Ann. Inst. Statist. Math.
Vol. 49, No. 2, 255–269 (1997)

TESTING DEPARTURES FROM GAMMA, RAYLEIGH AND TRUNCATED NORMAL DISTRIBUTIONS

JOAN DEL CASTILLO AND PEDRO PUIG

Unitat d'Estadistica i Investigació Operativa, Departament de Matemàtiques, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain

(Received May 15, 1995; revised June 27, 1996)

Abstract. This paper provides necessary and sufficient conditions for a solution to likelihood equations for an exponential family of distributions, which includes Gamma, Rayleigh and singly truncated normal distributions. Furthermore, the maximum likelihood estimator is obtained as a limit case when the equations have no solution. These results provide a way to test departures from Rayleigh and singly truncated normal distributions using the likelihood ratio test. A new easy way to test departures from a Gamma distribution is also introduced.

Key words and phrases: Exponential families, Gamma, Rayleigh and singly truncated normal distributions.

1. Introduction

Several kinds of statistical models are used to approximate data on non-negative real numbers. The Gamma, Rayleigh and singly truncated normal distributions are some examples which at the same time are exponential families. These are distributions of maximum entropy associated with some sufficient statistics and therefore they frequently appear in many problems.

The main goal of this paper is to consider a larger family that includes these three families in order to get a more general way of fitting non-negative data and, moreover, to be able to test departures from the three classical families against this more general family, using the maximum likelihood estimators and likelihood ratio test.

Barndorff-Nielsen and Cox (1979) proposed to introduce the statistic x^2 as a dispersion measure in order to test whether the data are consistent with the assumption that they come from a Gamma distribution, which is an exponential family generated by the statistics $(x, \mathrm{Log}(x))$. We consider the exponential family, say $\boldsymbol{P}$, generated by the statistic $(x, x^2, \mathrm{Log}(x))$. This family is a system of asymmetric bell-shaped distributions which include and generalise the Gamma, Rayleigh and singly truncated normal distributions. The family $\boldsymbol{P}$ had already been introduced by Toranzos (1952) as the solution to a differential equation, in a similar way to Pearson's system (see also Johnson and Kotz (1970)).

Toranzos (1952) uses a recurrent formula for the moments, given in (3.1), to provide an estimation of the parameters of the family $\boldsymbol{P}$. This formula is closely related to Ricatti's differential equation and it can be used to describe and compute the cumulant generating function of the distributions. However, the estimation method is unsatisfactory because it can produce estimates outside the domain of the parameters, as we shall illustrate with an example.

Family $\boldsymbol{P}$ is also a particular case of one of the four types of exponential families introduced in Cobb *et al.* (1983), as the stationary probability density functions of a nonlinear diffusion process. These families are also the solutions of differential equations, in a similar way to Pearson's system. The paper by Cobb *et al.* (1983) gives recurrent formulae for the moments and, from them, exhibits consistent asymptotically normal estimators of the parameters.

Here we study the likelihood equations of the family $\boldsymbol{P}$ from the point of view of exponential families. In this context, see Barndorff-Nielsen (1978), for a regular or a steep exponential family there is one and only one solution to likelihood equations, but $\boldsymbol{P}$ is a non-steep family. The main results of the paper, given in Theorems 4.1 and 4.2, provide necessary and sufficient conditions for a solution to likelihood equations for $\boldsymbol{P}$ and for certain sub-families $\boldsymbol{P}_\nu$ with ν fixed. When ν is larger than 1, these sub-families have an increasing failure rate and can be used in survival analysis, see Ross (1983). Moreover, when ν is equal to 1 we have the truncated normal family and therefore our results generalise some of Castillo's (1994).

In Subsection 6.1, we give an algorithm to apply the Newton-Raphson method to the likelihood equations in the reduced form given in (3.11). We include a graph, in terms of the coefficient of variation and the log-ratio between arithmetical and geometrical means, which can be used to find an initial value for the parameters. Estimators (3.2) from recursive moment equations could also be considered as starting points when they belong to the natural domain of parameters D.

Using maximum likelihood estimation for family $\boldsymbol{P}$, we can test departures from Rayleigh or from truncated normal distributions with the likelihood ratio test, which is highly accurate (see McCullagh (1987) or Barndorff-Nielsen and Cox (1989)). However, the Gamma distribution appears as a limit case in family $\boldsymbol{P}$ and therefore the standard theory of likelihood does not work properly. In this case, Theorem 5.1 provides a new easy way to test departures from Gamma distribution which is closely related to the main results of the paper. Subsection 6.2 shows that this new test can be more powerful than the classical EDF-tests A^2, W^2 and U^2 for some alternatives even not included in family $\boldsymbol{P}$.

2. The statistical model

Some families of probability density functions are introduced by generating a differential equation of the form:

$$\frac{y'}{y} = -\frac{g(x)}{v(x)}, \tag{2.1}$$

where $g(x)$ and $v(x)$ are polynomials. Pearson's system is a particular case, when $g(x)$ has degree 1 and $v(x)$ has degree 2. Cobb *et al.* (1983) consider four important cases on (2.1), assigning an arbitrary degree to $g(x)$ and setting $v(x)$ equal to 1, x, x^2 or $x(1-x)$. The family introduced by Toranzos (1952) is a particular case with $g(x)$ of degree 2 and $v(x) = x$.

The integral of rational functions can always be expressed as a linear combination of functions such as:

$$(x-a)^r, \qquad \mathrm{Log}(x-b), \qquad \mathrm{Atan}(x^2+px+q), \ldots. \tag{2.2}$$

Assuming the latter constants are known, the solution of (2.1) is always an exponential family distribution with sufficient statistics given by the functions in (2.2). In particular, the family introduced by Toranzos is the exponential family generated by the sufficient statistic $T(x) = (x, x^2, \mathrm{Log}(x))$, with respect to the Lebesgue measure defined only for non-negative values of x. We shall express these probability density functions in the form:

$$p(x;\alpha,\beta,\nu) = x^{\nu-1}\exp(-\alpha x - \beta x^2)/C(\alpha,\beta,\nu), \qquad (x>0), \tag{2.3}$$

where the normalizing constant is the Laplace Transform

$$C(\alpha,\beta,\nu) = \int_0^\infty x^{\nu-1}\exp(-\alpha x - \beta x^2)dx \tag{2.4}$$

which converges for $\alpha \in \mathbb{R}$, $\beta > 0$, $\nu > 0$ and $\alpha > 0$, $\beta = 0$, $\nu > 0$.

