On Approximating the Distribution of Quadratic Forms in Gamma Random Variables and Exponential Order Statistics

A. Akbar Mohsenipour

Department of Statistical and Actuarial Sciences, The University of Western Ontario
London, Ontario, Canada
amohsen@alumni.uwo.ca

Serge B. Provost

Department of Statistical and Actuarial Sciences, The University of Western Ontario
London, Ontario, Canada
provost@stats.uwo.ca

This paper proposes a moment-based approximation to the distribution of quadratic forms in gamma random variables. Quadratic forms in order statistics from an exponential population are considered as well. Actually, several test statistics can be expressed in terms of the latter. The density approximants are expressible as the product of a gamma type distributed base density function and a polynomial adjustment. Several illustrative examples are provided.

Keywords: Quadratic forms, Exponential random variables, Order statistics, Generalized gamma distribution, Moments, Density approximation.

2000 Mathematics Subject Classification: 62E17, 62M10

1. Introduction

Suppose that $Y_1 < \cdots < Y_n$ are order statistics from an exponential distribution with mean θ. Several tests of fit with respect to the exponential distribution are based on certain quadratic forms in the Y_i's divided by an estimate of the scaling parameter. These tests based on correlations involving the Y_i's are considered in [4, 5, 7]. Some distributional limit theorems such as those that are discussed in [1] in connection with a certain empirical quantile process, involve quadratic forms in exponential random variables. Moreover, three test statistics that can be expressed as quadratic forms in exponential random variables, are described in [2].

A representation of quadratic forms in gamma random variables as well as a derivation of their moments are provided in Section 2. A closed form representation of the moments of quadratic forms in order statistics from an exponential population is determined in Section 3. Five numerical examples illustrate the results.

2. Quadratic forms in gamma random variables

Let $\mathbf{X} = (X_1, \ldots, X_n)'$ denote a random vector whose components are independently distributed gamma random variables with parameters α and β whose density function is given by

$$\psi(x) = \frac{x^{\alpha-1} e^{-x/\beta}}{\Gamma(\alpha)\beta^\alpha} \, \mathscr{I}_{\mathscr{R}^+}(x) \, , \, \alpha > 0, \, \beta > 0, \tag{2.1}$$

where $\mathscr{I}_{\mathscr{R}^+}(x)$ denotes the indicator function on the set of positive real numbers. Consider the quadratic form

$$Q(\mathbf{X}) = Q(X_1, \ldots, X_n) = \mathbf{X}'A\mathbf{X} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}X_iX_j,$$

where $A = (a_{ij})$ is an $n \times n$ symmetric matrix and $\mathbf{X}'$ denotes the transpose of $\mathbf{X}$. We note that if A is not symmetric, it suffices to replace it by $(A + A')/2$. Letting $\prod_{i,j}^{n}$ denote the double product $\prod_{i=1}^{n}\prod_{j=1}^{n}$, it follows from the multinomial expansion that

$$Q(\mathbf{X})^m = \left(\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}X_iX_j\right)^m = \sum_{(m)} m!\left(\prod_{i,j}^{n} \frac{a_{ij}^{m_{ij}}}{m_{ij}!}\right)\prod_{\ell=1}^{n} X_\ell^{\delta_\ell}, \quad m = 0, 1, 2, \ldots, \tag{2.2}$$

where $\sum_{(m)}$ denotes the sum over all the partitions of m into n^2 terms such that $m_{11} + m_{12} + \cdots + m_{nn} = m$ with $0 \le m_{ij} \le m$, the m_{ij}'s being nonnegative integers, and $\delta_\ell = \sum_{j=0}^{n}(m_{\ell j} + m_{j\ell})$. We note that the following identity is useful for computing sums over partitions:

$$\sum_{(p)} \varphi(p_1, \ldots, p_r) = \sum_{p_1=0}^{p}\sum_{p_2=0}^{p-p_1}\cdots\sum_{p_{r-1}=0}^{p-p_1-\cdots-p_{r-2}} \varphi\left(p_1, p_2, \ldots, p_{r-1}, p - \sum_{i=1}^{r-1} p_i\right),$$

where the p_i's are such that $p_1 + \cdots + p_r = p$, $p_i = 0, 1, \ldots, p$; $i = 1, 2, \ldots, r$.

