Bernoulli **5**(6), 1999, 951–968

Limit laws for exponential families

AUGUST A. BALKEMA[1], CLAUDIA KLÜPPELBERG[2] and SIDNEY I. RESNICK[3]

[1]*Department of Mathematics, University of Amsterdam, NL-1018 TV Amsterdam, Netherlands. E-mail: guus@wins.uva.nl*
[2]*Center for Mathematical Sciences, Munich University of Technology, D-80290 Munich, Germany. E-mail: cklu@ma.tum.de*
[3]*Operations Research and Industrial Engineering, Cornell University, Ithaca, NY 14853-7501 USA. E-mail: sid@orie.cornell.edu*

For a real random variable X with distribution function F, define

$$\Lambda := \{\lambda \in \mathbb{R} : K(\lambda) := \mathrm{E}e^{\lambda X} < \infty\}.$$

The distribution F generates a natural exponential family of distribution functions $\{F_\lambda, \lambda \in \Lambda\}$, where

$$\mathrm{d}F_\lambda(x) := e^{\lambda x}\mathrm{d}F(x)/K(\lambda), \qquad \lambda \in \Lambda.$$

We study the asymptotic behaviour of the distribution functions F_λ as λ increases to $\lambda_\infty := \sup \Lambda$. If $\lambda_\infty = \infty$ then $F_\lambda \downarrow 0$ pointwise on $\{F < 1\}$. It may still be possible to obtain a non-degenerate weak limit law $G(y) = \lim F_\lambda(a_\lambda y + b_\lambda)$ by choosing suitable scaling and centring constants $a_\lambda > 0$ and b_λ, and in this case either G is a Gaussian distribution or G has a finite lower end-point $y_0 = \inf\{G > 0\}$ and $G(y - y_0)$ is a gamma distribution. Similarly, if λ_∞ is finite and does not belong to Λ then G is a Gaussian distribution or G has a finite upper end-point y_∞ and $1 - G(y_\infty - y)$ is a gamma distribution. The situation for sequences $\lambda_n \uparrow \lambda_\infty$ is entirely different: any distribution function may occur as the weak limit of a sequence $F_{\lambda_n}(a_n x + b_n)$.

Keywords: affine transformation; asymptotic normality; convergence of types; cumulant generating function; exponential family; Esscher transform; gamma distribution; Gaussian tail; limit law; normal distribution; moment generating function; power norming; semistable; stochastically compact; universal distributions

1. Introduction

Suppose X is a real random variable with distribution function (df) F. Let

$$\Lambda = \{\lambda \in \mathbb{R} | K(\lambda) := \mathrm{E}e^{\lambda X} < \infty\} \tag{1.1}$$

be the set where the moment generating function (mgf) $K(\lambda)$ of X is finite. The set Λ is a connected subset of $\mathbb{R}$ which contains the origin and on which the mgf $\lambda \mapsto K(\lambda)$ is continuous and strictly positive. Associated with F is the natural exponential family $\{F_\lambda, \lambda \in \Lambda\}$ where

$$\mathrm{d}F_\lambda(x) := \mathrm{e}^{\lambda x}\mathrm{d}F(x)/K(\lambda), \qquad \lambda \in \Lambda. \tag{1.2}$$

For convenience, we let X_λ be a random variable with distribution F_λ. We study the asymptotic behaviour of the dfs F_λ for $\lambda \to \lambda_\infty := \sup \Lambda$. Note that $\lambda_\infty \geqslant 0$. We assume $\lambda_\infty > 0$.

If $\lambda_\infty \in \Lambda$ then $F_\lambda \downarrow F_{\lambda_\infty}$ pointwise; see Corollary 2.2. If $\lambda_\infty \notin \Lambda$, then $F_\lambda \downarrow 1_{[x_\infty, \infty)}$, where $x_\infty = \sup\{F < 1\}$ is the *upper end-point* of the df F; see Proposition 2.3. In the latter case, the types in the exponential family $\{F_\lambda, \lambda \in \Lambda\}$, may have a *limit law* for $\lambda \to \lambda_\infty$. This means that it may sometimes be possible to normalize the variables X_λ of the exponential family by translation and positive scaling so that for some non-constant random variable Y,

$$A_\lambda X_\lambda := \frac{X_\lambda - b_\lambda}{a_\lambda} \xrightarrow{\mathrm{d}} Y, \qquad \lambda \to \lambda_\infty. \tag{1.3}$$

Here $\xrightarrow{\mathrm{d}}$ denotes convergence in distribution.

This paper determines the possible non-degenerate limit laws in (1.3). Our main result, Theorem 3.5, states that if there is a non-constant limit variable Y in (1.3), then one can choose the centring constants b_λ and scaling constants a_λ so that Y is a standard normal variable, or so that Y or $-Y$ has a gamma distribution. In a subsequent publication, we shall describe the domains of attraction of the limit laws.

This paper is partially motivated by Balkema *et al*. (1993), where it was found that asymptotic normality of F_λ has useful implications for the study of sums of independent random variables (rvs). A class of thin-tailed densities was identified which is closed under convolution. This closure property is dependent on the fact that each density of the family has an associated exponential family which is asymptotically normal. Rootzen (1987) and Davis and Resnick (1991) use related ideas for applications to extremes of moving averages. Feigin and Yashchin (1983) and Balkema *et al*. (1995) give Tauberian results based on the asymptotic normality of exponential families. If asymptotic normality of exponential families was useful for such things as convolution closure problems and Tauberian theory, we wondered what other weak limits could arise when converging to the boundary of Λ and what applications were possible when convergence was to a non-normal weak limit. The present paper is a first step in the exploration of applications of non-normal limits.

The importance of exponential families in statistics and for asymptotics in probability theory can hardly be overestimated. In analysis exponential families occur as Esscher transforms and are used in Laplace's principle and for saddlepoint approximations. Surveys of their use in statistics are given by Barndorff-Nielsen (1978), Barndorff-Nielsen and Cox (1994) and Brown (1986). For connections with saddlepoint approximations, see Barndorff-Nielsen and Klüppelberg (1999) and Jensen (1995). The limit behaviour of F_λ is of mathematical interest and, moreover, the exponential family offers an effective way to investigate the asymptotic behaviour of the mgf K and the cumulant generating function (cgf) $\kappa = \log K$.

Convergence in (1.3) depends on the behaviour of the cgf κ at λ_∞. The behaviour of the analytic function κ at a fixed point $\lambda_0 < \lambda_\infty$ is well known:

$$n\kappa\left(\lambda_0 + \frac{t}{\sqrt{n}}\right) - n\kappa(\lambda_0) - \mu t \sqrt{n} \to \frac{\sigma^2 t^2}{2}, \qquad n \to \infty. \tag{1.4}$$

Here $\mu = \mu_{\lambda_0} = \kappa'(\lambda_0)$ is the expectation of X_{λ_0} and $\sigma^2 = \sigma^2_{\lambda_0} = \kappa''(\lambda_0)$ the variance.

Relation (1.4) is the formula for the second derivative of κ at λ_0. It is also the central limit theorem for sums of independent observations from the df F_{λ_0} since $n(\kappa(\lambda_0 + \xi) - \kappa(\lambda_0))$ is the cgf of the df $F_{\lambda_0}^{*n}$. Teicher (1984) has investigated relation (1.4) for a sequence $\lambda_n \to \infty$, extending work of Feller (1969) on large deviations. More recently Broniatowski and Mason (1994) have looked at very large deviations. There the behaviour of the mgf for $\lambda \to \lambda_\infty$ plays a decisive role.

To understand the behaviour of the cgf κ for $\lambda \to \lambda_\infty$, assume existence of the following limit:

$$\kappa_\lambda^*(t) := \kappa(\lambda + t/\sigma) - \kappa(\lambda) - \mu t/\sigma \to \eta(t), \qquad \lambda \to \lambda_\infty. \tag{1.5}$$

The function κ_λ^* in (1.5) is the cgf of the standardized variable $X_\lambda^* = (X_\lambda - \mu_\lambda)/\sigma_\lambda$, where $\mu_\lambda = \kappa'(\lambda)$ is the expectation and σ_λ^2 the variance of X_λ; see Feigin and Yashchin (1983). It also describes the convex function κ around the point λ normalized so as to have a horizontal tangent at $t = 0$ and curvature 1. It is not surprising that the parabola $\eta(t) = t^2/2$ occurs as a limit – corresponding to the normal law for the limit variable Y in (1.3). The second limit function, the logarithm, corresponds to two families of gamma distributions. In Theorem 3.6 we prove that weak convergence (1.3) entails convergence of the cgfs. Hence we may use the first two moments of X_λ to normalize, thus obtaining the limit relation (1.5).

