IUG Journal of Natural Studies
Peer-reviewed Journal of Islamic University-Gaza

ISSN 2409-4587

IUGNES
Vol. 25, No 2, 2017, pp 239-243

Accepted on (14-03-2017)

Gamma Spacings as Finite Gamma Mixtures

Mohamed I. Riffi[1,*]

[1]Department of Mathematics, Faculty of Science, Islamic University of Gaza, Gaza Strip, Palestine

*** Corresponding author**
e-mail address: mriffi@iugaza.edu.ps

Abstract
This paper considers the distributions of spacings between successive order statistics corresponding to a random sample from a two-parameter gamma distribution $\Gamma(\alpha,\beta)$. We prove that when the shape parameter of the underlying distribution is a positive integer, these spacings can be expressed as finite gamma mixtures. We present exact formulas for computing the distributions of the spacings. Then we present a *Mathematica* program to implement the results.

Keywords:
Gamma distribution,
Gamma mixture,
Incomplete gamma function,
Order statistics,
Spacings.

1. Introduction:

Let $X_1, X_2, \cdots, X_n$ be a random sample of size n and $X_{(1)}, X_{(2)}, \cdots, X_{(n)}$ be the corresponding order statistics, where $X_{(1)} < X_{(2)} < \cdots < X_{(n)}$. Let $D_i = X_{(i+1)} - X_{(i)}$ be the distance between $X_{(i+1)}$ and $X_{(i)}$ for $i = 1,2,3,\ldots,n-1$. By convention, we set $D_0 = X_{(1)}$.

The random variables $D_0, D_1, \ldots, D_{n-1}$ are called the spacings between successive order statistics (Pyke, 1965; Casella et al., 2001). Order statistics and their spacings play an important role in mathematical statistics and other fields of applied probability. These spacings can be used to characterize some distributions like the uniform and exponential distributions (Pyke, 1965; Ahsanullah, 1978). They are very important in statistical inference. For example, they can be used to construct confidence intervals for the corresponding population. They also are used in goodness-of-fit testing (See for example Lockhart et al, 1986; Stephens, 1986). Therefore, these spacings may be of interest in their own right. In particular, gamma spacings play an important role in reliability, communication systems, queueing applications, and other areas in the engineering discipline.

For various properties of order statistics the reader may refer to (Arnold et al, 1992; David, 1981). Spacings and their properties are deeply reviewed in (Pyke, 1965). One also can refer to (DasGupta, 2011) for more information about spacings.

The random variable X is said to follow a gamma distribution with shape and scale parameters $\alpha > 0$ and $\beta > 0$, respectively, written for short as $X \sim \Gamma(\alpha,\beta)$, if the probability density function (pdf) of X is given by

$$f_X(x) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-x/\beta}, \quad 0 < x < \infty, \tag{1}$$

where $\Gamma(\alpha) = \int_0^{\infty} t^{\alpha-1} e^{-t} dt$ is the gamma function at $\alpha > 0$.

The cumulative distribution function (cdf) of $X \sim \Gamma(\alpha,\beta)$ at $t > 0$ is given by

$$F_X(t)=\frac{\gamma(\alpha,t/\beta)}{\Gamma(\alpha)}, \quad (2)$$

where $\gamma(\alpha,t/\beta)=\int_0^{t/\beta}x^{\alpha-1}e^{-x}dx$ is the lower incomplete gamma function at $(\alpha,t/\beta)$.

Many authors considered the representation of spacings in terms of normalized exponential random variables. (Pyke, 1965) showed that for any random variable X with failure rate function $\lambda_X(t)$, the spacings can be represented as

$$D_i=\frac{1}{(n-i)\lambda_X(a_i)}Y_i, \quad (3)$$

where $Y_0,Y_1,\ldots,Y_{n-1}$ are independent and identically distributed exponential random variables and $S_i<a_i<S_{i+1}$, with

$$S_i=\sum_{j=0}^{i}\frac{1}{n-j}Y_j,\quad i=0,1,\ldots,n-1. \quad (4)$$

(Weissman, 1978) showed that as the sample size n goes to infinity, the successive spacings between the k largest observations (k fixed) asymptotically become independent and exponentially distributed for a wide class of distributions that includes the gamma distribution.