Alternatively, symbolic computational packages such as Mathematica can readily generate the required partitions as well as expressing $Q(\mathbf{X})^m$ as a sum of products of powers of X_ℓ's. Then, assuming that the X_ℓ's are independently distributed with respective density functions $f_{X_\ell}(x_\ell)$, one can determine the mth moment of $Q(\mathbf{X})$ as follows:

$$\begin{aligned}
E(Q(\mathbf{X})^m) &= \int_0^\infty \int_0^\infty \cdots \int_0^\infty Q(\mathbf{x})^m f_{X_1,\ldots,X_n}(x_1,\ldots,x_n)\,\mathrm{d}x_1 \ldots \mathrm{d}x_n \\
&= \int_0^\infty \int_0^\infty \cdots \int_0^\infty \left(\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j\right)^m \prod_{\ell=1}^{n} f_{X_\ell}(x_\ell)\,\mathrm{d}x_1 \ldots \mathrm{d}x_n \\
&= \sum_{(m)} m!\left[\prod_{i,j}^{n} \frac{a_{ij}^{m_{ij}}}{m_{ij}!}\right] \int_0^\infty \int_0^\infty \cdots \int_0^\infty \left(\prod_{\ell=1}^{n} x_\ell^{\delta_\ell} f_{X_\ell}(x_\ell)\right)\mathrm{d}x_1 \ldots \mathrm{d}x_n \\
&= \sum_{(m)} m!\left[\prod_{i,j}^{n} \frac{a_{ij}^{m_{ij}}}{m_{ij}!}\right] \prod_{\ell=1}^{n}\left(\int_0^\infty x_\ell^{\delta_\ell} f_{X_\ell}(x_\ell)\,\mathrm{d}x_\ell\right).
\end{aligned}$$

Thus, when the X_ℓ's constitute a simple random sample from a gamma distribution with parameters α and β, one has

$$
\begin{aligned}
E(Q(\mathbf{X})^m) &= m! \sum_{(m)} \left[\prod_{i,j}^n \frac{a_{ij}^{m_{ij}}}{m_{ij}!} \right] \prod_{\ell=1}^n \left(\int_0^\infty \frac{x_\ell^{\delta_\ell + \alpha - 1} e^{-x_\ell/\beta}}{\Gamma(\alpha)\beta^\alpha} \right) dx_1 \ldots dx_n \\
&= m! \, \Gamma(\alpha)^{-n} \sum_{(m)} \left[\prod_{i,j}^n \frac{a_{ij}^{m_{ij}}}{m_{ij}!} \right] \beta^{\Sigma_{\ell=1}^n \delta_\ell} \prod_{\ell=1}^n \Gamma(\alpha + \delta_\ell) \\
&\equiv \mu_m \, .
\end{aligned}
\tag{2.3}
$$

Accordingly, when the components of the random vector $\mathbf{X}$ are exponentially distributed with parameter β and density function

$$
f(x) = \frac{1}{\beta} e^{-x/\beta} \, \mathscr{I}_{\mathscr{R}^+}(x) \, , \ \beta > 0,
\tag{2.4}
$$

the mth moment of $Q(\mathbf{X})$ is

$$
E(Q(\mathbf{X})^m) = m! \sum_{(m)} \left[\prod_{i,j}^n \frac{a_{ij}^{m_{ij}}}{m_{ij}!} \right] \beta^{\Sigma_{\ell=1}^n \delta_\ell} \prod_{\ell=1}^n \Gamma(1 + \delta_\ell).
\tag{2.5}
$$

Given the moments of such quadratic forms, approximations to their distribution can be obtained by making use of the methodologies discussed in [8]. A polynomially-adjusted density approximation, which is described in the Appendix A, is seen to provide accurate percentiles.

Example 2.1. Consider the quadratic form $Q_1(\mathbf{X}) = \mathbf{X}'A\mathbf{X}$ where $\mathbf{X} = (X_1, \ldots, X_5)$, and

$$
A = \begin{pmatrix} 4 & 3 & 2 & 1 & 0 \\ 3 & 0 & 2 & 0 & 1 \\ 2 & 2 & 0 & 3 & 2 \\ 1 & 0 & 3 & 1 & 0 \\ 0 & 1 & 2 & 0 & 6 \end{pmatrix},
$$

the X_i's being independently and exponentially distributed with parameter $\beta = 3$.

