Domains of attraction for exponential families

August A. Balkema, Claudia Klüppelberg, Sidney I. Resnick [a,b,c]

[a] *Department of Mathematics, University of Amsterdam, CNL-1018TV Amsterdam, CNL-1018TV Amsterdam, Holland, guus@wins.uva.nl.* [1]

[b] *Center of Mathematical Sciences, Munich University of Technology, D-80290 Munich, Germany, cklu@ma.tum.de, www.ma.tum.de/stat/*

[c] *Operations Research and Industrial Engineering, Cornell University, Ithaca, New York 14853-7501 USA, sid@orie.cornell.edu* [2]

Abstract

With the df F of the rv X we associate the natural exponential family of df's F_λ where

$$dF_\lambda(x) = e^{\lambda x} dF(x)/Ee^{\lambda X}$$

for $\lambda \in \Lambda := \{\lambda \in \mathbb{R} \mid Ee^{\lambda X} < \infty\}$. Assume $\lambda_\infty = \sup \Lambda \leq \infty$ does not lie in Λ. Let $\lambda \uparrow \lambda_\infty$, then non-degenerate limit laws for the normalised distributions $F_\lambda(a_\lambda x + b_\lambda)$ are the normal and gamma distributions. Their domains of attractions are determined. Applications to saddlepoint and gamma approximations are considered.

Key words:
asymptotic normality, convex conjugate, domain of attraction, exponential family, Esscher transform, gamma distribution, Laplace transform, normal distribution, regular variation, self-neglecting, strongly unimodal.
1991 MSC: 60F05, 60E05, 44A10.

---

1 Introduction.

A one-dimensional exponential family may sometimes be normalised by translation and scaling to yield a non-degenerate limit law in which case the only possible limit laws are the normal distribution and the gamma distributions on $(0, \infty)$ and on $(-\infty, 0)$. See Balkema, Klüppelberg and Resnick [2], hereafter

[1] A.B. thanks the Munich University of Technology for their hospitality.
[2] S.R. was partially supported by NSF Grant DMS-97-04982 at Cornell University.

abbreviated to BKR99. These are exactly the distributions which generate exponential families which are stable under a continuous one-parameter group of positive affine transformations (Bar-Lev and Casalis [4,5]). In the present paper we describe the domains of attraction of these limit distributions and give a new characterisation of a stable exponential family.

With the (random variable) rv X with distribution function (df) F, we associate the *natural exponential family* consisting of rv's X^λ with df's $dF_\lambda(x) = e^{\lambda x} dF(x)/K(\lambda)$. The constant $K(\lambda) = \int e^{\lambda x} dF(x)$ is the *moment generating function (mgf)* of X evaluated at λ. The domain of K is the interval

$$\Lambda = \{\lambda \in \mathbb{R} \mid K(\lambda) = Ee^{\lambda X} < \infty\}$$

containing the origin. The mgf $\lambda \mapsto K(\lambda)$ is continuous and strictly positive on Λ. The important *cumulant generating function (cgf)* $\kappa = \log K$ is a convex analytic function on the interior of Λ. We study the asymptotic behaviour of the df's F_λ for $\lambda \to \lambda_\infty := \sup \Lambda$, where $\lambda \to \lambda_\infty$ means that $\lambda \in \Lambda$ and λ converges to λ_∞ from below.

If $\lambda_\infty \in \Lambda$, then $F_\lambda \downarrow F_{\lambda_\infty}$ pointwise. In the more interesting case that $\lambda_\infty \notin \Lambda$, we have $F_\lambda(x) \downarrow 0$ for every $x < x_\infty$ where $x_\infty = \sup\{F < 1\}$ is the *upper endpoint* of the df F. Then $X^\lambda \to x_\infty$ in probability and it may be possible to normalize the rv's X^λ by changing location and scale so that for some $a_\lambda > 0$ and $b_\lambda \in \mathbb{R}$

$$A_\lambda X^\lambda := \frac{X^\lambda - b_\lambda}{a_\lambda} \quad \Rightarrow \quad V \qquad \lambda \to \lambda_\infty. \tag{1.1}$$

Here $\Rightarrow$ denotes convergence in law to a *non-degenerate* rv. The *upper endpoint* λ_∞ of Λ always satisfies $\lambda_\infty \geq 0$ and we are interested in the case $\lambda_\infty \notin \Lambda$ and hence $\lambda_\infty > 0$.

Theorem 3.5 of BKR99 [2] characterizes possible non-degenerate limits in (1.1) and Section 2 gives the necessary background. If V is a non-degenerate limit variable in (1.1), then there exist centering $b_\lambda \in \mathbb{R}$ and scaling $a_\lambda > 0$ so that V is a standard normal variable, or so that V or $-V$ has a gamma distribution on $(0, \infty)$. Thus, the class of limit distributions, the *extended gamma family*, (BKR99 [2], Example 2.9) can be written (see Section 2) as a continuous three parameter family, indexed by shape, location and scale.

When (1.1) holds, the rv X or df F generating the exponential family X^λ *lies in the domain of attraction of* the limit rv V or of its df G and we write $X \in \mathcal{D}(V)$ or $F \in \mathcal{D}(G)$. The aim of this paper is to characterize the domains of attraction of the limit distributions.

Weak convergence of the normalised rv's in the exponential family to a non-degenerate limit implies convergence of the mgf's (BKR99 [2], Theorem 3.6) and this implies convergence of all moments. In Section 3 we show that asymptotic normality is equivalent to convergence to zero of the third moment of X^λ standardized to mean 0 and variance 1. The normal domain is also characterized in terms of the cgf κ, and the limit in (1.1) is normal if and only if its cgf κ satisfies that $1/\sqrt{\kappa''}$ is *self-neglecting*; that is

$$\kappa''\left(\lambda + \frac{x}{\sqrt{\kappa''(\lambda)}}\right) \sim \kappa''(\lambda) \quad \lambda \to \lambda_\infty$$

for all $x \in \mathbb{R}$. The domains of attraction of the gamma distributions are the subject of Section 4. For F to be in the domain of a gamma with positive shape parameter requires $\lambda_\infty < \infty$ and then a necessary and sufficient condition is

$$\lim_{\lambda \to \lambda_\infty} (\lambda_\infty - \lambda)\kappa'(\lambda) = c > 0.$$

An equivalent condition is that $E((\lambda_\infty - \lambda)X^\lambda)$ converges to the first moment of the gamma distribution. Criteria for a negative shape parameter are comparable and also discussed in Section 4.

Domain of attraction conditions can be expressed either in terms of the df F or in terms of the cgf κ and hence the results developed here have applications to the asymptotic theory of transforms and their inversion, and for saddlepoint approximations. In the latter case, distributions in the domain of attraction of the normal law have the property that the saddlepoint approximation of the density becomes exact in the limit; see Barndorff-Nielsen and Klüppelberg [6,7]. This is further discussed in Section 5. The theory of this paper sheds new light on the subject of regular variation and on a class of distribution functions with very thin tails. Anticipated investigations of the boundary behaviour of multivariate Laplace transforms will require good understanding of the asymptotics of univariate exponential families.

Whether one develops results about exponential families in terms of measures, df's or rv's, is largely a matter of taste and habit. Many of our results are expressed in terms of the rv's X^λ, which we have injected into the exposition as a convenience.

2 Background.

2.1 *The class of limit distributions.*

To aid understanding of the limit distributions, we introduce the cgf's:

$$\varphi_\beta(\xi) = \begin{cases} -\beta^{-2}\log(1-\xi\beta) - \xi/\beta & \text{if } \beta \neq 0,\ \xi\beta < 1 \\ \xi^2/2 & \text{if } \beta = 0. \end{cases}$$

The functions φ_β depend continuously on the parameter β as one sees either by using l'Hospital's rule or by noting that $\varphi_\beta(0) = \varphi'_\beta(0) = 0$ and

$$\varphi''_\beta(\xi) = \frac{1}{(1-\beta\xi)^2} \qquad \beta\xi < 1.$$

A gamma distributed rv Z with parameter α has density $z^{\alpha-1}e^{-z}/\Gamma(\alpha)$ for $z > 0$ and cgf $\lambda \mapsto -\alpha\log(1-\lambda)$ on $(-\infty, 1)$. Set $\alpha = 1/\beta^2$. Then $Y_\beta = \beta Z - 1/\beta$ has cgf φ_β for $\beta \neq 0$. For $\beta \to 0$ the variable Y_β converges in distribution to the standard normal rv Y_0 with density $e^{-y^2/2}/\sqrt{2\pi}$ for $y \in \mathbb{R}$. The *extended gamma family* is the set of the probability distributions of the variables $aY_\beta + b$ with $a > 0$ and $\beta, b \in \mathbb{R}$. This is a continuous three parameter family. It is also the set of possible non-degenerate limit laws in (1.1).

If we think of $X \mapsto X^\lambda$ as a transformation, the next result shows that the extended gamma family is the set of fixed points.

Theorem 2.1 *Let Y^λ, $\lambda \in \Lambda$, be the exponential family generated by Y. The exponential family is stable in the sense that the standardised variables*

$$V_\lambda = \frac{Y^\lambda - \mu}{\sigma}, \qquad \mu = \kappa'(\lambda) = EY^\lambda,\ \sigma^2 = \kappa''(\lambda) = \mathrm{Var}(Y^\lambda), \qquad \lambda \in \Lambda$$

all have the same distribution if and only if there is a non-empty open interval $J \subset \Lambda$ so that EV_λ^3 does not depend on λ for $\lambda \in J$. Furthermore, the only stable exponential families are those generated by the rv's in the extended gamma family.

