# Information geodesics for gamma models of communication clustering*

C.T.J. Dodson
Department of Mathematics,
University of Manchester Institute of Science and Technology, Manchester M60 1QD, UK

July 20, 1999

**Abstract**

Recently, Akyildiz called for further work on non-Poisson models for communication arrivals in distributed networks such as cellular phone systems. The basic 'random' model for stochastic events is the Poisson process; for events on a line this results in an exponential distribution of intervals between events. Network designers and managers need to monitor and quantify call clustering in order to optimize resource usage; the natural reference state from which to measure departures is that arising from a Poisson process of calls. Here we consider gamma distributions, which contain exponential distributions as a special case. The surface representing gamma models has a natural Riemannian information metric and we obtain some geodesic sprays for this metric. The exponential distributions form a 1-dimensional subspace of the 2-dimensional space of all gamma distributions, so we have an isometric embedding of the random model as a subspace of the gamma models. This geometry may provide an appropriate structure on which to represent clustering as quantifiable departures from randomness and on which to impose dynamic control algorithms to optimize traffic at receiving nodes in distributed communication networks. In practice, we may expect correlation between call arrival times and call duration, reflecting for example peaks of different users of internet services. This would give rise to a twisted product of two surfaces with the twisting controlled by the correlation. Though bivariate gamma models do exist, such as Kibble's, none has tractable information geometry nor sufficiently general marginal gammas, but a simulation method of approach is suggested.

KEYWORDS: GAMMA MODELS, INFORMATION GEOMETRY, COMMUNICATION NETWORKS, ARRIVAL TIME

## Introduction

In his recent detailed reviews of the modeling and management of communication networks, Akyildiz [Akyildiz,1998a, Akyildiz,1998b] drew attention to the need for further work on non-Poisson models for communication arrivals. This article is a response which offers a family of alternatives to Poisson models, and provides a geometric framework in which to study their applicability.

The standard reference structure for queuing problems, such as call arrivals, is a Poisson process with an average interval $\tau$ of time between successive events. In this model the probability that, in a given time interval of length $t$, exactly $m$ calls arrive is

$$P_m = \left(\frac{t}{\tau}\right)^m \frac{1}{m!} e^{-t/\tau} \quad \text{for } m = 0, 1, 2, \ldots \tag{1}$$

Here, the mean and the variance of the *number* of events in such intervals is $t/\tau$. From $P_0$, the probability that no calls arrive in time interval $t$, we obtain the probability density function for the continuous random variable $t$, which represents the lengths of intervals between successive calls,

$$f_{random}(t; \tau) = \frac{1}{\tau} e^{-t/\tau} \quad \text{for } t \in [0, \infty) \tag{2}$$

*To appear in *J. Statistical Computation and Simulation* with variance $Var(t) = \tau^2$. It is normal experience that call arrivals display *clustering* and so depart from the Poisson model (1) of arrivals and hence from the exponential distribution (2) for intervals between calls. We may imagine also situations when calls arrived more regularly than Poisson models, this would indicate temporal *dispersion*—the opposite of clustering, having a variance lower than Poisson.

We describe a family of parametric statistical models that includes (2) as a special case; the other models in the family represent clustering or dispersion of events. This provides a parametric framework on which to represent call arrivals at a communication node. We derive the Riemannian information metric on the parameter space and show that the maximum entropy subspace corresponds to the Poisson case (1). There is in fact a second space that forms a twisted product with the above gamma manifold; together they represent the correlation between call arrival and duration. The importance of curvature in topology and physical applications of Riemannian geometry suggest that it will have significance also in models for call arrivals. The Riemannian geometry setting offers an environment in which optimization of node traffic can be handled naturally.

