Statistics and Applications {ISSN 2454-7395 (online)}
Volume 19, No.2, 2021(New Series), pp 217 – 230

Uniform spacings — a Bird's-Eye View

Sthitadhi Das[1] **and Saran Ishika Maiti**[1]
[1] *Department of Statistics, Visva-Bharati, Santiniketan, India*

Received: 17 July 2020; Revised: 20 December 2020; Accepted: 23 December 2020

Abstract

For a distribution, spacing is defined as the gap between order statistics. In characterization of any distribution, spacings play a pivotal role. Spacing originating from uniform distribution is called uniform spacing. Identical distribution of the first and any k-th spacings for some $k = 2, \cdots, n$ of a sample of size n guarantees a uniform distribution structure of parent population, subject to some underlying conditions. The uniqueness and tractability of uniform spacings propelled them as the focal point of many statistical investigations. However, for the regular statistics practitioners, the theory of spacings remain outside frontiers. In an effort to fill the lacuna, this article presents a succinct and lucid review of related results and applications of uniform spacings.

Key words: Order statistics; Spacings; Uniform distribution; Exponential distribution; Characterization of distribution.

AMS Subject Classifications: 60E05; 62G30; 62E10

1. Introduction

Spacing literally means *gaps* or *distance between two successive points*. In statistics, spacings gauge the distance between two successive order statistics. Let $X_1, X_2, \cdots, X_n$ be a set of independent and identical random variables from a continuous distribution function F with support $[a, b]$. Let the corresponding order statistics be $a < X_{1:n} < X_{2:n} < \cdots < X_{n:n} < b$. By j-th spacing $Y_{j:n}$, we mean

$$Y_{j:n} = X_{j+1:n} - X_{j:n}\,;\ j = 0, 1, 2, \cdots, n. \tag{1}$$

In particular, we assume two marginal order statistics as $X_{0:n} = a$ and $X_{n+1:n} = b$. Thus, in particular, $Y_{0:n} = X_{1:n} - a$, $Y_{n:n} = b - X_{n:n}$. So clearly, for n random variables there would be $(n+1)$ gaps or spacings. Theory of spacings gained steam in many fields of statistics —- goodness of fit tests, statistical estimation theory, reliability analysis, survival analysis and applications to name a few. For inciting readers' interest, a quick flavour on applicability of spacings may be presented from its latest advancement in estimation theory.

In estimating an unknown parameter $\theta \in \Theta$, under a distribution $F_\theta(.)$, Maximum Likelihood Estimation (MLE) is a widely used technique. Moreover, MLE is asymptotically unbiased and efficient under some regularity conditions. As an alternative to MLE some

Corresponding Author: Saran Ishika Maiti
Email: saranishika.maiti@visva-bharati.ac.in

authors proposed estimation process based on spacings. Cheng and Amin (1983) suggested the Maximum Product Spacing Estimator(MPSE) by maximizing

$$G = \left(\prod_{i=0}^{n} Y_{i:n}\right)^{\frac{1}{n+1}}$$

where $Y_{i:n}$ being the i-th spacing. The estimator of θ by maximizing G is known as the MPSE of θ and denoted by $\hat{\theta_n}$. The MPS estimator of the underlying distribution F_θ would be $F_{\hat{\theta_n}}$. MPSE is specially suited to the cases where one of the parameters is an unknown shift origin. This occurs, for example, in the three parameter lognormal, gamma and Weibull models. For such J-shaped distributions, under the condition of shape parameter less than unity, no stationary point can yield a consistent MLE due to unboundedness of likelihood equations. Thus not only global but also local maximum likelihood estimator breaks down. In fact, Johnson and Kotz (1976) recommended in the three parametric gamma/Weibull distribution that MLE should not be used if shape parameter < 2. On the contrary to MLE, MPS estimation too ensures consistent estimators but under much more general conditions than ML estimation. Also Cheng and Amin (1983) showed that MPSE is asymptotically normal and asymptotically efficient as MLE when both exist.

Later Renneby (1984) proposed independently another attractive alternative method as an approximation for the Kullback-Leibler measure of information. This spacing estimator is called Maximum Spacing Estimator (MSE), that can be obtained by maximizing

$$S = \frac{1}{n-1}\sum_{j=0}^{n} ln[(n+1)(F_\theta(X_{j+1:n}) - F_\theta(X_{j:n}))]$$

where $F(X_{0:n}) = 0$ and $F(X_{n+1:n}) = 1$. Maximizing S would furnish MSE of θ. MSE is conducive in tracking the true distribution from the angle of empirical distribution function. In order to address on consistency of MPSE/MSE under much weaker regularity conditions as compared to those required in MLE, Shao and Hahn (1999) can be endorsed. The regularity conditions proposed by them are very general in the sense that they cover most of the known counter examples against the universal appeal of the ML method.

Regarding the robustness of spacing estimator, both MPSE and MSE are minimax robust in Hellinger metric neighbourhoods of the given parametric model. A threadbare discussion about the robustness on general m-th order spacing with respect to certain information measure was accommodated in Ekstrorm (2001). He introduced a class of estimation methods, ensuring asymptotically efficient and robust estimator. MSE/MPSE method includes as a special case.

