

Chapter 2

The Univariate Gaussian and Related Distributions



2.1. Introduction

It is assumed that the reader has had adequate exposure to basic concepts in Probability, Statistics, Calculus and Linear Algebra. This chapter will serve as a review of the basic ideas about the univariate Gaussian, or normal, distribution as well as related distributions. We will begin with a discussion of the univariate Gaussian density. We will adopt the following notation: real scalar mathematical or random variables will be denoted by lower-case letters such as x, y, z , whereas vector or matrix-variate mathematical or random variables will be denoted by capital letters such as X, Y, Z, \dots . Statisticians usually employ the double notation X and x where it is claimed that x is a realization of X . Since x can vary, it is a variable in the mathematical sense. Treating mathematical and random variables the same way will simplify the notation and possibly reduce the confusion. Complex variables will be written with a tilde such as $\tilde{x}, \tilde{y}, \tilde{X}, \tilde{Y}$, etc. Constant scalars and matrices will be written without a tilde unless for stressing that the constant matrix is in the complex domain. In such a case, a tilde will be also be utilized for the constant. Constant matrices will be denoted by A, B, C, \dots .

The numbering will first indicate the chapter and then the section. For example, Eq. (2.1.9) will be the ninth equation in Sect. 2.1 of this chapter. Local numbering for sub-sections will be indicated as (i), (ii), and so on.

Let x_1 be a real univariate Gaussian, or normal, random variable whose parameters are μ_1 and σ_1^2 ; this will be written as $x_1 \sim N_1(\mu_1, \sigma_1^2)$, the associated density being given by

$$f(x_1) = \frac{1}{\sigma_1 \sqrt{2\pi}} e^{-\frac{1}{2\sigma_1^2}(x_1 - \mu_1)^2}, \quad -\infty < x_1 < \infty, \quad -\infty < \mu_1 < \infty, \quad \sigma_1 > 0.$$

In this instance, the subscript 1 in $N_1(\cdot)$ refers to the univariate case. Incidentally, a density is a real-valued scalar function of x such that $f(x) \geq 0$ for all x and $\int_x f(x) dx = 1$. The moment generating function (mgf) of this Gaussian random variable x_1 , with t_1 as

its parameter, is given by the following expected value, where $E[\cdot]$ denotes the expected value of $[\cdot]$:

$$E[e^{t_1 x_1}] = \int_{-\infty}^{\infty} e^{t_1 x_1} f(x_1) dx_1 = e^{t_1 \mu_1 + \frac{1}{2} t_1^2 \sigma_1^2}. \quad (2.1.1)$$

2.1a. The Complex Scalar Gaussian Variable

Let $\tilde{x} = x_1 + ix_2$, $i = \sqrt{-1}$, x_1, x_2 real scalar variables. Let $E[x_1] = \mu_1$, $E[x_2] = \mu_2$, $\text{Var}(x_1) = \sigma_1^2$, $\text{Var}(x_2) = \sigma_2^2$, $\text{Cov}(x_1, x_2) = \sigma_{12}$. By definition, the variance of the complex random variable \tilde{x} is defined as

$$\text{Var}(\tilde{x}) = E[\tilde{x} - E(\tilde{x})][\tilde{x} - E(\tilde{x})]^*$$

where $*$ indicates a conjugate transpose in general; in this case, it simply means the conjugate since \tilde{x} is a scalar. Since $\tilde{x} - E(\tilde{x}) = x_1 + ix_2 - \mu_1 - i\mu_2 = (x_1 - \mu_1) + i(x_2 - \mu_2)$ and $[\tilde{x} - E(\tilde{x})]^* = (x_1 - \mu_1) - i(x_2 - \mu_2)$,

$$\begin{aligned} \text{Var}(\tilde{x}) &= E[\tilde{x} - E(\tilde{x})][\tilde{x} - E(\tilde{x})]^* \\ &= E[(x_1 - \mu_1) + i(x_2 - \mu_2)][(x_1 - \mu_1) - i(x_2 - \mu_2)] = E[(x_1 - \mu_1)^2 + (x_2 - \mu_2)^2] \\ &= \sigma_1^2 + \sigma_2^2 \equiv \sigma^2. \end{aligned} \quad (i)$$

Observe that $\text{Cov}(x_1, x_2)$ does not appear in the scalar case. However, the covariance will be present in the vector/matrix case as will be explained in the coming chapters. The complex Gaussian density is given by

$$f(\tilde{x}) = \frac{1}{\pi \sigma^2} e^{-\frac{1}{\sigma^2} (\tilde{x} - \tilde{\mu})^* (\tilde{x} - \tilde{\mu})} \quad (ii)$$

for $\tilde{x} = x_1 + ix_2$, $\tilde{\mu} = \mu_1 + i\mu_2$, $-\infty < x_j < \infty$, $-\infty < \mu_j < \infty$, $\sigma^2 > 0$, $j = 1, 2$. We will write this as $\tilde{x} \sim \tilde{N}_1(\tilde{\mu}, \sigma^2)$. It can be shown that the two parameters appearing in the density in (ii) are the mean value of \tilde{x} and the variance of \tilde{x} . We now establish that the density in (ii) is equivalent to a real bivariate Gaussian density with $\sigma_1^2 = \frac{1}{2}\sigma^2$, $\sigma_2^2 = \frac{1}{2}\sigma^2$ and zero correlation. In the real bivariate normal density, the exponent is the following, with Σ as given below:

$$\begin{aligned} &-\frac{1}{2}[(x_1 - \mu_1), (x_2 - \mu_2)] \Sigma^{-1} \begin{bmatrix} x_1 - \mu_1 \\ x_2 - \mu_2 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \frac{1}{2}\sigma^2 & 0 \\ 0 & \frac{1}{2}\sigma^2 \end{bmatrix} \\ &= -\frac{1}{\sigma^2} \{(x_1 - \mu_1)^2 + (x_2 - \mu_2)^2\} = -\frac{1}{\sigma^2} (\tilde{x} - \tilde{\mu})^* (\tilde{x} - \tilde{\mu}). \end{aligned}$$

This exponent agrees with that appearing in the complex case. Now, consider the constant part in the real bivariate case:

$$(2\pi)|\Sigma|^{\frac{1}{2}} = (2\pi) \left| \begin{array}{cc} \frac{1}{2}\sigma^2 & 0 \\ 0 & \frac{1}{2}\sigma^2 \end{array} \right|^{\frac{1}{2}} = \pi\sigma^2,$$

which also coincides with that of the complex Gaussian. Hence, a complex scalar Gaussian is equivalent to a real bivariate Gaussian case whose parameters are as described above.

Let us consider the mgf of the complex Gaussian scalar case. Let $\tilde{t} = t_1 + it_2$, $i = \sqrt{-1}$, with t_1 and t_2 being real parameters, so that $\tilde{t}^* = \bar{\tilde{t}} = t_1 - it_2$ is the conjugate of \tilde{t} . Then $\tilde{t}^*\tilde{x} = t_1x_1 + t_2x_2 + i(t_1x_2 - t_2x_1)$. Note that $t_1x_1 + t_2x_2$ contains the necessary number of parameters (that is, 2) and hence, to be consistent with the definition of the mgf in a real bivariate case, the imaginary part should not be taken into account; thus, we should define the mgf as $E[e^{\Re(\tilde{t}^*\tilde{x})}]$, where $\Re(\cdot)$ denotes the real part of (\cdot) . Accordingly, in the complex case, the mgf is obtained as follows:

$$\begin{aligned} M_{\tilde{x}}(\tilde{t}) &= E[e^{\Re(\tilde{t}^*\tilde{x})}] \\ &= \frac{1}{\pi\sigma^2} \int_{\tilde{x}} e^{\Re(\tilde{t}^*\tilde{x}) - \frac{1}{\sigma^2}(\tilde{x}-\tilde{\mu})^*(\tilde{x}-\tilde{\mu})} d\tilde{x} \\ &= \frac{e^{\Re(\tilde{t}^*\tilde{\mu})}}{\pi\sigma^2} \int_{\tilde{x}} e^{\Re(\tilde{t}^*(\tilde{x}-\tilde{\mu})) - \frac{1}{\sigma^2}(\tilde{x}-\tilde{\mu})^*(\tilde{x}-\tilde{\mu})} d\tilde{x}. \end{aligned}$$

Let us simplify the exponent:

$$\begin{aligned} \Re(\tilde{t}^*(\tilde{x}-\tilde{\mu})) - \frac{1}{\sigma^2}(\tilde{x}-\tilde{\mu})^*(\tilde{x}-\tilde{\mu}) \\ &= -\left\{ \frac{1}{\sigma^2}(x_1-\mu_1)^2 + \frac{1}{\sigma^2}(x_2-\mu_2)^2 - t_1(x_1-\mu_1) - t_2(x_2-\mu_2) \right\} \\ &= \frac{\sigma^2}{4}(t_1^2+t_2^2) - \left\{ (y_1 - \frac{\sigma}{2}t_1)^2 + (y_2 - \frac{\sigma}{2}t_2)^2 \right\} \end{aligned}$$

where $y_1 = \frac{x_1-\mu_1}{\sigma}$, $y_2 = \frac{x_2-\mu_2}{\sigma}$, $dy_j = \frac{1}{\sigma}dx_j$, $j = 1, 2$. But

$$\frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-(y_j - \frac{\sigma}{2}t_j)^2} dy_j = 1, \quad j = 1, 2.$$

Hence,

$$M_{\tilde{x}}(\tilde{t}) = e^{\Re(\tilde{t}^*\tilde{\mu}) + \frac{\sigma^2}{4}\tilde{t}^*\tilde{t}} = e^{t_1\mu_1 + t_2\mu_2 + \frac{\sigma^2}{4}(t_1^2 + t_2^2)}, \quad (2.1a.1)$$

which is the mgf of the equivalent real bivariate Gaussian distribution.

Note 2.1.1. A statistical density is invariably a real-valued scalar function of the variables involved, be they scalar, vector or matrix variables, real or complex.

2.1.1. Linear functions of Gaussian variables in the real domain

If x_1, \dots, x_k are statistically independently distributed real scalar Gaussian variables with parameters (μ_j, σ_j^2) , $j = 1, \dots, k$ and if a_1, \dots, a_k are real scalar constants then the mgf of a linear function $u = a_1x_1 + \dots + a_kx_k$ is given by

$$\begin{aligned} M_u(t) &= E[e^{tu}] = E[e^{ta_1x_1 + \dots + ta_kx_k}] = M_{x_1}(a_1t) \cdots M_{x_k}(a_kt), \text{ as } M_{ax}(t) = M_x(at), \\ &= e^{(ta_1\mu_1 + \dots + ta_k\mu_k) + \frac{1}{2}t^2(a_1^2\sigma_1^2 + \dots + a_k^2\sigma_k^2)} \\ &= e^{t(\sum_{j=1}^k a_j\mu_j) + \frac{1}{2}t^2(\sum_{j=1}^k a_j^2\sigma_j^2)}, \end{aligned}$$

which is the mgf of a real normal random variable whose parameters are $(\sum_{j=1}^k a_j\mu_j, \sum_{j=1}^k a_j^2\sigma_j^2)$. Hence, the following result:

Theorem 2.1.1. *Let the real scalar random variable x_j have a real univariate normal (Gaussian) distribution, that is, $x_j \sim N_1(\mu_j, \sigma_j^2)$, $j = 1, \dots, k$ and let x_1, \dots, x_k be statistically independently distributed. Then, any linear function $u = a_1x_1 + \dots + a_kx_k$, where a_1, \dots, a_k are real constants, has a real normal distribution with mean value $\sum_{j=1}^k a_j\mu_j$ and variance $\sum_{j=1}^k a_j^2\sigma_j^2$, that is, $u \sim N_1(\sum_{j=1}^k a_j\mu_j, \sum_{j=1}^k a_j^2\sigma_j^2)$.*

Vector/matrix notation enables one to express this result in a more convenient form. Let

$$L = \begin{bmatrix} a_1 \\ \vdots \\ a_k \end{bmatrix}, \quad \mu = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_k \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_k^2 \end{bmatrix}.$$

Then denoting the transposes by primes, $u = L'X = X'L$, $E(u) = L'\mu = \mu'L$, and

$$\begin{aligned} \text{Var}(u) &= E[(u - E(u))(u - E(u))'] = L'E[(X - E(X))(X - E(X))']L \\ &= L'\text{Cov}(X)L = L'\Sigma L \end{aligned}$$

where, in this case, Σ is the diagonal matrix $\text{diag}(\sigma_1^2, \dots, \sigma_k^2)$. If x_1, \dots, x_k is a simple random sample from x_1 , that is, from the normal population specified by the density of x_1 or, equivalently, if x_1, \dots, x_k are iid (independently and identically distributed) random

variables having as a common distribution that of x_1 , then $E(u) = \mu_1 L'J = \mu_1 J'L$ and $\text{Var}(u) = \sigma_1^2 L'L$ where u is the linear function defined in Theorem 2.1.1 and $J' = (1, 1, \dots, 1)$ is a vector of unities.

Example 2.1.1. Let $x_1 \sim N_1(-1, 1)$ and $x_2 \sim N_1(2, 2)$ be independently distributed real normal variables. Determine the density of the linear function $u = 5x_1 - 2x_2 + 7$.

Solution 2.1.1. Since u is a linear function of independently distributed real scalar normal variables, it is real scalar normal whose parameters $E(u)$ and $\text{Var}(u)$ are

$$\begin{aligned} E(u) &= 5E(x_1) - 2E(x_2) + 7 = 5(-1) - 2(2) + 7 = -2 \\ \text{Var}(u) &= 25\text{Var}(x_1) + 4\text{Var}(x_2) + 0 = 25(1) + 4(2) = 33, \end{aligned}$$

the covariance being zero since x_1 and x_2 are independently distributed. Thus, $u \sim N_1(-2, 33)$.

2.1a.1. Linear functions in the complex domain

We can also look into the distribution of linear functions of independently distributed complex Gaussian variables. Let a be a constant and \tilde{x} a complex random variable, where a may be real or complex. Then, from the definition of the variance in the complex domain, one has

$$\begin{aligned} \text{Var}(a\tilde{x}) &= E[(a\tilde{x} - E(a\tilde{x}))(a\tilde{x} - E(a\tilde{x}))^*] = aE[(\tilde{x} - E(\tilde{x}))(\tilde{x} - E(\tilde{x}))^*]a^* \\ &= a\text{Var}(\tilde{x})a^* = aa^*\text{Var}(\tilde{x}) = |a|^2\text{Var}(\tilde{x}) = |a|^2\sigma^2 \end{aligned}$$

when the variance of \tilde{x} is σ^2 , where $|a|$ denotes the absolute value of a . As well, $E[a\tilde{x}] = aE[\tilde{x}] = a\tilde{\mu}$. Then, a companion to Theorem 2.1.1 is obtained.

Theorem 2.1a.1. Let $\tilde{x}_1, \dots, \tilde{x}_k$ be independently distributed scalar complex Gaussian variables, $\tilde{x}_j \sim \tilde{N}_1(\tilde{\mu}_j, \sigma_j^2)$, $j = 1, \dots, k$. Let a_1, \dots, a_k be real or complex constants and $\tilde{u} = a_1\tilde{x}_1 + \dots + a_k\tilde{x}_k$ be a linear function. Then, \tilde{u} has a univariate complex Gaussian distribution given by $\tilde{u} \sim \tilde{N}_1(\sum_{j=1}^k a_j\tilde{\mu}_j, \sum_{j=1}^k |a_j|^2\text{Var}(\tilde{x}_j))$.

Example 2.1a.1. Let $\tilde{x}_1, \tilde{x}_2, \tilde{x}_3$ be independently distributed complex Gaussian univariate random variables with expected values $\tilde{\mu}_1 = -1 + 2i$, $\tilde{\mu}_2 = i$, $\tilde{\mu}_3 = -1 - i$ respectively. Let $\tilde{x}_j = x_{1j} + ix_{2j}$, $j = 1, 2, 3$. Let $[\text{Var}(x_{1j}), \text{Var}(x_{2j})] = [(1, 1), (1, 2), (2, 3)]$, respectively. Let $a_1 = 1 + i$, $a_2 = 2 - 3i$, $a_3 = 2 + i$, $a_4 = 3 + 2i$. Determine the density of the linear function $\tilde{u} = a_1\tilde{x}_1 + a_2\tilde{x}_2 + a_3\tilde{x}_3 + a_4$.

Solution 2.1a.1.

$$\begin{aligned}
E(\tilde{u}) &= a_1 E(\tilde{x}_1) + a_2 E(\tilde{x}_2) + a_3 E(\tilde{x}_3) + a_4 \\
&= (1+i)(-1+2i) + (2-3i)(i) + (2+i)(-1-i) + (3+2i) \\
&= (-3+i) + (3+2i) + (-1-3i) + (3+2i) = 2+2i; \\
\text{Var}(\tilde{u}) &= |a_1|^2 \text{Var}(\tilde{x}_1) + |a_2|^2 \text{Var}(\tilde{x}_2) + |a_3|^2 \text{Var}(\tilde{x}_3)
\end{aligned}$$

and the covariances are equal to zero since the variables are independently distributed. Note that $\tilde{x}_1 = x_{11} + ix_{21}$ and hence, for example,

$$\begin{aligned}
\text{Var}(\tilde{x}_1) &= E[(\tilde{x}_1 - E(\tilde{x}_1))(\tilde{x}_1 - E(\tilde{x}_1))^*] \\
&= E\{[(x_{11} - E(x_{11})) + i(x_{21} - E(x_{21}))][(x_{11} - E(x_{11})) - i(x_{21} - E(x_{21}))]\} \\
&= E[(x_{11} - E(x_{11}))^2] - (i)^2 E[(x_{21} - E(x_{21}))^2] \\
&= \text{Var}(x_{11}) + \text{Var}(x_{21}) = 1 + 1 = 2.
\end{aligned}$$

Similarly, $\text{Var}(\tilde{x}_2) = 1 + 2 = 3$, $\text{Var}(\tilde{x}_3) = 2 + 3 = 5$. Moreover, $|a_1|^2 = (1)^2 + (1)^2 = 2$, $|a_2|^2 = (2)^2 + (3)^2 = 13$, $|a_3|^2 = (2)^2 + (1)^2 = 5$. Accordingly, $\text{Var}(\tilde{u}) = 2(2) + (13)(3) + (5)(5) = 68$. Thus, $\tilde{u} \sim \tilde{N}_1(2+2i, 68)$. Note that the constant a_4 only affects the mean value. Had a_4 been absent from \tilde{u} , its mean value would have been real and equal to -1 .

2.1.2. The chisquare distribution in the real domain

Suppose that x_1 follows a real standard normal distribution, that is, $x_1 \sim N_1(0, 1)$, whose mean value is zero and variance, 1. What is then the density of x_1^2 , the square of a real standard normal variable? Let the distribution function or cumulative distribution function of x_1 be $F_{x_1}(t) = \text{Pr}\{x_1 \leq t\}$ and that of $y_1 = x_1^2$ be $F_{y_1}(t) = \text{Pr}\{y_1 \leq t\}$. Note that since $y_1 > 0$, t must be positive. Then,

$$\begin{aligned}
F_{y_1}(t) &= \text{Pr}\{y_1 \leq t\} = \text{Pr}\{x_1^2 \leq t\} = \text{Pr}\{|x_1| \leq \sqrt{t}\} = \text{Pr}\{-\sqrt{t} \leq x_1 \leq \sqrt{t}\} \\
&= \text{Pr}\{x_1 \leq \sqrt{t}\} - \text{Pr}\{x_1 \leq -\sqrt{t}\} = F_{x_1}(\sqrt{t}) - F_{x_1}(-\sqrt{t}). \tag{i}
\end{aligned}$$

Denoting the density of y_1 by $g(y_1)$, this density at $y_1 = t$ is available by differentiating the distribution function $F_{y_1}(t)$ with respect to t . As for the density of x_1 , which is the standard normal density, it can be obtained by differentiating $F_{x_1}(\sqrt{t})$ with respect to \sqrt{t} . Thus, differentiating (i) throughout with respect to t , we have

$$\begin{aligned}
g(t)|_{t=y_1} &= \left[\frac{d}{dt} F_{x_1}(\sqrt{t}) - \frac{d}{dt} F_{x_1}(-\sqrt{t}) \right] \Big|_{t=y_1} \\
&= \left[\frac{d}{d\sqrt{t}} F_{x_1}(\sqrt{t}) \frac{d\sqrt{t}}{dt} - \frac{d}{d\sqrt{t}} F_{x_1}(-\sqrt{t}) \frac{d(-\sqrt{t})}{dt} \right] \Big|_{t=y_1} \\
&= \frac{1}{2} t^{\frac{1}{2}-1} \frac{1}{\sqrt{(2\pi)}} e^{-\frac{1}{2}t} \Big|_{t=y_1} + \frac{1}{2} t^{\frac{1}{2}-1} \frac{1}{\sqrt{(2\pi)}} e^{-\frac{1}{2}t} \Big|_{t=y_1} \\
&= \frac{1}{2^{\frac{1}{2}} \Gamma(1/2)} y_1^{\frac{1}{2}-1} e^{-\frac{1}{2}y_1}, \quad 0 \leq y_1 < \infty, \quad \text{with } \Gamma(1/2) = \sqrt{\pi}. \quad (ii)
\end{aligned}$$

Accordingly, the density of $y_1 = x_1^2$ or the square of a real standard normal variable, is a two-parameter real gamma with $\alpha = \frac{1}{2}$ and $\beta = 2$ or a real chisquare with one degree of freedom. A two-parameter real gamma density with the parameters (α, β) is given by

$$f_1(y_1) = \frac{1}{\beta^\alpha \Gamma(\alpha)} y_1^{\alpha-1} e^{-\frac{y_1}{\beta}}, \quad 0 \leq y_1 < \infty, \quad \alpha > 0, \quad \beta > 0, \quad (2.1.2)$$

and $f_1(y_1) = 0$ elsewhere. When $\alpha = \frac{n}{2}$ and $\beta = 2$, we have a *real chisquare density with n degrees of freedom*. Hence, the following result:

Theorem 2.1.2. *The square of a real scalar standard normal random variable is a real chisquare variable with one degree of freedom. A real chisquare with n degrees of freedom has the density given in (2.1.2) with $\alpha = \frac{n}{2}$ and $\beta = 2$.*

A real scalar chisquare random variable with m degrees of freedom is denoted as χ_m^2 . From (2.1.2), by computing the mgf we can see that the mgf of a real scalar gamma random variable y is $M_y(t) = (1 - \beta t)^{-\alpha}$ for $1 - \beta t > 0$. Hence, a real chisquare with m degrees of freedom has the mgf $M_{\chi_m^2}(t) = (1 - 2t)^{-\frac{m}{2}}$ for $1 - 2t > 0$. The condition $1 - \beta t > 0$ is required for the convergence of the integral when evaluating the mgf of a real gamma random variable. If $y_j \sim \chi_{m_j}^2$, $j = 1, \dots, k$ and if y_1, \dots, y_k are independently distributed, then the sum $y = y_1 + \dots + y_k \sim \chi_{m_1 + \dots + m_k}^2$, a real chisquare with $m_1 + \dots + m_k$ degrees of freedom, with mgf $M_y(t) = (1 - 2t)^{-\frac{1}{2}(m_1 + \dots + m_k)}$ for $1 - 2t > 0$.

Example 2.1.2. Let $x_1 \sim N_1(-1, 4)$, $x_2 \sim N_1(2, 2)$ be independently distributed. Let $u = x_1^2 + 2x_2^2 + 2x_1 - 8x_2 + 5$. Compute the density of u .

Solution 2.1.2.

$$\begin{aligned}
u &= x_1^2 + 2x_2^2 + 2x_1 - 8x_2 + 5 = (x_1 + 1)^2 + 2(x_2 - 2)^2 - 4 \\
&= 4 \left[\frac{(x_1 + 1)^2}{4} + \frac{(x_2 - 2)^2}{2} \right] - 4.
\end{aligned}$$

Since $x_1 \sim N_1(-1, 4)$ and $x_2 \sim N_1(2, 2)$ are independently distributed, so are $\frac{(x_1+1)^2}{4} \sim \chi_1^2$ and $\frac{(x_2-2)^2}{2} \sim \chi_1^2$, and hence the sum is a real χ_2^2 random variable. Then, $u = 4y - 4$ with $y = \chi_2^2$. But the density of y , denoted by $f_y(y)$, is

$$f_y(y) = \frac{1}{2}e^{-\frac{y}{2}}, \quad 0 \leq y < \infty,$$

and $f_y(y) = 0$ elsewhere. Then, $z = 4y$ has the density

$$f_z(z) = \frac{1}{8}e^{-\frac{z}{8}}, \quad 0 \leq z < \infty,$$

and $f_z(z) = 0$ elsewhere. However, since $u = z - 4$, its density is

$$f_u(u) = \frac{1}{8}e^{-\frac{(u+4)}{8}}, \quad -4 \leq u < \infty,$$

and zero elsewhere.

2.1a.2. The chisquare distribution in the complex domain

Let us consider the distribution of $\tilde{z}_1 \tilde{z}_1^*$ of a scalar standard complex normal variable \tilde{z}_1 . The density of \tilde{z}_1 is given by

$$f_{\tilde{z}_1}(\tilde{z}_1) = \frac{1}{\pi}e^{-\tilde{z}_1^* \tilde{z}_1}, \quad \tilde{z}_1 = z_{11} + iz_{12}, \quad -\infty < z_{1j} < \infty, \quad j = 1, 2.$$

Let $\tilde{u}_1 = \tilde{z}_1^* \tilde{z}_1$. Note that $\tilde{z}_1^* \tilde{z}_1$ is real and hence we may associate a real parameter t to the mgf. Note that $\tilde{z}_1 \tilde{z}_1^*$ in the scalar complex case corresponds to z^2 in the real scalar case where $z \sim N_1(0, 1)$. Then, the mgf of \tilde{u}_1 is given by

$$M_{\tilde{u}_1}(t) = E[e^{\Re(\tilde{t}\tilde{u}_1)}] = \frac{1}{\pi} \int_{\tilde{z}_1} e^{-\Re[(1-t)\tilde{z}_1^* \tilde{z}_1]} d\tilde{z}_1.$$

However, $\tilde{z}_1^* \tilde{z}_1 = z_{11}^2 + z_{12}^2$ as $\tilde{z}_1 = z_{11} + iz_{12}$, $i = \sqrt{-1}$, where z_{11} and z_{12} are real. Thus, the above integral gives $(1-t)^{-\frac{1}{2}}(1-t)^{-\frac{1}{2}} = (1-t)^{-1}$ for $1-t > 0$, which is the mgf of a real scalar gamma variable with parameters $\alpha = 1$ and $\beta = 1$. Let $\tilde{z}_j \sim \tilde{N}_1(\tilde{\mu}_j, \sigma^2)$, $j = 1, \dots, k$, be scalar complex normal random variables that are independently distributed. Letting

$$\tilde{u} = \sum_{j=1}^k \left(\frac{\tilde{z}_j - \tilde{\mu}_j}{\sigma_j} \right)^* \left(\frac{\tilde{z}_j - \tilde{\mu}_j}{\sigma_j} \right) \sim \text{real scalar gamma with parameters } \alpha = k, \beta = 1,$$

whose density is

$$f_{\tilde{u}}(u) = \frac{1}{\Gamma(k)} u^{k-1} e^{-u}, \quad 0 \leq u < \infty, \quad k = 1, 2, \dots, \quad (2.1a.2)$$

\tilde{u} is referred to as a scalar chisquare in the complex domain having k degrees of freedom, which is denoted $\tilde{u} \sim \tilde{\chi}_k^2$. Hence, the following result:

Theorem 2.1a.2. Let $\tilde{z}_j \sim \tilde{N}_1(\tilde{\mu}_j, \sigma_j^2)$, $j = 1, \dots, k$, be independently distributed and $\tilde{u} = \sum_{j=1}^k (\frac{\tilde{z}_j - \tilde{\mu}_j}{\sigma_j})^* (\frac{\tilde{z}_j - \tilde{\mu}_j}{\sigma_j})$. Then \tilde{u} is called a scalar chisquare having k degrees of freedom in the complex domain whose density as given in (2.1a.2) is that of a real scalar gamma random variable with parameters $\alpha = k$ and $\beta = 1$.