Statistical applications have motivated interest in exponential families closed under certain transformation groups. Lehmann (1983) mentions exponential location families. Casalis (1991) classifies natural exponential families on $\mathbb{R}^d$ which are invariant under certain groups of affine transformations, and Bar-Lev and Casalis (1994; 1998) describe exponential families G_γ, $\gamma \in \Gamma$, on $\mathbb{R}$ which are invariant under certain groups of affine transformations A^t, $t \in \mathbb{R}$. For each $t \in \mathbb{R}$ there exists $\gamma \in \Gamma$ so that $G_\gamma(x) = G(A^{-t}(x))$.

The paper is organized as follows. In Section 2 we first prove certain continuity results. From these we derive a stability property for the limit variable Y which allows us to obtain in Section 3 the possible limit distributions, the normal and gamma distributions. Section 4 comments briefly on limit relation (1.3) when convergence is only along sequences $\lambda_n \uparrow \lambda_\infty$, which makes the situation complex since then the cgfs need not converge. Example 4.6 shows that the Cauchy distribution may occur as weak limit and Theorem 4.8 shows that the behaviour of the convex function κ may be quite bizarre.

In a later paper we shall describe domains of attraction and give an application to saddlepoint approximations.

This paper treats the asymptotic behaviour of the exponential family in the neighbourhood of the upper end-point of Λ. The transformation $X' = -X$ allows us to translate these results into statements about the asymptotic behaviour in the neighbourhood of the lower end-point, $\inf \Lambda$. If Y is a limit variable for the exponential family generated by X in the upper end-point then $-Y$ is a limit variable for the exponential family generated by $-X$ in the lower end-point.

Obviously the multivariate case is the really interesting situation. The setting there is simple: the cgf of a random vector is a convex function defined (finite) on a convex subset $\Lambda \subset \mathbb{R}^d$. For simplicity assume Λ is open. The cgf is analytic. What is its behaviour as one approaches the boundary? Normalize the cgf for $\lambda_0 \in \Lambda$ so that the tangent hyperplane in λ_0 is horizontal and the second derivative is the standard inner product. The associated random vector X_λ^* has zero expectation vector and the identity matrix as covariance. What happens to the distribution of X_λ^* as λ approaches a point on the boundary of Λ or tends to infinity? Do there exist non-degenerate limit laws? Do the mgfs of X_λ^* converge? Is it possible that X_λ^* does not converge in distribution but that X_λ does converge for some other normalization?

This paper will answer some of these questions in the univariate case.

2. Stability of the limit laws

Random variables arising from a limit procedure frequently satisfy a stability condition. For the df G of the limit Y of the exponential family in (1.3), the stability relation takes on the form

$$G_\gamma(x) = G(ax + b), \qquad a > 0,\ b \in \mathbb{R}. \tag{2.1}$$

Indeed, G satisfies a large number of such relations. The random variables Y_γ in the exponential family of the limit variable Y all are of the same type! The essential step in establishing this stability for the limit variable is Proposition 2.12.

We start by studying the behaviour of $\{F_\lambda\}$ as $\lambda \uparrow \lambda_\infty = \sup \Lambda$ without using any normalization. We then consider the following question. Suppose a sequence of dfs F_n converges weakly to a non-degenerate df F. Let $G_n = (F_n)_{\lambda_n}$ be a df in the exponential family of F_n and suppose $G_n \to G$ weakly. What is the relation between the limit distributions F and G? We answer this question in Theorem 2.7. In the second part of this section we consider weak limit behaviour under positive affine transformations and consider $F_\lambda(a_\lambda x + b_\lambda)$ for $\lambda \to \lambda_\infty$. The norming constants $a_\lambda > 0$ and b_λ may be chosen to vary continuously with λ. The limit distribution will depend on the normalization. By Khinchine's convergence of types theorem different non-degenerate limit distributions will belong to the same type.

Proposition 2.1. *For any fixed x for which* $0 < F(x) < 1$ *the function* $\lambda \mapsto F_\lambda(x)$ *is strictly decreasing on* Λ *and continuous. It is also true that* $\lambda \mapsto F_\lambda(x-)$ *is strictly decreasing and continuous on* Λ.

Proof. For monotonicity see Brown (1986, Corollary 2.22). For continuity, if $\lambda \to \lambda_0$, convergence of

$$\int 1_{(-\infty,x]} e^{\lambda u} dF(u) = K(\lambda)F_\lambda(x) \to K(\lambda_0)F_{\lambda_0}(x)$$

follows by dominated convergence with dominating function $e^{\alpha u} \vee e^{\beta u}$, with $\alpha, \beta \in \Lambda$. The continuity of the mgf K on Λ gives $F_\lambda(x) \to F_{\lambda_0}(x)$. □

Corollary 2.2. *If* $\lambda_\infty = \sup\Lambda \in \Lambda$, *then* $F_\lambda \downarrow F_{\lambda_\infty}$ for $\lambda \uparrow \lambda_\infty$.

The interesting case is when the upper end-point λ_∞ does not lie in Λ.

Proposition 2.3. *Suppose $\lambda_\infty \notin \Lambda$. Let $x_\infty = \sup\{F < 1\} \leqslant \infty$ denote the upper end-point of the df F. Then $F_\lambda \downarrow 1_{[x_\infty, \infty)}$.*

Proof. If λ_∞ is finite, then $x_\infty = \infty$ and $K(\lambda) \to \infty$ for $\lambda \to \lambda_\infty$. (Otherwise $K(\lambda_\infty - 0) < \infty$ and $\lambda_\infty \in \Lambda$ by Fatou's lemma.) Hence for any $x_1 > 0$,

$$F_\lambda(x_1) = \int_{-\infty}^{x_1} \mathrm{e}^{\lambda x} \mathrm{d}F(x)/K(\lambda) \leqslant \mathrm{e}^{\lambda_\infty x_1}/K(\lambda) \to 0, \qquad \lambda \to \lambda_\infty.$$

Now assume $\lambda_\infty = \infty$. If F is degenerate the statement is obvious. Otherwise choose $x_1 < x_2 < x_\infty$ so that $F(x_1) > 0$. Then $1 - F(x_2) = p > 0$, and

$$\frac{1 - F_\lambda(x_2)}{F_\lambda(x_1)} = \frac{\displaystyle\int_{(x_2,\infty)} \mathrm{e}^{\lambda x} \mathrm{d}F(x)}{\displaystyle\int_{(-\infty,x_1]} \mathrm{e}^{\lambda x} \mathrm{d}F(x)} \geqslant \frac{\mathrm{e}^{\lambda x_2} p}{\mathrm{e}^{\lambda x_1}} \to \infty, \qquad \lambda \to \infty.$$

Since $1 - F_\lambda(x_2) \leqslant 1$, we have $F_\lambda(x_1) \to 0$. □

For convenience, we associate to each $\lambda \in \Lambda$ an rv X_λ with df F_λ. We shall write $X_\lambda =: E_\lambda X$ where E_λ denotes the *Esscher operator*. The Esscher operators E_λ satisfy the additive law

$$E_\mu E_\lambda = E_{\lambda+\mu}, \qquad \lambda, \lambda + \mu \in \Lambda.$$

Now suppose X_n are rvs and $Y_n = E_{\gamma_n} X_n$ for some sequence γ_n. Let X_n converge to X_0 in distribution and Y_n to Y_0. Does it follow that $\gamma_n \to \gamma_0$ and $Y \stackrel{\mathrm{d}}{=} E_\gamma X_0$?

Proposition 2.4. *Suppose $X_n \stackrel{\mathrm{d}}{\to} X_0$ and $\gamma_n \to \gamma_0$. Let X_n have mgf K_n for $n \geqslant 0$. Assume that $K_n(\gamma_n)$ is finite for $n \geqslant 1$ and write $Y_n = E_{\gamma_n} X_n$.*

(a) *If $K_n(\gamma_n) \to K_0(\gamma_0) < \infty$, then $Y_n \stackrel{\mathrm{d}}{\to} Y_0$.*
(b) *If $Y_n \stackrel{\mathrm{d}}{\to} Y$ for some rv Y, then $Y = E_{\gamma_0} X_0$ and $K_n(\gamma_n) \to K_0(\gamma_0) < \infty$.*

Proof. Let $a_n = K_n(\gamma_n)$ and let π_n be the distribution of X_n for $n \geqslant 0$ and ρ_n that of Y_n. Then $\mathrm{d}\mu_n(x) = \mathrm{e}^{\gamma_n x} \,\mathrm{d}\pi_n(x) = a_n \,\mathrm{d}\rho_n(x)$. Convergence of $\int \varphi \,\mathrm{d}\mu_n \to \int \varphi \,\mathrm{d}\mu_0$ holds for continuous functions φ with compact support. This means that $\mu_n \to \mu_0$ vaguely. To prove (a), note that if $a_n \to a_0 < \infty$, then $\mu_n \to \mu_0$ weakly and hence $\rho_n \to \mu_0/a_0$ weakly.

