UC Santa Barbara
UC Santa Barbara Previously Published Works

Title

Correlated gamma variables in the analysis of microbial densities in water

Permalink

https://escholarship.org/uc/item/6qr1r17n

Journal

Advances in Water Resources, 28(4)

ISSN

0309-1708

Authors

Loáiciga, Hugo A
Leipnik, Roy B

Publication Date

2005-04-01

DOI

10.1016/j.advwatres.2004.11.008

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Advances in Water Resources 28 (2005) 329–335
Advances in Water Resources
www.elsevier.com/locate/advwatres

Correlated gamma variables in the analysis of microbial densities in water

Hugo A. Loáiciga [a,*], Roy B. Leipnik [b]

[a] Department of Geography, University of California, Santa Barbara, CA 93106-4060, USA
[b] Department of Mathematics, University of California, Santa Barbara, CA 93106-4060 ,USA

Received 19 August 2004; received in revised form 9 November 2004; accepted 18 November 2004
Available online 21 January 2005

Abstract

The probability density function (p.d.f.) of the ratio of two correlated gamma variables is derived and used to fit aquatic microbial-density data. The ratio p.d.f. is tackled by first taking the Fourier transform of a generalized Kibble–Gaver, unsymmetrical, characteristic function (c.f.) to obtain the corresponding bivariate p.d.f. of two correlated gamma variables with different shape and scale parameters. The ratio p.d.f. follows by weighted integration of the bivariate p.d.f. The derivation of the gamma bivariate and ratio p.d.f.s relies on the use of weighted Laguerre–Charlier polynomials that lead to p.d.f.s amenable to computation. The bivariate gamma p.d.f. and the ratio p.d.f. of correlated gamma variables are useful statistical tools in the analysis of skewed water-resources data. Computational examples illustrate the calculation of bivariate p.d.f.s for positive and negative correlation and the fitting of the ratio p.d.f. to correlated bacterial densities in stream water.
© 2004 Elsevier Ltd. All rights reserved.

Keywords: Gamma variables; Correlation; Microbial densities; Ratio distribution

1. Gamma variables and probability density functions (p.d.f.s)

Consider two gamma random variables X_1 and X_2, each of three parameters, whose marginal p.d.f.s are given by Eq. (1) (b_1, b_2, α_1, α_2, γ, ξ_1, ξ_2 are parameters in Eq. (1), and $x'_j = x_j - \xi_j$, $j = 1, 2$):

$$f(x_j) = \frac{x_j'^{\gamma\alpha_j - 1} e^{-\frac{x_j}{b_j}}}{\Gamma(\gamma\alpha_j) b_j^{\gamma\alpha_j}} \quad j = 1, 2 \tag{1}$$

$x_j \geqslant \xi_j$, in which $\gamma\alpha_1$ and $\gamma\alpha_2$ are the marginal shapes of the p.d.f.s of X_1 and X_2, respectively, with $\alpha_1, \alpha_2 \geqslant 0$; γ is a (collective) shape parameter of the bivariate distribution of X_1 and X_2, $\gamma > 0$; $b_1, b_2 > 0$ are scale parameters, and ξ_1, ξ_2 are location parameters. The role of the parameter γ is further elaborated upon in Section 2.

Gamma p.d.f.s have been widely used in water resources analysis (see a recent example in [24]). They are flexible and mathematically simple p.d.f.s that can fit skewed data. One frequently used variant of the gamma p.d.f. is the log-gamma p.d.f., in which the logarithm of a random variable is assumed to follow a gamma p.d.f. The log-gamma p.d.f., also called log-Pearson p.d.f. (with two or three parameters), is widely used in the United States to describe the distribution of peak annual streamflow (see, e.g., [18]).

The study of bivariate gamma p.d.f.s involving two correlated gamma variables has received much less attention in the water resources literature. By and large, bivariate (and multivariate) correlated water-resources data are handled using normal or log-normal p.d.f.s that are mathematically simple and computationally straightforward (see, however, [27]). Certainly,

* Corresponding author. Tel./fax: +1 805 964 3252.
E-mail address: hugo@geog.ucsb.edu (H.A. Loáiciga).

0309-1708/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.advwatres.2004.11.008

non-negative, skewed, data may not be well fit by the normal p.d.f. This situation calls for p.d.f.s of more general applicability, the gamma p.d.f. being one plausible alternative. This article pursues the case of correlated gamma variables that occur in pairs, and whose physical significance is expressed by their ratio, that is, by $Z = X_1/X_2$. In this work, X_1 and X_2 may, in general, feature different shape and scale parameters, i.e., α_1 may or may not equal α_2, and, likewise, β_1 may or may not equal β_2. In this respect, this article's results generalize—in nontrivial fashion—those of Xekalaki et al. [32], who, using Kibble's [15] symmetric bivariate gamma p.d.f., assumed that $\alpha_1 = \alpha_2$ and $b_1 = b_2$ in the derivation of the distribution of the ratio Z. The derivation of the p.d.f. of Z, $g(z)$, requires the bivariate p.d.f. of X_1 and X_2, $f(X_1, X_2)$, as it is shown below. A novel technique is used in this work to derive $f(X_1, X_2)$ and $g(z)$ in terms of infinite series amenable to computation. Several examples demonstrate the calculation of $f(X_1, X_2)$ for positive and negative correlation. The fitting of $g(z)$ is illustrated with microbial densities in stream water. Other applications are surely possible.