Let $\Theta = \{\theta = (\alpha,\beta,\nu)' : \alpha \in \mathbb{R}, \beta > 0, \nu > 0\}$, $\Theta_0 = \{\theta = (\alpha, 0, \nu)' : \alpha > 0, \nu > 0\}$ and $D = \Theta \cup \Theta_0$. We call D the *natural domain of parameters* for the statistical model $\boldsymbol{P}$ of distributions given by (2.3). By $\boldsymbol{P}_{..\nu}$ (or simply $\boldsymbol{P}_\nu$), we denote the distribution of $\boldsymbol{P}$ with ν fixed and, in the same way, $\boldsymbol{P}_{\alpha..}$ and $\boldsymbol{P}_{.\beta.}$ are the distributions of $\boldsymbol{P}$ with α or β fixed.

The $\boldsymbol{P}$-model includes the Gamma distributions, $\boldsymbol{P}_{.0.}$, as the limit case $\beta = 0$. Then the x^2 statistic can be used as a dispersion measure in order to test whether the data are consistent with the assumption that they come from a Gamma distribution. This idea was considered in Barndorff-Nielsen and Cox (1979).

The sub-model $\boldsymbol{P}_{0..}$, corresponding to $\alpha = 0$, is just the Rayleigh family of distributions used in quality control. In fact the model $\boldsymbol{P}$ is the set of all distributions that can be obtained by *exponential tilting* from the family of Rayleigh (see Barndorff-Nielsen and Cox (1989)).

Another particular case is the sub-model $\boldsymbol{P}_1$, the left truncated normal distributions. The moment-generating function of $\boldsymbol{P}_n$, where n is a positive integer, can be obtained by taking derivatives of the moment-generating function of $\boldsymbol{P}_1$.

The shape of the probability density functions depends on the values of the parameters. When $\nu > 1$, they are unimodal and vanish at zero and at infinity; when $\nu < 1$, they tend to infinity at zero and are always decreasing or have a local maximum depending on the values of α and β.

The sub-model $\boldsymbol{P}_\nu$ with $\nu > 1$ is useful in survival analysis because in this situation the logarithm of the probability density functions is concave and the $\boldsymbol{P}_\nu$ are distributions of increasing failure rate (IFR) random variables (see Ross (1983)).

3. The estimating problem

In the same way as Cobb *et al.* (1983), from (2.3), (2.4) and integration by parts we find the following recurrent formula for the theoretical moments μ_k:

$$2\beta\mu_{k+2} + \alpha\mu_{k+1} = (\nu + k)\mu_k, \qquad k = 0, 1, 2, \dots \tag{3.1}$$

where $\mu_0 = 1$. Note that this allows us to calculate all the theoretical moments once μ_1 is known.

Toranzos (1952) has already remarked that by identifying the theoretical with the empirical moments and successively giving the values 0, 1, 2, to k we may fit the parameters by solving a linear system of equations. Given a random sample $x = (x_1, \dots, x_n)'$ with empirical moments $m_k = \frac{1}{n}\sum x_i^k$, the solution of the linear system is given by:

$$\begin{aligned}\tilde{\alpha} &= (2m_2m_3 - m_1m_4 - m_1m_2^2)/\Delta,\\ \tilde{\beta} &= (m_1^2m_2 - 2m_2^2 + m_1m_3)/(2\Delta),\\ \tilde{\nu} &= (-2m_2^3 + 3m_1m_2m_3 - m_1^2m_4)/\Delta,\end{aligned} \tag{3.2}$$

where $\Delta = m_2^3 - 2m_1m_2m_3 + m_3^2 + m_1^2m_4 - m_2m_4$. This method of estimation is not the method of moments. Note that we use four moments to estimate three parameters. These estimators are consistent and asymptotically normal (see Cobb *et al.* (1983)). This method is simple but not very good because very often the estimated parameters do not belong to D, as we shall see in Section 5. Moreover, since $\boldsymbol{P}$ is an exponential family, the maximum likelihood estimate is the best estimation procedure. For instance, it depends on T, which is a sufficient statistic, whereas the first moments are not.

Let $K(\alpha, \beta, \nu) = \mathrm{Log}(C(\alpha, \beta, \nu))$ be the *cumulant function* of $\boldsymbol{P}$. The log-likelihood function of the model can then be written:

$$l(\theta) = l(\alpha, \beta, \nu; x) = n(-\alpha t_1 - \beta t_2 + (\nu - 1)t_3 - K(\alpha, \beta, \nu)) \tag{3.3}$$

where $(t_1, t_2, t_3)' = t = \sum T(x_i)/n$. Note that $t_1 = m_1$, $t_2 = m_2$, $t_3 = \sum \mathrm{Log}(x_i)/n$. We define the mean value mapping: $\tau = (\tau_1, \tau_2, \tau_3)'$, on Θ, by $\tau(\theta) = E_\theta(T)$. This mapping is one to one and continuously differentiable both ways between Θ and $T = \tau(\Theta)$ (see Barndorff-Nielsen (1978)).

Since $E_\theta(\frac{\partial l}{\partial \theta}) = 0$, the likelihood equations of the model $\boldsymbol{P}$ are

$$\tau(\theta) = t, \qquad \theta \in \Theta. \tag{3.4}$$

When $t \in T$, equation (3.4) have only one solution. Note that with our definitions $\tau(\theta) \neq \frac{\partial K}{\partial \theta}$ because of the signs that appear in (2.3). However, Fisher information matrix $F(\theta) = E_\theta[-\frac{\partial^2 l}{\partial\theta\partial\theta'}] = \frac{\partial^2 K}{\partial\theta\partial\theta'}$, agrees with the natural parameterizations in exponential families.

