Stability for multivariate exponential families

August A. Balkema   Claudia Klüppelberg*   Sidney I. Resnick

January 18, 2000

Abstract

A random vector X generates a natural exponential family of vectors X^λ, $\lambda \in \Lambda$, where Λ is the set where the moment generating function (mgf) $K(\lambda) = Ee^{\lambda X}$ is finite. Assume that Λ is open and X non-degenerate. Suppose there exist affine transformations $\alpha_\lambda(x) = A_\lambda x + a_\lambda$ depending continuously on the parameter λ and a non-degenerate vector Y so that

$$\alpha_\lambda^{-1}(X^\lambda) \Rightarrow Y$$

when λ diverges. In this paper it will be shown that the limit vector satisfies a stability relation. Some examples of this limit relation are presented.

AMS Subject Classifications: Primary 60F05; secondary 60E05,44A10

Keywords: Esscher, exponential family, Laplace transform, multivariate, stability

*Author for correspondence.

1 Introduction

Exponential group families are of special interest to statisticians. Recently there has also been interest in this subject from the quarter of mathematical physics. The book Faraut and Korányi [1994] gives a complete classification of all exponential families which are generated by Lebesgue measure on proper open symmetric cones invariant under a group of linear transformations. Exponential families are useful tools in the theory of Laplace transformations.

In this paper we consider multivariate exponential families which can be normalised to converge to a non-degenerate limit. The limit vector generates an exponential family which is stable in the sense that all members of the family are of the same type. In the univariate case the limit distributions are normal or gamma. The significance of such a limit theorem for the theory of Laplace transforms was perceived by Feigin and Yashchin [1976]. Nagaev treated the univariate case in a series of papers dating back to the early sixties. A general univariate theory can be found in two papers by Balkema, Klüppelberg and Resnick. These will be referred to as BKR [1999a] and BKR [1999b] in the sequel. The statistical theory for densities in the strong domain of attraction of the multivariate normal distribution has been treated in Barndorff-Nielsen and Klüppelberg [1999]. The present paper presents a number of general results for the multivariate case. We shall give concrete results for two multivariate generalisations of the gamma distribution: The uniform distribution on the unit ball, and the multivariate Laplace distribution e^{-r}/c. The latter distribution and its generalisations have also been considered by Nagaev and Zaigraev [1999].

Let us now be more specific. With a non-degenerate random vector Z is associated an open convex set D, the *convex domain* of Z. The closure of D is the smallest closed convex set which contains almost every realisation of Z. In this paper we assume that $P\{Z \in D\} = 1$. Let the mgf K of Z converge on a neighbourhood of the origin. The *domain* of K is the convex set $\Lambda = \{\lambda \mid K(\lambda) = e^{\lambda X} < \infty\}$. In this paper we assume that Λ is open.

Let Z^λ, $\lambda \in \Lambda$, be the exponential family generated by Z. We are interested in the asymptotic behaviour of the distribution of X^λ as λ tends to the boundary of Λ or to ∞. The Ansatz of this paper is that there exists a non-degenerate limit vector W. So

$$W_\lambda := \alpha_\lambda^{-1}(Z^\lambda) \Rightarrow W \qquad \lambda \to \partial. \tag{1.1}$$

This paper then addresses two questions: What can one say about the limit distribution? Given the limit law, what can one say about its domain of attraction?

Similar questions have been asked about limit laws for sums, giving rise to the concept of stable distributions, and about limit laws for maxima and residual life times. In all these cases we start with a family of variables which diverges. By a proper choice of norming constants such a family may be stabilised to converge to a non-degenerate limit law. The limit variable W exhibits a certain stability property. The main result of this paper is a similar stability property for the limit vector in (1.1).

The paper is organised as follows. The first part, sections 2 to 4, treats basics and proves stability of the limit family in (1.1). The middle part, sections 5 to 9, develops some ideas about the limit relation. In the last part we treat some special cases.

Notation This paper is concerned with probability distributions rather than random variables. The vectors Z^λ are only introduced as a notational convenience. We use the notation $X \equiv Y$ or $X = Y$ to express equality of distribution and $Y_n \Rightarrow Y$ to denote weak convergence of the associated probability distributions.

The basic set-up is a natural exponential family of probability distributions π^λ, $\lambda \in \Lambda$, on a euclidean space E. Here $d\pi^\lambda(z) = e^{\lambda z} d\pi(z)/K(\lambda)$ is the distribution of Z^λ. The convex domain D of Z is a subset of E, the domain Λ of the mgf K is a subset of the dual space E^T.

The group of all affine transformations on the d-dimensional space E

$$\alpha(x) = Ax + a \qquad \det A \neq 0,\ a \in E \tag{1.2}$$

is denoted by $\mathcal{A} = \mathcal{A}(d)$, and $\mathcal{A}^T$ denotes the group of affine transformations $\xi \mapsto \xi A + a^T$ on the dual space E^T. Affine transformations on E may be represented by square matrices of size $d+1$ with top row $(1, 0, \ldots, 0)$. The transpose matrix represents an element of $\mathcal{A}^T$.

The set $\mathcal{A}$ is an open subset of a euclidean space. Let $\mathcal{K}$ be a compact subgroup of $\mathcal{A}$. Then there exists a continuous function $N : \mathcal{A} \to [0, \infty)$ with the properties:

1) $\{N = 0\} = \mathcal{K}$;

2) $\{N \leq c\}$ is compact for each c;

3) $N(\alpha) = N(\alpha^{-1})$.

