\mathrm{e}^{-z/(\theta\theta_i)},z>0$

3. 应用

例 1：设 $X_{(1)},X_{(2)},\cdots,X_{(r)}$ 为来自两参数指数分布总体 $X\sim\mathrm{Exp}(\mu,1/\theta)$ 的容量为 n 的前 r 个次序统计量， X 的分布函数 $F(x)$ 与密度函数分别为：

$$F(x)=1-\exp\left(-\frac{x-\mu}{\theta}\right),f(x)=\frac{1}{\theta}\exp\left(-\frac{x-\mu}{\theta}\right),x\ge\mu,0\le\mu<+\infty,\theta>0$$

(1) 求参数 μ,θ 的极大似然估计，并考察估计量的性质；(2) 求参数 μ,θ 的置信水平 $1-\alpha$ 的区间估计。

解：记次序统计量为 $X_{(1)}\le X_{(2)}\le\cdots\le X_{(r)}$，次序观察值为 $x_{(1)}\le x_{(2)}\le\cdots\le x_{(r)}$.

(1) 似然函数为： $L(\mu,\theta)=\frac{n!}{(n-r)!}\prod_{i=1}^{r}\left[\frac{1}{\theta}\exp\left(-\frac{x_{(i)}-\mu}{\theta}\right)\right]\cdot\exp\left[-(n-r)\frac{x_{(r)}-\mu}{\theta}\right]$

$$=\frac{n!}{(n-r)!}\frac{1}{\theta^r}\exp\left\{-\frac{1}{\theta}\left[\sum_{i=1}^{r}\left(x_{(i)}-\mu\right)+(n-r)\left(x_{(r)}-\mu\right)\right]\right\}$$

$$=\frac{n!}{(n-r)!}\frac{1}{\theta^r}\exp\left\{-\frac{1}{\theta}\left[\sum_{i=1}^{r}x_{(i)}+(n-r)x_{(r)}\right]+n\frac{\mu}{\theta}\right\}$$

$$\ln L(\mu,\theta)=\ln\frac{n!}{(n-r)!}-r\ln\theta-\frac{1}{\theta}\left[\sum_{i=1}^{r}x_{(i)}+(n-r)x_{(r)}\right]+n\frac{\mu}{\theta}$$

$$\frac{\partial\ln L(\mu,\theta)}{\partial\mu}=\frac{n}{\theta}>0$$

即似然函数 $L(\mu,\theta)$ 对 μ 严格单调增加，考虑到 $\mu\le x_{(1)}\le x_{(2)}\le\cdots\le x_{(r)}$

于是 μ 的极大似然估计为： $\hat{\mu}=X_{(1)}$

又
$$\frac{\partial\ln L(\mu,\theta)}{\partial\theta}=-\frac{r}{\theta}+\frac{1}{\theta^2}\left[\sum_{i=1}^{r}x_{(i)}+(n-r)x_{(r)}\right]-n\frac{\mu}{\theta^2}$$

$$\frac{\partial\ln L\left(x_{(1)},\theta\right)}{\partial\theta}=-\frac{r}{\theta}+\frac{1}{\theta^2}\left[\sum_{i=1}^{r}x_{(i)}+(n-r)x_{(r)}\right]-n\frac{x_{(1)}}{\theta^2}$$

令 $\frac{\partial\ln L\left(x_{(1)},\theta\right)}{\partial\theta}=0$，得如下方程： $-\frac{r}{\theta}+\frac{1}{\theta^2}\left[\sum_{i=1}^{r}x_{(i)}+(n-r)x_{(r)}\right]-n\frac{x_{(1)}}{\theta^2}=0$

从中可解得参数 θ 的极大似然估计为： $\hat{\theta}=\frac{1}{r}\left[\sum_{i=1}^{r}X_{(i)}+(n-r)X_{(r)}-nX_{(1)}\right]$

记 $Y_{(i)}=\frac{X_{(i)}-\mu}{\theta},i=1,2,\cdots,n$， 则 $Y_{(1)}\le Y_{(2)}\le\cdots\le Y_{(r)}$ 与标准指数分布 $\mathrm{Exp}(1)$ 总体容量为 n 的前 r 个次序统计量同分布。