The problem we tackle in this paper is determining the distributions of the D_i,s when the underlying population of the sample is $\Gamma(\alpha,\beta)$, where $\alpha=1,2,\ldots$.

We will show that the D_i's have gamma mixtures with $k=(\alpha-1)(n-i)+1$ components. The jth component has $\Gamma(j,\beta/(n-i))$ distribution, where $j=1,2,\ldots,k$. In other words, We will show that D_i, for $i=1,2,\cdots n-1$, can be represented as a gamma mixture; that is, a weighted sum of gamma distributions with scale parameters $\beta/(n-i)$, where β is the scale parameter of the underlying population.

That is, we will show that the pdf of D_i is given by

$$f_{D_i}(r\mid\beta)=\sum_{j=1}^{k}w_j\Gamma_j(j,\beta/(n-i)), \quad (5)$$

where $w_j>0$ and $\sum_j^k w_j=1$.

2. Distributions of the D_i's:

Let $X_1,X_2,\ldots,X_n$ be a random sample of size n from a gamma distribution $\Gamma(\alpha,\beta)$ with shape and scale parameters α and β, respectively.

Recall that the joint pdf of the ith and jth order statistics $X_{(i)}$ and $X_{(j)}, 1\le i<j\le n$, is given by (DasGupta, 2011)

$$f_{i,j}(x,y)=\frac{n!}{(i-1)!(j-1-i)!(n-j)!}f(x)f(y)F^{i-1}(x)\times[F(y)-F(x)]^{j-1-i}[1-F(y)]^{n-j}, \quad (6)$$

for $-\infty<x<y<\infty$.

When $j=i+1$, (6) reduces to

$$f_{i,i+1}(x,y)=\frac{n!}{(i-1)!(n-i-1)!}f(x)f(y)F^{i-1}(x)[1-F(y)]^{n-i-1} \quad (7)$$

It follows from (7) that the pdf of D_i at $r>0$ is then the integral

$$f_{D_i}(r\mid\alpha,\beta)=\int_0^{\infty}f_{i,i+1}(x,x+r)dx \quad (8)$$

Theorem 1 *The spacing D_i between $X_{(i+1)}$ and $X_{(i)}$, for each $i=1,2,\ldots n-1$, has a finite gamma mixture distribution.*

$$f_{D_i}(r\mid\beta)=\sum_{j=1}^{k}w_j\Gamma_j(j,\beta/(n-i)), \quad (9)$$

where $k=(\alpha-1)(n-i)+1$ is the number of components, $w_j>0$ with $\sum_j^k w_j=1$, and $\Gamma_j(j,\beta/(n-i))$ is the gamma pdf of shape parameter j and scale parameter $\beta/(n-i)$, for all $j=1,2,\ldots,k$.

Proof. We use the following expansion of $\gamma(\alpha,x)$, namely,

$$\gamma(\alpha,x)=\Gamma(\alpha)\left[1-e^{-x}\sum_{i=0}^{\alpha-1}\frac{x^i}{i!}\right] \quad (10)$$

when $\alpha=1,2,\ldots$ (see Temme, 1994; Gradshteyn et al., 2000; Gautshi, 2003), expand sums by using the multinomial theorem, namely, for a positive integer k and a non-negative integer n,

$$(x_1+x_2+\cdots+x_k)^m=\sum_{b_1+b_2+\cdots+b_k=m}\binom{m}{b_1,b_2,\ldots,b_k}\prod_{j=1}^{k}x_j^{b_j}, \quad (11)$$

and then use integration by parts and mathematical induction.