Example 2.1a.2. Let $\tilde{x}_1 \sim \tilde{N}_1(i, 2)$, $\tilde{x}_2 \sim \tilde{N}_1(1 - i, 1)$ be independently distributed complex Gaussian univariate random variables. Let $\tilde{u} = \tilde{x}_1^* \tilde{x}_1 + 2\tilde{x}_2^* \tilde{x}_2 - 2\tilde{x}_2^* - 2\tilde{x}_2 + i(\tilde{x}_1 + 2\tilde{x}_2^*) - i(\tilde{x}_1^* + 2\tilde{x}_2) + 5$. Evaluate the density of \tilde{u} .

Solution 2.1a.2. Let us simplify \tilde{u} , keeping in mind the parameters in the densities of \tilde{x}_1 and \tilde{x}_2 . Since terms of the type $\tilde{x}_1^* \tilde{x}_1$ and $\tilde{x}_2^* \tilde{x}_2$ are present in \tilde{u} , we may simplify into factors involving \tilde{x}_j^* and \tilde{x}_j for $j = 1, 2$. From the density of \tilde{x}_1 we have

$$\frac{(\tilde{x}_1 - i)^* (\tilde{x}_1 - i)}{2} \sim \tilde{\chi}_1^2$$

where

$$(\tilde{x}_1 - i)^* (\tilde{x}_1 - i) = (\tilde{x}_1^* + i)(\tilde{x}_1 - i) = \tilde{x}_1^* \tilde{x}_1 + i\tilde{x}_1 - i\tilde{x}_1^* + 1. \quad (i)$$

After removing the elements in (i) from \tilde{u} , the remainder is

$$\begin{aligned} & 2\tilde{x}_2^* \tilde{x}_2 - 2\tilde{x}_2^* - 2\tilde{x}_2 + 2i\tilde{x}_2^* - 2i\tilde{x}_2 + 4 \\ &= 2[(\tilde{x}_2 - 1)^* (\tilde{x}_2 - 1) - i\tilde{x}_2 + i\tilde{x}_2^* + 1] \\ &= 2[(\tilde{x}_2 - 1 + i)^* (\tilde{x}_2 - 1 + i)]. \end{aligned}$$

Accordingly,

$$\begin{aligned} \tilde{u} &= 2 \left[\frac{(\tilde{x}_1 - i)^* (\tilde{x}_1 - i)}{2} + (\tilde{x}_2 - 1 + i)^* (\tilde{x}_2 - 1 + i) \right] \\ &= 2[\tilde{\chi}_1^2 + \tilde{\chi}_1^2] = 2\tilde{\chi}_2^2 \end{aligned}$$

where $\tilde{\chi}_2^2$ is a scalar chisquare of degree 2 in the complex domain or, equivalently, a real scalar gamma with parameters $(\alpha = 2, \beta = 1)$. Letting $y = \tilde{\chi}_2^2$, the density of y , denoted by $f_y(y)$, is

$$f_y(y) = y e^{-y}, \quad 0 \leq y < \infty,$$

and $f_y(y) = 0$ elsewhere. Then, the density of $u = 2y$, denoted by $f_u(u)$, which is given by

$$f_u(u) = \frac{u}{2} e^{-\frac{u}{2}}, \quad 0 \leq u < \infty,$$

and $f_u(u) = 0$ elsewhere, is that of a real scalar gamma with the parameters ($\alpha = 2$, $\beta = 2$).

2.1.3. The type-2 beta and F distributions in the real domain

What about the distribution of the ratio of two independently distributed real scalar chisquare random variables? Let $y_1 \sim \chi_m^2$ and $y_2 \sim \chi_n^2$, that is, y_1 and y_2 are real chisquare random variables with m and n degrees of freedom respectively, and assume that y_1 and y_2 are independently distributed. Let us determine the density of $u = y_1/y_2$. Let $v = y_2$ and consider the transformation (y_1, y_2) onto (u, v) . Noting that

$$\frac{\partial u}{\partial y_1} = \frac{1}{y_2}, \quad \frac{\partial v}{\partial y_2} = 1, \quad \frac{\partial v}{\partial y_1} = 0,$$

one has

$$\begin{aligned} du \wedge dv &= \begin{vmatrix} \frac{\partial u}{\partial y_1} & \frac{\partial u}{\partial y_2} \\ \frac{\partial v}{\partial y_1} & \frac{\partial v}{\partial y_2} \end{vmatrix} dy_1 \wedge dy_2 = \begin{vmatrix} \frac{1}{y_2} & * \\ 0 & 1 \end{vmatrix} dy_1 \wedge dy_2 \\ &= \frac{1}{y_2} dy_1 \wedge dy_2 \Rightarrow dy_1 \wedge dy_2 = v du \wedge dv \end{aligned}$$

where the asterisk indicates the presence of some element in which we are not interested owing to the triangular pattern for the Jacobian matrix. Letting the joint density of y_1 and y_2 be denoted by $f_{12}(y_1, y_2)$, one has

$$f_{12}(y_1, y_2) = \frac{1}{2^{\frac{m+n}{2}} \Gamma(\frac{m}{2}) \Gamma(\frac{n}{2})} y_1^{\frac{m}{2}-1} y_2^{\frac{n}{2}-1} e^{-\frac{y_1+y_2}{2}}$$

for $0 \leq y_1 < \infty$, $0 \leq y_2 < \infty$, $m, n = 1, 2, \dots$, and $f_{12}(y_1, y_2) = 0$ elsewhere. Let the joint density of u and v be denoted by $g_{12}(u, v)$ and the marginal density of u be denoted by $g_1(u)$. Then,

$$\begin{aligned}
 g_{12}(u, v) &= c v(uv)^{\frac{m}{2}-1} v^{\frac{n}{2}-1} e^{-\frac{1}{2}(uv+v)}, \quad c = \frac{1}{2^{\frac{m+n}{2}} \Gamma(\frac{m}{2}) \Gamma(\frac{n}{2})}, \\
 g_1(u) &= c u^{\frac{m}{2}-1} \int_{v=0}^{\infty} v^{\frac{m+n}{2}-1} e^{-v\frac{(1+u)}{2}} dv \\
 &= c u^{\frac{m}{2}-1} \Gamma\left(\frac{m+n}{2}\right) \left(\frac{1+u}{2}\right)^{-\frac{m+n}{2}} \\
 &= \frac{\Gamma(\frac{m+n}{2})}{\Gamma(\frac{m}{2}) \Gamma(\frac{n}{2})} u^{\frac{m}{2}-1} (1+u)^{-\frac{m+n}{2}} \tag{2.1.3}
 \end{aligned}$$

for $m, n = 1, 2, \dots$, $0 \leq u < \infty$ and $g_1(u) = 0$ elsewhere. Note that $g_1(u)$ is a type-2 real scalar beta density. Hence, we have the following result:

Theorem 2.1.3. *Let the real scalar $y_1 \sim \chi_m^2$ and $y_2 \sim \chi_n^2$ be independently distributed, then the ratio $u = \frac{y_1}{y_2}$ is a type-2 real scalar beta random variable with the parameters $\frac{m}{2}$ and $\frac{n}{2}$ where $m, n = 1, 2, \dots$, whose density is provided in (2.1.3).*

This result also holds for general real scalar gamma random variables $x_1 > 0$ and $x_2 > 0$ with parameters (α_1, β) and (α_2, β) , respectively, where β is a common scale parameter and it is assumed that x_1 and x_2 are independently distributed. Then, $u = \frac{x_1}{x_2}$ is a type-2 beta with parameters α_1 and α_2 .

If u as defined in Theorem 2.1.3 is replaced by $\frac{m}{n} F_{m,n}$ or $F = F_{m,n} = \frac{\chi_m^2/m}{\chi_n^2/n} = \frac{n}{m} u$ is known as the F -random variable with m and n degrees of freedom, where the degrees of freedom indicate those of the numerator and denominator chisquare random variables which are independently distributed. Denoting the density of F by $f_F(F)$ we have the following result:

Theorem 2.1.4. *Letting $F = F_{m,n} = \frac{\chi_m^2/m}{\chi_n^2/n}$ where the two real scalar chisquares are independently distributed, the real scalar F -density is given by*

$$f_F(F) = \frac{\Gamma(\frac{m+n}{2})}{\Gamma(\frac{m}{2}) \Gamma(\frac{n}{2})} \left(\frac{m}{n}\right)^{\frac{m}{2}} \frac{F^{\frac{m}{2}-1}}{\left(1 + \frac{m}{n} F\right)^{\frac{m+n}{2}}} \tag{2.1.4}$$

for $0 \leq F < \infty$, $m, n = 1, 2, \dots$, and $f_F(F) = 0$ elsewhere.

Example 2.1.3. Let x_1 and x_2 be independently distributed real scalar gamma random variables with parameters (α_1, β) and (α_2, β) , respectively, β being a common parameter, whose densities are as specified in (2.1.2). Let $u_1 = \frac{x_1}{x_1+x_2}$, $u_2 = \frac{x_1}{x_2}$, $u_3 = x_1 + x_2$.

Show that (1): u_3 has a real scalar gamma density as given in (2.1.2) with the parameters $(\alpha_1 + \alpha_2, \beta)$; (2): u_1 and u_3 as well as u_2 and u_3 are independently distributed; (3): u_2 is a real scalar type-2 beta with parameters (α_1, α_2) whose density is specified in (2.1.3); (4): u_1 has a real scalar type-1 beta density given as

$$f_1(u_1) = \frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} u_1^{\alpha_1-1} (1 - u_1)^{\alpha_2-1}, \quad 0 \leq u_1 \leq 1, \quad (2.1.5)$$

for $\Re(\alpha_1) > 0$, $\Re(\alpha_2) > 0$ and zero elsewhere. [In a statistical density, the parameters are usually real; however, since the integrals exist for complex parameters, the conditions are given for complex parameters as the real parts of α_1 and α_2 , which must be positive. When they are real, the conditions will be simply $\alpha_1 > 0$ and $\alpha_2 > 0$.]

Solution 2.1.3. Since x_1 and x_2 are independently distributed, their joint density is the product of the marginal densities, which is given by

$$f_{12}(x_1, x_2) = c x_1^{\alpha_1-1} x_2^{\alpha_2-1} e^{-\frac{1}{\beta}(x_1+x_2)}, \quad 0 \leq x_j < \infty, \quad j = 1, 2, \quad (i)$$

for $\Re(\alpha_j) > 0$, $\Re(\beta) > 0$, $j = 1, 2$ and zero elsewhere, where

$$c = \frac{1}{\beta^{\alpha_1+\alpha_2} \Gamma(\alpha_1) \Gamma(\alpha_2)}.$$

Since the sum $x_1 + x_2$ is present in the exponent and both x_1 and x_2 are positive, a convenient transformation is $x_1 = r \cos^2 \theta$, $x_2 = r \sin^2 \theta$, $0 \leq r < \infty$, $0 \leq \theta \leq \frac{\pi}{2}$. Then, the Jacobian is available from the detailed derivation of Jacobian given in the beginning of Sect. 2.1.3 or from Example 1.6.1. That is,

$$dx_1 \wedge dx_2 = 2r \sin \theta \cos \theta dr \wedge d\theta. \quad (ii)$$

Then from (i) and (ii), the joint density of r and θ , denoted by $f_{r,\theta}(r, \theta)$, is the following:

$$f_{r,\theta}(r, \theta) = c (\cos^2 \theta)^{\alpha_1-1} (\sin^2 \theta)^{\alpha_2-1} 2 \cos \theta \sin \theta r^{\alpha_1+\alpha_2-1} e^{-\frac{1}{\beta}r} \quad (iii)$$

and zero elsewhere. As $f_{r,\theta}(r, \theta)$ is a product of positive integrable functions involving solely r and θ , r and θ are independently distributed. Since $u_3 = x_1 + x_2 = r \cos^2 \theta + r \sin^2 \theta = r$ is solely a function of r and $u_1 = \frac{x_1}{x_1+x_2} = \cos^2 \theta$ and $u_2 = \frac{\cos^2 \theta}{\sin^2 \theta}$ are solely functions of θ , it follows that u_1 and u_3 as well as u_2 and u_3 are independently distributed. From (iii), upon multiplying and dividing by $\Gamma(\alpha_1 + \alpha_2)$, we obtain the density of u_3 as

$$f_1(u_3) = \frac{1}{\beta^{\alpha_1+\alpha_2} \Gamma(\alpha_1 + \alpha_2)} u_3^{\alpha_1+\alpha_2-1} e^{-\frac{u_3}{\beta}}, \quad 0 \leq u_3 < \infty, \quad (iv)$$

and zero elsewhere, which is a real scalar gamma density with parameters $(\alpha_1 + \alpha_2, \beta)$. From (iii), the density of θ , denoted by $f_2(\theta)$, is

$$f_2(\theta) = c (\cos^2 \theta)^{\alpha_1-1} (\sin^2 \theta)^{\alpha_2-1}, \quad 0 \leq \theta \leq \frac{\pi}{2} \quad (v)$$

and zero elsewhere, for $\Re(\alpha_j) > 0, j = 1, 2, \dots$. From this result, we can obtain the density of $u_1 = \cos^2 \theta$. Then, $du_1 = -2 \cos \theta \sin \theta d\theta$. Moreover, when $\theta \rightarrow 0, u_1 \rightarrow 1$ and when $\theta \rightarrow \frac{\pi}{2}, u_1 \rightarrow 0$. Hence, the minus sign in the Jacobian is needed to obtain the limits in the natural order, $0 \leq u_1 \leq 1$. Substituting in (v), the density of u_1 denoted by $f_3(u_1)$, is as given in (2.1.5), u_1 being a real scalar type-1 beta random variable with parameters (α_1, α_2) . Now, observe that

$$u_2 = \frac{\cos^2 \theta}{\sin^2 \theta} = \frac{\cos^2 \theta}{1 - \cos^2 \theta} = \frac{u_1}{1 - u_1}. \quad (vi)$$

Given the density of u_1 as specified in (2.1.5), we can obtain the density of u_2 as follows. As $u_2 = \frac{u_1}{1-u_1}$, we have $u_1 = \frac{u_2}{1+u_2} \Rightarrow du_1 = \frac{1}{(1+u_2)^2} du_2$; then substituting these values in the density of u_1 , we have the following density for u_2 :

$$f_3(u_2) = \frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} u_2^{\alpha_1-1} (1 + u_2)^{-(\alpha_1+\alpha_2)}, \quad 0 \leq u_2 < \infty, \quad (2.1.6)$$

and zero elsewhere, for $\Re(\alpha_j) > 0, j = 1, 2$, which is a real scalar type-2 beta density with parameters (α_1, α_2) . The results associated with the densities (2.1.5) and (2.1.6) are now stated as a theorem.

Theorem 2.1.5. *Let x_1 and x_2 be independently distributed real scalar gamma random variables with the parameters $(\alpha_1, \beta), (\alpha_2, \beta)$, respectively, β being a common scale parameter. [If $x_1 \sim \chi_m^2$ and $x_2 \sim \chi_n^2$, then $\alpha_1 = \frac{m}{2}, \alpha_2 = \frac{n}{2}$ and $\beta = 2$.] Then $u_1 = \frac{x_1}{x_1+x_2}$ is a real scalar type-1 beta whose density is as specified in (2.1.5) with the parameters (α_1, α_2) , and $u_2 = \frac{x_1}{x_2}$ is a real scalar type-2 beta whose density is as given in (2.1.6) with the parameters (α_1, α_2) .*

2.1a.3. The type-2 beta and F distributions in the complex domain

It follows that in the complex domain, if $\tilde{\chi}_m^2$ and $\tilde{\chi}_n^2$ are independently distributed, then the sum is a chisquare with $m + n$ degrees of freedom, that is, $\tilde{\chi}_m^2 + \tilde{\chi}_n^2 = \tilde{\chi}_{m+n}^2$. We now look into type-2 beta variables and F -variables and their connection to chisquare variables in the complex domain. Since, in the complex domain, the chisquares are actually real variables, the density of the ratio of two independently distributed chisquares with m and n degrees of freedom in the complex domain, remains the same as the density given in

(2.1.3) with $\frac{m}{2}$ and $\frac{n}{2}$ replaced by m and n , respectively. Thus, letting $\tilde{u} = \tilde{\chi}_m^2 / \tilde{\chi}_n^2$ where the two chisquares in the complex domain are independently distributed, the density of \tilde{u} , denoted by $\tilde{g}_1(u)$, is

$$\tilde{g}_1(u) = \frac{\Gamma(m+n)}{\Gamma(m)\Gamma(n)} u^{m-1} (1+u)^{-(m+n)} \quad (2.1a.3)$$

for $0 \leq u < \infty$, $m, n = 1, 2, \dots$, and $\tilde{g}_1(u) = 0$ elsewhere.

Theorem 2.1a.3. *Let $\tilde{y}_1 \sim \tilde{\chi}_m^2$ and $\tilde{y}_2 \sim \tilde{\chi}_n^2$ be independently distributed where \tilde{y}_1 and \tilde{y}_2 are in the complex domain; then, $\tilde{u} = \frac{\tilde{y}_1}{\tilde{y}_2}$ is a real type-2 beta whose density is given in (2.1a.3).*

If the F random variable in the complex domain is defined as $\tilde{F}_{m,n} = \frac{\tilde{\chi}_m^2/m}{\tilde{\chi}_n^2/n}$ where the two chisquares in the complex domain are independently distributed, then the density of \tilde{F} is that of the real F -density with m and n replaced by $2m$ and $2n$ in (2.1.4), respectively.

Theorem 2.1a.4. *Let $\tilde{F} = \tilde{F}_{m,n} = \frac{\tilde{\chi}_m^2/m}{\tilde{\chi}_n^2/n}$ where the two chisquares in the complex domain are independently distributed; then, \tilde{F} is referred to as an F random variable in the complex domain and it has a real F -density with the parameters m and n , which is given by*

$$\tilde{g}_2(F) = \frac{\Gamma(m+n)}{\Gamma(m)\Gamma(n)} \left(\frac{m}{n}\right)^m F^{m-1} \left(1 + \frac{m}{n}F\right)^{-(m+n)} \quad (2.1a.4)$$

for $0 \leq F < \infty$, $m, n = 1, 2, \dots$, and $\tilde{g}_2(F) = 0$ elsewhere.

A type-1 beta representation in the complex domain can similarly be obtained from Theorem 2.1a.3. This will be stated as a theorem.

Theorem 2.1a.5. *Let $\tilde{x}_1 \sim \tilde{\chi}_m^2$ and $\tilde{x}_2 \sim \tilde{\chi}_n^2$ be independently distributed scalar chisquare variables in the complex domain with m and n degrees of freedom, respectively. Let $\tilde{u}_1 = \frac{\tilde{x}_1}{\tilde{x}_1 + \tilde{x}_2}$, which is a real variable that we will call u_1 . Then, \tilde{u}_1 is a scalar type-1 beta random variable in the complex domain with the parameters m, n , whose real scalar density is*

$$\tilde{f}_1(\tilde{u}_1) = \frac{\Gamma(m+n)}{\Gamma(m)\Gamma(n)} u_1^{m-1} (1-u_1)^{n-1}, \quad 0 \leq u_1 \leq 1, \quad (2.1a.5)$$

and zero elsewhere, for $m, n = 1, 2, \dots$

2.1.4. Power transformation of type-1 and type-2 beta random variables

Let us make a power transformation of the type $u_1 = ay^\delta$, $a > 0$, $\delta > 0$. Then, $du_1 = a\delta y^{\delta-1}dy$. For convenience, let the parameters in (2.1.5) be $\alpha_1 = \alpha$ and $\alpha_2 = \beta$. Then, the density given in (2.1.5) becomes

$$f_{11}(y) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \delta a^\alpha y^{\alpha\delta-1} (1 - ay^\delta)^{\beta-1}, \quad 0 \leq y \leq \frac{1}{a^{\frac{1}{\delta}}}, \quad (2.1.7)$$

and zero elsewhere, for $a > 0$, $\delta > 0$, $\Re(\alpha) > 0$, $\Re(\beta) > 0$. We can extend the support to $-a^{-\frac{1}{\delta}} \leq y \leq a^{-\frac{1}{\delta}}$ by replacing y by $|y|$ and multiplying the normalizing constant by $\frac{1}{2}$. Such power transformed models are useful in practical applications. Observe that a power transformation has the following effect: for $y < 1$, the density is reduced if $\delta > 1$ or raised if $\delta < 1$, whereas for $y > 1$, the density increases if $\delta > 1$ or diminishes if $\delta < 1$. For instance, the particular case $\alpha = 1$ is highly useful in reliability theory and stress-strength analysis. Thus, letting $\alpha = 1$ in the original real scalar type-1 beta density (2.1.7) and denoting the resulting density by $f_{12}(y)$, one has

$$f_{12}(y) = a\delta\beta y^{\delta-1} (1 - ay^\delta)^{\beta-1}, \quad 0 \leq y \leq a^{-\frac{1}{\delta}}, \quad (2.1.8)$$

for $a > 0$, $\delta > 0$, $\Re(\beta) > 0$, and zero elsewhere. In the model in (2.1.8), the reliability, that is, $Pr\{y \geq t\}$, for some t , can be easily determined. As well, the hazard function $\frac{f_{12}(y=t)}{Pr\{y \geq t\}}$ is readily available. Actually, the reliability or survival function is

$$Pr\{y \geq t\} = (1 - at^\delta)^\beta, \quad a > 0, \delta > 0, t > 0, \beta > 0, \quad (i)$$

and the hazard function is

$$\frac{f_{12}(y=t)}{Pr\{y \geq t\}} = \frac{a\delta\beta t^{\delta-1}}{1 - at^\delta}. \quad (ii)$$

Observe that the free parameters a , δ and β allow for much versatility in model building situations. If $\beta = 1$ in the real scalar type-1 beta model in (2.1.7), then the density reduces to $\alpha y^{\alpha-1}$, $0 \leq y \leq 1$, $\alpha > 0$, which is a simple power function. The most popular power function model in the statistical literature is the Weibull model, which is a power transformed exponential density. Consider the real scalar exponential density

$$g(x) = \theta e^{-\theta x}, \quad \theta > 0, x \geq 0, \quad (iii)$$

and zero elsewhere, and let $x = y^\delta$, $\delta > 0$. Then the model in (iii) becomes the real scalar Weibull density, denoted by $g_1(y)$:

$$g_1(y) = \theta\delta y^{\delta-1} e^{-\theta y^\delta}, \quad \theta > 0, \delta > 0, y \geq 0, \quad (iv)$$

and zero elsewhere.

Now, let us consider power transformations in a real scalar type-2 beta density given in (2.1.6). For convenience let $\alpha_1 = \alpha$ and $\alpha_2 = \beta$. Letting $y_2 = ay^\delta$, $a > 0$, $\delta > 0$, the model specified by (2.1.6) then becomes

$$f_{21}(y) = a^\alpha \delta \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{\alpha\delta-1} (1 + ay^\delta)^{-(\alpha+\beta)} \quad (v)$$

for $a > 0$, $\delta > 0$, $\Re(\alpha) > 0$, $\Re(\beta) > 0$, and zero elsewhere. As in the type-1 beta case, the most interesting special case occurs when $\alpha = 1$. Denoting the resulting density by $f_{22}(y)$, we have

$$f_{22}(y) = a\delta\beta y^{\delta-1} (1 + ay^\delta)^{-(\beta+1)}, \quad 0 \leq y < \infty, \quad (2.1.9)$$

for $a > 0$, $\delta > 0$, $\Re(\beta) > 0$, $\alpha = 1$, and zero elsewhere. In this case as well, the reliability and hazard functions can easily be determined:

$$\text{Reliability function} = Pr\{y \geq t\} = (1 + at^\delta)^{-\beta}, \quad (vi)$$

$$\text{Hazard function} = \frac{f_{22}(y=t)}{Pr\{y \geq t\}} = \frac{a\delta\beta t^{\delta-1}}{1 + at^\delta}. \quad (vii)$$

Again, for application purposes, the forms in (vi) and (vii) are seen to be very versatile due to the presence of the free parameters a , δ and β .

2.1.5. Exponentiation of real scalar type-1 and type-2 beta variables

Let us consider the real scalar type-1 beta model in (2.1.5) where, for convenience, we let $\alpha_1 = \alpha$ and $\alpha_2 = \beta$. Letting $u_1 = ae^{-by}$, we denote the resulting density by $f_{13}(y)$ where

$$f_{13}(y) = a^\alpha b \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} e^{-b\alpha y} (1 - ae^{-by})^{\beta-1}, \quad y \geq \ln a^{\frac{1}{b}}, \quad (2.1.10)$$

for $a > 0$, $b > 0$, $\Re(\alpha) > 0$, $\Re(\beta) > 0$, and zero elsewhere. Again, for practical application the special case $\alpha = 1$ is the most useful one. Let the density corresponding to this special case be denoted by $f_{14}(y)$. Then,

$$f_{14}(y) = ab\beta e^{-by} (1 - ae^{-by})^{\beta-1}, \quad y \geq \ln a^{\frac{1}{b}}, \quad (i)$$

for $a > 0$, $b > 0$, $\beta > 0$, and zero elsewhere. In this case,

$$\text{Reliability function} = Pr\{y \geq t\} = (1 - ae^{-bt})^\beta, \quad (ii)$$

$$\text{Hazard function} = \frac{f_{14}(y=t)}{Pr\{y \geq t\}} = \frac{ab\beta e^{-bt}}{[1 - ae^{-bt}]}. \quad (iii)$$

Now, consider exponentiating a real scalar type-2 beta random variable whose density is given in (2.1.6). For convenience, we will let the parameters be $(\alpha_1 = \alpha$ and $\alpha_2 = \beta)$. Letting $u_2 = e^{-by}$ in (2.1.6), we obtain the following density:

$$f_{21}(y) = a^\alpha b \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} e^{-b\alpha y} (1 + ae^{-by})^{-(\alpha+\beta)}, \quad -\infty < y < \infty, \quad (2.1.11)$$

for $a > 0$, $b > 0$, $\Re(\alpha) > 0$, $\Re(\beta) > 0$, and zero elsewhere. The model in (2.1.11) is in fact the *generalized logistic model* introduced by Mathai and Provost (2006). For the special case $\alpha = 1$, $\beta = 1$, $a = 1$, $b = 1$ in (2.1.11), we have the following density:

$$f_{22}(y) = \frac{e^{-y}}{(1 + e^{-y})^2} = \frac{e^y}{(1 + e^y)^2}, \quad -\infty < y < \infty. \quad (iv)$$

This is the famous *logistic model* which is utilized in industrial applications.

2.1.6. The Student- t distribution in the real domain

A real Student- t variable with ν degrees of freedom, denoted by t_ν , is defined as $t_\nu = \frac{z}{\sqrt{\chi_\nu^2/\nu}}$ where $z \sim N_1(0, 1)$ and χ_ν^2 is a real scalar chisquare with ν degrees of freedom, z and χ_ν^2 being independently distributed. It follows from the definition of a real $F_{m,n}$ random variable, that $t_\nu^2 = \frac{z^2}{\chi_\nu^2/\nu} = F_{1,\nu}$, an F random variable with 1 and ν degrees of freedom. Thus, the density of t_ν^2 is available from that of an $F_{1,\nu}$. On substituting the values $m = 1$, $n = \nu$ in the F -density appearing in (2.1.4), we obtain the density of $t^2 = w$, denoted by $f_w(w)$, as

$$f_w(w) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi}\Gamma(\frac{\nu}{2})} \left(\frac{1}{\nu}\right)^{\frac{1}{2}} \frac{w^{\frac{1}{2}-1}}{(1 + \frac{w}{\nu})^{\frac{\nu+1}{2}}}, \quad 0 \leq w < \infty, \quad (2.1.12)$$

for $w = t^2$, $\nu = 1, 2, \dots$ and $f_w(w) = 0$ elsewhere. Since $w = t^2$, then the part of the density for $t > 0$ is available from (2.1.12) by observing that $\frac{1}{2}w^{\frac{1}{2}-1}dw = dt$ for $t > 0$. Hence for $t > 0$ that part of the Student- t density is available from (2.1.12) as

$$f_{1t}(t) = 2 \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi\nu}\Gamma(\frac{\nu}{2})} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}}, \quad 0 \leq t < \infty, \quad (2.1.13)$$

and zero elsewhere. Since (2.1.13) is symmetric, we extend it over $(-\infty, \infty)$ and so, obtain the real Student- t density, denoted by $f_t(t)$. This is stated in the next theorem.