## Information theory

Jaynes [Jaynes,1957] provided the foundation for information theoretic methods in, among other things, Bayes hypothesis testing—as described by Tribus et al. [Tribus,1961, Tribus et al.,1964]. Given a set of observed values $\{< g_\alpha(t) > | \alpha = 1, 2, \ldots, N\}$ for functions $\{g_\alpha | \alpha = 1, 2, \ldots, N\}$ of the random variable $t$, we seek a 'least prejudiced' set of probability values $\{p_i | i = 1, 2, \ldots, n\}$ satisfying

$$< g_\alpha(t) > \;=\; \sum_{i=1}^{i=n} p_i\, g_\alpha(t_i) \quad \text{for } \alpha = 1, 2, \ldots, N \tag{3}$$

$$1 \;=\; \sum_{i=1}^{i=n} p_i \tag{4}$$

Jaynes showed that this occurs if we choose those $p_i$ that maximize Shannon's entropy function

$$S = -\sum_{i=1}^{i=n} p_i \log(p_i) \tag{5}$$

It turns out [Tribus et al.,1964] that if we have no data on observed functions of $t$, (so the set of equations (3) is empty) then the maximum entropy choice gives the exponential distribution (2). If we have estimates of the first two moments of the distribution of $t$, then we obtain the (truncated) Gaussian. If we have estimates of the mean and mean logarithm of $t$, then the maximum entropy choice is the gamma distribution.

In the sequel we shall consider the particular case of the gamma distribution for several reasons:

- the exponential distributions (2) form a subclass of gamma distributions and exponential distributions are complementary to Poisson processes—the standard random reference processes
- stochastic porous media have been modeled using gamma distributions [Dodson and Sampson,1997].
- the Riemannian geometry of gamma manifolds is tractable

## Gamma models and their geometry

We shall consider the family of gamma distributions with event space $\Omega = \mathbb{R}^+$, parameters $\tau, \nu \in \mathbb{R}^+$ and probability density functions given by

$$f(t; \tau, \nu) = \left(\frac{\nu}{\tau}\right)^\nu \frac{t^{\nu-1}}{\Gamma(\nu)}\, e^{-t\nu/\tau} \tag{6}$$

where $\Gamma$ is the gamma function defined by

$$\Gamma(\nu) = \int_0^\infty s^{\nu-1}\, e^{-s}\, ds, \quad (\text{for positive integer } n,\ \Gamma(n) = (n-1)!)$$

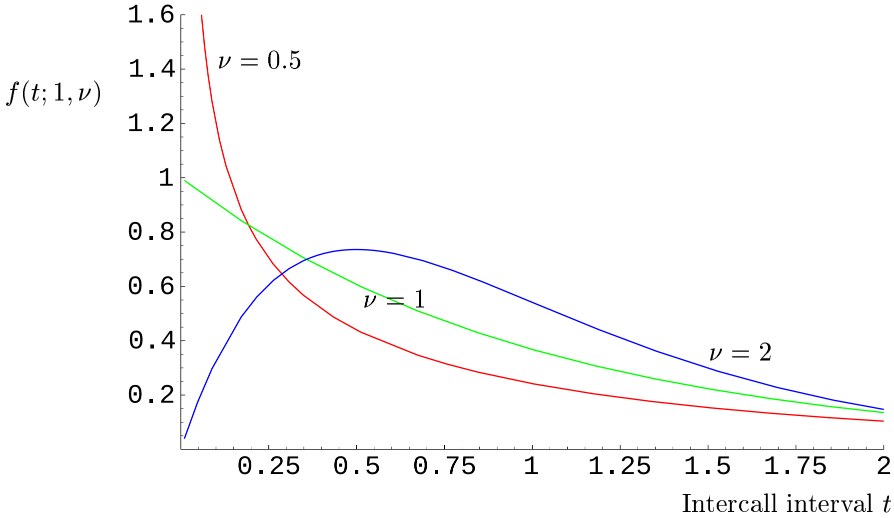


Figure 1: *Probability density functions, $f(t;\tau,\nu)$, for gamma distributions of intercall intervals $t$ with unit mean $\tau = 1$, and $\nu = 0.5,\ 1,\ 2$. The case $\nu = 1$ corresponds to an exponential distribution from an underlying Poisson process.*

Then $< t >= \tau$ is the mean and $Var(t) = \tau^2/\nu$ is the variance; the spread and shape is controlled by, $1/\nu$, the square of the coefficient of variation.