The models $\boldsymbol{P}$ and $\boldsymbol{P}_\nu$ are invariant when the scale of the data changes. That is, if X is a random variable with probability density function $p(x; \alpha, \beta, \nu)$, then

λX, for $\lambda > 0$, has probability density function $p(x; \alpha/\lambda, \beta/\lambda^2, \nu)$. This fact allow us to reduce the likelihood equations. To this end, we introduce the functions:

$$\varphi_\nu^s(\gamma) = \int_0^\infty x^{\nu-1} \log^s(x) \exp(-x^2 - \gamma x) dx, \qquad \nu > 0, \ s \geq 0, \ \gamma \in \mathbb{R}. \tag{3.5}$$

Notice the differentiation properties of these functions

$$\frac{\partial}{\partial \gamma} \varphi_\nu^s(\gamma) = -\varphi_{\nu+1}^s(\gamma), \qquad \frac{\partial}{\partial \nu} \varphi_\nu^s(\gamma) = \varphi_\nu^{s+1}(\gamma). \tag{3.6}$$

In numerical approaches the following expression can be used, obtained from (3.5) using Taylor's expansion for $\exp(-\gamma x)$ and taking integrals:

$$\varphi_\nu^s(\gamma) = \frac{1}{2^{s+1}} \sum_{k=0}^\infty \frac{(-\gamma)^k}{k!} \Gamma^{(s)} \left(\frac{\nu + k}{2} \right), \tag{3.7}$$

where $\Gamma^{(s)}$ is the s-th derivative of the Gamma Function.

We denote φ_ν^0 by φ_ν. Notice that $\varphi_\nu(\gamma) = C(\gamma, 1, \nu)$. Then using (3.1) with $k = 0$ we find the recurrent formula:

$$2\varphi_{\nu+2}(\gamma) + \gamma \varphi_{\nu+1}(\gamma) - \nu \varphi_\nu(\gamma) = 0. \tag{3.8}$$

This formula and the properties (3.6) state that φ_ν is a solution of the Ricatti equation:

$$2\ddot{\varphi}_\nu(\gamma) - \gamma \dot{\varphi}_\nu(\gamma) - \nu \varphi_\nu(\gamma) = 0,$$

where dot means differentiation with respect to γ. The solution can be expressed in terms of Kummer's Function $R(a, b, c)$ according to:

$$\varphi_\nu(\gamma) = \sqrt{\pi}\Gamma(\nu) 2^{-\nu} \cdot \left\{ R\left(\frac{\nu}{2}, \frac{1}{2}, \frac{\gamma^2}{4}\right) \Big/ \Gamma\left(\frac{\nu+1}{2}\right) - \gamma R\left(\frac{\nu+1}{2}, \frac{3}{2}, \frac{\gamma^2}{4}\right) \Big/ \Gamma\left(\frac{\nu}{2}\right) \right\}.$$

This is the solution of Kummer's Differential Equation, $c\frac{d^2R}{dc^2} + (b - c)\frac{dR}{dc} - aR = 0$. For properties and more information about this function, see Abramowitz and Stegun (1972).

The asymptotic behaviour of $\varphi_\nu(\gamma)$ when γ tends to $\pm\infty$ and of $\varphi_\nu^1(\gamma)$ when γ tends to $-\infty$ will be needed to prove the main results. The following asymptotic expansions can be obtained using Watson's lemma (see Appendix) and the classical methods described in Barndorff-Nielsen and Cox (1989) and Breitung (1994):

$$\varphi_\nu(\gamma) \sim \frac{\Gamma(\nu)}{\gamma^\nu} \left\{ 1 - \frac{\nu(\nu+1)}{\gamma^2} + o\left(\frac{1}{\gamma^2}\right) \right\}, \qquad \text{as} \quad \gamma \to +\infty,$$

$$\varphi_\nu(\gamma) \sim \sqrt{\pi} \left(-\frac{\gamma}{2}\right)^{\nu-1} \exp\left(\frac{\gamma^2}{4}\right) \left\{ 1 + \frac{(\nu-1)(\nu-2)}{\gamma^2} + o\left(\frac{1}{\gamma^2}\right) \right\}, \qquad \text{as} \quad \gamma \to -\infty, \tag{3.9}$$

$$\varphi_\nu^1(\gamma) \sim \sqrt{\pi}\left(-\frac{\gamma}{2}\right)^{\nu-1} \exp\left(\frac{\gamma^2}{4}\right)$$
$$\cdot\left\{\mathrm{Log}\left(-\frac{\gamma}{2}\right) + \frac{(\nu-1)(\nu-2)}{\gamma^2}\,\mathrm{Log}(-\gamma) + O\left(\frac{1}{\gamma^2}\right)\right\},$$
$$\text{as} \quad \gamma \to -\infty.$$

From (2.4) the cumulant function is now:

$$K(\alpha,\beta,\nu) = -\frac{\nu}{2}\,\mathrm{Log}(\beta) + \mathrm{Log}[\varphi_\nu(\alpha/\sqrt{\beta})].$$

Taking derivatives, the likelihood equations are:

$$\tau_1(\alpha,\beta,\nu) = \frac{1}{\sqrt{\beta}}\frac{\varphi_{\nu+1}}{\varphi_\nu}(\alpha/\sqrt{\beta}) = t_1,$$
$$\tau_2(\alpha,\beta,\nu) = \frac{\nu}{2\beta} - \frac{\alpha}{2\beta^{3/2}}\frac{\varphi_{\nu+1}}{\varphi_\nu}(\alpha/\sqrt{\beta}) = t_2, \tag{3.10}$$
$$\tau_3(\alpha,\beta,\nu) = \mathrm{Log}(1/\sqrt{\beta}) + \frac{\varphi_\nu^1}{\varphi_\nu}(\alpha/\sqrt{\beta}) = t_3.$$