The number $N(\alpha)$ is not a norm but $N(\alpha, \beta) = N(\alpha^{-1}\beta)$ measures the "distance" between α and β modulo $\mathcal{K}$. If the distance is large then one of the terms is large:

Lemma 1.1 *For each $c_0 > 0$ there exists c_1 so that*

$$N(\alpha, \beta) > c_1 \ \Rightarrow\ N(\alpha) > c_0 \text{ or } N(\beta) > c_0.$$

Proof The set $\{N \leq c_0\}$ is compact and the product is continuous so $N(\alpha,\beta) \leq c_1$ if $N(\alpha), N(\beta) \leq c_0$. ¶

For a sequence (ω_n) of points in an open convex set Ω we write $\omega_n \to \partial$ or $\omega_n \to \partial_\Omega$ if any compact subset of Ω contains only finitely many terms of the sequence. Similarly for $f : \Omega \to M$ and a a point in the metric space M we write

$$f(\omega) \to a \qquad \omega \to \partial \tag{1.3}$$

if for each $\epsilon > 0$ there exists a compact set F in Ω so that $d(f(\omega), a) < \epsilon$ for $\omega \in \Omega \setminus F$. (The space $\Omega \cup \{\partial\}$ is the Alexandrov one point compactification of the locally compact space Ω.)

2 Exponential families

Given a non-zero Radon measure μ on the finite dimensional vector space E define measures μ_λ by

$$d\mu_\lambda(x) = e^{\lambda x} d\mu(x) \qquad \lambda \in E^T.$$

Set $K(\lambda) = \mu_\lambda(E) = \int e^{\lambda x} d\mu(x) \leq \infty$. The *domain* of K is the set Λ where K is finite. For any $\lambda \in \Lambda$ introduce a random vector Z^λ with probability distribution $\pi^\lambda = \mu_\lambda / K(\lambda)$. The family of vectors Z^λ, $\lambda \in \Lambda$, is called *the natural exponential family generated by* μ. If μ is the distribution of a vector Z we speak of the *exponential family generated by* Z.

The measures $c\mu_\lambda$, $c > 0$ and $\lambda \in E^T$, all generate the same exponential family apart from a trivial translation. So we may always assume that the exponential family is generated by a random vector $Z = Z^0$. Then Λ contains the origin. The vectors Z and $Z - p$ generate the same exponential family apart from a trivial translation of the convex set D. So we may assume that $EX = 0 \in D$ if we wish.

It is well known that Λ is convex. The mgf K is lower semi-continuous (by Fatou's Lemma). If Λ is open then $K(\lambda) \to \infty$ when λ tends to a point in $\partial\Lambda$. If the origin lies in the convex domain D of Z then $K(\lambda) \to \infty$ for $\lambda \to \infty$. This gives:

Proposition 2.1 *Suppose the domain Λ of the mgf K is open and the origin lies in the convex domain D of the exponential family. Then $K(\lambda) \to \infty$ for $\lambda \to \partial$.*

The mgf K is C^∞ (even analytic) on the interior of Λ. For interior points λ of Λ the moments of Z^λ are finite and may be computed by differentiation $EZ^\lambda_{i_1} \cdots X^\lambda_{i_n} = \partial_{i_1} \cdots \partial_{i_n} K(\lambda)$. The cumulant generating function (cgf) $\kappa = \log K$ yields the reduced moments:

$$\kappa'(\lambda) := EZ^\lambda \qquad \kappa''(\lambda) = \mathrm{var}(Z^\lambda). \tag{2.1}$$

So if Z is non-degenerate and Λ is open then κ is a strictly convex C^∞ function on Λ. The relations (2.1) are well known for $\lambda = 0$ and they follow for arbitrary λ by observing that $K_\lambda(\xi) = Ee^{\xi Z^\lambda} = \int e^{\xi x} e^{\lambda x} d\mu(x) = K(\lambda+\xi)/K(\lambda)$ and hence

$$\kappa_\lambda(\xi) = \kappa(\lambda+\xi) - \kappa(\lambda) \qquad \lambda, \lambda+\xi \in \Lambda. \tag{2.2}$$

Introduce the Esscher transforms E^λ, writing $Z^\lambda = E^\lambda Z$. Then

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

For future use note:

Lemma 2.2 *Let Z have cgf κ. Then the vector $\beta(Z^\lambda) = BZ^\lambda + b$ has cgf*

$$\kappa(\lambda+\xi B) - \kappa(\lambda) + \xi b.$$

Lemma 2.3 *Let X in E have cgf κ with domain $\Lambda \subset E^T$. Let $\beta(x) = Bx + b$ where $B : E \to F$ is a linear surjection and $b \in F$. Then $B^T : F^T \to E^T$ is a linear injection. For $\theta B = B^T(\theta) \in \Lambda$*

$$\beta(E^{\theta B} X) \equiv E^\theta(\beta(X)).$$

If π is a non-degenerate probability measure on E and the domain Λ of the exponential family $[\pi] := \{\pi^\lambda \mid \int e^{\lambda x} d\pi(x) < \infty\}$ is open then $[\pi]$ is a closed subset of the space of all non-degenerate distributions on E and is homeomorphic to Λ. See Theorem 8.3 in Barndorff-Nielsen [1978]. A corresponding result holds for general Λ.

Proposition 2.4 *Let π be a non-degenerate probability measure on E with cgf κ. Let Λ be the domain of κ and let $[\pi] = \{\pi^\lambda \mid \lambda \in \Lambda\}$ be the exponential family. Then*

1) $[\pi]$ is homeomorphic to the graph of κ in $\Lambda \times \mathbf{R}$;

2) $[\pi]$ is a closed set in the space of non-degenerate probability measures on E.

Corollary 2.5 *If $\lambda_n \to \partial$ and $\beta_n^{-1}(Z^{\lambda_n}) \Rightarrow W$ with W non-degenerate then β_n diverges in $\mathcal{A}$.*

We end this section with a few remarks on the limit relation (1.1).