For (b), suppose $a_{k_n} \to a \in [0, \infty]$. Then $\mathrm{e}^{\gamma_0 x} \,\mathrm{d}\pi_0(x) = a \,\mathrm{d}\rho(x)$. It follows that a is finite and positive, and that $a_n \to a = \int \mathrm{e}^{\gamma_0 x} \,\mathrm{d}\pi(x)$. □

Example 2.5. *The sequence $K_n(\gamma_n)$ in Proposition 2.4 may converge to a finite limit $a \neq K_0(\gamma_0)$.*

Take $\gamma_n = 1$ for all n and let ρ_n have mass $\frac{1}{2}$ in the two points 0 and $x_n = n$. Then $X_n \stackrel{\mathrm{d}}{\to} X_0 \equiv 0$ and $a_n = 1/(\frac{1}{2} + e^{-n}/2) \to a = 2 > 1 = a_0$.

Example 2.6. *It may happen that* $X_n \xrightarrow{d} X$, $Y_n = E_{\gamma_n} X_n \xrightarrow{d} Y$ *and* $\gamma_n \to \infty$.

Let μ be a finite measure which charges both $(-\infty, 0)$ and $(0, \infty)$. Let π_n be the probability measure $c_n(1 \wedge \mathrm{e}^{-nx})\,\mathrm{d}\mu(x)$ for $n \geqslant 0$. Take $\gamma_n = n$. The rv $Y_n = E_{\gamma_n} X_n$ has distribution $\mathrm{d}\rho_n(x) = b_n(\mathrm{e}^{nx} \wedge 1)\,\mathrm{d}\mu(x)$. It is clear that X_n converges in distribution to an rv X with probability distribution $\mathrm{d}\pi = c1_{(-\infty,0]}\,\mathrm{d}\mu$ and Y_n to an rv Y with distribution $\mathrm{d}\rho = b1_{[0,\infty)}\,\mathrm{d}\mu$.

We can now prove a kind of convergence of types theorem where 'type' has to be interpreted as belonging to the same exponential family.

Theorem 2.7. *Let* $Y_n = E_{\gamma_n} X_n$ *for* $n \geqslant 1$ *and* $a_n = \mathrm{E}\mathrm{e}^{\gamma_n X_n}$. *Suppose* $X_n \xrightarrow{d} X$ *with* X *non-constant, and* $Y_n \xrightarrow{d} Y$.

If (γ_n) *is bounded, then* $\gamma_n \to \gamma$, $a_n \to a = \mathrm{E}\mathrm{e}^{\gamma X} < \infty$ *and* $Y = E_\gamma X$.

If $\sup \gamma_n = \infty$ *then* $\gamma_n \to \infty$ *and there exists a point* $c \in \mathbb{R}$ *such that* $X \leqslant c \leqslant Y$ a.s.

If $\inf \gamma_n = -\infty$ *then* $\gamma_n \to -\infty$ *and there exists a point* c *such that* $Y \leqslant c \leqslant X$ a.s.

Proof. First consider the case $\gamma_n \to \infty$. Suppose the distributions overlap: there exist $a < b$ so that $P\{X > b\} > 0$ and $P\{Y < a\} > 0$. Let δ denote the minimum of these two positive numbers. Then $P\{X_n > b\}$ and $P\{Y_n < a\}$ eventually exceed $\delta/2$. Thus eventually

$$\frac{\delta/2}{1-\delta/2} \leqslant \frac{P\{Y_n < a\}}{P\{Y_n > b\}} \leqslant \frac{\mathrm{e}^{\gamma_n a} P\{X_n < a\}}{\mathrm{e}^{\gamma_n b} P\{X_n > b\}} \leqslant \mathrm{e}^{-\gamma_n(b-a)} \frac{1-\delta/2}{\delta/2}.$$

This contradicts the assumption that $\gamma_n \to \infty$.

The case $\gamma_n \to -\infty$ is treated in the same way.

There are three mutually exclusive alternatives: either (i) $X \leqslant c \leqslant Y$, or (ii) $Y \leqslant c \leqslant X$, or (iii) neither (i) nor (ii) holds. Hence the sequence γ_n is bounded, or it diverges to $+\infty$ or it diverges to $-\infty$. If (γ_n) is bounded, then by Proposition 2.4, γ_n converges to some value γ since the Esscher transforms $E_\alpha X$ and $E_\beta X$ are different for $\alpha \neq \beta$ if X is not constant. □

Now return to the exponential family $\{X_\lambda, \lambda \in \Lambda\}$ and assume that $\lambda_\infty \notin \Lambda$. To obtain a non-degenerate limit distribution for the variables X_λ in the case $\lambda_\infty \notin \Lambda$, we have to normalize these variables, so assume (1.3) holds, $(X_\lambda - b_\lambda)/a_\lambda \xrightarrow{d} Y$ for some non-constant random variable Y. By Proposition 2.1 the family $\{F_\lambda\}$ of X_λ is weakly continuous in λ. This makes it possible to choose the coefficients $a_\lambda > 0$ and $b_\lambda \in \mathbb{R}$ to be continuous on Λ.

Lemma 2.8. *The constants* a_λ *and* b_λ *in (1.3) can be chosen to be continuous functions of* λ *on the set* Λ.

Proof. Write $Y = \psi(U)$ with ψ increasing and U uniform (0, 1). One may take for ψ the left-continuous inverse of the df of Y. Choose $p \in (0, \frac{1}{2})$ so small that $\psi(p) < \psi(1-p)$. Set

$$b := \int_{p/2}^{1-p/2} \psi(u)\,\mathrm{d}u, \qquad a := \int_{1-p}^{1-p/2} \psi(u)\,\mathrm{d}u - \int_{p/2}^{p} \psi(u)\,\mathrm{d}u.$$

Let Y' denote the normalized variable $(Y - b)/a$. Similarly, write $X_\lambda = \varphi_\lambda(U)$ and define the smoothed median b_λ and smoothed range a_λ as above with φ_λ replacing ψ. Then $a_\lambda > 0$ eventually and convergence $A_\lambda X_\lambda \xrightarrow{d} Y$ for some family of normalizations A_λ implies convergence $(X_\lambda - b_\lambda)/a_\lambda \xrightarrow{d} Y'$. (The norming constants a and b depend continuously on the increasing function ψ and hence on the df.) Weak continuity of $\lambda \mapsto \varphi_\lambda$ is equivalent to weak continuity of the exponential family F_λ and implies continuity of the norming constants a_λ and b_λ. □

We will need the fact that Esscher operators react in a simple way with scaling and translation:

$$E_{\lambda/a}(aX + b) \stackrel{d}{=} aE_\lambda X + b \tag{2.2}$$

for $\lambda \in \Lambda$, $a > 0$ and $b \in \mathbb{R}$. This follows since both sides of (2.2) have the same mgf $z \mapsto e^{bz} K(az + \lambda)/K(\lambda)$.

We now discuss the stability property of the limit variable Y in (1.3). Let $M(\gamma) = Ee^{\gamma Y}$ be the mgf of Y, and $\{Y_\gamma, \gamma \in \Gamma\}$ the associated exponential family with $\Gamma = \{M < \infty\}$. We shall see below that there exist many pairs (γ, C) with $\gamma \in \Gamma$ and C in the group $\mathscr{G}$ of positive affine transformations $x \mapsto C(x) = (x - b)/a$ with $a > 0$ and $b \in \mathbb{R}$ which satisfy the stability relation

$$E_\gamma Y \stackrel{d}{=} CY. \tag{2.3}$$

Example 2.9. *The extended gamma family.*

The following variables satisfy (2.3) for all γ for which the mgf of Y is finite:

(a) If Y is distributed as $N(\mu, \sigma^2)$ then $Y_\gamma \stackrel{d}{=} Y + \sigma^2\gamma$ for $\gamma \in \Gamma = \mathbb{R}$.

(b) The standard exponential rv satisfies the relation $Y_\gamma \stackrel{d}{=} Y/(1-\gamma)$ for $\gamma < 1$. Similarly, $Z = -Y$ satisfies the relation $Z_\gamma \stackrel{d}{=} Z/(1+\gamma)$ for $\gamma > -1$.

(c) More generally, if Y (or $-Z$) has a gamma density $x^{s-1}e^{-x}/\Gamma(s)$ on $(0, \infty)$ then $Y_\gamma \stackrel{d}{=} Y/(1-\gamma)$ for $\gamma < 1$ (and $Z_\gamma \stackrel{d}{=} Z/(1+\gamma)$ for $\gamma > -1$).

These rvs generate exponential families whose dfs are all of the same type.