Theorem 2.1.6. Consider a real scalar standard normal variable z , which is divided by the square root of a real chisquare variable with ν degrees of freedom divided by its

number of degrees of freedom ν , that is, $t = \frac{z}{\sqrt{\chi_{\nu}^2/\nu}}$, where z and χ_{ν}^2 are independently distributed; then t is known as the real scalar Student- t variable and its density is given by

$$f_t(t) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi\nu}\Gamma(\frac{\nu}{2})} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}}, \quad -\infty < t < \infty, \quad (2.1.14)$$

for $\nu = 1, 2, \dots$

2.1a.4. The Student- t distribution in the complex domain

Let $\tilde{z} \sim \tilde{N}_1(0, 1)$ and $\tilde{y} \sim \tilde{\chi}_{\nu}^2$ in the complex domain or equivalently \tilde{y} is distributed as a real gamma with the parameters ($\alpha = \nu, \beta = 1$), and let these random variables be independently distributed. Then, we will define Student- t with ν degrees of freedom in the complex domain as follows:

$$\tilde{t} = \tilde{t}_{\nu} = \frac{|\tilde{z}|}{\sqrt{\tilde{\chi}_{\nu}^2/\nu}}, \quad |\tilde{z}| = (z_1^2 + z_2^2)^{\frac{1}{2}}, \quad \tilde{z} = z_1 + iz_2$$

with z_1, z_2 real and $i = \sqrt{-1}$. What is then the density of \tilde{t}_{ν} ? The joint density of \tilde{z} and \tilde{y} , denoted by $\tilde{f}(\tilde{y}, \tilde{z})$, is

$$\tilde{f}(\tilde{z}, \tilde{y})d\tilde{y} \wedge d\tilde{z} = \frac{1}{\pi\Gamma(\nu)} y^{\nu-1} e^{-y-|\tilde{z}|^2} d\tilde{y} \wedge d\tilde{z}.$$

Let $\tilde{z} = z_1 + iz_2$, $i = \sqrt{-1}$, where $z_1 = r \cos \theta$ and $z_2 = r \sin \theta$, $0 \leq r < \infty$, $0 \leq \theta \leq 2\pi$. Then, $dz_1 \wedge dz_2 = r dr \wedge d\theta$, and the joint density of r and \tilde{y} , denoted by $f_1(r, y)$, is the following after integrating out θ , observing that y has a real gamma density:

$$f_1(r, y)dr \wedge dy = \frac{2}{\Gamma(\nu)} y^{\nu-1} e^{-y-r^2} r dr \wedge dy.$$

Let $u = t^2 = \frac{vr^2}{y}$ and $y = w$. Then, $du \wedge dw = \frac{2vr}{w} dr \wedge dy$ and so, $r dr \wedge dy = \frac{w}{2\nu} du \wedge dw$. Letting the joint density of u and w be denoted by $f_2(u, w)$, we have

$$f_2(u, w) = \frac{1}{\nu\Gamma(\nu)} w^{\nu} e^{-(w+\frac{uw}{\nu})}$$

and the marginal density of u , denoted by $g(u)$, is as follows:

$$g(u) = \int_{w=0}^{\infty} f_2(u, w) dw = \int_0^{\infty} \frac{w^{\nu}}{\Gamma(\nu+1)} e^{-w(1+\frac{u}{\nu})} dw = \left(1 + \frac{u}{\nu}\right)^{-(\nu+1)}$$

for $0 \leq u < \infty$, $\nu = 1, 2, \dots$, $u = t^2$ and zero elsewhere. Thus the part of the density of t , for $t > 0$ denoted by $f_{1t}(t)$ is as follows, observing that $du = 2tdt$ for $t > 0$:

$$f_{1t}(t) = 2t \left(1 + \frac{t^2}{\nu}\right)^{-(\nu+1)}, \quad 0 \leq t < \infty, \quad \nu = 1, 2, \dots \quad (2.1a.6)$$

Extending this density over the real line, we obtain the following density of \tilde{t} in the complex case:

$$\tilde{f}_\nu(t) = |t| \left(1 + \frac{t^2}{\nu}\right)^{-(\nu+1)}, \quad -\infty < t < \infty, \quad \nu = 1, \dots \quad (2.1a.7)$$

Thus, the following result:

Theorem 2.1a.6. *Let $\tilde{z} \sim \tilde{N}_1(0, 1)$, $\tilde{y} \sim \tilde{\chi}_\nu^2$, a scalar chisquare in the complex domain and let \tilde{z} and \tilde{y} in the complex domain be independently distributed. Consider the real variable $t = t_\nu = \frac{|\tilde{z}|}{\sqrt{\tilde{y}/\nu}}$. Then this t will be called a Student- t with ν degrees of freedom in the complex domain and its density is given by (2.1a.7).*

2.1.7. The Cauchy distribution in the real domain

We have already seen a ratio distribution in Sect. 2.1.3, namely the real type-2 beta distribution and, as particular cases, the real F-distribution and the real t^2 distribution. We now consider a ratio of two independently distributed real standard normal variables. Let $z_1 \sim N_1(0, 1)$ and $z_2 \sim N_1(0, 1)$ be independently distributed. The joint density of z_1 and z_2 , denoted by $f(z_1, z_2)$, is given by

$$f(z_1, z_2) = \frac{1}{2\pi} e^{-\frac{1}{2}(z_1^2 + z_2^2)}, \quad -\infty < z_j < \infty, \quad j = 1, 2.$$

Consider the quadrant $z_1 > 0, z_2 > 0$ and the transformation $u = \frac{z_1}{z_2}$, $v = z_2$. Then $dz_1 \wedge dz_2 = vdu \wedge dv$, see Sect. 2.1.3. Note that $u > 0$ covers the quadrants $z_1 > 0, z_2 > 0$ and $z_1 < 0, z_2 < 0$. The part of the density of u in the quadrant $u > 0, v > 0$, denoted as $g(u, v)$, is given by

$$g(u, v) = \frac{v}{2\pi} e^{-\frac{1}{2}v^2(1+u^2)}$$

and that part of the marginal density of u , denoted by $g_1(u)$, is

$$g_1(u) = \frac{1}{2\pi} \int_0^\infty v e^{-v^2 \frac{(1+u^2)}{2}} dv = \frac{1}{2\pi(1+u^2)}.$$

The other two quadrants $z_1 > 0, z_2 < 0$ and $z_1 < 0, z_2 > 0$, which correspond to $u < 0$, will yield the same form as above. Accordingly, the density of the ratio $u = \frac{z_1}{z_2}$, known as the real Cauchy density, is as specified in the next theorem.

Theorem 2.1.7. *Consider the independently distributed real standard normal variables $z_1 \sim N_1(0, 1)$ and $z_2 \sim N_1(0, 1)$. Then the ratio $u = \frac{z_1}{z_2}$ has the real Cauchy distribution having the following density:*

$$g_u(u) = \frac{1}{\pi(1+u^2)}, \quad -\infty < u < \infty. \quad (2.1.15)$$

By integrating out in each interval $(-\infty, 0)$ and $(0, \infty)$, with the help of a type-2 beta integral, it can be established that (2.1.15) is indeed a density. Since $g_u(u)$ is symmetric, $Pr\{u \leq 0\} = Pr\{u \geq 0\} = \frac{1}{2}$, and one could posit that the mean value of u may be zero. However, observe that

$$\int_0^{\infty} \frac{u}{1+u^2} du = \frac{1}{2} \ln(1+u^2) \Big|_0^{\infty} \rightarrow \infty.$$

Thus, $E(u)$, the mean value of a real Cauchy random variable, does not exist, which implies that the higher moments do not exist either.

Exercises 2.1

2.1.1. Consider someone throwing dart at a board to hit a point on the board. Taking this target point as the origin, consider a rectangular coordinate system. If (x, y) is a point of hit, then compute the densities of x and y under the following assumptions: (1): There is no bias in the horizontal and vertical directions or x and y are independently distributed; (2): The joint density is a function of the distance from the origin $\sqrt{x^2 + y^2}$. That is, if $f_1(x)$ and $f_2(y)$ are the densities of x and y then it is given that $f_1(x)f_2(y) = g(\sqrt{x^2 + y^2})$ where f_1, f_2, g are unknown functions. Show that f_1 and f_2 are identical and real normal densities.

2.1.2. Generalize Exercise 2.1.1 to 3-dimensional Euclidean space or

$$g(\sqrt{x^2 + y^2 + z^2}) = f_1(x)f_2(y)f_3(z).$$

2.1.3. Generalize Exercise 2.1.2 to k -space, $k \geq 3$.

2.1.4. Let $f(x)$ be an arbitrary density. Then Shannon's measure of entropy or uncertainty is $S = -k \int_x f(x) \ln f(x) dx$ where k is a constant. Optimize S , subject to the conditions

(a): $\int_{-\infty}^{\infty} f(x)dx = 1$; (b): Condition in (a) plus $\int_{-\infty}^{\infty} xf(x)dx = \text{given quantity}$; (c): The conditions in (b) plus $\int_{-\infty}^{\infty} x^2 f(x)dx = \text{a given quantity}$. Show that under (a), f is a uniform density; under (b), f is an exponential density and under (c), f is a Gaussian density. Hint: Use Calculus of Variation.

2.1.5. Let the error of measurement ϵ satisfy the following conditions: (1) $\epsilon = \epsilon_1 + \epsilon_2 + \dots$ or it is a sum of infinitely many infinitesimal contributions ϵ_j 's where the ϵ_j 's are independently distributed. (2): Suppose that ϵ_j can only take two values δ with probability $\frac{1}{2}$ and $-\delta$ with probability $\frac{1}{2}$ for all j . (3): $\text{Var}(\epsilon) = \sigma^2 < \infty$. Then show that this error density is real Gaussian. Hint: Use mgf. [This is Gauss' derivation of the normal law and hence it is called the error curve or Gaussian density also.]

2.1.6. The pathway model of Mathai (2005) has the following form in the case of real positive scalar variable x :

$$f_1(x) = c_1 x^\gamma [1 - a(1 - q)x^\delta]^{-\frac{1}{1-q}}, q < 1, 0 \leq x \leq [a(1 - q)]^{-\frac{1}{\delta}},$$

for $\delta > 0, a > 0, \gamma > -1$ and $f_1(x) = 0$ elsewhere. Show that this generalized type-1 beta form changes to generalized type-2 beta form for $q > 1$,

$$f_2(x) = c_2 x^\gamma [1 + a(q - 1)x^\delta]^{-\frac{1}{q-1}}, q > 1, x \geq 0, \delta > 0, a > 0$$

and $f_2(x) = 0$ elsewhere, and for $q \rightarrow 1$, the model goes into a generalized gamma form given by

$$f_3(x) = c_3 x^\gamma e^{-ax^\delta}, a > 0, \delta > 0, x \geq 0$$

and zero elsewhere. Evaluate the normalizing constants c_1, c_2, c_3 . All models are available either from $f_1(x)$ or from $f_2(x)$ where q is the pathway parameter.

2.1.7. Make a transformation $x = e^{-t}$ in the generalized gamma model of $f_3(x)$ of Exercise 2.1.6. Show that an extreme-value density for t is available.

2.1.8. Consider the type-2 beta model

$$f(x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1 + x)^{-(\alpha+\beta)}, x \geq 0, \Re(\alpha) > 0, \Re(\beta) > 0$$

and zero elsewhere. Make the transformation $x = e^y$ and then show that y has a generalized logistic distribution and as a particular case there one gets the logistic density.

2.1.9. Show that for $0 \leq x < \infty, \beta > 0, f(x) = c[1 + e^{\alpha+\beta x}]^{-1}$ is a density, which is known as Fermi-Dirac density. Evaluate the normalizing constant c .

2.1.10. Let $f(x) = c[e^{\alpha+\beta x} - 1]^{-1}$ for $0 \leq x < \infty$, $\beta > 0$. Show that $f(x)$ is a density, known as Bose-Einstein density. Evaluate the normalizing constant c .

2.1.11. Evaluate the incomplete gamma integral $\gamma(\alpha; b) = \int_0^b x^{\alpha-1} e^{-x} dx$ and show that it can be written in terms of the confluent hypergeometric series

$${}_1F_1(\beta; \delta; y) = \sum_{k=0}^{\infty} \frac{(\beta)_k}{(\delta)_k} \frac{y^k}{k!},$$

$(\alpha)_k = \alpha(\alpha + 1) \cdots (\alpha + k - 1)$, $\alpha \neq 0$, $(\alpha)_0 = 1$ is the Pochhammer symbol. Evaluate the normalizing constant c if $f(x) = cx^{\alpha-1} e^{-x}$, $0 \leq x \leq a$, $\alpha > 0$ and zero elsewhere, is a density.

2.1.12. Evaluate the incomplete beta integral $b(\alpha; \beta; b) = \int_0^b x^{\alpha-1} (1-x)^{\beta-1}$, $\alpha > 0$, $\beta > 0$, $0 \leq b \leq 1$. Show that it is available in terms of a Gauss' hypergeometric series of the form ${}_2F_1(a, b; c; z) = \sum_{k=0}^{\infty} \frac{(a)_k (b)_k}{(c)_k} \frac{z^k}{k!}$, $|z| < 1$.

2.1.13. For the pathway model in Exercise 2.1.6 compute the reliability function $Pr\{x \geq t\}$ when $\gamma = 0$ for all the cases $q < 1$, $q > 1$, $q \rightarrow 1$.

2.1.14. *Weibull density:* In the generalized gamma density $f(x) = cx^{\gamma-1} e^{-ax^\delta}$, $x \geq 0$, $\gamma > 0$, $a > 0$, $\delta > 0$ and zero elsewhere, if $\delta = \gamma$ then $f(x)$ is called a Weibull density. For a Weibull density, evaluate the hazard function $h(t) = f(t)/Pr\{x \geq t\}$.

2.1.15. Consider a type-1 beta density $f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1}$, $0 \leq x \leq 1$, $\alpha > 0$, $\beta > 0$ and zero elsewhere. Let $\alpha = 1$. Consider a power transformation $x = y^\delta$, $\delta > 0$. Let this model be $g(y)$. Compute the reliability function $Pr\{y \geq t\}$ and the hazard function $h(t) = g(t)/Pr\{y \geq t\}$.

2.1.16. Verify that if z is a real standard normal variable, $E(e^{tz^2}) = (1-2t)^{-1/2}$, $t < 1/2$, which is the mgf of a chi-square random variable having one degree of freedom. Owing to the uniqueness of the mgf, this result establishes that $z^2 \sim \chi_1^2$.

2.2. Quadratic Forms, Chisquaredness and Independence in the Real Domain

Let x_1, \dots, x_p be iid (independently and identically distributed) real scalar random variables distributed as $N_1(0, 1)$ and X be a $p \times 1$ vector whose components are x_1, \dots, x_p , that is, $X' = (x_1, \dots, x_p)$. Consider the real quadratic form $u_1 = X'AX$ for some $p \times p$ real constant symmetric matrix $A = A'$. Then, we have the following result:

Theorem 2.2.1. *The quadratic form $u_1 = X'AX$, $A = A'$, where the components of X are iid $N_1(0, 1)$, is distributed as a real chisquare with r , $r \leq p$, degrees of freedom if and only if A is idempotent, that is, $A = A^2$, and A of rank r .*

Proof: When $A = A'$ is real, there exists an orthonormal matrix P , $PP' = I$, $P'P = I$, such that $P'AP = \text{diag}(\lambda_1, \dots, \lambda_p)$, where the λ_j 's are the eigenvalues of A . Consider the transformation $X = PY$ or $Y = P'X$. Then

$$X'AX = Y'P'APY = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \dots + \lambda_p y_p^2 \quad (i)$$

where y_1, \dots, y_p are the components of Y and $\lambda_1, \dots, \lambda_p$ are the eigenvalues of A . We have already shown in Theorem 2.1.1 that all linear functions of independent real normal variables are also real normal and hence, all the y_j 's are normally distributed. The expectation of Y is $E[Y] = E[P'X] = P'E(X) = P'O = O$ and the covariance matrix associated with Y is

$$\text{Cov}(Y) = E[Y - E(Y)][Y - E(Y)]' = E[YY'] = P'\text{Cov}(X)P = P'IP = P'P = I$$

which means that the y_j 's are real standard normal variables that are mutually independently distributed. Hence, $y_j^2 \sim \chi_1^2$ or each y_j^2 is a real chisquare with one degree of freedom each and the y_j 's are all mutually independently distributed. If $A = A^2$ and the rank of A is r , then r of the eigenvalues of A are unities and the remaining ones are equal to zero as the eigenvalues of an idempotent matrix can only be equal to zero or one, the number of ones being equal to the rank of the idempotent matrix. Then the representation in (i) becomes sum of r independently distributed real chisquares of one degree of freedom each and hence the sum is a real chisquare of r degrees of freedom. Hence, the sufficiency of the result is proved. For the necessity, we assume that $X'AX \sim \chi_r^2$ and we must prove that $A = A^2$ and A is of rank r . Note that it is assumed throughout that $A = A'$. If $X'AX$ is a real chisquare having r degrees of freedom, then the mgf of $u_1 = X'AX$ is given by $M_{u_1}(t) = (1 - 2t)^{-\frac{r}{2}}$. From the representation given in (i), the mgf's are as follows: $M_{y_j^2}(t) = (1 - 2t)^{-\frac{1}{2}} \Rightarrow M_{\lambda_j y_j^2}(t) = (1 - 2\lambda_j t)^{-\frac{1}{2}}$, $j = 1, \dots, p$, the y_j 's being independently distributed. Thus, the mgf of the right-hand side of (i) is $M_{u_1}(t) = \prod_{j=1}^p (1 - 2\lambda_j t)^{-\frac{1}{2}}$. Hence, we have

$$(1 - 2t)^{-\frac{r}{2}} = \prod_{j=1}^p (1 - 2\lambda_j t)^{-\frac{1}{2}}, \quad 1 - 2t > 0, \quad 1 - 2\lambda_j t > 0, \quad j = 1, \dots, p. \quad (ii)$$

Taking the natural logarithm of each side of (ii), expanding the terms and then comparing the coefficients of $\frac{(2t)^n}{n}$ on both sides for $n = 1, 2, \dots$, we obtain equations of the type

$$r = \sum_{j=1}^p \lambda_j = \sum_{j=1}^p \lambda_j^2 = \sum_{j=1}^p \lambda_j^3 = \dots \quad (iii)$$

The only solution resulting from (iii) is that r of the λ_j 's are unities and the remaining ones are zeros. This result, combined with the property that $A = A'$ guarantees that A is idempotent of rank r .

Observe that the eigenvalues of a matrix being ones and zeros need not imply that the matrix is idempotent; take for instance triangular matrices whose diagonal elements are unities and zeros. However, this property combined with the symmetry assumption will guarantee that the matrix is idempotent.

Corollary 2.2.1. *If the simple random sample or the iid variables came from a real $N_1(0, \sigma^2)$ distribution, then the modification needed in Theorem 2.2.1 is that $\frac{1}{\sigma^2} X'AX \sim \chi_r^2$, $A = A'$, if and only if $A = A^2$ and A is of rank r .*

The above result, Theorem 2.2.1, coupled with another result on the independence of quadratic forms, are quite useful in the areas of Design of Experiment, Analysis of Variance and Regression Analysis, as well as in model building and hypotheses testing situations. This result on the independence of quadratic forms is stated next.

Theorem 2.2.2. *Let x_1, \dots, x_p be iid variables from a real $N_1(0, 1)$ population. Consider two real quadratic forms $u_1 = X'AX$, $A = A'$ and $u_2 = X'BX$, $B = B'$, where the components of the $p \times 1$ vector X are the x_1, \dots, x_p . Then, u_1 and u_2 are independently distributed if and only if $AB = O$.*

Proof: Let us assume that $AB = O$. Then $AB = O = O' = (AB)' = B'A' = BA$. When $AB = BA$, there exists a single orthonormal matrix P , $PP' = I$, $P'P = I$, such that both the quadratic forms are reduced to their canonical forms by the same P . Let

$$u_1 = X'AX = \lambda_1 y_1^2 + \dots + \lambda_p y_p^2 \quad (i)$$

and

$$u_2 = X'BX = \nu_1 y_1^2 + \dots + \nu_p y_p^2 \quad (ii)$$

where $\lambda_1, \dots, \lambda_p$ are the eigenvalues of A and v_1, \dots, v_p are the eigenvalues of B . Since $A = A'$, the eigenvalues λ_j 's are all real. Moreover,

$$AB = O \Rightarrow P'ABP = P'APP'BP = D_1D_2 = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_p \end{bmatrix} \begin{bmatrix} v_1 & 0 & \dots & 0 \\ 0 & v_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & v_p \end{bmatrix} = O, \quad (iii)$$

which means that $\lambda_j v_j = 0$ for all $j = 1, \dots, p$. Thus, whenever a λ_j is not zero, the corresponding v_j is zero and vice versa. Accordingly, the λ_j 's and v_j 's are separated in (i) and (ii), that is, the independent components are mathematically separated and hence u_1 and u_2 are statistically independently distributed. The converse which can be stated as follows: if u_1 and u_2 are independently distributed, $A = A'$, $B = B'$ and the x_j 's are real iid $N_1(0, 1)$, then $AB = O$, is more difficult to establish. The proof which requires additional properties of matrices, will not be herein presented. Note that there are several incorrect or incomplete “proofs” in the literature. A correct derivation may be found in Mathai and Provost (1992).

When x_1, \dots, x_p are iid $N_1(0, \sigma^2)$, the above result on the independence of quadratic forms still holds since the independence is not altered by multiplying the quadratic forms by $\frac{1}{\sigma^2}$.

Example 2.2.1. Construct two 3×3 matrices A and B such that $A = A'$, $B = B'$ [both are symmetric], $A = A^2$ [A is idempotent], $AB = O$ [A and B are orthogonal to each other], and A has rank 2. Then (1): verify Theorem 2.2.1; (2): verify Theorem 2.2.2.

Solution 2.2.1. Consider the following matrices:

$$A = \begin{bmatrix} \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 1 & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

Note that both A and B are symmetric, that is, $A = A'$, $B = B'$. Further, the rank of A is 2 since the first and second row vectors are linearly independent and the third row is a multiple of the first one. Note that $A^2 = A$ and $AB = O$. Now, consider the quadratic forms $u = X'AX$ and $v = X'BX$. Then $u = \frac{1}{2}x_1^2 + x_2^2 + \frac{1}{2}x_3^2 - x_1x_3 = x_2^2 + [\frac{1}{\sqrt{2}}(x_1 - x_3)]^2$. Our initial assumption is that $x_j \sim N_1(0, 1)$, $j = 1, 2, 3$ and the x_j 's are independently distributed. Let $y_1 = \frac{1}{\sqrt{2}}(x_1 - x_3)$. Then, $E[y_1] = 0$, $\text{Var}(y_1) = +\frac{1}{2}[\text{Var}(x_1) + \text{Var}(x_3)] =$

$\frac{1}{2}[1 + 1] = 1$. Since y_1 is a linear function of normal variables, y_1 is normal with the parameters $E[y_1] = 0$ and $\text{Var}(y_1) = 1$, that is, $y_1 \sim N_1(0, 1)$, and hence $y_1^2 \sim \chi_1^2$; as well, $x_2^2 \sim \chi_1^2$. Thus, $u \sim \chi_2^2$ since x_2 and y_1 are independently distributed given that the variables are separated, noting that y_1 does not involve x_2 . This verifies Theorem 2.2.1. Now, having already determined that $AB = O$, it remains to show that u and v are independently distributed where $v = X'BX = x_1^2 + x_3^2 + 2x_1x_3 = (x_1 + x_3)^2$. Let $y_2 = \frac{1}{\sqrt{2}}(x_1 + x_3) \Rightarrow y_2 \sim N_1(0, 1)$ as y_2 is a linear function of normal variables and hence normal with parameters $E[y_2] = 0$ and $\text{Var}(y_2) = 1$. On noting that v does not contain x_2 , we need only consider the parts of u and v containing x_1 and x_3 . Thus, our question reduces to: are y_1 and y_2 independently distributed? Since both y_1 and y_2 are linear functions of normal variables, both y_1 and y_2 are normal. Since the covariance between y_1 and y_2 , that is, $\text{Cov}(y_1, y_2) = \frac{1}{2}\text{Cov}(x_1 - x_3, x_1 + x_3) = \frac{1}{2}[\text{Var}(x_1) - \text{Var}(x_3)] = \frac{1}{2}[1 - 1] = 0$, the two normal variables are uncorrelated and hence, independently distributed. That is, y_1 and y_2 are independently distributed, thereby implying that u and v are also independently distributed, which verifies Theorem 2.2.2.

2.2a. Hermitian Forms, Chisquaredness and Independence in the Complex Domain

Let $\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_k$ be independently and identically distributed standard univariate Gaussian variables in the complex domain and let \tilde{X} be a $k \times 1$ vector whose components are $\tilde{x}_1, \dots, \tilde{x}_k$. Consider the Hermitian form $\tilde{X}^*A\tilde{X}$, $A = A^*$ (Hermitian) where A is a $k \times k$ constant Hermitian matrix. Then, there exists a unitary matrix Q , $QQ^* = I$, $Q^*Q = I$, such that $Q^*AQ = \text{diag}(\lambda_1, \dots, \lambda_k)$. Note that the λ_j 's are real since A is Hermitian. Consider the transformation $\tilde{X} = Q\tilde{Y}$. Then,

$$\tilde{X}^*A\tilde{X} = \lambda_1|\tilde{y}_1|^2 + \dots + \lambda_k|\tilde{y}_k|^2$$

where the \tilde{y}_j 's are iid standard normal in the complex domain, $\tilde{y}_j \sim \tilde{N}_1(0, 1)$, $j = 1, \dots, k$. Then, $\tilde{y}_j^*\tilde{y}_j = |\tilde{y}_j|^2 \sim \tilde{\chi}_1^2$, a chisquare having one degree of freedom in the complex domain or, equivalently, a real gamma random variable with the parameters ($\alpha = 1, \beta = 1$), the $|\tilde{y}_j|^2$'s being independently distributed for $j = 1, \dots, k$. Thus, we can state the following result whose proof parallels that in the real case.

Theorem 2.2a.1. *Let $\tilde{x}_1, \dots, \tilde{x}_k$ be iid $\tilde{N}_1(0, 1)$ variables in the complex domain. Consider the Hermitian form $u = \tilde{X}^*A\tilde{X}$, $A = A^*$ where \tilde{X} is a $k \times 1$ vector whose components are $\tilde{x}_1, \dots, \tilde{x}_k$. Then, u is distributed as a chisquare in the complex domain with r degrees of freedom or a real gamma with the parameters ($\alpha = r, \beta = 1$), if and only if A is of rank r and $A = A^2$.*

Theorem 2.2a.2. Let the \tilde{x}_j 's and \tilde{X} be as in Theorem 2.2a.1. Consider two Hermitian forms $u_1 = \tilde{X}^* A \tilde{X}$, $A = A^*$ and $u_2 = \tilde{X}^* B \tilde{X}$, $B = B^*$. Then, u_1 and u_2 are independently distributed if and only if $AB = O$ (null matrix).

Example 2.2a.1. Construct two 3×3 Hermitian matrices A and B , that is $A = A^*$, $B = B^*$, such that $A = A^2$ [idempotent] and is of rank 2 with $AB = O$. Then (1): verify Theorems 2.2a.1, (2): verify Theorem 2.2a.2.