The special case $\nu = 1$ corresponds to the situation when $t$ represents the random or Poisson process in (2) with mean intercall interval $\tau$. In fact, if we put $\lambda = \nu/\tau$, then for integer $\nu = 1, 2, \ldots,$ (6) gives the distribution of intervals between two (Poisson process) events separated by $(\nu - 1)$ other intermediate events. Namely, the gamma distribution (6) is the $\nu$-fold convolution of the exponential distribution; it has traditionally been called also the Pearson Type III distribution [Haight,1967]. Thus, the family of gamma distributions can model a range of stochastic processes corresponding to non-independent clumped or clustered events, for $\nu < 1$, and dispersed events, for $\nu > 1$, as well as the random case. Figure 1 shows a family of gamma distributions, all of unit mean, with $\nu = 0.5,\ 1,\ 2$.

In fact, early this century Erlang (cf. Brockmeyer et al. [Brockmeyer et al.,1960] for reference to English versions of his articles) used gamma distributions as a generalization of the exponential distribution to model telephone call arrivals. There the $\nu$-fold convolution interpretation of (6) was seen as important and viewed as a realisation of a 'censored' Poisson process, namely, $(\nu - 1)$ consecutive points being removed from a standard Poisson process. That would of course correspond to call arrival dispersion for $\nu > 1$. The name Erlang distribution is given to the case when $\nu$ is an integer in (6); several recent applications may be found in the literature [Miller and Bhat,1997, Luceno,1996, Pape,1996].

Shannon's information theoretic entropy or 'uncertainty' for such stochastic processes [Jaynes,1957, Tribus,1961, Tribus et al.,1964] is given, up to a factor, by the negative of the expectation of the logarithm of the probability density function (6), that is

$$S_f(\tau,\nu) = -\int_0^\infty \log(f(t;\tau,\nu))\, f(t;\tau,\nu)\, dt \quad (7)$$

$$= \nu + (1-\nu)\frac{\Gamma'(\nu)}{\Gamma(\nu)} + \log\frac{\tau\,\Gamma(\nu)}{\nu} \quad (8)$$

In particular, at unit mean, the maximum entropy (or maximum uncertainty) occurs at $\nu = 1$, which is the random case, and then $S_f(\tau, 1) = 1 + \log\tau$. Figure 2 shows a plot of $S_f(\tau,\nu)$, for the case of unit mean $\tau = 1$. Another way of saying this is that the Poisson points on a line are as disorderly as possible [Haight,1967]; in fact, among all homogeneous point processes with a given density, the Poisson process has maximum entropy [Rényi,1964].

The 'maximum likelihood' estimates $\hat{\tau}, \hat{\nu}$ of $\tau, \nu$ can be expressed in terms of the mean and mean logarithm of a set of independent observations $X = \{X_1, X_2, \ldots, X_n\}$. These estimates are obtained in terms of the

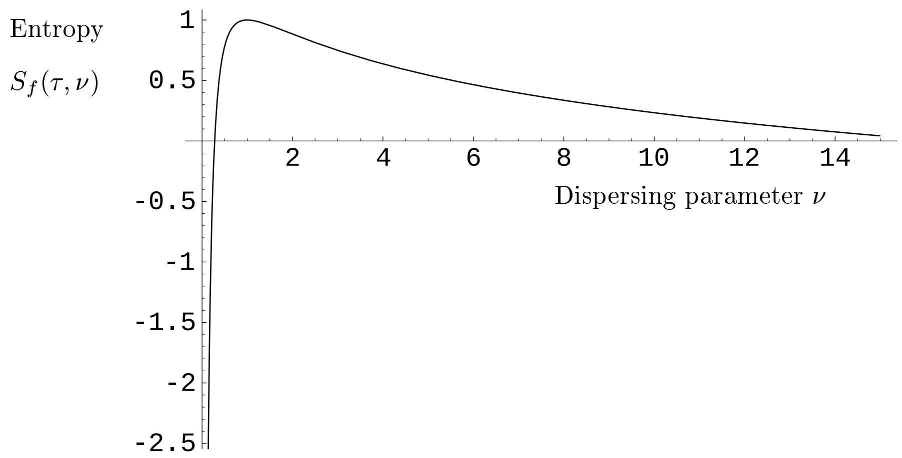


Figure 2: *Information theoretic entropy $S_f(\tau,\nu)$, for gamma distributions of intercall intervals $t$ with unit mean $\tau = 1$. The case $\nu = 1$ corresponds to an exponential distribution from an underlying Poisson process and $\nu < 1$ corresponds to clustering.*