Since the gamma distribution with shape parameter s is asymptotically normal for $c = 1/s \to 0$ we have a continuous three-parameter family of dfs $H_c(ax + b)$, $a > 0$, b and c real. Here H_0 is the standard normal df, $H_{-c}(x) = 1 - H_c(-x)$ for $c > 0$, and H_c is the df of the normalized gamma variable $V_c = (Y - s)/\sqrt{s}$, with $c = 1/s$, where Y has density $x^{s-1}e^{-x}/\Gamma(s)$ on $(0, \infty)$.

Our main result states that this three-parameter *extended gamma family* is the set of limit laws for exponential families, both for $\lambda \to \sup\Lambda$ and for $\lambda \to \inf\Lambda$.

Note the resemblance to extreme value limit theory where there also is a continuous three-parameter family of limit distributions; see de Haan (1970, p. 104). This resemblance is not due to some innate relation between extremes and exponential families, but results from the structure of the group $\mathscr{G}$ of positive affine transformations on $\mathbb{R}$. The group $\mathscr{G}$

has two kinds of elements: translations, and multiplications with a given centre. The normal distributions are stable for translations; the gamma distributions with a given endpoint are stable for multiplications having the end-point as centre. The extended gamma family reflects this structure.

We now want to show that the limit variable Y in (1.3) has to satisfy a number of stability relations of the form (2.3).

With the positive affine transformation A in $\mathscr{G}$, given by $Ax = (x-b)/a$, we associate the point $(\log a, b)$ in the plane. It is then natural to set

$$\|A\| := \|(\log a, b)\|_2 = \sqrt{(\log a)^2 + b^2}. \tag{2.4}$$

The function $\|\cdot\|$ is not a norm on the group $\mathscr{G}$, in particular $\|A^{-1}\| \neq \|A\|$, but it does describe the topology of $\mathscr{G}$ adequately for our purpose.

Proposition 2.10. *Let U_γ, $\gamma \in \Gamma$, be the exponential family generated by the non-constant rv U. Suppose $C_n \in \mathscr{G}$, $\gamma_n \in \Gamma$, $C_n U_{\gamma_n} \stackrel{d}{\to} Z$ with Z non-constant, $\gamma_n \to \gamma > 0$. Then*

$$\|C_n\| \to \infty \text{ if and only if } \gamma \notin \Gamma.$$

Proof. Suppose $\gamma \in \Gamma$. Then $E_{\gamma_n} U \stackrel{d}{\to} E_\gamma U$ by Proposition 2.1. The convergence of types theorem implies that $Z = CU_\gamma$ and $C_n \to C$ with $\|C\| < \infty$. For the converse, assume that (C_n) contains a convergent subsequence, say $C_n \to C$ as $n \to \infty$. Then $E_{\gamma_n} U \stackrel{d}{\to} C^{-1} Z$. Proposition 2.4 implies $\gamma \in \Gamma$. □

In order to characterize the possible limit distributions in (1.3), we need equation (2.3) to hold for a large collection of γ-values.

Lemma 2.11. *Suppose (1.3) holds. For any $r > 0$ and $\mu \in [0, \lambda_\infty)$ there exists $\lambda \in (\mu, \lambda_\infty)$ such that*

$$\|A_\mu A_\lambda^{-1}\| \vee (\lambda - \mu) a_\mu = r. \tag{2.5}$$

Proof. Write $Y^\lambda = A_\lambda X_\lambda$. (The upper index notation is used here to avoid confusion with the exponential family generated by the variable Y.) Fix $\mu \in \Lambda$. Use (2.2) and write

$$A_\mu A_\lambda^{-1} Y^\lambda \stackrel{d}{=} E_{(\lambda-\mu)a_\mu} Y^\mu, \qquad \lambda \in \Lambda. \tag{2.6}$$

By assumption $Y^\lambda \stackrel{d}{\to} Y$ for $\lambda \to \lambda_\infty$. Apply Proposition 2.10 with $E_{(\lambda-\mu)a_\mu} Y^\mu$ in the role of U_{γ_n} and $(A_\mu A_\lambda^{-1})^{-1}$ in the role of C_n to conclude that $\|(A_\mu A_\lambda^{-1})^{-1}\| \to \infty$ as $\lambda \to \lambda_\infty$. Check that $\|C_n\| \to \infty$ if $\|C_n^{-1}\| \to \infty$. By Lemma 2.8 the quantity $\|A_\mu A_\lambda^{-1}\|$ varies continuously from 0 to ∞ as λ increases from μ to λ_∞. So the leftmost term in (2.5) will equal r before λ reaches the value λ_∞. □

Fix $r > 0$. Let $\mu_n \to \lambda_\infty$ and choose $\lambda_n > \mu_n$ as in Lemma 2.11. Choose a subsequence $k_1 < k_2 < \cdots$ so that

$$C_{k_n} := A_{\mu_{k_n}} A_{\lambda_{k_n}}^{-1} \to C \in \mathscr{G}, \qquad \gamma_{k_n} := (\lambda_{k_n} - \mu_{k_n}) a_{\mu_{k_n}} \to \gamma.$$

This is possible since $\|C_n\|$ and γ_n are bounded by r. Then $C_{k_n} Y^{\lambda_{k_n}} \stackrel{\mathrm{d}}{=} E_{\gamma_{k_n}} Y^{\mu_{k_n}}$ by (2.6). Theorem 2.7 gives $CY \stackrel{\mathrm{d}}{=} Y_\gamma$ with $\|C\| \vee \gamma = r$ by continuity. This establishes the next result.

Proposition 2.12. *If (1.3) holds and Y is non-degenerate then for each* $r > 0$ *there exists a constant* $\gamma > 0$ *and a positive affine transformation C with* $\|C\| \vee \gamma = r$ *such that (2.3) holds:* $E_\gamma Y \stackrel{\mathrm{d}}{=} CY$.

The question whether *all* distributions in the exponential family of the limit distribution are of the same type will be settled by algebraic arguments in the next section.

3. Solutions of the stability equation

The stability equation (2.3), $E_\gamma Y \stackrel{\mathrm{d}}{=} CY$, allows us to determine the possible limit laws for the exponential family X_λ for $\lambda \to \lambda_\infty$.

For statistical applications it is of importance to characterize exponential families which are invariant under a given group $\mathscr{H}$ of transformations. Lehmann (1983, p. 35) observes that the normal distributions with fixed variance form the only natural exponential family which also is a location family. For natural exponential families Casalis (1991), in a very readable paper, has solved the characterization problem when $\mathscr{H}$ is a group of translations on $\mathbb{R}^d$ and for some other classical groups of affine transformations on $\mathbb{R}^d$. Bar-Lev and Casalis (1994; 1998) solve the problem for the case when $\mathscr{H}$ is a subgroup of the group of affine transformations on $\mathbb{R}$. We are grateful to a referee of a previous version of this paper for pointing out these two references. The second paper contains full proofs and hence we restrict ourselves here to a short exposition of the results of this paper which are relevant to us.

If a natural exponential family Y_ξ, $\xi \in \Gamma$, is invariant under a group $\mathscr{H}$ of positive affine transformations, and the dfs are non-degenerate, then $\mathscr{H}$ is a closed commutative subgroup of $\mathscr{G}$. If $\mathscr{H}$ is the group of translations then Y has a Gaussian distribution. If $\mathscr{H}$ is the group of all multiplications with centre c then there exists a constant $d \neq 0$ such that $(Y - c)/d$ has a gamma distribution on $(0, \infty)$.

In the present paper we are concerned with the more elementary question of describing all dfs G which satisfy one or more stability relations of the form (2.1).

Example 3.1. There exist rvs V which satisfy the relation $V_\gamma \stackrel{\mathrm{d}}{=} V + \beta$ only if γ and β are integers.

To see this, let V be the random integer with distribution

$$P\{V = k\} = p_k = \mathrm{e}^{-k^2/2}/c, \qquad k \in \mathbb{Z}, \tag{3.1}$$

with c a norming constant. The rv V_ξ has distribution $P\{V_\xi = k\} = \mathrm{e}^{-(k-\xi)^2/2}/C(\xi)$. If γ is an integer then $V_\gamma \stackrel{\mathrm{d}}{=} V + \gamma$, but if γ is not an integer then V and V_γ are not of the same type.

Let $\mathscr{F}(\gamma, C)$ denote the set of all dfs G which satisfy the relation $G(ax+b) = G_\gamma(x)$ for $Cx = (x-a)/b$. Let Y_ξ, $\xi \in \Gamma$, be the natural exponential family generated by the rv Y with df G. Given the df G, what one can say about the set $\mathscr{G}(G)$ of all $C \in \mathscr{G}$ for which there exists $\gamma \in \Gamma$ so that $G \in \mathscr{F}(\gamma, C)$?