Proof The condition is obviously necessary. We may assume that J contains the origin and that $EY = 0$ and $EY^2 = 1$. Then $V_0 = Y^0 = Y$. Let κ_λ be the cgf of Y^λ. The cgf of V_λ is

$$\log Ee^{\xi(Y^\lambda - \mu)/\sigma} = \kappa_\lambda(\xi/\sigma) - \xi\mu/\sigma = \kappa(\lambda + \xi/\sigma) - \kappa(\lambda) - \xi\mu/\sigma.$$

If $V_\lambda \stackrel{\mathrm{d}}{=} Y$ it has cgf κ and hence

$$\kappa(\lambda + \xi/\sigma) = \kappa(\lambda) + \kappa(\xi) + \xi\mu/\sigma. \tag{2.1}$$

Taking the second derivative with respect to ξ we obtain

$$\sigma^2(\lambda+\xi/\sigma)/\sigma^2 = \sigma^2(\xi), \qquad \sigma = \sigma(\lambda). \tag{2.2}$$

Write $\tau = 1/\sigma$ to get $\tau(\lambda + \tau\xi)/\tau = \tau(\xi)$ and differentiate this expression to find

$$\tau'(\lambda + \tau\xi) = \tau'(\xi). \tag{2.3}$$

Setting $\xi = 0$ we see that $\tau'(\lambda)$ is constant in λ. Note that

$$\tau'(\lambda) = (1/\sigma(\lambda))' = -\frac{\kappa'''(\lambda)}{2\sigma^3} = -\frac{1}{2}E(\frac{Y^\lambda - \mu}{\sigma})^3. \tag{2.4}$$

If EV_λ^3 is constant, then (2.4) (2.3),(2.2),(2.1) all hold and $V_\lambda \stackrel{d}{=} Y$.

If $\tau'(\lambda) = EV_\lambda^3 = -\beta$ does not depend on λ then from (2.4) $\tau(\lambda) = 1 - \beta\lambda$, (since $\sigma(0) = 1$) and $\kappa''(\lambda) = 1/(1-\beta\lambda)^2$ on J. Hence, by analytic continuation $\kappa(\lambda) = \varphi_\beta(\lambda)$ for $\beta\lambda < 1$. □

2.2 Tail equivalence.

The description of the domains of attraction of the limit laws relates the behaviour of the df F in the neighbourhood of x_∞ with the behaviour of the mgf K and of the exponential family F_λ in the neighbourhood of λ_∞. When considering such asymptotic behavior, the notion of tail equivalence provides an invariance classifier. Two df's F and G are *tail equivalent* if they have the same upper endpoint $x_\infty = \sup\{F < 1\} = \sup\{G < 1\}$, if they are continuous in the upper endpoint, $F(x_\infty-) = G(x_\infty-) = 1$, and if

$$1 - G(x) \sim 1 - F(x) \qquad x \to x_\infty \tag{2.5}$$

in the sense that the quotient tends to 1. Again we mean by $x \to \xi_\infty$ that x converges to x_∞ from below.

Theorem 2.2 *Let X have df F which is continuous in its upper endpoint $x_\infty = \sup\{F < 1\}$ and mgf K with domain Λ with $\lambda_\infty \notin \Lambda$. Suppose Y is another rv with df G and mgf M and that F and G are tail equivalent. Then*

1) K and M have the same upper endpoint λ_∞ and $M(\lambda) \sim K(\lambda)$ for $\lambda \to \lambda_\infty$;

2) $G_\lambda(x) - F_\lambda(x) \to 0$ uniformly in $x \in \mathbb{R}$ for $\lambda \to \lambda_\infty$;

3) $(1 - G_\lambda(x))/(1 - F_\lambda(x)) \to 1$ *uniformly in* $x < x_\infty$ *for* $\lambda \to \lambda_\infty$.

Furthermore, if X^λ *and* Y^λ *are the associated exponential families, then for any sequence* $\lambda_n \to \lambda_\infty$ *and any rv* V

$$\frac{X^{\lambda_n} - b_n}{a_n} \Rightarrow V \quad \iff \quad \frac{Y^{\lambda_n} - b_n}{a_n} \Rightarrow V.$$

Proof Note that $\lambda_\infty(G) = \lambda_\infty(F)$ since for $\lambda > 0$ the integral $\int e^{\lambda x}(1 - G(x))dx = M(\lambda)/\lambda$ converges if and only if this holds for the corresponding integral for the df F. Since 3) implies 2) it remains to prove the asymptotic equivalences in 1) and 3).

Now first assume that $G \equiv F$ on an interval $[x_0, x_\infty)$. Then

$$K(\lambda)(1 - F_\lambda(x)) = \int_{(x,\infty)} e^{\lambda t} dF(t) = M(\lambda)(1 - G_\lambda(x)) \quad x_0 \leq x < x_\infty. \tag{2.6}$$

For fixed $x < x_\infty$ the central term grows without bound for $\lambda \to \lambda_\infty$ since $K(\lambda) \uparrow \infty$ if $\lambda_\infty \notin \Lambda$. Hence $M(\lambda) \uparrow \infty$ and $G_\lambda(x) \to 0$ for $x < x_\infty$ by the first sentence of the proof. Since this also holds for $F_\lambda(x)$ we see that $M(\lambda) \sim K(\lambda)$ for $\lambda \to \lambda_\infty$. This in turn implies 3) by (2.6) for $x_0 \leq x < x_\infty$.

Now let $\varepsilon > 0$. There exists a constant $x_0 < x_\infty$ so that

$$|\log(1 - G(x)) - \log(1 - F(x))| < \varepsilon \qquad x \in [x_0, x_\infty).$$

Let $F^* = F 1_{[x_0,\infty)}$ and define G^* similarly. The inequality

$$\int h(t) dG^*(t) < e^\varepsilon \int h(t) dF^*(t)$$

holds for all indicator functions $h = 1_{[x,\infty)}$. Hence it holds for all non-negative increasing functions. Take $h(t) = e^{\lambda t}$ with $\lambda \in (0, \lambda_\infty)$ to conclude that $M^*(\lambda) < e^\varepsilon K^*(\lambda)$ and take $h(t) = e^{\lambda t} 1_{[x,\infty)}(t)$ to conclude that $M^*(\lambda)(1 - G^*_\lambda(x)) < e^\varepsilon K^*(\lambda)(1 - F^*_\lambda(x))$ for $\lambda \in (0, \lambda_\infty)$ and all x. By symmetry the inequalities hold if we interchange F and G (and K and M). Hence we conclude that $|\log M^*(\lambda) - \log K^*(\lambda)| < \varepsilon$ for $\lambda \in (0, \lambda_\infty)$ and $|\log(1 - G^*_\lambda(x)) - \log(1 - F^*_\lambda(x))| < 2\varepsilon$ for $\lambda \in (0, \lambda_\infty)$ and all $x < x_\infty$.

Since $F^* \equiv F$ and $G^* \equiv G$ on $[x_0, x_\infty)$ we know that $K^*(\lambda) \sim K(\lambda)$ and $M^*(\lambda) \sim M(\lambda)$. Thus $|\log M(\lambda) - \log K(\lambda)| < 2\varepsilon$ for $\lambda \in (\lambda_1, \lambda_\infty)$. Similarly $|\log(1 - F_\lambda(x)) - \log(1 - G_\lambda(x))| < 3\varepsilon$ for $\lambda \in (\lambda_2, \lambda_\infty)$ and all $x < x_\infty$. Since ε is arbitrary this proves the asymptotics in 1) and 3).

By 2) weak convergence $G_{\lambda_n}(a_n x + b_n) \to H(x)$ holds if and only if $F_{\lambda_n}(a_n x + b_n)$ converges weakly to $H(x)$. □

2.3 Convergence preservation.

Assume (1.1), so that a limit law exists and define the notational convenience, called the *Esscher transform*, by

$$E^{\lambda}X = X^{\lambda} \qquad \lambda \in \Lambda. \tag{2.7}$$

The semi-group property holds: $E^{\alpha}E^{\beta} = E^{\alpha+\beta}$ if α, β and $\alpha+\beta$ lie in Λ.