Now, equation (3.10) can be written in a more suitable form:

$$\tau_1(\alpha,\beta,\nu) = t_1,$$
$$\mathrm{CV}^2(\alpha,\beta,\nu) = c^2,$$
$$\mathrm{RM}(\alpha,\beta,\nu) = r,$$

where $c^2 = t_2/t_1^2 - 1$ is the empirical squared coefficient of variation, $r = \mathrm{Log}(t_1) - t_3$ is the logarithm of the ratio between the arithmetical and geometrical means, $\mathrm{CV}^2 = \tau_2/\tau_1^2 - 1$ and $\mathrm{RM} = \mathrm{Log}(\tau_1) - \tau_3$. Note that CV^2 is the theoretical squared coefficient of variation and RM is the log-ratio between the theoretical arithmetical and geometrical means. Due to the invariance with respect to changes of scale, CV^2 and RM depend only on ν and $\gamma = \alpha/\sqrt{\beta}$, and we can reduce (3.10) to

$$\mathrm{CV}_*^2(\gamma,\nu) = \mathrm{CV}^2(\gamma,1,\nu) = \frac{1}{2}\left(\nu\left[\frac{\varphi_\nu}{\varphi_{\nu+1}}(\gamma)\right]^2 - \gamma\frac{\varphi_\nu}{\varphi_{\nu+1}}(\gamma)\right) - 1 = c^2,$$
$$\mathrm{RM}_*(\gamma,\nu) = \mathrm{RM}(\gamma,1,\nu) = \mathrm{Log}\left(\frac{\varphi_{\nu+1}}{\varphi_\nu}(\gamma)\right) - \frac{\varphi_\nu^1}{\varphi_\nu}(\gamma) = r. \tag{3.11}$$

Once we solve the system of equation (3.11) we can obtain the estimates of the original parameters α and β by using the first equation in (3.10) and the definition of γ.

In the same way we can obtain the likelihood equations for the family $\boldsymbol{P}_\nu$. In fact these are the first two in (3.10), and the reduced form is only the first equation in (3.11).

Notice that family $\boldsymbol{P}$ is not *regular* because $D \neq \Theta$ (D is not open), but it is also *non-steep*, as we shall see in the next section. Cobb *et al.* (1983), do not seem to have been aware of the fact that the family $\boldsymbol{P}$ is not steep and hence the likelihood equations may have no solution. The same situation occurs for families $\boldsymbol{P}_\nu$. For the Normal Truncated Family $\boldsymbol{P}_1$, in particular, this problem is studied in Castillo (1994).

4. Characterisation of the domain of means

Let $\boldsymbol{X} = \mathbb{R}^+$ be the sample space. Given any sample $x = (x_1, \ldots, x_n)'$ the statistic $t = \sum T(x_i)/n$ belongs to the closed convex hull of $T(\boldsymbol{X})$, which we denote by $\boldsymbol{C}$. Moreover, any neighbourhood of any point in $\boldsymbol{C}$ has non-zero probability of including t when the sample x comes from a distribution of $\boldsymbol{P}$. Proposition 4.1, proved in the Appendix, characterizes $\boldsymbol{C}$ for $\boldsymbol{P}$ and, in the same way, the set $\boldsymbol{C}_\nu$ for $\boldsymbol{P}_\nu$.

PROPOSITION 4.1. *The closed convex hull of $T(\boldsymbol{X})$, for the model $\boldsymbol{P}$, is the set of points $t = (t_1, t_2, t_3)'$ defined by $\boldsymbol{C} = \{t : t_1 \geq 0, t_2 \geq t_1^2, \mathrm{Log}(t_1) - t_3 \geq 0\}$. In the same way, for the model $\boldsymbol{P}_\nu$ and its corresponding sufficient statistic, the closed convex hull is $\boldsymbol{C}_\nu = \{t : t_1 \geq 0, t_2 \geq t_1^2\}$.*

Notice that for any sample of non-negative numbers the mean, the mean of squares and the mean of the logarithm of the data define a point in $\boldsymbol{C}$.

On the other hand, for a full exponential family we can solve the likelihood equation (3.4) only when t belongs to the domain of means T, always enclosed in $\boldsymbol{C}$. An exponential family is said to be *steep* if $|\tau(\theta)|$ tends to infinity as θ tends to any boundary point of the natural domain of parameters. In this situation the domain of means is exactly equal to the interior of $\boldsymbol{C}$ and the likelihood equations always have a solution (see Barndorff-Nielsen (1978)).

Proposition 4.2, proved in the Appendix, shows that the family $\boldsymbol{P}$ is a non-steep family.

PROPOSITION 4.2. *The functions τ_i defined on Θ extend continuously to Θ_0. Moreover, if $\alpha > 0$ and β tends to 0 then*

$$\tau_1(\alpha, \beta, \nu) \to \nu/\alpha, \qquad \tau_2(\alpha, \beta, \nu) \to \nu(\nu + 1)/\alpha^2, \qquad \tau_3(\alpha, \beta, \nu) \to \frac{\Gamma'}{\Gamma}(\nu) - \mathrm{Log}(\alpha).$$

Moreover, as a consequence of this proposition, it follows that $\boldsymbol{P}_\nu$ is a non-steep family for any $\nu > 0$; $\boldsymbol{P}_{\alpha..}$ is regular for $\alpha \leq 0$ and non-steep for $\alpha > 0$; and for any $\beta \geq 0$, $\boldsymbol{P}_{.\beta.}$ is always regular.

The following theorems determine the domain of means for families $\boldsymbol{P}_\nu$ and $\boldsymbol{P}$, and so they establish the conditions under which the likelihood equations have a solution. The proofs can be found in the Appendix.

THEOREM 4.1. *For families $\boldsymbol{P}_\nu$, likelihood equations have solution iff the coefficient of variation of the sample satisfies $c^2 < 1/\nu$.*

Note that for the Normal Truncated Family the theorem says that likelihood equations have a solution iff $c^2 < 1$ (see Castillo (1994)).

Let $\gamma(\nu) = \gamma_c(\nu)$ be the solution of the likelihood equations established by Theorem 4.1. Proposition 4.3, proved in the Appendix, shows us that $\gamma(\nu)$ tends to ∞ when ν tends to $1/c^2$.

PROPOSITION 4.3. *Given* $\nu < 1/c^2$, *let* $\gamma(\nu) = \gamma_c(\nu)$ *be the solution of* $\mathrm{CV}_*^2(\gamma, \nu) = c^2$ *established by Theorem* 4.1. *Then,* $\gamma(\nu)$ *tends to* ∞ *when* ν *tends to* $1/c^2$ *and tends to* $-\infty$ *when* ν *tends to* 0.