Following the footsteps of MPSE and MSE, Ghosh *et al.* (2001) introduced a general class of maximum spacing estimator.

$$T(\theta) = \sum_{i=1}^{n} h(n(F_\theta(X_{(i)}) - F_\theta(X_{(i-1)}))$$

where $h : (0, \infty) \to \mathbb{R}$ is a strict convex function. For better understanding a handy example might be referred below.

Example 1: Let $X_{(1)}, X_{(2)}, \cdots, X_{(n)}$ be ordered observations from $U(0,\theta)$, $\theta \in (0,\infty)$. The MLE of θ is $\hat{\theta} = X_{(n)}$. Note that here MLE fails to be asymptotic efficient as the regularity conditions for Cramer Rao lower bound of an estimate are not met. On the contrary, generalized spacing estimate of θ is $\hat{\theta_{GSE}} = \frac{n}{n-1}X_{(n)}$ which is obtained by minimising $T(\theta) = \sum_{i=1}^{n} h(n(F_\theta(X_{(i)}) - F_\theta(X_{(i-1)}))$ with $h(x) = -logx$. Both MLE and GSE are consistent but their asymptotic distributions are different.

$$n(\theta - \theta_{MLE}) \xrightarrow{d} Exponential(0,\theta)$$
$$n(\theta - \theta_{GSE}) \xrightarrow{d} Exponential(-\theta,\theta)$$

The two asymptotic distributions have the same variance, but the first has expectation θ while the second has expectation zero and thus is centered better. Moreover,
$E(n(\theta - \theta_{MLE}))^2 \to 2\theta^2$ while $E(n(\theta - \theta_{MSE}))^2 \to \theta^2$. This yields MLE less admissible. Eventually, generalized spacing estimate is also the UMVUE of θ.

Generalized spacing estimator, under the assumption of existence of continuous derivative of p.d.f. with respect to θ, is consistent, asymptotically normal and robust (Ghosh *et al.* (2001)). These estimators are not always explicitly obtainable but can always be computed through numerical methods. Theory of spacings evolved its journey primarily in goodness of fit test and characterization of distributions. Characterization of distribution, through the view point of spacing, stems from the concept of elementary uniform spacings. Due to its comprehensiveness and computational tractability, uniform spacings are considered as benchmarks in spacing theory.

Under the set up, mentioned in equation (1), if F_θ is considered as $uniform(0,1)$ then $\{Y_{j:n,j\geq 1}\}$ is the sequence of uniform spacing variables. Clearly the sum $Y_{0:n} + Y_{1:n} + Y_{2:n} + \cdots + Y_{n:n} = 1$. Due to this linear constraint the random vector $\mathbf{Y} = (Y_{0,n}, Y_{1:n}, Y_{2:n}, \cdots, Y_{n:n})$ has a singular distribution. Moreover, $\mathbf{Y}$ has the joint probability density function as $f_{\mathbf{Y}}(y_{0:n}, y_{1:n}, y_{2:n}, \cdots, y_{n:n}) = n!$ if $y_{i:n} \geq 0$ $\forall i$. Also, the distribution of $\mathbf{Y}$ affirms that distribution function is unaltered under any permutation of the co-ordinates. Using this fact, p.d.f of $Y_{i:n}$ can be easily computed. The p.d.f. is $f_{Y_{i:n}}(x) = n(1-x)^{n-1}$, $\forall i$ where $0 < Y_{i:n} < 1$. Clearly, this form is a $beta(1,n)$ distribution.

As the order statistics from any absolutely continuous distribution with distribution function $F(x)$ can be transformed by order preserving probability integral transformation $u = F(x)$ to the order statistics from a uniform distribution, spacing from any continuous distribution can be explained through uniform spacings. This signifies the prime importance of uniform spacings in distribution free interval estimation and many nonparametric applications.

Theory of spacings received its pioneering thrust from Greenwood's (1946) foundational work, although some ground works by Bortkiewicz(1915) and Morant (1920) left a stamp in literature. In those primitive studies, hints of uniform spacings emerged as the distribution of intervals between successive events of Poisson process given the number of events in a specified interval. The basic methodology on the characterization through uniform spacings was documented in the literature by Darling(1953). The most general method on limit theorems of spacings was disseminated by Lecam(1958). A little later, R.Pyke's classic paper 'Spacing' (1965) grabbed the readers' attention wholly on the wide applicability of spacing theory in the context of distribution free goodness of fit tests as well as characterization problem. Pyke's article, much inclined to theory of uniform spacings, unfurled the idea of construction and limiting theory of spacing at length. Further, Wichura (1968) and Bickel (1969) generalized Le Cam's result and thus provided a concise collection of limit theorems in the context of uniform spacings.

More recently, Ali and Mead (1969), Ahsanullah (1989), Gather *et al.* (1968) and Hamedani and Volkmer (2005) did a good deal of work on spacings. Specifically, Huang *et al.* (1979) established that under the assumptions of continuity and super-additivity the identical distribution of the first and the k-th ($k = 2, \cdots, n$) spacing characterizes uniform distribution.