Theorem 2.3 *Suppose as in (1.1) $V_{\lambda} := A_{\lambda}X^{\lambda} \Rightarrow V$ as $\lambda \to \lambda_{\infty}$, with V non-constant. Then for all $\gamma \in \mathbb{R}$ for which $Ee^{\gamma V}$ is finite, we have $E^{\gamma}V_{\lambda} \Rightarrow E^{\gamma}V$ and $Ee^{\gamma V_{\lambda}} \to Ee^{\gamma V}$ as $\lambda \to \lambda_{\infty}$,*

Proof Convergence of the mgf's is proved in BKR99 [2], Theorem 3.6. Observe that the set $\Gamma = \{\gamma \in \mathbb{R} \mid Ee^{\gamma V} < \infty\}$ is open if V is normal or if V or $-V$ has a gamma distribution. Hence convergence of the mgf's implies for any $\gamma \in \Gamma$

$$\int \varphi(x)e^{\gamma x}d\pi_{\lambda}(x) \to \int \varphi(x)e^{\gamma x}d\pi(x) \qquad \lambda \to \lambda_{\infty}$$

for all continuous bounded functions φ on $\mathbb{R}$. Here π is the distribution of V and π_{λ} the distribution of V_{λ}. This gives the asserted weak convergence of the exponential families. □

2.4 Densities.

Suppose X has density f and set $f_{\lambda}(x) = e^{\lambda x}f(x)/K(\lambda)$ for the density of X^{λ}. Assume

$$g_{\lambda}(c) = a_{\lambda}f_{\lambda}(a_{\lambda}c + b_{\lambda}) \to g(c) > 0 \qquad \lambda \to \lambda_{\infty} \tag{2.8}$$

in some point c. Write $c_{\lambda} = a_{\lambda}c + b_{\lambda}$. This yields an asymptotic expression for the mgf:

$$K(\lambda) \sim a_{\lambda}f(c_{\lambda})e^{\lambda c_{\lambda}}/g(c) \qquad \lambda \to \lambda_{\infty}. \tag{2.9}$$

The seminal work by Feigin and Yashchin [14] discusses this asymptotic relation. They considered the exponential family of rv's Y^{λ} generated by the measure with density $f^*(y) = 1-F(y)$. The density f^*_{λ} of Y^{λ} is $e^{\lambda y}(1-F(y))/K^*(\lambda)$ where $K^*(\lambda) = \int e^{\lambda y}(1-F(y)dy = K(\lambda)/\lambda$ by partial integration. Theorem 1 of [14] gives the Tauberian relation

$$1 - F(c^*_{\lambda}) \sim K(\lambda)e^{-\lambda c^*_{\lambda}}g^*(c)/(\lambda a^*_{\lambda}) \quad \lambda \to \lambda_{\infty} \tag{2.10}$$

provided that $a_\lambda^* f_\lambda^*(a_\lambda^* c + b_\lambda^*) \to g^*(c)$.

Since weak convergence of rv's Y^λ, properly normalised, is assumed in their results, the theory developed in BKR99 [2] shows that only the normal and the gamma densities can occur as limit in these asymptotic relations.

3 The domain of attraction of the normal law.

Let $U = N_{01}$ denote a standard normal random variable with probability distribution γ_{01} and density $g_{01}(u) = (1/\sqrt{2\pi})\exp(-u^2/2)$, $u \in \mathbb{R}$. A sequence of rv's X_n, $n \in \mathbb{N}$, is *asymptotically normal (AN)* if there exist $a_n > 0$ and $b_n \in \mathbb{R}$ so that $(X_n - b_n)/a_n \Rightarrow U$. Asymptotic normality of X_n does not imply that the second moments exist, and if these exist this does not imply that $U_n := (X_n - E(X_n))/\sqrt{\mathrm{Var}(X_n)} \Rightarrow U$.

Set $U_n = (X_n - \mu_n)/\sigma_n$, where $\mu_n = E(X_n)$ and $\sigma_n^2 = \mathrm{Var}(X_n)$, assuming these moments exist. Set $\kappa_n = \log K_n$ for the cgf of the standardised rv U_n. With this notation, consider the following statements, whose equivalence is not apriori obvious.

AN1) $(X_n - b_n)/a_n \Rightarrow U$;

AN2) $U_n \Rightarrow U$;

AN3) $EU_n^3 \to 0$;

AN4) $EU_n^k \to EU^k$ for $k \in \mathbb{N}$;

AN5) $\kappa_n(\tau) \to \tau^2/2$ uniformly on bounded intervals.

The limit relation AN5) implies that $Eh(U_n) \to Eh(U)$ for all continuous functions of exponential growth and will be called *strong asymptotic normality*. We investigate the equivalence of these statements applied to exponential families by replacing X_n by X^λ and U_n by $U_\lambda = (X^\lambda - E(X^\lambda))/\sqrt{\mathrm{Var}(X^\lambda)}$.

Theorem 3.1 *For the exponential family X^λ the limit relations AN1)–AN5) are equivalent.*

Since convergence of mgf's implies weak convergence and convergence of moments, it suffices to prove AN1)⇒AN5) and AN3)⇒AN1). For the implication AN1)⇒AN5), use Theorem 2.3 and the fact that pointwise convergence of mgfs implies uniform convergence on bounded sets. The proof that AN3)⇒AN1) will be supplied after a discussion of asymptotically parabolic functions.

If X has cgf κ then the cgf of X^λ satisfies

$$\kappa_\lambda(\xi) = \kappa(\lambda + \xi) - \kappa(\lambda). \tag{3.1}$$

For each $\lambda \in (0, \lambda_\infty)$ the cgf κ_λ exists on a neighbourhood of the origin. We may compute the moments of X^λ by differentiating the cgf:

$$\mu(\lambda) = E\,X^\lambda = \kappa'(\lambda) \qquad \sigma^2(\lambda) = \mathrm{Var}(X_\lambda) = \kappa''(\lambda) \quad 0 < \lambda < \lambda_\infty. \tag{3.2}$$

In particular, if F is non-degenerate then so is F_λ , and we see that the function κ then has a strictly positive second derivative. The following result is implicit in the proof of Corollary 1 in Feigin and Yashchin [14]. See also Balkema, Klüppelberg and Resnick [1], henceforth abbreviated as BKR93.

Proposition 3.2 *Let X have cgf κ with upper endpoint λ_∞. If the function σ in (3.2) satisfies the relation*

$$\sigma(\lambda + x/\sigma(\lambda))/\sigma(\lambda) \to 1 \qquad \lambda \to \lambda_\infty \tag{3.3}$$

for each $x \in \mathbb{R}$, then the family X^λ is strongly asymptotically normal; i.e. $AN5$) holds.

Proof Since the cgf κ of X is a convex analytic function so is the cgf γ_λ of the normalized variable U_λ given by (3.8). Relation (3.1) gives

$$\begin{aligned}\gamma_\lambda(\xi) &= \kappa_\lambda(\xi/\sigma(\lambda)) - \xi\mu(\lambda)/\sigma(\lambda)\\ &= \kappa(\lambda + \xi/\sigma(\lambda)) - \kappa(\lambda) - \xi\mu(\lambda)/\sigma(\lambda).\end{aligned} \tag{3.4}$$

Note that we have normalized the convex function κ to make $\gamma_\lambda(0) = 0,\ \gamma_\lambda'(0) = 0,\ \gamma''{}_\lambda(0) = 1$. The condition (3.3) is assumed to hold pointwise. By continuity of the function σ it will hold uniformly on bounded sets by Bloom's theorem (see Bingham, Goldie and Teugels (henceforth BGT) [9], Section 2.11. It thus implies that the second derivative of γ_λ will be close to 1 uniformly on any bounded interval around the origin, and hence $\gamma_\lambda(\xi) \to \xi^2/2$ uniformly on bounded intervals, which implies AN5).

If $\Lambda \neq \mathbb{R}$ one has to check that $\gamma_\lambda(\xi)$ is well-defined in the sense that for any ξ the point $\lambda + \xi/\sigma(\lambda)$ lies in Λ eventually. Note that $\sigma^2(\lambda) \to \infty$ if $\lambda_\infty < \infty$ since $\lambda_\infty \notin \Lambda$ then implies $\kappa(\lambda) \to \infty$. Below we shall see that (3.3) implies that $1/\sigma(\lambda) = o(\lambda)$ if $\lambda_\infty = \infty$ and $1/\sigma(\lambda) = o(\lambda_\infty - \lambda)$ if $\lambda_\infty < \infty$. This ensures that $\lambda + \xi/\sigma(\lambda) \in \Lambda$ eventually. □

To prove the converse of Proposition 3.2, we need the following concept: A positive function s is *self-neglecting* at $t_\infty \leq \infty$ if it is defined on a left neighbourhood of t_∞ and if

$$s(t+xs(t))/s(t) \to 1, \qquad t \to t_\infty, \tag{3.5}$$

uniformly on bounded x-intervals. If $t_\infty < \infty$ we also require that $s(t) \to 0$ for $t \to t_\infty$.

If $t_\infty = \infty$ and the first derivative of s exists and vanishes at ∞ then the function s is self-neglecting. If $t_\infty < \infty$ then s is self-neglecting if both s and s' vanish at t_∞. Any self-neglecting function is asymptotic to such a function. Hence if s is self-neglecting , then $s(t) = o(t)$ if $t_\infty = \infty$ and $s(t) = o(t_\infty - t)$ if $t_\infty < \infty$. A function which is asymptotic to a self-neglecting function is self-neglecting. For a continuous function s it suffices to assume pointwise convergence in (3.5) by Bloom's theorem (BGT [9], Section 2.11). The condition in Proposition 3.2 is formulated as: The function $s(\lambda) = 1/\sigma(\lambda)$ should be self-neglecting for $\lambda \to \lambda_\infty$.

A function ψ is *asymptotically parabolic at* $t_\infty \leq \infty$ if it is defined, convex and C^2 on a left neighbourhood of t_∞ with $\psi'' > 0$ and if $s = 1/\sqrt{\psi''}$ is self-neglecting at t_∞ (cf. BKR93 [1]).

