On The Spacings of Exponential and Uniform order Statistics Prof. Mohamed I. Riffi *

$$[a,b] \label{eq:ab} .(-\ \mbox{$\frac{1}{2}$}\ ,+\mbox{$\frac{1}{2}$}\)$$

ABSTRACT

Order statistics, spacings of order statistics, characterization of distributions, adjacent order statistics.

^{*} Mathematics Department, Islamic University of Gaza, Palestine P.O. Box 108. riffim@mail.iugaza.edu.

INTRODUCTION:

Let X_1, X_2, K , X_n be independent random variables constituting a random sample of size n, n = 1, 2, 3, K from a distribution with absolutely continuous cumulative distribution function (cdf) F; i.e., F has probability density function (pdf) f with respect to the Lebesgue measure. Let $X_{1:n} < X_{2:n} < L < X_{n:n}$ be the order statistics of these iid $D_1 := X_{1:n}$ random variables. and $D_i := X_{i:n} - X_{i-1:n}, i = 1,2,K,n$; i.e., D_i is the difference between the two adjacent order statistics $X_{i:n}$ and $X_{i-1:n}$ with the convention $X_{0:n} = 0$. The random variables D_1, D_2, K, D_n are called the spacings of adjacent order statistics (Casella et al. [4]). These spacings evidently give some idea about the "spread" of a distribution. They also can be used to characterize some distributions like the uniform and exponential distributions. They also can be used to construct confidence intervals for the corresponding population. Therefore, these spacings may be of interest in their own right.

In Section 2, we use the joint distribution of order statistics to derive the cdf of the random variables, D_1, D_2, K , D_n in terms of F and f of the distribution of the random sample.

In Section 3, we study the properties of these spacings in the exponential population case. We write their distributions in a closed form.

In Section 4, we derive the distributions of these spacings in closed form and study their properties.

We will assume throughout the paper, unless otherwise stated, that we are sampling from a distribution with support (- , +). We will use F_i and f_i to denote, respectively, the cdf and pdf of D_i . We will also use $f_{i,j}$

to denote the pdf of $X_{j:n}$ - $X_{i:n}$. We also will use $X \stackrel{d}{=} Y$ to denote that the two random variables X and Y have the same distribution. We will use the symbol W to remark the end of a proof.

DISTRIBUTIONS OF D_i **S**:

Recall that the pdf of the ith order statistic $X_{i:n}$ is given by

$$f_{i}(x) = i \xi^{II} \frac{1}{2\pi i} f(x) F^{i-1}(x) [1 - F(x)]^{n-i}, - \Psi < x < \Psi , \quad (2.1)$$

the joint pdf of $X_{i:n}$ and $X_{j:n}$, $1 \ \pounds \ i \le j \ \pounds \ n$, is given by

$$f_{i,j}(x,y) = \frac{n!}{(i-1)!(j-1-i)!(n-j)!} f(x) f(y) F^{i-1}(x),$$

$$(2.2)$$

$$[F(y) - F(x)]^{j-1-i} [1 - F(y)]^{n-j},$$

for - Y < x < y < Y,

and the joint pdf of $X_{1:n}, X_{2:n}, K, X_{n:n}$ is given by

$$f_{1,2,K,n}(x_1,x_2,K,x_n) = n! f(x_1) f(x_2) L f(x_n), (2.3)$$

for - Ψ < x_1 < x_2 < L < x_n < Ψ .

Lemma 2.1. The cdf F_i^{c} of D_i is given for i = 2, 3, K, n by

$$F_{i}(r) = 1 - \underbrace{(i-1)\xi_{i-1}^{\frac{1}{2}}}_{=\frac{1}{2}} f(x)F^{i-2}(x)[1 - F(x+r)]^{n-i+1} dx, (2.4)$$

for r > 0; and it is given for i = 1 by

$$F_1(r) = 1 - [1 - F(r)]^n,$$
 (2.5)

for r > 0.

Proof. See Riffi [5] for a proof. W

Lemma 2.2. If the support of the underlying distribution is [a, b], where

-
$$\Psi$$
 < a < b < Ψ , then for any 0 < r < b and i = 2,3,K , n ,

$$F_{i}(r) = 1 - \int_{a}^{b-r} (i-1) \xi^{n} \int_{\frac{\pi}{a}}^{\frac{\pi}{a}} f(x) F^{i-1}(x) [1 - F(x+r)]^{n-i} dx (2.6)$$

and P_1^{\bullet} is equal to the cdf of the first order statistics.

