Probability in the Engineering and Informational Sciences, **32**, 2018, 306–322.
doi:10.1017/S0269964817000134

ORDERING PROPERTIES OF SPACINGS FROM HETEROGENEOUS GEOMETRIC SAMPLES

WEIYONG DING
School of Mathematics and Statistics, Jiangsu Normal University, Xuzhou 221116, China

YIYING ZHANG
Department of Statistics and Actuarial Science, The University of Hong Kong, Pokfulam Road, Hong Kong

PENG ZHAO
School of Mathematics and Statistics, Jiangsu Normal University, Xuzhou 221116, China
E-mail: zhaop@jsnu.edu.cn

In the reliability context, the geometric distribution is a natural choice to model the lifetimes of some equipment and components when they are measured by the number of completed cycles of operation or strokes, or in case of periodic monitoring of continuous data. This paper aims at investigating how the heterogeneity among the parameters affects some characteristics such as the distribution and hazard rate functions of spacings arising from independent heterogeneous geometric random variables. First, refined representations of the distribution functions are provided for both the second spacing and sample range from heterogeneous geometric sample. Second, stochastic comparisons are carried out on the second spacings and sample ranges for two sets of independent and heterogeneous geometric random variables in the sense of the usual stochastic and hazard rate orderings. The results established here not only fill the gap on the study of stochastic properties of spacings from heterogeneous geometric samples, but also are expected to be applied in the fields of statistics and reliability.

Keywords: geometric distribution, hazard rate ordering, majorization, sample range, second spacing, usual stochastic ordering

1. INTRODUCTION

As the discrete counterpart of the exponential distribution, the geometric random variable X with parameter $0 < q < 1$, is given by the probability mass function

$$\mathsf{P}(X = k) = pq^{k-1}, \quad k = 1, 2, \ldots,$$

where $p = 1 - q$. This model is characterized by the lack of memory and the constant hazard rate properties. It is well known that the exponential distribution is widely referenced

 0269-9648/17 $25.00

probability law as the simplest choice used in reliability and life testing for continuous data. When the lifetimes of some equipment and components are measured by the number of completed cycles of operation or strokes, or in case of periodic monitoring of continuous data, the geometric distribution is a natural choice instead of exponential distribution. In the literature, the geometric random variable plays a special role in stochastic modeling and has received much attention. For instance, [7] modeled the fatigue failure of aircraft wing by geometric distribution. [9] studied the order statistics from geometric distribution and their relation to inverse sampling and reliability of redundant systems. [2] pointed out the maximum order statistics from heterogenous geometric samples could be used in the engineering models, such as wireless broadcast transmission systems.

Let $X_1, \ldots, X_n$ be independent random variables with possible different distributions. Denote the ith order statistic by $X_{i:n}$, and the ith spacing by $X_{i:n} - X_{i-1:n}$, $i = 1, \ldots, n$, with $X_{0:n} \equiv 0$, and the sample range by $X_{n:n} - X_{1:n}$. Spacings and their functions are of great interest in many areas. For instance, the sample range plays an important role as one of the criteria for comparing variabilities among distributions in statistics, while spacings are used in goodness-of-fit tests, outlier detection and characterization of distributions. In the context of reliability, spacings correspond to times elapsed between failures of components in a system, etc. Therefore, it is important to study the stochastic properties of spacings under different models. In the past decades, a surge of research has appeared on stochastic comparisons of spacings and sample ranges arising from independent exponential samples. Especially, considerable attention has been paid recently to studying how the heterogeneity affects some distribution characteristics of exponential spacings and ranges, including their distribution functions and the (reversed) hazard rate functions. Interested readers may refer to [5,6] for an elaborate review. However, if the random sample comes from discrete variables, the distribution theory of spacings becomes very complicated due to the existence of ties, and few literature was devoted to studying this topic.

For ease of reference, let us recall some stochastic orderings for discrete random variables and the notion of majorization that will be used in the sequel. Let X and Y be two non-negative discrete random variables taking values on the set of natural numbers $\mathbb{N}$. Let $\mathbb{N}_+ = \mathbb{N} \setminus \{0\}$, $\mathbb{R} = (-\infty, +\infty)$ and $\mathbb{R}_- = (-\infty, 0]$. The notation "$\overset{\text{sgn}}{=}$" is used to express that the both sides of the equity have the same sign. Throughout this paper, the term *increasing* is used for *monotone non-decreasing* and *decreasing* for *monotone non-increasing*.

Definition 1.1: X is said to be greater than Y in:

(i) *the usual stochastic ordering* (denoted by $X \geq_{\text{st}} Y$) if $\mathsf{P}(X \geq k) \geq \mathsf{P}(Y \geq k)$ for all $k \in \mathbb{N}$;

(ii) *the hazard rate ordering* (denoted by $X \geq_{\text{hr}} Y$) if $\mathsf{P}(X \geq k)/\mathsf{P}(Y \geq k)$ is increasing in $k \in \mathbb{N}$;

(iii) *the reversed hazard rate ordering* (denoted by $X \geq_{\text{rh}} Y$) if $\mathsf{P}(X \leq k)/\mathsf{P}(Y \leq k)$ is increasing in $k \in \mathbb{N}$;

(iv) *the likelihood ratio ordering* (denoted by $X \geq_{\text{lr}}$) if $\mathsf{P}(X = k)/\mathsf{P}(Y = k)$ is increasing in $k \in \mathbb{N}$.

It is well known that both of the hazard rate ordering and reversed hazard rate ordering imply the usual stochastic ordering. For comprehensive details on the theory of stochastic orderings and their applications, one may refer to [12]).

The concept of *majorization*, which is quite useful in establishing various inequalities, will be also employed to describe the difference in parameter vectors of random geometric samples in our discussion. Let $x_{1:n} \le \cdots \le x_{n:n}$ be the increasing arrangement of the components of the vector $\boldsymbol{x} = (x_1, \ldots, x_n)$.

DEFINITION 1.2: A vector $\boldsymbol{x}$ is said to *majorize* another vector $\boldsymbol{y} = (y_1, \ldots, y_n)$ (denoted by $\boldsymbol{x} \overset{\text{m}}{\succeq} \boldsymbol{y}$), if $\sum_{i=1}^{n} x_{i:n} = \sum_{i=1}^{n} y_{i:n}$ and $\sum_{i=1}^{j} x_{i:n} \le \sum_{i=1}^{j} y_{i:n}$ for $j = 1, \ldots, n-1$.

Recall that a function $\varphi : \mathbb{R}^n \mapsto \mathbb{R}$ is said to be *Schur-convex* [*Schur-concave*] if $\varphi(\boldsymbol{x}) \ge [\le]\varphi(\boldsymbol{y})$ whenever $\boldsymbol{x} \overset{\text{m}}{\succeq} \boldsymbol{y}$ for $\boldsymbol{x}$ and $\boldsymbol{y}$ in $\mathbb{R}^n$. For a comprehensive exposition on theory and application of the majorization order as well as Schur-convex (Schur-concave) function, one may refer to [10].

The following two lemmas concerning majorization and Schur-convex functions are quite useful in developing our main results in the sequel.

LEMMA 1.3: [10] Let $\phi(\boldsymbol{x})$ be symmetric and have continuous partial derivatives for $\boldsymbol{x} = (x_1, \cdots, x_n) \in \mathcal{D}^n$, $\mathcal{D}$ is an open interval. Then $\phi : \mathcal{D}^n \mapsto \mathbb{R}$ is Schur-convex [Schur-concave], if and only if

$$(x_i - x_j)\left(\frac{\partial \phi(\boldsymbol{x})}{\partial x_i} - \frac{\partial \phi(\boldsymbol{x})}{\partial x_j}\right) \ge [\le]0,$$

for $\boldsymbol{x} \in \mathcal{D}^n$ such that $x_i \ne x_j$ with $1 \le i < j \le n$.

LEMMA 1.4: [3] If $I \subset \mathbb{R}$ is an interval and $g : I \to \mathbb{R}$ is a convex function, then $\phi(\boldsymbol{x}) = \sum_{i=1}^{m} g(x_i)$ is a Schur-convex function of $\boldsymbol{x} = (x_1, \ldots, x_m)$ on I^n.