There is a certain duality between the convex sets D and Λ if Λ is open. For any point $\lambda \in \Lambda$ let $z(\lambda) = EZ^\lambda$ denote the centre of gravity of the probability measure π^λ on D. If Z has density f then Z^λ has density $f^\lambda = e^\lambda f/K(\lambda)$. The factor $e^{\lambda z}/K(\lambda)$ introduces an exponential bias in the distribution of Z which shifts the center of gravity roughly in the direction of λ. The vectors λ and $z(\lambda)$ are coupled by a Legendre transform. Since the function κ is smooth and strictly convex the map $\lambda \mapsto z(\lambda) = \kappa'(\lambda)$ is a diffeomorphism of the open set Λ onto D. The correspondence need not extend to the boundaries.

This paper is concerned with the limit relation (1.1). We assume that Z and W are non-degenerate and that the domain Λ of the mgf is open. These conditions are imposed for the sake of simplicity. The condition of weak convergence for $\lambda \to \partial$ is a harsh condition but it leads to an interesting and useful basic theory.

Exponential families are most simply expressed in terms of densities. In statistical applications one is often interested in *strong convergence*:

$$g_\lambda(y) = |\det(A_\lambda)| f^\lambda(\alpha_\lambda(y))) \to g(y) \text{ uniformly on compact sets in } Q \text{ and in } \mathbf{L}^1.$$

The normalised densities converge uniformly on compact subsets of the convex domain Q of W. Here g is the density of the limit vector W, f the density of Z and $f^\lambda = e^\lambda f/K(\lambda)$.

We shall also consider vague convergence.

3 Stable exponential families

This section investigates stable exponential families.

Definition 3.1 *The random vector W or the exponential family W^θ, $\theta \in \Theta$, generated by W is called* stable *if Θ is open and if all vectors W^θ are of the same type.*

Remark 3.2 It suffices that the domain Θ contains an interior point θ_0.

A basic question for stable exponential families is: How are the index λ and the affine transformations linked? First let us make some remarks.

If X^μ is distributed like $\alpha(X^\lambda)$ for some $\lambda \neq \mu$ then α maps the exponential family bijectively into itself. Indeed let $\alpha(x) = Ax + a$. Then Lemma 2.3 gives for any ρ in Λ

$$\alpha^{-1}(X^\rho) = \alpha^{-1}(E^{\rho-\mu}X^\mu) = E^{(\rho-\mu)A}\alpha^{-1}X^\mu = E^{(\rho-\mu)A+\lambda}X.$$

So α^{-1} maps the exponential family into itself. By the same argument so does α.

Definition 3.3 *An element $\sigma \in \mathcal{A}$ is a* symmetry *of the non-degenerate random vector X if $\sigma X \equiv X$. The set of all symmetries is a closed compact group $\mathcal{K}$, the* symmetry group of the random vector X. *Let $[X] = \{X^\lambda \mid \lambda \in \Lambda\}$ denote the exponential family generated by X. The* symmetry group of the exponential family $[X]$ *is the set*

$$\mathcal{G} = \{\alpha \in \mathcal{A} \mid [\alpha X] = [X]\}. \tag{3.1}$$

The map $\alpha(x) = Ax + b \mapsto A$ from $\mathcal{A}(d)$ to $\mathcal{GL}(d)$ is a homomorphism. For any subgroup $\mathcal{G}$ of the affine group $\mathcal{A}$ we define the linear group $\mathcal{G}_0$ as the corresponding subgroup of $\mathcal{GL}$.

Proposition 3.4 *The set $\mathcal{G}$ in (3.1) is a closed subgroup of $\mathcal{A}$ which contains the compact symmetry group $\mathcal{K}$ of Y.*

There is a simple condition for stability of an exponential family in terms of the groups introduced above.

Define $\varphi : \mathcal{G} \to E^T$ by the relation $X^{\varphi(\alpha)} \equiv \alpha X$. Casalis [1990] has shown that

$$\varphi(\alpha\beta) = \varphi(\alpha)B + \varphi(\beta) \qquad \alpha, \beta \in \mathcal{G},\ \beta(x) = Bx + b.$$

This implies that the map

$$\alpha \sim \begin{pmatrix} 1 & 0 \\ a & A \end{pmatrix} \mapsto \alpha^* \sim \begin{pmatrix} 1 & \varphi(\alpha) \\ 0 & A \end{pmatrix}$$

is a homomorphism from $\mathcal{G}$ onto some group $\mathcal{G}^*$ in $\mathcal{A}^T$ which extends the canonical isomorphism $A \mapsto A^T$ from the linear group $\mathcal{G}_0$ introduced in definition 3.3 above. We shall call $\mathcal{G}^*$ the *Casalis group* associated with the exponential family $[X]$ and the homonorphism $\alpha \mapsto \alpha^*$ the *Casalis map*. Note that the Casalis map is continuous. So the Casalis group $\mathcal{G}^*$ is a Lie group and the Casalis map is C^∞. Since the symmetry group $\mathcal{K}$ is a subgroup of the linear group $\mathcal{G}_0$ the Casalis map is injective. By definition it is onto. So it is an isomorphism. We obtain the following result.

Theorem 3.5 *If W is stable the domain Θ of the cgf η of W is an orbit of the Casalis group $\mathcal{G}^*$ of the exponential family W^θ, $\theta \in \Theta$.*

This implies

Theorem 3.6 *Let W be a non-degenerate random vector in $\mathbf{R}^d$ with symmetry group $\mathcal{K}$. Let $\mathcal{G}$ be the symmetry group of the exponential family $[W] = \{W^\theta \mid \theta \in \Theta\}$. Here Θ is the domain of the mgf of W. If $\dim \mathcal{G} = \dim \mathcal{K} + d$ then W is stable.*

Corollary 3.7 *If there exist independent vectors $\theta_1, \ldots, \theta_d$ in Θ so that the vectors $W^{t\theta_i}$, $t \in [0,1]$, $i = 1, \ldots, d$, are all of the same type then W is stable.*

4 Stability of the limit

This section contains our main result.