The current article surveys the theoretical developments in the context of uniform spacings that exist in literature so far. This review, mostly, presents a bunch of useful results without delving into the intricate mathematical exposition. Of most interest of this review is results on uniform spacing as they arise in characterization of distributions. Authors' objective is to popularize the results of uniform spacing in characterization of distribution, crafted under a milder tone of discussion. The rest of the article is outlined as follows. Section 2 contains main results related to characterization, based on uniform spacing. Some preliminary ideas on uniform characterization are also included. Additionally, results on ordered uniform spacings are mentioned. In third section, results on asymptotic properties on uniform spacing are documented. Finally, a short conclusion ends the article.

2. Main Results

2.1. Genesis of uniform spacings

This subsection discusses some preliminaries that explore a few basic construction techniques of uniform spacings. Recalling the setting, already mentioned in Introduction, $Y_{1:n}, Y_{2:n}, \cdots, Y_{n:n}$ being the spacings formed from uniform$(0,1)$, the following results are presented.

Result 1: As $f_{Y_{i:n}}(x) = n(1-x)^{n-1}$ for $0 < Y_{i:n} < 1$ and $f_{Y_{i:n},Y_{j:n}}(x,y) = n(n-1)(1-x-y)^{n-1}$ standard technique nails down to deduce

$$E(Y_{i:n}) = \frac{1}{n+1},\ V(Y_{i:n}) = \frac{n}{(n+1)^2(n+2)},\ Cov(Y_{i:n}, Y_{j:n}) = -\frac{1}{(n+1)^2(n+2)}$$

Higher order moments of uniform spacings were derived by Greenwood (1946); Renyi(1953).

Result 2: For positive constants v_i; $i = 1, 2, \ldots, r$ with $r \leq n$ and $\sum_{i=1}^{r} v_i \leq 1$ the joint survival function of $Y_{1:n}, Y_{2:n}, \cdots, Y_{n:n}$ is given by

$$Prob(Y_{1:n} > v_1 \cdots Y_{r:n} > v_r) = (1 - \sum_{i=1}^{r} v_i)^{n-1}.$$

This leads that under $n \to \infty$,

$$Prob(nY_{1:n} > v_1 \cdots, nY_{r:n} > v_r) = \prod_{i=1}^{r}\{exp(-v_i)\}, v_1, \cdots, v_r > 0.$$

So the limiting distribution of $nY_{i:n}$ is Exponential(1).

The following theorem exudes a connective relation for the characterization of uniform spacings variables through exponential distribution.

Theorem 1: Let $E_1, E_2, \cdots, E_n, E_{n+1}$ be a sequence of iid exponential variables. $Y_{1:n}, \cdots, Y_{n+1:n}$ is distributed as $\frac{E_1}{\sum_{i=1}^{n+1} E_i}, \cdots, \frac{E_{n=1}}{\sum_{i=1}^{n+1} E_i}$. Furthermore, let G_{n+1} be a gamma variable with parameter $(n+1)$. Then $Y_{1:n}G_{n+1}, \cdots, Y_{n+1:n}G_{n+1}$ is distributed as $E_1, E_2, \cdots, E_{n+1}$.

As we know that sum of i.i.d. exponentials follow Gamma distribution. Ratio of an exponential random variable and Gamma variable (of which numerator is a member) lies between 0 to 1, Tacitly, from basic sampling distribution theory each ratio $\frac{E_i}{\sum_{i=1}^{n+1} E_i}$ follows a $beta(1, n)$ distribution. Theorem 1 is an important theorem as a lot of results follow from it. Any standard, degree course text book would be a sufficient resource of those basic results.

Further, Theorem 1 could be applied in order to generate spacing variables from Uniform(0,1) directly. By this we mean that it is not necessary to generate $U_1, U_2, \cdots, U_n$ first and then apply some sorting method, rather generating a bunch of uniform spacing variables at first hand. First we generate iid exponential random variables $E_1, E_2, \cdots, E_n$. Next we would compute the sum of all these random variables G. Then using Theorem 1 we could run a loop of continuation as $U_{(j-1)} + \frac{E_j}{G} = U_{(j)}$.

2.2. Results on characterization of uniform spacings

Characterization of any distribution is a certain distributional property of statistic/statistics that uniquely ascertains the probability structure of underlying distribution. Characterization of uniform distribution can be studied via spacings. An insightful investigation was done by Huang *et al.* (1979) where they asserted the identical distributions of two or more spacings characterize an uniform parent distribution. Keeping in mind that all uniform spacings are identically distributed as $beta(1, n)$, one can characterize the parent cdf uniquely with the aid of some distributional properties of the spacings. Naturally the question arises if identical distributions of two or more spacings are sufficient to characterize a uniform distribution. To address this question Huang *et al.* (1979) assumed that under the condition of (i) continuity, (ii) super-additivity (or, sub-additivity), and (iii) boundedness of support of F identical distributions of $Y_{1:n}$ and $Y_{k:n}$ for $k = 1, \ldots, n$ characterize a uniform distribution. Before unveiling the crux of the result, let us have a sneak peek on what super-additivity (or, sub-additivity) of F is.