By the above arguments any asymptotically parabolic function ψ satisfies

$$\psi(t+xs(t)) = \psi(t) + xs(t)\psi'(t) + x^2/2 + o(1) \qquad t \to t_\infty \tag{3.6}$$

uniformly on bounded x-intervals. For asymptotic normality of the exponential family X^λ it thus suffices that the cgf κ be asymptotically parabolic at λ_∞. Condition (3.6) on the cgf implies that the cgf's $\gamma_\lambda(\xi)$ of the standardised rv's U_λ converge to the standard normal cgf $\xi^2/2$ as $\lambda \to \lambda_\infty$.

Consider the following list of statements for $\lambda \to \lambda_\infty$:

AP1) κ is asymptotically parabolic at λ_∞;

AP2) $s = 1/\sqrt{\kappa''}$ is self-neglecting at λ_∞;

AP3) the derivative of $s(\lambda) = 1/\sigma(\lambda)$ vanishes at λ_∞, and so does $s(\lambda)$ if $\lambda_\infty < \infty$.

We arrive at another central result of this section.

Theorem 3.3 *Let X have cgf κ with upper endpoint λ_∞. The exponential family X^λ is asymptotically normal if and only if κ is asymptotically parabolic at λ_∞. Moreover, the statements AP1)-AP3) are equivalent.*

Proof The implications AP3) $\Rightarrow$ AP2) $\Rightarrow$ AP1) hold from the discussion after (3.5). Now assume that X^λ is asymptotically normal. Then the family is strongly asymptotically normal (Theorem 3.1) and hence all moments converge and, in particular,

$$EU_\lambda^3 \to 0 \qquad \lambda \to \lambda_\infty \tag{3.7}$$

holds. Also $\lambda_\infty \notin \Lambda$. So $\kappa(\lambda) \to \infty$ and $1/\sigma(\lambda)$ vanishes for $\lambda \to \lambda_\infty$ if $\lambda_\infty < \infty$. Furthermore, by Theorem 5.4 of BKR93 [1] it follows from relation AP3) that $\kappa(\lambda_\infty) = \infty$ if $\lambda_\infty < \infty$ as an asymptotically parabolic function. Since $\kappa'''(\lambda) = E(X^\lambda - \mu(\lambda))^3$ and $(1/\sigma(\lambda))' = -\kappa'''(\lambda)/(2\sigma^3(\lambda))$, AP3) is equivalent to the condition that $\kappa(\lambda_\infty) = \infty$ if $\lambda_\infty < \infty$ and (3.7). □

Relation AP1) implies strong asymptotic normality of the exponential family (Proposition 3.2). Hence in the context of exponential families, AN3) $\Rightarrow$ AN1) and the proof of Theorem 3.1 is complete.

3.1 Complements.

3.1.1 Higher derivatives; asymptotic equivalence.

The condition that a function is asymptotically parabolic is a condition on the second derivative. For cgf's this condition also determines the asymptotic behaviour of the higher derivatives.

Proposition 3.4 *Suppose the cgf κ is asymptotically parabolic at λ_∞. Define $\sigma(\lambda) = \sqrt{\kappa''(\lambda)}$ as in (3.2). Then for all integers $n > 2$:*

$$\kappa^{(n)}(\lambda)/\sigma^n(\lambda) \to 0 \qquad \lambda \to \lambda_\infty.$$

Proof Strong asymptotic normality of the associated exponential family implies $EU_\lambda^n \to EU^n$ for all $n \geq 1$. Hence the cgf γ_λ of U_λ has the property $\gamma_\lambda^{(n)}(0) \to \gamma^{(n)}(0)$ where $\gamma(\xi) = \xi^2/2$. (The relation also follows directly from the normal convergence of analytic functions.) □

Cumulant generating functions are C^∞ and convex. Given an asymptotically parabolic function ψ it is not hard to construct a convex C^∞ function which is asymptotic to ψ but which itself is not asymptotically parabolic. For cgf's this is not possible.

Proposition 3.5 *Let the rv X have mgf $K = e^\kappa$ with upper endpoint λ_∞. Suppose $K(\lambda) \sim e^{\psi(\lambda)}$ for $\lambda \to \lambda_\infty$ where ψ is asymptotically parabolic at λ_∞. Then*

1) $\kappa''(\lambda) \sim \psi''(\lambda)$ *for* $\lambda \to \lambda_\infty$*;*

2) κ *is asymptotically parabolic at* λ_∞*.*

Proof Set $b(\lambda) = \psi'(\lambda)$ and $a(\lambda) = \sqrt{\psi''(\lambda)}$. Then $1/a(\lambda)$ is self-neglecting and

$$\kappa(\lambda + \xi/a(\lambda)) - \kappa(\lambda) - b(\lambda)\xi/a(\lambda) \to \xi^2/2 \qquad x \in \mathbb{R}$$

since this holds for ψ, and the difference $\kappa(\lambda) - \psi(\lambda) = o(1)$.

It follows that $(X^\lambda - b(\lambda))/a(\lambda) \Rightarrow U$. So 2) holds by Theorem 3.2. One even has strong asymptotic normality, which implies convergence in law of the standardised rv's $U_\lambda = (X^\lambda - \mu(\lambda))/\sigma(\lambda)$. Khinchine's convergence of types theorem (see Feller [15], Lemma VIII.2.1) then gives $\sigma(\lambda) \sim a(\lambda)$ which is 1). □

3.1.2 Densities and the domain of attraction of the normal law.

If the exponential family is generated by a rv X with a continuous density f then each rv X^λ of the exponential family has a continuous density f_λ. Asymptotic normality of the exponential family does not imply convergence of the densities. Let g_λ denote the density of the standardised rv

$$U_\lambda = (X^\lambda - \mu(\lambda))/\sigma(\lambda) \tag{3.8}$$

with μ and σ the mean and variance of X^λ; see (3.2). Consider the following statements about convergence of the densities for $\lambda \to \lambda_\infty$:

D1) $g_\lambda \to g_{01}$ in $\mathcal{L}^1$;

D2) $g_\lambda \to g_{01}$ uniformly on $\mathbb{R}$;

D3) for all $M > 1$

$$\sup_{u \in \mathbb{R}} e^{M|u|}|g_\lambda(u) - g_{01}(u)| \to 0 \quad \lambda \to \lambda_\infty. \tag{3.9}$$

We shall prove the following: If X has a density f, which is *strongly unimodal*, i.e. $f = e^{-\varphi}$ with φ convex, then asymptotic normality of the exponential family is equivalent with D1)–D3).

We previously characterised the domain of attraction of the normal law for exponential families in terms of transforms and also gave conditions on the upper tail of F or the density f which guarantees that the associated exponential family is asymptotically normal. Here we collate some results about such conditions.

Theorem 3.6 *Suppose X has a bounded density f with upper endpoint x_∞, positive on a left neighbourhood of x_∞. Then the following hold:*

1) If $f(x) \sim e^{-\psi(x)}$ for $x \to x_\infty$, where ψ is asymptotically parabolic, then the exponential family X^λ is asymptotically normal and D3) holds.

2) If f is C^3 on a left neighbourhood of the upper endpoint x_∞, $f(x) \to 0$ as $x \to x_\infty$ and the function $\psi = -\log f$ satisfies $\psi'' > 0$ and $\psi'''(x)/(\psi''(x))^{3/2} \to 0$ as $x \to x_\infty$, then D3) holds.

Proof 1) First assume $f = e^{-\psi}$. Let $x_0 < x_\infty$. Then $\lambda = \psi'(x_0)$ is the slope of the convex function ψ in x_0. Set $a_0 = 1/\sqrt{\psi''(x_0)}$. Then (3.6) gives

$$\psi_\lambda(u) := \psi(x_0 + a_0 u) - \psi(x_0) - \lambda a_0 u \to u^2/2 \qquad x_0 \to x_\infty. \tag{3.10}$$

The density h_λ of $(X^\lambda - x_0)/a_0$ is $c_0 e^{-\psi_\lambda(u)}$ for some normalising constant $c_0 = c(\lambda) > 0$. Thus $h_\lambda \to g_{01}$ by (3.10) and convexity of ψ_λ. The convexity also gives (3.9). See BKR93 [1], Theorem 6.4 for further details.

2) The conditions imply that ψ is asymptotically parabolic. □

In Balkema, Klüppelberg and Stadtmüller [3] a number of Tauberian conditions were formulated which ensure that the rv X with asymptotically parabolic cgf κ has a density f with Gaussian tail. This means that the upper endpoint x_∞ is infinite and $f \sim e^{-\psi}$ for some asymptotically parabolic function ψ. The results of that paper were formulated in the framework of densities with upper endpoint $x_\infty = \infty$ but the theorem below remains valid in the case where $x_\infty < \infty$. For the proof of this theorem we refer to Section 2 in [3].

Theorem 3.7 *Let X have a strongly unimodal density f.*

1) The conditions D1) – D3) are equivalent.

2) If the exponential family X^λ is asymptotically normal, then f is bounded and $f(x) \sim e^{-\psi(x)}$ for $x \to x_\infty$ where ψ is asymptotically parabolic.

For the last statement, note that the weakest condition D1) implies asymptotic normality. Hence by Theorem 3.7 $f(x) \sim e^{-\psi(x)}$ for $x \to x_\infty$ with ψ asymptotically parabolic and D3) holds by Theorem 3.6.

3.1.3 Distributions and the domain of attraction of the normal law.

Comparable results to those stated in the previous subsection are valid in terms of the upper tail of the df F. We first note the following fact, which is an immediate consequence of the definition.