Since the exponential distribution has a semi-infinite support and all the elements of A are non-negative, a generalized gamma distribution can be used as base density to determine an approximate distribution for $Q_1(\mathbf{X})$. The proposed methodology comprises the following steps:

- The moments of $Q_1(\mathbf{X})$ are determined from Equation (2.3) wherein in this case $n = 5$ and $\beta = 3$.
- Consider the following generalized gamma density function as a base density:

$$
\psi(z) = \frac{\gamma}{\beta^{\alpha\gamma}\Gamma(\alpha)} z^{\alpha\gamma - 1} e^{-(z/\beta)^\gamma} \, \mathscr{I}_{(0,\infty)}(z), \quad \alpha > 0, \ \beta > 0, \ \gamma > 0.
\tag{2.6}
$$

- The parameters α, β and γ are determined by solving simultaneously the following nonlinear equations

$$\mu_j = m_j \quad \text{for } j = 1, 2, 3,$$

where

$$m_j = \frac{\beta^j \Gamma(\alpha + j/\gamma)}{\Gamma(\alpha)}, \ j = 0, 1, \ldots$$

are the moments associated with a generalized gamma density function and μ_j can be determined from Equation (2.3).

- A polynomial adjustment of degree d can be made as explained in the Appendix A. The resulting density approximation is

$$f_d(z) = \varphi(z) \sum_{j=0}^{d} \xi_j z^j;$$

in this case, we set $d = 7$.

Table 1. Approximate cdf of $Q_1(\mathbf{X})$ corresponding to certain percentiles obtained by simulation (Sim%).

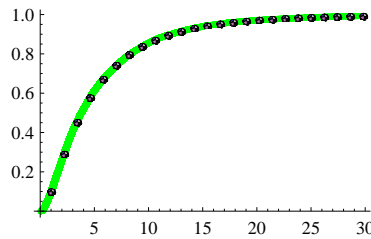| CDF	Sim%	G. Gam. Poly
0.01	0.2562	0.008006
0.05	0.6255	0.045688
0.50	3.8274	0.501751
0.95	16.261	0.949537
0.99	27.589	0.989890

Fig. 1. Simulated cdf of $Q_1(\mathbf{X})$ and 7th degree polynomially adjusted generalized gamma cdf approximation (dots).

Certain values of the resulting approximate distribution function of $Q_1(\mathbf{X})$ are displayed in Table 1 where G. Gam. Poly denotes the cdf obtained from the polynomially adjusted generalized gamma density function. The percentiles were determined by simulation on the basis of 1,000,000 replications. The plots shown in Figure 1 confirm that the polynomially adjusted generalized gamma distribution provides a very accurate approximation to the distribution of $Q_1(\mathbf{X})$.

Remark 2.1. Referring to Equation (2.1), when $\alpha_i = v_i/2$, $i = 1, 2, \ldots, n$ and $\beta = 2$, the ith component of the random vector $\mathbf{X} = (X_1, \ldots, X_n)'$ has a chi-square distribution with v_i degrees of freedom and the representation of the mth moment of $Q(\mathbf{X})$ given in Equation (2.3) applies.

Remark 2.2. When the matrix A in the quadratic form $Q(\mathbf{X}) = \mathbf{X}'A\mathbf{X}$ contains negative elements, one can utilize the density function of the difference of two gamma random variables as base density in order to determine an approximation to the distribution of $Q(\mathbf{X})$. Such a density function can be determined as follows.

Let Y_1 and Y_2 be independently distributed random variables with parameters α_1, β_1 and α_2, β_2, respectively. By making use of binomial expansion of $(Y_1 - Y_2)^h$, $h = 1, 2, 3, 4$, one can determine the first four raw moments of $Y_1 - Y_2$, which are

$$E(Y_1 - Y_2) = \alpha_1 \beta_1 - \alpha_2 \beta_2$$
$$E(Y_1 - Y_2)^2 = \alpha_1 (1 + \alpha_1) \beta_1^2 - 2\alpha_1 \alpha_2 \beta_1 \beta_2 + \alpha_2 (1 + \alpha_2) \beta_2^2$$
$$E(Y_1 - Y_2)^3 = \alpha_1 (1 + \alpha_1) (2 + \alpha_1) \beta_1^3 - \alpha_2 \beta_2 (3\alpha_1 (1 + \alpha_1) \beta_1^2 - 3\alpha_1 (1 + \alpha_2) \beta_1 \beta_2$$
$$+ (1 + \alpha_2) (2 + \alpha_2) \beta_2^2)$$
$$E(Y_1 - Y_2)^4 = \alpha_1 (1 + \alpha_1) (2 + \alpha_1) (3 + \alpha_1) \beta_1^4 + \alpha_2 (1 + \alpha_2) (2 + \alpha_2) (3 + \alpha_2) \beta_2^4$$
$$- 2\alpha_1 \alpha_2 \beta_1 \beta_2 (2 (1 + \alpha_1) (2 + \alpha_1) \beta_1^2 - 3 (1 + \alpha_1) (1 + \alpha_2) \beta_1 \beta_2$$
$$+ 2 (1 + \alpha_2) (2 + \alpha_2) \beta_2^2) \tag{2.7}$$

Now, on equating these moments to those obtained from (2.3), one can solve the resulting system of equations for $\alpha_1, \beta_1, \alpha_2$ and β_2, which can be achieved by making use of symbolic computational packages such as Maple and *Mathematica*.