2. Characteristic functions (c.f.s) and correlation

The c.f.s (ϕ) of X_1 and X_2 are obtained by taking the Fourier transforms of Eq. (1) to yield

$$\phi_j(t_j) = \mathrm{e}^{\mathrm{i}\xi_j t_j}(1 - \mathrm{i}t_j b_j)^{-\gamma\alpha_j} \quad j = 1,2 \tag{2}$$

Assume that X_1 and X_2 are correlated variables, with the density correlation ρ defined in terms of their means μ_1, μ_2 and variances σ_1^2 and σ_2^2 as usual:

$$\rho \equiv \frac{\mu_{1,1}}{\sigma_1\sigma_2} = \frac{E[(x_1 - \mu_1)(x_2 - \mu_2)]}{\sigma_1\sigma_2} \tag{3}$$

The particular form of the bivariate gamma p.d.f. of correlated X_1 and X_2 is derived in Section 4.

Consider the following bivariate c.f. for the gamma variables X_1 and X_2 described above:

$$\phi(t_1, t_2) = \mathrm{e}^{\mathrm{i}(\xi_1 t_1 + \xi_2 t_2)}[(1 - \mathrm{i}t_1 b_1)^{\alpha_1}(1 - \mathrm{i}t_2 b_2)^{\alpha_2} + \beta t_1 t_2]^{-\gamma} \tag{4}$$

which produces the marginal c.f.s in Eq. (2). The parameter β in Eq. (4) is related to the density correlation ρ by the expression:

$$\rho = \frac{\beta\gamma}{\sigma_1\sigma_2} = \frac{\beta}{b_1 b_2\sqrt{\alpha_1\alpha_2}} \tag{5}$$

Clearly, $\operatorname{sgn}\rho = \operatorname{sgn}\beta$, and $|\rho| \leqslant 1$.

The proposed c.f. (4) is unsymmetrical, that is, it has unequal shape and scale parameters. Furthermore, the parameter β is related to the correlation between the variables X_1 and X_2, while γ captures the joint shape of their bivariate distribution. The unsymmetrical bivariate c.f. (4) is a generalization of the symmetrical bivariate results by Kibble [15] and Gaver [10]. The former, in particular, assumed that $\alpha_1 = \alpha_2$ and $b_1 = b_2$. Other statistical gamma models are the symmetrical, one-parameter, bivariate chi-square by Krishnaiah and Rao [19] and the symmetric multivariate gamma model of Krishnamoorthy–Parthasarathy [20], based on determinants. In the unsymmetrical realm, bivariate gamma distributions, ratio distributions, and their moments have been derived by Sarmanov [29], David and Fix [6], Gunst and Webster [12], and Prékova and Szántai [27]. The cited unsymmetrical distributions are both complicated in appearance and lacking in parameters. Accessibility of the broad family of unsymmetrical bivariate gamma distributions implied by the c.f. (4) relies on the introduction of Laguerre–Charlier polynomials and separable form of the joint distribution (see Sections 3 and 4). The basic paper of Meixner [25] and the statistical papers of Lancaster [21,22] used these polynomials and forms, but with fewer results and no computational testing of equations. Interestingly, the Krishnamoorthy–Parthasarathy [20] paper had a brief account of an unsymmetrical extension that does not appear to have been pursued. Becker and Roux [1] suggested a nice bivariate extension of the gamma distribution.

Other bivariate gamma and related distributions can be found in [31,5,28,21,8,26,30,7,17].

3. Laguerre–Charlier (LC) polynomials

Recall the Charlier polynomial $C_n(\lambda, x)$ [2,3,9]:

$$\begin{aligned} C_n(\lambda,x) &= \sum_{k=0}^{n}(-1)^k\binom{n}{k}\binom{\lambda}{k}\frac{k!}{x^k} \\ &= (-x)^{-n}n!\sum_{k=0}^{n}\binom{\lambda}{n-k}\frac{x^k(-1)^k}{k!} \end{aligned} \tag{6}$$

$n = 0, 1, 2, \ldots$; $-\infty \leqslant x \leqslant \infty$, $x \neq 0$, in which:

$$\binom{\lambda}{k} = \frac{\lambda(\lambda-1)\cdots(\lambda-k+1)}{k!} \tag{7}$$

is Newton's binomial coefficient. In terms of the confluent hypergeometric function ${}_1F_1(a, c; x)$, Eq. (6) may be rewritten in the following form:

$$C_n(v,x) = \frac{\Gamma(v+1)}{\Gamma(v-n+1)}(-x)^{-n}{}_1F_1(-n, v-n+1; x) \quad n = 0,1,2,\ldots \tag{8}$$

Using generalized Laguerre polynomials $L_j^{\alpha}(x)$, the conversion formula is

$$C_n(v,x) = (-x)^{-n}n!L_n^{(v-n)}(x) \tag{9}$$

which is why the $C_n(\lambda, x)$ are sometimes called Laguerre–Charlier (LC) polynomials. The extensive known identities for ${}_1F_1(a, c; x)$ and $L_j^{\alpha}(x)$ can therefore be tapped to express properties of the Charlier polynomials. Eqs. (6)–(9) constitute a set of useful formulas in the derivation of a suite of gamma results that follow.