Solution 2.2a.1. Consider the following matrices

$$A = \begin{bmatrix} \frac{1}{2} & 0 & -\frac{(1+i)}{\sqrt{8}} \\ 0 & 1 & 0 \\ -\frac{(1-i)}{\sqrt{8}} & 0 & \frac{1}{2} \end{bmatrix}, \quad B = \begin{bmatrix} \frac{1}{2} & 0 & \frac{(1+i)}{\sqrt{8}} \\ 0 & 0 & 0 \\ \frac{(1-i)}{\sqrt{8}} & 0 & \frac{1}{2} \end{bmatrix}.$$

It can be readily verified that $A = A^*$, $B = B^*$, $A = A^2$, $AB = O$. Further, on multiplying the first row of A by $-\frac{2(1-i)}{\sqrt{8}}$, we obtain the third row, and since the third row is a multiple of the first one and the first and second rows are linearly independent, the rank of A is 2. Our initial assumption is that $\tilde{x}_j \sim \tilde{N}_1(0, 1)$, $j = 1, 2, 3$, that is, they are univariate complex Gaussian, and they are independently distributed. Then, $\tilde{x}_j^* \tilde{x}_j \sim \tilde{\chi}_1^2$, a chisquare with one degree of freedom in the complex domain or a real gamma random variable with the parameters ($\alpha = 1, \beta = 1$) for each $j = 1, 2, 3$. Let us consider the Hermitian forms $u = \tilde{X}^* A \tilde{X}$ and $v = \tilde{X}^* B \tilde{X}$, $\tilde{X}' = (\tilde{x}_1, \tilde{x}_2, \tilde{x}_3)$. Then

$$\begin{aligned} u &= \frac{1}{2} \tilde{x}_1^* \tilde{x}_1 - \frac{(1+i)}{\sqrt{8}} \tilde{x}_1^* \tilde{x}_3 - \frac{(1-i)}{\sqrt{8}} \tilde{x}_3^* \tilde{x}_1 + \frac{1}{2} \tilde{x}_3^* \tilde{x}_3 + \tilde{x}_2^* \tilde{x}_2 \\ &= \tilde{x}_2^* \tilde{x}_2 + \frac{1}{2} [\tilde{x}_1^* \tilde{x}_1 - 4 \frac{(1+i)}{\sqrt{8}} \tilde{x}_1^* \tilde{x}_3 + \tilde{x}_3^* \tilde{x}_3] \\ &= \tilde{\chi}_1^2 + \left[\frac{1}{\sqrt{2}} (\tilde{y}_1 - \tilde{x}_3) \right]^* \left[\frac{1}{\sqrt{2}} (\tilde{y}_1 - \tilde{x}_3) \right] \end{aligned} \quad (i)$$

where

$$\begin{aligned} \tilde{y}_1 &= 2 \frac{(1+i)}{\sqrt{8}} \tilde{x}_1 \Rightarrow E[\tilde{y}_1] = 0, \quad \text{Var}(\tilde{y}_1) = \left| 2 \frac{(1+i)}{\sqrt{8}} \right|^2 \text{Var}(\tilde{x}_1) \\ \text{Var}(\tilde{y}_1) &= E \left\{ \left[2 \frac{(1+i)}{\sqrt{8}} \right]^* \left[2 \frac{(1+i)}{\sqrt{8}} \right] \tilde{x}_1^* \tilde{x}_1 \right\} = E\{\tilde{x}_1^* \tilde{x}_1\} = \text{Var}(\tilde{x}_1) = 1. \end{aligned} \quad (ii)$$

Since \tilde{y}_1 is a linear function of \tilde{x}_1 , it is a univariate normal in the complex domain with parameters 0 and 1 or $\tilde{y}_1 \sim \tilde{N}_1(0, 1)$. The part not containing the $\tilde{\chi}_1^2$ in (i) can be written as follows:

$$\left[\frac{1}{\sqrt{2}} (\tilde{y}_1 - \tilde{x}_3) \right]^* \left[\frac{1}{\sqrt{2}} (\tilde{y}_1 - \tilde{x}_3) \right] \sim \tilde{\chi}_1^2 \quad (iii)$$

since $\tilde{y}_1 - \tilde{x}_3 \sim \tilde{N}_1(0, 2)$ as $\tilde{y}_1 - \tilde{x}_3$ is a linear function of the normal variables \tilde{y}_1 and \tilde{x}_3 . Therefore $u = \tilde{\chi}_1^2 + \tilde{\chi}_1^2 = \tilde{\chi}_2^2$, that is, a chisquare having two degrees of freedom in the complex domain or a real gamma with the parameters ($\alpha = 2, \beta = 1$). Observe that the two chisquares are independently distributed because one of them contains only \tilde{x}_2 and the other, \tilde{x}_1 and \tilde{x}_3 . This establishes (1). In order to verify (2), we first note that the Hermitian form v can be expressed as follows:

$$v = \frac{1}{2}\tilde{x}_1^*\tilde{x}_1 + \frac{(1+i)}{\sqrt{8}}\tilde{x}_1^*\tilde{x}_3 + \frac{(1-i)}{\sqrt{8}}\tilde{x}_3^*\tilde{x}_1 + \frac{1}{2}\tilde{x}_3^*\tilde{x}_3$$

which can be written in the following form by making use of steps similar to those leading to (iii):

$$v = \left[2\frac{(1+i)}{\sqrt{8}}\frac{\tilde{x}_1}{\sqrt{2}} + \frac{\tilde{x}_3}{\sqrt{2}}\right]^* \left[2\frac{(1+i)}{\sqrt{8}}\frac{\tilde{x}_1}{\sqrt{2}} + \frac{\tilde{x}_3}{\sqrt{2}}\right] = \tilde{\chi}_1^2 \quad (iv)$$

or v is a chisquare with one degree of freedom in the complex domain. Observe that \tilde{x}_2 is absent in (iv), so that we need only compare the terms containing \tilde{x}_1 and \tilde{x}_3 in (iii) and (iv). These terms are $\tilde{y}_2 = 2\frac{(1+i)}{\sqrt{8}}\tilde{x}_1 + \tilde{x}_3$ and $\tilde{y}_3 = 2\frac{(1+i)}{\sqrt{8}}\tilde{x}_1 - \tilde{x}_3$. Noting that the covariance between \tilde{y}_2 and \tilde{y}_3 is zero:

$$\text{Cov}(\tilde{y}_2, \tilde{y}_3) = \left|2\frac{(1+i)}{\sqrt{8}}\right|^2 \text{Var}(\tilde{x}_1) - \text{Var}(\tilde{x}_3) = 1 - 1 = 0,$$

and that \tilde{y}_2 and \tilde{y}_3 are linear functions of normal variables and hence normal, the fact that they are uncorrelated implies that they are independently distributed. Thus, u and v are indeed independently distributed, which establishes (2).

2.2.1. Extensions of the results in the real domain

Let $x_j \sim N_1(\mu_j, \sigma_j^2)$, $j = 1, \dots, k$, be independently distributed. Then, $\frac{x_j}{\sigma_j} \sim N_1\left(\frac{\mu_j}{\sigma_j}, 1\right)$, $\sigma_j > 0$, $j = 1, \dots, k$. Let

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix}, \quad \mu = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_k \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_k^2 \end{bmatrix}, \quad \Sigma^{\frac{1}{2}} = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_k \end{bmatrix}.$$

Then, let

$$Y = \Sigma^{-\frac{1}{2}}X = \begin{bmatrix} y_1 \\ \vdots \\ y_k \end{bmatrix}, \quad E[Y] = \Sigma^{-\frac{1}{2}}E[X] = \Sigma^{-\frac{1}{2}}\mu.$$

If $\mu = O$, it has already been shown that $Y'Y \sim \chi_k^2$. If $\mu \neq O$, then $Y'Y \sim \chi_k^2(\lambda)$, $\lambda = \frac{1}{2}\mu'\Sigma^{-1}\mu$. It is assumed that the noncentral chisquare distribution has already been discussed in a basic course in Statistics. It is defined for instance in Mathai and Haubold (2017a, 2017b) and will be briefly discussed in Sect. 2.3.1. If $\mu = O$, then for any $k \times k$ symmetric matrix $A = A'$, $Y'AY \sim \chi_r^2$ if and only if $A = A^2$ and A is of rank r . This result has already been established. Now, if $\mu \neq O$, then $X'AX = Y'\Sigma^{\frac{1}{2}}A\Sigma^{\frac{1}{2}}Y \sim \chi_k^2$ if and only if $\Sigma^{\frac{1}{2}}A\Sigma^{\frac{1}{2}} = \Sigma^{\frac{1}{2}}A\Sigma A\Sigma^{\frac{1}{2}} \Rightarrow A = A\Sigma A$ and $\Sigma^{\frac{1}{2}}A\Sigma^{\frac{1}{2}}$ is of rank r or A is of rank r since $\Sigma > O$. Hence, we have the following result:

Theorem 2.2.3. *Let the real scalars $x_j \sim N_1(\mu_j, \sigma_j^2)$, $j = 1, \dots, k$, be independently distributed. Let*

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix}, Y = \begin{bmatrix} y_1 \\ \vdots \\ y_k \end{bmatrix} = \Sigma^{-\frac{1}{2}}X, E(Y) = \Sigma^{-\frac{1}{2}}\mu, \mu = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_k \end{bmatrix}.$$

Then for any $k \times k$ symmetric matrix $A = A'$,

$$X'AX = Y'\Sigma^{\frac{1}{2}}A\Sigma^{\frac{1}{2}}Y \sim \begin{cases} \chi_r^2 & \text{if } \mu = O \\ \chi_k^2(\lambda) & \text{if } \mu \neq O, \lambda = \frac{1}{2}\mu'\Sigma^{-\frac{1}{2}}A\Sigma^{-\frac{1}{2}}\mu \end{cases}$$

if and only if $A = A\Sigma A$ and A is of rank r .

Independence is not altered if the variables are relocated. Consider two quadratic forms $X'AX$ and $X'BX$, $A = A'$, $B = B'$. Then, $X'AX = Y'\Sigma^{\frac{1}{2}}A\Sigma^{\frac{1}{2}}Y$ and $X'BX = Y'\Sigma^{\frac{1}{2}}B\Sigma^{\frac{1}{2}}Y$, and we have the following result:

Theorem 2.2.4. *Let x_j , X , Y , Σ be as defined in Theorem 2.2.3. Then, the quadratic forms $X'AX = Y'\Sigma^{\frac{1}{2}}A\Sigma^{\frac{1}{2}}Y$ and $X'BX = Y'\Sigma^{\frac{1}{2}}B\Sigma^{\frac{1}{2}}Y$, $A = A'$, $B = B'$, are independently distributed if and only if $A\Sigma B = O$.*

Let X , A and Σ be as defined in Theorem 2.2.3, Z be a standard normal vector whose components z_i , $i = 1, \dots, k$ are iid $N_1(0, 1)$, and P be an orthonormal matrix such that $P'\Sigma^{\frac{1}{2}}A\Sigma^{\frac{1}{2}}P = \text{diag}(\lambda_1, \dots, \lambda_k)$; then, a general quadratic form $X'AX$ can be expressed as follows:

$$X'AX = (Z'\Sigma^{\frac{1}{2}} + \mu')A(\Sigma^{\frac{1}{2}}Z + \mu) = (Z + \Sigma^{-\frac{1}{2}}\mu)'PP'\Sigma^{\frac{1}{2}}A\Sigma^{\frac{1}{2}}PP'(Z + \Sigma^{-\frac{1}{2}}\mu)$$

where $\lambda_1, \dots, \lambda_k$ are the eigenvalues of $\Sigma^{\frac{1}{2}} A \Sigma^{\frac{1}{2}}$. Hence, the following decomposition of the quadratic form:

$$X'AX = \lambda_1(u_1 + b_1)^2 + \dots + \lambda_k(u_k + b_k)^2, \quad (2.2.1)$$

where $\begin{bmatrix} b_1 \\ \vdots \\ b_k \end{bmatrix} = P' \Sigma^{-\frac{1}{2}} \mu$, $\begin{bmatrix} u_1 \\ \vdots \\ u_k \end{bmatrix} = P' \begin{bmatrix} z_1 \\ \vdots \\ z_k \end{bmatrix}$, and hence the u_i 's are iid $N_1(0, 1)$.

Thus, $X'AX$ can be expressed as a linear combination of independently distributed non-central chisquare random variables, each having one degree of freedom, whose non-centrality parameters are respectively $b_j^2/2$, $j = 1, \dots, k$. Of course, the k chisquares will be central when $\mu = O$.

2.2a.1. Extensions of the results in the complex domain

Let the complex scalar variables $\tilde{x}_j \sim \tilde{N}_1(\tilde{\mu}_j, \sigma_j^2)$, $j = 1, \dots, k$, be independently distributed and $\Sigma = \text{diag}(\sigma_1^2, \dots, \sigma_k^2)$. As well, let

$$\tilde{X} = \begin{bmatrix} \tilde{x}_1 \\ \vdots \\ \tilde{x}_k \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_k^2 \end{bmatrix}, \quad \Sigma^{-\frac{1}{2}} \tilde{X} = \tilde{Y} = \begin{bmatrix} \tilde{y}_1 \\ \vdots \\ \tilde{y}_k \end{bmatrix}, \quad \tilde{\mu} = \begin{bmatrix} \tilde{\mu}_1 \\ \vdots \\ \tilde{\mu}_k \end{bmatrix}$$

where $\Sigma^{\frac{1}{2}}$ is the Hermitian positive definite square root of Σ . In this case, $\tilde{y}_j \sim \tilde{N}_1(\frac{\tilde{\mu}_j}{\sigma_j}, 1)$, $j = 1, \dots, k$ and the \tilde{y}_j 's are assumed to be independently distributed. Hence, for any Hermitian form $\tilde{X}^* A \tilde{X}$, $A = A^*$, we have $\tilde{X}^* A \tilde{X} = \tilde{Y}^* \Sigma^{\frac{1}{2}} A \Sigma^{\frac{1}{2}} \tilde{Y}$. Hence if $\tilde{\mu} = O$ (null vector), then from the previous result on chisquaredness, we have:

Theorem 2.2a.3. *Let \tilde{X} , Σ , \tilde{Y} , $\tilde{\mu}$ be as defined above. Let $u = \tilde{X}^* A \tilde{X}$, $A = A^*$ be a Hermitian form. Then $u \sim \tilde{\chi}_r^2$ in the complex domain if and only if A is of rank r , $\tilde{\mu} = O$ and $A = A \Sigma A$. [A chisquare with r degrees of freedom in the complex domain is a real gamma with parameters $(\alpha = r, \beta = 1)$.]*

If $\tilde{\mu} \neq O$, then we have a noncentral chisquare in the complex domain. A result on the independence of Hermitian forms can be obtained as well.

Theorem 2.2a.4. *Let \tilde{X} , \tilde{Y} , Σ be as defined above. Consider the Hermitian forms $u_1 = \tilde{X}^* A \tilde{X}$, $A = A^*$ and $u_2 = \tilde{X}^* B \tilde{X}$, $B = B^*$. Then u_1 and u_2 are independently distributed if and only if $A \Sigma B = O$.*

The proofs of Theorems 2.2a.3 and 2.2a.4 parallel those presented in the real case and are hence omitted.

Exercises 2.2

2.2.1. Give a proof to the second part of Theorem 2.2.2, namely, given that $X'AX$, $A = A'$ and $X'BX$, $B = B'$ are independently distributed where the components of the $p \times 1$ vector X are mutually independently distributed as real standard normal variables, then show that $AB = O$.

2.2.2. Let the real scalar $x_j \sim N_1(0, \sigma^2)$, $\sigma^2 > 0$, $j = 1, 2, \dots, k$ and be independently distributed. Let $X' = (x_1, \dots, x_k)$ or X is the $k \times 1$ vector where the elements are x_1, \dots, x_k . Then the joint density of the real scalar variables x_1, \dots, x_k , denoted by $f(X)$, is

$$f(X) = \frac{1}{(\sqrt{2\pi})^k} e^{-\frac{1}{2\sigma^2} X'X}, \quad -\infty < x_j < \infty, \quad j = 1, \dots, k.$$

Consider the quadratic form $u = X'AX$, $A = A'$ and X is as defined above. (1): Compute the mgf of u ; (2): Compute the density of u if A is of rank r and all eigenvalues of A are equal to $\lambda > 0$; (3): If the eigenvalues are $\lambda > 0$ for m of the eigenvalues and the remaining n of them are $\lambda < 0$, $m + n = r$, compute the density of u .

2.2.3. In Exercise 2.2.2 compute the density of u if (1): r_1 of the eigenvalues are λ_1 each and r_2 of the eigenvalues are λ_2 each, $r_1 + r_2 = r$. Consider all situations $\lambda_1 > 0$, $\lambda_2 > 0$ etc.

2.2.4. In Exercise 2.2.2 compute the density of u for the general case with no restrictions on the eigenvalues.

2.2.5. Let $x_j \sim N_1(0, \sigma^2)$, $j = 1, 2$ and be independently distributed. Let $X' = (x_1, x_2)$. Let $u = X'AX$ where $A = A'$. Compute the density of u if the eigenvalues of A are (1): 2 and 1, (2): 2 and -1 ; (3): Construct a real 2×2 matrix $A = A'$ where the eigenvalues are 2 and 1.

2.2.6. Show that the results on chisquaredness and independence in the real or complex domain need not hold if $A \neq A^*$, $B \neq B^*$.

2.2.7. Construct a 2×2 Hermitian matrix $A = A^*$ such that $A = A^2$ and verify Theorem 2.2a.3. Construct 2×2 Hermitian matrices A and B such that $AB = O$, and then verify Theorem 2.2a.4.

2.2.8. Let $\tilde{x}_1, \dots, \tilde{x}_m$ be a simple random sample of size m from a complex normal population $\tilde{N}_1(\tilde{\mu}_1, \sigma_1^2)$. Let $\tilde{y}_1, \dots, \tilde{y}_n$ be iid $\tilde{N}(\tilde{\mu}_2, \sigma_2^2)$. Let the two complex normal populations be independent. Let

$$s_1^2 = \sum_{j=1}^m (\tilde{x}_j - \bar{\tilde{x}})^* (\tilde{x}_j - \bar{\tilde{x}}) / \sigma_1^2, \quad s_2^2 = \sum_{j=1}^n (\tilde{y}_j - \bar{\tilde{y}})^* (\tilde{y}_j - \bar{\tilde{y}}) / \sigma_2^2,$$

$$s_{11}^2 = \frac{1}{\sigma_1^2} \sum_{j=1}^m (\tilde{x}_j - \tilde{\mu}_1)^* (\tilde{x}_j - \tilde{\mu}_1), \quad s_{21}^2 = \frac{1}{\sigma_2^2} \sum_{j=1}^n (\tilde{y}_j - \tilde{\mu}_2)^* (\tilde{y}_j - \tilde{\mu}_2)$$

Then, show that

$$\frac{s_{11}^2/m}{s_{21}^2/n} \sim \tilde{F}_{m,n}, \quad \frac{s_1^2/(m-1)}{s_2^2/(n-1)} \sim \tilde{F}_{m-1,n-1}$$

for $\sigma_1^2 = \sigma_2^2$.

2.2.9. In Exercise 2.2.8 show that $\frac{s_{11}^2}{s_{21}^2}$ is a type-2 beta with the parameters m and n , and $\frac{s_1^2}{s_2^2}$ is a type-2 beta with the parameters $m-1$ and $n-1$ for $\sigma_1^2 = \sigma_2^2$.

2.2.10. In Exercise 2.2.8 if $\sigma_1^2 = \sigma_2^2 = \sigma^2$ then show that

$$\frac{1}{\sigma^2} \left[\sum_{j=1}^m (\tilde{x}_j - \bar{\tilde{x}})^* (\tilde{x}_j - \bar{\tilde{x}}) + \sum_{j=1}^n (\tilde{y}_j - \bar{\tilde{y}})^* (\tilde{y}_j - \bar{\tilde{y}}) \right] \sim \tilde{\chi}_{m+n-2}^2.$$

2.2.11. In Exercise 2.2.10 if $\bar{\tilde{x}}$ and $\bar{\tilde{y}}$ are replaced by $\tilde{\mu}_1$ and $\tilde{\mu}_2$ respectively then show that the degrees of freedom of the chisquare is $m+n$.

2.2.12. Derive the representation of the general quadratic form $X'AX$ given in (2.2.1).

2.3. Simple Random Samples from Real Populations and Sampling Distributions

For practical applications, an important result is that on the independence of the sample mean and sample variance when the sample comes from a normal (Gaussian) population. Let x_1, \dots, x_n be a simple random sample of size n from a real $N_1(\mu_1, \sigma_1^2)$ or, equivalently, x_1, \dots, x_n are iid $N_1(\mu_1, \sigma_1^2)$. Recall that we have established that any linear function $L'X = X'L$, $L' = (a_1, \dots, a_n)$, $X' = (x_1, \dots, x_n)$ remains normally distributed (Theorem 2.1.1). Now, consider two linear forms $y_1 = L_1'X$, $y_2 = L_2'X$, with $L_1' = (a_1, \dots, a_n)$, $L_2' = (b_1, \dots, b_n)$ where a_1, \dots, a_n , b_1, \dots, b_n are real scalar constants. Let us examine the conditions that are required for assessing the independence of

the linear forms y_1 and y_2 . Since x_1, \dots, x_n are iid, we can determine the joint mgf of x_1, \dots, x_n . We take a $n \times 1$ parameter vector T , $T' = (t_1, \dots, t_n)$ where the t_j 's are scalar parameters. Then, by definition, the joint mgf is given by

$$E[e^{T'X}] = \prod_{j=1}^n M_{x_j}(t_j) = \prod_{j=1}^n e^{t_j\mu_1 + \frac{1}{2}t_j^2\sigma_1^2} = e^{\mu_1 T'J + \frac{\sigma_1^2}{2} T'T} \quad (2.3.1)$$

since the x_j 's are iid, $J' = (1, \dots, 1)$. Since every linear function of x_1, \dots, x_n is a univariate normal, we have $y_1 \sim N_1(\mu_1 L_1' J, \frac{\sigma_1^2 t_1^2}{2} L_1' L_1)$ and hence the mgf of y_1 , taking t_1 as the parameter for the mgf, is $M_{y_1}(t_1) = e^{t_1\mu_1 L_1' J + \frac{\sigma_1^2 t_1^2}{2} L_1' L_1}$. Now, let us consider the joint mgf of y_1 and y_2 taking t_1 and t_2 as the respective parameters. Let the joint mgf be denoted by $M_{y_1, y_2}(t_1, t_2)$. Then,

$$\begin{aligned} M_{y_1, y_2}(t_1, t_2) &= E[e^{t_1 y_1 + t_2 y_2}] = E[e^{(t_1 L_1' + t_2 L_2') X}] \\ &= e^{\mu_1 (t_1 L_1' + t_2 L_2') J + \frac{\sigma_1^2}{2} (t_1 L_1 + t_2 L_2)' (t_1 L_1 + t_2 L_2)} \\ &= e^{\mu_1 (L_1' + L_2') J + \frac{\sigma_1^2}{2} (t_1^2 L_1' L_1 + t_2^2 L_2' L_2 + 2t_1 t_2 L_1' L_2)} \\ &= M_{y_1}(t_1) M_{y_2}(t_2) e^{\sigma_1^2 t_1 t_2 L_1' L_2}. \end{aligned}$$

Hence, the last factor on the right-hand side has to vanish for y_1 and y_2 to be independently distributed, and this can happen if and only if $L_1' L_2 = L_2' L_1 = 0$ since t_1 and t_2 are arbitrary. Thus, we have the following result:

Theorem 2.3.1. *Let x_1, \dots, x_n be iid $N_1(\mu_1, \sigma_1^2)$. Let $y_1 = L_1' X$ and $y_2 = L_2' X$ where $X' = (x_1, \dots, x_n)$, $L_1' = (a_1, \dots, a_n)$ and $L_2' = (b_1, \dots, b_n)$, the a_j 's and b_j 's being scalar constants. Then, y_1 and y_2 are independently distributed if and only if $L_1' L_2 = L_2' L_1 = 0$.*

Example 2.3.1. Let x_1, x_2, x_3, x_4 be a simple random sample of size 4 from a real normal population $N_1(\mu = 1, \sigma^2 = 2)$. Consider the following statistics: (1): u_1, v_1, w_1 , (2): u_2, v_2, w_2 . Check for independence of various statistics in (1): and (2): where

$$\begin{aligned} u_1 &= \bar{x} = \frac{1}{4}(x_1 + x_2 + x_3 + x_4), \quad v_1 = 2x_1 - 3x_2 + x_3 + x_4, \quad w_1 = x_1 - x_2 + x_3 - x_4; \\ u_2 &= \bar{x} = \frac{1}{4}(x_1 + x_2 + x_3 + x_4), \quad v_2 = x_1 - x_2 + x_3 - x_4, \quad w_2 = x_1 - x_2 - x_3 + x_4. \end{aligned}$$

Solution 2.3.1. Let $X' = (x_1, x_2, x_3, x_4)$ and let the coefficient vectors in (1) be denoted by L_1, L_2, L_3 and those in (2) be denoted by M_1, M_2, M_3 . Thus they are as follows :

$$L_1 = \frac{1}{4} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, L_2 = \begin{bmatrix} 2 \\ -3 \\ 1 \\ 1 \end{bmatrix}, L_3 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix} \Rightarrow L_1' L_2 = \frac{1}{4}, L_1' L_3 = 0, L_2' L_3 = 5.$$

This means that u_1 and w_1 are independently distributed and that the other pairs are not independently distributed. The coefficient vectors in (2) are

$$M_1 = \frac{1}{4} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, M_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}, M_3 = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} \Rightarrow M_1' M_2 = 0, M_1' M_3 = 0, M_2' M_3 = 0.$$

This means that all the pairs are independently distributed, that is, u_2 , v_2 and w_2 are mutually independently distributed.

We can extend Theorem 2.3.1 to sets of linear functions. Let $Y_1 = AX$ and $Y_2 = BX$ where A of dimension $m_1 \times n$, $m_1 \leq n$ and B of dimension $m_2 \times n$, $m_2 \leq n$ are constant matrices and $X' = (x_1, \dots, x_n)$ where the x_j 's are iid $N_1(\mu_1, \sigma_1^2)$. Let the parameter vectors T_1 and T_2 be of dimensions $m_1 \times 1$ and $m_2 \times 1$, respectively. Then, the mgf of Y_1 is $M_{Y_1}(T_1) = E[e^{T_1' Y_1}] = E[e^{T_1' A_1 X}]$, which can be evaluated by integration over the joint density of x_1, \dots, x_n , individually, or over the vector $X' = (x_1, \dots, x_n)$ with $E[X'] = [\mu_1, \mu_1, \dots, \mu_1] = \mu_1 [1, 1, \dots, 1] = \mu_1 J'$, $J' = [1, \dots, 1] \Rightarrow E[Y_1] = \mu_1 A_1 J$. The mgf of Y_1 is then

$$M_{Y_1}(T_1) = E[e^{\mu_1 T_1' A_1 J + T_1' A_1 [X - E(X)]}] = E[e^{\mu_1 T_1' A_1 J + T_1' A_1 Z}], \quad Z = X - E(X), \quad (i)$$

and the exponent in the expected value, not containing μ_1 , simplifies to

$$-\frac{1}{2\sigma_1^2} \{Z'Z - 2\sigma_1^2 T_1' A_1 Z\} = -\frac{1}{2\sigma_1^2} \{(Z' - \sigma_1 T_1' A_1)(Z - \sigma_1 A_1' T_1) - \sigma_1^2 T_1' A_1 A_1' T_1\}.$$

Integration over Z or individually over the elements of Z , that is, z_1, \dots, z_n , yields 1 since the total probability is 1, which leaves the factor not containing Z . Thus,

$$M_{Y_1}(T_1) = e^{\mu_1 T_1' A_1 J + \frac{1}{2} T_1' A_1 A_1' T_1}, \quad (ii)$$

and similarly,

$$M_{Y_2}(T_2) = e^{\mu_1 T_2' A_2 J + \frac{1}{2} T_2' A_2 A_2' T_2}. \quad (iii)$$

The joint mgf of Y_1 and Y_2 is then

$$\begin{aligned} M_{Y_1, Y_2}(T_1, T_2) &= e^{\mu_1(T_1' A_1 J + T_2' A_2 J) + \frac{1}{2}(T_1' A_1 + T_2' A_2)(T_1' A_1 + T_2' A_2)'} \\ &= M_{Y_1}(T_1) M_{Y_2}(T_2) e^{T_1' A_1 A_2' T_2}. \end{aligned} \quad (iv)$$

Accordingly, Y_1 and Y_2 will be independent if and only if $A_1 A_2' = O \Rightarrow A_2 A_1' = O$ since T_1 and T_2 are arbitrary parameter vectors, the two null matrices having different orders. Then, we have

Theorem 2.3.2. Let $Y_1 = A_1 X$ and $Y_2 = A_2 X$, with $X' = (x_1, \dots, x_n)$, the x_j 's being iid $N_1(\mu_1, \sigma_1^2)$, $j = 1, \dots, n$, be two sets of linear forms where A_1 is $m_1 \times n$ and A_2 is $m_2 \times n$, $m_1 \leq n$, $m_2 \leq n$, are constant matrices. Then, Y_1 and Y_2 are independently distributed if and only if $A_1 A_2' = O$ or $A_2 A_1' = O$.