To the best of our knowledge, there are few papers in the literature handling the stochastic properties of order statistics arising from heterogeneous geometric samples such as [1,8,13], but none of their work exists for spacings. To fill this gap, this paper focuses on providing refined representations of the distribution functions of the second spacing and sample range arising from heterogeneous geometric sample, and then investigating how the heterogeneity affects some characteristics such as the corresponding distribution and hazard rate functions. Let $X_1, \ldots, X_n$ be a sequence of independent geometric variables with respective parameters $q_1, \ldots, q_n$, and let $Y_1, \ldots, Y_n$ be another sequence of independent geometric variables with respective parameters $q_1^*, \ldots, q_n^*$. Denote $\boldsymbol{q} = (q_1, \ldots, q_n)$ and $\log \boldsymbol{q} = (\log q_1, \ldots, \log q_n)$. For the case that both samples are heterogeneous, we prove that the majorization order between $\log \boldsymbol{q}$ and $\log \boldsymbol{q}^*$ implies the usual stochastic ordering between $X_{2:n} - X_{1:n}$ and $Y_{2:n} - Y_{1:n}$, and the majorization order between $\boldsymbol{q}$ and $\boldsymbol{q}^*$ implies the usual stochastic ordering between $X_{n:n} - X_{1:n}$ and $Y_{n:n} - Y_{1:n}$. Let

$$\tilde{q} = \left(\frac{1}{n}\sum_{i=1}^{n}\prod_{j\ne i}^{n} q_j\right)^{1/(n-1)} \quad \text{and} \quad \bar{q} = \frac{1}{n}\sum_{i=1}^{n} q_i.$$

Under the assumption that $q_1^* = \cdots = q_n^* = q$, it is further shown that, in the sense of hazard rate ordering, $X_{2:n} - X_{1:n}$ is larger than $Y_{2:n} - Y_{1:n}$ if $q \le \tilde{q}$, whereas $X_{2:n} - X_{1:n}$ is smaller than $Y_{2:n} - Y_{1:n}$ if $q \ge q_{\max} = \max_{1 \le i \le n} q_i$, and $X_{n:n} - X_{1:n}$ is larger than $Y_{n:n} - Y_{1:n}$ according to the usual stochastic ordering if $q \le \bar{q}$.

The remainder of the paper is rolled out as follows. Section 2 discusses the ordering properties of the second spacing from independent heterogeneous geometric variables in the sense of the hazard rate and usual stochastic orderings. In Section 3, sufficient conditions are provided for the usual stochastic ordering between the sample ranges arising from two sets of independent and heterogenous/homogeneous geometric variables. Section 4 concludes the paper.

2. SECOND SPACING

In this section, stochastic comparisons are carried out on the second spacings arising from independent heterogeneous geometric samples. We start our discussion by presenting the explicit representation of survival function of second spacing arising from independent heterogeneous geometric sample.

PROPOSITION 2.1: Let $X_1, \ldots, X_n$ be a sequence of independent geometric random variables with respective parameters $q_1, \ldots, q_n$. Then, it follows that, for $k \in \mathbb{N}_+$

$$\mathsf{P}(X_{2:n} - X_{1:n} \geq k) = \frac{\prod_{i=1}^n q_i^k}{1 - \prod_{i=1}^n q_i} \sum_{i=1}^n \frac{1 - q_i}{q_i^k},$$

and

$$\mathsf{P}(X_{2:n} - X_{1:n} = 0) = 1 - \frac{\prod_{i=1}^n q_i}{1 - \prod_{i=1}^n q_i} \sum_{i=1}^n \frac{1 - q_i}{q_i}.$$

PROOF: Note that, for $k \in \mathbb{N}_+$,

$$\begin{aligned} \mathsf{P}(X_{2:n} - X_{1:n} \geq k) &= \mathsf{P}(X_{2:n} - X_{1:n} \geq k, X_{2:n} - X_{1:n} > 0) \\ &= \sum_{i=1}^n \mathsf{P}(X_{2:n} - X_{1:n} \geq k, X_{2:n} - X_{1:n} > 0, X_{1:n} = X_i) \\ &= \sum_{i=1}^n \mathsf{P}(X_{2:n} - X_{1:n} \geq k | X_{2:n} - X_{1:n} > 0, X_{1:n} = X_i) \\ &\quad \times \mathsf{P}(X_{2:n} - X_{1:n} > 0, X_{1:n} = X_i). \end{aligned} \quad \textbf{(1)}$$

First, it can be observed that, for any fixed $i \in \{1, \ldots, n\}$

$$\begin{aligned} \mathsf{P}(X_{2:n} - X_{1:n} > 0, X_{1:n} = X_i) &= \mathsf{P}(X_j > X_i, j \neq i) \\ &= \sum_{k=1}^{\infty} \mathsf{P}(X_j > X_i = k, j \neq i) \\ &= \sum_{k=1}^{\infty} (1 - q_i) q_i^{k-1} \prod_{j \neq i}^n q_j^k \\ &= \frac{1 - q_i}{q_i} \sum_{k=1}^{\infty} \prod_{j=1}^n q_j^k \\ &= \frac{1 - q_i}{q_i} \frac{\prod_{j=1}^n q_j}{1 - \prod_{j=1}^n q_j}. \end{aligned} \quad \textbf{(2)}$$

Notice that the random variable $Z^{[i]} := [X_{2:n} - X_{1:n} | X_{2:n} - X_{1:n} > 0, X_{1:n} = X_i]$ is the smallest order statistic from the independent random variables $Z_j^{[i]} := [X_j - X_i | X_j > X_i]$

https://doi.org/10.1017/S0269964817000134 Published online by Cambridge University Press

for $j = 1, \ldots, n$ and $j \neq i$. From the memoryless property of the geometric distribution, it is known that $Z_j^{[i]}$ has the same distribution as X_j for $j = 1, \ldots, n$ and $j \neq i$. Let $X_{j:n-1}^{[i]}$ be the jth order statistic from $\{X_1, \ldots, X_n\} \setminus \{X_i\}$. Therefore, $Z^{[i]}$ has the same survival function as $X_{1:n-1}^{[i]}$, that is,

$$\mathsf{P}(Z^{[i]} \geq k) = \mathsf{P}(X_{1:n-1}^{[i]} \geq k) = \prod_{j \neq i}^{n} q_j^{k-1}. \tag{3}$$

For $k = 1, 2, \ldots$, substituting (2) and (3) into (1), we get

$$\mathsf{P}(X_{2:n} - X_{1:n} \geq k) = \sum_{i=1}^{n} \prod_{j \neq i} q_i^{k-1} \frac{1-q_i}{q_i} \frac{\prod_{i=1}^{n} q_i}{1 - \prod_{i=1}^{n} q_i} = \frac{\prod_{i=1}^{n} q_i^k}{1 - \prod_{i=1}^{n} q_i} \sum_{i=1}^{n} \frac{1-q_i}{q_i^k}.$$

On the other hand,

$$\begin{aligned}\mathsf{P}(X_{2:n} - X_{1:n} = 0) &= 1 - \mathsf{P}(X_{2:n} - X_{1:n} \geq 1) \\ &= 1 - \frac{\prod_{i=1}^{n} q_i}{1 - \prod_{i=1}^{n} q_i} \sum_{i=1}^{n} \frac{1-q_i}{q_i}.\end{aligned}$$

Thus, the proof is completed. ■

Now, we are ready to state the first main result of this section.