4. Bivariate gamma probability density function

The bivariate p.d.f of X_1 and X_2 is obtained by taking the Fourier transform of the proposed c.f. (4):

$$f(x_1,x_2) = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}(t_1x_1+t_2x_2)}\phi(t_1,t_2)\,\mathrm{d}t_1\,\mathrm{d}t_2 \tag{10}$$

Substitution of Eq. (4) in (10) and using the LC-polynomials, yields, after integration, the bivariate, 8-parameter, gamma p.d.f. (with $x_j' = x_j - \xi_j$, $\lambda_j = \alpha_j(n+\gamma)$; $\lambda_j' = \lambda_j - 1$, $j = 1, 2$):

$$f(x_1,x_2) = \sum_{n=0}^{\infty}(-1)^n\binom{-\gamma}{n}\left(\frac{\beta}{b_1b_2}\right)^n \frac{\left(\frac{x_1'}{b_1}\right)^{\lambda_1'}\left(\frac{x_2'}{b_2}\right)^{\lambda_2'}\mathrm{e}^{-\left(\frac{x_1'}{b_1}+\frac{x_2'}{b_2}\right)}}{\Gamma(\lambda_1)\Gamma(\lambda_2)b_1b_2} \times C_n\left(\lambda_1',\frac{x_1'}{b_1}\right)C_n\left(\lambda_2',\frac{x_2'}{b_2}\right) \tag{11}$$

The bivariate gamma p.d.f. is a separable series of weighted LC polynomials. A (very lengthy) proof that the expression in Eq. (11) is a valid p.d.f. (i.e., non-negative and integrates to one) is available from the authors upon request.

Figs. 1–3 show the plots of calculated gamma p.d.f.s using Eq. (11). The values of the parameters used in Fig. 1 were $b_1 = b_2 = 1$, $\alpha_1 = \alpha_2 = 3$, $\beta = 1.5$, $\gamma = 1$ (which implies that $\rho = 0.5$), and $\xi_1 = \xi_2 = 0$. Notice in Fig. 1 the symmetry of the calculated p.d.f., a consequence of the choice $\alpha_1\gamma = \alpha_2\gamma$. Fig. 2 shows a second (unsymmetrical) bivariate p.d.f., in which case $b_1 = b_2 = 1$, $\alpha_1 = 1$, $\alpha_2 = 3$, $\beta = 1.5$, $\gamma = 1.5$ ($\rho = 0.87$). The ratio of the marginal shape parameters was $\alpha_1\gamma/\alpha_2\gamma = 1/3$. In Fig. 3 we show an (unsymmetrical) p.d.f. whose parameters are $b_1 = 1.5$, $b_2 = 2$, $\alpha_1 = 2$, $\alpha_2 = 1.5$, $\beta = -2$, $\gamma = 0.70$. In this last instance, the correlation is negative, $\rho = -2/(3\sqrt{3}) = -0.385$. The data shown in the graphs of Figs. 1–3 were produced with a computer program written in FORTRAN 90 interfaced with the International Mathematical Subroutine Library (IMSL). The IMSL returns values of special functions (gamma, binomial coefficients, etc.) that appear in the equations.

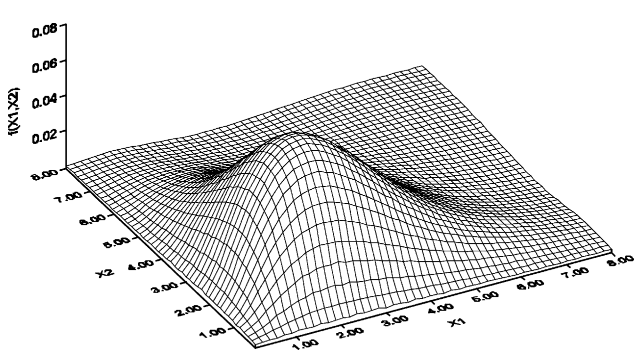

Fig. 1. An example of the bivariate gamma p.d.f. with $b_1 = b_2 = 1$, $\alpha_1 = \alpha_2 = 3$, $\beta = 1.5$, $\gamma = 1$, $\rho = 0.5$, and $\xi_1 = \xi_2 = 0$.