又
$$X_{(i)}=\mu+\theta Y_{(i)},i=1,2,\cdots,r$$

易知 $Y_{(1)}\sim\mathrm{Exp}(n)$，则
$$E\left(Y_{(1)}\right)=\frac{1}{n},D\left(Y_{(1)}\right)=\frac{1}{n^2}$$

由此
$$E(\hat{\mu})=E\left(X_{(1)}\right)=\mu+\frac{\theta}{n},\lim_{n\to+\infty}E(\hat{\mu})=\mu,D(\hat{\mu})=D\left(X_{(1)}\right)=\frac{\theta^2}{n^2},\lim_{n\to+\infty}D(\hat{\mu})=0$$

即 $\hat{\mu}$ 为参数 μ 的近似无偏与相合估计。

记 $Y_{(0)}=0$，注意到：
$$\hat{\theta}=\frac{\theta}{r}\left[\sum_{i=1}^{r}Y_{(i)}+(n-r)Y_{(r)}-nY_{(1)}\right]$$
$$=\frac{\theta}{r}\left[\sum_{i=1}^{r}(n-i+1)\left(Y_{(i)}-Y_{(i-1)}\right)-nY_{(1)}\right]$$
$$=\frac{\theta}{r}\sum_{i=2}^{r}(n-i+1)\left(Y_{(i)}-Y_{(i-1)}\right)$$

由定理 1 知： $(n-i+1)\left(Y_{(i)}-Y_{(i-1)}\right),i=1,2,\cdots,r$ 相互独立且同服从 $\mathrm{Exp}(1)$

进而知： $2(n-i+1)\left(Y_{(i)}-Y_{(i-1)}\right),i=1,2,\cdots,r$ 相互独立且同服从 $\chi^2(2)$

则
$$E\left(\hat{\theta}\right)=\frac{\theta}{2r}\sum_{i=2}^{r}E\left[2(n-i+1)\left(Y_{(i)}-Y_{(i-1)}\right)\right]=\frac{r-1}{r}\theta$$

$$D\left(\hat{\theta}\right)=\frac{\theta^2}{4r^2}\sum_{i=2}^{r}D\left[2(n-i+1)\left(Y_{(i)}-Y_{(i-1)}\right)\right]=\frac{r-1}{r^2}\theta^2$$

当 n,r 很大时， $\hat{\theta}$ 为参数 θ 的近似无偏与相合估计。

欲使 $\hat{\theta}_1 = c\hat{\theta}$ 为 θ 的无偏估计，即 $E\left(\hat{\theta}_1\right) = \theta$，$E\left(\hat{\theta}_1\right) = cE\left(\hat{\theta}\right) = c\dfrac{r-1}{r}\theta = \theta$，取 $c = \dfrac{r}{r-1}$

$$D\left(\hat{\theta}_1\right) = D\left(c\hat{\theta}\right) = \frac{r^2}{\left(r-1\right)^2}\frac{r-1}{r^2}\theta^2 = \frac{\theta^2}{r-1}$$

当 n,r 很大时，$\hat{\theta}_1$ 为参数 θ 的无偏与相合估计。

(2) 由于 $2\sum_{i=2}^{r}(n-i+1)\left(Y_{(i)} - Y_{(i-1)}\right) = \dfrac{2}{\theta}\sum_{i=2}^{r}(n-i+1)\left(X_{(i)} - X_{(i-1)}\right) \sim \chi^2\left(2(r-1)\right)$

给定置信水平 $1-\alpha$，有：

$$P\left(\chi^2_{1-\alpha/2}\left(2(r-1)\right) \le \frac{2}{\theta}\sum_{i=2}^{r}(n-i+1)\left(X_{(i)} - X_{(i-1)}\right) \le \chi^2_{\alpha/2}\left(2(r-1)\right)\right) = 1-\alpha$$

进而得 θ 的置信水平 $1-\alpha$ 的置信区间为：

$$\left[\frac{2\sum_{i=2}^{r}(n-i+1)\left(X_{(i)} - X_{(i-1)}\right)}{\chi^2_{\alpha/2}\left(2(r-1)\right)}, \frac{2\sum_{i=2}^{r}(n-i+1)\left(X_{(i)} - X_{(i-1)}\right)}{\chi^2_{1-\alpha/2}\left(2(r-1)\right)}\right]$$