properties of $X$ by maximizing the 'log-likelihood' function

$$\log lik_X(\tau,\nu) = \log\left(\prod_{i=1}^{n} p(X_i;\tau,\nu)\right)$$

with the following result

$$\hat{\tau} \;=\; <X>=\frac{1}{n}\sum_{i=1}^{n} X_i \tag{9}$$

$$\log\hat{\nu} - \frac{\Gamma'(\hat{\nu})}{\Gamma(\hat{\nu})} \;=\; <\log X> - \log <X> \tag{10}$$

where $<\log X>=\frac{1}{n}\sum_{i=1}^{n}\log X_i$.
The usual Riemannian information metric on the parameter space $\mathcal{S} = \{(\tau,\nu)\in\mathbb{R}^+\times\mathbb{R}^+\}$ is given by

$$ds^2_{\mathcal{S}} = \frac{\nu}{\tau^2}\,d\tau^2 + \left(\psi'(\nu) - \frac{1}{\nu}\right)\,d\nu^2 \quad \text{for } \tau,\nu\in\mathbb{R}^+, \tag{11}$$

where $\psi(\nu) = \frac{\Gamma'(\nu)}{\Gamma(\nu)}$ is the digamma function, the logarithmic derivative of the gamma function. For more details see [Lauritzen,1987, Dodson,1998, Dodson and Poston,1991]. The 1-dimensional subspace parametrized by $\nu = 1$ corresponds to the available 'random' processes.
Given any path through the parameter space $\mathcal{S}$ of gamma models, it determines a curve, parametrized by $t$ in some interval $a \le t \le b$, given by

$$c : [a,b] \to \mathcal{S} : t \mapsto (c_1(t), c_2(t)) \tag{12}$$

and its tangent vector $\dot{c}(t) = (\dot{c}_1(t),\dot{c}_2(t))$ has norm $||\dot{c}||$ given via (11) by

$$||\dot{c}(t)||^2 = \frac{c_2(t)}{c_1(t)^2}\,\dot{c}_1(t)^2 + \left(\psi'(c_2(t)) - \frac{1}{c_2(t)}\right)\,\dot{c}_2(t)^2 \tag{13}$$

so the information length of the curve is

$$L_c(a,b) = \int_a^b ||\dot{c}(t)||\,dt. \tag{14}$$

We see that the curve corresponding to Poisson call arrivals has $c(t) = (t, 1)$, which has $t = \tau$ and $\nu = 1 = constant$, the information length is $\log \frac{b}{a}$.

Since arc length is often difficult to evaluate analytically, we sometimes use the 'energy' of the curve instead of length for comparison of information cost differences between nearby curves. Energy is given by integrating the *square* of the norm of $\dot{c}$

$$E_c(a,b) = \int_a^b ||\dot{c}(t)||^2\, dt. \tag{15}$$

so in the case of the curve $c(t) = (t, 1)$, the energy is $\frac{b-a}{ab}$. It is easily shown that a curve of constant $\tau$ has $c(t) = (constant, t)$ where $t = \nu$ and $\dot{c}(t) = (0, 1)$; this has energy $\log \frac{a}{b} + \psi'(b) - \psi'(a)$. Locally, minimal paths in $\mathcal{S}$ are given by the autoparallel curves or geodesics [Dodson and Poston,1991] defined by (11). Some examples of geodesic sprays in the vicinities of the points

$$(\tau, \nu) = (1, 0.5),\ (1, 1),\ (1, 2)$$

are shown in Figure 3. Figure 4 shows a family of curves between fixed endpoints, around an approximately geodesic curve which crosses the line $\nu = 1$ corresponding to all Poisson processes for call arrivals. The geodesics were computed using Gray's Mathematica code for surfaces [Gray,1998].

In the management of a communication network, the control of call clustering is an important activity that presumably must be responsive to capacity and tariffs. The information metric on the space of gamma models offers a way to make decisions based on maximum likelihood conditions and provides a geometric space on which to represent the dynamics. Thus, given some initial state, rate processes specified in terms of parameters $\tau, \nu$ or in terms of mean and variance for example, appear as differential equations with solution curves the possible trajectories of the communication parameters.