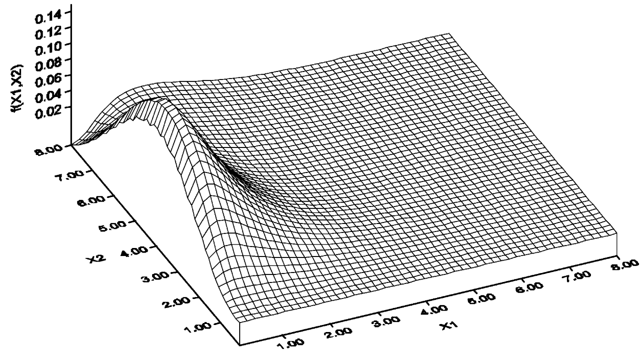

Fig. 2. An example of the bivariate gamma p.d.f. with $b_1 = b_2 = 1$, $\alpha_1 = \alpha_2 = 3$, $\beta = 1.5$, $\gamma = 1.5$, $\rho = 0.87$, $\xi_1 = \xi_2 = 0$.

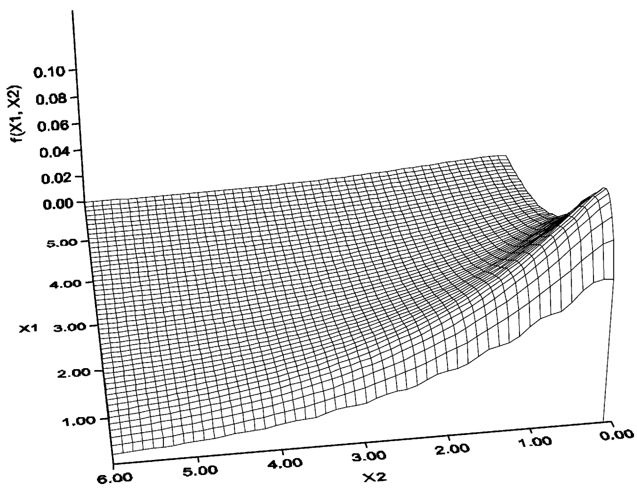

Fig. 3. An example of the bivariate gamma p.d.f., $b_1 = 1.5$, $b_2 = 2$, $\alpha_1 = 2$, $\alpha_2 = 1.5$, $\beta = -2$, $\gamma = 0.70$, $\rho = -0.385$, $\xi_1 = \xi_2 = 0$.

5. Moments of the bivariate gamma p.d.f.

Let $\mu_{n,m}$, $n, m = 0, 1, 2, \ldots$, be the (central) moments about the mean of X_1 and X_2 according to the p.d.f. (11) Then

$$\mu_{n,m} = E[(X_1-\mu_1)^n(X_2-\mu_2)^m] \tag{12}$$

In addition, let $\mu_{n,m}'$, $n, m = 0, 1, 2, \ldots$, be the central moments of $X_j' = X_j - \xi_j$, $j = 1, 2$. Note that $\mu_j' \equiv E(X_j') = \mu_j - \xi_j$, with $\mu_j \equiv E(X_j)$, $j = 1, 2$. To obtain $\mu_{n,m}$ it is advantageous to derive $\mu'_{n,m}$ and use the fact that $\mu_{n,m} = \mu'_{n,m}$. To this end, let the non-central moments $\mu''_{n,m}$ of X'_1 and X'_2 be defined as follows:

$$\mu''_{n,m} = E[X'^n_1 X'^m_2] \tag{13}$$

The non-central moments in Eq. (13) are related to the c.f. ϕ_0 of X'_1 and X'_2 by the following formula:

$$\mu''_{n,m} = (-\mathrm{i})^{n+m} \frac{\partial^{n+m}\phi_0}{\partial t_1^n \partial t_2^m}\bigg|_{t_1=t_2=0} \tag{14}$$

where, from Eq. (4):

$$\phi_0(t_1,t_2) = \lfloor (1-\mathrm{i}t_1 b_1)^{\alpha_1}(1-\mathrm{i}t_2 b_2)^{\alpha_2} + \beta t_1 t_2 \rfloor^{-\gamma} \tag{15}$$

Once the non-central moments $\mu''_{n,m}$ are known, the central moments $\mu_{n,m} = \mu'_{n,m}$ follow straightforwardly. For example, $\mu_{1,1} = \mu'_{1,1} = \mu''_{1,1,} - \mu'_1\mu'_2$. For the variance, we have that $\sigma_1^2 \equiv \mu_{2,0} = \mu''_{2,0} - \mu'^2_1$, etc. The first, second, and third moments were obtained by this method to yield:

$$E(X_j) = \alpha_j b_j \gamma + \xi_j \quad j = 1,2 \tag{16}$$

$$\sigma_j^2 = \alpha_j b_j^2 \gamma \quad j = 1,2 \tag{17}$$

$$\mu_{1,1} = \beta\gamma \tag{18}$$

$$\rho \equiv \frac{\mu_{1,1}}{\sigma_1\sigma_2} = \frac{\beta\gamma}{b_1 b_2\sqrt{\alpha_1\alpha_2\gamma^2}} = \frac{\beta}{b_1 b_2 \sqrt{\alpha_1\alpha_2}} \tag{19}$$

$$\mu_{3,0} = 2b_1^3\gamma\alpha_1 \tag{20}$$

$$\mu_{0,3} = 2b_2^3\gamma\alpha_2 \tag{21}$$

$$\mu_{2,1} = 2\beta\gamma\alpha_1 b_1 = 2\gamma\sqrt{\alpha_1^3\alpha_2} b_1^2 b_2 \rho \tag{22}$$

$$\mu_{1,2} = 2\beta\gamma\alpha_2 b_2 = 2\gamma\sqrt{\alpha_1\alpha_2^3} b_1 b_2^2 \rho \tag{23}$$