It follows from the results derived in [9] that the density function of $Q = Y_1 - Y_2$ where Y_1 and Y_2 are independently distributed gamma random variables with parameters α_1, β_1 and α_2, β_2, respectively, can be expressed as

$$h_n(q) \, \mathscr{I}_{(-\infty,0)}(q) + h_p(q) \, \mathscr{I}_{[0,\infty)}(q) \tag{2.8}$$

with

$$h_n(q) = \int_0^\infty f_1(y) f_2(y - q) \, dy$$
$$= \frac{\beta_1^{-\alpha_1} \beta_2^{-\alpha_2}}{\Gamma(\alpha_2)} e^{q/\beta_2} \vartheta^{-(\alpha_1+\alpha_2)/2} e^{-\vartheta q/2} (-q)^{(\alpha_1+\alpha_2-2)/2}$$
$$\times W_{(\alpha_2-\alpha_1)/2,(1-\alpha_1-\alpha_2)/2}(-\vartheta q)$$

and

$$h_p(q) = \int_q^\infty f_1(y) f_2(y - q) \, dy$$
$$= \frac{\beta_1^{-\alpha_1} \beta_2^{-\alpha_2}}{\Gamma(\alpha_1)} e^{q/\beta_2} \vartheta^{-(\alpha_1+\alpha_2)/2} e^{-\vartheta q/2} q^{(\alpha_1+\alpha_2-2)/2}$$
$$\times W_{(\alpha_1-\alpha_2)/2,(1-\alpha_1-\alpha_2)/2}(\vartheta q)$$

where $\vartheta = (\beta_1 + \beta_2)/(\beta_1 \beta_2)$, $\vartheta q \neq 0$, $\alpha_1 > 0$, $\alpha_2 > 0$, $\beta_1 > 0$, $\beta_2 > 0$, $(1 - \alpha_1 - \alpha_2)$ is not a negative integer or zero and $W(\cdot)$ denotes the Whittaker function, which on making use of some identities

given in Sections 9.220 and 9.210.1 of [3], is seen to have the following representation:

$$W_{l,\mu}(z) = \frac{\Gamma(-2\mu)}{\Gamma(\frac{1}{2} - \mu - l)} z^{\mu + \frac{1}{2}} e^{-z/2} {}_1F_1\left(\mu - l + \frac{1}{2}; 2\mu + 1; z\right)$$

$$+ \frac{\Gamma(2\mu)}{\Gamma(\frac{1}{2} + \mu - l)} z^{-\mu + \frac{1}{2}} e^{-z/2} {}_1F_1\left(-\mu - l + \frac{1}{2}; -2\mu + 1; z\right)$$

where ${}_1F_1(a, b, z) = \sum_{k=0}^{\infty} \frac{\Gamma(a+k)\Gamma(b) z^k}{\Gamma(a)\Gamma(b+k) k!}$.

Example 2.2. Consider the quadratic form $Q_2(\mathbf{X}) = \mathbf{X}'A\mathbf{X}$ where $\mathbf{X} = (X_1, X_2, X_3)'$ is a vector of independently distributed chi-square random variables having 4, 3 and 5 degrees of freedom, respectively, and

$$A = \begin{pmatrix} 4 & 1 & -2 \\ 1 & 0 & 2 \\ -2 & 2 & -4 \end{pmatrix}.$$

In light of Remark 2.2, one can determine an approximation to the distribution function of $Q_2(\mathbf{X})$ by following the steps described in Example 2.1, the base density being given by (2.8) in this instance. In Figure 2, this approximation is superimposed on the simulated distribution function which was determined from 1,000,000 replications.
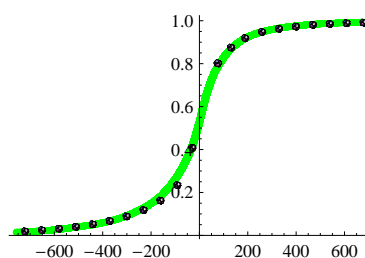

Fig. 2. Simulated cdf of $Q_2(\mathbf{X})$ and cdf approximation obtained from the difference of two gamma random variables (dots).