This result and the asymptotic expansions (3.9) allow us to approximate $\gamma(\nu)$ for values of ν near $1/c^2$ by:

$$\gamma(\nu) \sim \sqrt{\nu+1}\sqrt{\frac{c^2\nu(\nu+2)+\nu^2+4\nu+2}{c^2\nu - 1}}.$$

THEOREM 4.2. *For family* $\boldsymbol{P}$, *likelihood equations have a solution if and only if the following inequality holds:*

$$\frac{\Gamma'}{\Gamma}(1/c^2) - \log(1/c^2) + r > 0, \tag{4.1}$$

where r *is the log-ratio of means and* c^2 *is the squared coefficient of variation of the sample.*

Note that for the Gamma family $\boldsymbol{P}_{.0.}$, the theoretical squared coefficient of variation, CV^2, is given by $\mathrm{CV}^2 = 1/\nu$. Moreover, the likelihood equations in terms of α and CV^2 are:

$$\begin{aligned} &\frac{1}{\alpha \mathrm{CV}^2} = t_1, \\ &\frac{\Gamma'}{\Gamma}(1/\mathrm{CV}^2) - \mathrm{Log}(1/\mathrm{CV}^2) + r = 0. \end{aligned} \tag{4.2}$$

Theorems 4.1 and 4.2 provide an easy way to determine whether equations (3.11) have a solution. If the likelihood equations have no solution, because the log-likelihood function (3.3) always has a maximum, it is found in Θ_0. From Proposition 4.2, we know that (3.3) restricted to Θ_0 is the log-likelihood function for the Gamma family. Then, when (4.1) does not hold, the maximum likelihood estimate for the family $\boldsymbol{P}$ can be obtained as a solution of (4.2), i.e., as a maximum likelihood estimate for the Gamma family.

Similarly, for family $\boldsymbol{P}_\nu$, if the condition of Theorem 4.1 does not hold, then the maximum likelihood estimate can be found by solving the first equation in (4.2), that is, by finding the maximum likelihood estimate for the family $\boldsymbol{P}_{.0\nu}$. For the Normal Truncated Family this is the Exponential Distribution.

5. Testing hypotheses

To test departures from Rayleigh or singly truncated normal distributions, we can use the classical likelihood ratio test. The null hypothesis is established by $\theta_0 = (0, \beta, \nu)$ for Rayleigh distribution, and by $\theta_0 = (\alpha, \beta, 1)$ for singly truncated normal distribution. In both cases θ_0 are interior points of the natural domain of parameters.

The situation is different when testing departures from Gamma distribution because $\theta_0 = (\alpha, 0, \nu)$ is a boundary point. Then, the asymptotic distribution of the likelihood ratio test statistic is not a chi-squared (see Self and Liang (1987)). We propose to use the fact that condition (4.1) is important in discerning between the Gamma family and the general family $\boldsymbol{P}$. Then, for testing the hypothesis $\beta = 0$ versus $\beta > 0$, the asymptotic behaviour of the left hand side of (4.1), when the null hypothesis is true, can be employed. This behaviour is established by the following result.

THEOREM 5.1. *Given a random sample $x = (x_1, \ldots, x_n)'$ of a Gamma distribution, the statistic:*

$$\frac{\sqrt{n}}{\sigma(c^2)}\left(\frac{\Gamma'}{\Gamma}(1/c^2) - \log(1/c^2) + r\right) \tag{5.1}$$

tends in law to an $N(0,1)$ as n tends to ∞, where:

$$\sigma^2(c^2) = \psi'\left(\frac{1}{c^2}\right)\left(2\psi'\left(\frac{1}{c^2}\right)\left(\frac{1}{c^4} + \frac{1}{c^2}\right) - \frac{4}{c^2} - 5\right) + 2 + 3c^2 \quad \textit{and}$$

$$\psi(x) = \frac{\Gamma'}{\Gamma}(x).$$

The proof is given in the Appendix.

Note that for the Gamma distribution, the statistics c^2 and r are invariant with respect to changes of scale and thus the statistic in (5.1) depends only on ν. Notice that Theorem 5.1 provides a very easy way to test Gamma distribution against the family $\boldsymbol{P}$ alternative without finding the likelihood estimators. The following table displays the upper tail 5% points of the distribution of the statistic in (5.1), based on 10000 samples of size $n = 5, 10, 20, 50, 100, 200, 650$, and $\nu = 1$, 1.5, 3, 5, 10, 15, 20.

Table 1 shows that the convergence to a standard normal is very slow for practical purposes. For small samples it is necessary to evaluate specifically the percentiles. Better approximations for the distribution function of the statistic in (5.1) may be obtained by using higher order expansions.

Table 1. Estimated upper tail 5% points, for the statistic given in Theorem 5.1.

N	$\nu = 1$	$\nu = 1.5$	$\nu = 3$	$\nu = 5$	$\nu = 10$	$\nu = 15$	$\nu = 20$
5	2.76	2.37	1.70	1.54	1.30	1.22	1.20
10	2.94	2.56	2.07	1.84	1.63	1.54	1.49
20	2.52	2.33	1.99	1.89	1.71	1.68	1.67
50	2.12	2.06	1.96	1.81	1.74	1.70	1.67
100	2.09	2.00	1.85	1.77	1.77	1.74	1.63
200	1.89	1.86	1.74	1.76	1.70	1.70	1.71
650	1.77	1.77	1.73	1.71	1.67	1.67	1.67
∞	1.64	1.64	1.64	1.64	1.64	1.64	1.64

6. Some examples

6.1 *Fitting distribution of divorces using the family P*

The distribution of divorces by duration of the dissolved union, in France in 1900 (see Pressat (1966)), is a classical example in which the data has traditionally be fitted to a Gamma distribution, but family $\boldsymbol{P}$ can do better. The 650 times of duration of the dissolved unions described in the book by Pressat have an arithmetical mean $m_1 = 11.99384$, a squared coefficient of variation $c^2 = 0.366657$ and an $r = 0.211235$. Moreover the moments of orders two, three and four are $m_2 = 196.5969$, $m_3 = 3931.461$ and $m_4 = 89575.643$. The method of estimation given in (3.1) using the first four moments is inadequate because it gives negative estimations of β and ν. The evaluation of the left part of (4.1) gives the value 0.016 and it shows that equation (3.10) have a solution. In order to solve them, we use the reduced equation (3.11) and the Newton-Raphson method. The steps to improve it are the following:

Step 0. Choose a starting point (γ_0, ν_0) near the solution. Given c^2 and r of the sample, Fig. 1 can help us to do it.