The correlation (19) and cross-skewness in Eqs. (22) and (23) vanish when $\beta = 0$, as required for uncorrelated variables. If desired, Eqs. (16)–(23) may be used to obtain estimates of b_1, b_2, α_1, α_2, β, and γ from the sample moments. The lower bounds ξ_1, ξ_2 are zero in most applications (although they need not be), usually a physical feasibility requirement. Such is the case of microbial densities in stream water, illustrated in Section 7.

6. The ratio p.d.f. and its moments

The p.d.f $g(z)$ of $Z = X_1/X_2$ is derived from the following relationship between $g(z)$ and the bivariate density $f(x_1, x_2)$ (see, e.g., [14], where, without loss of generality the location parameters are set equal to zero, $\xi_1 = \xi_2 = 0$):

$$g(z) = \int_0^\infty x_2 f(x_2 z, x_2)\,\mathrm{d}x_2 \tag{24}$$

Substitution of the joint p.d.f. (11) into (24) followed by termwise integration (using (6) twice) yields the p.d.f. $g(z)$, $z \geqslant 0$:

$$g(z) = \sum_{n=0}^{\infty}\sum_{k=0}^{n}\sum_{j=0}^{n} \frac{n!^2(-1)^{n+k+j}}{j!k!}\binom{-\gamma}{n}\binom{\lambda_1}{n-k}\binom{\lambda_2}{n-j} \times \left(\frac{\beta}{b_1^{\alpha_1}b_2^{\alpha_2}}\right)^n \left(\frac{b_1^{-(\gamma\alpha_1+k)}}{b_2^{\gamma\alpha_2+j}}\right)\cdot\frac{\Gamma(\lambda_{1,2})\cdot z^{\lambda_1+k-n-1}}{z'^{\lambda_{1,2}}\Gamma(\lambda_1)\Gamma(\lambda_2)} \tag{25}$$

in which $\lambda_j = \alpha_j(n + \gamma)$, $j = 1, 2$; $\lambda_{1,2} = \lambda_1 + \lambda_2 + k + j - 2n$, $\lambda'_j = \lambda_j - 1$, $j = 1, 2$; and

$$z' = \frac{z}{b_1} + \frac{1}{b_2} \tag{26}$$

Note the dependence of λ_j on j and n and of $\lambda_{1,2}$ on j, k, n, which lengthens the calculations of $g(z)$. Xekalaki et al. [32] used Kibble's symmetric bivariate gamma p.d.f. (with $\alpha_1 = \alpha_2$ and $b_1 = b_2$) in association with Eq. (24) to derive the symmetric version of Eq. (25), which is in this case reducible to closed form. Kotlarski [16] studied the p.d.f of the ratio Z of two positive-valued, correlated, random variables and its relationship with the F and other known distributions. Hinkley [13] derived the p.d.f. of Z for two correlated normal variables.

Consider next the (non-central) moments of the ratio $Z = X_1/X_2$, defined as follows:

$$\mu'_{Z,p} = \int_0^\infty z^p g(z)\,\mathrm{d}z = \int_0^\infty z^p\,\mathrm{d}z \int_0^\infty x_2 f(x_2 z, x_2)\,\mathrm{d}x_2 \tag{27}$$

Letting Ω_{njk} be the numerical coefficient in (25), we rewrite (25) in the form:

$$g(z) = \sum_{njk} \Omega_{njk} z^{\lambda_1+k-n-1}\left(\frac{z}{b_1}+\frac{1}{b_2}\right)^{-\lambda_{1,2}} \tag{28}$$

Using (28) in (27) produces (with $\lambda''_1 = \lambda_1 + k - n + p$):

$$\mu'_{Z,p} = \sum_{njk} \Omega_{njk} b_2^{\lambda_{1,2}} \left(\frac{b_1}{b_2}\right)^{\lambda''_1} B[\lambda''_1, \lambda_{1,2} - \lambda''_1] \tag{29}$$

in which $B[u, v]$ is the Euler beta function. These are defined either by a trigonometric integral ([11, Eq. 8.381.4], if $\mathrm{Re}(u + v) \geqslant 2$) or an algebraic integral ([11, Eq. 8.380.1], if $\mathrm{Re}\,u > 0$, $\mathrm{Re}\,v > 0$). Here, $u \geqslant \alpha_1\gamma > 0$; $u + v = \lambda_{1,2} \geqslant (\alpha_1 + \alpha_2)\gamma$, but $v = \lambda_2 + j - n - p$ can become negative for some u, n, p. This would be avoided if

$$p < \alpha_2\gamma \tag{30}$$

In the example of Section 7, $u + v \geqslant (\alpha_1 + \alpha_2)\gamma > 3.7$, so all the terms in the series (29) are well defined. Note that the dependences of the series (29) on p, b_1, b_2 are complicated, although in terms of p and b_1/b_2 it is simply $(b_1/b_2)^p$.