3. Quadratic forms in order statistics from an exponential population

In this section, we derive the moments of the quadratic form $Q(\mathbf{X}) = \mathbf{X}'A\mathbf{X}$ where $\mathbf{X}$ is a vector of order statistics $X_1 = Y_{r_1:n}$, $X_2 = Y_{r_1+r_2:n}$ and $X_k = Y_{r_1+\cdots+r_k:n}$ obtained from a simple random sample of n observations generated from a standard exponential distribution (with density $g(y) = e^{-y} \mathscr{I}_{\mathscr{R}^+}(x)$).

In this case, the joint density of $X_1, \ldots, X_k$ is

$$f(x_1, \ldots, x_k) = \frac{\Gamma(n+1)}{\sum_{j=1}^{k+1} \Gamma(r_j)} \left(\prod_{j=1}^{k} e^{-x_j}\right) \left(1 - e^{-x_1}\right)^{r_1 - 1}$$

$$\times \prod_{i=2}^{k} \left(e^{-x_{i-1}} - e^{-x_i}\right)^{r_i - 1} \left(e^{-x_k}\right)^{r_{k+1} - 1} \tag{3.1}$$

whenever $0 < x_1 < \ldots < x_k < \infty$ with $r_{k+1} = n + 1 - \sum_{j=1}^{k} r_j$, and 0, otherwise.

Consider the transformation $z_1 = 1 - e^{x_1}$ and $z_j = e^{-x_{j-1}} - e^{-x_j}$ for $j = 2, \ldots, k$. The inverse transformation is then

$$x_j = -\ln(1 - z_1 - \cdots - z_j)$$

for $j = 1, \ldots, k$, and its Jacobian is

$$\prod_{j=1}^{k} (1 - z_1 - \cdots - z_j)^{-1} = \prod_{j=1}^{k} e^{x_j} > 0. \tag{3.2}$$

Noting that $e^{-x_k} = 1 - z_1 - \cdots - z_k$, the joint density of $Z_1, \ldots, Z_k$ is seen to be

$$h(z_1, \ldots, z_k) = \frac{\Gamma(n+1)}{\prod_{j=1}^{k+1} \Gamma(r_j)} \left(\prod_{j=1}^{k} z_j^{r_j-1} \right) (1 - z_1 - \cdots - z_k)^{r_{k+1}-1} \tag{3.3}$$

whenever $0 < z_j < 1$, $i = 1, \ldots, k$, and $\sum_{i=1}^{k} z_i \leq 1$, and 0 otherwise. Thus the random vector $\mathbf{Z} = (Z_1, \ldots, Z_k)'$ has a type-one Dirichlet distribution with parameters $r_1, r_2, \ldots, r_{k+1}$.

In view of (3.2), the joint moment-generating function of $\mathbf{U} = (-X_1, \ldots, -X_k)'$ evaluated at the point $\mathbf{t} = (t_1, \ldots, t_k)$ can be expressed as

$$M_{\mathbf{U}}(\mathbf{t}) = E\left(e^{t_1 \ln(1-Z_1) + \cdots + t_k \ln(1-Z_1 - \cdots - Z_k)} \right) = E\left((1 - Z_1)^{t_1} \cdots (1 - Z_1 - \cdots - Z_k)^{t_k} \right)$$

$$= \frac{\Gamma(n+1)}{\prod_{j=1}^{k+1} \Gamma(r_j)} \int \cdots \int (1 - z_1)^{t_1} (1 - z_1 - z_2)^{t_2} \cdots (1 - z_1 - \cdots - z_k)^{t_k}$$

$$\times z_1^{r_1-1} z_2^{r_2-1} \cdots z_k^{r_{k+1}-1} dz_k \cdots dz_2 \, dz_1 \tag{3.4}$$

where the domain of integration is $0 < z_i < 1$, $i = 1, \ldots, k$, with $\sum_{i=1}^{k} z_i \leq 1$. Integrating out z_k and making the change of variables $w = z_k / (1 - z_1 - \cdots - z_{k-1})$ yields

$$\int_0^{1-z_1-\cdots-z_{k-1}} z_k^{r_k-1} (1 - z_1 - \cdots - z_k)^{r_{k+1}+r_k-1} dz_k$$

$$= (1 - z_1 - \cdots - z_{k-1})^{r_k + r_{k+1} + t_k - 1} \int_0^1 w^{r_k-1} (1 - w)^{r_{k+1}+t_k-1} dw$$

$$= (1 - z_1 - \cdots - z_{k-1})^{r_k + r_{k+1} + t_k - 1} \frac{\Gamma(r_k)\Gamma(r_{k+1}+t_k)}{\Gamma(r_k) + r_{k+1} + t_k}.$$