The Gaussian curvature of the surface $\mathcal{S}$ is[1]

$$K(\tau, \nu) = \frac{\psi'(\nu) + \nu\, \psi''(\nu)}{4(\nu\, \psi'(\nu) - 1)^2} \quad \text{for } \tau, \nu \in \mathbb{R}^+ \tag{16}$$

$$K(\tau, \nu) \quad \rightarrow \quad -\frac{1}{4} \text{ as } \nu \rightarrow 0 \tag{17}$$

$$K(\tau, \nu) \quad \rightarrow \quad -\frac{1}{2} \text{ as } \nu \rightarrow \infty. \tag{18}$$

The importance of curvature in topological and physical applications of Riemannian geometry [Dodson and Poston,1991] suggest that it will have significance also in models for call arrivals. We see in the next section that there is a second surface that forms a twisted product with the above gamma manifold; together they represent the bivariate gamma models for correlation between call arrival and duration. The properties of this twisted product manifold are as yet unknown.

We should mention also the possibility of autoregressive behaviour and the existence of a new body of geometrical results. Recent work [van Garderen,1999] provides exact geometric expressions for autoregressive time series as examples of curved exponential models, building on that for autoregressive moving average models [Ravishanker et al.,1990].

## Coupling call arrival and duration statistics

The standard probabilistic model for the duration $d$ of calls is again an exponential distribution with mean $< d >= \delta$,

$$h_{random}(d; \delta) = \frac{1}{\delta}\, e^{-d/\delta} \quad \text{for } d \in [0, \infty) \tag{19}$$

with variance $Var(d) = 1/\delta^2$. Now, this is embedded in the family of gamma distributions

$$h(d; \delta, \epsilon) = \left(\frac{\epsilon}{\delta}\right)^{\epsilon} \frac{d^{\epsilon-1}}{\Gamma(\epsilon)}\, e^{-d\epsilon/\delta} \tag{20}$$