Proposition 3.2. *Suppose Y has df G. Any positive affine map $A \in \mathscr{G}(G)$ determines a bijection $\xi \to \xi_* = A_*(\xi) = \xi a + \alpha$ on Γ by $AY_\xi \stackrel{d}{=} Y_{\xi_*}$.*

Proof. Let $\xi \in \Gamma$. Then (2.2) determines an element $\xi_* = A_* \xi$ in Γ by

$$AE_\xi Y \stackrel{d}{=} E_{a\xi} AY = E_{a\xi} E_\alpha Y = E_{a\xi+\alpha} Y = Y_{\xi_*}, \qquad \xi \in \Gamma. \tag{3.2}$$

In the same way the inverse relation $A^{-1} Y_\alpha \stackrel{d}{=} Y$ defines a map A_*^{-1} on Γ. Apply A^{-1} to (3.2) to see that $(A^{-1})_* = (A_*)^{-1}$. Hence A_* is a bijection on Γ. □

Corollary 3.3. *The set $\mathscr{G}(G)$ is a group and the map $C \mapsto C_*$ is a homomorphism of $\mathscr{G}(G)$ into the group of positive affine bijections on Γ.*

Proof. Apply B to (3.2) to see that $(BA)_* = B_* A_*$. □

The exponential family generated by G is invariant under the group $\mathscr{H} = \mathscr{G}(G)$. Bar-Lev and Casalis show that the group $\mathscr{H}$ is a discrete group A^k, $k \in \mathbb{Z}$, or the continuous group A^t of all translations or all multiplications with a common centre c.

Proposition 3.4. *Let $\mathscr{G}(G)$ be a continuous one-parameter group A^t, $t \in \mathbb{R}$. For any $\xi \in \Gamma$ define $\xi(t) \in \Gamma$ by $A^t Y_\xi \stackrel{d}{=} Y_{\xi(t)}$. The map $t \mapsto \xi(t)$ is a homomorphism from $\mathbb{R}$ to Γ.*

Proof. Observe that $\xi(t) = (A_*)^t \xi$ for all $t \in \mathbb{R}$ by a continuity argument. So the map $t \to \xi(t)$ is continuous and strictly monotone. It is onto since $\xi(n) \to \gamma \in \Gamma$ would give the sequence $(A^n Y_\xi)_n$ a non-degenerate limit Y_γ. So $\xi(t)$ tends to an end-point of Γ for $t \to \infty$ and to the other endpoint for $t \to -\infty$ by monotonicity. □

For the normal distribution $\mathscr{G}(G)$ is the group of translations, and for the gamma distribution on $(0, \infty)$ or on $(-\infty, 0)$ it is the group of all multiplications with centre $c = 0$. Hence $\mathscr{G}(G)$ is a continuous one-parameter group for all distributions in the extended gamma family. The formulae below give explicit expressions for $C_*^t \xi$ when $Y_\beta \stackrel{d}{=} CY_\alpha$ with $Cx = (x-b)/a$. Suppose $\alpha < \beta$. A straightforward calculation gives, for $\gamma_0 \in \Gamma$ and any integer t,

$$C^t Y_{\gamma_0} \stackrel{d}{=} Y_{\gamma(t)}, \qquad \gamma(t) = C_*^t \gamma_0 = \begin{cases} \gamma_0 + (\beta - \alpha)t, & \text{if } a = 1, \\ a^t \gamma_0 + \dfrac{a^t - 1}{a - 1}(\beta - a\alpha), & \text{if } a \neq 1. \end{cases} \tag{3.3}$$

If $\mathscr{G}(G)$ is a continuous group then (3.3) holds for all $t \in \mathbb{R}$. In that case $\gamma(t)$ is an increasing bijection from $\mathbb{R}$ to Γ by Proposition 3.4. It satisfies the differential equation

$$\ddot{\gamma} = (\log a)\dot{\gamma}, \qquad \gamma(0) = \gamma_0, \gamma(1) = a(\gamma_0 - \alpha) + \beta.$$

We now return to the basic limit relation (1.3). Let the limit Y have df G. By Proposition 2.12 the set $\mathscr{G}(G)$ is not discrete. Hence $\mathscr{G}(G)$ is a continuous one-parameter subgroup of $\mathscr{G}$. Then from Bar-Lev and Casalis (1994; 1998) we have the following:

Theorem 3.5. *Let F_λ, $\lambda \in \Lambda$, be the exponential family (1.2). If $\lambda_\infty = \sup \Lambda$ does not belong to Λ and if there exist constants $a_\lambda > 0$ and $b_\lambda \in \mathbb{R}$ such that $F_\lambda(a_\lambda x + b_\lambda) \to G(x)$ weakly for some non-degenerate df G, then G belongs to the extended gamma family introduced in Example 2.9.*

The exponential families of gamma distributions are generated by Radon measures with densities $x^s 1_{(0,\infty)}$ on $\mathbb{R}$, with $s > -1$. They converge to the Gaussian exponential family if $s \to \infty$ provided we apply a proper normalization. What happens if $s \to -1$? For $s \leqslant -1$ the measure $x^s 1_{(0,\infty)}(x)\,dx$ is no longer a Radon measure on $\mathbb{R}$. However, one can truncate this measure and ask for the limit behaviour of the exponential family of probability measures with densities $f_\gamma(x) = c(s, \gamma)e^{\gamma x} x^s 1_{[1,\infty)}(x)$, as $\gamma \uparrow \gamma_\infty = 0$. For $s < -1$ the answer is simple: γ_∞ lies in Γ and Corollary 2.2 applies. If $s = -1$ the situation is more delicate. There exists a non-degenerate limit distribution, but only under nonlinear normalization. See Example 4.10.

Weak convergence in (1.3) implies convergence of the mgfs. The significance of this result will become apparent in the next section.

Theorem 3.6. *Let (1.3) hold. Suppose Y is non-constant and $\lambda_\infty \notin \Lambda$. Then the mgfs of the normalized variables $A_\lambda X_\lambda$ converge to the mgf of the limit variable Y on the interval $\Gamma = \{\gamma : Ee^{\gamma Y} < \infty\}$.*

Proof. Let $\gamma \in \Gamma$. There exists a unique positive affine transformation $C = C(\gamma)$ such that $CY \stackrel{d}{=} Y_\gamma$ by Proposition 3.4.

Let $\mu_n \uparrow \lambda_\infty$ and set $Y_n = A_{\mu_n} X_{\mu_n}$. We write $A_\lambda x = (x - b_\lambda)/a_\lambda$ and assume that $a_\lambda > 0$ and b_λ depend continuously on λ. We claim that there is a sequence $\lambda_n \to \lambda_\infty$ such that $\gamma_n = (\lambda_n - \mu_n)a_{\mu_n} \to \gamma$ and $C_n = A_{\mu_n} A_{\lambda_n}^{-1} \to C$.

First assume $\gamma < 0$. Set $r := \sup\{\|C(\xi)\| \,|\gamma \leqslant \xi \leqslant 0\}$. Then $\|A_{\mu_n}\| \to \infty$ by Proposition 2.10 and hence $\|A_{\mu_n} A_0^{-1}\| > r + 1$ and $\mu_n > 0$ for $n \geqslant n_0$. Let $n \geqslant n_0$. Let λ decrease from μ_n to 0. By continuity there is a maximal value λ_n for which $\gamma_n := (\lambda_n - \mu_n)a_{\mu_n} = \gamma$ or $\|C_n\| := \|A_{\mu_n} A_{\lambda_n}^{-1}\| = r + 1$. (In the latter case $\gamma_n \in [\gamma, 0]$.) Note that $\lambda_n \to \lambda_\infty$ since $\lambda_{k_n} \to \lambda \in [(0, \lambda_\infty)$ implies that $\|C_n\| \to \infty$ by Proposition 2.10. Now assume $\gamma_n \to \beta$ and $C_n \to B$. (Take subsequences if need be.) Then $E_{\gamma_n} Y_n \stackrel{d}{=} C_n Y^{\lambda_n} \stackrel{d}{\to} BY$ by (2.6). Hence $BY \stackrel{d}{=} E_\beta Y$ and $Ee^{\gamma_n Y_n} \to Ee^{\beta Y}$ by Proposition 2.4. So $B = C(\beta)$, and $\beta \in [\gamma, 0]$ implies $\|B\| \leqslant r$. Hence eventually $\|C_n\| < r + 1$ which implies $\gamma_n = \gamma$. Thus we see that $\gamma_n \to \gamma$. The proof for $\gamma \geqslant 0$ is similar. □

Corollary 3.7. *If (1.3) holds, convergence to a non-degenerate limit still takes place if F_λ is centred and scaled by expectation and standard deviation.*

4. Sequential limits

In this section we only assume that the limit relation (1.3) holds for a sequence $\lambda_n \uparrow \lambda_\infty$. As above, $\lambda_\infty = \sup\Lambda \notin \Lambda$. We adapt the notation slightly. V is a non-degenerate rv such that

$$V_n := A_n X_{\lambda_n} = (X_{\lambda_n} - b_n)/a_n \xrightarrow{\mathrm{d}} V. \tag{4.1}$$

We treat two questions. First, what information does the sequence (A_n) give about the distribution of the limit variable V? Second, what limit distributions are possible in (4.1)?