If ψ is asymptotically parabolic at $t_\infty > 0$, then so is $\psi(t) + \log t$.

Theorem 3.8 *Suppose the df F has upper endpoint x_∞ and tail $1 - F(x) \sim e^{-\psi(x)}$ for $x \to x_\infty$ where ψ is asymptotically parabolic at x_∞. Then the associated exponential family is asymptotically normal.*

Proof Define the bounded density $f^*(x) = e^x(1-F(x))/c$. Then $f^*(x) \sim e^{-\phi(x)}$ where $\phi(x) = \psi(x) - x + \log c$ is asymptotically parabolic since $\phi'' = \psi''$. So from Theorem 3.6, the exponential family generated by f^* is asymptotically normal. Let K be the mgf of F. We then have from Theorem 3.3 that

$$\log \int e^{\lambda x} f^*(x) dx = \log \int e^{(\lambda+1)x}(1 - F(x))dx/c = \log \left(\frac{K(1+\lambda)}{(1+\lambda)c} \right)$$

is asymptotically parabolic. This implies that $\log K(1 + \lambda)$ is asymptotically parabolic and hence so is $\log K$. □

The converse is false. Asymptotic normality of an exponential family does not imply that the underlying df has a tail $1-F(x) \sim e^{-\psi(x)}$ with ψ asymptotically parabolic. The tail need not even be asymptotically continuous.

Example 3.9 The Poisson distributions form an exponential family which is well known to be asymptotically normal. The tail of a Poisson distribution with expectation 1 is very irregular: $(1 - F(n-))/(1 - F(n)) \sim n \not\to 1$ for $n \to \infty$. □

One can introduce measures with increasingly smooth densities by setting $f_1 = 1 - F$ and $f_{n+1}(x) = \int_x^\infty f_n(t)dt$. The cgf's corresponding to f_n are $\kappa(\lambda) - n \log \lambda$, and these are asymptotically parabolical if and only if κ is. If F is the Poisson distribution with expectation 1 then none of the densities f_n is of the form $f_n \sim e^{-\psi_n}$ with ψ_n asymptotically parabolic, even though they all generate exponential families which are asymptotically normal.

3.1.4 Asymptotically parabolic functions.

Here are some examples of asymptotically parabolic functions. We seek functions which are convex and unbounded at their upper endpoint. The function x^2 is asymptotically parabolic at infinity, and so are the functions x^α for $\alpha > 1$, $x - x^\alpha$ for $\alpha \in (0, 1)$ and e^{x^α} for $\alpha > 0$. Positive linear combinations of such functions are again asymptotically parabolic. The functions $1/(c - x)^\alpha$ with $\alpha > 0$ and $|\log(c-x)|^\beta$ for $\beta > 1$ are asymptotically parabolic at the point c.

Not every asymptotically parabolic function is the cgf of a probability measure. Cgf's are very special convex functions. A mgf is totally positive, its derivatives are all strictly positive on Λ, and it extends to an analytic function on the vertical strip $\{\Re z \in \Lambda\}$. So one may ask which of the functions in the example above is asymptotic to a cumulant generating function. The final result of the section addresses this question.

We shall make use of a beautiful result which links the asymptotic behaviour of a density and its mgf. This result is based on the conjugate Legendre transform ψ^* of a convex function ψ with domain D

$$\psi^*(t) = \sup\{xt - \psi(x) \mid x \in D\}. \tag{3.11}$$

If $f = e^{-\psi}$ is a strongly unimodal density and ψ is asymptotically parabolic, then (2.9) with $c = 0$ gives

$$K(\lambda) \sim \sqrt{2\pi} a_\lambda f(b_\lambda) e^{\lambda b_\lambda} \sim \sqrt{2\pi}\sigma(\lambda) e^{\psi^*(\lambda)} \qquad \lambda \to \lambda_\infty \tag{3.12}$$

if we choose b_λ so that $\psi'(b_\lambda) = \lambda$, thus maximising $\lambda x - \psi(x)$ in (3.11). In that case $a_\lambda \sim \sigma(\lambda)$. One can get rid of the factor $\sqrt{2\pi}\sigma(\lambda)$ in (3.12) since this function is practically constant (flat) on intervals of length $\sigma(\lambda)$.

Theorem 3.10 *Let φ be asymptotically parabolic in λ_∞. Then there exists a rv X with mgf K so that $K(\lambda) \sim e^{\varphi(\lambda)}$ for $\lambda \to \lambda_\infty$. We may choose X to have a strongly unimodal density.*

Proof We may assume that φ is convex and that φ'' is continuous and strictly positive. Let $t_\infty = \sup\{\varphi'(\lambda) \mid \lambda < \lambda_\infty\}$ and let $\psi(t) = \varphi^*(t)$ be the conjugate Legendre transform of $\varphi(\lambda)$. The function ψ is defined on a left neighbourhood of t_∞ and is asymptotically parabolic in t_∞ by Theorem 5.3 in BKR93 [1] with scale function $a(t) = 1/\sqrt{\psi''(t)}$. Now apply Theorem 6.6 in [1] with a bounded density $f \sim \gamma e^{-\psi}$ where $\gamma(t) = 1/(\sqrt{2\pi}a(t))$. The function γ is flat (see [1], p.580) for a since a is self-neglecting. This implies that we may choose f strongly unimodal. Note that $\psi^* = \varphi^{**} = \varphi$. Hence the mgf K of f satisfies $K(\lambda) \sim e^{\varphi(\lambda)}$ by relation (6.6) in [1]. □

4 Domains of attraction for the gamma limits.

For the domains of attraction of the gamma limits there is a simple and complete description in terms of regular variation. In fact the limit theory for exponential families with a gamma limit leads to a novel approach to regular variation. We shall obtain a new derivation of Karamata's Tauberian theorem. It will also be seen that smoothly varying functions occur naturally in the limit theory of exponential families.

For the definition and properties of regular variation we refer to BGT [9], Feller [15], Embrechts, Klüppelberg, and Mikosch [13], Geluk and de Haan [16] or Resnick [18];

Let γ_α for $\alpha > 0$ denote the probability distribution on $(0, \infty)$ with density

$$g_\alpha(y) = y^{\alpha-1}e^{-y}/\Gamma(\alpha) \qquad y > 0. \tag{4.1}$$

The mgf $K(\lambda) = 1/(1-\lambda)^\alpha$ of the distribution γ_α is finite on $(-\infty, 1)$. The gamma variable V with density (4.1) satisfies a stability relation. For a normal rv the Esscher transform has the effect of a translation, for a gamma rv the Esscher transform has the effect of a multiplication:

$$E^\xi V = V^\xi \stackrel{\mathrm{d}}{=} \frac{V}{1-\xi} \qquad \xi < 1. \tag{4.2}$$

We are interested in rv's in the domain of attraction of V and of the rv $\overline{V}$ with probability distribution $\overline{\gamma}_\alpha$, mgf $\overline{K}(\lambda) = 1/(1+\lambda)^\alpha$, $\lambda > -1$, and density

$$\overline{g}_\alpha(y) = (-y)^{\alpha-1}e^{y}/\Gamma(\alpha) \qquad y < 0. \tag{4.3}$$

The following is a first important result of this section.

Proposition 4.1 *1) If $X \in \mathcal{D}(\gamma_\alpha)$, then $\lambda_\infty < \infty$ and*

$$(\lambda_\infty - \lambda)X^\lambda \Rightarrow V \qquad \lambda \to \lambda_\infty\,. \tag{4.4}$$

2) If $X \in \mathcal{D}(\overline{\gamma}_\alpha)$, then $x_\infty < \infty$ and

$$\lambda(X^\lambda - x_\infty) \Rightarrow \overline{V} \qquad \lambda \to \infty\,. \tag{4.5}$$

Proof We make use of the following fact (see (2.2) in BKR99 [2]):
If $Ax = ax + b$ for $a > 0, b \in \mathbb{R}$, and $\lambda \in \Lambda$, then

$$AE^\lambda X = E^{\lambda/a}AX. \tag{4.6}$$

1) There exist (BKR99 [2], Lemma 2.8) positive affine transformations A_λ depending continuously on the parameter λ, so that as $\lambda \to \lambda_\infty$

$$U_\lambda := A_\lambda X^\lambda \Rightarrow V.$$

Let $\xi = 1/2$. For some $\lambda_0 < \lambda_\infty$, $Ee^{\xi U_\lambda}$ is finite for $\lambda \in [\lambda_0, \lambda_\infty)$. Use (4.6) to see that it is possible to choose $\lambda_0 < \lambda_1 < \cdots$ and positive affine transformations