[1] This corrects the published expression

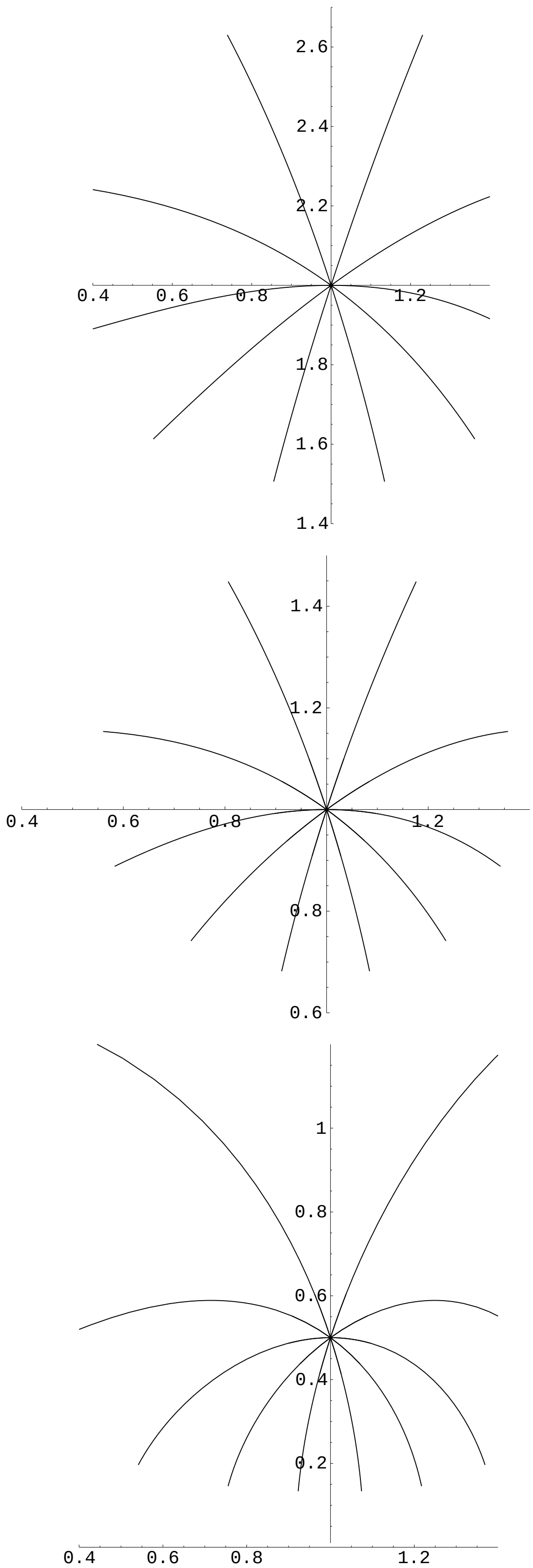


Figure 3: *Geodesic sprays in the gamma manifold, radiating from the points with unit mean* $\tau = 1$*, and* $\nu = 0.5,\ 1,\ 2$ *increasing vertically. The case* $\nu = 1$ *corresponds to an exponential distribution from an underlying Poisson process;* $\nu < 1$ *corresponds to call clustering and* $\nu > 1$ *corresponds to call dispersion processes. The Riemannian metric is given by* $ds_{\mathcal{S}}^2 = \frac{\nu}{\tau^2}\, d\tau^2 + \left(\psi'(\nu) - \frac{1}{\nu}\right)\, d\nu^2$.

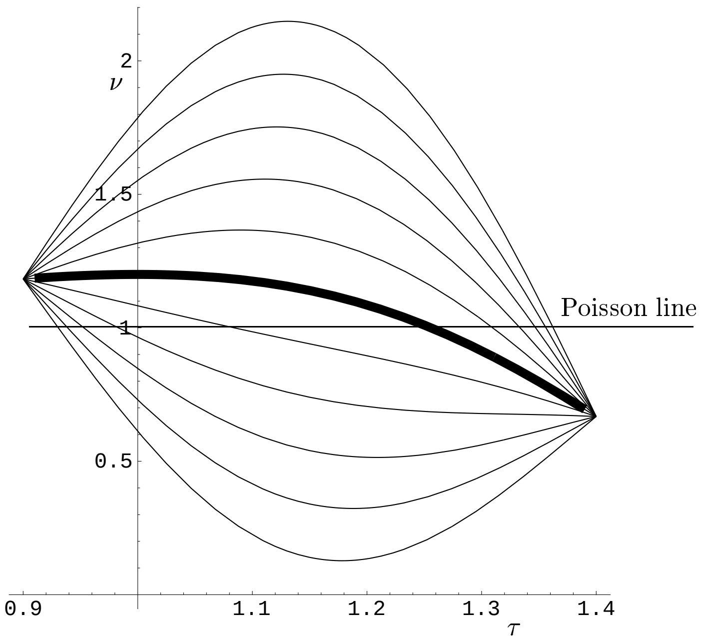


Figure 4: *Families of curves between two endpoints in the gamma manifold. The central bold curve approximates a geodesic that crosses the line $\nu = 1$ corresponding to the set of all Poisson processes for call arrivals.*

Then $< d >= \delta$ and $Var(d) = \delta^2/\epsilon$ and we can consider the geometry of this family in a similar way to that for the intercall intervals. Of course, interesting geometry of the product family would depend on correlation between call arrival intervals and call duration intervals.

The management and performance optimization of communication networks may then be represented through the product geometry of the two stochastic processes: call arrival and call duration. Presumably, initial sampling of a network would reveal whether there was some 'twisting' of the two geometries together as a result of correlation between call frequency and duration. A well known but apparently difficult problem is to find a closed analytic representation of a correlated process for two variables with different marginal gamma distributions. Kibble's bivariate gamma distribution has been used in a variety of applications [Kibble,1941], but from our viewpoint it suffers from the disadvantage that its two marginal gamma distributions have a common dispersion parameter $\nu = \epsilon$. In our case, Kibble's density function takes the following form for $0 \leq \rho < 1$

$$g(t,d) = \left(\frac{\nu^2}{\tau\delta}\right)^{\nu} \left(\frac{td\tau\delta}{\nu^2\rho}\right)^{(\nu-1)/2} \frac{e^{-\nu(t/\tau+d/\delta)/(1-\rho)}}{(1-\rho)\,\Gamma(\nu)} I_{\nu-1}\left(\frac{2\nu\sqrt{\rho td/\tau\delta}}{1-\rho}\right) \tag{21}$$

where $I_{\nu-1}$ is the modified Bessel function of the first kind of order $(\nu-1)$, which appears here in the form

$$I_{\nu-1}\left(\frac{2\nu\sqrt{\rho td/\tau\delta}}{1-\rho}\right) = \sum_{k=0}^{\infty} \frac{(\rho td/\tau\delta)^{k+(\nu-1)/2}}{k!\,\Gamma(\nu+k)} \left(\frac{\nu}{1-\rho}\right)^{\nu-1+2k} \tag{22}$$

Simulation of correlated call arrival intervals and call duration intervals could provide sequences of coupled processes with gamma distributions as marginal distributions, parametrized by the correlation. This would allow an experimental approach to the analytic problem and linking with such real data as may be obtained from existing networks.