THEOREM 2.2: Let $X_1, \ldots, X_n$ be a sequence of independent geometric random variables with respective parameters $q_1, \ldots, q_n$, and let $Y_1, \ldots, Y_n$ be another sequence of independent geometric random variables with respective parameters $q_1^*, \ldots, q_n^*$. Then,

$$\log \boldsymbol{q} \overset{\mathrm{m}}{\succeq} \log \boldsymbol{q}^* \quad \Longrightarrow \quad X_{2:n} - X_{1:n} \geq_{\mathrm{st}} Y_{2:n} - Y_{1:n}.$$

PROOF: By making the transformations $\lambda_i = \log q_i$ and $\lambda_i^* = \log q_i^*$ for $i = 1, \ldots, n$, we have $\boldsymbol{\lambda} \succeq_{\mathrm{m}} \boldsymbol{\lambda}^*$. It is trivial that $\mathsf{P}(X_{2:n} - X_{1:n} \geq 0) = \mathsf{P}(Y_{2:n} - Y_{1:n} \geq 0) = 1$. Then, for any $k \in \mathbb{N}_+$, the survival functions of $X_{2:n} - X_{1:n}$ and $Y_{2:n} - Y_{1:n}$ can be written, respectively, as

$$\mathsf{P}(X_{2:n} - X_{1:n} \geq k) = \frac{e^{k \sum_{i=1}^{n} \lambda_i}}{1 - e^{\sum_{i=1}^{n} \lambda_i}} \sum_{i=1}^{n} \frac{1 - e^{\lambda_i}}{e^{k\lambda_i}}$$

and

$$\mathsf{P}(Y_{2:n} - Y_{1:n} \geq k) = \frac{e^{k \sum_{i=1}^{n} \lambda_i^*}}{1 - e^{\sum_{i=1}^{n} \lambda_i^*}} \sum_{i=1}^{n} \frac{1 - e^{\lambda_i^*}}{e^{k\lambda_i^*}}.$$

We need to show the inequality

$$\sum_{i=1}^{n} \frac{1 - e^{\lambda_i}}{e^{k\lambda_i}} \geq \sum_{i=1}^{n} \frac{1 - e^{\lambda_i^*}}{e^{k\lambda_i^*}}.$$

As a result, it is enough to show that

$$\phi_1(\boldsymbol{\lambda}) = \sum_{i=1}^{n} \frac{1 - e^{\lambda_i}}{e^{k\lambda_i}}$$

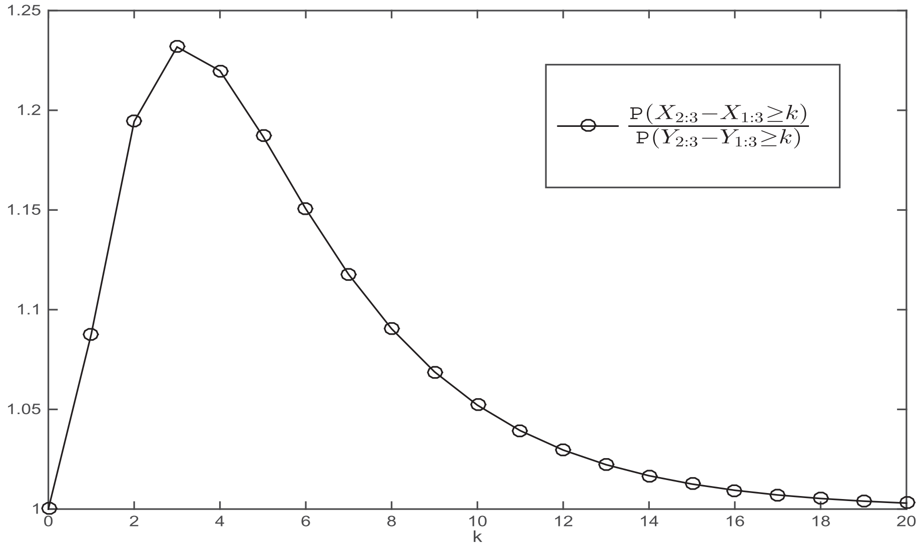

FIGURE 1. Ratio of the survival functions of $X_{2:3} - X_{1:3}$ and $Y_{2:3} - Y_{1:3}$.

is Schur-convex in $\boldsymbol{\lambda} \in \mathbb{R}_-^n$. Note that, for all $x \in \mathbb{R}_-$,

$$\frac{\mathrm{d}^2(1-e^x)e^{-kx}}{\mathrm{d}x^2} = k^2e^{-kx} - (k-1)^2e^{(1-k)x} \geq 0,$$

which means $(1-e^x)e^{-kx}$ is convex in $x \in \mathbb{R}_-$. Now, upon using Lemma 1.4, the desired result follows. ■

As the illustration of Theorem 2.2 above, the next example shows that the result in Theorem 2.2 cannot be strengthened to the hazard rate ordering.

Example 2.3: Let X_1, X_2, and X_3 be independent geometric random variables with parameters $(q_1, q_2, q_3) = (0.15, 0.2, 0.6)$, and let Y_1, Y_2, and Y_3 be another set of independent geometric variables with parameters $(q_1^*, q_2^*, q_3^*) = (0.15, 0.3, 0.4)$. Clearly, it can be verified that $(\log(0.15), \log(0.2), \log(0.6)) \succeq_{\mathrm{m}} (\log(0.15), \log(0.3), \log(0.4))$. The ratio of survival functions of $X_{2:3} - X_{1:3}$ and $Y_{2:3} - Y_{1:3}$ is plotted in Figure 1, from which it can be seen that the ratio is always larger than 1, but the corresponding curve is firstly increasing and then decreasing in $k \in \mathbb{N}_+$. This implies that there is usual stochastic ordering but no hazard rate ordering between $X_{2:3} - X_{1:3}$ and $Y_{2:3} - Y_{1:3}$.

However, the hazard rate ordering between the second spacings can be established in the case wherein one random sample is heterogeneous and the other is homogeneous. They are presented in the following two results.

THEOREM 2.4: Let $X_1, \ldots, X_n$ be a sequence of independent geometric random variables with respective parameters $q_1, \ldots, q_n$, and let $Y_1, \ldots, Y_n$ be another sequence of independent and identically distributed (i.i.d.) geometric random variables with a common parameter q. Then, if $q \leq \tilde{q}$, we have $X_{2:n} - X_{1:n} \geq_{\mathrm{hr}} Y_{2:n} - Y_{1:n}$, here $\tilde{q} = ((1/n)\sum_{i=1}^n \prod_{j\neq i}^n q_j)^{1/(n-1)}$.

https://doi.org/10.1017/S0269964817000134 Published online by Cambridge University Press

Proof: Denote by $\lambda_X(k)$ and $\lambda_Y(k)$ the hazard rate functions of $X_{2:n} - X_{1:n}$ and $Y_{2:n} - Y_{1:n}$, respectively. Then, from Proposition 2.1 we have

$$\lambda_X(0) = \frac{\mathsf{P}(X_{2:n} - X_{1:n} = 0)}{\mathsf{P}(X_{2:n} - X_{1:n} \geq 0)} = 1 - \frac{\prod_{i=1}^n q_i}{1 - \prod_{i=1}^n q_i} \sum_{i=1}^n \frac{1-q_i}{q_i},$$

and for $k \in \mathbb{N}_+$,

$$\begin{aligned}\lambda_X(k) &= \frac{\mathsf{P}(X_{2:n} - X_{1:n} \geq k) - \mathsf{P}(X_{2:n} - X_{1:n} \geq k+1)}{\mathsf{P}(X_{2:n} - X_{1:n} \geq k)} \\ &= 1 - \frac{\prod_{i=1}^n q_i \sum_{i=1}^n \frac{1-q_i}{q_i^{k+1}}}{\sum_{i=1}^n \frac{1-q_i}{q_i^k}} = 1 - \frac{\sum_{i=1}^n \frac{1-q_i}{q_i^k} \prod_{j\neq i}^n q_j}{\sum_{i=1}^n \frac{1-q_i}{q_i^k}}.\end{aligned}$$