Definition 1: A distribution function F is super-additive for all $x, y \in \mathbb{R}$ and $x, y, x+y \in support(F)$ if

$$F(x+y) \geq (\leq)\ F(x) + F(y).$$

The final statement of the result comes as follows.

Result 3: Continuity and super-additivity of F, under which $Y_{0;n}$ and $Y_{k;n}$ for some $k = 1, 2, \ldots, n$, have identical distribution, characterize a uniform distribution.

Instead of super-additivity, the bounded support of F too leads to the similar characterization.

Result 4: If F has bounded support and continuous density, then the identical distribution of $Y_{0;n}$ and $Y_{k;n}$ characterizes a uniform distribution.

Next lemma is conducive in establishing some properties of F which emerge as immediate consequences of the identical distribution of $Y_{0;n}$ and $Y_{k;n}$ for some $k = 1, \ldots, n-1$.

Lemma 1: If F is continuous and the spacings $Y_{0;n}$ and $Y_{k;n}$ for some $k = 1, \ldots, n$ have identical distribution, then (a) F(0)=0, (b) if $F(x_1) = F(x_2)$ for some $0 < x_1 < x_2$, then $F(x_1) = 1$, and (c) $F(x) > 0$ for all positive x.

Remark 1: Using this lemma, we can conclude that for some $k = 1, \ldots, n-1$ when $Y_{0;n}$ and $Y_{k;n}$ have identical distribution, the support of F is either a finite interval $[0, a]$ or an infinite one $[0, \infty)$.

The forthcoming results are going to be discussed upon the condition that support of F is either finite or some other stronger conditions implying the support to be finite.

Remark 2: (1) If F is *sub-additive* and *support*(F) is finite, then F is uniform on $[0, a]$, for some $a > 0$.

(2) Neither the identical distribution of $Y_{0;n}$ and $Y_{k;n}$ nor that of $Y_{k;n}$ and $Y_{j;n}$ for some $1 \leq k < j \leq n-1$ solely guarantees that the parent distribution is uniform distribution on $[0, a]$, $a > 0$.

Clearly, all symmetric distribution functions F will satisfy the identical distribution of $Y_{0;n}$ and $Y_{n;n}$ as well as the identical distribution of $Y_{k;n}$ and $Y_{n-k;n}$, $k = 1, \ldots, n-1$. This type of identical distribution of spacings has been utilized in characterization theory by other authors as well. Some of the profound works include characterizations of exponential and geometric distribution by Puri *et al.* (1970), only exponential by Ahsanullah (1976) and the fairly recent work on general class of continuous distributions by Mirakhmedov *et al.* (2013).

Another meticulous finding of Huang *et al.* would surely grip the researchers' attention.

Theorem 2: Let F be a continuous distribution function of a bounded variable X. $Y_{0;n}$ and $Y_{1;n}$ have identical distribution. Moreover, if F has a density f which is continuous on $(0, a)$ with finite limits $f(0+)$ and $f(a-)$, then F is uniform on $[0, a]$.

Remark 3: As the byproduct of Theorem 8, one can list few remarks as mentioned below.

(1) Theorem does not state that for any arbitrary k, $Y_{0;n}$ and $Y_{k;n}$ have identical distribution.

(2) If F has bounded support $[0, a]$ and the regularity conditions mentioned in the theorems hold then the identical distribution of $Y_{n;n}$ and $Y_{n-1;n}$ leads to the fact that F has uniform distribution on $[0, a]$.

(3) However in Theorem 2 some smoothness conditions on F is required, otherwise it might be misleading. For example, if a random sample $X_1, \ldots, X_n$ is drawn from a Bernoulli

distribution with the probability of success $n/(n+1)$, then $Y_{0;n}$ and $Y_{1;n}$ have identical distribution, even though the parent distribution is not uniform.

Again taking cue from Huan *et al.* (1979) two more further explorations are mention-worthy in characterization theory by uniform spacing.

Result 5: (Ahsanullah (1989))
Let F be absolutely continuous with density function f, $F(0) = 0$, $F(1) = 1$, and f is monotonic on $(0, 1)$. Then $F \sim U[0, 1]$ iff there exists a pair (r, n), $2 \leq r \leq n$, such that

$$X_{r;n} - X_{r-1;n} \sim X_{r-1;n} - X_{r-2;n}.$$

Let F be absolutely continuous, symmetric, either super-additive or sub-additive, $F(0+) = 0$, $F(1) = 1$. Then $F \sim U[0, 1]$ iff $X_{n;n} - X_{1;n} \sim X_{n-1;n}$ for some $n \geq 2$.

Result 6: (Madreimov, Petumin (1983))
Let F be continuous and let $X_1, \ldots, X_n, X \sim F$ be independent random variables. Then

(1) $F \sim U[0, 1]$ iff $E(X_{n,n} - X_{i,n}) = Probability(X \in (X_{i,n}\ ,\ X_{n,n}))$ for all $i \in \mathbb{N}$ and $n \geq i$.