Some special cases have been studied by others for various applications. Downton [Downton,1970] used a bivariate exponential distribution and compared it in application to reliability theory with Marshall and Olkin's distribution [Marshall and Olkin,1967]. Ong [Ong,1995] has provided some expansions for the corresponding cumulative distribution function. Shi and Lai [Shi and Lai,1998] provided the Fisher information matrix components for the bivariate exponential distribution as a special case of Kibble's bivariate gamma distribution (21) but still had to resort to a numerical evaluation of one of the integrals. So it would be difficult to apply the Fisher matrix to develop the Riemannian geometry of the parametric models even in the correlated exponential case.

A related problem, to find a multivariate Poisson-type distribution with Poisson marginals, has also received attention [Haight,1967] and more recently for example [Foss and Zuyev,1996, Adamidis and Loukas,1996, Wesolowski,1995, Lee,1996, Lai,1995, Kocherlakota,1995]. Guldberg [Guldberg,1934] solved the bivariate case (cf. also eg. [Holgate,1964]) in the form

$$P_{m,n} = e^{(-\lambda-\mu-\eta)} \frac{(\lambda-\eta)^m(\mu-\eta)^n}{m!n!} \left(1 + \frac{mn}{(\lambda-\eta)(\mu-\eta)}\frac{\eta}{1!} + \frac{m(m-1)n(n-1)}{(\lambda-\eta)^2(\mu-\eta)^2}\frac{\eta^2}{2!} + \cdots\right) \tag{23}$$

for $m, n = 0, 1, 2, \ldots$ In another direction [Lampard,1968] there is a generalized renewal process model for coupled successive intervals arising from two Poisson processes and Fisher information measure has been used to define an efficiency criterion for periodic observation of gamma-distributed intervals [Miller and Bhat,1997].

## Acknowledgement

The author wishes to thank UMIST and the Peter Allen fund for support with a travel grant, and C.D. Lai and S.H. Ong for helpful comments.

## References

[Adamidis and Loukas,1996] K. Adamidis and S. Loukas. Estimation in a truncated bivariate Poisson distribution using the EM algorithm. *Commun. Statist. Theory and Methods* 25, 10 (1996) 2215-2222.

[Akyildiz,1998a] Akyildiz I. Mobility management in current and future communication networks. *IEEE Network Mag.* 12, 4 (1998) 39-49.

[Akyildiz,1998b] Akyildiz I. Performance modeling of next generation wireless systems, Keynote Address, Conference on Simulation Methods and Applications, 1-3 November 1998, Orlando, Florida.

[Brockmeyer et al.,1960] Brockmeyer E., H.L. Halstrøm and A. Jensen. **The Life and Works of A.K. Erlang** Danish Academy of Technical Sciences, Copenhagen, 1960.

[Dodson,1998] Dodson C.T.J. Gamma manifolds and stochastic geometry. In: **Proceedings of the Workshop on Recent Topics in Differential Geometry**, Santiago de Compostela 16-19 July 1997. *Public. Depto. Geometría y Topología* 89 (1998) 85-92.

[Dodson and Poston,1991] C.T.J. Dodson and T. Poston. **Tensor Geometry** Graduate Texts in Mathematics 130, Second edition, Springer-Verlag, New York, 1991.

[Dodson and Sampson,1997] Dodson C.T.J. and W.W. Sampson. Modeling a class of stochastic porous media. *Appl. Math. Lett.* 10, 2 (1997) 87-89.

[Downton,1970] Downton F. Bivariate exponential distributions in reliability theory. *J. Royal Statist. Soc. Series B* 32 (1970) 408-417.

[Foss and Zuyev,1996] Foss S.G. and S.A. Zuyev. On a Voronoi aggregative process related to a bivariate Poisson process. *Adv. Appl. Prob.* 28, 4 (1996) 965-981.