Similarly, it can be observed that

$$\lambda_Y(0) = 1 - n\frac{1-q}{q}\frac{q^n}{1-q^n} = 1 - n\frac{q^{n-1}-q^n}{1-q^n},$$

and for $k \in \mathbb{N}_+$, $\lambda_Y(k) = 1 - q^{n-1}$. To establish the desired result, it needs to prove that $\lambda_X(k) \leq \lambda_Y(k)$ for $k \in \mathbb{N}$, which is equivalent to showing that

$$\frac{\sum_{i=1}^n \frac{1-q_i}{q_i^k} \prod_{j\neq i}^n q_j}{\sum_{i=1}^n \frac{1-q_i}{q_i^k}} \geq q^{n-1}, \quad \text{for } k \geq 1, \tag{4}$$

and

$$\frac{\prod_{i=1}^n q_i}{1 - \prod_{i=1}^n q_i} \sum_{i=1}^n \frac{1-q_i}{q_i} \geq n\frac{q^{n-1}-q^n}{1-q^n}. \tag{5}$$

We first prove the inequality (4). Without loss of generality, it is assumed that $q_1 \geq \cdots \geq q_n$. By noticing that $(1-t)/t^k$ is decreasing in $t \in (0,1)$ for any $k \in \mathbb{N}_+$, we have

$$\frac{1-q_1}{q_1^k} \leq \frac{1-q_2}{q_2^k} \leq \cdots \leq \frac{1-q_n}{q_n^k}.$$

Further, based on the fact

$$\prod_{j\neq 1}^n q_j \leq \prod_{j\neq 2}^n q_j \leq \cdots \leq \prod_{j\neq n}^n q_j$$

and the *Čebyšev's sum inequality* (see [4]), it then follows that

$$\sum_{i=1}^n \frac{1-q_i}{q_i^k} \prod_{j\neq i}^n q_j \geq \sum_{i=1}^n \frac{1-q_i}{q_i^k} \left(\frac{1}{n} \sum_{i=1}^n \prod_{j\neq i}^n q_j \right).$$

Hence, it suffices to show that

$$\frac{1}{n} \sum_{i=1}^n \prod_{j\neq i}^n q_j \geq q^{n-1},$$

which is readily guaranteed by the assumption.

https://doi.org/10.1017/S0269964817000134 Published online by Cambridge University Press

Next, we prove the inequality (5). From the well-known *MacLaurin's inequalities* (see [4]), we have $\tilde{q} \geq \left(\prod_{i=1}^n q_i\right)^{1/n}$, or equivalently, $\tilde{q}^n \geq \prod_{i=1}^n q_i$. For any fixed $0 < \alpha \leq 1$, one can see that the function $\frac{\alpha - t}{1-t}$ is decreasing in $t \in (0,1)$. Then, the left-hand side of inequality (5) satisfies

$$\frac{\prod_{i=1}^n q_i}{1 - \prod_{i=1}^n q_i} \sum_{i=1}^n \frac{1-q_i}{q_i} = \frac{\prod_{i=1}^n q_i}{1 - \prod_{i=1}^n q_i} \sum_{i=1}^n \left(\frac{1}{q_i} - 1\right)$$
$$= n \frac{\tilde{q}^{n-1} - \prod_{i=1}^n q_i}{1 - \prod_{i=1}^n q_i} \geq n \frac{\tilde{q}^{n-1} - \tilde{q}^n}{1 - \tilde{q}^n}.$$

From the assumption $q \leq \tilde{q}$, inequality (5) follows immediately by observing the function $(x^{n-1} - x^n)/(1 - x^n)$ is increasing in $x \in (0,1)$. Thus, the proof is completed. ■

The following corollary, easily seen from the proof of Theorem 2.4, can be used to compare the second spacings arising from two sets of homogeneous geometric samples in the sense of the hazard rate order.

COROLLARY 2.5: *Let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ be two sets of i.i.d. geometric variables with respective parameters q and q^* such that $q > q^*$. Then, $X_{2:n} - X_{1:n} \geq_{\mathrm{hr}} Y_{2:n} - Y_{1:n}$.*

THEOREM 2.6: Let $X_1, \ldots, X_n$ be a sequence of independent geometric variables with respective parameters $q_1, \ldots, q_n$, and let $Y_1, \ldots, Y_n$ be another sequence of i.i.d. geometric variables with a common parameter q. Then, if $q \geq q_{\max}$, we have $X_{2:n} - X_{1:n} \leq_{\mathrm{hr}} Y_{2:n} - Y_{1:n}$, here $q_{\max} = \max_{1 \leq i \leq n} q_i$.

PROOF: According to the proof of Theorem 2.4, it suffices for us to show

$$\frac{\sum_{i=1}^n \frac{1-q_i}{q_i^k} \prod_{j \neq i}^n q_j}{\sum_{i=1}^n \frac{1-q_i}{q_i^k}} \leq q^{n-1}, \quad \text{for } k \geq 1 \tag{6}$$

and

$$\frac{\prod_{i=1}^n q_i}{1 - \prod_{i=1}^n q_i} \sum_{i=1}^n \frac{1-q_i}{q_i} \leq n \frac{q^{n-1} - q^n}{1 - q^n}. \tag{7}$$

The inequality in (6) is apparent, and hence we only need to show the inequality in (7).

Let

$$\Delta(q_1, \ldots, q_n) = \frac{\prod_{i=1}^n q_i}{1 - \prod_{i=1}^n q_i} \sum_{i=1}^n \frac{1-q_i}{q_i}.$$

In the following, we will prove that $\Delta(q_1, \ldots, q_n)$ is increasing in each component $q_i \in (0,1)$, for $i = 1, \ldots, n$, which in turn implies that

$$\Delta(q_1, \ldots, q_n) \leq \Delta(q, \ldots, q) = n \frac{q^{n-1} - q^n}{1 - q^n}.$$

Upon taking the partial derivative of $\Delta(q_1,\ldots,q_n)$ with respect to q_1, we have

$$\begin{aligned}\frac{\partial\Delta(q_1,\ldots,q_n)}{\partial q_1} &= \frac{\prod_{i=2}^n q_i(1-\prod_{i=1}^n q_i)+\prod_{i=1}^n q_i\prod_{i=2}^n q_i}{\left(1-\prod_{i=1}^n q_i\right)^2}\sum_{i=1}^n\frac{1-q_i}{q_i} \\ &\quad -\frac{\prod_{i=1}^n q_i}{1-\prod_{i=1}^n q_i}\times\frac{1}{q_1^2} \\ &= \frac{\prod_{i=2}^n q_i}{\left(1-\prod_{i=1}^n q_i\right)^2}\sum_{i=1}^n\frac{1-q_i}{q_i}-\frac{\prod_{i=2}^n q_i}{q_1(1-\prod_{i=1}^n q_i)} \\ &\overset{\text{sgn}}{=} \sum_{i=1}^n\frac{1-q_i}{q_i}-\frac{1-\prod_{i=1}^n q_i}{q_1} \\ &= \sum_{i=2}^n\frac{1-q_i}{q_i}+\prod_{i=2}^n q_i-1 \\ &= \sum_{i=2}^n\frac{1}{q_i}+\prod_{i=2}^n q_i-n.\end{aligned}$$

Let

$$\Psi(q_2,\ldots,q_n)=\sum_{i=2}^n\frac{1}{q_i}+\prod_{i=2}^n q_i-n.$$

It suffices to show $\Psi(q_2,\ldots,q_n)\geq 0$ for all $0\leq q_j\leq 1$, $j=2,\ldots,n$. For any index $j\in\{2,3,\ldots,n\}$, we have

$$\frac{\partial\Psi(q_2,\ldots,q_n)}{\partial q_j}=-\frac{1}{q_j^2}+\prod_{i\neq 1,j}^n q_i\leq 0.$$

Therefore, $\Psi(q_2,\ldots,q_n)$ is decreasing in $q_j\in(0,1)$, $j=2,\ldots,n$. Then, it holds that

$$\Psi(q_2,\ldots,q_n)\geq\Psi(1,\ldots,1)=0.$$

Thus, we know that $\Delta(q_1,\ldots,q_n)$ is increasing in q_1. The proof of the increasing of $\Delta(q_1,\ldots,q_n)$ in q_i, $i=2,\ldots,n-1$, is similar. Thus, we complete the proof. ■

As a direct consequence of Theorems 2.4 and 2.6, we can obtain an upper and lower bounds on the hazard rate function of the second spacing from any heterogeneous geometric sample in terms of that from i.i.d. geometric variables. We now present a numerical example to illustrate this fact.