(2) $F \sim U[0, 1]$ iff there exists a pair (i, j), $1 \leq i < j \leq n$, such that

$$E(X_{i,n}) = P(X \in (X_{j-1,n}\ ,\ X_{j,n}))\ for\ all\ n \geq 2.$$

Characterization of the uniform distribution is further extended by Hamedani and Volkmer (2005) in the probability structure of more than one step spacing variables.

Result 7: Let us imagine $(s - r)$ th step spacing, (unlike the distance between consecutive order statistics) $X_{s:n} - X_{r:n}$, $s > r$. If F is uniform$(0, a)$, $X_{s:n} - X_{r:n} \sim X_{s+k:n} - X_{r+k:n}\ \forall 1 \leq r < s < s + k \leq n$.

Special case of the above result appears when $s = r + 1$ and $k = 1$.

Theorem 3: Let X be a positive-valued random variable having an absolutely continuous cdf F. If the corresponding pdf f is monotone on support(F) which is an interval, and the above result holds on $s = r + 1$, $k = 1$ for some r, then X has a uniform distribution on *support*(F).

Remark 4: (1) Sometimes Result 7 and the consequent theorem might hold for some $s = r + 1$ and $k = 1$ for some r but it still does not guarantee that F is uniformly distributed. For example, whenever f is symmetric on $[a, b]$ $(i.e.\ f(x) = f(b - a - x))$ for all x, then Theorem 3 holds for the choices $r = k = 1$ and $n = 3$. That is, the monotonicity of f on $[a, b]$ is important to satisfy the condition of the theorem.

(2) If the assumption that support(F) fails to be an interval the above theorem fails as well. Here is a thought provoking example from Hamidani *et al.* Consider the following

pdf

$$f(x) = \begin{cases} \frac{3}{2} & \text{if } x \in \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right] \\ 0 & \text{otherwise} \end{cases}$$

Then f is monotone on support(F)=$\left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right]$. Since f is symmetric as $f(1-x) = f(x)$, Result 7 holds for a particular choice, say, $r = k = 1$ and $n = 3$, but clearly F is not $U(0,1)$.

2.3. Results on ordered uniform spacings

Arranging the spacings $Y_{1:n}, \cdots, Y_{n+1:n}$ in increasing order we obtain ordered uniform spacings. In Levy(1939); Renyi(1953); Barton and David (1956); Pyke (1965) and Devroye (1981) distributional and asymptotic results on smallest and largest uniform spacings, are discussed at length. Relatively little was done in the context of general ordered uniform spacings. A concrete idea on ordered uniform spacings is developed in Bairamov (2010).

Let us denote the ordered spacings as $\Delta_{0:n} < \Delta_{1:n} < \cdots < \Delta_{n+1:n}$ where $\Delta_{i:n}$ being the i-th largest spacing. Tacitly, for $(n+1)$ number of spacings we would have same number of ordered spacings.

Result 8: $(n-k-1)(\Delta_{k+1:n} - \Delta_{k:n}) \stackrel{d}{=} \Delta_{1:n}$ $(k = 0, \cdots, n)$ where $\stackrel{d}{=}$ means that the statistics are distributionally same (Pyke (1965)).

In particular, distribution of the k-th $(1 \leq k \leq n+1)$ ordered uniform spacing $\Delta_{k:n}$ is deduced by Bairamov *et al.* (2010). Let the domain of uniform distribution $(0,1)$ be presented as the sum of non overlapping intervals

$$(0,1) = I_{1,n} \cup I_{2,n} \cup \cdots \cup I_{n+1,n}$$

where $I_{1,n} = (0, \frac{1}{n+1})$ and $I_{m,n} = (\frac{1}{n+3-m}, \frac{1}{n+2-m})$.

Theorem 4: The distribution of the k-th $(1 \leq k \leq n+1)$ ordered uniform spacing is presented by

$$P\{\Delta_{k:n} \leq x\} = 0 \ \ (x < 0)$$
$$P\{\Delta_{k:n} \leq x\} = 1 \ \ (x(n+2-k) \geq 1)$$

and for $x \in I_{m,n}, m = 1, 2 \cdots, k$

$$P\{\Delta_{k:n} > x\} = (-1)^{k-1}(n+1)\binom{n}{k-1} \sum_{i=m}^{k} \frac{(-1)^{i-1}}{(n+2-i)} \binom{k-1}{i-1} (1 - x(n+2-i))^n.$$

As a corollary of the last theorem the distribution of minimal ordered spacing and maximal spacing can be presented.

Corollary 1: The distribution of the minimal spacing is given by

$$P\{\Delta_{1,n} > x\} = (1 - x(n+1))^n, \ x \in \ I_{1,n}$$

Simultaneously, the distribution of maximal spacing is given by

$$P\{\Delta_{n+1,n} > x\} = (-1)^n(n+1)\sum_{i=m}^{n+1}\frac{(-1)^{i-1}}{n+2-i}\binom{n}{i-1}(1-x(n+2-i))^n$$

.

Theorem 4 allows the readers to compute the expectation of k-th ordered uniform spacing which is $\frac{1}{n+1}\sum_{i=n+2-k}^{n+1}\frac{1}{i}$ $(k = 1, \cdots, n+1)$.