[Gray,1998] Gray A. **Modern Differential Geometry of Curves and Surfaces** $2^{nd}$ Edition, CRC Press, Boca Raton 1998.

[Griffiths,1969] Griffiths R.C. The canonical correlation coefficients of bivariate gamma distributions. *Annals Math. Statist.* 40, 4 (1969) 1401-1408.

[Guldberg,1934] Guldberg A. On discontinuous frequency functions of two variables. *Skandinavisk Aktuarietidskrift* 17 (1934) 89-117.

[Haight,1967] Haight F.A. **Handbook of the Poisson Distribution** J. Wiley, New York, 1967.

[Holgate,1964] Holgate P. Estimation for the bivariate Poisson process. *Biometrika* 51 (1964) 241-245.

[Jaynes,1957] Jaynes E.T. Information theory and statistical inference. *The Physical Review* 106 (1957) 620-630 and 108 (1957) 171-190.

[Kibble,1941] Kibble W.F. A two variate gamma-type distribution. *Sankhyā* 5 (1941) 137-150.

[Kocherlakota,1995] Kocherlakota S. Discrete bivariate weighted distributions under multiplicative weight function. *Commun. Statist. Theory and Methods* 24, 2 (1995) 553-551.

[Lai,1995] Lai C.D. Constructions of bivariate distributions by a generalized trivariate reduction technique. *Statistics and Probability Letters* 25, 3 (1995) 265-270.

[Lampard,1968] Lampard D.G. A stochastic process whose successive intervals between events form a first order Markov chain–I. *J. Appl. Prob.* 5 (1968) 648-668.

[Lauritzen,1987] Lauritzen S.L. Statistical Manifolds. In **Differential Geometry in Statistical Inference**, Institute of Mathematical Statistics Lecture Notes, Volume 10, Berkeley 1987, pp 163-218.

[Lee,1996] Lee M.L.T. Properties and applications of the Sarmanov family of bivariate distributions. *Commun. Statist. Theory and Methods* 25, 6 (1996) 1207-1222.

[Luceno,1996] Luceno A. A generalized Erlang distribution showing overdispersion. *Statistics and Probability Letters* 28, 4 (1996) 375-386.

[Marshall and Olkin,1967] Marshall A.W. and I. Olkin. A generalized bivariate exponential distribution. *J. Appl. Prob.* 4 (1967) 291-302.

[Miller and Bhat,1997] Miller G.K. and U.N. Bhat. Estimation for renewal processes with unobservable gamma or Erlang interarrival times. *J. Statistical Planning and Inference* 61, 2 (1997) 355-372.

[Ong,1995] Ong S.H. Computation of bivariate-gamma and inverted-beta distribution functions. *J. Statistal Computation and Simulation* 51, 2-4 (1995) 153-163.

[Pape,1996] Pape E.S. Erlang sampling of an alternating Poisson process. *Commun. Statist. Theory and Methods* 25, 6 (1996) 1283-1288.

[Ravishanker et al.,1990] Ravishanker N., E.L. Melnick and C-L. Tsai. Differential geometry of ARMA models. *J. Time Series Analysis* 11 (1990) 259-274.

[Rényi,1964] Rényi A. On an extremal property of the Poisson process. *Annals Inst. Stat. Math.* 16 (1964) 129-133.

[Shi and Lai,1998] Shi D. and C.D. Lai. Fisher information for Downton's bivariate exponential distribution. *J. Statistical Computation and Simulation* 60, 2 (1998) 123-127.

[Tribus,1961] Tribus M. **Thermostatics and Thermodynamics** D. Van Nostrand and Co., Princeton N.J., 1961.)

[Tribus et al.,1964] Tribus M., R. Evans and G. Crellin. The use of entropy in hypothesis testing. In Proc. **Tenth National Symposium on Reliability and Quality Control** 7-9 January 1964.

[van Garderen,1999] van Garderen K.J. Exact geometry of autoregressive models. *J. Time Series Analysis* 20, 1 (1999) 1-21.

[Wesolowski,1995] Wesolowski J. Bivariate discrete measures via a power series conditional distribution and a regression. *J. Multivariate Analysis* 55, 2 (1995) 219-229.