Example 2.7: Let X_1, X_2, and X_3 be independent geometric variables with parameters $(q_1,q_2,q_3)=(0.5,0.6,0.7)$. Then

$$\tilde{q}=\sqrt{\frac{1}{3}\sum_{i=1}^3\prod_{j\neq i}^3 q_j}=0.5972.$$

Now, Let Y_1, Y_2, and Y_3 be a set of i.i.d. geometric variables with common parameter $q=0.59$, and Z_1, Z_2, and Z_3 be another set of i.i.d. geometric variables with common parameter $q_{\max}=0.7$. Figure 2 plots the hazard rate functions of $X_{2:3}-X_{1:3}$, $Y_{2:3}-Y_{1:3}$, and $Z_{2:3}-Z_{1:3}$, that is, $\lambda_X(k)$, $\lambda_Y(k)$, and $\lambda_Z(k)$, $k\in\mathbb{N}$. It can be seen that, $\lambda_X(k)$ ranges between $\lambda_Y(k)$ and $\lambda_Z(k)$, and $\lambda_Y(k)$ is always above $\lambda_X(k)$ and $\lambda_Z(k)$. This implies $Z_{2:3}-Z_{1:3}\geq_{\text{hr}}X_{2:3}-X_{1:3}\geq_{\text{hr}}Y_{2:3}-Y_{1:3}$.

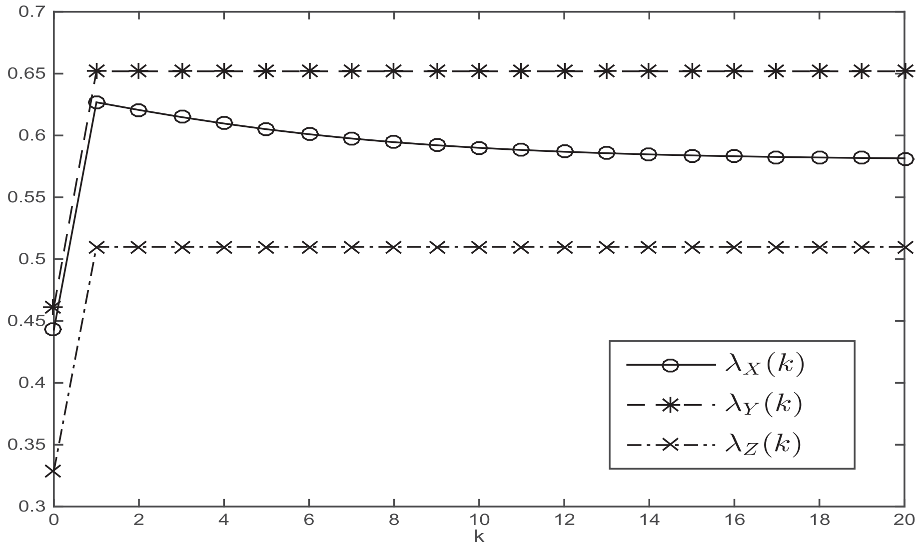

FIGURE **2.** Hazard rate functions of $X_{2:3} - X_{1:3}$, $Y_{2:3} - Y_{1:3}$, and $Z_{2:3} - Z_{1:3}$.

3. SAMPLE RANGE

This section carries out stochastic comparisons of sample ranges arising from heterogeneous or homogeneous geometric samples. In what follows, we let $\mathcal{N}_n = \{1, 2, \ldots, n\}$ and denote the cardinality of some finite set $\mathcal{I} \subset \mathcal{N}_n$ by $|\mathcal{I}|$. Let $X_1, \ldots, X_n$ be a sequence of independent geometric variables with respective parameters $q_1, \ldots, q_n$. Denote by L the length of the first tie, that is, $L := |\{1 \le i \le n, X_i = X_{1:n}\}|$. Define the event $E_{\mathcal{I}}$ as

$$E_{\mathcal{I}} := \{X_j = X_{1:n} < \min_{l \in \mathcal{N}_n \setminus \mathcal{I}} X_l, j \in \mathcal{I}\}.$$

Besides, the event $\{X_j = X_{1:n} = i < \min_{l \in \mathcal{N}_n \setminus \mathcal{I}} X_l, j \in \mathcal{I}\}$ is abbreviated to $\{E_{\mathcal{I}} = i\}$ for some $i \in \mathbb{N}_+$. The following result provides an explicit representation for the distribution function of sample range from a sequence of independent and heterogeneous geometric variables.

PROPOSITION 3.1: Let $X_1, \ldots, X_n$ be a sequence of independent geometric random variables with respective parameters $q_1, \ldots, q_n$. Then, it follows that

$$\mathsf{P}(X_{n:n} - X_{1:n} \le k) = \frac{\prod_{j=1}^n (1 - q_j^{k+1}) - \prod_{j=1}^n (q_j - q_j^{k+1})}{1 - \prod_{j=1}^n q_j}, \quad k \in \mathbb{N}. \tag{8}$$

PROOF: For some fixed $|\mathcal{I}| = m$, we have

$$\begin{aligned}\mathsf{P}(X_{n:n} - X_{1:n} \le k, L = m, E_{\mathcal{I}}) &= \sum_{i=1}^{\infty} \mathsf{P}(X_{n:n} - X_{1:n} \le k, L = m, E_{\mathcal{I}} = i)\\ &= \sum_{i=1}^{\infty} \prod_{j \in \mathcal{I}} (1 - q_j) q_j^{i-1} \prod_{j \in \mathcal{N}_n \setminus \mathcal{I}} \mathsf{P}(i + 1 \le X_j \le k + i)\\ &= \sum_{i=1}^{\infty} \prod_{j \in \mathcal{I}} (1 - q_j) q_j^{i-1} \prod_{j \in \mathcal{N}_n \setminus \mathcal{I}} (q_j^i - q_j^{k+i})\end{aligned}$$

https://doi.org/10.1017/S0269964817000134 Published online by Cambridge University Press

$$= \prod_{j\in\mathcal{I}} \frac{1-q_j}{q_j} \prod_{j\in\mathcal{N}_n\setminus\mathcal{I}} (1-q_j^k) \sum_{i=1}^{\infty} \prod_{j=1}^{n} q_j^i$$

$$= \frac{\prod_{j=1}^n q_j}{1-\prod_{j=1}^n q_j} \prod_{j\in\mathcal{I}} \frac{1-q_j}{q_j} \prod_{j\in\mathcal{N}_n\setminus\mathcal{I}} (1-q_j^k)$$

$$= \frac{\prod_{j=1}^n (1-q_j)}{1-\prod_{j=1}^n q_j} \prod_{j\in\mathcal{N}_n\setminus\mathcal{I}} \frac{q_j-q_j^{k+1}}{1-q_j}.$$

Consequently, the distribution of $X_{n:n} - X_{1:n}$ can be written as

$$\begin{aligned}
\mathsf{P}(X_{n:n} - X_{1:n} \le k) &= \sum_{m=1}^{n} \mathsf{P}(X_{n:n} - X_{1:n} \le k, L = m) \\
&= \sum_{m=1}^{n} \sum_{\mathcal{I}\subseteq\mathcal{N}_n, |\mathcal{I}|=m} \mathsf{P}(X_{n:n} - X_{1:n} \le k, L = m, E_{\mathcal{I}}) \\
&= \sum_{m=1}^{n} \sum_{\mathcal{I}\subseteq\mathcal{N}_n, |\mathcal{I}|=m} \sum_{i=1}^{\infty} \mathsf{P}(X_{n:n} - X_{1:n} \le k, L = m, E_{\mathcal{I}} = i) \\
&= \frac{\prod_{j=1}^n (1-q_j)}{1-\prod_{j=1}^n q_j} \sum_{m=1}^{n} \sum_{\mathcal{I}\subseteq\mathcal{N}_n, |\mathcal{I}|=m} \prod_{j\in\mathcal{N}_n\setminus\mathcal{I}} \frac{q_j-q_j^{k+1}}{1-q_j} \qquad (9) \\
&= \frac{\prod_{j=1}^n (1-q_j)}{1-\prod_{j=1}^n q_j} \left(\prod_{j=1}^{n} \left(1 + \frac{q_j-q_j^{k+1}}{1-q_j}\right) - \prod_{j=1}^{n} \frac{q_j-q_j^{k+1}}{1-q_j} \right) \\
&= \frac{\prod_{j=1}^n (1-q_j)}{1-\prod_{j=1}^n q_j} \left(\prod_{j=1}^{n} \frac{1-q_j^{k+1}}{1-q_j} - \prod_{j=1}^{n} \frac{q_j-q_j^{k+1}}{1-q_j} \right) \\
&= \frac{\prod_{j=1}^n (1-q_j^{k+1}) - \prod_{j=1}^n (q_j-q_j^{k+1})}{1-\prod_{j=1}^n q_j}.
\end{aligned}$$

Thus, the proof is completed. ∎

Now, we present a result stating that more heterogeneity among the geometric random variables will result in larger sample range in the sense of the usual stochastic ordering.