3. Limiting Results on Uniform Spacings

Limiting theories for spacings depicts some of the more interesting results. Here, we present several results from the limiting theory of uniform spacings chronologically. Mostly every case affirms on the asymptotic distribution as normal distribution.

Result 9: Levy (1939) obtained the limiting distributions of the maximal spacing $\Delta_{n+1,n}$ = maximum ordered spacing from the uniform distribution on $[-1, 1]$ as

$$P\{\frac{n\Delta_{n+1,n}}{\log n} \leq x\} \rightarrow exp(-exp(-x));\ \ x \in \mathbb{R}$$

Result 10: Devroye (1981) established that

$$\lim_{n\rightarrow\infty} \sup\left[\frac{n\Delta_{n+1,n}}{2\log\log n}\right] = 1\ a.s.$$

Result 11: Let $F(x)$ be a continuous distribution function. If $X_{1:n}, \ldots, X_{n:n}$ is an ordered sample of n values from the population whose distribution function is $F(x)$ then the random variable

$$\omega_n = \frac{1}{2}\sum_{i=1}^{n}\left|F(x_{i:n}) - F(x_{i-1:n}) - \frac{1}{n+1}\right|$$

is asymptotically normally distributed with mean $E(\omega_n)$ and variance $var(\omega_n)$ (Sherman (1947)), *i.e.*, the standardized random variable

$$\frac{\omega_n - E(\omega_n)}{\sqrt{var(\omega_n)}}$$

approaches towards a Standard Normal variate as $n \rightarrow \infty$.

Result 12: Kimball (1947) proved the asymptotic normal distribution of

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

which is also a measure of deviation from uniform spacing.

Result 13: Moran (1972) considered a similar statistic

$$\beta_n = \sum_{i=1}^{n+1} \left(F(X_{i,n}) - F(X_{i-1,n})\right)^2$$

and proved that β_n is asymptotically normal.

All of these results on asymptotic theory of uniform spacings could serve as handy tools for constructing the goodness of fit tests as the distributions are asymptotically normal. For the exact expressions of expectation and variance, readers are recommended to go through the respective articles. Recently, Eryilmaz and Stepanov (2008) studied runs based uniform order statistics and developed even more complex limit results related to uniform spacings from the results obtained for runs. One may be interested in the asymptotic behavior of ordered uniform spacing $\Delta_{k,n}$. A solid and meticulous discussion is found in Bairamov *et al.* (2010). In this article we restrain from mentioning all those critical results as that might spoil the flow of simplicity of the article. Instead we can concentrate our attention on the asymptotic behavior of the expectation of $\Delta_{k,n}$ (k th ordered uniform spacing) which is relatively simpler.

Result 14: (1) For finite $k \geq 1$ and $n \to \infty$,

$$E(\Delta_{k:n}) = O\left(\frac{1}{n^2}\right) \to 0.$$

(2) For finite $k \geq 1$ and $n \to \infty$ we have

$$E(\Delta_{n+2-k:n}) \sim \frac{logn}{n} \to 0.$$

(3) If $k = k_n \to \infty$ such that $k_n = o(n)$, then

$$E(\Delta_{k_n:n}) \sim \frac{k_n}{n^2} \to 0.$$

(4) If $k = k_n \to \infty$ such that $k_n = o(n)$, then

$$E(\Delta_{n+2-k_n:n}) \sim \frac{log(n/k_n)}{n} \to 0.$$

The recent trend in characterization through spacings is escalated by investigations in probability distribution of adjacent spacings. By adjacent spacing we mean the neighborhood around an order statistic, *i.e.*, the points encapsulated in $(X_{k:n} - d, X_{k:n})$ or $(X_{k:n}, X_{k:n} + d)$ where the d may or may not be dependent on n. Pakes and Steutel (1997); Balakrishnan and Stepanov(2005); Dembinska *et al.* (2007); Dembinska and Balakrishnan (2010) are few worth references. Further Nagaraja *et al.* (2014) accelerated this route by discussing joint limiting distribution of adjacent spacings $(Y_{k:n}, \cdots, Y_{k+r:n})$ and $(Y_{k:n}, \cdots, Y_{k-s:n})$ around three types of order statistics in particular – central, intermediate or an extreme order statistic. When $n \to \infty$, these three different scenarios arise and (i) Central case where $\frac{k}{n} \to p,\ 0 < p < 1$,

(ii) Intermediate case where $k, n-k \to \infty$ and $\frac{k}{n} \to 0$, (iii) Extreme case where k or $n-k$ is kept fixed. Borrowing the knowledge of extreme value theory they showed that in the first two cases $(r+s)$ spacings converge weakly to a batch of i.i.d. standard exponential random variables but in extreme case, this weak convergence would hold only in the domain of attraction of Gumbel (heavy tail distribution) or Weibull type (short ended, finite end point distribution).