Proposition 4.1. *Suppose (4.1) holds. If* $A_{n+1}A_n^{-1} \to \mathrm{id}$ *then (1.3) holds: There exist functions* $a(\lambda) > 0$ *and* $b(\lambda)$ *so that* $(X_\lambda - b(\lambda))/a(\lambda) \xrightarrow{\mathrm{d}} V$.

Proof. Khinchine's convergence of types theorem implies $A_n X_{\lambda_{n+1}} \xrightarrow{\mathrm{d}} V$. Set $A(\lambda) = A_n$ for $\lambda_n \leqslant \lambda < \lambda_{n+1}$. Monotonicity of $\lambda \mapsto F_\lambda(x)$ (see Proposition 2.1) ensures that $A(\lambda)X_\lambda \xrightarrow{\mathrm{d}} V$ for $\lambda \to \lambda_\infty$. □

Now assume $A_{n+1}A_n^{-1} \to C \neq \mathrm{id}$. In the asymptotic theory of sums or maxima this implies that the limit, if it exists, is semi-stable (see Hazod and Scheffler 1993). For exponential families semi-stability means that the limit distribution belongs to $\mathscr{F}(\gamma, C)$ for some $\gamma \neq 0$. However, the situation for exponential families is more complex than for sums or maxima. We shall investigate the behaviour for translations.

Example 4.2. *An integer-valued limit variable.*

Let the random integer X have a log-concave distribution with $P\{X = k\} = p_k = \mathrm{e}^{-\alpha_k} > 0$ for all k. So $p_{k-1}p_{k+1} \leqslant p_k^2$. Assume $p_{k+1}p_{k-1}/p_k^2 \to \mathrm{e}^{-\gamma} \in (0, 1)$ for $k \to \infty$. Let $\rho_k = (\alpha_{k-1} + \alpha_{k+1})/2$. Then $X_{\rho_n} - n \xrightarrow{\mathrm{d}} V$ with $P\{V = k\} = \mathrm{e}^{-\gamma k^2/2}/c(\gamma)$; see (3.1). Now suppose $(\lambda_n - \rho_n)/\gamma \to \beta \in \mathbb{R}$. Then $X_{\lambda_n} - n \xrightarrow{\mathrm{d}} V'$, where V' has distribution $P\{V' = k\} = \mathrm{e}^{-\gamma(k-\beta)^2/2}/c(\beta)$. All limit variables V' belong to $\mathscr{F}(\gamma, C)$ where C is the translation $Cx = x + 1$.

The exponential family F_λ of the rv X in the example above gives rise to a one-parameter exponential family of limit distributions G_λ, $\lambda \in \mathbb{R}$. We are only interested in limit *types*. Since $G_{\lambda+\gamma}(x) = G_\lambda(x-1)$ the limit types in this example form a compact family. Topologically this family is a circle. Let $[F]$ denote the type of the df F. As in the case of semi-stable limit distributions for sums and maxima one may describe the behaviour of the family of types $[F_\lambda]$, $\lambda \in \Lambda$, for $\lambda \to \lambda_\infty$ as a curve which spirals to a limiting circle in the space of distribution types.

For exponential families there is an additional limit family. This limit family has no counterpart in the asymptotic theory of sums or maxima.

Choose the weights p_k above so that $\delta_k = p_{k+1}p_{k-1}/p_k^2 \to 0$. The possible non-constant limit distributions of the sequence $X_{\lambda_n} - n$ are then members of the exponential family of Bernoulli variables, $P\{V_\xi = 1\} = 1/(1 + \mathrm{e}^{-\xi}) = 1 - P\{V_\xi = 0\}$, together with the constant variable $V \equiv 0$. Any two-valued random variable can occur in the limit. There are no other limits. In the sequence $p_k(\lambda) = p_k e^{\lambda k}/K(\lambda)$ the maximum occurs in $k(\lambda) \to \infty$, and since $\delta_k \to 0$ there is at most one other point which makes a non-negligible contribution.

We shall now adapt this example so as to obtain a compact 'circle' of non-degenerate limit types.

Example 4.3. *Here we shall exhibit an rv X such that the set of limit types of the exponential family is a 'circle' consisting of the types of the following rvs: U_γ, $\gamma \in \mathbb{R}$, E, W_γ, $-1 < \gamma < 1$, $-E$. Here E is exponentially distributed, U_γ is the exponential family generated by the uniform* (0, 1) *rv U and W_γ is the exponential family generated by the rv W with Laplace density* $e^{-|x|}/2$.

Let X have density $f = e^{-\varphi}$, where $\varphi : \mathbb{R} \to \mathbb{R}$ is a convex function which is piecewise linear on each interval $[k, k+1]$ with slope α_k such that $\alpha_{k+1} - \alpha_k \to \infty$ for $k \to \infty$. Let $\lambda_n \to \lambda_\infty$ and set $\psi_n(x) = \varphi(x) - \lambda_n x$. There are two cases of interest:

(a) $\psi'_n(k_n + \frac{1}{2}) \to \beta \in \mathbb{R}$ for some integer sequence $k_n \to \infty$. Then $Y_{\lambda_n} - k_n \xrightarrow{d} E_\beta U$, where U is uniformly distributed on (0, 1).

(b) ψ_n is minimal in $k_n \to \infty$ and $\psi'_n(k_n \pm \frac{1}{2}) = \pm\tau_n^\pm$, where $\sigma_n = \tau_n^+ \wedge \tau_n^- \to \infty$ and $\sigma_n/(\tau_n^+ \vee \tau_n^-) \to \beta \in [0, 1]$. Then $\sigma_n(Y_{\lambda_n} - k_n)$ converges to an rv Y with density $(e^{x/\beta^-} 1_{(-\infty,0)} + e^{-x/\beta^+} 1_{(0,\infty)})/c$ with $\beta^+ \vee \beta^- = 1$, $\beta^+ \wedge \beta^- = \beta$ and $c = \beta^+ + \beta^-$. If $\beta = 0$ then Y or $-Y$ is exponential, otherwise Y belongs to the exponential family generated by a Laplace density.

The rv X with density $f = e^{-\varphi}$ in the example above has the following property. There is a continuous family of non-degenerate limit distributions G^θ, $0 \leqslant \theta \leqslant 2\pi$, all of different type, except that $G^0 = G^{2\pi}$; a continuous curve $A : \Lambda \to \mathscr{G}$; and a continuous strictly increasing function $\chi : \Lambda \to \mathbb{R}$ tending to infinity for $\lambda \to \lambda_\infty$ so that for each $\theta \in [0, 2\pi]$

$$A(\lambda_n)X_{\lambda_n} \xrightarrow{d} V^\theta \sim G^\theta$$

whenever $\lambda_n \uparrow \lambda_\infty$ and $e^{i\chi(\lambda_n)} \to e^{i\theta}$.

In particular, the set of dfs F_λ, $\lambda \in \Lambda \cap [0, \infty)$, is stochastically compact: any sequence F_{λ_n} contains a subsequence which may be normed to converge weakly to a non-degenerate limit distribution; see de Haan and Resnick (1984). The family of all possible limit variables $aV^\theta + b$ is closed under the Esscher transform. It contains certain rvs from the extended gamma family but also bounded rvs and unbounded rvs which are not semi-stable.

These two examples give an indication of the behaviour of the sequence X_{λ_n} under the condition that $A_n A_{n+1}^{-1} \to C \neq \text{id}$. However, in order that the limit distribution in (4.1) belongs to the class $\mathscr{F}(\gamma, C)$ it is not necessary that the sequence $A_n A_{n+1}^{-1}$ converge. Large gaps may occur. From Section 2 we know that the limit V belongs to $\mathscr{F}(\gamma, C)$ if there exist integer sequences $q_n \to \infty$ and $k_n > q_n$ so that $A_{q_n} A_{k_n}^{-1} \to C$ and $(\lambda_{k_n} - \lambda_{q_n})a_{q_n} \to \gamma$. We therefore introduce the set $\mathscr{H}_0$ of all $C \in \mathscr{G}$, $C \neq \text{id}$, which are the limit of a sequence $C_n = A_{q_n} A_{k_n}^{-1}$ with $k_n > q_n \to \infty$.