Hence, for $X : \Gamma(\alpha,\beta)$

$$F_X(x) = \frac{\gamma(\alpha, x/\beta)}{\Gamma(\alpha)} = 1 - e^{-x/\beta}\sum_{i=0}^{\alpha-1}\frac{(x/\beta)^i}{i!} \quad \text{and}$$

$$1 - F_X((x+r)/\beta) = \frac{\gamma(\alpha,(x+r)/\beta)}{\Gamma(\alpha)} = e^{-(x+r)/\beta}\sum_{i=0}^{\alpha-1}\frac{((x+r)/\beta)^i}{i!}.$$

Substituting in (7), we get for $i = 1,2,\ldots n-1$

$$f_{i,i+1}(x,x+r) = \delta x^{\alpha-1}e^{-r/\beta}e^{-2x/\beta}(r+x)^{\alpha-1}\left(1-e^{-\frac{x}{\beta}}\sum_{i=0}^{\alpha-1}\frac{\left(\frac{x}{\beta}\right)^i}{i!}\right)^{i-1} \times\left(e^{-\frac{r+x}{\beta}}\sum_{i=0}^{\alpha-1}\frac{\left(\frac{r+x}{\beta}\right)^i}{i!}\right)^{n-i-1}, \quad (12)$$

where $\delta = \beta^{-2\alpha}n!/\Gamma(n-i)\Gamma(i)\Gamma(\alpha)^2$.

$$S_1 = \left(1-e^{-\frac{x}{\beta}}\sum_{k=0}^{\alpha-1}\frac{\left(\frac{x}{\beta}\right)^k}{k!}\right)^{i-1}$$
$$= \sum_{a_1+a_2+\cdots+a_{\alpha+1}=i-1}\binom{i-1}{a_1,a_2,\ldots,a_{\alpha+1}}(1)^{a_1}(-e^{-x/\beta})^{a_2}(-\frac{x}{\beta}e^{-x/\beta})^{a_3}\cdots \times\left(-\frac{x^{\alpha-1}}{(\alpha-1)!\beta^{\alpha-1}}e^{-x/\beta}\right)^{a_{\alpha+1}}$$
$$= \sum_{a_1+a_2+\cdots+a_{\alpha+1}=i-1}(-1)^{a_2+a_3+\cdots+a_{\alpha+1}}\binom{i-1}{a_1,a_2,\ldots,a_{\alpha+1}} \times\frac{e^{-(a_2+a_3\cdots+a_{\alpha+1})x/\beta}}{(2!)^{a_4}(3!)^{a_5}\cdots((\alpha-1)!)^{a_{\alpha+1}}}\frac{x^{a_3+2a_4+\cdots+(\alpha-1)a_{\alpha+1}}}{\beta^{a_3+2a_4+\cdots+(\alpha-1)a_{\alpha+1}}}. \quad (13)$$

Similarly,

$$S_2 = \left(e^{-\frac{x+r}{\beta}}\sum_{k=0}^{\alpha-1}\frac{\left(\frac{(x+r)}{\beta}\right)^k}{k!}\right)^{n-i-1}$$
$$= e^{-(n-i-1)(x+r)/\beta}\sum_{b_1+b_2+\cdots+b_\alpha=n-i-1}\binom{n-i-1}{b_1,b_2,\ldots,b_\alpha}(1)^{b_1}\left(\frac{(x+r)}{1!\beta}\right)^{b_2} \times\left(\frac{(x+r)^2}{2!\beta^2}\right)^{b_3}\cdots\left(\frac{(x+r)^{\alpha-1}}{(\alpha-1)!\beta^{\alpha-1}}\right)^{b_\alpha}$$
$$= e^{-(n-i-1)(x+r)/\beta}\sum_{b_1+b_2+\cdots+b_\alpha=n-i-1}\binom{n-i-1}{b_1,b_2,\ldots,b_\alpha} \times\frac{(x+r)^{b_2+2b_3+\cdots+(\alpha-1)b_\alpha}}{(1!)^{b_2}(2!)^{b_3}(3!)^{b_4}\cdots((\alpha-1)!)^{b_\alpha}}\frac{1}{\beta^{b_2+2b_3+\cdots+(\alpha-1)b_\alpha}}. \quad (14)$$