4. Conclusion

The main purpose of this review article is to elucidate some instructive results related to uniform spacings arising in the context of characterization theory, without much dipping down to mathematical complexity. Other than the usual uniform and exponential spacings, some relevant investigations were done by several authors on the characterization of the Gamma, Normal and Weibull distributions as well as on some discrete distributions like Geometric, Poisson, Negative Binomial distributions etc. Still characterization of uniform distribution through spacings draws fundamental importance in a wide variety of fields. Apart from characterization theory; spacing, specially uniform spacing, is applicable in few other fields as well e.g. distribution-free goodness of fit test, information theory, time series analysis etc. A couple of topics of interest might be worth mentioning here.

Goodness of fit test is used to verify if the distribution function is equal to a specified one. By probability integral transformation, any specified continuous distribution function can be converted to uniform(0,1). Thus the null hypothesis of interest boils down to $H_0 : F(x) = x,\ 0 \le x \le 1$. To test this H_o, Greenwood (1946) introduced a statistic on the basis of sum of squares of disjoint uniform spacings. The statistic, called Greenwood statistic, is $G = \frac{1}{n}\sum_0^{n-1}(nY_{(i:n)})^2$. By virtue of Pitman asymptotic efficiency (AE) Greenwood test is proved as optimal among the goodness of fit tests within the class of symmetric tests.

All the more, Greenwood statistic could be generalized by $G_n = n^{-1}\sum_{i=0}^{n-1} h(nY_{(i:n)})$ where $h(.)$ is a function that satisfies some mild regularity conditions. Clearly, the choice of $h(x) = x^2$ would turn it to Greenwood's original statistic. $h(x) = x^r$ for $r > 0$ was proposed by Kimball (1950) who obtained the limiting distribution of statistic under H_0. Another popular choice of $h(.)$ is $h(x) = logx$, suggested by Darling (1953).

Periodogram analysis is an effective device in frequency analysis of time series. For a stationary, random time series periodogram ordinates are exponentially distributed and independent. Actually ordinates have the same distribution as the uniform distributions of the spacings. Thus to test for the peak , *i.e.*, the largest ordinate of periodogram test statistic might be considered as largest uniform spacing (For details see Durbin (1960)). Also, in the context of distribution of serial/auto correlation coefficients; any order of serial correlation coefficient can be put in the form of linear functions of spacings. The joint distribution of several linear functions of the spacings was derived by Watson (1956) which could explain out the distribution of serial correlations.

In applied economics, auction theory holds current trend of attraction. Motivated by the upsurge of auctions in online advertisements, like auction through eBay and Amazon, the query on expected revenues in auctions is quite of interest in recent years. Uniform spacings (or spacings as a whole) might come out as an effective tool in such stochastic auction theory where the following spacings–$Y_{2:n}$ and $Y_{n:n}$ would represent auction rents in buyers' auction and reverse auction in the second-price business auction under identical bids. One might be inquisitive on stochastic modeling of second price (explained by second order uniform spacing) or if the expected revenues depending on the number of bidders.

So far what we presented here is all about univariate spacings. An intrinsic question might trigger regarding the exact distribution of bivariate/multivariate uniform spacings. There are myriad examples in which samples are drawn from bivariate/multivariate set-up for which it is pertinent to study the spacings of the observations. As a stepping stone, one can start with multivariate uniform distribution and investigate on multivariate uniform spacings accordingly. Barton and David (1962) studied the distribution on spacings computed on the random points drawn on the two-dimensional plane but still now number of organized works in multivariate spacings is almost nil, probably due to its degree of computational difficulty.

Acknowledgement

This research work has been made under the financial support from the major research project (EMR/2017/005421) sanctioned by Science and Engineering Board(SERB), DST, Govt. of India. Both of the authors are indebted to the Editor, unknown referee and Prof. Jyotirmoy Sarkar, IUPUI,USA for their constructive suggestions which led to much improvement of this article.

References

Ahsanullah, M. (1976). On a characterization of the exponential distribution by order statistics. *Journal of Applied Probability*, **13**, 818-822.

Ahsanullah, M. (1989). On characterizations of the uniform distribution based on functions of order statistics. *Aligarh Journal of Statistics*, **9**, 1-6.

Ali, Mir M., and Mead E. R. (1969). On the distribution of several linear combinations of order statistics from the uniform distribution. *Bulletin of the Institute of Statistical Research and Training*, **3**, 22-41.

Bairamov, I., Alexandre B., and Stepanov A. (2010). Limit results for ordered uniform spacings. *Statistical Papers*, **51(1)**, 227-240.

Balakrishnan, N., Pakes, A. Gn., and Stepanov, A. (2005). On the number and sum of near-record observations. *Advances in Applied Probability*, **37(3)**, 765-780.

Barton, D. E., and F. N. David (1956). Some notes on ordered random intervals. *Journal of the Royal Statistical Society: Series B (Methodological)*, **18(1)**, 79-94.

Bickel, P. J. (1969). A distribution free version of the Smirnov two sample test in the p-variate case. *The Annals of Mathematical Statistics*, **40(1)**, 1-23.

Bortkiewicz, L. Von (1915). *Ü*ber die Zeitfällinger Ereignisse. *Bulletine of International Statistical Institute*, **20(2)**, 30-111.