Then, integrating the terms involving z_{k-1} from 0 to $1 - z_1 - \cdots - z_{k-2}$, one has

$$(1 - z_1 - \cdots - z_{k-2})^{r_{k+1}+r_k+r_{k-1}+t_k+t_{k-1}-1} \frac{\Gamma(r_{k-1})\Gamma(r_{k+1}+r_k+t_k+t_{k-1})}{\Gamma(r_{k+1}+r_k+r_{k-1}+t_k+t_{k-1})}$$

and integrating successively the terms involving $z_{k-2}, \ldots, z_2$ and z_1, one obtains

$$M_{\mathbf{U}}(\mathbf{t}) = \frac{\Gamma(n+1)}{\Gamma(r_{k+1})} \prod_{j=1}^{k} \frac{\Gamma(r_{k+1} + \cdots + r_{j+1} + t_k + \cdots + t_j)}{\Gamma(r_{k+1} + \cdots + r_j + t_k + \cdots + t_j)}. \tag{3.5}$$

Accordingly,

$$E(X_1^{\delta_1} X_2^{\delta_2} \cdots X_k^{\delta_k}) = (-1)^{\delta_1+\delta_2+\cdots+\delta_k} \frac{\partial^{\delta_1+\delta_2+\cdots+\delta_k} M_{\mathbf{U}}(\mathbf{t})}{\partial^{\delta_1} t_1 \partial^{\delta_2} t_2 \ldots \partial^{\delta_k} t_k}\bigg|_{\mathbf{t}=\mathbf{0}}, \tag{3.6}$$

and in light of Equations (2.2), (3.5) and (3.6), the mth moment of the quadratic form $Q(\mathbf{X})$ can be evaluated as follows:

$$E(Q(\mathbf{X})^m) = \sum_{(m)} m! \left[\prod_{i,j}^{k} \frac{a_{ij}^{m_{ij}}}{m_{ij}!} \right] (-1)^{\delta_1+\delta_2+\cdots+\delta_k} \frac{\partial^{\delta_1+\delta_2+\cdots+\delta_k} M_{\mathbf{U}}(\mathbf{t})}{\partial^{\delta_1} t_1 \partial^{\delta_2} t_2 \ldots \partial^{\delta_k} t_k}\bigg|_{\mathbf{t}=\mathbf{0}} \equiv \mu_m^*. \tag{3.7}$$

For computational purposes, it is simpler to make use of the joint cumulant generating function of $\mathbf{U} = (-X_1, \ldots, -X_k)'$, which is

$$C_{\mathbf{U}}^*(\mathbf{t}) = \ln[\Gamma(n+1)] - \ln[\Gamma(r_{k+1})] + \sum_{i=1}^{k} \{\ln(r_{k+1} + \cdots + r_{j+1} + t_k$$
$$+ \cdots + t_j) - \ln(r_{k+1} + \cdots + r_j + t_k + \cdots + t_j)\}. \tag{3.8}$$

The joint cumulants of $-X_1, \ldots, -X_k$ of orders $\xi_1, \ldots, \xi_k$ are then given by

$$\kappa_{\mathbf{U}}^*(\xi_1, \ldots, \xi_k) = \frac{\partial^{\xi_1+\cdots+\xi_k}}{\partial^{\xi_1} t_1 \cdots \partial^{\xi_k} t_k} C_{\mathbf{U}}^*(\mathbf{t})\,|_{\mathbf{t}=\mathbf{0}} = p! \sum_{\ell=0}^{v-1} (-1/(x+\ell))^{p+1} \tag{3.9}$$

where $v = \sum_{j=1}^{\lambda} r_j$, λ is the position of the first non null component in $\xi = (\xi_1, \ldots, \xi_k)'$, $x = n+1-v$ and $p = (\sum_{j=1}^{k} \xi_j) - 1$. On making use of a recursive relationship given in [11], which can also be deduced for instance from Theorem 3.2b.2 in [6], one can determine the mth moment of $Q(\mathbf{X})$ as follows:

$$\mu_m^* = \sum_{i=0}^{m-1} \frac{(m-1)!}{(m-1-i)!\,i!}\, \kappa_U^*(m-i)\,\mu_i^*, \tag{3.10}$$

where $\kappa_U^*(m-i)$ is as specified by (3.9).