Theorem 3.2: Let $X_1, \ldots, X_n$ be a sequence of independent geometric variables with respective parameters $q_1, \ldots, q_n$, and let $Y_1, \ldots, Y_n$ be another sequence of independent geometric variables with respective parameters $q_1^*, \ldots, q_n^*$. Then,

$$\boldsymbol{q} \overset{\text{m}}{\succeq} \boldsymbol{q}^* \quad \Longrightarrow \quad X_{n:n} - X_{1:n} \ge_{\text{st}} Y_{n:n} - Y_{1:n}.$$

Proof: In light of Proposition 3.1, it suffices to prove that

$$\Phi(\boldsymbol{q}) = \frac{\prod_{l=1}^n (1-q_l^{k+1}) - \prod_{l=1}^n (q_l-q_l^{k+1})}{1-\prod_{l=1}^n q_l}$$

is Schur-concave in $\boldsymbol{q} = (q_1, \ldots, q_n) \in (0,1)^n$. From Lemma 1.3, we need to prove

$$(q_i - q_j)\left(\frac{\partial \Phi(\boldsymbol{q})}{\partial q_i} - \frac{\partial \Phi(\boldsymbol{q})}{\partial q_j}\right) \le 0, \quad \text{for any } 1 \le i \ne j \le n.$$

Note that, for $i \ne j$,

$$\begin{aligned}
&\left(1 - \prod_{l=1}^n q_l\right)^2 \left(\frac{\partial \Phi(\boldsymbol{q})}{\partial q_i} - \frac{\partial \Phi(\boldsymbol{q})}{\partial q_j}\right) \\
&= \left[\prod_{l=1}^n (1 - q_l^{k+1}) - \prod_{l=1}^n (q_l - q_l^{k+1})\right] (q_j - q_i) \prod_{l \ne i,j} q_l \\
&\quad + \Bigg[(k+1) q_j^k \prod_{l \ne j}^n (1 - q_l^{k+1}) + \big(1 - (k+1) q_j^k\big) \prod_{l \ne j}^n (q_l - q_l^{k+1}) \\
&\quad - (k+1) q_i^k \prod_{l \ne i}^n (1 - q_l^{k+1}) - \big(1 - (k+1) q_i^k\big) \prod_{l \ne i}^n (q_l - q_l^{k+1})\Bigg] \left(1 - \prod_{l=1}^n q_l\right) \\
&= \Delta_1(k) + (k+1)\left(1 - \prod_{l=1}^n q_l\right) \Delta_2(k),
\end{aligned}$$

where

$$\begin{aligned}
\Delta_1(k) = &\left[\prod_{l=1}^n (1 - q_l^{k+1}) - \prod_{l=1}^n (q_l - q_l^{k+1})\right] (q_j - q_i) \prod_{l \ne i,j} q_l \\
&+ \left[\prod_{l \ne j}^n (q_l - q_l^{k+1}) - \prod_{l \ne i}^n (q_l - q_l^{k+1})\right] \left(1 - \prod_{l=1}^n q_l\right)
\end{aligned}$$

and

$$\Delta_2(k) = q_j^k \prod_{l \ne j}^n (1 - q_l^{k+1}) - q_j^k \prod_{l \ne j}^n (q_l - q_l^{k+1}) - q_i^k \prod_{l \ne i}^n (1 - q_l^{k+1}) + q_i^k \prod_{l \ne i}^n (q_l - q_l^{k+1}).$$

Without loss of generality, assume that $q_i \le q_j$. Then, we need to show that $\Delta_1(k) \ge 0$ and $\Delta_2(k) \ge 0$. First, it is easy to check that $\Delta_1(0) = \prod_{l=1}^n (1 - q_l)(q_j - q_i) \ge 0$. For $k \in \mathbb{N}_+$, from the observation

$$\prod_{l=1}^n (1 - q_l^{k+1}) - \prod_{l=1}^n (q_l - q_l^{k+1}) = \prod_{l=1}^n (1 - q_l^{k+1}) - \prod_{l=1}^n q_l \prod_{l=1}^n (1 - q_l^k) \ge (1 - \prod_{l=1}^n q_l) \prod_{l=1}^n (1 - q_l^k),$$

we have

$$\begin{aligned}
\Delta_1(k) \ge &\left(1 - \prod_{l=1}^n q_l\right) \prod_{l=1}^n (1 - q_l^k)\,(q_j - q_i) \prod_{l \ne i,j} q_l \\
&+ \left[\prod_{l \ne j}^n (q_l - q_l^{k+1}) - \prod_{l \ne i}^n (q_l - q_l^{k+1})\right] \left(1 - \prod_{l=1}^n q_l\right)
\end{aligned}$$

https://doi.org/10.1017/S0269964817000134 Published online by Cambridge University Press

$$
\begin{aligned}
&\overset{\text{sgn}}{=} \prod_{l=1}^{n}(1-q_l^k)\,(q_j-q_i)\prod_{l\neq i,j} q_l + \left[\prod_{l\neq j}^{n}(q_l-q_l^{k+1}) - \prod_{l\neq i}^{n}(q_l-q_l^{k+1})\right] \\
&= \prod_{l\neq i,j}^{n}(q_l-q_l^{k+1})(1-q_i^k)(1-q_j^k)(q_j-q_i) + \prod_{l\neq i,j}^{n}(q_l-q_l^{k+1})(q_i-q_i^{k+1}-q_j+q_j^{k+1}) \\
&\overset{\text{sgn}}{=} (1-q_i^k)(1-q_j^k)(q_j-q_i) + q_i - q_i^{k+1} - q_j + q_j^{k+1} \\
&= q_j^k q_i - q_i^{k+1}q_j^k + q_i^k q_j^{k+1} - q_i^k q_j \\
&= q_i q_j\left[q_j^{k-1}(1-q_i^k) - q_i^{k-1}(1-q_j^k)\right] \geq 0.
\end{aligned}
$$

Next, we prove $\Delta_2(k) \geq 0$, which can be acquired from

$$
\begin{aligned}
\Delta_2(k) &= \left[q_j^k(1-q_i^{k+1}) - q_i^k(1-q_j^{k+1})\right]\prod_{l\neq i,j}^{n}(1-q_l^{k+1}) \\
&\qquad - \left[q_j^k(q_i-q_i^{k+1}) - q_i^k(q_j-q_j^{k+1})\right]\prod_{l\neq i,j}^{n}(q_l-q_l^{k+1}) \\
&\geq \left[q_j^k(1-q_i^{k+1}) - q_i^k(1-q_j^{k+1})\right]\prod_{l\neq i,j}^{n}(q_l-q_l^{k+1}) \\
&\qquad - \left[q_j^k(q_i-q_i^{k+1}) - q_i^k(q_j-q_j^{k+1})\right]\prod_{l\neq i,j}^{n}(q_l-q_l^{k+1}) \\
&\overset{\text{sgn}}{=} q_j^k(1-q_i^{k+1}) - q_i^k(1-q_j^{k+1}) - q_j^k(q_i-q_i^{k+1}) + q_i^k(q_j-q_j^{k+1}) \\
&= q_j^k(1-q_i) - q_i^k(1-q_j) \geq 0.
\end{aligned}
$$