Step 1. Compute the quantities,

$$d = \varphi_{\nu_0}(\gamma_0), \qquad e = \varphi_{\nu_0+1}(\gamma_0), \qquad f = \varphi^1_{\nu_0}(\gamma_0),$$
$$g = \varphi^1_{\nu_0+1}(\gamma_0), \qquad h = \varphi^2_{\nu_0}(\gamma_0).$$

Step 2. Compute the values of the function and its derivatives,

$$F_1 = \frac{\nu_0 d^2}{2e^2} - \frac{\gamma_0 d}{2e} - c^2 - 1, \qquad F_2 = \operatorname{Log}\left(\frac{e}{d}\right) - \left(\frac{f}{d}\right) - r,$$
$$F_{11} = \frac{\gamma_0}{2} + \frac{d^3\nu_0^2}{2e^3} - \frac{3d^2\nu_0\gamma_0}{4e^2} - \frac{d(4\nu_0 - \gamma_0^2 + 2)}{4e},$$
$$F_{12} = \frac{d(d + 2f\nu_0 + g\gamma_0)}{2e^2} - \frac{d^2 g\gamma_0}{e^3} - \frac{f\gamma_0}{2e},$$
$$F_{21} = \frac{\gamma_0}{2} + \frac{e+g}{d} - \frac{ef}{d^2} - \frac{\nu_0 d}{2e}, \qquad F_{22} = \frac{g}{e} + \frac{f^2 - d(f+h)}{d^2}.$$

These have been deduced using (3.6) and (3.8).

Step 3. Calculate the point,

$$\gamma_1 = \gamma_0 - \frac{F_{22}F_1 - F_{12}F_2}{F_{11}F_{22} - F_{12}F_{21}}, \qquad \nu_1 = \nu_0 - \frac{F_{11}F_2 - F_{21}F_1}{F_{12}F_{22} - F_{12}F_{21}}.$$

Now we shall iterate steps 1 to 3 in order to obtain $(\gamma_2, \nu_2), (\gamma_3, \nu_3), \ldots$ until the sequence stabilizes.

φ_ν, φ^1_ν, φ^2_ν in Step 1 can be evaluated by numerical integration of (3.5) or using (3.7). If the expansion (3.7) is used then it is necessary to compute the gamma function and its first two derivatives. Suitable algorithms can be found in Abramowitz and Stegun (1972).

We used the program MATHEMATICA to follow the steps described above, and the evaluations of Step 1 was performed using numerical integration with the procedure NIntegrate.

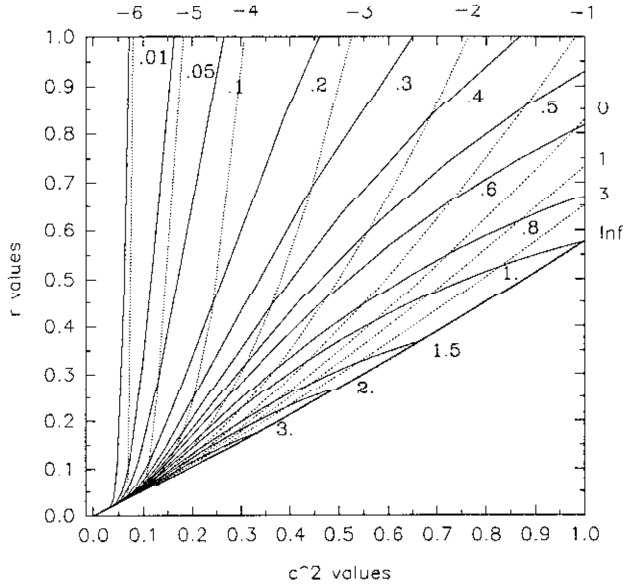

Fig. 1. Solutions of (3.11) for values of r and c^2. Dotted lines and numbers outside the square are the γ-values. Continuous lines and numbers inside the square are the ν-values.

For the data of duration times in divorced couples, solving the equation (3.11), starting with the point $\nu_0 = 2$, $\gamma_0 = 3$ we obtain the values $\nu = 2.15023$ and $\gamma = 3.11197$ in 3 iterations. These values correspond to $\alpha = .125749$ and $\beta = .00163282$. If we fit these data to the Gamma distribution, we also obtain a good result which passes, for instance, the χ^2 test. However it does not pass specific tests for the Gamma distribution like the EDF-tests A^2, W^2 or U^2 (see D'Agostino and Stephens (1986)). In the same way, if we calculate the expression (5.1), the result is 2.1517. Looking the upper tail 5% points of the Table 1 for $n = 650$, it rules out the hypothesis that the data belong to a Gamma distribution.

6.2 *Testing Gamma distribution against mixtures*

The statistic introduced by Theorem 5.1 is conceptually adequate for testing Gamma distribution against alternatives in family $\boldsymbol{P}$. It is not an "omnibus" test like the EDF-tests A^2, W^2 or U^2, but for certain alternatives, some of them even not included in family $\boldsymbol{P}$, it seems to have good power. This is the situation when the alternative is a mixture of a Gamma distribution with parameters ν, α and a small proportion of a Gamma distribution with the same α and the smaller parameter of shape ν.

We have considered a Gamma with $\nu = 10$ and $\alpha = 1$, contaminated by an 2.5% of a Gamma with $\alpha = 1$ and $\nu = 1, 2, 3, 4, 5, 6$. Table 2 shows the powers based on 10000 samples of size $n = 200$ obtained when using the test of Theorem 5.1 (test P) and the EDF-tests A^2, W^2 and U^2. The level of significance considered is 0.05 and the critical point for test P showed in Table 1 in 1.70. The

Table 2. Estimated powers for the tests P, A^2, W^2 and U^2 for different values of the parameter ν.