Example 2.3.2. Consider a simple random sample of size 4 from a real scalar normal population $N_1(\mu_1 = 0, \sigma_1^2 = 4)$. Let $X' = (x_1, x_2, x_3, x_4)$. Verify whether the sets of linear functions $U = A_1 X$, $V = A_2 X$, $W = A_3 X$ are pairwise independent, where

$$A_1 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & -1 & 1 & 3 \\ 1 & 2 & -1 & -2 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 1 & -1 & -1 & 1 \\ -1 & -1 & 1 & 1 \end{bmatrix}.$$

Solution 2.3.2. Taking the products, we have $A_1 A_2' \neq O$, $A_1 A_3' = O$, $A_2 A_3' \neq O$. Hence, the pair U and W are independently distributed and other pairs are not.

We can apply Theorems 2.3.1 and 2.3.2 to prove several results involving sample statistics. For instance, let x_1, \dots, x_n be iid $N_1(\mu_1, \sigma_1^2)$ or a simple random sample of size n from a real $N_1(\mu_1, \sigma_1^2)$ and $\bar{x} = \frac{1}{n}(x_1 + \dots + x_n)$. Consider the vectors

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \mu = \begin{bmatrix} \mu_1 \\ \mu_1 \\ \vdots \\ \mu_1 \end{bmatrix}, \quad \bar{X} = \begin{bmatrix} \bar{x} \\ \bar{x} \\ \vdots \\ \bar{x} \end{bmatrix}.$$

Note that when the x_j 's are iid $N_1(\mu_1, \sigma_1^2)$, $x_j - \mu_1 \sim N_1(0, \sigma_1^2)$, and that since $X - \bar{X} = (X - \mu) - (\bar{X} - \mu)$, we may take x_j 's as coming from $N_1(0, \sigma_1^2)$ for all operations involving (X, \bar{X}) . Moreover, $\bar{x} = \frac{1}{n} J' X$, $J' = (1, 1, \dots, 1)$ where J is a $n \times 1$ vector of unities. Then,

$$X - \bar{X} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} - \begin{bmatrix} \bar{x} \\ \vdots \\ \bar{x} \end{bmatrix} = \begin{bmatrix} x_1 - \frac{1}{n}J'X \\ x_2 - \frac{1}{n}J'X \\ \vdots \\ x_n - \frac{1}{n}J'X \end{bmatrix} = (I - \frac{1}{n}JJ')X, \quad (i)$$

and on letting $A = \frac{1}{n}JJ'$, we have

$$A = A^2, \quad I - A = (I - A)^2, \quad A(I - A) = O. \quad (ii)$$

Also note that

$$(X - \bar{X})'(X - \bar{X}) = \sum_{j=1}^n (x_j - \bar{x})^2 \text{ and } s^2 = \frac{1}{n} \sum_{j=1}^n (x_j - \bar{x})^2 \quad (iii)$$

where s^2 is the sample variance and $\frac{1}{n}J'X = \bar{x}$ is the sample mean. Now, observe that in light of Theorem 2.3.2, $Y_1 = (I - A)X$ and $Y_2 = AX$ are independently distributed, which implies that $X - \bar{X}$ and \bar{X} are independently distributed. But \bar{X} contains only $\bar{x} = \frac{1}{n}(x_1 + \dots + x_n)$ and hence $X - \bar{X}$ and \bar{x} are independently distributed. We now will make use of the following result: If w_1 and w_2 are independently distributed real scalar random variables, then the pairs (w_1, w_2^2) , (w_1^2, w_2) , (w_1^2, w_2^2) are independently distributed when w_1 and w_2 are real scalar random variables; the converses need not be true. For example, w_1^2 and w_2^2 being independently distributed need not imply the independence of w_1 and w_2 . If w_1 and w_2 are real vectors or matrices and if w_1 and w_2 are independently distributed then the following pairs are also independently distributed wherever the quantities are defined: $(w_1, w_2 w_2')$, $(w_1, w_2' w_2)$, $(w_1 w_1', w_2)$, $(w_1' w_1, w_2)$, $(w_1' w_1, w_2' w_2)$. It then follows from (iii) that \bar{x} and $(X - \bar{X})'(X - \bar{X}) = \sum_{j=1}^n (x_j - \bar{x})^2$ are independently distributed. Hence, the following result:

Theorem 2.3.3. *Let x_1, \dots, x_n be iid $N_1(\mu_1, \sigma_1^2)$ or a simple random sample of size n from a univariate real normal population $N_1(\mu_1, \sigma_1^2)$. Let $\bar{x} = \frac{1}{n}(x_1 + \dots + x_n)$ be the sample mean and $s^2 = \frac{1}{n} \sum_{j=1}^n (x_j - \bar{x})^2$ be the sample variance. Then \bar{x} and s^2 are independently distributed.*

This result has several corollaries. When x_1, \dots, x_n are iid $N_1(\mu_1, \sigma_1^2)$, then the sample sum of products, which is also referred to as the corrected sample sum of products (corrected in the sense that \bar{x} is subtracted), is given by

$$\sum_{j=1}^n (x_j - \bar{x})^2 = X'(I - A)X, \quad A = \frac{1}{n} \begin{bmatrix} 1 & 1 & \dots & 1 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & \dots & 1 & 1 \end{bmatrix}$$

where both A and $I - A$ are idempotent. In this case, $\text{tr}(A) = \frac{1}{n}(1 + \dots + 1) = 1$ and $\text{tr}(I - A) = n - 1$ and hence, the ranks of A and $I - A$ are 1 and $n - 1$, respectively. When a matrix is idempotent, its eigenvalues are either zero or one, the number of ones corresponding to its rank. As has already been pointed out, when X and \bar{X} are involved, it can be equivalently assumed that the sample is coming from a $N_1(0, \sigma_1^2)$ population. Hence

$$\frac{ns^2}{\sigma_1^2} = \frac{1}{\sigma_1^2} \sum_{j=1}^n (x_j - \bar{x})^2 \sim \chi_{n-1}^2 \quad (2.3.2)$$

is a real chisquare with $n - 1$ (the rank of the idempotent matrix of the quadratic form) degrees of freedom as per Theorem 2.2.1. Observe that when the sample comes from a real $N_1(\mu_1, \sigma_1^2)$ distribution, we have $\bar{x} \sim N_1(\mu_1, \frac{\sigma_1^2}{n})$ so that $z = \frac{\sqrt{n}(\bar{x} - \mu_1)}{\sigma_1} \sim N_1(0, 1)$ or z is a real standard normal, and that $\frac{(n-1)s_1^2}{\sigma_1^2} \sim \chi_{n-1}^2$ where $s_1^2 = \frac{\sum_{j=1}^n (x_j - \bar{x})^2}{n-1}$. Recall that \bar{x} and s^2 are independently distributed. Hence, the ratio

$$\frac{z}{s_1/\sigma_1} \sim t_{n-1}$$

has a real Student-t distribution with $n - 1$ degrees of freedom, where $z = \frac{\sqrt{n}(\bar{x} - \mu_1)}{\sigma_1}$ and $\frac{z}{s_1/\sigma_1} = \frac{\sqrt{n}(\bar{x} - \mu_1)}{s_1}$. Hence, we have the following result:

Theorem 2.3.4. *Let x_1, \dots, x_n be iid $N_1(\mu_1, \sigma_1^2)$. Let $\bar{x} = \frac{1}{n}(x_1 + \dots + x_n)$ and $s_1^2 = \frac{\sum_{j=1}^n (x_j - \bar{x})^2}{n-1}$. Then,*

$$\frac{\sqrt{n}(\bar{x} - \mu_1)}{s_1} \sim t_{n-1} \quad (2.3.3)$$

where t_{n-1} is a real Student-t with $n - 1$ degrees of freedom.

It should also be noted that when x_1, \dots, x_n are iid $N_1(\mu_1, \sigma_1^2)$, then

$$\frac{\sum_{j=1}^n (x_j - \mu_1)^2}{\sigma_1^2} \sim \chi_n^2, \quad \frac{\sum_{j=1}^n (x_j - \bar{x})^2}{\sigma_1^2} \sim \chi_{n-1}^2, \quad \frac{\sqrt{n}(\bar{x} - \mu_1)^2}{\sigma_1^2} \sim \chi_1^2,$$

wherefrom the following decomposition is obtained:

$$\frac{1}{\sigma_1^2} \sum_{j=1}^n (x_j - \mu_1)^2 = \frac{1}{\sigma_1^2} \left[\sum_{j=1}^n (x_j - \bar{x})^2 + n(\bar{x} - \mu_1)^2 \right] \Rightarrow \chi_n^2 = \chi_{n-1}^2 + \chi_1^2, \quad (2.3.4)$$

the two chisquare random variables on the right-hand side of the last equation being independently distributed.

2.3a. Simple Random Samples from a Complex Gaussian Population

The definition of a simple random sample from any population remains the same as in the real case. A set of complex scalar random variables $\tilde{x}_1, \dots, \tilde{x}_n$, which are iid as $\tilde{N}_1(\tilde{\mu}_1, \sigma_1^2)$ is called a simple random sample from this complex Gaussian population. Let \tilde{X} be the $n \times 1$ vector whose components are these sample variables, $\bar{\tilde{x}} = \frac{1}{n}(\tilde{x}_1 + \dots + \tilde{x}_n)$ denote the sample average, and $\bar{\tilde{X}}' = (\bar{\tilde{x}}, \dots, \bar{\tilde{x}})$ be the $1 \times n$ vector of sample means; then \tilde{X} , $\bar{\tilde{X}}$ and the sample sum of products matrix \tilde{s} are respectively,

$$\tilde{X} = \begin{bmatrix} \tilde{x}_1 \\ \vdots \\ \tilde{x}_n \end{bmatrix}, \quad \bar{\tilde{X}} = \begin{bmatrix} \bar{\tilde{x}} \\ \vdots \\ \bar{\tilde{x}} \end{bmatrix} \quad \text{and} \quad \tilde{s} = (\tilde{X} - \bar{\tilde{X}})^*(\tilde{X} - \bar{\tilde{X}}).$$

These quantities can be simplified as follows with the help of the vector of unities $J' = (1, 1, \dots, 1)$: $\bar{\tilde{x}} = \frac{1}{n}J'\tilde{X}$, $\tilde{X} - \bar{\tilde{X}} = [I - \frac{1}{n}JJ']\tilde{X}$, $\bar{\tilde{X}} = \frac{1}{n}JJ'\tilde{X}$, $\tilde{s} = \tilde{X}^*[I - \frac{1}{n}JJ']\tilde{X}$. Consider the Hermitian form

$$\tilde{X}^*[I - \frac{1}{n}JJ']\tilde{X} = \sum_{j=1}^n (\tilde{x}_j - \bar{\tilde{x}})^*(\tilde{x}_j - \bar{\tilde{x}}) = n\tilde{s}^2$$

where \tilde{s}^2 is the sample variance in the complex scalar case, given a simple random sample of size n .

Consider the linear forms $\tilde{u}_1 = L_1^*\tilde{X} = \bar{a}_1\tilde{x}_1 + \dots + \bar{a}_n\tilde{x}_n$ and $\tilde{u}_2 = L_2^*\tilde{X} = \bar{b}_1\tilde{x}_1 + \dots + \bar{b}_n\tilde{x}_n$ where the a_j 's and b_j 's are scalar constants that may be real or complex, \bar{a}_j and \bar{b}_j denoting the complex conjugates of a_j and b_j , respectively.

$$E[\tilde{X}]' = (\tilde{\mu}_1)[1, 1, \dots, 1] = (\tilde{\mu}_1)J', \quad J' = [1, \dots, 1],$$

since the \tilde{x}_j 's are iid $\tilde{N}_1(\tilde{\mu}_1, \sigma_1^2)$, $j = 1, \dots, n$. The mgf of \tilde{u}_1 and \tilde{u}_2 , denoted by $M_{\tilde{u}_j}(\tilde{t}_j)$, $j = 1, 2$ and the joint mgf of \tilde{u}_1 and \tilde{u}_2 , denoted by $M_{\tilde{u}_1, \tilde{u}_2}(\tilde{t}_1, \tilde{t}_2)$ are the following, where $\Re(\cdot)$ denotes the real part of (\cdot) :

$$\begin{aligned} M_{\tilde{u}_1}(\tilde{t}_1) &= E[e^{\Re(\tilde{t}_1^* \tilde{u}_1)}] = E[e^{\Re(\tilde{t}_1^* L_1^* \tilde{X})}] \\ &= e^{\Re(\tilde{\mu}_1 \tilde{t}_1^* L_1^* J)} E[e^{\Re(\tilde{t}_1^* L_1^* (\tilde{X} - E(\tilde{X})))}] = e^{\Re(\tilde{\mu}_1 \tilde{t}_1^* L_1^* J)} e^{\frac{\sigma_1^2}{4} \tilde{t}_1^* L_1^* L_1 \tilde{t}_1} \end{aligned} \quad (i)$$

$$M_{\tilde{u}_2}(\tilde{t}_2) = e^{\Re(\tilde{\mu}_1 \tilde{t}_2^* L_2^* J) + \frac{\sigma_1^2}{4} \tilde{t}_2^* L_2^* L_2 \tilde{t}_2} \quad (ii)$$

$$M_{\tilde{u}_1, \tilde{u}_2}(\tilde{t}_1, \tilde{t}_2) = M_{\tilde{u}_1}(\tilde{t}_1) M_{\tilde{u}_2}(\tilde{t}_2) e^{\frac{\sigma_1^2}{2} \tilde{t}_1^* L_1^* L_2 \tilde{t}_2}. \quad (iii)$$

Consequently, \tilde{u}_1 and \tilde{u}_2 are independently distributed if and only if the exponential part is 1 or equivalently $\tilde{t}_1^* L_1^* L_2 \tilde{t}_2 = 0$. Since \tilde{t}_1 and \tilde{t}_2 are arbitrary, this means $L_1^* L_2 = 0 \Rightarrow L_2^* L_1 = 0$. Then we have the following result:

Theorem 2.3a.1. *Let $\tilde{x}_1, \dots, \tilde{x}_n$ be a simple random sample of size n from a univariate complex Gaussian population $\tilde{N}_1(\tilde{\mu}_1, \sigma_1^2)$. Consider the linear forms $\tilde{u}_1 = L_1^* \tilde{X}$ and $\tilde{u}_2 = L_2^* \tilde{X}$ where L_1, L_2 and \tilde{X} are the previously defined $n \times 1$ vectors, and a star denotes the conjugate transpose. Then, \tilde{u}_1 and \tilde{u}_2 are independently distributed if and only if $L_1^* L_2 = 0$.*

Example 2.3a.1. Let \tilde{x}_j , $j = 1, 2, 3, 4$ be iid univariate complex normal $\tilde{N}_1(\tilde{\mu}_1, \sigma_1^2)$. Consider the linear forms

$$\begin{aligned} \tilde{u}_1 &= L_1^* \tilde{X} = (1 + i)\tilde{x}_1 + 2i\tilde{x}_2 - (1 - i)\tilde{x}_3 + 2\tilde{x}_4 \\ \tilde{u}_2 &= L_2^* \tilde{X} = (1 + i)\tilde{x}_1 + (2 + 3i)\tilde{x}_2 - (1 - i)\tilde{x}_3 - i\tilde{x}_4 \\ \tilde{u}_3 &= L_3^* \tilde{X} = -(1 + i)\tilde{x}_1 + i\tilde{x}_2 + (1 - i)\tilde{x}_3 + \tilde{x}_4. \end{aligned}$$

Verify whether the three linear forms are pairwise independent.

Solution 2.3a.1. With the usual notations, the coefficient vectors are as follows:

$$\begin{aligned} L_1^* &= [1 + i, 2i, -1 + i, 2] \Rightarrow L_1' = [1 - i, -2i, -1 - i, 2] \\ L_2^* &= [1 + i, 2 + 3i, 1 - i, -i] \Rightarrow L_2' = [1 - i, 2 - 3i, 1 + i, i] \\ L_3^* &= [-(1 + i), i, 1 - i, 1] \Rightarrow L_3' = [-(1 - i), -i, 1 + i, 1]. \end{aligned}$$

Taking the products we have $L_1^* L_2 = 6 + 6i \neq 0$, $L_1^* L_3 = 0$, $L_2^* L_3 = 3 - 3i \neq 0$. Hence, only \tilde{u}_1 and \tilde{u}_3 are independently distributed.

We can extend the result stated in Theorem 2.3a.1 to sets of linear forms. Let $\tilde{U}_1 = A_1 \tilde{X}$ and $\tilde{U}_2 = A_2 \tilde{X}$ where A_1 and A_2 are constant matrices that may or may not be in

the complex domain, A_1 is $m_1 \times n$ and A_2 is $m_2 \times n$, with $m_1 \leq n$, $m_2 \leq n$. As was previously the case, $\tilde{X}' = (\tilde{x}_1, \dots, \tilde{x}_n)$, \tilde{x}_j , $j = 1, \dots, n$, are iid $\tilde{N}_1(\tilde{\mu}_1, \sigma_1^2)$. Let \tilde{T}_1 and \tilde{T}_2 be parameter vectors of orders $m_1 \times 1$ and $m_2 \times 1$, respectively. Then, on following the steps leading to (iii), the mgf of \tilde{U}_1 and \tilde{U}_2 and their joint mgf are obtained as follows:

$$M_{\tilde{U}_1}(\tilde{T}_1) = E[e^{\Re(\tilde{T}_1^* A_1 \tilde{X})}] = e^{\Re(\tilde{\mu}_1 \tilde{T}_1^* A_1 J) + \frac{\sigma_1^2}{4} \tilde{T}_1^* A_1 A_1^* \tilde{T}_1} \quad (iv)$$

$$M_{\tilde{U}_2}(\tilde{T}_2) = e^{\Re(\tilde{\mu}_1 \tilde{T}_2^* A_2 J) + \frac{\sigma_1^2}{4} \tilde{T}_2^* A_2 A_2^* \tilde{T}_2} \quad (v)$$

$$M_{\tilde{U}_1, \tilde{U}_2}(\tilde{T}_1, \tilde{T}_2) = M_{\tilde{U}_1}(\tilde{T}_1) M_{\tilde{U}_2}(\tilde{T}_2) e^{\frac{\sigma_1^2}{2} \tilde{T}_1^* A_1 A_2^* \tilde{T}_2}. \quad (vi)$$

Since \tilde{T}_1 and \tilde{T}_2 are arbitrary, the exponential part in (vi) is 1 if and only if $A_1 A_2^* = O$ or $A_2 A_1^* = O$, the two null matrices having different orders. Then, we have:

Theorem 2.3a.2. *Let the \tilde{x}_j 's and \tilde{X} be as defined in Theorem 2.3a.1. Let A_1 be a $m_1 \times n$ constant matrix and A_2 be a $m_2 \times n$ constant matrix, $m_1 \leq n$, $m_2 \leq n$, and the constant matrices may or may not be in the complex domain. Consider the general linear forms $\tilde{U}_1 = A_1 \tilde{X}$ and $\tilde{U}_2 = A_2 \tilde{X}$. Then \tilde{U}_1 and \tilde{U}_2 are independently distributed if and only if $A_1 A_2^* = O$ or, equivalently, $A_2 A_1^* = O$.*

Example 2.3a.2. Let \tilde{x}_j , $j = 1, 2, 3, 4$, be iid univariate complex Gaussian $\tilde{N}_1(\tilde{\mu}_1, \sigma_1^2)$. Consider the following sets of linear forms $\tilde{U}_1 = A_1 \tilde{X}$, $\tilde{U}_2 = A_2 \tilde{X}$, $\tilde{U}_3 = A_3 \tilde{X}$ with $X' = (x_1, x_2, x_3, x_4)$, where

$$A_1 = \begin{bmatrix} 2+3i & 2+3i & 2+3i & 2+3i \\ 1+i & -(1+i) & -(1+i) & 1+i \end{bmatrix}$$

$$A_2 = \begin{bmatrix} 2i & -2i & 2i & -2i \\ 1-i & 1-i & -1+i & -1+i \\ -(1+2i) & 1+2i & -(1+2i) & 1+2i \end{bmatrix}, \quad A_3 = \begin{bmatrix} 1+2i & 0 & 1-2i & -2 \\ -2 & -1+i & -1+i & 2i \end{bmatrix}.$$

Verify whether the pairs in $\tilde{U}_1, \tilde{U}_2, \tilde{U}_3$ are independently distributed.

Solution 2.3a.2. Since the products $A_1 A_2^* = O$, $A_1 A_3^* \neq O$, $A_2 A_3^* \neq O$, only \tilde{U}_1 and \tilde{U}_2 are independently distributed.

As a corollary of Theorem 2.3a.2, one has that the sample mean $\bar{\tilde{x}}$ and the sample sum of products \tilde{s} are also independently distributed in the complex Gaussian case, a result parallel to the corresponding one in the real case. This can be seen by taking $A_1 = \frac{1}{n} J J'$ and $A_2 = I - \frac{1}{n} J J'$. Then, since $A_1 = A_1^2$, $A_2 = A_2^2$, and $A_1 A_2 = O$, we have

$\frac{1}{\sigma_1^2} \tilde{X}^* A_1 \tilde{X} \sim \tilde{\chi}_1^2$ for $\tilde{\mu}_1 = 0$ and $\frac{1}{\sigma_1^2} \tilde{X}^* A_2 \tilde{X} \sim \tilde{\chi}_{n-1}^2$, and both of these chisquares in the complex domain are independently distributed. Then,

$$\frac{1}{\sigma_1^2} n\tilde{s} = \frac{1}{\sigma_1^2} \sum_{j=1}^n (\tilde{x}_j - \bar{\tilde{x}})^* (\tilde{x}_j - \bar{\tilde{x}}) \sim \tilde{\chi}_{n-1}^2. \quad (2.3a.1)$$

The Student-t with $n - 1$ degrees of freedom can be defined in terms of the standardized sample mean and sample variance in the complex case.

2.3.1. Noncentral chisquare having n degrees of freedom in the real domain

Let $x_j \sim N_1(\mu_j, \sigma_j^2)$, $j = 1, \dots, n$ and the x_j 's be independently distributed. Then, $\frac{x_j - \mu_j}{\sigma_j} \sim N_1(0, 1)$ and $\sum_{j=1}^n \frac{(x_j - \mu_j)^2}{\sigma_j^2} \sim \chi_n^2$ where χ_n^2 is a real chisquare with n degrees of freedom. Then, when at least one of the μ_j 's is nonzero, $\sum_{j=1}^n \frac{x_j^2}{\sigma_j^2}$ is referred to as a *real non-central chisquare with n degrees of freedom and non-centrality parameter λ , which is denoted $\chi_n^2(\lambda)$* , where

$$\lambda = \frac{1}{2} \mu' \Sigma^{-1} \mu, \quad \mu = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_n \end{bmatrix}, \quad \text{and} \quad \Sigma = \begin{bmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_n^2 \end{bmatrix}.$$

Let $u = \sum_{j=1}^n \frac{x_j^2}{\sigma_j^2}$. In order to derive the distribution of u , let us determine its mgf. Since u is a function of the x_j 's where $x_j \sim N_1(\mu_j, \sigma_j^2)$, $j = 1, \dots, n$, we can integrate out over the joint density of the x_j 's. Then, with t as the mgf parameter,

$$\begin{aligned} M_u(t) &= E[e^{tu}] \\ &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} e^{t \sum_{j=1}^n \frac{x_j^2}{\sigma_j^2} - \frac{1}{2} \sum_{j=1}^n \frac{(x_j - \mu_j)^2}{\sigma_j^2}} dx_1 \wedge \dots \wedge dx_n. \end{aligned}$$

The exponent, excluding $-\frac{1}{2}$ can be simplified as follows:

$$-2t \sum_{j=1}^n \frac{x_j^2}{\sigma_j^2} + \sum_{j=1}^n \frac{(x_j - \mu_j)^2}{\sigma_j^2} = (1 - 2t) \sum_{j=1}^n \frac{x_j^2}{\sigma_j^2} - 2 \sum_{j=1}^n \frac{\mu_j x_j}{\sigma_j^2} + \sum_{j=1}^n \frac{\mu_j^2}{\sigma_j^2}.$$

Let $y_j = \sqrt{(1-2t)}x_j$. Then, $(1-2t)^{-\frac{n}{2}}dy_1 \wedge \dots \wedge dy_n = dx_1 \wedge \dots \wedge dx_n$, and

$$\begin{aligned} (1-2t) \sum_{j=1}^n \left(\frac{x_j^2}{\sigma_j^2} \right) - 2 \sum_{j=1}^n \left(\frac{\mu_j x_j}{\sigma_j^2} \right) &= \sum_{j=1}^n \frac{y_j^2}{\sigma_j^2} - 2 \sum_{j=1}^n \frac{\mu_j y_j}{\sigma_j^2 \sqrt{(1-2t)}} \\ &\quad + \sum_{j=1}^n \left(\frac{\mu_j}{\sigma_j \sqrt{(1-2t)}} \right)^2 - \sum_{j=1}^n \frac{\mu_j^2}{\sigma_j^2 (1-2t)} \\ &= \sum_{j=1}^n \left(\frac{y_j - \frac{\mu_j}{\sqrt{(1-2t)}}}{\sigma_j} \right)^2 - \sum_{j=1}^n \frac{\mu_j^2}{\sigma_j^2 (1-2t)}. \end{aligned}$$

But

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} e^{-\sum_{j=1}^n \frac{1}{2\sigma_j^2} \left(y_j - \frac{\mu_j}{\sqrt{(1-2t)}} \right)^2} dy_1 \wedge \dots \wedge dy_n = 1.$$

Hence, for $\lambda = \frac{1}{2} \sum_{j=1}^n \frac{\mu_j^2}{\sigma_j^2} = \frac{1}{2} \mu' \Sigma^{-1} \mu$,

$$\begin{aligned} M_u(t) &= \frac{1}{(1-2t)^{\frac{n}{2}}} [e^{-\lambda + \frac{\lambda}{(1-2t)}}] \\ &= \sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!} \frac{1}{(1-2t)^{\frac{n}{2}+k}}. \end{aligned} \tag{2.3.5}$$

However, $(1-2t)^{-\left(\frac{n}{2}+k\right)}$ is the mgf of a real scalar gamma with parameters $(\alpha = \frac{n}{2}+k, \beta = 2)$ or a real chisquare with $n+2k$ degrees of freedom or χ_{n+2k}^2 . Hence, the density of a non-central chisquare with n degrees of freedom and non-centrality parameter λ , denoted by $g_{u,\lambda}(u)$, is obtained by term by term inversion as follow:

$$g_{u,\lambda}(u) = \sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!} \frac{u^{\frac{n+2k}{2}-1} e^{-\frac{u}{2}}}{2^{\frac{n+2k}{2}} \Gamma\left(\frac{n}{2}+k\right)} \tag{2.3.6}$$

$$= \frac{u^{\frac{n}{2}-1} e^{-\frac{u}{2}}}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} \sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!} \frac{u^k}{\left(\frac{n}{2}\right)_k} \tag{2.3.7}$$

where $\left(\frac{n}{2}\right)_k$ is the Pochhammer symbol given by

$$(a)_k = a(a+1) \dots (a+k-1), \quad a \neq 0, \quad (a)_0 \text{ being equal to } 1, \tag{2.3.8}$$

and, in general, $\Gamma(\alpha + k) = \Gamma(\alpha)(\alpha)_k$ for $k = 1, 2, \dots$, whenever the gamma functions are defined. Hence, provided $(1 - 2t) > 0$, (2.3.6) can be looked upon as a weighted sum of chisquare densities whose weights are Poisson distributed, that is, (2.3.6) is a Poisson mixture of chisquare densities. As well, we can view (2.3.7) as a chisquare density having n degrees of freedom appended with a Bessel series. In general, a Bessel series is of the form

$${}_0F_1(; b; x) = \sum_{k=0}^{\infty} \frac{1}{(b)_k} \frac{x^k}{k!}, \quad b \neq 0, -1, -2, \dots, \quad (2.3.9)$$

which is convergent for all x .