Theorem 4.1 *Suppose $\gamma_n \to \gamma_0$, $U_n \Rightarrow U_0$ and $V_n \Rightarrow V_0$ where U_0 and V_0 are non-degenerate. Let K_n be the mgf of U_n for $n \geq 0$. Assume that $K_n(\gamma_n)$ is finite and $V_n = U_n^{\gamma_n}$ for $n \geq 1$. Then $V_0 = U_0^{\gamma_0}$ and $K_n(\gamma_n) \to K_0(\gamma_0) < \infty$.*

Proof Let π_n denote the probability distribution of U_n and ρ_n the distribution of V_n for $n \geq 0$. Then $e^{\gamma_n x} d\pi_n(x) = a_n d\rho_n(x)$ for $n \geq 1$ where $a_n = K_n(\gamma_n)$. Observe that

$$a_n d\rho_n(x) = e^{\gamma_n x} d\pi_n(x) \to e^{\gamma_0 x} d\pi_0(x) =: d\mu(x) \quad \text{vaguely.}$$

Since $\rho_n \to \rho_0$ vaguely and ρ_0 and μ do not vanish we conclude that a_n converges to a constant $a_0 \in (0, \infty)$ and $d\mu = a_0 d\rho_0$. Then $\int e^{\gamma_0 x} d\pi_0(x) = a_0 < \infty$ and $d\rho_0(x) = e^{\gamma_0 x} d\pi_0(x)/a_0$ since $\rho_0 E = 1$. ¶

Now return to the basic relation (1.1). The Esscher transform $E^\sigma Y_\lambda$ exists if σ is small. It satisfies

$$Y_\lambda^\sigma \equiv \beta Y_\mu \qquad \beta = \alpha_\lambda^{-1} \alpha_\mu, \quad \mu = \lambda + \sigma A_\lambda^{-1} \tag{4.1}$$

provided σ is chosen so that μ lies in Λ.

Theorem 4.2 *Let Z and W be non-degenerate vectors. Let Z^λ, $\lambda \in \Lambda$, be the exponential family generated by Z. Assume that Λ is open. Let $\alpha : \Lambda \to \mathcal{A}$ be a continuous normalisation so that (1.1) holds. Then the exponential family generated by W is stable.*

Proof We shall prove that for each $\sigma \in E^T$ there exists a $t_0 > 0$ so that the vectors $W^{t\sigma}$, $0 \leq t \leq t_0$, all are of the same type. This implies that the exponential family W^θ, $\theta \in \Theta$, generated by W is stable by Corollary 3.7.

Let $\sigma \in E^T$ and let $\lambda_n \to \partial$. Fix $n \geq 1$. For $t \geq 0$ relation (4.1) gives

$$W_{\lambda_n}^{t\sigma} = \beta_n(t) W_{\mu_n(t)} \qquad \mu_n(t) = \lambda_n + t\sigma A_{\lambda_n}^{-1}$$

provided $\mu_n(t) \in \Lambda$. Since Λ is convex and open this will be the case for $t \in [0, c_n)$ for some maximal $c_n \leq \infty$. Note that $\beta_n(0) = \mathrm{id}$ and $N(\beta_n(t)) \to \infty$ for $t \to c_n$ by Corollary 2.5. Choose $t_n \in (0, c_n)$ so that $N(\beta_n(t_n)) = 1$. If $t_n > 1$ then take $t_n = 1$.

Now assume $\mu_n = \mu_n(t_n) \to \partial$. Take a subsequence so that $\beta_n(t_n) \to \beta$ for some $\beta \in \mathcal{A}$ and $t_n \to t \in [0,1]$. Then $W_{\lambda_n} \Rightarrow W$, $t_n\sigma \to t\sigma$ and $W_{\lambda_n}^{t_n\sigma} \equiv \beta_n(t_n)W_{\mu_n(t_n)} \Rightarrow \beta W$ imply $W^{t\sigma} \equiv \beta W$ by Proposition 2.4. If $\mu_n \to \mu \in \Lambda$ for some subsequence then $N(\alpha_{\mu_n}) \to N(\alpha_\mu)$ and $N(\beta_n(t_n)) \to \infty$ by Lemma 1.1. This contradicts our choice of t_n. ¶

Remark 4.3 We may replace the continuity condition by a weaker condition. Suppose the limit vector in (1.1) has symmetry group $\mathcal{K}$. Let $\mathcal{A}/\mathcal{K}$ be the symmetric space of cosets $[\beta] = \beta\mathcal{K}$, $\beta \in \mathcal{A}$. It suffices that there exists a compact convex set $C \subset \Lambda$ so that α is continuous modulo $\mathcal{K}$ on $\Lambda \setminus C$, i.e. $[\alpha] : \Lambda \to \mathcal{A}/\mathcal{K}$ is continuous on $\Lambda \setminus C$.

We have now established our main result.

Lemma 4.4 *Suppose (1.1) holds. Let Θ be the domain of the mgf of the limit vector. Write $\alpha_\lambda(x) = A_\lambda x + a_\lambda$. Let $\lambda_n \to \partial$ and let $\theta_n \to \theta \in \Theta$. Define*

$$\mu_n(t) = \lambda_n + t\theta_n A_{\lambda_n}^{-1} \qquad \beta_n(t) = \alpha_{\lambda_n}^{-1}\alpha_{\mu_n(t)} \qquad t \geq 0.$$

Then the sequence $(\beta_n(t_n))$ in $\mathcal{A}$ is bounded for any sequence (t_n) in $[0,1]$ and $\mu_n(t_n) \to \partial$.