Proof. See Riffi [5] for a proof. W

Theorem 2.1. If the support of the underlying distribution is [0, + Y], then we have for i = 1, 2, K, n and r > 0

$$f_{i,j}^{0}(r) := f_{X_{j:n}-X_{i:n}}(r) = \int_{0}^{\infty} f_{i,j}(x,x+r)dx.$$
 (2.7)

Proof. Let S be the region bounded by $X_{j:n}$ 3 $X_{i:n}$ and $X_{j:n}$ £ $X_{j:n}$ + r. Then

On The Spacings of Exponential and Uniform order ...

$$F_{X_{j,n}-X_{j,n}}(r) = f_{i,j}(x,y)dydx$$

$$= \begin{cases} f_{i,j}(x,y)dydx \\ \vdots \\ f_{i,j}(x,y)dydx. \end{cases}$$

By using Leibnitz's rule, we get

We using Leibiniz's rule, we get
$$f_{i,j}^{0}(r) = \frac{\P}{\P r} F_{X_{j:n}-X_{i:n}}(r) = \frac{\P}{\P r} \int_{0}^{\frac{Y}{n}} f_{i,j}(x,y) dy dx$$

$$= \int_{0}^{\frac{Y}{n}} \frac{\P}{\P r} \int_{0}^{x+r} f_{i,j}(x,y) dy \frac{1}{2} dx$$

$$= \int_{0}^{x} f_{i,j}(x,x+r) dx. \qquad W$$
(2.8)

Theorem 2.2. If the support of the underlying distribution is of the form [a, b], then

$$f_{i}(r) = \int_{a}^{b-r} f_{i-1,i}(x, x+r) dx.$$
 (2.9)

Proof. Similar to the proof of the case when the support is [0, Y] . W

Applications to the Exponential Distribution:

In the following theorem, we derive the pdf of the random variable $X_{j:n}$ - $X_{i:n}$ for the order statistics resulting from a random sample of size n from an exponential distribution with parameter i > 0, where i < j. Then we will derive the pdf of the spacings D_i for i = 1, 2, K, n resulting from this random sample.

Theorem 3.1. Let $X_{1:n}, X_{2:n}, K$, $X_{n:n}$ be the order statistics of a random sample of size n from an exponential distribution with parameter l > 0. Then the pdf of $X_{j:n} - X_{i:n}$ for i and j in $\{1, 2, K, n\}$ such that i < j is given by:

$$f_{i,j}^{0}(r) = \frac{(n-i)!}{(n-j)!(j-i-1)!} I e^{-l(n-j+1)r} (1-e^{-lr})^{j-i-1}, r > 0. (2.10)$$

Proof. It follows from (2.2) that:

$$f_{i,j}(x,x+r) = c(r)e^{-i(n-i+1)x}(1-e^{-ix})^{i-1}$$

where the c(i) is given by

$$c(r) := \frac{n!}{(i-1)!(j-i-1)!(n-j)!} I^2 e^{-I(n-j+1)r} (1-e^{-Ir})^{j-i-1}.$$

By using Theorem 2.1 and integration by parts we get for r > 0

$$f_{i,j}^{0}(r) = \int_{0}^{r} f_{i,j}(x, x + r) dx$$

$$= c(r) \int_{0}^{r} e^{-l(n-i+1)} (1 - e^{-lx})^{i-1} dx$$

$$= c(r) \frac{(i-1)!(n-i)!}{ln!}$$

$$= \frac{(n-i)!}{(n-i)!(i-j-1)!} le^{-l(n-j+1)r} (1 - e^{-lr})^{j-i-1}.$$
W

Corollary 3.1. Let $X_{1:n} < X_{2:n} < L < X_{n:n}$ be the order statistics of a random sample of size n from an exponential distribution with parameter l > 0. Then the pdf of $X_{k+i:n} - X_{i:n}$ for i and $1 \pounds k < n$ such that $1 \pounds i < k+i \pounds n$ is given by:

$$f_{i,k+i}^{0}(r) = k \begin{cases} n - i \\ \frac{1}{2} \\ k \end{cases} e^{-i(n-k-i+1)r} (1 - e^{-ir})^{k-1}, r > 0. \quad (2.11)$$

Proof. The proof follows from Theorem 3.1 by replacing j with k+i. Remark 3.1. We note that when k=n-1 and i=1, we get the pdf of the sample range, namely

$$f_{1,n}^{0}(r) = (n-1)Ie^{-Ir}(1-e^{-Ir})^{n-2}$$
. W

Theorem 3.2. Let $X_{1:n} < X_{2:n} < L < X_{n:n}$ be the order statistics of a random sample of size n from an exponential distribution with parameter 1 > 0. Then $X_1 = X_2 = X_1 = X_2 = X_2 = X_3 = X_4 = X_$

1 > 0. Then $X_{j:n} - X_{i:n} \stackrel{d}{=} X_{j-i:n-j}$ for i and j in $\{1,2,K,n\}$ with i < j.