Full derivations of Eqs. (25) and (29) are available from the authors upon request. The conditional moments of X_1, X_2 with p.d.f. (11) and of the ratio Z (given

X_2) have also been determined exactly and are available from the authors.

7. An application to water-quality data

Water in Las Palmas Creek, Santa Barbara, CA, was tested to study the ratio of fecal coliforms (FC) to fecal streptococcus (FS) in it. FC and FS are enteric bacteria, that is, they live in the intestinal tract of warm-blooded animals, and are frequently used as indicators of fecal contamination of water bodies [4]. A total of thirty eight pairs of 100-milliliter water aliquots were collected in 1999 and 2000 [23]. In each pair of aliquots, one was analyzed for FC and the other for FS. The FC and FS values were well represented by univariate gamma distributions, which are useful on their own right to make assessments about their statistical and environmental significance. From each pair of FC and FS values, the corresponding ratio FC/FS was determined. This procedure yielded a FC/FS sample of thirty eight experimental values. The FC/FS ratio is of interest because, under suitable conditions, it may be used to discern the origin of enteric bacteria [23]. For example, in the Las Palmas Creek study, a FC/FS ratio in the interval [0, 0.4] was deemed of equine origin, while a FC/FS $\geqslant$ 3.0 was considered to be human in origin. The range 0.4 < FC/FS < 4.0 was associated with mixed origin (i.e., humans, horses, and wildlife in the area where water was tested).

Fig. 4 displays the calculated bivariate distribution $f(x_1, x_2)$, in which $X_1 \equiv$ FC and X_2 = FS. The bivariate distribution was calculated with Eq. (12) using a computer program written in FORTRAN 90 interfaced with the IMSL. In the case of bacterial concentrations in water, FC, FS $\geqslant 0$. Estimates of the six distribution parameters were obtained from the moments of X_1 and X_2, Eqs. (16)–(23), which yielded $\hat{b}_1 = \hat{b}_2 = 1.0$ (the raw data were normalized to render $\hat{b}_1 = \hat{b}_2 = 1.0$); $\hat{\alpha}_1 = 2.471$; $\hat{\alpha}_2 = 8.245$; $\hat{\beta} = 1.417$; $\hat{\gamma} = 0.35$, $\hat{\rho} = 0.40$. The bivariate distribution of FC and FS turned out to be quite asymmetrical in this case (see Fig. 4).

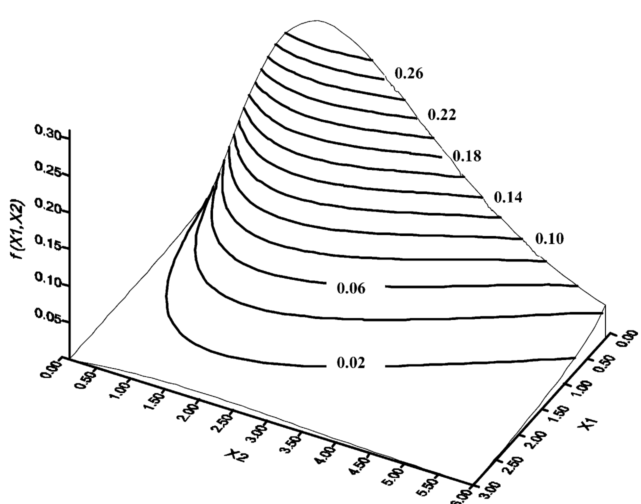

Fig. 4. The bivariate distribution of FC ($\equiv X_1$) and FS ($\equiv X_2$) in Las Palmas Creek, Santa Barbara, CA, 1999–2000. The bivariate distribution was calculated with Eq. (12). The parameters of the distribution are: $\hat{b}_1 = \hat{b}_2 = 1.0$; $\hat{\alpha}_1 = 2.471$; $\hat{\alpha}_2 = 8.245$; $\hat{\beta} = 1.417$; $\hat{\gamma} = 0.35$, $\hat{\rho} = 0.40$. The contour interval in this figure is 0.02.