记 $X_{(0)} = 0$，取定 $k, k = 1,2,\cdots,r-1$，通常可取 $k = \left[\dfrac{r}{2}\right]$

则
$$2\sum_{i=1}^{k}(n-i+1)\left(Y_{(i)} - Y_{(i-1)}\right) = \frac{2}{\theta}\left[\sum_{i=1}^{k}(n-i+1)\left(X_{(i)} - X_{(i-1)}\right) - n\mu\right] \sim \chi^2(2k)$$

$$2\sum_{i=k+1}^{r}(n-i+1)\left(Y_{(i)} - Y_{(i-1)}\right) = \frac{2}{\theta}\sum_{i=k+1}^{r}(n-i+1)\left(X_{(i)} - X_{(i-1)}\right) \sim \chi^2\left(2(r-k)\right)$$

进而
$$\frac{r-k}{k}\frac{\sum_{i=1}^{k}(n-i+1)\left(Y_{(i)} - Y_{(i-1)}\right)}{\sum_{i=k+1}^{r}(n-i+1)\left(Y_{(i)} - Y_{(i-1)}\right)} \sim F\left(2k, 2(r-k)\right)$$

即
$$\frac{r-k}{k}\frac{\sum_{i=1}^{k}(n-i+1)\left(X_{(i)} - X_{(i-1)}\right) - n\mu}{\sum_{i=k+1}^{r}(n-i+1)\left(X_{(i)} - X_{(i-1)}\right)} \sim F\left(2k, 2(r-k)\right)$$

给定置信水平 $1-\alpha$，有：

$$P\left(F_{1-\alpha/2}\left(2k, 2(r-k)\right) \le \frac{r-k}{k}\frac{\sum_{i=1}^{k}(n-i+1)\left(X_{(i)} - X_{(i-1)}\right) - n\mu}{\sum_{i=k+1}^{r}(n-i+1)\left(X_{(i)} - X_{(i-1)}\right)} \le F_{\alpha/2}\left(2k, 2(r-k)\right)\right) = 1-\alpha$$

进而 μ 的置信水平 $1-\alpha$ 的置信区间为 $\left[\hat{\mu}_1, \hat{\mu}_2\right]$。

其中

$$\hat{\mu}_1 = \frac{1}{n}\left[\sum_{i=1}^{k}(n-i+1)\left(X_{(i)} - X_{(i-1)}\right) - \frac{k}{r-k}F_{\alpha/2}\left(2k, 2(r-k)\right)\sum_{i=k+1}^{r}(n-i+1)\left(X_{(i)} - X_{(i-1)}\right)\right]$$

$$\hat{\mu}_2 = \frac{1}{n}\left[\sum_{i=1}^{k}(n-i+1)\left(X_{(i)} - X_{(i-1)}\right) - \frac{k}{r-k}F_{1-\alpha/2}\left(2k, 2(r-k)\right)\sum_{i=k+1}^{r}(n-i+1)\left(X_{(i)} - X_{(i-1)}\right)\right]$$

引理[11]：设两个相互独立的随机变量 $X_1 \sim \chi^2(n_1), X_2 \sim \chi^2(n_2)$，记 $Y_1 = \dfrac{X_2/n_2}{X_1/n_1}$，$Y_2 = X_1 + X_2$，则 $Y_1 = \dfrac{X_2/n_2}{X_1/n_1} \sim F(n_2, n_1), Y_2 \sim \chi^2(n_1+n_2)$，且 Y_1 与 Y_2 相互独立。