Similarly,

$$(x+r)^{\alpha+b_2+2b_3+\cdots+(\alpha-1)b_\alpha-1} = \sum_{h=0}^{B_2+\alpha-1}\binom{B_2+\alpha-1}{h}x^h r^{B_2+\alpha-h-1}, \quad (15)$$

where $B_2 = b_2 + 2b_3 + \cdots + (\alpha-1)b_\alpha$.
Therefore,

$$f_{i,i+1}(x,x+r) = \delta' \sum_{a_1+a_2+\cdots+a_{\alpha+1}=i-1}\ \sum_{b_1+b_2+\cdots+b_{\alpha+2}=n-i-1}\ \sum_{h=0}^{B_2+\alpha-1} e^{-r/\beta} \times\binom{i-1}{a_1,a_2,\ldots,a_{\alpha+1}}\binom{n-i-1}{b_1,b_2,\ldots,b_\alpha}\binom{B_2+\alpha-1}{h} \times\frac{e^{-A_1x/\beta}x^{A_2}}{A_3\beta^{A_2}}\times\frac{e^{-(n-i-1)(x+r)/\beta}}{B_3\beta^{B_2}}\times x^{\alpha+h-1}r^{B_2+\alpha-h-1}, \quad (16)$$

where $\delta' = (-1)^{a_2+a_3+\cdots+a_{\alpha+1}}\delta$ and

$$\begin{aligned} A_1 &= a_2 + a_3\cdots + a_{\alpha+1} + 2 \\ A_2 &= a_3 + 2a_4 + \cdots + (\alpha-1)a_{\alpha+1} \\ A_3 &= (1!)^{a_3}(2!)^{a_4}(3!)^{a_5}\cdots((\alpha-1)!)^{a_{\alpha+1}} \\ B_1 &= b_2 + b_3 + \cdots + b_\alpha \\ B_2 &= b_2 + 2b_3 + \cdots + (\alpha-1)b_\alpha \\ B_3 &= (1!)^{b_2}(2!)^{b_3}(3!)^{b_4}\cdots((\alpha-1)!)^{b_\alpha}. \end{aligned} \quad (17)$$

Therefore, the pdf of D_i at $r > 0$ is given by

$$f_{D_i}(r) = \delta' \sum_{a_1+a_2+\cdots+a_{\alpha+1}=i-1}\ \sum_{b_1+b_2+\cdots+b_{\alpha+2}=n-i-1}\ \sum_{h=0}^{B_2+\alpha-1}\frac{1}{A_3\beta^{A_2}B_3\beta^{B_2}} \times\binom{i-1}{a_1,a_2,\ldots,a_{\alpha+1}}\binom{n-i-1}{b_1,b_2,\ldots,b_\alpha}\binom{B_2+\alpha-1}{h}\times I(r), \quad (18)$$

where

$$\begin{aligned} I(r) &= \int_0^\infty r^{B_2+\alpha-h-1} e^{-r/\beta} x^{A_2+\alpha+h-1} e^{-A_1 x/\beta} e^{-(n-i-1)(x+r)/\beta} dx \\ &= \Gamma(A_2+\alpha+h) e^{-(n-i)r/\beta} r^{B_2+\alpha-h-1} (\beta/(A_1+n-i-1))^{A_2+\alpha+h}. \end{aligned} \quad (19)$$

Hence,

$$\begin{aligned} f_{D_i}(r) = {} & \delta' \sum_{a_1+a_2+\cdots+a_{\alpha+1}=i-1} \sum_{b_1+b_2+\cdots+b_{\alpha+2}=n-i-1} \sum_{h=0}^{B_2+\alpha-1} \frac{\Gamma(A_2+\alpha+h)}{A_3\beta^{A_2} B_3\beta^{B_2}} \\ & \times \binom{i-1}{a_1,a_2,\ldots,a_{\alpha+1}} \binom{n-i-1}{b_1,b_2,\ldots,b_\alpha} \binom{B_2+\alpha-1}{h} \\ & \times e^{-\frac{(n-i)r}{\beta}} r^{B_2+\alpha-h-1} (\beta/(A_1+n-i-1))^{A_2+\alpha+h}. \end{aligned} \quad (20)$$

Remark 1 2 *It is clear from (20) that* f_{D_i} *is a finite gamma mixture, where the* j *th component has shape parameter* j *and scale parameter* $\beta/(n-i)$*, for* $i = 1,2,\ldots,n-1$.