Cheng, R. C. H., and Amin, N. A. K. (1983). Estimating parameters in continuous univariate distributions with a shifted origin. *Journal of the Royal Statistical Society: Series B (Methodological)*, **45(3)**, 394-403.

Darling, Donald A. (1953). On a class of problems related to the random division of an interval. *The Annals of Mathematical Statistics*, **24**, 239-253.

David, H. A. and H. N. Nagaraja (2003). *Order Statistics*, John Wiley & Sons, New York.

Dembinska, A., Stepanov, A. and Wesolowski, J. (2007). How many observations fall in a neighborhood of an order statistic? *Communications in Statistics-Theory and Methods*, **36(5)**, 851-867.

Dembinska, A. and Balakrishnan, N. (2010). On the asymptotic independence of numbers of observations near order statistics. *Statistics*, **44(5)**, 517-528.

Devroye, Luc (1981). Laws of the iterated logarithm for order statistics of uniform spacings. *The Annals of Probability*, **9**, 860-867.

Durbin, J. (1960). Some methods of constructing exact tests. *Biometrika*, **48(1-2)**, 41-65.

Ekstrom, M. (2001). Consistency of generalized maximum spacing estimates, *Scandinavian Journal of Statistics*, **28(2)**, 343-354.

Eryilmaz, S., and Stepanov, A. (2008). Runs in an ordered sequence of random variables. *Metrika*, **67(3)**, 299-313.

Gather, U., Udo Kamps., and Schweitzer, N. (1998). Characterizations of distributions via identically distributed functions of order statistics. *Handbook of Statistics*, **16**, 257-290.

Ghosh, K., and Rao Jammalamadaka, S. (2001). A general estimation method using spacings. *Journal of Statistical Planning and Inference*, **93(1-2)**, 71-82.

Greenwood, M. (1946). The statistical study of infectious diseases. *Journal of the Royal Statistical Society*, **109(2)**, 85-110.

Hamedani, G. G., and Volkmer, H. W. (2005). Certain characterizations of the uniform dstribution. *Metrika*, **61(2)**, 117-125.

Huang, J. S., Arnold B. C., and Ghosh, M. (1979). On characterizations of the uniform distribution based on identically distributed spacings. *Sankhya: The Indian Journal of Statistics, Series B*, 109-115.

Johnson, N. J., and Kotz, S. (1970). *Continuous Univariate Distributions-2*. Houghton Mifflin, Boston.

Kimball, Bradford F. (1947). Some basic theorems for developing tests of fit for the case of the non-parametric probability distribution function. *The Annals of Mathematical Statistics*, **18(4)**, 540-548.

LeCam, L. (1958). Un théoreme sur la division d'un intervalle par les points pris au hasard. *Publications de Institut de Statistique de l'Université de Paris*, **7**, 7- 16.

Levy, P. (1939). Sur la division d'un segment par des points choisis au hasard. *Comptes rendus de l'Académie des Sciences* , **208**, 147-149.

Madreimov, I., and Petunin, Y. I. (1983). A characterization of the uniform distribution with the aid of order statistics. *Theory of Probability and Mathematical Statistics*, **27**, 105-110.

Mirakhmedov, Sherzod M., and Rao Jammalamadaka, S. (2013). Higher-order expansions and efficiencies of tests based on spacings. *Journal of Nonparametric Statistics*, **25(2)**, 339-359.

Moran, P. A. P. (1947). The random division of an interval. *Supplement to the Journal of the Royal Statistical Society*, **9(1)**, 92-98.

Morant, G. M. (1920). On random occurrences in space and time, when followed by a closed interval. *Biometrika*, **13**, 309-337.

Pakes, A. G., and Steutel, F. W. (1997). On the number of records near the maximum. *Australian Journal of Statistics*, **39(2)**, 179-192.

Puri, P. S., and Herman, R. (1970). A characterization based on the absolute difference of two iid random variables. *The Annals of Mathematical Statistics*, **41(6)**, 2113-2122.

Pyke, R. (1965). Spacings. *Journal of the Royal Statistical Society: Series B (Methodological)*, **27(3)**, 395-436.

Ranneby, B. (1984). The maximum spacing method. An estimation method related to the maximum likelihood method. *Scandinavian Journal of Statistics*, **11**, 93-112.

Renyi, A. (1953). On the theory of order statistics. *Acta Mathematica Academiae Scientiarum Hungarica*, **4(3-4)**, 191-231.

Shao, Y., and Hahn, M. G. (1999). Strong consistency of the maximum product of spacings estimates with applications in nonparametrics and in estimation of unimodal densities. *Annals of Institute of Statistical Mathematics*, **51**, 31-49.

Sherman, B. (1950). A random variable related to the spacing of sample values. *The Annals of Mathematical Statistics*, **21(3)**, 339-361.

Watson, G. S. (1956). On the joint distribution of the circular serial correlation coefficients. *Biometrika*, **43(1/2)**, 161-168.

Wichura, M. J. (1968). On the weak convergence of non-Borel probabilities on a metric space. Unpublished Dissertation, Columbia University, USA.