Example 3.1. Let the order statistics $X_1 \le \cdots \le X_5$ be obtained from a random sample of size 5 from an exponential distribution with parameter 1. Consider the quadratic form $Q_3(\mathbf{X}) = \mathbf{X}'A\mathbf{X}$ where $\mathbf{X} = (X_1, \ldots, X_5)'$ and

$$A = \begin{pmatrix} 1 & 1 & 1 & 3 & 0 \\ 1 & 0 & 2 & 0 & 1 \\ 1 & 2 & 0 & 4 & 2 \\ 3 & 0 & 4 & 1 & 0 \\ 0 & 1 & 2 & 0 & 2 \end{pmatrix}. \tag{3.11}$$

In this example, we approximate the distribution of $Q_3(\mathbf{X})$ whose support is non-negative by making use of a generalized gamma distribution. The moments of $Q_3(\mathbf{X})$ can be determined from Equation (3.7) or (3.10). The steps described in Example 2.1 were followed. The results included in Table 2 indicate that the generalized gamma density function provides an accurate approximation to

Table 2. Approximate cdf of $Q_3(\mathbf{X})$ obtained from a generalized gamma (G. Gamma) density function corresponding to certain percentiles obtained by simulation (Sim%).

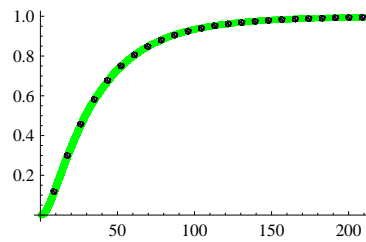| CDF	Sim%	G. Gamma
0.01	0.1337	0.009585
0.05	0.3183	0.049363
0.50	1.8023	0.500105
0.95	6.9692	0.949828
0.99	11.218	0.989847
0.999	18.358	0.998990

Fig. 3. Simulated cdf of $Q_3(\mathbf{X})$ and 7$^{\text{th}}$ degree polynomially adjusted generalized gamma cdf approximation (dots).

the distribution of $Q_3(\mathbf{X})$. The generalized gamma was adjusted with a seventh degree polynomial in order to determine the cdf plotted in Figure 3.

Example 3.2. Referring to Example 3.1, suppose that A is the matrix

$$\begin{pmatrix} -5 & 1 & 1 & 3 & 0 \\ 1 & 0 & -2 & 0 & -4 \\ 1 & -2 & 0 & 4 & 2 \\ 3 & 0 & 4 & 1 & 0 \\ 0 & -4 & 2 & 0 & -2 \end{pmatrix}.$$

In this case, a base density such as the one given in (2.8) is appropriate. Then, on following the steps described in Example 2.1, one can determine an approximation for $Q_4(\mathbf{X})$. Figure 4 suggests that the approximated cdf is quite accurate.
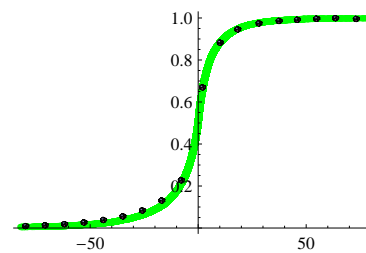

Fig. 4. Simulated cdf of $Q_4(\mathbf{X})$ and cdf approximation (dots).

Remark 3.1. More generally, when the order statistics $X_1 \leq \cdots \leq X_n$ are generated from an *Exponential*(β) random variable whose density function is as specified by Equation (2.4), one can represent $Q(\mathbf{X}) = \mathbf{X}'A\mathbf{X}$ as $Q(\mathbf{Y}) = \beta^2 (Y_1, \ldots, Y_n) A (Y_1, \ldots, Y_n)'$ where the Y_i's are order statistics from an *Exponential*(1) random variable. Once an approximate density is obtained for $(Y_1, \ldots, Y_n) A (Y_1, \ldots, Y_n)'$, a simple change of variables will yield the density function of $Q(\mathbf{X})$.

Example 3.3. Suppose that the order statistics $X_1 \leq \cdots \leq X_5$ are generated from a random sample of size 5 from an exponential distribution with parameter 4 and let Q_5 denote the quadratic form $\mathbf{X}'A\mathbf{X}$ with A as given in (3.11). Then, proceeding as in Example 3.1 and reexpressing the quadratic form in terms of *Exponential*(1) random variables as explained in Remark 3.1, one can approximate the density function of $Q_5(\mathbf{X})$ by making use of a polynomially-adjusted generalized gamma distribution. The results presented in Table 3 and Figure 5 indicate that approximate distribution is in close agreement with the simulated distribution which was determined on the basis of 1,000,000 replications.

Table 3. Approximate cdf of $Q_5(\mathbf{X})$ corresponding to certain percentiles obtained by simulation (Sim%).