2.3.1.1. Mean value and variance, real central and non-central chisquare

The mgf of a real χ_ν^2 is $(1 - 2t)^{-\frac{\nu}{2}}$, $1 - 2t > 0$. Thus,

$$E[\chi_\nu^2] = \frac{d}{dt}(1 - 2t)^{-\frac{\nu}{2}}|_{t=0} = \left(-\frac{\nu}{2}\right)(-2)(1 - 2t)^{-\frac{\nu}{2}-1}|_{t=0} = \nu$$

$$E[\chi_\nu^2]^2 = \frac{d^2}{dt^2}(1 - 2t)^{-\frac{\nu}{2}}|_{t=0} = \left(-\frac{\nu}{2}\right)(-2)\left(-\frac{\nu}{2} - 1\right)(-2) = \nu(\nu + 2).$$

That is,

$$E[\chi_\nu^2] = \nu \text{ and } \text{Var}(\chi_\nu^2) = \nu(\nu + 2) - \nu^2 = 2\nu. \quad (2.3.10)$$

What are then the mean and the variance of a real non-central $\chi_\nu^2(\lambda)$? They can be derived either from the mgf or from the density. Making use of the density, we have

$$E[\chi_\nu^2(\lambda)] = \sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!} \int_0^{\infty} u \frac{u^{\frac{\nu}{2}+k-1} e^{-\frac{u}{2}}}{2^{\frac{\nu}{2}+k} \Gamma(\frac{\nu}{2} + k)} du,$$

the integral part being equal to

$$\frac{\Gamma(\frac{\nu}{2} + k + 1) 2^{\frac{\nu}{2}+k+1}}{\Gamma(\frac{\nu}{2} + k) 2^{\frac{\nu}{2}+k}} = 2\left(\frac{\nu}{2} + k\right) = \nu + 2k.$$

Now, the remaining summation over k can be looked upon as the expected value of $\nu + 2k$ in a Poisson distribution. In this case, we can write the expected values as the expected value of a conditional expectation: $E[u] = E[E(u|k)]$, $u = \chi_\nu^2(\lambda)$. Thus,

$$E[\chi_\nu^2(\lambda)] = \nu + 2 \sum_{k=0}^{\infty} k \frac{\lambda^k e^{-\lambda}}{k!} = \nu + 2E[k] = \nu + 2\lambda.$$

Moreover,

$$E[\chi_v^2(\lambda)]^2 = \sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!} \int_0^{\infty} u^2 \frac{u^{\frac{v}{2}+k-1}}{2^{\frac{v}{2}+k} \Gamma(\frac{v}{2}+k)} du,$$

the integral part being

$$\begin{aligned} \frac{\Gamma(\frac{v}{2}+k+2)}{\Gamma(\frac{v}{2}+k)} \frac{2^{\frac{v}{2}+k+2}}{2^{\frac{v}{2}+k}} &= 2^2 \left(\frac{v}{2}+k+1\right) \left(\frac{v}{2}+k\right) \\ &= (v+2k+2)(v+2k) = v^2 + 2vk + 2v(k+1) + 4k(k+1). \end{aligned}$$

Since $E[k] = \lambda$, $E[k^2] = \lambda^2 + \lambda$ for a Poisson distribution,

$$E[\chi_v^2(\lambda)]^2 = v^2 + 2v + 4v\lambda + 4(\lambda^2 + 2\lambda).$$

Thus,

$$\begin{aligned} \text{Var}(\chi_v^2(\lambda)) &= E[\chi_v^2(\lambda)]^2 - [E(\chi_v^2(\lambda))]^2 \\ &= v^2 + 2v + 4v\lambda + 4(\lambda^2 + 2\lambda) - (v + 2\lambda)^2 \\ &= 2v + 8\lambda. \end{aligned}$$

To summarize,

$$E[\chi_v^2(\lambda)] = v + 2\lambda \text{ and } \text{Var}(\chi_v^2(\lambda)) = 2v + 8\lambda. \quad (2.3.11)$$

Example 2.3.3. Let $x_1 \sim N_1(-1, 2)$, $x_2 \sim N_1(1, 3)$ and $x_3 \sim N_1(-2, 2)$ be independently distributed and $u = \frac{x_1^2}{2} + \frac{x_2^2}{3} + \frac{x_3^2}{2}$. Provide explicit expressions for the density of u , $E[u]$ and $\text{Var}(u)$.

Solution 2.3.3. This u has a noncentral chisquare distribution with non-centrality parameter λ where

$$\lambda = \frac{1}{2} \left[\frac{\mu_1^2}{\sigma_1^2} + \frac{\mu_2^2}{\sigma_2^2} + \frac{\mu_3^2}{\sigma_3^2} \right] = \frac{1}{2} \left[\frac{(-1)^2}{2} + \frac{(1)^2}{3} + \frac{(-2)^2}{2} \right] = \frac{17}{12},$$

and the number of degrees of freedom is $n = 3 = v$. Thus, $u \sim \chi_3^2(\lambda)$ or a real noncentral chisquare with $v = 3$ degrees of freedom and non-centrality parameter $\frac{17}{12}$. Then $E[u] = E[\chi_3^2(\lambda)] = v + 2\lambda = 3 + 2(\frac{17}{12}) = \frac{35}{6}$. $\text{Var}(u) = \text{Var}(\chi_3^2(\lambda)) = 2v + 8\lambda = (2)(3) + 8(\frac{17}{12}) = \frac{52}{3}$. Let the density of u be denoted by $g(u)$. Then

$$\begin{aligned}
 g(u) &= \frac{u^{\frac{n}{2}-1} e^{-\frac{u}{2}}}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} \sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!} \frac{u^k}{(\frac{n}{2})_k} \\
 &= \frac{u^{\frac{1}{2}} e^{-\frac{u}{2}}}{\sqrt{2\pi}} \sum_{k=0}^{\infty} \frac{(17/12)^k e^{-17/12}}{k!} \frac{u^k}{(\frac{3}{2})_k}, \quad 0 \leq u < \infty,
 \end{aligned}$$

and zero elsewhere.

2.3a.1. Noncentral chisquare having n degrees of freedom in the complex domain

Let us now consider independently Gaussian distributed variables in the complex domain. Let the complex scalar variables $\tilde{x}_j \sim \tilde{N}(\tilde{\mu}_j, \sigma_j^2)$, $j = 1, \dots, n$, be independently distributed. Then, we have already established that $\sum_{j=1}^n (\frac{\tilde{x}_j - \tilde{\mu}_j}{\sigma_j})^* (\frac{\tilde{x}_j - \tilde{\mu}_j}{\sigma_j}) \sim \tilde{\chi}_n^2$, which is a chisquare variable having n degrees of freedom in the complex domain. If we let $\tilde{u} = \sum_{j=1}^n \frac{\tilde{x}_j^* \tilde{x}_j}{\sigma_j^2}$, then this u will be said to have a noncentral chisquare distribution with n degrees of freedom and non-centrality parameter λ in the complex domain where \tilde{x}_j^* is only the conjugate since it is a scalar quantity. Since, in this case, \tilde{u} is real, we may associate the mgf of \tilde{u} with a real parameter t . Now, proceeding as in the real case, we obtain the mgf of \tilde{u} , denoted by $M_{\tilde{u}}(t)$, as follows:

$$\begin{aligned}
 M_{\tilde{u}}(t) &= E[e^{t\tilde{u}}] = (1-t)^{-n} e^{-\lambda + \frac{\lambda}{1-t}}, \quad 1-t > 0, \quad \lambda = \sum_{j=1}^n \frac{\tilde{\mu}_j^* \tilde{\mu}_j}{\sigma_j^2} = \tilde{\mu}^* \Sigma^{-1} \tilde{\mu} \\
 &= \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} (1-t)^{-(n+k)}. \tag{2.3a.2}
 \end{aligned}$$

Note that the inverse corresponding to $(1-t)^{-(n+k)}$ is a chisquare density in the complex domain with $n+k$ degrees of freedom, and that part of the density, denoted by $f_1(u)$, is

$$f_1(u) = \frac{1}{\Gamma(n+k)} u^{n+k-1} e^{-u} = \frac{1}{\Gamma(n)(n)_k} u^{n-1} u^k e^{-u}.$$

Thus, the noncentral chisquare density with n degrees of freedom in the complex domain, that is, $u = \tilde{\chi}_n^2(\lambda)$, denoted by $f_{u,\lambda}(u)$, is

$$f_{u,\lambda}(u) = \frac{u^{n-1}}{\Gamma(n)} e^{-u} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} \frac{u^k}{(n)_k}, \tag{2.3a.3}$$

which, referring to Eqs. (2.3.5)–(2.3.9) in connection with a non-central chisquare in the real domain, can also be represented in various ways.

Example 2.3a.3. Let $\tilde{x}_1 \sim \tilde{N}_1(1 + i, 2)$, $\tilde{x}_2 \sim \tilde{N}_1(2 + i, 4)$ and $\tilde{x}_3 \sim \tilde{N}_1(1 - i, 2)$ be independently distributed univariate complex Gaussian random variables and

$$\tilde{u} = \frac{\tilde{x}_1^* \tilde{x}_1}{\sigma_1^2} + \frac{\tilde{x}_2^* \tilde{x}_2}{\sigma_2^2} + \frac{\tilde{x}_3^* \tilde{x}_3}{\sigma_3^2} = \frac{\tilde{x}_1^* \tilde{x}_1}{2} + \frac{\tilde{x}_2^* \tilde{x}_2}{4} + \frac{\tilde{x}_3^* \tilde{x}_3}{2}.$$

Compute $E[\tilde{u}]$ and $\text{Var}(\tilde{u})$ and provide an explicit representation of the density of \tilde{u} .

Solution 2.3a.3. In this case, \tilde{u} has a noncentral chisquare distribution with degrees of freedom $\nu = n = 3$ and non-centrality parameter λ given by

$$\lambda = \frac{\tilde{\mu}_1^* \tilde{\mu}_1}{\sigma_1^2} + \frac{\tilde{\mu}_2^* \tilde{\mu}_2}{\sigma_2^2} + \frac{\tilde{\mu}_3^* \tilde{\mu}_3}{\sigma_3^2} = \frac{[(1)^2 + (1)^2]}{2} + \frac{[(2)^2 + (1)^2]}{4} + \frac{[(1)^2 + (-1)^2]}{2} = 1 + \frac{5}{4} + 1 = \frac{13}{4}.$$

The density, denoted by $g_1(\tilde{u})$, is given in (i). In this case \tilde{u} will be a real gamma with the parameters $(\alpha = n, \beta = 1)$ to which a Poisson series is appended:

$$g_1(u) = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} \frac{u^{n+k-1} e^{-u}}{\Gamma(n+k)}, \quad 0 \leq u < \infty, \quad (i)$$

and zero elsewhere. Then, the expected value of u and the variance of u are available from (i) by direct integration.

$$E[u] = \int_0^{\infty} u g_1(u) du = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} \frac{\Gamma(n+k+1)}{\Gamma(n+k)}.$$

But $\frac{\Gamma(n+k+1)}{\Gamma(n+k)} = n+k$ and the summation over k can be taken as the expected value of $n+k$ in a Poisson distribution. Thus, $E[\tilde{\chi}_n^2(\lambda)] = n + E[k] = n + \lambda$. Now, in the expected value of

$$[\tilde{\chi}_n^2(\lambda)][\tilde{\chi}_n^2(\lambda)]^* = [\tilde{\chi}_n^2(\lambda)]^2 = [u]^2,$$

which is real in this case, the integral part over u gives

$$\frac{\Gamma(n+k+2)}{\Gamma(n+k)} = (n+k+1)(n+k) = n^2 + 2nk + k^2 + n + k$$

with expected value $n^2 + 2n\lambda + n + \lambda + (\lambda^2 + \lambda)$. Hence,

$$\text{Var}(\tilde{\chi}_n^2(\lambda)) = E[u - E(u)][u - E(u)]^* = \text{Var}(u) = n^2 + 2n\lambda + n + \lambda + (\lambda^2 + \lambda) - (n + \lambda)^2,$$

which simplifies to $n + 2\lambda$. Accordingly,

$$E[\tilde{\chi}_n^2(\lambda)] = n + \lambda \text{ and } \text{Var}(\tilde{\chi}_n^2(\lambda)) = n + 2\lambda, \quad (ii)$$

so that

$$E[u] = n + \lambda = 3 + \frac{13}{4} = \frac{25}{4} \text{ and } \text{Var}(u) = n + 2\lambda = 3 + \frac{13}{2} = \frac{19}{2}. \quad (iii)$$

The explicit form of the density is then

$$g_1(u) = \frac{u^2 e^{-u}}{2} \sum_{k=0}^{\infty} \frac{(13/4)^k e^{-13/4} u^k}{k! (3)_k}, \quad 0 \leq u < \infty, \quad (iv)$$

and zero elsewhere.

Exercises 2.3

2.3.1. Let x_1, \dots, x_n be iid variables with common density a real gamma density with the parameters α and β or with the mgf $(1 - \beta t)^{-\alpha}$, $1 - \beta t > 0$, $\alpha > 0$, $\beta > 0$. Let $u_1 = x_1 + \dots + x_n$, $u_2 = \frac{1}{n}(x_1 + \dots + x_n)$, $u_3 = u_2 - \alpha\beta$, $u_4 = \frac{\sqrt{nu_3}}{\beta\sqrt{\alpha}}$. Evaluate the mgfs and thereby the densities of u_1, u_2, u_3, u_4 . Show that they are all gamma densities for all finite n , may be relocated. Show that when $n \rightarrow \infty$, $u_4 \rightarrow N_1(0, 1)$ or u_4 goes to a real standard normal when n goes to infinity.

2.3.2. Let x_1, \dots, x_n be a simple random sample of size n from a real population with mean value μ and variance $\sigma^2 < \infty$, $\sigma > 0$. Then the central limit theorem says that $\frac{\sqrt{n}(\bar{x} - \mu)}{\sigma} \rightarrow N_1(0, 1)$ as $n \rightarrow \infty$, where $\bar{x} = \frac{1}{n}(x_1 + \dots + x_n)$. Translate this statement for

(1): binomial probability function $f_1(x) = \binom{n}{x} p^x (1-p)^{n-x}$, $x = 0, 1, \dots, n$, $0 < p < 1$

and $f_1(x) = 0$ elsewhere; (2): negative binomial probability law $f_2(x) = \binom{x-1}{k-1} p^k (1-p)^{x-k}$, $x = k, k+1, \dots$, $0 < p < 1$ and zero elsewhere; (3): geometric probability law $f_3(x) = p(1-p)^{x-1}$, $x = 1, 2, \dots$, $0 < p < 1$ and $f_3(x) = 0$ elsewhere; (4): Poisson probability law $f_4(x) = \frac{\lambda^x}{x!} e^{-\lambda}$, $x = 0, 1, \dots$, $\lambda > 0$ and $f_4(x) = 0$ elsewhere.

2.3.3. Repeat Exercise 2.3.2 if the population is (1): $g_1(x) = c_1 x^{\gamma-1} e^{-ax^\delta}$, $x \geq 0$, $\delta > 0$, $a > 0$, $\gamma > 0$ and $g_1(x) = 0$ elsewhere; (2): The real pathway model $g_2(x) = c_q x^\gamma [1 - a(1-q)x^\delta]^{1/q}$, $a > 0$, $\delta > 0$, $1 - a(1-q)x^\delta > 0$ and for the cases $q < 1$, $q > 1$, $q \rightarrow 1$, and $g_2(x) = 0$ elsewhere.

2.3.4. Let $x \sim N_1(\mu_1, \sigma_1^2)$, $y \sim N_1(\mu_2, \sigma_2^2)$ be real Gaussian and be independently distributed. Let $x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}$ be simple random samples from x and y respectively. Let $u_1 = \sum_{j=1}^{n_1} (x_j - \mu_1)^2$, $u_2 = \sum_{j=1}^{n_2} (y_j - \mu_2)^2$, $u_3 = 2x_1 - 3x_2 + y_1 - y_2 + 2y_3$,

$$u_4 = \frac{\sum_{j=1}^{n_1} (x_j - \mu_1)^2 / \sigma_1^2}{\sum_{j=1}^{n_2} (y_j - \mu_2)^2 / \sigma_2^2}, u_5 = \frac{\sum_{j=1}^{n_1} (x_j - \bar{x})^2 / \sigma_1^2}{\sum_{j=1}^{n_2} (y_j - \bar{y})^2 / \sigma_2^2}.$$

Compute the densities of u_1, u_2, u_3, u_4, u_5 .

2.3.5. In Exercise 2.3.4 if $\sigma_1^2 = \sigma_2^2 = \sigma^2$ compute the densities of u_3, u_4, u_5 there, and $u_6 = \sum_{j=1}^{n_1} (x_j - \bar{x})^2 + \sum_{j=1}^{n_2} (y_j - \bar{y})^2$ if (1): $n_1 = n_2$, (2): $n_1 \neq n_2$.

2.3.6. For the noncentral chisquare in the complex case, discussed in (2.3a.5) evaluate the mean value and the variance.

2.3.7. For the complex case, starting with the mgf, derive the noncentral chisquare density and show that it agrees with that given in (2.3a.3).

2.3.8. Give the detailed proofs of the independence of linear forms and sets of linear forms in the complex Gaussian case.

2.4. Distributions of Products and Ratios and Connection to Fractional Calculus

Distributions of products and ratios of real scalar random variables are connected to numerous topics including Krätzel integrals and transforms, reaction-rate probability integrals in nuclear reaction-rate theory, the inverse Gaussian distribution, integrals occurring in fractional calculus, Kobayashi integrals and Bayesian structures. Let $x_1 > 0$ and $x_2 > 0$ be real scalar positive random variables that are independently distributed with density functions $f_1(x_1)$ and $f_2(x_2)$, respectively. We respectively denote the product and ratio of these variables by $u_2 = x_1 x_2$ and $u_1 = \frac{x_2}{x_1}$. What are then the densities of u_1 and u_2 ? We first consider the density of the product. Let $u_2 = x_1 x_2$ and $v = x_2$. Then $x_1 = \frac{u_2}{v}$ and $x_2 = v$, $dx_1 \wedge dx_2 = \frac{1}{v} du \wedge dv$. Let the joint density of u_2 and v be denoted by $g(u_2, v)$ and the marginal density of u_2 by $g_2(u_2)$. Then

$$g(u_2, v) = \frac{1}{v} f_1\left(\frac{u_2}{v}\right) f_2(v) \text{ and } g_2(u_2) = \int_v \frac{1}{v} f_2\left(\frac{u_2}{v}\right) f_1(v) dv. \quad (2.4.1)$$

For example, let f_1 and f_2 be generalized gamma densities, in which case

$$f_j(x_j) = \frac{a_j^{\frac{\gamma_j}{\delta_j}}}{\Gamma\left(\frac{\gamma_j}{\delta_j}\right)} x_j^{\gamma_j - 1} e^{-a_j x_j^{\delta_j}}, \quad a_j > 0, \delta_j > 0, \gamma_j > 0, x_j \geq 0 \quad j = 1, 2,$$

and $f_j(x_j) = 0$ elsewhere. Then,

$$\begin{aligned}
 g_2(u_2) &= c \int_0^\infty \left(\frac{1}{v}\right) \left(\frac{u_2}{v}\right)^{\gamma_1-1} v^{\gamma_2-1} \\
 &\quad \times e^{-a_2 v^{\delta_2} - a_1 \left(\frac{u_2}{v}\right)^{\delta_1}} dv, \quad c = \prod_{j=1}^2 \frac{a_j^{\frac{\gamma_j}{\delta_j}}}{\Gamma\left(\frac{\gamma_j}{\delta_j}\right)} \\
 &= c u_2^{\gamma_1-1} \int_0^\infty v^{\gamma_2-\gamma_1-1} e^{-a_2 v^{\delta_2} - a_1 (u_2^{\delta_1}/v^{\delta_1})} dv. \tag{2.4.2}
 \end{aligned}$$

The integral in (2.4.2) is connected to several topics. For $\delta_1 = 1, \delta_2 = 1$, this integral is the basic Krätzel integral and Krätzel transform, see Mathai and Haubold (2020). When $\delta_2 = 1, \delta_1 = \frac{1}{2}$, the integral in (2.4.2) is the basic reaction-rate probability integral in nuclear reaction-rate theory, see Mathai and Haubold (1988). For $\delta_1 = 1, \delta_2 = 1$, the integrand in (2.4.2), once normalized, is the inverse Gaussian density for appropriate values of $\gamma_2 - \gamma_1 - 1$. Observe that (2.4.2) is also connected to the Bayesian structure of unconditional densities if the conditional and marginal densities belong to generalized gamma family of densities. When $\delta_2 = 1$, the integral is a mgf of the remaining part with a_2 as the mgf parameter (It is therefore the Laplace transform of the remaining part of the function).

Now, let us consider different f_1 and f_2 . Let $f_1(x_1)$ be a real type-1 beta density with the parameters $(\gamma + 1, \alpha)$, $\Re(\alpha) > 0, \Re(\gamma) > -1$ (in statistical problems, the parameters are real but in this case the results hold as well for complex parameters; accordingly, the conditions are stated for complex parameters), that is, the density of x_1 is

$$f_1(x_1) = \frac{\Gamma(\gamma + 1 + \alpha)}{\Gamma(\gamma + 1)\Gamma(\alpha)} x_1^\gamma (1 - x_1)^{\alpha-1}, \quad 0 \leq x_1 \leq 1, \quad \alpha > 0, \quad \gamma > -1,$$

and $f_1(x_1) = 0$ elsewhere. Let $f_2(x_2) = f(x_2)$ where f is an arbitrary density. Then, the density of u_2 is given by

$$\begin{aligned}
 g_2(u_2) &= c \frac{1}{\Gamma(\alpha)} \int_v \frac{1}{v} \left(\frac{u_2}{v}\right)^\gamma \left(1 - \frac{u_2}{v}\right)^{\alpha-1} f(v) dv, \quad c = \frac{\Gamma(\gamma + 1 + \alpha)}{\Gamma(\gamma + 1)} \\
 &= c \frac{u_2^\gamma}{\Gamma(\alpha)} \int_{v \geq u_2 > 0} v^{-\gamma-\alpha} (v - u_2)^{\alpha-1} f(v) dv \tag{2.4.3}
 \end{aligned}$$

$$= c K_{2,u_2,\gamma}^{-\alpha} f \tag{2.4.4}$$

where $K_{2,u_2,\gamma}^{-\alpha} f$ is the Erdélyi-Kober fractional integral of order α of the second kind, with parameter γ in the real scalar variable case. Hence, if f is an arbitrary density, then $K_{2,u_2,\gamma}^{-\alpha} f$ is $\frac{\Gamma(\gamma+1)}{\Gamma(\gamma+1+\alpha)} g_2(u_2)$ or a constant multiple of the density of a product of independently distributed real scalar positive random variables where one of them has a real type-1 beta density and the other has an arbitrary density. When f_1 and f_2 are densities, then $g_2(u_2)$ has the structure

$$g_2(u_2) = \int_v^1 \frac{1}{v} f_1\left(\frac{u_2}{v}\right) f_2(v) dv. \quad (i)$$

Whether or not f_1 and f_2 are densities, the structure in (i) is called *the Mellin convolution of a product* in the sense that if we take the Mellin transform of g_2 , with Mellin parameter s , then

$$M_{g_2}(s) = M_{f_1}(s) M_{f_2}(s) \quad (2.4.5)$$

where $M_{g_2}(s) = \int_0^\infty u_2^{s-1} g_2(u_2) du_2$,

$$M_{f_1}(s) = \int_0^\infty x_1^{s-1} f_1(x_1) dx_1 \text{ and } M_{f_2}(s) = \int_0^\infty x_2^{s-1} f_2(x_2) dx_2,$$

whenever the Mellin transforms exist. Here (2.4.5) is the Mellin convolution of a product property. In statistical terms, when f_1 and f_2 are densities and when x_1 and x_2 are independently distributed, we have

$$E[u_2^{s-1}] = E[x_1^{s-1}] E[x_2^{s-1}] \quad (2.4.6)$$

whenever the expected values exist. Taking different forms of f_1 and f_2 , where f_1 has a factor $\frac{(1-x_1)^{\alpha-1}}{\Gamma(\alpha)}$ for $0 \leq x_1 \leq 1$, $\Re(\alpha) > 0$, it can be shown that the structure appearing in (i) produces all the various fractional integrals of the second kind of order α available in the literature for the real scalar variable case, such as the Riemann-Liouville fractional integral, Weyl fractional integral, etc. Connections of distributions of products and ratios to fractional integrals were established in a series of papers which appeared in *Linear Algebra and its Applications*, see Mathai (2013, 2014, 2015).

Now, let us consider the density of a ratio. Again, let $x_1 > 0$, $x_2 > 0$ be independently distributed real scalar random variables with density functions $f_1(x_1)$ and $f_2(x_2)$, respectively. Let $u_1 = \frac{x_2}{x_1}$ and let $v = x_2$. Then $dx_1 \wedge dx_2 = -\frac{v}{u_1^2} du_1 \wedge dv$. If we take $x_1 = v$, the Jacobian will be only v and not $-\frac{v}{u_1^2}$ and the final structure will be different. However, the first transformation is required in order to establish connections to fractional integral

of the first kind. If f_1 and f_2 are generalized gamma densities as described earlier and if $x_1 = v$, then the marginal density of u_1 , denoted by $g_1(u_1)$, will be as follows:

$$g_1(u_1) = c \int_v v v^{\gamma_1-1} (u_1 v)^{\gamma_2-1} e^{-a_1 v^{\delta_1} - a_2 (u_1 v)^{\delta_2}} dv, \quad c = \prod_{j=1}^2 \frac{a_j^{\frac{\gamma_j}{\delta_j}}}{\Gamma(\frac{\gamma_j}{\delta_j})},$$

$$= c u_1^{\gamma_2-1} \int_{v=0}^{\infty} v^{\gamma_1+\gamma_2-1} e^{-a_1 v^{\delta_1} - a_2 (u_1 v)^{\delta_2}} dv \quad (2.4.7)$$

$$= \frac{c}{\delta} \Gamma\left(\frac{\gamma_1 + \gamma_2}{\delta}\right) u_1^{\gamma_2-1} (a_1 + a_2 u_1^{\delta})^{-\frac{\gamma_1+\gamma_2}{\delta}}, \quad \text{for } \delta_1 = \delta_2 = \delta. \quad (2.4.8)$$

On the other hand, if $x_2 = v$, then the Jacobian is $-\frac{v}{u_1^2}$ and the marginal density, again denoted by $g_1(u_1)$, will be as follows when f_1 and f_2 are gamma densities:

$$g_1(u_1) = c \int_v \left(\frac{v}{u_1^2}\right) \left(\frac{v}{u_1}\right)^{\gamma_1-1} v^{\gamma_2-1} e^{-a_1 \left(\frac{v}{u_1}\right)^{\delta_1} - a_2 v^{\delta_2}} dv$$

$$= c u_1^{-\gamma_1-1} \int_{v=0}^{\infty} v^{\gamma_1+\gamma_2-1} e^{-a_2 v^{\delta_2} - a_1 \left(\frac{v}{u_1}\right)^{\delta_1}} dv.$$

This is one of the representations of the density of a product discussed earlier, which is also connected to Krätzel integral, reaction-rate probability integral, and so on. Now, let us consider a type-1 beta density for x_1 with the parameters (γ, α) having the following density:

$$f_1(x_1) = \frac{\Gamma(\gamma + \alpha)}{\Gamma(\gamma)\Gamma(\alpha)} x_1^{\gamma-1} (1 - x_1)^{\alpha-1}, \quad 0 \leq x_1 \leq 1,$$

for $\gamma > 0$, $\alpha > 0$ and $f_1(x_1) = 0$ elsewhere. Let $f_2(x_2) = f(x_2)$ where f is an arbitrary density. Letting $u_1 = \frac{x_2}{x_1}$ and $x_2 = v$, the density of u_1 , again denoted by $g_1(u_1)$, is

$$g_1(u_1) = \frac{\Gamma(\gamma + \alpha)}{\Gamma(\gamma)\Gamma(\alpha)} \int_v \frac{v}{u_1^2} \left(\frac{v}{u_1}\right)^{\gamma-1} \left(1 - \frac{v}{u_1}\right)^{\alpha-1} f(v) dv$$

$$= \frac{\Gamma(\gamma + \alpha)}{\Gamma(\gamma)} \frac{u_1^{-\gamma-\alpha}}{\Gamma(\alpha)} \int_{v \leq u_1} v^{\gamma} (u_1 - v)^{\alpha-1} f(v) dv \quad (2.4.9)$$

$$= \frac{\Gamma(\gamma + \alpha)}{\Gamma(\gamma)} K_{1,u_1,\gamma}^{-\alpha} f, \quad \Re(\alpha) > 0, \quad (2.4.10)$$

where $K_{1,u_1,\gamma}^{-\alpha} f$ is Erdélyi-Kober fractional integral of the first kind of order α and parameter γ . If f_1 and f_2 are densities, this Erdélyi-Kober fractional integral of the first kind is a constant multiple of the density of a ratio $g_1(u_1)$. In statistical terms,

$$u_1 = \frac{x_2}{x_1} \Rightarrow E[u_1^{s-1}] = E[x_2^{s-1}]E[x_1^{-s+1}] \text{ with } E[x_1^{-s+1}] = E[x_1^{(2-s)-1}] \Rightarrow$$

$$M_{g_1}(s) = M_{f_1}(2-s)M_{f_2}(s), \quad (2.4.11)$$

which is the Mellin convolution of a ratio. Whether or not f_1 and f_2 are densities, (2.4.11) is taken as the Mellin convolution of a ratio and it cannot be given statistical interpretations when f_1 and f_2 are not densities. For example, let $f_1(x_1) = x_1^{-\alpha} \frac{(1-x_1)^{\alpha-1}}{\Gamma(\alpha)}$ and $f_2(x_2) = x_2^\alpha f(x_2)$ where $f(x_2)$ is an arbitrary function. Then the Mellin convolution of a ratio, as in (2.4.11), again denoted by $g_1(u_1)$, is given by

$$g_1(u_1) = \int_{v \leq u_1} \frac{(u_1 - v)^{\alpha-1}}{\Gamma(\alpha)} f(v) dv, \quad \Re(\alpha) > 0. \quad (2.4.12)$$

This is *Riemann-Liouville fractional integral of the first kind of order α* if v is bounded below; when v is not bounded below, then it is *Weyl fractional integral of the first kind of order α* . An introduction to fractional calculus is presented in Mathai and Haubold (2018). The densities of u_1 and u_2 are connected to various problems in different areas for different functions f_1 and f_2 .