ν	P	A^2	W^2	U^2
1	98.0	95.2	93.3	92.3
2	92.6	83.3	76.5	72.8
3	77.5	58.6	49.3	42.2
4	52.2	29.7	23.6	20.0
5	28.1	13.2	11.2	9.5
6	14.3	7.0	6.2	6.2

critical points for the EDF-tests A^2, W^2 and U^2 can be found in D'Agostino and Stephens (1986), and they are respectively 0.754, 0.127 and 0.117.

Table 2 shows that, for the case considered, test P is always more powerful than the EDF-tests.

Acknowledgements

We thank the referees for valuable comments and suggestions that led to a substantial improvement in this paper.

Appendix

Watson's lemma. Given is a locally integrable function $f(t)$ on $(0,\infty)$ bounded for finite t and $f(t) = O(e^{at})$, $t \to \infty$, and $f(t) \sim \sum c_m t^{Am}$, $t \to 0+$, with $-1 < A0 < A1 < \cdots < An \to \infty$. Then, as $\lambda \to \infty$

$$\int_0^\infty e^{-\gamma t} f(t)dt \sim \sum_{m=0}^{\infty} \frac{c_m \Gamma(Am+1)}{\lambda^{Am+1}}.$$

PROOF. See Breitung ((1994), pp. 47–49).

PROOF OF PROPOSITION 4.1. Given two points in $T(X)$ any convex combination of them is in $\boldsymbol{C}$ because x, x^2 and $-\operatorname{Log}(x)$ are convex functions. On the other hand, given $t \in T(X)$, Bolzano's theorem shows that the equation $a^2 \operatorname{Log}(t_1 - \sqrt{t_2 - t_1^2}/a) + \operatorname{Log}(t_1 + a\sqrt{t_2 - t_1^2}) - t_3(a^2+1) = 0$ has a solution, a, in $[0, \sqrt{t_2 - t_1^2}/t_1]$. Therefore, taking $\lambda = a^2/(1+a^2)$ and x, y in $T(X)$ with its first components given by $x_1 = t_1 - \sqrt{t_2 - t_1^2}/a$, $y_1 = t_1 + a\sqrt{t_2 - t_1^2}$, we find $t = \lambda x + (1-\lambda)y$.

To characterize $\boldsymbol{C}_\nu$ it is enough to take x_1 and y_1, as before, for any $a > 0$. □

PROOF OF PROPOSITION 4.2. We have only to pass limits inside the integrals. Notice the following inequality:

$$|f(x)\exp(-\alpha x - \beta x^2 + (\nu-1)\operatorname{Log}(x))| \le |f(x)|\exp(-\alpha x + (\nu-1)\operatorname{Log}(x)),$$

where $\alpha > 0$, $x > 0$ and $f(x)$ are the functions 1, x, x^2 and $\text{Log}(x)$. The functions in the right part of the inequality are integrable and using the Lebesgue bounded convergence theorem the result holds. The values of the limits can easily be obtained by integration and they correspond to the Gamma Distribution. □

COROLLARY A.1. *If (γ_n, ν_n) tends to $(+\infty, \nu)$ then*

$$\text{CV}_*^2(\gamma_n, \nu_n) = \text{CV}^2(1, 1/\gamma_n^2, \nu_n) \to 1/\nu$$
$$\text{RM}_*(\gamma_n, \nu_n) = \text{RM}(1, 1/\gamma_n^2, \nu_n) \to \text{Log}(\nu) - \frac{\Gamma'}{\Gamma}(\nu).$$

PROOF OF THEOREM 4.1. For the family $\boldsymbol{P}_\nu$, we consider the first equation in (3.11). Now CV_*^2 is only a function of γ because ν is fixed. The asymptotic behaviour (3.9) and some algebra show us that $\text{CV}_*^2(\gamma, \nu) \sim (2/\gamma^2) + o(1/\gamma^2)$ when γ tends to $-\infty$, which implies that $\text{CV}_*^2(\gamma, \nu)$ tends to 0 when γ tends to $-\infty$, for any ν. Moreover $\text{CV}_*^2(\gamma, \nu)$ tends to $1/\nu$ when γ tends to $+\infty$, as a consequence of Corollary A.1. Then, theorem of Bolzano establishes that the equation has a solution if $c^2 < 1/\nu$, and this is unique because $\tau(\theta)$ is 1-1. Suppose now that there is a solution for a random sample with $c^2 \geq 1/\nu$. Then, there is also a solution for a random sample with $c^2 = 1/\nu$ because the domain T is a connected set. Now we have one point in Θ and another in Θ_0 (see Proposition 4.2) that have the same τ-image. This is a contradiction because, as a consequence of the Inverse Function Theorem, τ has an inverse on neighbourhoods of $t \in T$. □

COROLLARY A.2. *Given $\nu > 0$, the map $\text{CV}_*^2(\cdot, \nu) : \mathbb{R} \to (0, 1/\nu)$ is differentiable and one to one. Moreover, when γ tends to ∞ $\text{CV}_*^2(\gamma, \nu)$ tends to $1/\nu$ and when γ tends to $-\infty$ then $\text{CV}_*^2(\gamma, \nu)$ tends to 0.*

PROOF OF PROPOSITION 4.3. By definition $\text{CV}_*^2(\gamma(\nu), \nu) = c^2$. If $\gamma(\nu)$ tends to $\gamma_* < \infty$ when ν tends to $1/c^2$, then $\text{CV}_*^2(\gamma(\nu), \nu)$ tends to $\text{CV}_*^2(\gamma_*, 1/c^2)$ and it is lower than c^2 by Corollary A.2, thus yielding a contradiction.