$B_n x = (x - b_n)/a_n$ so that the variables $Z_n = U_{\lambda_n}$ satisfy $B_{n+1}E^{\xi}Z_n = Z_{n+1}$ and $Z_0 = B_0 E^{\lambda_0} X$. Then $B_n x \to Qx = x/2$ by (4.2). This means that $a_n \to 2$ and $b_n \to 0$. Observe that from repeated use of (4.6),

$$Z_{n+1} = B_{n+1}E^{\xi}B_n E^{\xi} \cdots B_1 E^{\xi} A_{\lambda_0} E^{\lambda_0} X =: D_n E^{\xi_n} X$$

with

$$\xi_n = \lambda_0 + \xi/a_0 + \cdots + \xi/(a_0 \cdots a_n) \uparrow \xi_\infty < \infty$$

since $a_n \to 2$, and $D_n = B_n \circ \cdots \circ B_0$. Set $D_n x = c_n x + d_n$. Then $c_n = 1/(a_0 \cdots a_n) \to 0$ and hence $\|D_n\| := \sqrt{(\log c_n)^2 + d_n^2} \to \infty$ and therefore, by BKR99 [2], Proposition 2.10, $\xi_\infty = \lambda_\infty \notin \Lambda$. Since $a_n \to 2$, we have

$$\lambda_\infty - \xi_n \sim \xi/(a_0 \cdots a_n) \sim c_n/2.$$

The relation $D_n x = c_n x + d_n$ gives $d_n = D_n(0) = B_n(D_{n-1}(0)) = B_n(d_{n-1}) = (d_{n-1} - b_n)/a_n$. Due to $b_n \to 0$ and $a_n \to 2$, we get $d_n \to 0$ and

$$(\lambda_\infty - \xi_n)E^{\xi_n} X \Rightarrow V.$$

Finally write $\lambda = \xi_n + \theta_n(\lambda_\infty - \xi_n)$ for $\lambda \in [\xi_n, \xi_{n+1})$. Then $\theta_n = \theta_n(\lambda) \in [0, 2/3]$ eventually, $V_\lambda \stackrel{\mathrm{d}}{=} B_n^{\theta_n} Z_n$ with $Q_{\theta_n} B_n^{\theta_n} \to \mathrm{id}$ uniformly in $\theta_n \in [0, 2/3]$. This implies

$$(\lambda_\infty - \lambda)X^\lambda = (1 - \theta_n)(\lambda_\infty - \xi_n)E^{\theta_n} U_{\lambda_n} = (1 - \theta_n)E^{\theta_n} Z_n \Rightarrow V$$

which is the desired relation (4.4).

2) The proof is similar. Take $\xi = 1$. Then $B_n x \to 2x$ and $D_n x = c_n(x + \delta_n)$ with c_n as above and

$$\delta_n = \frac{d_n}{c_n} = \frac{d_{n-1} - b_n}{a_n c_n} = \delta_{n-1} - \frac{b_n}{c_{n-1}} \to \delta_\infty < \infty.$$

We thus find $\xi_n \sim 2\xi/(a_0 \cdots a_{n-1}) \sim c_n \sim \lambda_n$ and $D_n E^{\xi_n} X \Rightarrow \overline{V}$ gives

$$\lambda_n(X_{\lambda_n} - x_\infty) \Rightarrow \overline{V}.$$

Assume $x_\infty = 0$ for simplicity. Set $Z_n = \lambda_n X_{\lambda_n}$. Then $Z_n \Rightarrow \overline{V}$ implies $\theta E^\theta Z_n \Rightarrow \overline{V}$ uniformly in $\theta \in [1, 3]$. Hence writing $\lambda = \theta_n \lambda_n$ for $\lambda_n \le \lambda < \lambda_{n+1}$ we find

$$\lambda X^\lambda = \theta_n E^{\theta_n} Z_n \Rightarrow \overline{V}.$$

For general x_∞ one obtains (4.5). □

4.1 Regular variation

By Proposition 4.1,1), if $F \in \mathcal{D}(\gamma_\alpha)$ then λ_∞ is finite. The measure $d\mu(y) = e^{\lambda_\infty y} dF(y)$ has infinite total mass, since $\lambda_\infty \notin \Lambda$ implies $K(\lambda) \to \infty$ for

$\lambda \to \lambda_\infty$. Note however that

$$\widehat{M}(\tau) = \int e^{\tau y} d\mu(y) = K(\lambda_\infty + \tau) < \infty \qquad \tau_0 < \tau < 0 \tag{4.7}$$

for some $\tau_0 < 0$.

The exponential family generated by the Radon measure μ consists of rv's Y^τ with distribution

$$dG_\tau(y) = e^{\tau y} d\mu(y)/\widehat{M}(\tau) \qquad \tau_0 < \tau < 0.$$

This is also the exponential family generated by the df F up to a shift in the parametrization: $G_\tau = F_\lambda$ for $\lambda = \lambda_\infty + \tau$.

The df $M(y) = \mu((-\infty, y])$ of the measure μ plays a key role in the description of $\mathcal{D}(\gamma_\alpha)$. Consider the following examples.

Example 4.2 (i) Let μ be a Radon measure on $\mathbb{R}$ with density m which vanishes off $[0, \infty)$. Suppose $m(x) \to 1$ for $x \to \infty$. Let Y^τ, $\tau < \tau_\infty = 0$, be the exponential family generated by μ. The rv Y^τ has density $e^{\tau y} m(y)/\widehat{M}(\tau)$. Set $\xi = -\tau$. The normalized rv $V_\tau = \xi Y^\tau$ has density $e^{-y} m(y/\xi)/(\xi \widehat{M}(\tau))$ which converges to the standard exponential density for $\xi \downarrow 0$ since $m(y/\xi) \to 1$. Note that $\widehat{M}(\tau) \sim 1/\xi$ for $\tau \uparrow 0$ and $M(y) = \mu((-\infty, y]) \sim y$ for $y \to \infty$.

(ii) More generally, start with a measure μ on $\mathbb{R}$ with distribution function $M(y) = \mu((-\infty, y])$ which varies regularly at ∞ with exponent $\alpha > 0$. Assume that $\int e^{\lambda_0 y} d\mu(y)$ is finite for some $\lambda_0 < 0$. The corresponding exponential family Y^λ, $\lambda_0 \leq \lambda < 0$, with distribution

$$d\pi_\lambda(y) = e^{\lambda y} d\mu(y)/\widehat{M}(\lambda) \qquad \widehat{M}(\lambda) = \int e^{\lambda y} d\mu(y) \qquad \lambda_0 \leq \lambda < 0$$

satisfies $V_\lambda = (\lambda_\infty - \lambda) X^\lambda \Rightarrow V$ with $\lambda_\infty = 0$.
Proof Regular variation with exponent α implies for $\beta > \alpha$ that $M(y) = o(y^\beta)$ for $y \to \infty$. Hence $e^{\lambda y} d\mu(y)$ is a finite measure for $\lambda_0 \leq \lambda < 0$. For $\xi > 0$ define the measure μ_ξ with df $\mu_\xi((-\infty, y]) = \mu((-\infty, y/\xi]) = M(y/\xi)$. Let $A(\xi) = M(1/\xi)\Gamma(\alpha + 1)$. Then

$$\frac{M(y/\xi)}{A(\xi)} \to \frac{y_+^\alpha}{\Gamma(\alpha+1)} \qquad \text{weakly on } \mathbb{R} \text{ for } \xi \downarrow 0.$$

Note that for $y \leq 0$,

$$\frac{M(y/\xi)}{M(1/\xi)} \leq \frac{M(0)}{M(1/\xi)} \to 0,$$

as $\xi \downarrow 0$, since $M(1/\xi) \to \infty$ as a consequence of regular variation. The finite measures $d\nu_\xi(y) = e^{-y}d\mu_\xi(y)/A(\xi)$, $\xi > 0$, satisfy

$$d\nu_\xi(y) \to e^{-y}y_+^{\alpha-1}dy/\Gamma(\alpha) \qquad \xi \downarrow 0$$

vaguely on $[-\infty, \infty)$ and even weakly since $\nu_\xi(\mathbb{R}) \to \int e^{-y}dy_+^\alpha/\Gamma(\alpha+1) = 1$ because of the relation $M(y) = o(y^\beta)$ for $\beta > \alpha$ mentioned above. We conclude that $A(\xi) \sim \widehat{M}(-\xi) = \int e^{-\xi y}d\mu(y)$ for $\xi \downarrow 0$ and hence for $\lambda = -\xi$ the probability measure $e^{-y}d\mu_\xi(y)/\widehat{M}(\lambda)$ of ξY^λ tends to γ_α weakly for $\lambda \uparrow 0$. □

The ideas of these examples suggest the general results of the next setion.

4.2 Domain of attraction of the positive gamma law.

Suppose $\lambda \to \lambda_\infty < \infty$. Let the limit variable $V > 0$ have distribution γ_a. As in the case of a normal limit distribution a number of limit relations turn out to be equivalent for a gamma limit:

Theorem 4.3 *Let V have probability distribution γ_α on $[0, \infty)$ for some parameter $\alpha > 0$. Let the rv X with df F have mgf K with upper endpoint $\lambda_\infty < \infty$. Let*

$$M(y) = \mu((-\infty, y]) = \int_{(-\infty,y]} e^{\lambda_\infty x}dF(x).$$

Then the following statements are equivalent for $\lambda \to \lambda_\infty$.

G1) $V_\lambda = (\lambda_\infty - \lambda)X^\lambda \Rightarrow V$;

G2) $EV_\lambda = (\lambda_\infty - \lambda)EX^\lambda \to EV = \alpha$;

G3) $EV_\lambda^n \to EV^n$ for $n \in \mathbb{N}$;

G4) $K_\lambda(\tau) = Ee^{\tau V_\lambda} \to 1/(1-\tau)^\alpha$ for $\tau < 1$.