例 2：设 $X_{(1)}, X_{(2)}, \cdots, X_{(r)}$ 为来自三参数威布尔分布总体 X 的容量为 n 的前 r（$5 \le r \le n$）次序统计量，X 的分布函数为：$F(x) = 1 - \exp\left[-\left(\dfrac{x-\mu}{\beta}\right)^m\right]$，$x \ge \mu \ge 0$，$\beta > 0$，$m > 0$。记 $Y = \left(\dfrac{X-\mu}{\beta}\right)^m$，$Y_{(i)} = \left(\dfrac{X_{(i)}-\mu}{\beta}\right)^m, i = 1, 2, \cdots, r$，以及 $X_{(0)} = \mu$，$Y_{(0)} = 0$，$T_\mu = \dfrac{\ln\left(X_{(2)}-\mu\right) - \ln\left(X_{(1)}-\mu\right)}{\ln\left(X_{(3)}-\mu\right) - \ln\left(X_{(1)}-\mu\right)}$，

$T_{\mu,m} = \dfrac{r-k}{k-3} \dfrac{\sum_{i=4}^{k}(n-i+1)\left[\left(X_{(i)}-\mu\right)^m - \left(X_{(i-1)}-\mu\right)^m\right]}{\sum_{i=k+1}^{r}(n-i+1)\left[\left(X_{(i)}-\mu\right)^m - \left(X_{(i-1)}-\mu\right)^m\right]}$，$T_{\mu,m,\beta} = \dfrac{2}{\beta^m} \sum_{i=4}^{r}(n-i+1)\left[\left(X_{(i)}-\mu\right)^m - \left(X_{(i-1)}-\mu\right)^m\right]$，其中 $4 \le k \le r-1$，则有如下结论：(1) $\ln Y_{(2)} - \ln Y_{(1)}$ 与 $2\sum_{i=1}^{r}(n-i+1)\left(Y_{(i)} - Y_{(i-1)}\right) \sim \chi^2(2r)$ 相互独立；(2) T_μ 为枢轴量，$T_{\mu,m} \sim F(2(k-3), 2(r-k))$，$T_{\mu,m,\beta} \sim \chi^2(2(r-3))$，且 $T_\mu, T_{\mu,m}, T_{\mu,m,\beta}$ 三者相互独立。

证明：(1) 易见 $Y \sim \mathrm{Exp}(1)$，而 $Y_{(1)} \le Y_{(2)} \le \cdots \le Y_{(r)}$ 与标准指数分布 $\mathrm{Exp}(1)$ 总体容量为 n 的前 r 个次序统计量同分布。由定理 1 知：

$$(n-i+1)\left(Y_{(i)} - Y_{(i-1)}\right), i = 1, 2, \cdots, r \text{ 相互独立且同服从 } \mathrm{Exp}(1)$$

$$2(n-i+1)\left(Y_{(i)} - Y_{(i-1)}\right), i = 1, 2, \cdots, r \text{ 相互独立且同服从 } \chi^2(2)$$

由引理知：$2\left[nY_{(1)} + (n-1)\left(Y_{(2)} - Y_{(1)}\right)\right] \sim \chi^2(4)$，$\dfrac{n-1}{n}\dfrac{Y_{(2)} - Y_{(1)}}{Y_{(1)}} \sim F(2,2)$，且两者相互独立。

又
$$\ln Y_{(2)} - \ln Y_{(1)} = \ln\left(1 + \frac{Y_{(2)} - Y_{(1)}}{Y_{(1)}}\right)$$

则 $\ln Y_{(2)} - \ln Y_{(1)}$ 与 $2\left[nY_{(1)} + (n-1)\left(Y_{(2)} - Y_{(1)}\right)\right]$ 相互独立。

又 $\ln Y_{(2)} - \ln Y_{(1)}$ 与 $2\sum_{i=3}^{r}(n-i+1)\left(Y_{(i)} - Y_{(i-1)}\right) \sim \chi^2(2(r-2))$ 相互独立。

记 $Z_1 = \ln Y_{(2)} - \ln Y_{(1)}$，$Z_2 = 2\left[nY_{(1)} + (n-1)\left(Y_{(2)} - Y_{(1)}\right)\right]$，$Z_3 = 2\sum_{i=3}^{r}(n-i+1)\left(Y_{(i)} - Y_{(i-1)}\right)$

对 $-\infty < z_1 < +\infty, z_2, z_3 > 0$ 有：

$$\begin{aligned} P(Z_1 \le z_1, Z_2 \le z_2, Z_3 \le z_3) &= P\left((Z_1 \le z_1, Z_2 \le z_2) \cap (Z_3 \le z_3)\right) \\ &= P(Z_1 \le z_1, Z_2 \le z_2) P(Z_3 \le z_3) \\ &= P(Z_1 \le z_1) P(Z_2 \le z_2) P(Z_3 \le z_3) \end{aligned}$$