Proof. The proof follows from Theorem 3.1 which gives the pdf of $X_{i:n}$ - $X_{i:n}$ for i < j as

$$f_{i,j}^{0}(r) = \frac{(n-i)!}{(n-j)!(j-i-1)!} I e^{-l(n-j+1)r} (1-e^{-lr})^{j-i-1}, r > 0. \quad (2.12)$$

But (2.1) gives the pdf of the (j - i) th order statistic resulting from a random sample of size n - i as

$$f_{j-1}(x) = (j-1) \begin{cases} f^{j-1} & \text{if } (x)F^{j-j-1}(x)[1-F(x)]^{p-j}, - \frac{1}{2} < x < \frac{1}{2} \end{cases}$$
 (2.13)

Substituting $f(x) = 1e^{-1x}$ and $F(x) = 1 - e^{-1x}$ in (2.13), the assertion follows. W

Theorem 3.3. Let $X_{1:n} < X_{2:n} < L < X_{n:n}$ be the order statistics of a random sample of size \$n\$ from an exponential distribution with parameter I > 0. Then the D_i s are independent and exponentially distributed with parameters I(n - i + 1) for each i = 1, 2, K, n, respectively. In other words, the pdf, I_i , of I_i is given by

$$f_i(r) = I(n - i + 1)e^{-I(n - i + 1)r}, r > 0.$$
 (2.14)

Proof. For the case when i = 1, we have

$$f_1(r) = f_1(r) = nle^{-lr}e^{-l(n-1)r} = nle^{-lnr}, r > 0. (2.15)$$

If
$$i \ge 1$$
, we let $c_i(n) := \frac{n!}{(i-2)!(n-i)!}$. Then for $r \ge 0$,

where we have used integration by parts repeatedly.

To prove that D_1, D_2, K , D_n are mutually independent, we see that the inverse of the transformation that gives each D_i as a function of $X_{1:n}, X_{2:n}, K$, $X_{n:n}$ is given by $X_{i:n} = D_1 + D_2 + L + D_i$, for i = 1, 2, K, n.

The Jacobian of the inverse transformation is equal to 1 since it is the determinant of an n' n triangular matrix. Therefore, by using (3.2), we see that

$$I_{1,2,K,n}^{n}(x_{1},x_{2},K,x_{n}) = I^{n}n!e^{-\delta_{-1!!}i!\pi_{n}}I^{(n+1-i)x_{i}}$$

$$= I^{n}n! \underbrace{e^{-I(n+1-i)x_{i}}}_{1!! i!! n}, i = 1,2,K,n.$$

Since the joint pdf of the D_i s factors into the product of functions of each variable alone, they are mutually independent (See Theorem 4.6.4 of Casella et al. [4]). W

Corollary 3.2. Let $X_{1:n}, X_{2:n}, K, X_{n:n}$ be the order statistics of a random sample of size n from an exponential distribution with parameter l > 0. Let the random variables $D_i^0, i = 1, 2, K, n$ be defined by $D_i^0 := (n - i + 1)D_i$. Then D_i^0, D_i^0, K, D_n^0 are iid with exponential distribution with parameter l.

Proof. The inverse of the transformation r = g(x) = (n - i + 1)x is given by

$$g^{-1}(r) = r/(n-i+1)$$
 with derivative $\frac{d}{dr}g^{-1}(r) = 1/(n-i+1)$. Hence, the pdf of each \mathcal{D}_{i}^{n} is given by

$$f_{p_{j}}(r) = f_{i}(g^{-1}(r)) \mid \frac{d}{dr}g^{-1}(r) \mid$$

$$= Ie^{-Ir}, r > 0.$$

This proves the claim of the corollary. W

APPLICATIONS TO THE UNIFORM DISTRIBUTION

Let $X_{1:n}, X_{2:n}, K$, $X_{n:n}$ be the order statistics of a random sample of size n from a uniform distribution on the finite interval [a,b]. We will denote by $D_{i,j}$ for $1 \, \pounds \, i < j \, \pounds \, n$ to the spacings between $X_{j:n}$ and

 $X_{i:n}$. We also let the pdf of $D_{i,j}$ be denoted by $f_{i,j}^{0}$. When j=i+1, we simply let $D_{j}=X_{j:n}$ - $X_{i:n}$ for $1 < i \pm n$ and $D_{1}=X_{1:n}$.