G5) M varies regularly at ∞ with exponent α;

G6) K varies regularly at λ_∞ with exponent $-\alpha$; that is

$$\lim_{t\downarrow 0} \frac{K(\lambda_\infty - tx)}{K(\lambda_\infty - t)} = x^{-\alpha} \quad x > 0.$$

G7) the df M and the mgf K are asymptotically related for $\xi = \lambda_\infty - \lambda \downarrow 0$:

$$\frac{M(y/\xi)}{K(\lambda_\infty - \xi)} \to \frac{y_+^\alpha}{\Gamma(\alpha+1)} \quad \text{weakly on } \mathbb{R}. \tag{4.8}$$

Proof We proceed in six steps. Set $\xi = \lambda_\infty - \lambda \downarrow 0$.

G6) $\Longleftrightarrow$ G4) since the mgf of V_λ is $\tau \to K(\lambda_\infty - \xi + \tau\xi)/K(\lambda_\infty - \xi)$.

G4) $\Rightarrow$ G1) Convergence of mgf's implies weak convergence.

G1) $\Rightarrow$ G7) The rv V_ξ has distribution $d\pi_\xi^*(y) = e^{-y}d\mu_\xi(y)/K(\lambda)$, where again $\mu_\xi((-\infty, y]) = \mu((-\infty, y/\xi])$. Then

$$d\pi_\xi^*(y) \to e^{-y}y^{\alpha-1}dy/\Gamma(\alpha) \qquad \xi \downarrow 0. \tag{4.9}$$

Multiply by e^y and integrate over $(-\infty, y]$. Since μ_ξ has df $M(y/\xi)$ we obtain (4.8).

G7) $\Rightarrow$ G5) is obvious.

G7) $\Rightarrow$ G6) by symmetry: $K(\lambda_\infty - \eta\xi)/M(1/\xi) \to \Gamma(\alpha+1)/\eta^\alpha$ on $(0, \infty)$.

G5) $\Rightarrow$ G1) is proved in the Example 4.2 (ii) above.

So we have established G1) $\Rightarrow$ G7) $\Rightarrow$ G5) $\Rightarrow$ G1) and G1) $\Rightarrow$ G7) $\Rightarrow$ G6) $\Rightarrow$ G4) $\Rightarrow$ G1); i.e. the equivalence of $G1)$ and $G4) - G7)$. Note that G4) implies G2) and G3), full equivalence is established in Theorem 4.6. □

Remark 4.4 (a) Note that we have proven Karamata's celebrated Tauberian theorem G6) $\Rightarrow$ G5).

(b) We have also proven that weak convergence implies convergence of the mgf's for exponential families with limit distribution γ_α; cf. Theorem 2.3. □

Proposition 4.5 *If $F \in \mathcal{D}(\gamma_\alpha)$, then the mgf K varies smoothly at $\lambda_\infty < \infty$ with exponent $-\alpha$; i.e. $\kappa'(\lambda) \to -\alpha$ and $\kappa^{(n)}(\lambda) \to 0$ for all $n \geq 2$.*

Proof Let $\kappa = \log K$ denote the cgf in the theorem above and set

$$\varphi(t) = \kappa(\lambda_\infty - e^{-t}).$$

Then regular variation of K with exponent $-\alpha$ in λ_∞ just means that for $t \to \infty$

$$\varphi(t+x) - \varphi(t) = \log\left(\frac{K(\lambda_\infty - e^{-(t+x)})}{K(\lambda_\infty - e^{-t})}\right) \to \log(e^{\alpha x}) = \alpha x \tag{4.10}$$

uniformly on bounded x-intervals in $\mathbb{R}$. The function φ is analytic and hence $\varphi'(t) \to \alpha$ and $\varphi^{(n)}(t) \to 0$ for $n \geq 2$. This means that the mgf K varies smoothly at λ_∞. See BGT [9], Section 1.8. □

Theorem 4.6 *Suppose X is a rv with cgf κ with upper endpoint $\lambda_\infty < \infty$. Then $X \in \mathcal{D}(\gamma_\alpha)$ if and only if $(\lambda_\infty - \lambda)EX^\lambda \to \alpha$ for $\lambda \to \lambda_\infty$.*

Proof Necessity of the condition has been proved above: G1) $\Rightarrow$ G2). For sufficiency we use Proposition 4.5 and note that the condition can be formulated in terms of the function $\widehat{M}(\tau) = K(\lambda_\infty + \tau)$ for $\tau < 0$ (see (4.7)) as

$$|\tau|\widehat{M}'(\tau)/\widehat{M}(\tau) \to -\alpha \qquad \tau \uparrow 0.$$

This is the well-known von Mises sufficient condition for regular variation with exponent $-\alpha$, giving that K is regularly varying; i.e. G6). See BGT [9]. □

4.3 Domain of attraction of the negative gamma law.

The theory for the domain of attraction is, in this case, even simpler.

Let X have df $F \in \mathcal{D}(\overline{\gamma}_\alpha)$ and mgf K. By Proposition 4.1,2), the upper endpoint x_∞ of F is finite and we may assume $x_\infty = 0$. Since F is continuous at its upper endpoint the mgf $K(\lambda)$ vanishes for $\lambda \to \infty$. The probability measure η of $-X$ has df $H(y) = 1 - F(-y-)$. The positive rv $-\lambda X^\lambda$ has probability distribution

$$e^{-y}d(\eta_\lambda)(y)/K(\lambda)$$

and η_λ has df $H(\cdot/\lambda)$. The following two weak limit relations for $\lambda \to \infty$ are equivalent:

$$\begin{aligned} e^{-y}d(\eta_\lambda)(y)/K(\lambda) &\to e^{-y}y^{\alpha-1}dy/\Gamma(\alpha) \\ H(y/\lambda)/K(\lambda) &\to y^\alpha/\Gamma(\alpha+1). \end{aligned}$$

Theorem 4.7 *Let $\overline{V}$ have probability distribution $\overline{\gamma}_\alpha$ on $(-\infty, 0]$ for some $\alpha > 0$. Let X have df F with $x_\infty < \infty$ and mgf K with $\lambda_\infty = \infty$. The following statements are equivalent:*

$\overline{G}$1) $\overline{V}_\lambda = \lambda(X^\lambda - x_\infty) \Rightarrow \overline{V}$ *for* $\lambda \to \infty$;

$\overline{G}$2) $E\overline{V}_\lambda = \lambda(x_\infty - EX^\lambda) \to E\overline{V} = \alpha \qquad \lambda \to \infty$;

$\overline{G}$3) $E\overline{V}_\lambda^n \to E\overline{V}^n$ *for* $n \in \mathbb{N}$;

$\overline{G}$4) $K_\lambda(\xi) = Ee^{\xi V_\lambda} \to \dfrac{1}{(1+\xi)^\alpha}$ *for* $\xi > -1$ *for* $\lambda \to \infty$;

$\overline{G}$5) $1 - F$ *varies regularly with exponent* α *in* x_∞;

$\overline{G}$6) $e^{-x_\infty \lambda}K(\lambda)$ *varies regularly in* ∞ *with exponent* $-\alpha$;

$\overline{G}7$) *the tail* $1 - F$ *and the mgf* K *are asymptotically related: for* $x > 0$

$$\frac{1 - F(x_\infty - x/\lambda)}{e^{-\lambda x_\infty} K(\lambda)} \to \frac{x^\alpha}{\Gamma(\alpha+1)} \qquad \lambda \to \infty.$$

The proof is similar to that of Theorem 4.3 and therefore omitted.

Similarly setting $\varphi(t) = \kappa(e^t)$ we find $\varphi(t+x) - \varphi(t) \to -\alpha x$ which proves

Proposition 4.8 *Let* $F \in \mathcal{D}(\overline{\gamma}_\alpha)$ *with mgf* K*, then* K *varies smoothly in* ∞ *with exponent* $-\alpha$*.*

5 Applications

We show that limit laws for exponential families can be applied to prove tail-accuracy of certain approximating densities. For densities in in the domain of attraction of the normal law results of this kind and some statistical examples are in Barndorff-Nielsen and Klüppelberg [6]; first multivariate results can be found in [7].

5.1 Convolution closure properties.

Consider the convolution of df's and densities from the domains of attraction giving emphasis to convolving the positive gamma and the normal distribution; the negative gamma distribution can be treated analogously to the positive one. The parameter α appears in the domain of attraction of a gamma distribution (4.1), and we denote the corresponding domain of attraction by $\mathcal{D}(\alpha)$ for $\alpha > 0$. The normal distribution as a member of the extended gamma family corresponds to $\alpha = \infty$; hence we denote its domain of attraction by $\mathcal{D}(\infty)$.

Proposition 5.1 *Suppose both* F *and* $G \in \mathcal{D}(\infty)$ *and assume their mgf's have the same upper endpoint. Then* $F * G \in \mathcal{D}(\infty)$*.*

Proof Notice that the cgf of $F * G$ is the sum $\kappa = \kappa_F + \kappa_G$ of the factors, and the variances add. Hence $\sigma \geq \sigma_F, \sigma_G$. If $1/\sigma_F$ and $1/\sigma_G$ are self-neglecting, then so is $1/\sigma$. The result follows then by Theorem 3.3. □

In BKR93 [1] a slightly more general class of densities than in Section 2.4 has been introduced aiming at convolution closure.