Relation (2.6) gives

$$C_n V_{k_n} \stackrel{d}{=} E_{\gamma_n} V_{q_n}, \tag{4.2}$$

where we write $C_n = A_{q_n} A_{k_n}^{-1}$ as above and $\gamma_n = (\lambda_{k_n} - \lambda_{q_n}) a_{q_n}$. Since we assume that $C_n \to C$, Theorem 2.7 applies: $\gamma_n \to \gamma \in [0, \infty]$. If $\gamma = \infty$ there exists a constant c such that $V \leqslant c \leqslant CV$ and the rv V is bounded. If γ is finite then $CV \stackrel{d}{=} V_\gamma$ and V belongs to the set $\mathscr{F}(\gamma, C)$.

This yields the following dichotomy:

Theorem 4.4. *If $\mathscr{H}_0$ is non-empty then either V is bounded and $\gamma = \infty$ for each $C \in \mathscr{H}_0$, or V is unbounded and γ is finite for each $C \in \mathscr{H}_0$.*

Our next result extends Proposition 4.1 and is a partial converse to Theorem 3.5.

Theorem 4.5. *Suppose (4.1) holds and the sequence $(\|A_n A_{n+1}^{-1}\|)$ is bounded. If the limit variable belongs to the extended gamma family then (1.3) holds.*

Proof. The limit variable V satisfies the stability relations $C^t V \stackrel{d}{=} E_{\gamma(t)} V$, $t \in \mathbb{R}$; see Proposition 3.4. Write $(\lambda_{n+1} - \lambda_n) a_n = \gamma_n = \gamma(t_n)$. The sequence (t_n) is bounded. Equivalent are $t_{k_n} \to t_0$ and $\gamma_{k_n} \to \gamma_0 = \gamma(t_0)$. Indeed,

$$A_{k_n} A_{k_n+1}^{-1} V_{k_n+1} = A_{k_n} E_{\lambda_{k_n+1} - \lambda_{k_n}} X_{\lambda_n} = E_{\gamma(t_{k_n})} V_{k_n} \stackrel{d}{\to} E_{\gamma_0} V = C^{t_0} V.$$

Hence $A_{k_n} A_{k_n+1}^{-1} \to C^{t_0}$ and by Proposition 2.4 $\mathrm{E}e^{\gamma_{k_n} V_{k_n}} \to \mathrm{E}e^{\gamma_0} V$.

Now define

$$A(\lambda) := C^{-s} A_n, \qquad \lambda = \lambda_n + \gamma(s)/a_n, \qquad 0 \leqslant s < t_n.$$

We have to prove that $A(\mu_n) X_{\mu_n} \stackrel{d}{\to} V$ for any sequence $\mu_n \uparrow \lambda_\infty$. It suffices to consider sequences $\mu_n = \lambda_{j_n} + \alpha_n / a_{j_n}$ with $\alpha_n = \gamma(s_n) \to \alpha_0 = \gamma(s_0)$ for $0 \leqslant s_n < t_{j_n}$. Then

$$A(\mu_n) X_{\mu_n} \stackrel{d}{=} C^{-s_n} A_{j_n} E_{\alpha_n / a_{j_n}} X_{\lambda_{j_n}} = C^{-s_n} E_{\alpha_n} V_{j_n}.$$

Now observe $s_n \to s_0$, $\alpha_n \to \alpha_0$. The bound $0 \leqslant \alpha_n \leqslant \gamma_n$ implies $\mathrm{E}e^{\alpha_n V_{jn}} \to \mathrm{E}e^{\alpha_0 V}$ and hence $E_{\alpha_n} V_{j_n} \stackrel{d}{=} E_{\alpha_0} V$ and $A(\mu_n) X_{\mu_n} \stackrel{d}{\to} C^{-s_0} E_{\gamma(s_0)} V \stackrel{d}{=} V$. □

The condition that the limit variable belongs to the extended gamma family is less restrictive than it seems. Since the sequence $A_n A_{n+1}^{-1}$ is bounded, the set $\mathscr{H}_0$ is non-empty. Hence the condition will be satisfied if (a) $\mathscr{H}_0$ contains a sequence $C_n \to \mathrm{id}$, or (b) V is unbounded and $\mathscr{H}_0$ is not contained in a discrete subgroup C^k, $k \in \mathbb{Z}$, of $\mathscr{G}$.

Without conditions on the sequence A_n every limit law is possible in (4.1).

Example 4.6. *A Cauchy-distributed limit variable V is possible in (4.1).*

To exhibit this, we shall construct an rv X with density f so that $X_{\lambda_n} - \lambda_n$ converges to an rv V with density $1/(\pi(1 + x^2))$ for $\lambda_n = n^2$.

Let I_n be the interval $[-\sqrt{n}, \sqrt{n}]$ and define

$$h_n(u) = \frac{e^{u^2/2}}{\pi(1 + u^2)} 1_{I_n}(u), \qquad n \geqslant 1.$$

Now introduce h as maximum of translates of the functions h_n: set $h(n^2 + u) = h_n(u)$ for $n \geqslant 1$ and $u \in I_n$ and set $h(x) = 0$ elsewhere. Similarly, define $h^*(n^2 + u) = \mathrm{e}^n$ for $n \geqslant 1$ and $u \in I_n$ and $h^*(x) = 0$ elsewhere. Then $h \leqslant h^*$. Define $g(x) := h(x)\mathrm{e}^{-x^2/2}$ and $g_\lambda(x) := g(x)\mathrm{e}^{\lambda x}/\mathrm{e}^{\lambda^2/2}$. A simple computation gives

$$g_n^*(u) := g_{\lambda_n}(n^2 + u) = h(n^2 + u)\mathrm{e}^{-u^2/2}.$$

This means that $g_n^*(u) = 1/(\pi(1 + u^2))$ on I_n and $g_n^*(u) = 0$ for $\sqrt{n} < |u| \leqslant n$. The tails of g_n^* are negligible: Lemma 4.7 below implies that

$$g_n^*(u) \leqslant h^*(n^2 + u)\mathrm{e}^{-u^2/2} \leqslant \mathrm{e}^{-u^2/6}, \qquad n \geqslant 4,\ |u| \geqslant n.$$

Hence $\|g_n^*\|_1 \to 1$. Now let the rv X have density $f = g/c$, with $c = \|g\|_1$. Then $X_{\lambda_n} - n^2$ has density g_n^*/c_n for $n \geqslant 1$ where $c_n = \|g_n^*\|_1 \to 1$.

Lemma 4.7. *The function h^* in Example 4.6 satisfies the inequality*

$$h^*(n^2 + u) \leqslant \mathrm{e}^{u^2/3}, \qquad |u| \geqslant n,\ n \geqslant 4.$$

Proof. Introduce the concave piecewise linear function ψ: $[0, \infty) \to [0, \infty)$ with the value n in $n^2 - n$ for $n \geqslant 1$. Then $h^* \leqslant \mathrm{e}^\psi$ and $\psi(m^2 + u) \leqslant u^2/3$ for $|u| \geqslant m$ and $m \geqslant 4$. (The inequality holds in $u = \pm m$ and $\psi'(m^2 + u) = 1/2(m + 1) \leqslant 2u/3$ in $u = m + 0$.) □

Doeblin introduced the concept of universal distributions in his study of the asymptotic behaviour of sums of independently and identically distributed rvs. Let S_n be the sum of the first n terms of a sequence of independent samples from the df F. The distribution F is universal if for each rv V there exists a subsequence $k_1 < k_2 < \cdots$ and a sequence of positive affine normalizations A_n such that $A_n S_{k_n} \xrightarrow{\mathrm{d}} V$. Doeblin (1946) established the existence of universal distributions. See Feller (1966, Section XVII.9) for details. One can introduce a similar concept for exponential families. An exponential family X_λ, $\lambda \in \Lambda$, is *universal* if for each rv V there exists a sequence $\lambda_n \uparrow \lambda_\infty = \sup \Lambda$ and a sequence of positive affine transformations A_n such that $A_n X_{\lambda_n} \xrightarrow{\mathrm{d}} V$. □

Theorem 4.8. *Universal exponential families exist.*

Proof. First note that there exists a sequence of dfs Q_n on $\mathbb{R}$ which is dense in the space of all probability distributions with the topology of weak convergence; see Parthasarathy (1967, Theorem II.6.2). We can choose the dfs Q_n to have a continuous density q_n which is bounded by $\mathrm{e}^{n/2}$ and which vanishes outside the interval $I_n = [-\sqrt{n}, \sqrt{n}]$. The construction of Example 4.6 yields an rv X with density f such that $X_{\lambda_n} - \lambda_n$ has density g_n^*/c_n, where g_n^* agrees with q_n on $[-n, n]$ and $c_n \to 1$ since the function g_n^* is bounded by $\mathrm{e}^{-x^2/6}$ outside the interval $[-n, n]$ for $n \geqslant 4$. As in the example, we take $\lambda_n = n^2$.