Proof Let c_0 be the maximum of N over the compact set of all $\alpha \in \mathcal{A}$ for which there exists $t \in [0,1]$ such that $W^{t\theta} \equiv \alpha W$. Let $c_1 = c_0 + 1$. Suppose $\beta_n(t_n)$ is unbounded for some sequence (t_n) in $[0,1]$. Choose a subsequence so that $N(\beta_n(t_n)) \to \infty$. By continuity there exist $r_n \in (0, t_n)$ so that $N(\beta_n(r_n)) \to c_1$. Take a subsequence so that $\beta_n(r_n) \to \beta$ and $r_n \to r$. Then $N(\beta) = c_1$ and $W^{r\theta} \equiv \beta W$ as in the proof of Theorem 4.2. By definition $N(\beta) \leq c_0$. Contradiction. ¶

Weak convergence in (1.1) implies convergence of the cgf's.

Theorem 4.5 *Suppose (1.1) holds. Let η_λ denote the cgf of the normalised variable $W_\lambda = \alpha_\lambda^{-1}(Z^\lambda)$ and let Θ be the domain of the cgf η of the limit variable W. Then $\eta_\lambda \to \eta$ uniformly on compact subsets of Θ for $\lambda \to \partial$.*

Proof Let $\theta_n \to \theta \in \Theta$, $\lambda_n \to \partial$. The sequence $(\beta_n(1))$ in Lemma 4.4 is relatively compact and all limit points β satisfy $W^\theta \equiv \beta W$ and $W_{\lambda_n}^{\theta_n} \Rightarrow W^\theta$. Theorem 4.1 gives $\eta_{\lambda_n}(\theta_n) \to \eta(\theta)$. ¶

5 The support of a stable family

The support of a stable exponential family with symmetry group $\mathcal{G}$ is an atom in the σ-field of $\mathcal{G}$ invariant sets. This explains why the multivariate theory for exponential families is similar to the univariate theory for limits of sums and maxima rather than the multivariate theory. See Balkema and Qi [1997] for related results on multivariate limit laws of residual life times.

Given a closed subgroup $\mathcal{G}$ of $\mathcal{A}$ introduce the σ-algebra $\mathcal{E} = \mathcal{E}(\mathcal{G})$ of invariant Borel sets of the space E. It is generated by the sets

$$\mathcal{G}K = \{\alpha(x) \mid \alpha \in \mathcal{G}, x \in K\} \qquad K \subset E, K \text{compact}.$$

Theorem 5.1 *Let the exponential family W^θ, $\theta \in \Theta$, be invariant under the closed subgroup $\mathcal{G}$. Then the support S of W is an atom of $\mathcal{E}$.*

Proof First note that S is the support of W^θ for every $\theta \in \Theta$. So $\beta W \equiv W^\theta$ implies that $\gamma S = S$. The support is invariant. Since it is σ-compact it is a set in $\mathcal{E}$.

Suppose there exists a set $S_0 \in \mathcal{E}$ so that $P\{W \in S_0\} = p_0 \in (0,1)$. Then $P\{W \in S_1\} = p_1 = 1 - p_0$ for $S_1 = S \setminus S_0$. Note that

$$Ee^{\theta W} 1_{\{W \in S_0\}} + Ee^{\theta W} 1_{\{W \in S_1\}} = Ee^{\theta W} = K(\theta).$$

Let $W^\theta \equiv \beta W$. Then for $i = 0, 1$

$$p_i = P\{\beta W \in S_i\} = P\{W^\theta \in S_i\} = Ee^{\theta W} 1_{\{W \in S_i\}}/K(\theta).$$

So $K(\theta)$ is the mgf of the conditional distribution of W given $W \in S_i$ for $i = 0$ and $i = 1$. Hence these conditional distributions agree. Contradiction. ¶

6 Domains of attraction

As in the limit theory for sums and maxima one is not only interested in the limit laws, but even more in a description of their domains of attraction.

Definition 6.1 *Let W be a non-degenerate random vector which generates a stable exponential family W^θ, $\theta \in \Theta$. The exponential family Z^λ, $\lambda \in \Lambda$, belongs to the* domain of attraction *of the exponential family W^θ and we write $Z \in \mathcal{D}(W)$ if Λ is open and if (1.1) holds for some family* *of normalisations $\alpha_\lambda \in \mathcal{A}$. (We do not assume continuous dependence on λ here.) Similarly we say that a random vector, a measure, a density, a mgf or cgf belongs to the domain of attraction if this holds for the exponential families generated by these objects.*

Two measures which are asymptotically equal belong to the same domains of attraction.

Theorem 6.2 *Suppose (1.1) holds for the random vector Z with mgf K with domain Λ. Let $\tilde{Z}$ have mgf $\tilde{K}$ with domain Λ and suppose $\tilde{K}(\lambda) \sim K(\lambda)$ for $\lambda \to \partial$. Then $\alpha_\lambda^{-1}(\tilde{Z}_\lambda) \Rightarrow W$.*

Proof Let the limit vector W have cgf η with domain Θ. The vector $W_\lambda = \alpha_\lambda^{-1}(Z^\lambda)$ has cgf

$$\eta_\lambda(\xi) := \kappa(\lambda + \xi B_\lambda) - \kappa(\lambda) + \xi b_\lambda \qquad \alpha_\lambda^{-1}(y) = B_\lambda y + b_\lambda$$

by Lemma 2.2. The vector $\tilde{W}_\lambda$ has cgf $\tilde{\eta}_\lambda$ with κ replaced by $\tilde{\kappa}$ in the formula above. By assumption $(\tilde{\kappa} - \kappa)(\lambda) \to 0$ for $\lambda \to \partial$. This implies $(\tilde{\eta}_{\lambda_n} - \eta_{\lambda_n})(\theta_n) \to 0$ for $\theta \in \Theta$ and $\lambda_n \to \partial$ by Lemma 4.4. So $\eta_\lambda \to \eta$ on Θ implies $\tilde{\eta}_\lambda \to \tilde{\eta}$ on Θ. Convergence of the mgfs implies convergence in law. ¶

The asymptotic behaviour of the mgf $K(\lambda)$ for $\lambda \to \partial$ is determined by the asymptotic behaviour of the distribution of Z for $z \to \partial_D$.