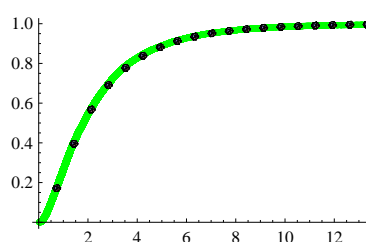| *CDF*	Sim%	G. Gamma
0.01	0.1179	0.009580
0.05	0.2797	0.049251
0.50	1.5857	0.500469
0.95	6.1811	0.500469
0.99	10.054	0.990157
0.999	16.615	0.999056

Fig. 5. Simulated cdf of $Q_5(\mathbf{X})$ and polynomially-adjusted generalized gamma cdf approximation (dots).

Acknowledgments

References

[1] E. del Barrio, E. Giné and F. Utzet, Asymptotics for L_2 functionals of the empirical quantile process, with applications to tests of fit based on weighted Wasserstein distances, *Bernoulli*, **11** (2005) 131–189.

[2] S. G. Donald and H. J. Paarsch, Superconsistent estimation and inference in structural econometric models using extreme order statistics, *Journal of Econometrics*, **109** (2002) 305–340.

[3] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series, Products, Corrected and Enlarged Edition*, (Academic Press, New York 1980).

[4] H. O. Hartley and R. C. Pfaffenberger, Quadratic forms in order statistics used as goodness-of-fit criteria, *Biometrika*, **59** (1972) 605–611.

[5] R. A. Lockhart, The asymptotic distribution of the correlation coefficient in testing fit to the exponential distribution, *Canad. J. Statist*, **13** (1985) 253–256.

[6] A. M. Mathai, and S. B. Provost, *Quadratic Forms in Random Variables, Theory and Applications*, (Marcel Dekker Inc., New York, 1992).

[7] C. G. McLaren and R. A. Lockhart, On the asymptotic efficiency of correlation tests of fit, *Canad. J. Statist*, **15** (1987) 159–167.

[8] A. A. Mohsenipour and S. B. Provost, On approximating the distributions of ratios and differences of noncentral quadratic forms in normal vectors, *Journal of Statistical Research*, **44** (2010) 315–334.

[9] A. A. Mohsenipour and S. B. Provost, Approximating the distributions of singular quadratic expressions and their ratios, *Journal of the Iranian Statistical Society*, **11** (2012) 147–171.

[10] S. B. Provost, Moment-based density approximants, *The Mathematica Journal*, **9** (2005) 727–756.

[11] P. J. Smith, A recursive formulation of the old problem of obtaining moments from cumulants and vice versa, *The American Statistician*, **49** (1995) 217–219.

Appendix A. Polynomially adjusted density functions

As explained in [10], an initial density approximation can be adjusted by means of a polynomial whose coefficients are such that the first d moments of the approximation coincide with those of a target distribution.

For instance, on the basis of the first d moments of a positive definite quadratic form $Q(\mathbf{X})$, a density approximation of the following form can be assumed for $Q(\mathbf{X})$:

$$f_d(x) = \varphi(x) \sum_{j=0}^{d} \xi_j x^j$$

where $\varphi(x)$ is an initial density approximant referred to as base density function, which could be for instance a gamma density.

The polynomial coefficients, ξ_j, are determined by equating the hth moment of $Q(\mathbf{X})$ denoted by μ_h to the hth moment of the approximate distribution specified by $f_d(x)$ for $h = 0, 1, \ldots, d$, that is,

$$\mu_h = \int_p^q x^h \varphi(x) \sum_{j=0}^{d} \xi_j x^j \, \mathrm{d}x = \sum_{j=0}^{d} \xi_j \int_p^q x^{h+j} \varphi(x) \, \mathrm{d}x$$

$$= \sum_{j=0}^{d} \xi_j \, m_{h+j}, \qquad h = 0, 1, \ldots, d,$$

where m_{h+j} is the $(h+j)^{\text{th}}$ moment determined from $\varphi(x)$. This leads to a linear system of $(d+1)$ equations in $(d+1)$ unknowns whose solution is

$$
\begin{pmatrix} \xi_0 \\ \xi_1 \\ \vdots \\ \xi_d \end{pmatrix} = \begin{pmatrix} m_0 & m_1 & \cdots & m_d \\ m_1 & m_2 & \cdots & m_{d+1} \\ \cdots & \cdots & \cdots & \cdots \\ m_d & m_{d+1} & \cdots & m_{2d} \end{pmatrix}^{-1} \begin{pmatrix} \mu_0 \\ \mu_1 \\ \vdots \\ \mu_d \end{pmatrix}.
$$

The resulting representation of the density function of $Q(\mathbf{X})$ is referred to as a dth degree polynomially adjusted density approximant. As long as higher moments are available and the calculations can be carried out with sufficient precision, more accurate approximations can be obtained by making use of additional moments.