To determine the number of components, we need to look at the range of the power of r. Note first that $\max(B_2) = (\alpha-1)(n-i-1)$. This max is achieved when $b_1 = b_2 = \cdots = b_{\alpha-1} = 0$ and $b_\alpha = n-i-1$. Therefore, the power of r have a maximum of $\max(B_2)+\alpha-1$. That is the maximum of the power of r is equal to $(\alpha-1)(n-i-1)+(\alpha-1) = (\alpha-1)(n-i)$. Therefore, the power of r ranges from 0 to $(\alpha-1)(n-i)$ and the number of components described by (20) is $k = (\alpha-1)(n-i)+1$.

Examples

Here, we use a simple *Mathematica* program to implement the result.

Example 3 *Let us consider the case when* $\alpha = 2, n = 3, i = 1$*, and* $\beta > 0$.

Now $a_1 = a_2 = a_3 = i-1 = 0$ implies $A_1 = 2, A_2 = 0, A_3 = 1$. Also, $b_1 + b_2 = n-i-1 = 1$ implies $b_1 = 0, b_2 = 1$ or $b_1 = 1, b_2 = 0$. This implies that $B_2 = 0 \text{or} 1$ and $B_3 = 1$. Note that, depending on B_2, h ranges from 0 to $B_2+\alpha-1=1$. Note also that $\delta' = 3!/\beta^4$. Considering all possible values of B_2 and h, we get the following terms

$$\frac{3!}{\beta^4}\left(\frac{\beta}{3}\right)^2 re^{-2r/\beta}, \frac{3!}{\beta^4} 2\left(\frac{\beta}{3}\right)^3 e^{-2r/\beta},$$
$$\frac{3!}{\beta^4}\frac{1}{\beta}\left(\frac{\beta}{3}\right)^2 r^2 e^{-2r/\beta}, \frac{3!}{\beta^4}\frac{4}{\beta}\left(\frac{\beta}{3}\right)^3 re^{-2r/\beta},$$
$$\frac{3!}{\beta^4}\frac{3!}{\beta}\left(\frac{\beta}{3}\right)^4 e^{-2r/\beta}.$$

Summing the above terms, we get the pdf of D_1

$$f_{D_1}(r) = \frac{8e^{-\frac{2r}{\beta}}}{9\beta} + \frac{14re^{-\frac{2r}{\beta}}}{9\beta^2} + \frac{2r^2e^{-\frac{2r}{\beta}}}{3\beta^3}. \quad (21)$$

The components proportions are $w_1 = 8/18, w_2 = 7/18$, and $w_3 = 3/18$. Therefore,

$$f_{D_1}(r) = \frac{4}{9}\Gamma(1,\beta/2) + \frac{7}{18}\Gamma(2,\beta/2) + \frac{1}{6}\Gamma(3,\beta/2). \quad (22)$$

Remark 4 *Note in Example 3 that*

$$\int_0^\infty \left(\frac{8e^{-\frac{2r}{\beta}}}{9\beta} + \frac{14re^{-\frac{2r}{\beta}}}{9\beta^2} + \frac{2r^2e^{-\frac{2r}{\beta}}}{3\beta^3} \right) dr = 1. \quad (23)$$

Note also that the components proportions can be obtained by integrating each term with respect to r; i.e.,

$$\begin{aligned} w_1 &= \int_0^\infty \frac{8e^{-\frac{2r}{\beta}}}{9\beta} dr = \frac{4}{9}, \\ w_2 &= \int_0^\infty \frac{14re^{-\frac{2r}{\beta}}}{9\beta^2} dr = \frac{7}{18}, \\ w_3 &= \int_0^\infty \frac{2r^2e^{-\frac{2r}{\beta}}}{3\beta^3} dr = \frac{1}{6}. \end{aligned} \quad (24)$$

Example 5 *Suppose that* $\alpha = 2, n = 4$*, and* $i = 1$.

$$\begin{aligned} f_{D_1}(r) &= \frac{3e^{-\frac{3r}{\beta}}(\beta+r)\left(13b^2+20\beta r+8r^2\right)}{32\beta^4} \\ &= \frac{3r^3e^{-\frac{3r}{\beta}}}{4\beta^4} + \frac{21r^2e^{-\frac{3r}{\beta}}}{8\beta^3} + \frac{99re^{-\frac{3r}{\beta}}}{32\beta^2} + \frac{39e^{-\frac{3r}{\beta}}}{32\beta}. \end{aligned} \quad (25)$$