Thus, we finish the proof. ■

Here is an example as the illustration of Theorem 3.2

Example 3.3: Let (X_1, X_2, X_3) be a vector of independent geometric variables with parameter vector $(q_1, q_2, q_3) = (0.8, 0.6, 0.2)$ and let (Y_1, Y_2, Y_3) be another vector of independent geometric variables with parameter vector $(q_1^*, q_2^*, q_3^*) = (0.6, 0.6, 0.4)$. Clearly, $(q_1, q_2, q_3) \overset{\text{m}}{\succeq} (q_1^*, q_2^*, q_3^*)$. The distribution functions of $X_{3:3} - X_{1:3}$ and $Y_{3:3} - Y_{1:3}$ are plotted in Figure 3, from which it is seen that the distribution function of $Y_{3:3} - Y_{1:3}$ is always larger than that of $X_{3:3} - X_{1:3}$. Hence, $X_{3:3} - X_{1:3} \geq_{\text{st}} Y_{3:3} - Y_{1:3}$.

In what follows, we turn to investigate the case wherein one sample is homogeneous and the other is heterogeneous. Before proceeding to the main result, we recall a very useful lemma due to [11].

Lemma 3.4: Let $\Theta \subset \mathbb{R}$ be a subset of a real line and U be a nonnegative random variable having a cumulative distribution function belonging to a stochastically ordered family $\mathcal{P} = \{\mathsf{H}(\cdot|\theta), \theta \in \Theta\}$; that is, for $\theta_1, \theta_2 \in \Theta$, $\mathsf{H}(\cdot|\theta_1) \geq_{\text{st}} \mathsf{H}(\cdot|\theta_2)$, whenever $\theta_1 < \theta_2$. Suppose a real function $\psi(u, \theta)$ on $\mathbb{R} \cdot \Theta$ is measurable in u for each θ such that $\mathsf{E}_\theta[\psi(U, \theta)]$ exists. Then, $\mathsf{E}_\theta[\psi(U, \theta)]$ is decreasing in θ, if $\psi(u, \theta)$ is decreasing in θ and increasing in u.

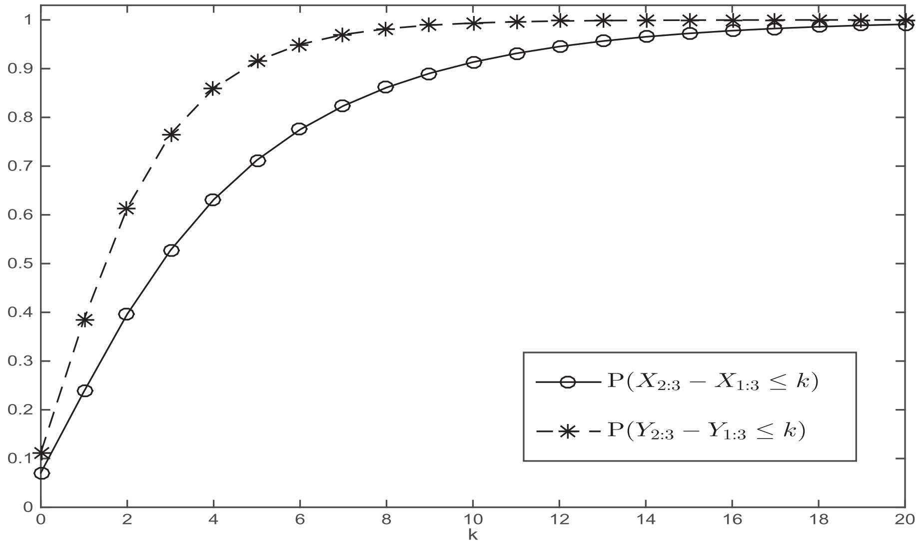

FIGURE **3.** Distribution functions of $X_{3:3} - X_{1:3}$ and $Y_{3:3} - Y_{1:3}$.

First, the monotonicity of the distribution function of sample range from i.i.d. geometric sample is investigated with respect to the parameter q. This reveals that the usual stochastic ordering of geometric random variables is closed under the formation of ranges, and hence is independent of interest.

LEMMA 3.5: Let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ be two sets of i.i.d. geometric variables with respective parameters q and q^* such that $q > q^*$. Then, $X_{n:n} - X_{1:n} \geq_{\text{st}} Y_{n:n} - Y_{1:n}$.

PROOF: According to Equation (9) in the proof of Proposition 3.1, the distribution function of the sample range arising from $X_1, \ldots, X_n$ can be written as, for $k \in \mathbb{N}$,

$$\begin{aligned}
\mathsf{P}(X_{n:n} - X_{1:n} \leq k) &= \sum_{m=1}^{n} \mathsf{P}(X_{n:n} - X_{1:n} \leq k, L = m) \\
&= \sum_{m=1}^{n} \sum_{\mathcal{I} \subseteq \mathcal{N}_n, |\mathcal{I}|=m} \mathsf{P}(X_{n:n} - X_{1:n} \leq k, L = m, E_{\mathcal{I}}) \\
&= \sum_{m=1}^{n} \left[\sum_{\mathcal{I} \subseteq \mathcal{N}_n, |\mathcal{I}|=m} \frac{q^n}{1-q^n} \left(\frac{1-q}{q} \right)^m \right] (1-q^k)^{n-m} \\
&= \sum_{m=1}^{n} \left[\binom{n}{m} \frac{q^n}{1-q^n} \left(\frac{1-q}{q} \right)^m \right] (1-q^k)^{n-m}.
\end{aligned}$$

Denote by $\mathsf{P}(L = m|q)$ the probability of the event $\{L = m\}$ (the length of the first tie equals m) when the underlying geometric random samples have common parameter q. Observing that the distribution function of the random variable L belongs to the family $\mathcal{P} = \{H(\cdot|q), q \in (0,1)\}$ with probability mass function

$$\mathsf{P}(L = m|q) = \binom{n}{m} \frac{q^n}{1-q^n} \left(\frac{1-q}{q} \right)^m,$$

https://doi.org/10.1017/S0269964817000134 Published online by Cambridge University Press

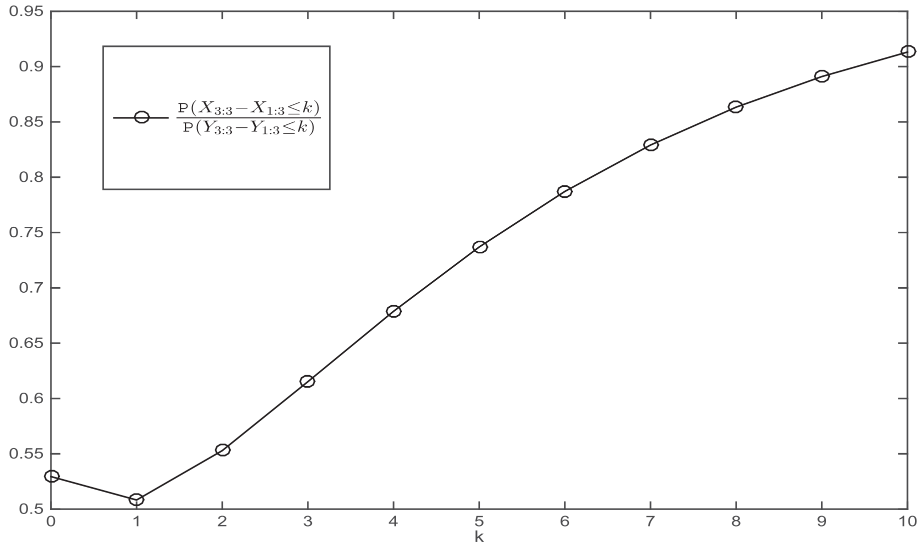

FIGURE 4. Ratio of the distribution functions of $X_{3:3} - X_{1:3}$ and $Y_{3:3} - Y_{1:3}$.