In the $p \times p$ matrix case in the complex domain, we will assume that the matrix is Hermitian positive definite. Note that when $p = 1$, Hermitian positive definite means a real positive variable. Hence in the scalar case, we will not discuss ratios and products in the complex domain since densities must be real-valued functions.

Exercises 2.4

2.4.1. Derive the density of (1): a real non-central F, where the numerator chisquare is non-central and the denominator chisquare is central, (2): a real doubly non-central F where both the chisquares are non-central with non-centrality parameters λ_1 and λ_2 respectively.

2.4.2. Let x_1 and x_2 be real gamma random variables with parameters (α_1, β) and (α_2, β) with the same β respectively and be independently distributed. Let $u_1 = \frac{x_1}{x_1+x_2}$, $u_2 = \frac{x_1}{x_2}$, $u_3 = x_1 + x_2$. Compute the densities of u_1, u_2, u_3 . Hint: Use the transformation $x_1 = r \cos^2 \theta$, $x_2 = r \sin^2 \theta$.

2.4.3. Let x_1 and x_2 be as defined as in Exercise 2.4.2. Let $u = x_1x_2$. Derive the density of u .

2.4.4. Let x_j have a real type-1 beta density with the parameters (α_j, β_j) , $j = 1, 2$ and be independently distributed. Let $u_1 = x_1x_2$, $u_2 = \frac{x_1}{x_2}$. Derive the densities of u_1 and u_2 . State the conditions under which these densities reduce to simpler known densities.

2.4.5. Evaluate (i): Weyl fractional integral of the second kind of order α if the arbitrary function is $f(v) = e^{-v}$; (ii) Riemann-Liouville fractional integral of the first kind of order α if the lower limit is 0 and the arbitrary function is $f(v) = v^\delta$.

2.4.6. In Exercise 2.4.2 show that (1): u_1 and u_3 are independently distributed; (2): u_2 and u_3 are independently distributed.

2.4.7. In Exercise 2.4.2 show that for arbitrary h , $E\left[\frac{x_1}{x_1+x_2}\right]^h = \frac{E(x_1^h)}{E(x_1+x_2)^h}$ and state the conditions for the existence of the moments. [Observe that, in general, $E\left(\frac{y_1}{y_2}\right)^h \neq \frac{E(y_1^h)}{E(y_2^h)}$ even if y_1 and y_2 are independently distributed.]

2.4.8. Derive the corresponding densities in Exercise 2.4.1 for the complex domain by taking the chisquares in the complex domain.

2.4.9. Extend the results in Exercise 2.4.2 to the complex domain by taking chisquare variables in the complex domain instead of gamma variables.

2.5. General Structures

2.5.1. Product of real scalar gamma variables

Let x_1, \dots, x_k be independently distributed real scalar gamma random variables with x_j having the density $f_j(x_j) = c_j x_j^{\alpha_j-1} e^{-\frac{x_j}{\beta_j}}$, $0 \leq x_j < \infty$, $\alpha_j > 0$, $\beta_j > 0$ and $f_j(x_j) = 0$ elsewhere. Consider the product $u = x_1x_2 \cdots x_k$. Such structures appear in many situations such as geometrical probability problems when we consider gamma distributed random points, see Mathai (1999). How can we determine the density of such a general structure? The transformation of variables technique is not a feasible procedure in this case. Since the x_j 's are positive, we may determine the Mellin transforms of the x_j 's with parameter s . Then, when $f_j(x_j)$ is a density, the Mellin transform $M_{f_j}(s)$, once

expressed in terms of an expected value, is $M_{f_j}(s) = E[x_j^{s-1}]$ whenever the expected value exists:

$$\begin{aligned} M_{f_j}(s) &= E[x_j^{s-1}] = \frac{1}{\beta_j^{\alpha_j} \Gamma(\beta_j)} \int_0^\infty x_j^{s-1} x_j^{\alpha_j-1} e^{-\frac{x_j}{\beta_j}} dx_j \\ &= \frac{\Gamma(\alpha_j + s - 1)}{\Gamma(\alpha_j)} \beta_j^{s-1}, \Re(\alpha_j + s - 1) > 0. \end{aligned}$$

Hence,

$$E[u^{s-1}] = \prod_{j=1}^k E[x_j^{s-1}] = \prod_{j=1}^k \frac{\Gamma(\alpha_j + s - 1)}{\Gamma(\alpha_j)} \beta_j^{s-1}, \Re(\alpha_j + s - 1) > 0, j = 1, \dots, k,$$

and the density of u is available from the inverse Mellin transform. If $g(u)$ is the density of u , then

$$g(u) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \left\{ \prod_{j=1}^k \frac{\Gamma(\alpha_j + s - 1)}{\Gamma(\alpha_j)} \beta_j^{s-1} \right\} u^{-s} dx, i = \sqrt{-1}. \quad (2.5.1)$$

This is a contour integral where c is any real number such that $c > -\Re(\alpha_j - 1)$, $j = 1, \dots, k$. The integral in (2.5.1) is available in terms of a known special function, namely Meijer's G-function. The G-function can be defined as follows:

$$\begin{aligned} G(z) &= G_{p,q}^{m,n}(z) = G_{p,q}^{m,n} \left[z \left| \begin{matrix} a_1, \dots, a_p \\ b_1, \dots, b_q \end{matrix} \right. \right] \\ &= \frac{1}{2\pi i} \int_L \frac{\{\prod_{j=1}^m \Gamma(b_j + s)\} \{\prod_{j=1}^n \Gamma(1 - a_j - s)\}}{\{\prod_{j=m+1}^q \Gamma(1 - b_j - s)\} \{\prod_{j=n+1}^p \Gamma(a_j + s)\}} z^{-s} ds, i = \sqrt{-1}. \end{aligned} \quad (2.5.2)$$

The existence conditions, different possible contours L , as well as properties and applications are discussed in Mathai (1993), Mathai and Saxena (1973, 1978), and Mathai et al. (2010). With the help of (2.5.2), we may now express (2.5.1) as follows in terms of a G-function:

$$g(u) = \left\{ \prod_{j=1}^k \frac{1}{\beta_j \Gamma(\alpha_j)} \right\} G_{0,k}^{k,0} \left[\frac{u}{\beta_1 \dots \beta_k} \left| \alpha_j - 1, j=1, \dots, k \right. \right] \quad (2.5.3)$$

for $0 \leq u < \infty$. Series and computable forms of a general G-function are provided in Mathai (1993). They are built-in functions in the symbolic computational packages *Mathematica* and *MAPLE*.

2.5.2. Product of real scalar type-1 beta variables

Let y_1, \dots, y_k be independently distributed real scalar type-1 beta random variables with the parameters (α_j, β_j) , $\alpha_j > 0$, $\beta_j > 0$, $j = 1, \dots, k$. Consider the product $u_1 = y_1 \cdots y_k$. Such a structure occurs in several contexts. It appears for instance in geometrical probability problems in connection with type-1 beta distributed random points. As well, when testing certain hypotheses on the parameters of one or more multivariate normal populations, the resulting likelihood ratio criteria, also known as λ -criteria, or one- to-one functions thereof, have the structure of a product of independently distributed real type-1 beta variables under the null hypothesis. The density of u_1 can be obtained by proceeding as in the previous section. Since the moment of u_1 of order $s - 1$ is

$$\begin{aligned} E[u_1^{s-1}] &= \prod_{j=1}^k E[y_j^{s-1}] \\ &= \prod_{j=1}^k \frac{\Gamma(\alpha_j + s - 1)}{\Gamma(\alpha_j)} \frac{\Gamma(\alpha_j + \beta_j)}{\Gamma(\alpha_j + \beta_j + s - 1)} \end{aligned} \quad (2.5.4)$$

for $\Re(\alpha_j + s - 1) > 0$, $j = 1, \dots, k$, Then, the density of u_1 , denoted by $g_1(u_1)$, is given by

$$\begin{aligned} g_1(u_1) &= \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} [E(u_1^{s-1})] u_1^{-s} ds \\ &= \left\{ \prod_{j=1}^k \frac{\Gamma(\alpha_j + \beta_j)}{\Gamma(\alpha_j)} \right\} \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \left\{ \prod_{j=1}^k \frac{\Gamma(\alpha_j + s - 1)}{\Gamma(\alpha_j + \beta_j + s - 1)} \right\} u_1^{-s} ds \\ &= \left\{ \prod_{j=1}^k \frac{\Gamma(\alpha_j + \beta_j)}{\Gamma(\alpha_j)} \right\} G_{k,k}^{k,0} \left[u_1 \middle| \begin{matrix} \alpha_j + \beta_j - 1, j=1, \dots, k \\ \alpha_j - 1, j=1, \dots, k \end{matrix} \right] \end{aligned} \quad (2.5.5)$$

for $0 \leq u_1 \leq 1$, $\Re(\alpha_j + s - 1) > 0$, $j = 1, \dots, k$.

2.5.3. Product of real scalar type-2 beta variables

Let $u_2 = z_1 z_2 \cdots z_k$ where the z_j 's are independently distributed real scalar type-2 beta random variables with the parameters (α_j, β_j) , $\alpha_j > 0$, $\beta_j > 0$, $j = 1, \dots, k$. Such products are encountered in several situations, including certain problems in geometrical probability that are discussed in Mathai (1999). Then,

$$E[z_j^{s-1}] = \frac{\Gamma(\alpha_j + s - 1) \Gamma(\beta_j - s + 1)}{\Gamma(\alpha_j) \Gamma(\beta_j)}, \quad -\Re(\alpha_j - 1) < \Re(s) < \Re(\beta_j + 1),$$

and

$$E[u_2^{s-1}] = \prod_{j=1}^k E[z_j^{s-1}] = \left\{ \prod_{j=1}^k [\Gamma(\alpha_j) \Gamma(\beta_j)]^{-1} \right\} \left\{ \prod_{j=1}^k \Gamma(\alpha_j + s - 1) \Gamma(\beta_j - s + 1) \right\}. \quad (2.5.6)$$

Hence, the density of u_2 , denoted by $g_2(u_2)$, is given by

$$\begin{aligned} g_2(u_2) &= \left\{ \prod_{j=1}^k [\Gamma(\alpha_j) \Gamma(\beta_j)]^{-1} \right\} \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \left\{ \prod_{j=1}^k \Gamma(\alpha_j + s - 1) \Gamma(\beta_j - s + 1) \right\} u_2^{-s} ds \\ &= \left\{ \prod_{j=1}^k [\Gamma(\alpha_j) \Gamma(\beta_j)]^{-1} \right\} G_{k,k}^{k,k} \left[u_2 \middle|_{\alpha_j-1, j=1, \dots, k}^{-\beta_j, j=1, \dots, k} \right], \quad u_2 \geq 0. \end{aligned} \quad (2.5.7)$$

2.5.4. General products and ratios

Let us consider a structure of the following form:

$$u_3 = \frac{t_1 \cdots t_r}{t_{r+1} \cdots t_k},$$

where the t_j 's are independently distributed real positive variables, such as real type-1 beta, real type-2 beta, and real gamma variables, where the expected values $E[t_j^{s-1}]$ for $j = 1, \dots, k$, will produce various types of gamma products, some containing $+s$ and others, $-s$, both in the numerator and in the denominator. Accordingly, we obtain a general structure such as that appearing in (2.5.2), and the density of u_3 , denoted by $g_3(u_3)$, will then be proportional to a general G-function.

2.5.5. The H-function

Let $u = v_1 v_2 \cdots v_k$ where the v_j 's are independently distributed generalized real gamma variables with densities

$$h_j(v_j) = \frac{a_j^{\frac{\gamma_j}{\delta_j}}}{\Gamma(\frac{\gamma_j}{\delta_j})} v_j^{\gamma_j-1} e^{-a_j v_j^{\delta_j}}, \quad v_j \geq 0,$$

for $a_j > 0$, $\delta_j > 0$, $\gamma_j > 0$, and $h_j(v_j) = 0$ elsewhere for $j = 1, \dots, k$. Then,

$$E[v_j^{s-1}] = \frac{\Gamma(\frac{\gamma_j+s-1}{\delta_j})}{\Gamma(\frac{\gamma_j}{\delta_j})} \frac{1}{a_j^{\frac{s-1}{\delta_j}}},$$

for $\Re(\gamma_j + s - 1) > 0$, $v_j \geq 0$, $\delta_j > 0$, $a_j > 0$, $\gamma_j > 0$, $j = 1, \dots, k$, and

$$E[u^{s-1}] = \left\{ \prod_{j=1}^k \frac{a_j^{\frac{1}{\delta_j}}}{\Gamma(\frac{\gamma_j}{\delta_j})} \right\} \left\{ \prod_{j=1}^k \Gamma(\frac{\gamma_j - 1}{\delta_j} + \frac{s}{\delta_j}) a_j^{-\frac{s}{\delta_j}} \right\}. \quad (2.5.8)$$

Thus, the density of u , denoted by $g_3(u)$, is given by

$$\begin{aligned} g_3(u) &= \left\{ \prod_{j=1}^k \frac{a_j^{\frac{1}{\delta_j}}}{\Gamma(\frac{\gamma_j}{\delta_j})} \right\} \frac{1}{2\pi i} \int_L \left\{ \prod_{j=1}^k \Gamma(\frac{\gamma_j - 1}{\delta_j} + \frac{s}{\delta_j}) a_j^{-\frac{s}{\delta_j}} \right\} u^{-s} ds \\ &= \left\{ \prod_{j=1}^k \frac{a_j^{\frac{1}{\delta_j}}}{\Gamma(\frac{\gamma_j}{\delta_j})} \right\} H_{0,k}^{k,0} \left[\left\{ \prod_{j=1}^k a_j^{\frac{1}{\delta_j}} \right\} u \middle|_{(\frac{\gamma_j-1}{\delta_j}, \frac{1}{\delta_j}), j=1,\dots,k} \right] \end{aligned} \quad (2.5.9)$$

for $u \geq 0$, $\Re(\gamma_j + s - 1) > 0$, $j = 1, \dots, k$, where L is a suitable contour and the general H-function is defined as follows:

$$\begin{aligned} H(z) &= H_{p,q}^{m,n}(z) = H_{p,q}^{m,n} \left[z \middle|_{(b_1, \beta_1), \dots, (b_q, \beta_q)}^{(a_1, \alpha_1), \dots, (a_p, \alpha_p)} \right] \\ &= \frac{1}{2\pi i} \int_L \frac{\{\prod_{j=1}^m \Gamma(b_j + \beta_j s)\} \{\prod_{j=1}^n \Gamma(1 - a_j - \alpha_j s)\}}{\{\prod_{j=m+1}^q \Gamma(1 - b_j - \beta_j s)\} \{\prod_{j=n+1}^p \Gamma(a_j + \alpha_j s)\}} z^{-s} ds \end{aligned} \quad (2.5.10)$$

where $\alpha_j > 0$, $j = 1, \dots, p$; $\beta_j > 0$, $j = 1, \dots, q$ are real and positive, b_j 's and a_j 's are complex numbers, the contour L separates the poles of $\Gamma(b_j + \beta_j s)$, $j = 1, \dots, m$, lying on one side of it and the poles of $\Gamma(1 - a_j - \alpha_j s)$, $j = 1, \dots, n$, which must lie on the other side. The existence conditions and the various types of possible contours are discussed in Mathai and Saxena (1978) and Mathai et al. (2010). Observe that we can consider arbitrary powers of the variables present in u , u_1 , u_2 and u_3 as introduced in Sects. 2.5.1–2.5.5; however, in this case, the densities of these various structures will be expressible in terms of H-functions rather than G-functions. In the G-function format as defined in (2.5.2), the complex variable s has ± 1 as its coefficients, whereas the coefficients of s in the H-function, that is, $\pm \alpha_j$, $\alpha_j > 0$ and $\pm \beta_j$, $\beta_j > 0$, are not restricted to unities.

We will give a simple illustrative example that requires the evaluation of an inverse Mellin transform. Let $f(x) = e^{-x}$, $x > 0$. Then, the Mellin transform is

$$M_f(s) = \int_0^{\infty} x^{s-1} e^{-x} dx = \Gamma(s), \quad \Re(s) > 0,$$

and it follows from the inversion formula that

$$f(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \Gamma(s) x^{-s} ds, \quad \Re(s) > 0, \quad i = \sqrt{-1}. \quad (2.5.11)$$

If $f(x)$ is unknown and we are told that the Mellin transform of a certain function is $\Gamma(s)$, then are we going to retrieve $f(x)$ as e^{-x} from the inversion formula? Let us explore this problem. The poles of $\Gamma(s)$ occur at $s = 0, -1, -2, \dots$. Thus, if we take c in the contour of integration as $c > 0$, this contour will enclose all the poles of $\Gamma(s)$. We may now apply Cauchy's residue theorem. By definition, the residue at $s = -\nu$, denoted by R_ν , is

$$R_\nu = \lim_{s \rightarrow -\nu} (s + \nu) \Gamma(s) x^{-s}.$$

We cannot substitute $s = -\nu$ to obtain the limit in this case. However, noting that

$$(s + \nu) \Gamma(s) x^{-s} = \frac{(s + \nu)(s + \nu - 1) \cdots s \Gamma(s) x^{-s}}{(s + \nu - 1) \cdots s} = \frac{\Gamma(s + \nu + 1) x^{-s}}{(s + \nu - 1) \cdots s},$$

which follows from the recursive relationship, $\alpha \Gamma(\alpha) = \Gamma(\alpha + 1)$, the limit can be taken:

$$\begin{aligned} \lim_{s \rightarrow -\nu} (s + \nu) \Gamma(s) x^{-s} &= \lim_{s \rightarrow -\nu} \frac{\Gamma(s + \nu + 1) x^{-s}}{(s + \nu - 1) \cdots s} \\ &= \frac{\Gamma(1) x^\nu}{(-1)(-2) \cdots (-\nu)} = \frac{(-1)^\nu x^\nu}{\nu!}. \end{aligned} \quad (2.5.12)$$

Hence, the sum of the residues is

$$\sum_{\nu=0}^{\infty} R_\nu = \sum_{\nu=0}^{\infty} \frac{(-1)^\nu x^\nu}{\nu!} = e^{-x},$$

and the function is recovered.

Note 2.5.1. Distributions of products and ratios of random variables in the complex domain could as well be worked out. However, since they may not necessarily have practical applications, they will not be discussed herein. Certain product and ratio distributions for variables in the complex domain which reduce to real variables, such as a chisquare in the complex domain, have already been previously discussed.

Exercises 2.5

2.5.1. Evaluate the density of $u = x_1x_2$ where the x_j 's are independently distributed real type-1 beta random variables with the parameters (α_j, β_j) , $\alpha_j > 0$, $\beta_j > 0$, $j = 1, 2$ by using Mellin and inverse Mellin transform technique. Evaluate the density for the case $\alpha_1 - \alpha_2 \neq \pm v$, $v = 0, 1, \dots$ so that the poles are simple.

2.5.2. Repeat Exercise 2.5.1 if x_j 's are (1): real type-2 beta random variables with parameters (α_j, β_j) , $\alpha_j > 0$, $\beta_j > 0$ and (2): real gamma random variables with the parameters (α_j, β_j) , $\alpha_j > 0$, $\beta_j > 0$, $j = 1, 2$.

2.5.3. Let $u = \frac{u_1}{u_2}$ where u_1 and u_2 are real positive random variables. Then the h -th moment, for arbitrary h , is $E[\frac{u_1}{u_2}]^h \neq \frac{E[u_1^h]}{E[u_2^h]}$ in general. Give two examples where $E[\frac{u_1}{u_2}]^h = \frac{E[u_1^h]}{E[u_2^h]}$.

2.5.4. $E[\frac{1}{u}]^h = E[u^{-h}] \neq \frac{1}{E[u^h]}$ in general. Give two examples where $E[\frac{1}{u}] = \frac{1}{E[u]}$.

2.5.5. Let $u = \frac{x_1x_2}{x_3x_4}$ where the x_j 's are independently distributed. Let x_1, x_3 be type-1 beta random variables, x_2 be a type-2 beta random variable, and x_4 be a gamma random variable with parameters (α_j, β_j) , $\alpha_j > 0$, $\beta_j > 0$, $j = 1, 2, 3, 4$. Determine the density of u .

2.6. A Collection of Random Variables

Let x_1, \dots, x_n be iid (independently and identically distributed) real scalar random variables with a common density denoted by $f(x)$, that is, assume that the sample comes from the population that is specified by $f(x)$. Let the common mean value be μ and the common variance be $\sigma^2 < \infty$, that is, $E(x_j) = \mu$ and $\text{Var}(x_j) = \sigma^2$, $j = 1, \dots, n$, where E denotes the expected value. Denoting the sample average by $\bar{x} = \frac{1}{n}(x_1 + \dots + x_n)$, what can be said about \bar{x} when $n \rightarrow \infty$? This is the type of questions that will be investigated in this section.

2.6.1. Chebyshev's inequality

For some $k > 0$, let us examine the probability content of $|x - \mu|$ where $\mu = E(x)$ and the variance of x is $\sigma^2 < \infty$. Consider the probability that the random variable x lies outside the interval $\mu - k\sigma < x < \mu + k\sigma$, that is k times the standard deviation σ away from the mean value μ . From the definition of the variance σ^2 for a real scalar random variable x ,

$$\begin{aligned}
\sigma^2 &= \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx \\
&= \int_{-\infty}^{\mu - k\sigma} (x - \mu)^2 f(x) dx + \int_{\mu - k\sigma}^{\mu + k\sigma} (x - \mu)^2 f(x) dx + \int_{\mu + k\sigma}^{\infty} (x - \mu)^2 f(x) dx \\
&\geq \int_{-\infty}^{\mu - k\sigma} (x - \mu)^2 f(x) dx + \int_{\mu + k\sigma}^{\infty} (x - \mu)^2 f(x) dx
\end{aligned}$$

since the probability content over the interval $\mu - k\sigma < x < \mu + k\sigma$ is omitted. Over this interval, the probability is either positive or zero, and hence the inequality. However, the intervals $(-\infty, \mu - k\sigma]$ and $[\mu + k\sigma, \infty)$, that is, $-\infty < x \leq \mu - k\sigma$ and $\mu + k\sigma \leq x < \infty$ or $-\infty < x - \mu \leq -k\sigma$ and $k\sigma \leq x - \mu < \infty$, can thus be described as the intervals for which $|x - \mu| \geq k\sigma$. In these intervals, the smallest value that $|x - \mu|$ can take on is $k\sigma$, $k > 0$ or equivalently, the smallest value that $|x - \mu|^2$ can assume is $(k\sigma)^2 = k^2\sigma^2$. Accordingly, the above inequality can be further sharpened as follows:

$$\begin{aligned}
\sigma^2 &\geq \int_{|x - \mu| \geq k\sigma} (x - \mu)^2 f(x) dx \geq \int_{|x - \mu| \geq k\sigma} (k\sigma)^2 f(x) dx \Rightarrow \\
\frac{\sigma^2}{k^2\sigma^2} &\geq \int_{|x - \mu| \geq k\sigma} f(x) dx \Rightarrow \\
\frac{1}{k^2} &\geq \int_{|x - \mu| \geq k\sigma} f(x) dx = Pr\{|x - \mu| \geq k\sigma\}, \text{ that is,} \\
\frac{1}{k^2} &\geq Pr\{|x - \mu| \geq k\sigma\},
\end{aligned}$$

which can be written as

$$Pr\{|x - \mu| \geq k\sigma\} \leq \frac{1}{k^2} \text{ or } Pr\{|x - \mu| < k\sigma\} \geq 1 - \frac{1}{k^2}. \quad (2.6.1)$$

If $k\sigma = k_1$, $k = \frac{k_1}{\sigma}$, and the above inequalities can be written as follows:

$$Pr\{|x - \mu| \geq k\} \leq \frac{\sigma^2}{k^2} \text{ or } Pr\{|x - \mu| < k\} \geq 1 - \frac{\sigma^2}{k^2}. \quad (2.6.2)$$

The inequalities (2.6.1) and (2.6.2) are known as Chebyshev's inequalities (also referred to as Chebycheff's inequalities). For example, when $k = 2$, Chebyshev's inequality states

that $Pr\{|x - \mu| < 2\sigma\} \geq 1 - \frac{1}{4} = 0.75$, which is not a very sharp probability limit. If $x \sim N_1(\mu, \sigma^2)$, then we know that

$$Pr\{|x - \mu| < 1.96\sigma\} \approx 0.95 \text{ and } Pr\{|x - \mu| < 3\sigma\} \approx 0.99.$$

Note that the bound 0.75 resulting from Chebyshev's inequality seriously underestimate the actual probability for a Gaussian variable x . However, what is astonishing about this inequality, is that the given probability bound holds for any distribution, whether it be continuous, discrete or mixed. Sharper bounds can of course be obtained for the probability content of the interval $[\mu - k\sigma, \mu + k\sigma]$ when the exact distribution of x is known.