Lemma 6.3 *Let μ be a finite measure with convex support C. Let $H = \{\xi_0 \leq c_0\}$ be a half space so that $C \setminus H$ is bounded and non-empty. Then there exists an open cone $\Gamma \subset \Lambda$ containing ξ_0 so that*

$$\mu_\gamma H / \mu_\gamma C \to 0 \qquad \gamma \in \Gamma,\ \gamma \to \infty.$$

Lemma 6.4 *Let μ be a finite measure with convex support C and ρ a finite measure with convex support A contained in the interior of C. Let the domain Λ of the mgf K_μ be open. Then*

$$K_\rho(\lambda)/K_\mu(\lambda) \to 0 \qquad \lambda \to \partial_\Lambda.$$

Theorem 6.5 *Let μ be a finite measures with convex domain D. Assume that μ lives on D and that the domain Λ of the mgf K of μ is open. Let $h : D \to [0, \infty)$ be a Borel function such that $\int h d\mu$ is finite and $h(x) \to 1$ for $x \to \partial_D$. Then the mgf K_0 of $d\mu_0 = h d\mu$ has domain Λ and K_0 and K are asymptotically equal for $\lambda \to \partial$.*

Proof Let $\epsilon > 0$. There is a compact convex set $A \subset D$ so that $|\log h(x)| < \epsilon$ for $x \in D \setminus A$. Let μ_ϵ agree with μ on A and with μ_0 on $D \setminus A$. Then the cgfs satisfy $|\kappa_\epsilon(\lambda) - \kappa(\lambda)| < \epsilon$ for all $\lambda \in \Lambda$ and $|\kappa_\epsilon(\lambda) - \kappa_0(\lambda)| < \epsilon$ eventually. So $|\kappa_0(\lambda) - \kappa(\lambda)| < 2\epsilon$ eventually. ¶

7 The construction of stable exponential families

Given a group $\mathcal{G}$ in $\mathcal{A}$ how does one construct a measure such that the natural exponential family associated with this measure is invariant under $\mathcal{G}$?

We give an example to describe a procedure.

Example 7.1 The two-dimensional commutative group of diagonal matrices with positive elements gives rise to the exponential families generated by the bivariate gamma densities on $(0,\infty)^2$. There is one other two-dimensional commuatative group $\mathcal{G}$ of linear transformations, with matrices

$$A = \begin{pmatrix} a & 0 \\ b & a \end{pmatrix} \qquad a > 0, b \in \mathbf{R}.$$

The right half plane is invariant and $\mathcal{G}$ is transitive on this half plane since $A(1,0) = (a,b)$. Assume there exists a density $f = e^{-\varphi}$ such that $f \circ A = f_\lambda$ for some vector λ depending on a and b. This means that $\varphi(x,y) + L(x,y) = \varphi(ax, bx+ay)$ for some linear function L depending on a and b. Taking derivatives with respect to x and y and writing φ_{ij} for $\partial_i\partial_j\varphi(ax, bx+ay)$ we find

$$\begin{aligned} \varphi_{xx} &= a^2\varphi_{11} + 2ab\varphi_{12} + b^2\varphi_{22} \\ \varphi_{xy} &= a^2\varphi_{12} + ab\varphi_{22} \\ \varphi_{xx} &= a^2\varphi_{22}. \end{aligned}$$

Now given x and y choose a and b so that $(ax, bx+ay) = (1,0)$. This gives $a = 1/x$ and $b = -y/x^2$. The three quantities $c_{i+j} = \varphi_{ij}(1,0)$ are constant. So we obtain

$$\begin{aligned} \varphi_{xx} &= \frac{c_2}{x^2} - \frac{2c_3 y}{x^3} + \frac{c_4 y^2}{x^4} \\ \varphi_{xy} &= \frac{c_3}{x^2} - \frac{c_4 y}{x^3} \\ \varphi_{xx} &= \frac{c_4}{x^2}. \end{aligned}$$

These functions have to satisfy the integrability conditions $\varphi_{xxy} = \varphi_{xyx}$ and $\varphi_{xyy} = \varphi_{yyx}$. This yields the solution

$$\varphi(x,y) = a_0 + a_1 x + a_2 y + a_3 \log x + a_4 y/x \qquad f(x,y) = c_0 x^{a_3} e^{a_1 x + a_2 y} e^{a_4 y/x}.$$

The corresponding measures on $(0,\infty)\times\mathbf{R}$ are invariant. However none of them is finite.

8 Vague convergence and continuation

Suppose (1.1) holds. There is a finite measure μ with convex domain D and a finite measure ρ with convex domain Q so that

$$\alpha_\lambda^{-1}(\mu_\lambda)/K(\lambda) \to \rho \text{ weakly} \qquad \lambda \to \partial.$$

Here $K(\lambda) = \mu_\lambda(E)$ and $d\mu_\lambda(x) = e^{\lambda x} d\mu$. So ρ is a probability measure.