则 Z_1, Z_2, Z_3 三者相互独立，进而 Z_1 与 $Z_2 + Z_3$ 也相互独立，即有：$\ln Y_{(2)} - \ln Y_{(1)}$ 与 $2\sum_{i=1}^{r}(n-i+1)\left(Y_{(i)} - Y_{(i-1)}\right) \sim \chi^2(2r)$ 相互独立。

(2) 又 $T_{\mu}=\dfrac{\ln\left(X_{(2)}-\mu\right)-\ln\left(X_{(1)}-\mu\right)}{\ln\left(X_{(3)}-\mu\right)-\ln\left(X_{(1)}-\mu\right)}=\dfrac{\ln Y_{(2)}-\ln Y_{(1)}}{\ln Y_{(3)}-\ln Y_{(1)}}$，易见其为枢轴量。

取定 $k, 4\le k\le r-1$，通常可取 $k=\left[\dfrac{r+4}{2}\right]$

则有：

$$\frac{2}{\beta^m}\sum_{i=4}^{k}(n-i+1)\left[\left(X_{(i)}-\mu\right)^m-\left(X_{(i-1)}-\mu\right)^m\right]\sim\chi^2\left(2(k-3)\right)$$

$$\frac{2}{\beta^m}\sum_{i=k+1}^{r}(n-i+1)\left[\left(X_{(i)}-\mu\right)^m-\left(X_{(i-1)}-\mu\right)^m\right]\sim\chi^2\left(2(r-k)\right)$$

并有：$T_{\mu,m}=\dfrac{r-k}{k-3}\dfrac{\sum_{i=4}^{k}(n-i+1)\left[\left(X_{(i)}-\mu\right)^m-\left(X_{(i-1)}-\mu\right)^m\right]}{\sum_{i=k+1}^{r}(n-i+1)\left[\left(X_{(i)}-\mu\right)^m-\left(X_{(i-1)}-\mu\right)^m\right]}\sim F\left(2(k-3),2(r-k)\right)$ 与

$T_{\mu,m,\beta}=\dfrac{2}{\beta^m}\sum_{i=4}^{r}(n-i+1)\left[\left(X_{(i)}-\mu\right)^m-\left(X_{(i-1)}-\mu\right)^m\right]\sim\chi^2\left(2(r-3)\right)$ 相互独立。

记 $Z_1=T_{\mu},Z_2=T_{\mu,m},Z_3=T_{\mu,m,\beta}$，对 $0<z_1<1,z_2,z_3>0$ 有：

$$\begin{aligned}P\left(Z_1\le z_1,Z_2\le z_2,Z_3\le z_3\right)&=P\left(\left(Z_1\le z_1\right)\cap\left(Z_2\le z_2,Z_3\le z_3\right)\right)\\&=P\left(Z_1\le z_1\right)P\left(Z_2\le z_2,Z_3\le z_3\right)\\&=P\left(Z_1\le z_1\right)P\left(Z_2\le z_2\right)P\left(Z_3\le z_3\right)\end{aligned}$$

则 Z_1,Z_2,Z_3 三者相互独立。

4. 结论

针对文献[1]-[4]中所提出的“指数分布抽样基本定理”进行了评说与完善，认为将其称为“基本定理”并不合适，真正对指数分布的统计推断起关键作用的是定理 1 及定理 2，同时简化了文献[4]中结论的相关证明，并修正了文献[4]中的一些错误。论文最后通过两参数指数分布与三参数威布尔分布总体说明“指数分布抽样分布基本定理”的应用。

致 谢

作者非常感谢各位审稿专家提出的宝贵建议！

基金项目

(1) 上海对外经贸大学“应用统计学”国家一流专业建设项目资助；(2) 上海对外经贸大学“高水平地方高校建设项目－创新人才培养－特定领域急需人才精准培养教学资源建设－AI 赋能教学‘1 + 1 + N’计划－商务统计”((A1A-7003-25-058-02))资助。

参考文献

[1] 李国安. 指数分布抽样基本定理及在指数分布参数统计推断中的应用[J]. 大学数学, 2016, 32(5): 30-36.