Example 6 *Suppose that* $\alpha = 3, n = 5, i = 2$*, and* $\beta > 0$*. In this case we have*

$$f_{D_2}(r) = \frac{543r^6 e^{-\frac{3r}{\beta}}}{80000\beta^7} + \frac{12687r^5 e^{-\frac{3r}{\beta}}}{160000\beta^6} + \frac{135447r^4 e^{-\frac{3r}{\beta}}}{320000\beta^5} + \frac{2078481r^3 e^{-\frac{3r}{\beta}}}{1600000\beta^4} + \frac{153589893r^2 e^{-\frac{3r}{\beta}}}{64000000\beta^3} + \frac{8051628663re^{-\frac{3r}{\beta}}}{3200000000\beta^2} + \frac{3737720079e^{-\frac{3r}{\beta}}}{3200000000\beta} \quad (26)$$

The proportions are $w_1 = 0.389346, w_2 = 0.27957, w_3 = 0.177766,$ $w_4 = 0.096226, w_5 = 0.0418046, w_6 = 0.0130525$, and $w_7 = 0.00223457$.

References:

Ahsanullah, M. (1978). On characterizations of exponential distribution by spacings. *Ann. Inst. Stat. Math.*, *30*(1), 163-166.

Arnold, B. C., Balakrishnan, N., & Nagaraja, H. N. (1992). *A First Course in Order Statistics*. New York: Wiley.

Casella, G., & Berger, R. L. (2001). *Statistical Inference*. (2nd ed.). Duxbury Press.

David, H. A. (1981). *Order Statistics*. (2nd ed.). New York: Wiley.

DasGupta, A. (2011). *Probability for statistics and machine learning. Fundamentals and advanced topics*. New York: Springer.

Gautshi, W., Harris F. E., & Temme, N. M. (2003). Expansions of the exponential integral in incomplete gamma functions. *Appl. Math. Lett.*, *16*, 1095-1099.

Gradshteyn, I. S., & Ryzhik, I. M. (2000). *Table of Integrals, Series, and Products*. (6th ed.). San Diego: Academic Press.

Lockhart, R. A., O'Reilly, F., & Stephens, M. A. (1986). Tests for the Weibull and Extreme-value distributions based on normalized spacings. *Naval Research Logistics Quarterly*, *33*, 413-421.

Pyke, R. (1965). Spacings, *Journal of the Royal Statistical Society*, *Series B*, *27*(3), 395-449.

Stephens, M. A. (1986). Tests for uniformity arising from a series of events. *Festschrift for Prof. H. Solomon's 65th birthday*, 197-219.

Temme, N. M. (1994). Computational aspects of incomplete gamma functions with large complex parameters. In R. V. M. Zahar (Ed.), *Proceedings of the conference on Approximation and computation*: *A fetschrift in honor of Walter Gautschi* (pp. 551-562). Cambridge, MA, USA: Birkhauser Boston Inc.

Weissman, I. (1978). Estimation of parameters and large quantiles based on the k largest observations. *JASA*, *73*, 812-815.

مسافات جاما كمزيج جاما المحدود

تأخذ هذه الورقة في الاعتبار توزيعات المباعدات بين الإحصاءات المرتبة المتعاقبة المقابلة لعينة عشوائية من توزيع جاما ذي المعلمتين $G(\alpha, \beta)$. نبرهن، عندما تكون معلمة الشكل للتوزيع الأساسي عددا صحيحا، على أنه يمكن التعبير هذه المباعدات كمزيج محدود من توزيعات جاما. نقدم معادلات مضبوطة لحساب توزيعات هذه المباعدات. ونقدم برنامجا باستخدام برنامج Mathematica للحصول على النتائج.

كلمات مفتاحية:
توزيع جاما،
مزيج جاما،
دالة جاما غير الكاملة،
الإحصاءات المرتبة،
المباعدات.