Let $d\tilde{\mu} = h d\mu$ for a Borel function h on D and $d\tilde{\rho} = h_0 d\rho$ for a continuous strictly positive function h_0 on Q, and assume that there exist positive constants $C(\lambda)$ so that

$$\alpha_\lambda^{-1}(\tilde{\mu}_\lambda)/C(\lambda) \to \tilde{\rho} \text{ vaguely} \qquad \lambda \to \partial. \tag{8.1}$$

This will be the case if

$$h(\alpha_\lambda(y))/c(\lambda) \to h_0 \text{ uniformly on compact sets in } Q. \tag{8.2}$$

Note that

$$c(\lambda) \sim |\det A_\lambda|/C(\lambda) \sim h(p_\lambda)/h_0(q) \qquad \lambda \to \partial$$

for $p_\lambda = \alpha_\lambda(q)$ and any $q \in Q$. So we find

Proposition 8.1 *Suppose (8.2) holds. Let the corresponding measures $d\tilde{\mu} = h d\mu$ and $d\tilde{\rho} = h_0 d\rho$ be finite. Weak convergence holds in (8.1) if and only if the mgf $\tilde{K}(\lambda) = \tilde{\mu}_\lambda(D)$ of $\tilde{\mu}$ satisfies the asymptotic equality*

$$\tilde{K}(\lambda) \sim \tilde{\rho}(Q) h(p_\lambda)/(h_0(q)|\det A_\lambda|) \qquad \lambda \to \partial.$$

Vague convergence in (8.1) implies vague convergence on any open subset of Q. Conversely if (8.1) holds on any non-empty open subset of Q then it holds on Q. This is the main result of the present section.

Suppose $\rho_n \to \rho$ vaguely on the open set W and $\rho W > 0$. If $c_n \rho_n \to \sigma$ vaguely on W then $c_n \to c \geq 0$ and $\sigma = c\rho$. This implies

Lemma 8.2 *Suppose $\beta_n \to \beta \in \mathcal{A}$, $\rho_n \to \rho$ vaguely on the open set U and $c_n \beta_n \rho_n \to \sigma$ vaguely on U. If $\rho(U \cap \beta U) > 0$ then $c_n \to c \geq 0$ and $\sigma = c\beta\rho$ on $U \cap \beta U$.*

Now suppose (8.1) holds on the non-empty open set $U \subset Q$. Let

$$\lambda_n \to \partial \quad \sigma_n \to \sigma \in \Gamma \quad \tau_n = \lambda_n + \sigma_n A_{\lambda_n}^{-1} \quad \beta_n = \alpha_{\lambda_n}^{-1} \alpha_{\tau_n}.$$

Since we assume that (1.1) holds we find $\tau_n \to \partial$ and $[\beta_n] \to [\beta]$. Then

$$\beta_n \rho_{\tau_n} = \alpha_{\lambda_n}^{-1}(\mu_{\tau_n}/C(\tau_n) = \alpha_{\lambda_n}^{-1}(E^{\sigma_n A_{\lambda_n}^{-1}}\mu_{\lambda_n})/C(\tau_n) = E^{\sigma_n}\rho_{\lambda_n} C(\lambda_n)/C(\tau_n).$$

So $E^{\sigma_n}\rho_{\lambda_n} \to E^{\sigma}\rho$ vaguely on U and $\beta_n\rho_{\tau_n} \to \beta\rho$ vaguely on βU. If $\rho(U \cap \beta U)$ is positive then $C(\tau_n)/C(\lambda_n) \to c \geq 0$ and $\beta\rho = cE^{\sigma}\rho$ on $U \cap \beta U$.

Theorem 8.3 *Suppose (1.1) holds. Let μ be a Radon measure on D and ρ a finite measure on a relatively compact open U in Q. Suppose $\rho_\lambda = \alpha_\lambda^{-1}(\mu_\lambda)/C(\lambda) \to \rho$ vaguely on U for $\lambda \to \partial$. Then ρ extends to a unique Radon measure $\bar{\rho}$ on Q such that $\rho_\lambda \to \bar{\rho}$ vaguely on Q.*

Corollary 8.4 *Let $h : D \to [0,\infty)$ be a Borel function and $h_0 : U \to (0,\infty)$ continuous. If $h \circ \alpha_\lambda^{-1} \to h_0$ uniformly on U then h_0 extends to a continuous function $\bar{h}_0$ on Q so that $h \circ \alpha_\lambda^{-1} \to \bar{h}_0$ uniformly on compact subsets of Q. The function $\bar{h}_0$ is a quasi-multiplieer for the symmetry group $\mathcal{G}$ of W^θ, $\theta \in \Theta$.*

Given convergence in (1.1) for a density f it is fairly easy to adapt f or the underlying Lebesgue measure so that vague convergence holds on some relativley compact open subset of Q. All that then remains to be done is to prove weak convergence!

9 The geometry

A *geometry* on D is a collection $\mathcal{E}$ of ellipsoids E_p, $p \in D$. A geometry on D determines a metric on D, the collection of flat functions, and the collection of measures which are asymptotically Lebesgue. Suppose (1.1) holds. There are three ways to generate the geometry $\mathcal{E}$.

1) Any non-degenerate random vector X in E determines a family of ellipsoids centered in EX. These are the level curves of the density of a gaussian variable with the same first two moments and can be represented as $E_r = \{EX + u \mid \Sigma^{-1}(u,u) \leq r^2\}$ where Σ is the covariance. Fix $r > 0$ and let $\mathcal{E}$ be the collection of the ellipsoids $E_{r\lambda}$ associated with the vectors Z^λ.

2) The cgf κ of Z is convex and analytic. It determines a quadratic form $\kappa''(\lambda)$ in each point $\kappa'(\lambda)$ of D and hence as in 1) a collection $\mathcal{E}$ of ellipsoids on D.

3) Let Q be the convex domain of the limit vector W in (1.1). Assume that W is centered and has unit variance. Let $B = B_r(0)$ be the closed ball of radius r. Define $\mathcal{E}$ to be the set of all ellipsoids $\alpha_\lambda(B)$.

Proposition 9.1 *Suppose (1.1) holds. For any compact convex sets $B \subset Q$ and $C \subset D$ there is a compact convex set $A \subset \Lambda$ so that $\alpha_\lambda(B) \subset D \setminus C$ for $\lambda \in \Lambda \setminus A$.*

The third method may be generalised. The set B may be any compact convex set in Q which is invariant under the symmetry group $\mathcal{K}$ of W. Thus if $D = (0, \infty)^d$, $Q = (-\infty, 1)^d$ and W is the vector with independent gamma components it is convenient to take B the cube $[1/2, 3/2]^d$.