[2] 李国安. 指数分布抽样基本定理及在四参数二元 Marshall-Olkin 型指数分布参数估计中的应用[J]. 统计研究, 2016, 33(7): 98-102.

[3] 李国安, 李建峰. 指数分布抽样基本定理及在三参数一般指数分布参数估计中的应用[J]. 数学的实践与认识, 2017, 47(3): 165-169.

[4] 李国安, 李穆真. 指数抽样分布定理及三个期望之极小方差无偏估计的有效性比较[J]. 纯粹数学与应用数学,

2017, 33(6): 568-577.

[5] Sukhatme, P.V. (1937) Tests of Significance for Samples of the X^2-Population with Two Degrees of Freedom. *Annals of Eugenics*, **8**, 52-56. https://doi.org/10.1111/j.1469-1809.1937.tb02159.x

[6] Rao, C.R. (1973) Linear Statistical Inference and Its Applications. 2nd Edition, Wiley, 43-46. https://doi.org/10.1002/9780470316436

[7] Balakrishnan, N. and Clifford Cohen, A. (1991) Order Statistics and Inference Estimation Methods. Academic Press, 34-36.

[8] 曹晋华, 程侃. 可靠性数学引论[M]. 北京: 科学出版社, 1986: 383-385.

[9] 魏宗舒. 概率论与数理统计教程[M]. 北京: 高等教育出版社, 1983: 477-481.

[10] 徐晓岭, 王蓉华, 顾蓓青. 概率论与数理统计[M]. 第二版. 上海: 上海交通大学出版社, 2021: 233-238.

[11] 王蓉华, 徐晓岭, 顾蓓青. 概率论与数理统计案例分析[M]. 上海: 上海交通大学出版社, 2023: 122-125, 136-143.

[12] 王蓉华, 徐晓岭, 顾蓓青. 概率论与数理统计案例分析与提高[M]. 上海: 上海交通大学出版社, 2025: 69, 72-73.

[13] Baklizi, A. (2016) Confidence Intervals for the Two-Parameter Exponential Reliability with Type II Censored Data. *American Journal of Mathematical and Management Sciences*, **35**, 297-308. https://doi.org/10.1080/01966324.2016.1202163

[14] 程绩, 李云飞. 不完全数据场合下双参数指数分布参数的区间估计[J]. 统计与决策, 2017(16): 15-18.

[15] Zhang, J. (2018) Minimum Volume Confidence Sets for Two-Parameter Exponential Distributions. *The American Statistician*, **72**, 213-218. https://doi.org/10.1080/00031305.2016.1264315

[16] Bobotas, P. (2018) Estimation of the Smallest Scale Parameter of Two-Parameter Exponential Distributions. *Communications in Statistics—Theory and Methods*, **48**, 2748-2765. https://doi.org/10.1080/03610926.2018.1472792

[17] Asgharzadeh, A., Bagheri, S.F., Ibrahim, N.A. and Abubakar, M.R. (2019) Optimal Confidence Regions for the Two-Parameter Exponential Distribution Based on Records. *Computational Statistics*, **35**, 309-326. https://doi.org/10.1007/s00180-019-00914-x

[18] Akbari, M. and Akbari, M. (2020) Some Applications of Near-Order Statistics in Two-Parameter Exponential Distribution. *Journal of Statistical Theory and Applications*, **19**, Article No. 21. https://doi.org/10.2991/jsta.d.200224.001

[19] El-Adll, M.E. (2019) Inference for the Two-Parameter Exponential Distribution with Generalized Order Statistics. *Mathematical Population Studies*, **28**, 1-23. https://doi.org/10.1080/08898480.2019.1681187

[20] Malekzadeh, A., Jafari, A.A. and Mahmoudi, S.M. (2021) Inference on the Mean Parameter of a Two-Parameter Exponential Distribution: Complete, Censored and Record Data. *Journal of Statistics and Management Systems*, **24**, 1233-1252. https://doi.org/10.1080/09720510.2020.1809118