Proposition 5.2 (BKR93 [1])
Let $f_i(t) = \gamma_i(t)e^{-\psi_i(t)}$, $t \leq t_{i\infty}$, *for* $i = 1, 2$, *where* ψ_i *are asymptotically parabolic with self-neglecting functions* $s_i = 1/\sqrt{\psi_i''}$ *and*

$$\gamma_i(t + xs_i(t))/\gamma_i(t) \to 1 \quad t \to t_{i\infty}.$$

Denote

$$f(t) = f_1 * f_2(t) = \int f_1(t-y)f_2(y)dy\,,$$

then $f(t) \sim \gamma(t)e^{-\psi(t)}$ *and* γ *and* ψ *have the same properties. Furthermore, they can be expressed in terms of the* γ_i *and* ψ_i*; see BKR93 [1] for details.*

We now turn to $\mathcal{D}(\alpha)$ for finite α and start with a convolution result, which may be compared to Cline [11], Theorem 3.4.

Proposition 5.3 *Suppose* $F \in \mathcal{D}(\alpha_1)$, $G \in \mathcal{D}(\alpha_2)$ *for* $\alpha_1, \alpha_2 < \infty$. *If the mgf's have the same upper endpoint* λ_∞ *then* $F * G \in \mathcal{D}(\alpha_1 + \alpha_2)$.

Proof The mgf of $F * G$ is the product of the mgf of F and the mgf of G, hence it varies regularly in λ_∞ with exponent $-(\alpha_1 + \alpha_2)$. Here we use G6) of Theorem 4.4. □

5.2 On the tail accuracy of the saddlepoint and gamma approximation

Let f be a density, defined and positive on an interval I that is unbounded above. The *(unnormalised) saddlepoint approximation* to $f(x)$ may be expressed as

$$f^\dagger(x) = \frac{1}{\sqrt{2\pi\kappa''(\lambda)}}\, e^{-(\lambda x - \kappa(\lambda))} \tag{5.1}$$

where κ denotes the cgf and λ is the saddlepoint, i.e. it satisfies $\kappa'(\lambda) = x$. The ratio $f^\dagger(x)/f(x)$ expresses the relative accuracy of the saddlepoint approximation and we obtain immediately from (3.12) that $f^\dagger(x) \sim f(x)$ as $x \to \infty$ and hence for the relative error

$$RE^\dagger(x) = \left|\log\left(f^\dagger(x)/f(x)\right)\right| \to 0 \qquad x \to \infty\,.$$

Now assume that $f(x) \sim e^{-x}\,x^{\alpha-1}\,\ell(x)$, $x \to \infty$, for $\alpha > 0$ and $\ell \in SV$ (i.e. $\lim_{x\to\infty} \ell(xt)/\ell(x) = 1$ for all $t > 0$). Then $\lambda_\infty = 1$ and $F \in \mathcal{D}(\alpha)$ by Theorem 4.3. Indeed, it has been shown already in Theorem 7.1 of Daniels [12] that the associated exponential family is asymptotically gamma. By an immediate consequence of smooth regular variation (cf. Proposition 4.8) we obtain for the derivatives of the mgf and the cgf

$$K^{(j)}(\lambda) \sim \frac{\Gamma(b+j)}{(1-\lambda)^{\alpha+j}} \ell\left(\frac{1}{1-\lambda}\right) \qquad j \in \mathbb{N}_0 \,,$$

$$\kappa^{(j)}(\lambda) \sim \frac{\alpha}{(1-\lambda)^j} \qquad j \in \mathbb{N}\,.$$

Furthermore, since $f(x)/\overline{F}(x) \to 1$ as $x \to \infty$, (5.2) implies that

$$f^{\dagger}(x) \sim \frac{1-\lambda}{\sqrt{2\pi\,\alpha}} \, \frac{\Gamma(\alpha)}{(1-\lambda)^{\alpha}} \;\; \ell\left(\frac{1}{1-\lambda}\right) \; e^{-\lambda x}$$

and for λ satisfying $\kappa'(\lambda) = \frac{\alpha}{1-\lambda}(1+o(1)) = x$ as $\lambda \to 1$ (see Theorem 4.6), we obtain

$$\begin{aligned} f^{\dagger}(x) &\sim \frac{\Gamma(\alpha)}{\sqrt{2\pi\,\alpha}} \, \left(\frac{x}{\alpha}\right)^{\alpha-1} \;\; \ell\left(\frac{x}{\alpha}\right) \; e^{-(x-\alpha)(1+o(1))} \\ &\sim \frac{\Gamma(\alpha)}{\sqrt{2\pi\,\alpha}} \, \alpha^{-(\alpha-1)} \, x^{\alpha-1} \, \ell(x) \, e^{-x} \, e^{\alpha} \\ &= \frac{\Gamma(\alpha)\, e^{\alpha}}{\sqrt{2\pi\,\alpha}\;\alpha^{\alpha-1}} \, f(x) \qquad x \to \infty\,. \end{aligned}$$

Hence $RE^{\dagger}(x)$ is bounded and independent of x.

On the other hand, for densities in the domain of attraction of a gamma distribution, a gamma approximation as e.g. suggested by Bower is more appropriate [cf. Beard, Pentikäinen and Pesonen [8], see also Jensen [17], equation (3.7)]. The gamma approximation is defined as follows.

$$f^{\dagger\dagger}(x) = \frac{\kappa'(\lambda)}{\kappa''(\lambda)} \, \gamma\left(\frac{(\kappa'(\lambda))^2}{\kappa''(\lambda)}\right) \; e^{-(\lambda x - \kappa(\lambda))} \tag{5.2}$$

where $\gamma(u) = u^{u-1}\, e^{-u}/\Gamma(u)$ and λ is such that $\kappa'(\lambda) = x$. We use Theorem 4.6 which gives $\kappa'(\lambda) = x \sim \alpha/(1-\lambda)$ and hence $\lambda = 1 - a/x\;(1+o(1))$, which implies that

$$f^{\dagger\dagger}(x) \sim (1-\lambda)\;\gamma(\alpha)\; e^{-x}\; e^{\alpha(1+o(1))}\;\frac{\Gamma(\alpha)}{(1-\lambda)^{\alpha}}\;\ell\left(\frac{1}{1-\lambda}\right)$$

$$\sim e^{-x} \left(\frac{\alpha}{1-\lambda}\right)^{\alpha-1} \ell\left(\frac{\alpha}{1-\lambda}\right)$$
$$\sim e^{-x}\, x^{\alpha-1}\, \ell(x)$$
$$= f(x)\,, \qquad x \to \infty\,.$$

Hence

$$RE^{\dagger\dagger}(x) = \left|\log\left(f^{\dagger\dagger}(x)/f(x)\right)\right| \to 0 \quad \text{as } x \to \infty\,,$$

i.e. the gamma approximation becomes exact in the tail.

References

[1] Balkema, A.A., Klüppelberg, C. and Resnick, S.I. (1993) Densities with Gaussian tails. *Proc. London Math. Soc.* **66**, 568-588.

[2] Balkema, A.A., Klüppelberg, C. and Resnick, S.I. (1999) Limit laws for exponential families. *Bernoulli* **5**, 951-968.

[3] Balkema, A.A., Klüppelberg, C. and Stadtmüller, U. (1995) Tauberian results for densities with Gaussian tails. *J. London Math. Soc.* **51**, 383-400.

[4] Bar-Lev, S.K. and Casalis, M. (1994) Les familles exponentielles naturelles reproduisantes. *C. R. Acad. Sci. Paris Ser. I Math.* **319** 1323–1326.

[5] Bar-Lev, S.K. and Casalis, M. (2002) A classification of reproducible Natural Exponential Families in the broad sense. *J. Theoret. Probability.* To appear.

[6] Barndorff-Nielsen, O.E. and Klüppelberg, C. (1992) A note on the tail accuracy of the univariate saddlepoint approximation. *Ann. Toulouse (Sér. 6)* **1**, 5-14.

[7] Barndorff-Nielsen, O.E. and Klüppelberg, C. (1999) Tail exactness of multivariate saddlepoint approximations. *Scand. J. Stat.* **26**, 1-12.

[8] Beard, R.E., Pentikäinen, T. and Pesonen, R. (1984) *Risk Theory*, 3rd edition. Chapman and Hall, London.

[9] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1987) *Regular Variation.* Cambridge University Press.

[10] Breiman, L. (1965) On some limit theorems similar to the arc-sine law. *Theor. Probab. Appl.* 10, 323-331.

[11] Cline, D.B.H. (1986) Convolution tails, product tails and domains of attraction. *Probab. Th. Rel. Fields* 72, 529-557.

[12] Daniels, H.E. (1954) Saddlepoint approximations in statistics. *Ann. Math. Statist.* 25, 631-650.

[13] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997) *Modelling Extremal Events for Insurance and Finance.* Springer, Berlin.

[14] Feigin, P. and Yashchin, E. (1983) On a strong Tauberian result. *Z. Wahrsch. verw. Gebiete* 65, 35-48.

[15] Feller, W. (1966) *An Introduction to Probability Theory and Its Applications.* Vol.II. Wiley, New York.

[16] Geluk, J.L. and de Haan, L. (1980) *Regular Variation, Extensions and Tauberian Theorems.* CWI, Mathematisch Centrum, Amsterdam.

[17] Jensen, J.L. (1988) Uniform saddlepoint approximations. *Adv. Appl. Prob.* 20, 622-634

[18] Resnick, S.I. (1987) *Extreme Values, Regular Variation, and Point Processes.* Springer, New York.