Let V be an rv with df Q. There is a sequence $k_n \uparrow \infty$ such that $Q_{k_n} \to Q$ weakly. Then $X_{\mu_n} - \mu_n \xrightarrow{\mathrm{d}} V$ if we choose $\mu_n = k_n^2$. □

Universal exponential families have the property that any df $Q(x)$ is limit of some sequence $F_{\lambda_n}(a_n x + b_n)$ with $\lambda_n \uparrow \lambda_\infty$. With more effort one can show that this is also possible under the additional restriction that the sequences λ_n are asymptotically dense: $\lambda_{n+1} - \lambda_n \to 0$. This result will be published elsewhere.

A further question of interest is whether there exist non-degenerate dfs F such that only degenerate limit distributions are possible in (4.1).

If the df F of X has a jump in its upper end-point x_∞ then $P\{X_\lambda = x_\infty\} \to 1$ and only a degenerate limit is possible in (4.1). Less trivial examples are the following:

Example 4.9. *Let F have density $f(x) = c\,\mathrm{e}^{-x}/x$ on $x \geqslant 1$. Then (4.1) will hold only for constant limit variables V.*

To show this we proceed as follows. $Z_\varepsilon = \varepsilon X_{1-\varepsilon}$ has density $f_\varepsilon(x) = c(\varepsilon)\,\mathrm{e}^{-x}/x$ on $[\varepsilon, \infty)$. If $\varepsilon \downarrow 0$ then $f_\varepsilon(x)/c(\varepsilon) \to \mathrm{e}^{-x}/x$ on $(0, \infty)$ and $c(\varepsilon) \sim 1/\log(1/\varepsilon)$. Hence $F_\varepsilon(\varepsilon^u) \to 1 - u$ for $u \in (0, 1)$ for $\varepsilon \downarrow 0$. Take $0 < u \ll 1$. The half-line $[\varepsilon^u, \infty)$ carries weight $u > 0$, but a large part, $1 - 2u$, of the probability lives on the relatively short interval $[0, \varepsilon^{2u}]$. Since $\varepsilon^{2u} = o(\varepsilon^u)$ there is an atom of weight $\geqslant 1 - 2u$ in the limit. Because $u > 0$ is arbitrary the limit can only be degenerate.

Note that $F_\varepsilon(\varepsilon^u) \to 1 - u$ implies that $\log(X_{1-\varepsilon})/\log(1/\varepsilon) \xrightarrow{\mathrm{d}} U$, where U is uniformly distributed on (0, 1). This means that the exponential family X_λ has a non-degenerate limit under *power norming*. To see this, take $c(\lambda) = -\log(1 - \lambda)$; then

$$X_\lambda^{1/c(\lambda)} \xrightarrow{\mathrm{d}} \mathrm{e}^U, \qquad \lambda \to 1.$$

For recent work on power norming for extremes, see Pancheva (1984) and Ravi (1991). With the arguments of Example 4.9 one may show:

Example 4.10. *Let F have density $f(x) = c\,\mathrm{e}^{-x}(\log x)^\alpha/x$ on* (e, ∞). *For $\alpha \geqslant -1$ the limit relation (4.1) has only constant limits. Let U be uniformly distributed on* [0, 1]. *If $\alpha > -1$ then*

$$\frac{\log X_\lambda}{\log(1/(1-\lambda))} \xrightarrow{\mathrm{d}} U^{1/(1+\alpha)}, \qquad \lambda \to \lambda_\infty = 1.$$

If $\alpha = -1$ *then even power norming yields only constant limit variables but*

$$(\log\log X_\lambda)/\log\log(1/(1-\lambda)) \xrightarrow{\mathrm{d}} U, \qquad \lambda \to 1.$$

Let $[F]$ denote the type of the non-degenerate df F. In this paper we have studied the behaviour of the curve $[F_\lambda]$ in the space of non-degenerate probability types. Introduce $\Xi(F)$ as the set of limits of sequences $[F_{\lambda_n}]$ with $\lambda_n \uparrow \lambda_\infty$. If $[F_\lambda]$ converges to a point $[G]$ then $\Xi(F)$ is a singleton and $[G]$ belongs to the one-parameter extended gamma-type family of Example 2.9. We have seen examples where the limit set $\Xi(F)$ is a circle, a line, the whole space of types (for universal exponential families) and the empty set (Examples 4.9 and 4.10).

The asymptotic behaviour of the tail of the df F is reflected in the asymptotic behaviour of the tail of the mgf. The exponential family F_λ for $\lambda \uparrow \lambda_\infty$ describes the tail behaviour of F. In terms of the cgf κ the exponential family consists of translates of the graph of this convex function κ. It is not clear how the wide range of behaviour of the curve $[F_\lambda]$ hinted at in this section is reflected in the asymptotic behaviour of the convex analytic function κ. More insight into this question should lead to a better understanding of the relation between the tail behaviour of a df and its mgf.

Acknowledgements

Research on this topic was started during a visit to Cornell University some years ago by A.A. Balkema and C. Klüppelberg. They would like to thank the Department of Operations Research and Industrial Engineering for its hospitality. Sidney Resnick was partially supported by NSF grant DMS-97-04982 at Cornell University. We would also like to thank Muriel Casalis for her very careful reading of our paper and for her detailed comments.

References

Balkema, A.A., Klüppelberg, C. and Resnick, S.I. (1993) Densities with Gaussian tails. *Proc. London Math. Soc. (3)*, **66**, 568–588.

Balkema, A.A., Klüppelberg, C. and Stadtmüller, U. (1995) Tauberian results for densities with Gaussian tails. *J. London Math. Soc. (2)*, **51**, 383–400.

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

Bar-Lev, S.K. and Casalis, M. (1998) A general property of reproducibility. *Ann. Probab.* Submitted.

Barndorff-Nielsen, O.E. (1978) *Information and Exponential Families in Statistical Theory.* Chichester: Wiley.

Barndorff-Nielsen, O.E. and Cox, D.R. (1994) *Inference and Asymptotics.* London: Chapman & Hall.

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

Broniatowski, M. and Mason, D. (1994) Extended large deviations. *J. Theoret. Probab.*, **7**, 647–666.

Brown, L. (1986) Fundamentals of statistical exponential families. IMS Lecture Notes Monogr. Ser., Vol. 9. Hayward, CA: Institute of Mathematical Statistics.

Casalis, M. (1991) Familles exponentielles naturelles sur $\mathbb{R}^d$ invariantes par un groupe. *Internat. Statist. Rev.*, **59**, 241–262.

Davis, R. and Resnick, S.I. (1991) Extremes of moving averages of rv's with finite endpoint. *Ann. Probab.*, **19**, 312–328.

de Haan, L. (1970) *On Regular Variation and Its Application to Weak Convergence of Sample Extremes*, CWI Tract 32. Amsterdam: Centrum voor Wiskunde en Informatica.

de Haan, L. and Resnick, S.I. (1984) Asymptotically balanced functions and stochastic compactness of sample extremes. *Ann. Probab.*, **12**, 588–608.

Doeblin, W. (1946) Sur l'ensemble de puissances d'une loi de probabilité. *Ann. École Normale Supérieure (III)*, **63**, 317–350.

Feigin, P. and Yashchin, E. (1983) On a strong Tauberian result. *Z. Wahrscheinlichkeitstheorie Verw. Geb.*, **65**, 35–48.

Feller, W. (1966) *An Introduction to Probability Theory*, Vol. 2. New York: Wiley.

Feller, W. (1969) Limit theorems for probabilities of large deviations. *Z. Wahrscheinlichkeitstheorie Verw. Geb.* **14**, 1–20.

Hazod, W. and Scheffler, H.P. (1993) The domains of partial attraction of probabilities on groups and on vector spaces. *J. Theoret. Probab.*, **6**, 175–186.

Jensen, J.L. (1995) *Saddlepoint Approximations.* Oxford: Oxford University Press.

Lehmann, E.L. (1983) *Theory of Point Estimation.* New York: Wiley.

Pancheva, E. (1985) Limit theorems for extreme order statistics under non-linear normalization. In V.V. Kalashnikov and V.M. Zolotarev (eds), *Stability Problems for Stochastic Models*, Lecture Notes in Math. 1155, pp. 284–309. Berlin: Springer-Verlag.

Parthasarathy, K.R. (1967) *Probability Measures on Metric Spaces*. New York: Academic Press.

Ravi S. (1991) Tail equivalence of df's belonging to the max domains of attraction of univariate p-max stable laws. *J. Indian Statist. Assoc.*, **29**, 69–77.

Rootzen, H. (1987) A ratio limit theorem for the tails of weighted sums. *Ann. Probab.*, **15**, 728–747.

Teicher, H. (1984) The CLT for Esscher transformed rv's. *Sankhyā Ser. A*, **46**, 35–40.

Received November 1996 and Revised November 1998