Theorem 4.1. Let $X_{1:n}, X_{2:n}, K$, $X_{n:n}$ be the order statistics of a random sample of size n from a uniform distribution on the finite interval [a, b]. Then the pdf $f_{i,j}^{0}$ of $D_{i,j}$ is given by

$$f_{i,j}^{0}(r) = \int_{a}^{b-r} f_{i,j}(x, x+r) dx$$

$$= (j-i) \begin{cases} n & \text{if } (b-a)^{-n} r^{j-i-1} (b-a-r)^{n-j+i}, \end{cases} (3.1)$$

for 0 < r < b- a and 1 £ <math>i < j £ n.

Proof. We use Theorem 2.2 to write $f_{i,j}(x, x + r)$ for $1 \, \pounds \, i < j \, \pounds \, n$ as

$$f_{i,j}(x_i x_i + r) = \frac{n!}{(i-1)!(j-i-1)!(n-j)!} (b-a)^{-n}(x-a)^{i-1} r^{j-i-1} (b-x-r)^{n-j}.$$

By integrating over the interval [a, b - r] with respect to x, we get

$$f_{i,j}(x,x+r)dx = \frac{n!}{(i-1)!(j-i-1)!(n-j)!}(b-a)^{-n}r^{j-i-1}, \text{ where}$$

$$I = \int_{a}^{b-r} (x-a)^{i-1}(b-x-r)^{n-j}dx$$

$$= \frac{(i-1)!(n-j)!}{(n-j+i)!}(b-a-r)^{n-j+i},$$

by using the integration by parts technique. Therefore,

$$\int_{a}^{b \cdot r} f_{i,j}(x,x+r) dx = (j-1) \int_{a}^{b} \frac{n}{j-1} dx = (j-1) \int_{a}^{b} \frac{n}{j-1} (b-a-r)^{n-j+1}, (3.2)$$

for 0 < r < b - a, i < j, as required. W

Corollary 4.1. Let $X_{1:n}, X_{2:n}, K, X_{n:n}$ be the order statistics of a random sample of size n from a uniform distribution on the interval [0,1].

Then
$$X_{j:n} - X_{i:n} \stackrel{d}{=} X_{j-i:n}$$
.

Proof. It follows from (3.2) that the pdf of $D_{i,j}$ is given by

$$f_{i,j}^{0}(r) = (j-1) \begin{cases} \frac{n}{2} & \text{if } (b-a)^{-n} r^{j-j-1} (b-a-r)^{n-j+1}, a < r < b \end{cases} (3.3)$$

It follows also from (3.1) that the pdf of the (j - i) th order statistic of a random sample of size n from a uniform distribution on the finite interval [a,b] is

$$f_{j-i}(r) = (j-i) \xi_{j-i}^{\frac{1}{2}} (b-a)^{-1} \left(\frac{r}{b-a}\right)^{j-i-1} \left(1 - \frac{r}{b-a}\right)^{n-j+1}$$

$$= (j-i) \xi_{j-i}^{\frac{1}{2}} (b-a)^{-n} r^{j-i-1} (b-a-r)^{n-j+i}, a < r < b.$$
(3.4)

It follows from (3.3) and (3.4) that the two pdfs are the same. Hence $X_{j:n} - X_{i:n} \stackrel{d}{=} X_{j-i:n}$. W

Corollary 4.2. Let $X_{1:n}, X_{2:n}, K, X_{n:n}$ be the order statistics of a random sample of size n from a uniform distribution on the finite interval [a,b]. Then the pdf $f_{i,i+k}^{0}$ of $D_{i,i+k}$ is given for $1 \pm i < n$ and $1 \pm k \pm n - i$ by

$$\int_{a}^{b-r} f_{i,i+k}(x,x+r) dx = k \int_{a}^{b} \frac{1}{x} (b-a)^{-n} r^{k-1} (b-a-r)^{n-k}, (3.5)$$

for 0 < r < b- a, $1 \pm i < n$, and 0 < k < n such that $1 < i + k \pm n$.