Notice that using (3.11) and the recurrent formula (3.8) the equality $\text{CV}_*^2(\gamma(\nu), \nu) = c^2$ can be expressed as:

$$4\frac{\varphi_{\nu+2}}{\varphi_{\nu+1}}(\gamma(\nu)) = \frac{8\nu(1+c^2)}{\gamma + \sqrt{\gamma^2 + 8\nu(1+c^2)}}. \tag{A.1}$$

Now, if $\gamma(\nu)$ tends to any finite value when ν tends to 0 this is impossible because the right part of (A.1) tends to zero and the left part tends to a certain positive value. However $\gamma(\nu)$ does not tend to $+\infty$ when ν tends to 0; using (3.9), we can see that $\frac{\varphi_{\nu+2}}{\varphi_{\nu+1}}(\gamma) \sim \frac{\nu+1}{\gamma} + o(1/\gamma^2)$ as γ tends to $+\infty$. Then, multiplying both parts of (A.1) by $\gamma(\nu)$ and taking ν tending to 0 would lead to a contradiction. □

PROOF OF THEOREM 4.2. Given c^2, for each ν fixed belonging to the open interval $(0, 1/c^2)$ we consider the solution of the first equation in (3.11), $\gamma(\nu)$. We substitute γ in the second equation by the function $\gamma(\nu)$ and the problem is reduced to determining whether the equation $\mathrm{RM}_*(\gamma(\nu),\nu) = r$ has a solution for ν in $(0, 1/c^2)$. Using expression (3.8) we can write $\frac{\varphi_{\nu+1}}{\varphi_\nu}(\gamma) = \frac{\nu}{\nu+1}\frac{2\varphi_{\nu+3}(\gamma)+\gamma\varphi_{\nu+2}(\gamma)}{2\varphi_{\nu+2}(\gamma)+\gamma\varphi_{\nu+1}(\gamma)}$, and deriving (3.8) with respect to ν, $\frac{\varphi_\nu^1}{\varphi_\nu}(\gamma) = -\frac{1}{\nu}+\frac{2\varphi_{\nu+2}^1(\gamma)+\gamma\varphi_{\nu+1}^1(\gamma)}{2\varphi_{\nu+2}(\gamma)+\gamma\varphi_{\nu+1}(\gamma)}$. If we replace these expressions in the second equation in (3.11) we can use Proposition 4.3, (3.9) and some algebra to see that $\mathrm{RM}_*(\gamma(\nu),\nu) \sim 1/\nu + O(1)$ as ν tends to 0. Then, $\mathrm{RM}_*(\gamma(\nu),\nu)$ tends to $+\infty$ when ν tends to zero. Moreover, $\mathrm{RM}_*(\gamma(\nu),\nu)$ tends to $-\frac{\Gamma'}{\Gamma}(1/c^2)+\mathrm{Log}(1/c^2)$ when ν tends to $1/c^2$ as a consequence of Corollary A.1. Then, the theorem of Bolzano establishes that there is a solution if $\frac{\Gamma'}{\Gamma}(1/c^2) - \mathrm{Log}(1/c^2) + r > 0$. The same topological argument used in the proof of Theorem 4.1 can be applied here, and we conclude with the proof of Theorem 4.2. □

PROOF OF THEOREM 5.1. Given a random sample $x = (x_1, \ldots, x_n)'$ coming from a Gamma distribution with parameters α, ν, it is easy to establish that the statistic $\sqrt{n}(t_1, t_2, t_3)$ converges in law to a trivariant normal with mean vector given by $E(t) = (\frac{\nu}{\alpha}, \frac{\nu(\nu+1)}{\alpha^2}, \frac{\Gamma'}{\Gamma}(\nu) - \log(\alpha))$ and covariance matrix $\sum = \{\sigma_{ij}\}$, where $\sigma_{11} = \nu/\alpha^2$, $\sigma_{22} = 2\nu(\nu+1)(2\nu+3)/\alpha^4$, $\sigma_{33} = \psi'(\nu)$, $\sigma_{12} = 2\nu(\nu+1)/\alpha^3$, $\sigma_{13} = 1/\alpha$ and $\sigma_{23} = (2\nu+1)/\alpha^2$. The left part of (4.1) can be written as a differentiable real valued function $\phi(t) = \phi(t_1, t_2, t_3)$. Now, the δ-method (see Rao (1973)) establishes that $\sqrt{n}\Phi(t_1, t_2, t_3)$ converges in law to a univariate normal with mean $\phi(E(t)) = 0$ and variance $d\phi(E(t))' \sum d\phi(E(t)) = \psi'(\nu)(2\psi'(\nu)(\nu^2 + \nu) - 4\nu - 5) + 2 + (3/\nu)$. The last expression is a continuous function of ν and $1/c^2$ converges in probability to ν. Then standardizing, this concludes the proof. □

REFERENCES

Abramowitz, M. and Stegun, I. A. (1972). *Handbook of Mathematical Functions*, Dover Publications, New York.

Barndorff-Nielsen, O. (1978). *Information and Exponential Families*, Wiley, Norwich.

Barndorff-Nielsen, O. and Cox, D. R. (1979). Edgeworth and saddle-point approximation with statistical application, *J. Roy. Statist. Soc. Ser. B*, **41**(3), 279–312.

Barndorff-Nielsen, O. and Cox, D. R. (1989). *Asymptotic Techniques for use in Statistics*, Chapman and Hall, London.

Breitung, K. W. (1994). *Asymptotic Approximations for Probability Integrals*, Springer, Berlin.

Castillo, J. (1994). The singly truncated normal distribution, a non-steep exponential family, *Ann. Inst. Statist. Math.*, **46**(1), 57–66.

Cobb, L., Koppstein, P. and Chen, N. H. (1983). Estimation and moment recursion relations for multimodal distributions of the exponential family, *J. Amer. Statist. Assoc.*, **78**(381), 124–130.

D'Agostino, R. and Stephens, M. A. (1986). *Goddness-of-Fit Techniques*, Marcel Dekker, New York.

Johnson, N. and Kotz, S. (1970). *Continuous Univariate Distributions-I*, Wiley, New York.

McCullagh, P. (1987). *Tensor Methods in Statistics*, Chapman and Hall, London.

Pressat, R. (1966). Principes d'analyse, *Cours d'analyse démographique de L'I.D.U.P.*, Editions de l'Institut d'études démographiques, Paris.

Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*, 2nd ed., Wiley, London.

Ross, S. M. (1983). *Stochastic Processes*, Wiley, New York.

Self, S. G. and Liang, K. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions, *J. Amer. Statist. Assoc.*, **82**(398), 605–610.

Toranzos, F. I. (1952). An asymmetric bell-shaped frequency curve, *Ann. Math. Statist.*, **23**, 467–469.