Fig. 5 shows the empirical (obs × 100) and calculated (model × 100) frequencies of the ratio FC/FS in Las Palmas Creek, Santa Barbara, CA (1999–2000). The empirical frequency in each range was calculated by dividing the number of observations within the range by the sample size (=38), and then scaling it by 100 for ease of interpretation. The model frequency in each range was calculated by integrating Eq. (26) within the range (a FORTRAN 90 program was written to this effect and special functions were evaluated with the IMSL), and then scaling it by 100 for ease of interpretation. In Fig. 5, the range labeled 0.1 equals the interval [0.0, 0.1], that labeled 0.2 =]0.1, 0.2], etc. The last range is $\geqslant$2.0. Fig. 5 shows an overall excellent agreement between the empirical and calculated probabilities. The observed and model-calculated probabilities $P(Z \leqslant 0.4)$ were 71.1% and 66.1%, respectively, which provides strong evidence of the predominance of equine fecal bacteria in Las Palmas Creek.

Table 1 contains the FC/FS data and shows the calculations associated with a chi-square goodness-of-fit test run on the observed and model probabilities graphed in Fig. 5. The chi-square statistic $\chi^2_{19}(0.05) = 30.14$ is

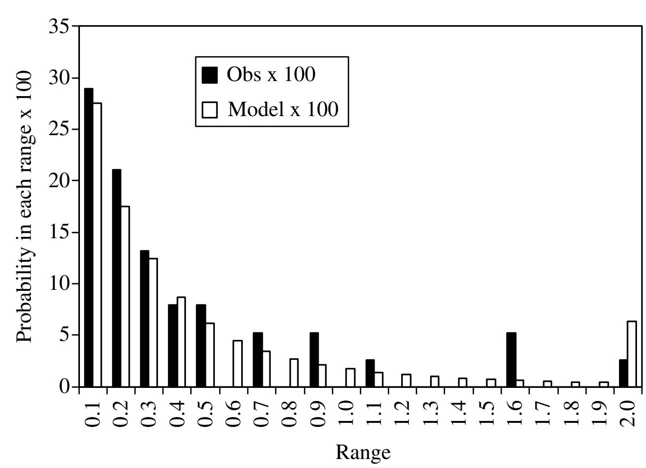

Fig. 5. Empirical (obs × 100) and calculated (model × 100) frequencies of the ratio FC/FS at Las Palmas Creek, Santa Barbara, CA (1999–2000). The empirical frequency in each range was calculated by dividing the number of observations within the range by the sample size (=38), and then scaling it by 100 for ease of interpretation. The model frequency in each range was calculated by integrating Eq. (26) within the range, and then scaling it by 100. The range labeled 0.1 equals the interval [0.0, 0.1], that labeled 0.2 =]0.1, 0.2], etc. The last range is $\geqslant$2.0. FC = fecal coliform concentration; FS = fecal streptococcus concentration.

Table 1
Empirical (FC/FS obs × 100) and calculated (FC/FS model × 100) frequencies of FC/FS in Las Palmas Creek (1999–2000), plus results of the chi-square goodness-of-fit test

Range	FC/FS obs × 100	FC/FS model × 100	$(n/100) \times (\text{obs-model})^2/\text{model}$
0.1	28.9	27.5	0.0289
0.2	21.1	17.5	0.274
0.3	13.2	12.4	0.0176
0.4	7.89	8.67	0.0263
0.5	7.89	6.15	0.188
0.6	0.00	4.50	1.710
0.7	5.26	3.40	0.388
0.8	0.00	2.65	1.01
0.9	5.26	2.11	1.787
1.0	0.00	1.71	0.650
1.1	2.63	1.41	0.401
1.2	0.00	1.17	0.445
1.3	0.00	0.983	0.374
1.4	0.00	0.832	0.316
1.5	0.00	0.708	0.269
1.6	5.26	0.607	13.55
1.7	0.00	0.524	0.199
1.8	0.00	0.454	0.173
1.9	0.00	0.395	0.150
2.0	2.63	6.33	0.821
Sum=	100.0	100.0	22.7

The range labeled 0.1 equals the interval [0.0, 0.1], that labeled 0.2 =]0.1, 0.2], etc. The last range is ⩾2.0. The empirical frequency in each range was calculated by dividing the number of observed values in the range by the sample size ($n = 38$) and then scaling it by 100. The model frequency in each range was calculated by integrating Eq. (26) within the range and then scaling it by 100. FC = fecal coliform concentration; FS = fecal streptococcus concentration.
[1]: the chi-squared statistic $\chi^2_{19}(0.05) = 30.14$ is larger than the test statistic 22.7. Thus, the null hypothesis of a gamma ratio distribution was not rejected at a 5% significance level.

larger than the test statistic 22.7. Thus, the null hypothesis of a gamma ratio distribution was not rejected at a 5% significance level. Notice that the test statistic was calculated using the FC/FS obs. × 100 and FC/FS model × 100 data shown in Table 1.

8. Summary

An bivariate, unsymmetrical, gamma p.d.f. and its moments were developed in this work. The derivation of the bivariate p.d.f. relied on an extension of Kibble's [15] symmetric characteristic function (c.f.) to the unsymmetrical case, coupled with advanced integration techniques applied to the Fourier transform of the c.f. and realized by Laguerre–Charlier (LC) polynomials. From the bivariate p.d.f., the distribution $g(z)$ of the ratio $Z = X_1/X_2$ of two correlated gamma variables and its moments were obtained, also. The derived bivariate gamma and ratio p.d.f.s were illustrated with symmetrical and unsymmetrical examples. A sample of bacterial densities in river water of fecal coliforms (FC) and fecal streptococcus (FS) showed the numerical and graphical fitting of the theoretical ratio p.d.f to such bacteriological data.