such that $\sum_{m=1}^{n} \mathsf{P}(L = m|q) = 1$, we get

$$\mathsf{P}(X_{n:n} - X_{1:n} \le k) = \sum_{m=1}^{n} \mathsf{P}(L = m|q)(1 - q^k)^{n-m}$$
$$= \mathsf{E}_q\left[\psi(L, q)\right],$$

here, $\psi(L, q) = (1 - q^k)^{n-L}$. On the one hand, for any $1 > q_1 > q_2 > 0$, it holds that

$$\frac{\mathsf{P}(L = m|q_1)}{\mathsf{P}(L = m|q_2)} = \frac{\binom{n}{m}\frac{q_1^n}{1-q_1^n}\left(\frac{1-q_1}{q_1}\right)^m}{\binom{n}{m}\frac{q_2^n}{1-q_2^n}\left(\frac{1-q_2}{q_2}\right)^m} = \frac{q_1^n(1 - q_2^n)}{q_2^n(1 - q_1^n)} \cdot \left[\frac{q_2(1 - q_1)}{q_1(1 - q_2)}\right]^m$$

is decreasing in $m \in \mathcal{N}_n$. Thus, we know that $\mathsf{H}(\cdot|q_1) \le_{\mathrm{lr}} \mathsf{H}(\cdot|q_2)$, which in turn implies that $\mathsf{H}(\cdot|q_1) \le_{\mathrm{st}} \mathsf{H}(\cdot|q_2)$. On the other hand, it is plain that $\psi(m, q) = (1 - q^k)^{n-m}$ is decreasing in $q \in (0, 1)$ and increasing in $m \in \mathcal{N}_n$ for any $k \in \mathbb{N}$. Upon applying Lemma 3.4, we know that $\mathsf{E}_q\left[\psi(L, q)\right]$ is decreasing in $q \in (0, 1)$. Hence, the proof is completed. ■

Combining Theorem 3.2 and Lemma 3.5, the following result can be obtained.

THEOREM 3.6: Let $X_1, \ldots, X_n$ be a sequence of independent geometric variables with respective parameters $q_1, \ldots, q_n$, and let $Y_1, \ldots, Y_n$ be another sequence of i.i.d. geometric variables with a common parameter q. Then, if $q \le \bar{q}$, we have $X_{n:n} - X_{1:n} \ge_{\mathrm{st}} Y_{n:n} - Y_{1:n}$.

PROOF: Let $Z_1, \ldots, Z_n$ be a sequence of i.i.d. geometric variables with a common parameter $\bar{q}$. From Theorem 3.2, it holds that $X_{n:n} - X_{1:n} \ge_{\mathrm{st}} Z_{n:n} - Z_{1:n}$. On the other hand, the result of Lemma 3.5 implies that $Z_{n:n} - Z_{1:n} \ge_{\mathrm{st}} Y_{n:n} - Y_{1:n}$ since $q \le \bar{q}$. Thus, $X_{n:n} - X_{1:n} \ge_{\mathrm{st}} Y_{n:n} - Y_{1:n}$. ■

At the end, we illustrate the result in Theorem 3.6 by one example, which also give a negative answer for the question whether the result can be improved to the reversed hazard rate ordering.

Example 3.7: Let X_1, X_2, X_3 be a sequence of independent geometric variables with respective parameters $(q_1, q_2, q_3) = (0.1, 0.6, 0.8)$, and let Y_1, Y_2, Y_3 be another sequence of independent geometric variables with parameters $(q_1^*, q_2^*, q_3^*) = (0.5, 0.5, 0.5)$. It is easy to see that $(0.1, 0.6, 0.8) \succeq_{\mathrm{m}} (0.5, 0.5, 0.5)$ and $q = q_1^* = q_2^* = q_3^* = 0.5 = (q_1 + q_2 + q_3)/3$. Figure 2 plots the ratio of distribution functions of $X_{3:3} - X_{1:3}$ and $Y_{3:3} - Y_{1:3}$, which indicates that the ratio is smaller than 1 but not monotone in $k \in \mathbb{N}$. Hence, $X_{n:n} - X_{1:n} \ngeq_{\mathrm{rh}} Y_{n:n} - Y_{1:n}$, while $X_{n:n} - X_{1:n} \geq_{\mathrm{st}} Y_{n:n} - Y_{1:n}$.

It is still an open problem whether $X_{n:n} - X_{1:n} \geq_{\mathrm{hr}} Y_{n:n} - Y_{1:n}$ holds.

4. CONCLUDING REMARKS

Spacings are of great interest in many areas such as statistics and reliability. Due to the complex distribution theory of spacings for discrete random variables, few work has been done on the stochastic properties of spacings from geometric random variables. To fill this gap, this paper studies the ordering properties of the second spacings and the sample ranges arising from independent and heterogeneous geometric variables in the sense of the usual stochastic and hazard rate orderings.

To conclude the paper, it is also of interest to study the ordering properties of the general spacings $X_{i:n} - X_{1:n}$ and the simple spacings $X_{i:n} - X_{i-1:n}$ from a heterogeneous geometric sample $X_1, \ldots, X_n$, for $i = 1, \ldots, n$, which are worth further investigating and left as open problems.

Acknowledgements

The authors are grateful to the AE and two anonymous referees for their insightful and constructive comments, which considerably resulted in an improvement in the presentation of this manuscript. This work was supported by the National Natural Science Foundation of China (grant no. 11422109), Natural Science Foundation of Hebei Province (grant no. A2014209240), Hong Kong PhD Fellowship Scheme (PF14-13413) by Hong Kong Research Grants Council, Natural Science Foundation of Jiangsu Province (grant no. BK20141145) and A Project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions.

References

1. Du, B., Zhao, P., & Balakrishnan, N. (2012). Likelihood ratio and hazard rate orderings of the maxima in two multiple-outlier geometric samples. *Probability in the Engineering and Informational Sciences* 26: 375–391.
2. Jeske, D.R. & Blessinger, T. (2004). Tunable approximations for the mean and variance of the maximum of heterogeneous geometrically distributed random variables. *The American Statistician* 58(4): 322–327.
3. Hardy, G.H., Littlewood J.E., & Pólya, G. (1929). Some simple inequalities satisfied by convex function. *Messenger of Mathematics* 58: 145–152.
4. Hardy, G.H., Littlewood J.E., & Pólya, G. (1934). *Inequalities*. Cambridge: Cambridge University Press.
5. Kochar, S.C. (2012). Stochastic comparisons of order statistics and spacings: A review. *ISRN Probability and Statistics*, vol. 2012, Article ID 839473, 47 pages.
6. Kochar, S.C. & Xu, M. (2011). Stochastic comparisons of spacings from heterogeneous samples. In Martin T. Wells and Ashis SenGupta (eds.), *Advances in directional and linear statistics*. Physica-Verlag HD, pp. 113–129.
7. Lundberg, B. (1955). Fatigue life of airplane structures. *Journal of the Aeronautical Science* 22: 394.
8. Mao, T. & Hu, T. (2010). Equivalent characterizations on ordering of order statistics and sample ranges. *Probability in the Engineering and Informational Sciences* 24: 245–262.

9. Margolin, B.H. & Winokur, H.S., Jr. (1967). Exact moments of the order statistics of the geometric distribution and their relation to inverse sampling and reliability of redundant systems. *Journal of the American Statistical Association* 62(319): 915–925.
10. Marshall, A.W., Olkin, I., & Arnold, B.C. (2011). *Inequalities: Theory of majorization and its applications*, 2nd edition. New York: Springer-Verlag.
11. Misra, N. & van der Meulen, E.C. (2003). On stochastic properties of m-spacings. *Journal of Statistical Planning and Inference* 115(2): 683–697.
12. Shaked, M. & Shanthikumar, J.G. (2007). *Stochastic orders*. New York: Springer-Verlag.
13. Xu, M. & Hu, T. (2011). Order statistics from heterogeneous negative binomial random variables. *Probability in the Engineering and Informational Sciences* 25: 435–448.