Proof. The proof follows directly from Theorem 4.1 by letting j = k + i. W

Proof. The pdf, $I_{i,i+k}^{0}$, of $X_{i+k,n}$ - $X_{i:n}$ is given by

On The Spacings of Exponential and Uniform order ...

$$f_{i,i+k}^{0}(x) = \int_{a}^{b-x} f_{i,i+k}(t,t+r)dt$$

$$= k \int_{a}^{\frac{1}{2}} (b-a)^{-n} x^{k-1} (b-a-x)^{n-k}, 0 < x < b-a.$$
(3.6)

The pdf of $X_{n:k}$ is given by (2.1) as

$$f_k(x) = k \frac{\partial^n}{\partial x^{\frac{1}{2}}} (b - a)^{-n} (x - a)^{k-1} (b - x)^{n-k}, a < x < b.$$

By using the transformation Y := X - a, we see that

$$f_Y(x) = k \xi_{\frac{1}{k}}^{\frac{n}{2}} (b - a)^{-n} x^{k-1} (b - a - x)^{n-k}, 0 < x < b - a. \quad (3.7)$$

Comparing (3.6) and (3.7), we see that both $X_{i+k:n}$ - $X_{i:n}$ and $X_{k:n} + a$ are identically distributed for $1 \pm i \pm n$ and $1 \pm k \pm n - 1$. W

Theorem 4.2. Let $X_{1:n}, X_{2:n}, K$, $X_{n:n}$ be the order statistics of a random sample of size n from a uniform distribution on the finite interval [a, b]. Then

Consequently, the pdf of D_i for i = 2, K, n is given by

$$f_i(r) = n(b - a)^{-n}(b - a - r)^{n-1}, 0 < r < b - a, \text{ and}$$
 (3.9)

$$f_1(r) = n(b - a)^{-n}(b - r)^{n-1}, a < r < b.$$
(3.10)

That is D_i , i=2,K, n are transformations of a beta distribution with parameters 1 and n; more precisely, the random variable $D_i := g(D_i)$, where $g(y) = \frac{X}{b-a}$ has a beta(1,n) distribution. D_1 is also a transformation of a beta(1,n) distribution, where the transformation in this case is given by $h(y) = \frac{X}{b}$.

Proof. We start with the pdf of D_i for i = 2, K, n

$$f_{i}(r) = \int_{a}^{b-r} f_{i-1,i}(x,x+r) dx, \text{ where}$$

$$f_{i-1,i}(x,x+r) = \frac{n!}{(i-2)!(n-i)!} f(x) f(x+r) F^{i-2}(x) [1 - F(x+r)]^{n-i},$$

for a < x < b, 0 < r < b - a.

Since f(x) = 1/(b - a), a < x < b

and F(x) = (x - a)/(b - a), a < x < b, we have

$$f_{i-1,i}(x,x+r) = \frac{n!}{(i-2)!(n-i)!}(b-a)^{-n}(x-a)^{i-2}(b-x-r)^{n-i},$$

for a < x < b, 0 < r < b - a.

It follows by integration by parts repeatedly that

The case when i=1 follows from our definition of $D_1 := X_{1:n}$ and that $f_{X_{1:n}}(r) = n(b-a)^{-n}(b-x)^{n-1}$. W

Corollary 4.3. Let $X_{1:n}, X_{2:n}, K, X_{n:n}$ be the order statistics of a random sample of size n from a uniform distribution on the interval [0,1]. Then

$$F_{D_i}(r) = \begin{cases} 50 & \text{if } r \pm 0 \\ -1 - (1 - r)^n & \text{if } 0 < r < 1 \\ 1 & \text{if } r = 1. \end{cases}$$
 (3.11)

Consequently, the pdf of D_i for i = 1, 2, K, n is given by

$$I_{I}(r) = n(1 - r)^{n-1}, 0 < r < 1.$$
 (3.12)

That is, D_i has a beta distribution with parameters 1 and n.

Proof. The proof follows directly from Theorem 4.2 by letting a=0 and b=1. W

REFERENCE:

- 1. Ahsanullah, M. (1995). Record Statistics. Nova Science Publishers, Commack, New York.
- 2. Arnold, B. C., Balakrishnan, N. and Nagaraja, H. N. (1992). A First Course in Order Statistics. Wiley, New York.
- 3. DasGupta, Rinott, Y., and Vidakovic, B. (1999). Stopping Times Related to Diagnostics and Outliers, Preprint.
- 4. George Casella, Berger, and Roger L. (1990). Statistical Inference. Duxbury Press.
- 5. Riffi, Mohamed I. (2002). Distributions of Spacings of Order Statistics and Their Ratios, Islamic University Journal, Preprint.