Acknowledgement

This work was supported in part by US Geological Survey Grant HQ-96-GR-02657 and by University of California Water Resources Grant WR-952.

References

[1] Becker DJ, Roux JJJ. A bivariate extension of the gamma distribution. South African Statist J 1981;15:1–12.
[2] Charlier CVL. Researches in the theory of probability. K Fysiogr Sällsk Handbook 1906;16.
[3] Charlier CVL. Applications de la théorie des probabilities á la astronomie. In: Borel M (Ed.), Traité du calcul des probabilities, vol. II, part IV. Paris: Gauthier-Villars; 1931.
[4] Clesceri LS, Greenberg AE, Eaton DW. In: Standard methods for the examination of water and wastewater. 20th ed.. Washington, DC: American Public Health Association; 1998.
[5] Cheriyan KC. A bivariate correlated gamma-type distribution function. J Indian Math Soc 1941;5:133–44.
[6] David FN, Fix E. Proceedings of the fourth Berkeley symposium vol. I. Berkeley, CA: University of California Press; 1961. p. 177–87.
[7] D'este GM. A Morgenstern-type bivariate gamma distribution. Biometrika 1981;68:339–40.
[8] Eagleson GK. Polynomial expansions of bivariate distributions. Ann Math Statist 1964;35:1208–15.
[9] Erdelyi A. Tables of integral transforms vol. I. New York: McGraw-Hill; 1954.
[10] Gaver DP. Multivariate gamma distributions generated by mixture. Sankhyâ A 1970;33:123–6.
[11] Gradshteyn IS, Ryzhik IM. In: Tables of integrals, series, and products. 5th ed.. San Diego, CA: Academic Press; 1994.
[12] Gunst RF, Webster TT. Density functions of the bivariate chi-square distribution. J Statist Comput Simulat 1973;2:275–88.
[13] Hinkley DV. On the ratio of two correlated normal random variables. Biometrika 1969;56:635–51.
[14] Kendall M. The advanced theory of statistics vols. I, II, 4th ed. London: Griffin; 1977.
[15] Kibble WF. A two-variate gamma type distribution. Sankhyâ 1941;5:137–50.
[16] Kotlarski I. On bivariate random variables where the quotient of their coordinates follows some known distributions. Ann Math Statist 1964:1673–84.
[17] Kotz S, Balakrishnan N, Johnson NL. Continuous multivariate distributions vol. I, 2nd ed. New York: John Wiley; 2000.
[18] Koutrouvelis IA, Canavos GC. Estimation in the Pearson type III distribution. Water Resour Res 1999;35:2693–704.
[19] Krishnaiah PR, Rao MM. Remarks on a multivariate gamma distribution. Amer Math Monthly 1961;68:392–6.
[20] Krishnamoorthy AS, Parthasarathy M. A multivariate gamma-type distribution. Ann Math Statist 1951;22:549–57.
[21] Lancaster HO. The structure of bivariate distributions. Ann Math Statist 1958;29:719–36.
[22] Lancaster HO. Joint probability distributions in the Meixner classes. J Roy Statist Soc B 1975:435–43.
[23] Loáiciga HA, Renehan S. Bacterial densities in Las Palmas Creek, Hope Ranch, California, 1999–2000. Report 2-2-2000, Depart-

ment of Geography, University of California, Santa Barbara, CA, 2000.

[24] Loáiciga HA. Residence time, groundwater age, and solute output in steady-state groundwater systems. Adv Water Resour 2004; 27:681–8.

[25] Meixner J. Orthogonale polynomsysteme mit einer besonderen gestalt der erzeugenden function. J London Math Soc 1934;9: 6–13.

[26] Moran PAP. Statistical inference with bivariate gamma distributions. Biometrika 1969;56:627–34.

[27] Prékopa A, Szántai T. A new multivariate gamma distribution and its fitting to empirical stream data. Water Resour Res 1978; 14:19–29.

[28] Ramabhadran YR. A multivariate gamma-type distribution. Sankhyâ 1951;11:45–6.

[29] Sarmanov S. A generalized symmetric gamma distribution. Doklady Akad Nauk SSSR 1968;179:1279–85.

[30] Vere-Jones D. The infinite divisibility of a bivariate gamma distribution. Sankhyâ A 1967;27:421–7.

[31] Wicksell SD. On correlation surfaces of type III. Biometrika 1933;25:121–33.

[32] Xekalaki E, Panaretos J, Psarakis S. A predictive model evaluation and selection approach—the correlated gamma ratio distribution. In: Panaretos J, editor. Stochastic musings: Perspectives from the pioneers of the late 20th century. USA: Laurence Erlbaum Publishing; 2003. p. 188–202.