These inequalities can be expressed in terms of generalized moments. Let $\mu_r^{\frac{1}{r}} = \{E|x - \mu|^r\}^{\frac{1}{r}}$, $r = 1, 2, \dots$, which happens to be a measure of scatter in x from the mean value μ . Given that

$$\mu_r = \int_{-\infty}^{\infty} |x - \mu|^r f(x) dx,$$

consider the probability content of the intervals specified by $|x - \mu| \geq k\mu_r^{\frac{1}{r}}$ for $k > 0$. Paralleling the derivations of (2.6.1) and (2.6.2), we have

$$\begin{aligned} \mu_r &\geq \int_{|x-\mu| \geq k\mu_r^{\frac{1}{r}}} |x - \mu|^r f(x) dx \geq \int_{|x-\mu| \geq k\mu_r^{\frac{1}{r}}} |(k\mu_r^{\frac{1}{r}})|^r f(x) dx \Rightarrow \\ Pr\{|x - \mu| \geq k\mu_r^{\frac{1}{r}}\} &\leq \frac{1}{k^r} \text{ or } Pr\{|x - \mu| < k\mu_r^{\frac{1}{r}}\} \geq 1 - \frac{1}{k^r}, \end{aligned} \quad (2.6.3)$$

which can also be written as

$$Pr\{|x - \mu| \geq k\} \leq \frac{\mu_r}{k^r} \text{ or } Pr\{|x - \mu| < k\} \geq 1 - \frac{\mu_r}{k^r}, \quad r = 1, 2, \dots \quad (2.6.4)$$

Note that when $r = 2$, $\mu_r = \sigma^2$, and Chebyshev's inequalities as specified in (2.6.1) and (2.6.2) are obtained from (2.6.3) and (2.6.4), respectively. If x is a real scalar positive random variables with $f(x) = 0$ for $x \leq 0$, we can then obtain similar inequalities in terms of the first moment μ . For $k > 0$,

$$\begin{aligned} \mu = E(x) &= \int_0^{\infty} xf(x) dx \text{ since } f(x) = 0 \text{ for } x \leq 0 \\ &= \int_0^k xf(x) dx + \int_k^{\infty} xf(x) dx \geq \int_k^{\infty} xf(x) dx \geq \int_k^{\infty} kf(x) dx \Rightarrow \\ \frac{\mu}{k} &\geq \int_k^{\infty} f(x) dx = Pr\{x \geq k\}. \end{aligned}$$

Accordingly, we have the following inequality for any real positive random variable x :

$$\Pr\{x \geq k\} \leq \frac{\mu}{k} \text{ for } x > 0, k > 0. \quad (2.6.5)$$

Suppose that our variable is $\bar{x} = \frac{1}{n}(x_1 + \cdots + x_n)$, where x_1, \dots, x_n are iid variables with common mean value μ and the common variance $\sigma^2 < \infty$. Then, since $\text{Var}(\bar{x}) = \frac{\sigma^2}{n}$ and $E(\bar{x}) = \mu$, Chebyshev's inequality states that

$$\Pr\{|\bar{x} - \mu| < k\} \geq 1 - \frac{\sigma^2}{n} \rightarrow 1 \text{ as } n \rightarrow \infty \quad (2.6.6)$$

or $\Pr\{|\bar{x} - \mu| \geq k\} \rightarrow 0$ as $n \rightarrow \infty$. However, since a probability cannot be greater than 1, $\Pr\{|\bar{x} - \mu| < k\} \rightarrow 1$ as $n \rightarrow \infty$. In other words, \bar{x} tends to μ with probability 1 as $n \rightarrow \infty$. This is referred to as the *Weak Law of Large Numbers*.

The Weak Law of Large Numbers

Let x_1, \dots, x_n be iid with common mean value μ and common variance $\sigma^2 < \infty$. Then, as $n \rightarrow \infty$,

$$\Pr\{\bar{x} \rightarrow \mu\} \rightarrow 1. \quad (2.6.7)$$

Another limiting property is known as the *Central Limit Theorem*. Let x_1, \dots, x_n be iid real scalar random variables with common mean value μ and common variance $\sigma^2 < \infty$. Letting $\bar{x} = \frac{1}{n}(x_1 + \cdots + x_n)$ denote the sample mean, the standardized sample mean is

$$u = \frac{\bar{x} - E(\bar{x})}{\sqrt{\text{Var}(\bar{x})}} = \frac{\sqrt{n}}{\sigma}(\bar{x} - \mu) = \frac{1}{\sigma\sqrt{n}}[(x_1 - \mu) + \cdots + (x_n - \mu)]. \quad (i)$$

Consider the characteristic function of $x - \mu$, that is,

$$\begin{aligned} \phi_{x-\mu}(t) &= E[e^{it(x-\mu)}] = 1 + \frac{it}{1!}E(x-\mu) + \frac{(it)^2}{2!}E(x-\mu)^2 + \cdots \\ &= 1 + 0 - \frac{t^2}{2!}E(x-\mu)^2 + \cdots = 1 + \frac{t}{1!}\phi^{(1)}(0) + \frac{t^2}{2!}\phi^{(2)}(0) + \cdots \quad (ii) \end{aligned}$$

where $\phi^{(r)}(0)$ is the r -th derivative of $\phi(t)$ with respect to t , evaluated at $t = 0$. Let us consider the characteristic function of our standardized sample mean u .

Making use of the last representation of u in (i), we have $\phi_{\frac{\sum_{j=1}^n(x_j-\mu)}{\sigma\sqrt{n}}}(t) = [\phi_{x_j-\mu}(\frac{t}{\sigma\sqrt{n}})]^n$ so that $\phi_u(t) = [\phi_{x_j-\mu}(\frac{t}{\sigma\sqrt{n}})]^n$ or $\ln \phi_u(t) = n \ln \phi_{x_j-\mu}(\frac{t}{\sigma\sqrt{n}})$. It then

follows from (ii) that

$$\begin{aligned} [\phi_{x_j-\mu}(\frac{t}{\sigma\sqrt{n}})] &= 1 + 0 - \frac{t^2}{2!} \frac{\sigma^2}{n\sigma^2} - i \frac{t^3}{3!} \frac{E(x_j - \mu)^3}{(\sigma\sqrt{n})^3} + \dots \\ &= 1 - \frac{t^2}{2n} + O\left(\frac{1}{n^{\frac{3}{2}}}\right). \end{aligned} \tag{iii}$$

Now noting that $\ln(1 - y) = -[y + \frac{y^2}{2} + \frac{y^3}{3} + \dots]$ whenever $|y| < 1$, we have

$$\ln \phi_{x_j-\mu}(\frac{t}{\sigma\sqrt{n}}) = -\frac{t^2}{2n} + O\left(\frac{1}{n^{\frac{3}{2}}}\right) \Rightarrow n \ln \phi_{x_j-\mu}(\frac{t}{\sigma\sqrt{n}}) = -\frac{t^2}{2} + O\left(\frac{1}{n^{\frac{1}{2}}}\right) \rightarrow -\frac{t^2}{2} \text{ as } n \rightarrow \infty.$$

Consequently, as $n \rightarrow \infty$,

$$\phi_u(t) = e^{-\frac{t^2}{2}} \Rightarrow u \rightarrow N_1(0, 1) \text{ as } n \rightarrow \infty.$$

This is known as the central limit theorem.

The Central Limit Theorem. Let x_1, \dots, x_n be iid real scalar random variables having common mean value μ and common variance $\sigma^2 < \infty$. Let the sample mean be $\bar{x} = \frac{1}{n}(x_1 + \dots + x_n)$ and u denote the standardized sample mean. Then

$$u = \frac{\bar{x} - E(\bar{x})}{\sqrt{\text{Var}(\bar{x})}} = \frac{\sqrt{n}}{\sigma}(\bar{x} - \mu) \rightarrow N_1(0, 1) \text{ as } n \rightarrow \infty. \tag{2.6.8}$$

Generalizations, extensions and more rigorous statements of this theorem are available in the literature. We have focussed on the substance of the result, assuming that a simple random sample is available and that the variance of the population is finite.

Exercises 2.6

2.6.1. For a binomial random variable with the probability function $f(x) = \binom{n}{x} p^x (1 - p)^{n-x}$, $0 < p < 1$, $x = 0, 1, \dots, n$, $n = 1, 2, \dots$ and zero elsewhere, show that the standardized binomial variable itself, namely $\frac{x-np}{\sqrt{np(1-p)}}$ goes to the standard normal when $n \rightarrow \infty$.

2.6.2. State the central limit theorem for the following real scalar populations by evaluating the mean value and variance there, assuming that a simple random sample is available: (1) Poisson random variable with parameter λ ; (2) Geometric random variable with parameter p ; (3) Negative binomial random variable with parameters (p, k) ; (4) Discrete

hypergeometric probability law with parameters (a, b, n) ; (5) Uniform density over $[a, b]$; (6) Exponential density with parameter θ ; (7) Gamma density with the parameters (α, β) ; (8) Type-1 beta random variable with the parameters (α, β) ; (9) Type-2 beta random variable with the parameters (α, β) .

2.6.3. State the central limit theorem for the following probability/density functions: (1):

$$f(x) = \begin{cases} 0.5, & x = 2 \\ 0.5, & x = 5 \end{cases}$$

and $f(x) = 0$ elsewhere; (2): $f(x) = 2e^{-2x}$, $0 \leq x < \infty$ and zero elsewhere; (3): $f(x) = 1$, $0 \leq x \leq 1$ and zero elsewhere. Assume that a simple random sample is available from each population.

2.6.4. Consider a real scalar gamma random variable x with the parameters (α, β) and show that $E(x) = \alpha\beta$ and variance of x is $\alpha\beta^2$. Assume a simple random sample x_1, \dots, x_n from this population. Derive the densities of (1): $x_1 + \dots + x_n$; (2): \bar{x} ; (3): $\bar{x} - \alpha\beta$; (4): Standardized sample mean \bar{x} . Show that the densities in all these cases are still gamma densities, may be relocated, for all finite values of n however large n may be.

2.6.5. Consider the density $f(x) = \frac{c}{x^\alpha}$, $1 \leq x < \infty$ and zero elsewhere, where c is the normalizing constant. Evaluate c stating the relevant conditions. State the central limit theorem for this population, stating the relevant conditions.

2.7. Parameter Estimation: Point Estimation

There exist several methods for estimating the parameters of a given density/probability function, based on a simple random sample of size n (iid variables from the population designated by the density/probability function). The most popular methods of point estimation are the method of maximum likelihood and the method of moments.

2.7.1. The method of moments and the method of maximum likelihood

The likelihood function $L(\theta)$ is the joint density/probability function of the sample values, at an observed sample point, x_1, \dots, x_n . As a function of θ , $L(\theta)$, or a one-to-one function thereof, is maximized in order to determine the most likely value of θ in terms of a function of the given sample. This estimation process is referred to as the method of maximum likelihood.

Let $m_r = \frac{\sum_{j=1}^n x_j^r}{n}$ denote the r -th integer moment of the sample, where x_1, \dots, x_n is the observed sample point, the corresponding population r -th moment being $E[x^r]$, where E denotes the expected value. According to the method of moments, the estimates of the parameters are obtained by solving $m_r = E[x^r]$, $r = 1, 2, \dots$.

For example, consider a $N_1(\mu, \sigma^2)$ population with density

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}, \quad -\infty < x < \infty, \quad -\infty < \mu < \infty, \quad \sigma > 0, \quad (2.7.1)$$

where μ and σ^2 are the parameters here. Let x_1, \dots, x_n be a simple random sample from this population. Then, the joint density of x_1, \dots, x_n , denoted by $L = L(x_1, \dots, x_n; \mu, \sigma^2)$, is

$$\begin{aligned} L &= \frac{1}{[\sqrt{2\pi}\sigma]^n} e^{-\frac{1}{2\sigma^2} \sum_{j=1}^n (x_j - \mu)^2} = \frac{1}{[\sqrt{2\pi}\sigma]^n} e^{-\frac{1}{2\sigma^2} [\sum_{j=1}^n (x_j - \bar{x})^2 + n(\bar{x} - \mu)^2]} \Rightarrow \\ \ln L &= -n \ln(\sqrt{2\pi}\sigma) - \frac{1}{2\sigma^2} \left[\sum_{j=1}^n (x_j - \bar{x})^2 + n(\bar{x} - \mu)^2 \right], \quad \bar{x} = \frac{1}{n}(x_1 + \dots + x_n). \end{aligned} \quad (2.7.2)$$

Maximizing L or $\ln L$, since L and $\ln L$ are one-to-one functions, with respect to μ and $\theta = \sigma^2$, and solving for μ and σ^2 produces the maximum likelihood estimators (MLE's). An observed value of the estimator is the corresponding estimate. It follows from a basic result in Calculus that the extrema of L can be determined by solving the equations

$$\frac{\partial}{\partial \mu} \ln L = 0 \quad (i)$$

and

$$\frac{\partial}{\partial \theta} \ln L = 0, \quad \theta = \sigma^2. \quad (ii)$$

Equation (i) produces the solution $\mu = \bar{x}$ so that \bar{x} is the MLE of μ . Note that \bar{x} is a random variable and that \bar{x} evaluated at a sample point or at a set of observations on x_1, \dots, x_n produces the corresponding estimate. We will denote both the estimator and estimate of μ by $\hat{\mu}$. As well, we will utilize the same abbreviation, namely, MLE for the maximum likelihood estimator and the corresponding estimate. Solving (ii) and substituting $\hat{\mu}$ to μ , we have $\hat{\theta} = \hat{\sigma}^2 = \frac{1}{n} \sum_{j=1}^n (x_j - \bar{x})^2 = s^2 =$ the sample variance as an estimate of $\theta = \sigma^2$. Does the point (\bar{x}, s^2) correspond to a local maximum or a local minimum or a saddle point? Since the matrix of second order partial derivatives at the point (\bar{x}, s^2) is

negative definite, the critical point (\bar{x}, s^2) corresponds to a maximum. Thus, in this case, $\hat{\mu} = \bar{x}$ and $\hat{\sigma}^2 = s^2$ are the maximum likelihood estimators/estimates of the parameters μ and σ^2 , respectively. If we were to differentiate with respect to σ instead of $\theta = \sigma^2$ in (ii), we would obtain the same estimators, since for any differentiable function $g(t)$, $\frac{d}{dt}g(t) = 0 \Rightarrow \frac{d}{d\phi(t)}g(t) = 0$ if $\frac{d}{dt}\phi(t) \neq 0$. In this instance, $\phi(\sigma) = \sigma^2$ and $\frac{d}{d\sigma}\sigma^2 \neq 0$.

For obtaining the moment estimates, we equate the sample integer moments to the corresponding population moments, that is, we let $m_r = E[x^r]$, $r = 1, 2$, two equations being required to estimate μ and σ^2 . Note that $m_1 = \bar{x}$ and $m_2 = \frac{1}{n} \sum_{j=1}^n x_j^2$. Then, consider the equations

$$\bar{x} = E[x] = \mu \quad \text{and} \quad \frac{1}{n} \sum_{j=1}^n x_j^2 = E[x^2] \Rightarrow s^2 = \sigma^2.$$

Thus, the moment estimators/estimates of μ and σ^2 , which are $\hat{\mu} = \bar{x}$ and $\hat{\sigma}^2 = s^2$, happen to be identical to the MLE's in this case.

Let us consider the type-1 beta population with parameters (α, β) whose density is

$$f_1(x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \quad 0 \leq x \leq 1, \quad \alpha > 0, \quad \beta > 0, \quad (2.7.3)$$

and zero otherwise. In this case, the likelihood function contains gamma functions and the derivatives of gamma functions involve psi and zeta functions. Accordingly, the maximum likelihood approach is not very convenient here. However, we can determine moment estimates without much difficulty from (2.7.3). The first two population integer moments are obtained directly from a representation of the h -th moment:

$$E[x^h] = \frac{\Gamma(\alpha + h)}{\Gamma(\alpha)} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha + \beta + h)}, \quad \Re(\alpha + h) > 0 \Rightarrow$$

$$E[x] = \frac{\Gamma(\alpha + 1)}{\Gamma(\alpha)} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha + \beta + 1)} = \frac{\alpha}{\alpha + \beta} \quad (2.7.4)$$

$$E[x^2] = \frac{\alpha(\alpha + 1)}{(\alpha + \beta)(\alpha + \beta + 1)} = E[x] \frac{\alpha + 1}{\alpha + \beta + 1}. \quad (2.7.5)$$

Equating the sample moments to the corresponding population moments, that is, letting $m_1 = E[x]$ and $m_2 = E[x^2]$, it follows from (2.7.4) that

$$\bar{x} = \frac{\alpha}{\alpha + \beta} \Rightarrow \frac{\beta}{\alpha} = \frac{1 - \bar{x}}{\bar{x}}; \quad (iii)$$

Then, from (2.7.5), we have

$$\frac{\frac{1}{n} \sum_{j=1}^n x_j^2}{\bar{x}} = \frac{\alpha + 1}{\alpha + \beta + 1} = \frac{1}{1 + \frac{\beta}{\alpha + 1}} \Rightarrow \frac{\bar{x} - \sum_{j=1}^n x_j^2/n}{\sum_{j=1}^n x_j^2/n} = \frac{\beta}{\alpha + 1}. \quad (iv)$$

The parameter β can be eliminated from (iii) and (iv), which yields an estimate of α ; $\hat{\beta}$ is then obtained from (iii). Thus, the moment estimates are available from the equations, $m_r = E[x^r]$, $r = 1, 2$, even though these equations are nonlinear in the parameters α and β . The method of maximum likelihood or the method of moments can similarly yield parameters estimates for populations that are otherwise distributed.

2.7.2. Bayes' estimates

This procedure is more relevant when the parameters in a given statistical density/probability function have their own distributions. For example, let the real scalar variable x be discrete having a binomial probability law for the fixed (given) parameter p , that is, let $f(x|p) = \binom{n}{x} p^x (1-p)^{n-x}$, $0 < p < 1$, $x = 0, 1, \dots, n$, $n = 1, 2, \dots$, and $f(x|p) = 0$ elsewhere be the conditional probability function. Let p have a prior type-1 beta density with known parameters α and β , that is, let the prior density of p be

$$g(p) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1} (1-p)^{\beta-1}, \quad 0 < p < 1, \quad \alpha > 0, \quad \beta > 0$$

and $g(p) = 0$ elsewhere. Then, the joint probability function $f(x, p) = f(x|p)g(p)$ and the unconditional probability function of x , denoted by $f_1(x)$, is as follows:

$$\begin{aligned} f_1(x) &= \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \binom{n}{x} \int_0^1 p^{\alpha+x-1} (1-p)^{\beta+n-x-1} dp \\ &= \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \binom{n}{x} \frac{\Gamma(\alpha + x)\Gamma(\beta + n - x)}{\Gamma(\alpha + \beta + n)}. \end{aligned}$$

Thus, the posterior density of p , given x , denoted by $g_1(p|x)$, is

$$g_1(p|x) = \frac{f(x, p)}{f_1(x)} = \frac{\Gamma(\alpha + \beta + n)}{\Gamma(\alpha + x)\Gamma(\beta + n - x)} p^{\alpha+x-1} (1-p)^{\beta+n-x-1}.$$

Accordingly, the expected value of p in this conditional distribution of p given x , which is called the posterior density of p , is known as the Bayes estimate of p :

$$\begin{aligned}
E[p|x] &= \frac{\Gamma(\alpha + \beta + n)}{\Gamma(\alpha + x)\Gamma(\beta + n - x)} \int_0^1 p p^{\alpha+x-1} (1-p)^{\beta+n-x-1} dp \\
&= \frac{\Gamma(\alpha + \beta + n)}{\Gamma(\alpha + x)\Gamma(\beta + n - x)} \frac{\Gamma(\alpha + x + 1)\Gamma(\beta + n - x)}{\Gamma(\alpha + \beta + n + 1)} \\
&= \frac{\Gamma(\alpha + x + 1)}{\Gamma(\alpha + x)} \frac{\Gamma(\alpha + \beta + n)}{\Gamma(\alpha + \beta + n + 1)} = \frac{\alpha + x}{\alpha + \beta + n}.
\end{aligned}$$

The prior estimate/estimator of p as obtained from the binomial distribution is $\frac{x}{n}$ and the posterior estimate or the Bayes estimate of p is

$$E[p|x] = \frac{\alpha + x}{\alpha + \beta + n},$$

so that $\frac{x}{n}$ is revised to $\frac{\alpha+x}{\alpha+\beta+n}$. In general, if the conditional density/probability function of x given θ is $f(x|\theta)$ and the prior density/probability function of θ is $g(\theta)$, then the posterior density of θ is $g_1(\theta|x)$ and $E[\theta|x]$ or *the expected value of θ in the conditional distribution of θ given x is the Bayes estimate of θ .*

2.7.3. Interval estimation

Before concluding this section, the concept of confidence intervals or interval estimation of a parameter will be briefly touched upon. For example, let x_1, \dots, x_n be iid $N_1(\mu, \sigma^2)$ and let $\bar{x} = \frac{1}{n}(x_1 + \dots + x_n)$. Then, $\bar{x} \sim N_1(\mu, \frac{\sigma^2}{n})$, $(\bar{x} - \mu) \sim N_1(0, \frac{\sigma^2}{n})$ and $z = \frac{\sqrt{n}}{\sigma}(\bar{x} - \mu) \sim N_1(0, 1)$. Since the standard normal density $N_1(0, 1)$ is free of any parameter, one can select two percentiles, say, a and b , from a standard normal table and make a probability statement such as $Pr\{a < z < b\} = 1 - \alpha$ for every given α ; for instance, $Pr\{-1.96 \leq z \leq 1.96\} \approx 0.95$ for $\alpha = 0.05$. Let $Pr\{-z_{\frac{\alpha}{2}} \leq z \leq z_{\frac{\alpha}{2}}\} = 1 - \alpha$ where $z_{\frac{\alpha}{2}}$ is such that $Pr(z > z_{\frac{\alpha}{2}}) = \frac{\alpha}{2}$. The following inequalities are mathematically equivalent and hence the probabilities associated with the corresponding intervals are equal:

$$\begin{aligned}
-z_{\frac{\alpha}{2}} \leq z \leq z_{\frac{\alpha}{2}} &\Leftrightarrow -z_{\frac{\alpha}{2}} \leq \frac{\sqrt{n}}{\sigma}(\bar{x} - \mu) \leq z_{\frac{\alpha}{2}} \\
&\Leftrightarrow \mu - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \leq \bar{x} \leq \mu + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \\
&\Leftrightarrow \bar{x} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}. \tag{i}
\end{aligned}$$

Accordingly,

$$Pr\{\mu - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \leq \bar{x} \leq \mu + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\} = 1 - \alpha \tag{ii}$$

⇔

$$Pr\{\bar{x} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\} = 1 - \alpha. \quad (iii)$$

Note that (iii) is not a usual probability statement as opposed to (ii), which is a probability statement on a random variable. In (iii), the interval $[\bar{x} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}]$ is random and μ is a constant. This can be given the interpretation that the random interval covers the parameter μ with probability $1 - \alpha$, which means that we are $100(1 - \alpha)\%$ confident that the random interval will cover the unknown parameter μ or that the random interval is an interval estimator and an observed value of the interval is the interval estimate of μ . We could construct such an interval only because the distribution of z is parameter-free, which enabled us to make a probability statement on μ .

In general, if $u = u(x_1, \dots, x_n, \theta)$ is a function of the sample values and the parameter θ (which may be a vector of parameters) and if the distribution of u is free of all parameter, then such a quantity is referred to as a *pivotal quantity*. Since the distribution of the pivotal quantity is parameter-free, we can find two numbers a and b such that $Pr\{a \leq u \leq b\} = 1 - \alpha$ for every given α . If it is possible to convert the statement $a \leq u \leq b$ into a mathematically equivalent statement of the type $u_1 \leq \theta \leq u_2$, so that $Pr\{u_1 \leq \theta \leq u_2\} = 1 - \alpha$ for every given α , then $[u_1, u_2]$ is called a $100(1 - \alpha)\%$ *confidence interval or interval estimate for θ* , u_1 and u_2 being referred to as the *lower confidence limit* and the *upper confidence limit*, and $1 - \alpha$ being called the *confidence coefficient*. Additional results on interval estimation and the construction of confidence intervals are, for instance, presented in Mathai and Haubold (2017b).

Exercises 2.7

2.7.1. Obtain the method of moments estimators for the parameters (α, β) in a real type-2 beta population. Assume that a simple random sample of size n is available.

2.7.2. Obtain the estimate/estimator of the parameters by the method of moments and the method of maximum likelihood in the real (1): exponential population with parameter θ , (2): Poisson population with parameter λ . Assume that a simple random sample of size n is available.

2.7.3. Let x_1, \dots, x_n be a simple random sample of size n from a point Bernoulli population $f_2(x) = p^x(1 - p)^{1-x}$, $x = 0, 1$, $0 < p < 1$ and zero elsewhere. Obtain the MLE as well as moment estimator for p . [Note: These will be the same estimators for p in all the populations based on Bernoulli trials, such as binomial population, geometric population, negative binomial population].

2.7.4. If possible, obtain moment estimators for the parameters of a real generalized gamma population, $f_3(x) = c x^{\alpha-1} e^{-bx^\delta}$, $\alpha > 0$, $b > 0$, $\delta > 0$, $x \geq 0$ and zero elsewhere, where c is the normalizing constant.

2.7.5. If possible, obtain the MLE of the parameters a, b in the following real uniform population $f_4(x) = \frac{1}{b-a}$, $b > a$, $a \leq x < b$ and zero elsewhere. What are the MLE if $a < x < b$? What are the moment estimators in these two situations?

2.7.6. Construct the Bayes' estimate/estimator of the parameter λ in a Poisson probability law if the prior density for λ is a gamma density with known parameters (α, β) .

2.7.7. By selecting the appropriate pivotal quantities, construct a 95% confidence interval for (1): Poisson parameter λ ; (2): Exponential parameter θ ; (3): Normal parameter σ^2 ; (4): θ in a uniform density $f(x) = \frac{1}{\theta}$, $0 \leq x \leq \theta$ and zero elsewhere.

References

A.M. Mathai (1993): *A Handbook of Generalized Special Functions for Statistical and Physical Sciences*, Oxford University Press, Oxford, 1993.

A.M. Mathai (1999): *Introduction to Geometrical Probability: Distributional Aspects and Applications*, Gordon and Breach, Amsterdam, 1999.

A.M. Mathai (2005): A pathway to matrix-variate gamma and normal densities, *Linear Algebra and its Applications*, **410**, 198-216.

A.M. Mathai (2013): Fractional integral operators in the complex matrix-variate case, *Linear Algebra and its Applications*, **439**, 2901-2913.

A.M. Mathai (2014): Fractional integral operators involving many matrix variables, *Linear Algebra and its Applications*, **446**, 196-215.

A.M. Mathai (2015): Fractional differential operators in the complex matrix-variate case, *Linear Algebra and its Applications*, **478**, 200-217.

A.M. Mathai and Hans J. Haubold (1988): *Modern Problems in Nuclear and Neutrino Astrophysics*, Akademie-Verlag, Berlin.

A.M. Mathai and Hans J. Haubold (2017a): *Linear Algebra, A course for Physicists and Engineers*, De Gruyter, Germany, 2017.

A.M. Mathai and Hans J. Haubold (2017b): *Probability and Statistics, A course for Physicists and Engineers*, De Gruyter, Germany, 2017.

A.M. Mathai and Hans J. Haubold (2018): *Erdélyi-Kober Fractional Calculus from a Statistical Perspective, Inspired by Solar Neutrino Physics*, Springer Briefs in Mathematical Physics, 31, Springer Nature, Singapore, 2018.

A.M. Mathai and Hans J. Haubold (2020): Mathematical aspects of Krätzel integral and Krätzel transform, *Mathematics*, **8(526)**; doi:10.3390/math8040526.

A.M. Mathai and Serge B. Provost (1992): *Quadratic Forms in Random Variables: Theory and Applications*, Marcel Dekker, New York, 1992.

A.M. Mathai and Serge B. Provost (2006): On q-logistic and related distributions, *IEEE Transactions on Reliability*, **55(2)**, 237-244.

A.M. Mathai and R.K. Saxena (1973): *Generalized Hypergeometric Functions*, Springer Lecture Notes No. 348, Heidelberg, 1973.

A.M. Mathai and R. K. Saxena (1978): *The H-function with Applications in Statistics and Other Disciplines*, Wiley Halsted, New York, 1978.

A.M. Mathai, R.K. Saxena and Hans J. Haubold (2010): *The H-function: Theory and Applications*, Springer, New York, 2010.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