In (1.1) we may replace α_λ by $\tilde{\alpha}_\lambda = \alpha_\lambda \epsilon_\lambda$ provided any sequence $\lambda_n \to \partial$ contains a subsequence λ'_n so that $\epsilon_{\lambda'_n} \to \epsilon \in \mathcal{K}$, where $\mathcal{K}$ is the symmetry group of W. If B is invariant under $\mathcal{K}$ the ellipsoids $\tilde{E}_\lambda$ and E_λ are asymptotic in the sense that for any $r > 1$ the ellipsoid $\tilde{E}_\lambda$ is eventually enclosed between two concentric scalings $E_\lambda(1/r)$ and $E_\lambda(r)$ of E_λ.

We shall assume that the collection $\mathcal{E}$ has the property of Proposition 9.1. Collections which are asymptotically equal for $p = \kappa'(\lambda) \to \partial$ define the same geometry.

Theorem 9.2 *The geometry depends only on the normalisations α_λ, $\lambda \in \Lambda$, and not on the distribution of Z or W.*

Definition 9.3 *A function $h : D \to [0, \infty)$ is* flat *if it is positive outside some compact convex subset C of D and if it is asymptotically constant on the ellipsoids in the geometry $\mathcal{E}$: If $p_n \to \partial_D$ then $\sup\{h(x)/h(y) \mid x, y \in E_{p_n}\} \to 1$. A measure μ on D is* asymptotically Lebesgue *if $1_B \alpha_\lambda^{-1}(\mu)/|\det A_\lambda|$ converges weakly to the uniform distribution on B. In particular $\mu(E_\lambda) \sim |E_\lambda| = |B||\det A_\lambda|$.*

Suppose (1.1) holds. Let $f : D \to [0, \infty)$ be a Borel function and $g : Q \to [0, \infty)$ be continuous with $g(q) > 0$ for some $q \in Q$. Set $g_\lambda = |\det A_\lambda|(e^\lambda f) \circ \alpha_\lambda$. Suppose

$$g_\lambda / c(\lambda) \to g \qquad \text{uniformly on compact sets in } Q. \tag{9.1}$$

Then $c(\lambda) \sim g_\lambda(q)/g(q)$. Let $\tilde{f} = hf$ with h flat. Then (9.1) holds for $\tilde{g}$ with the normalisation $\tilde{c}(\lambda) = \tilde{g}_\lambda(q)/g(q)$. If (9.1) holds in $\mathbf{L}^1$ then $g_\lambda/K(\lambda) \to g/\int g dz$ where $K(\lambda) = \int e^\lambda f dz$ is the mgf of the density f. Hence

$$K(\lambda) \sim g_\lambda(q) \int g dz / g(q) = (|\det A_\lambda| \int g dz / g(q))(e^\lambda f)(\alpha_\lambda(q)) \qquad \lambda \to \partial.$$

If (9.1) also holds in $\mathbf{L}^1$ for $\tilde{f}$ then

$$\tilde{K}(\lambda) \sim K(\lambda) h(\alpha_\lambda(q)). \tag{9.2}$$

In some cases the geometry is determined by the form of the set D or Λ. We give four examples

1) If $D = Q$ is a symmetric cone with symmetry group $\mathcal{G}$ and if the limit vector W has the characteristic function of Q as density or a gamma function then $\mathcal{E} = \{\gamma(B) \mid \gamma \in \mathcal{G}\}$ where B is a symmetric compact body in Q. See Faraut and Korányi [1994].

2) If D is an open body which contains the origin then each point $p \in D$ has the representation $p = (1-t)b$ with $t \in (0,1]$ and $b \in \partial D$. Call D *rotund* if the boundary is C^2 and has positive curvature. See Section 10 for details. Then for $p \neq 0$ there exists a unique ellipsoid E_p^*. which osculates ∂D in b. Take E_p to be the ellipsoid concentric to E_p^* scaled down by a factor 1/2. This yields the geometry associated with the uniform distribution on D.

3) If Λ is a rotund body which contains the origin then so is the interior of the polar set Λ^o. For $p = tb$, $b \in \Lambda^o$, $t > 0$, let E_p^* be the ellipsoid centered in p which osculates the convex set $b - \Lambda^o$ in the origin, and define E_p to be concentric to E_p^* and scaled by a factor 1/2. This geometry is associated with the vector Z with the Laplace density e^{-r}/c where r is the norm function of Λ^o.

4) The geometry on D is *asymptotically euclidean* if the map $p \mapsto Q_p$ is flat. Here Q_p is the quadratic function associated with the ellipsoid $E_p \in \mathcal{E}$. So

$$Q_{p'}[x,x]/Q_p[x,x] \to 1 \quad \text{uniformly in } x \neq 0 \qquad p_n \to \partial_D,\ p_n' \in E_{p_n}.$$

See Barndorff-Nielsen and Klüppelberg [1999]. This characterises the domain of attraction of the normal distribution.

On the unit disk there is the geometry associated with the uniform distribution, but there also is a host of distinct geometries on D associated with a normal limit, even if we restrict ourselves to densities with circular symmetry.

A geometry $\mathcal{E}$ determines a rough integer valued distance: Count the number of ellipsoids needed to form a chain from p_1 to p_2. This metric can be refined by using smaller ellipsoids. In the limit we obtain the Riemannian metric induced by the quadratic functions associated with the ellipsoids.

Define the *cumulant metric* on D as the Riemannian metric with the quadratic functions $Q_p = \kappa''(\lambda)$ for $p = \kappa'(\lambda)$. In the cumulant metric bounded sets are relatively compact.

Theorem 9.4 *Let d be the Riemannian metric on Λ induced by the quadratic forms $\kappa''(\lambda)^{-1}$, and let d_c be the cumulant metric on D. Then the Legendre transform is an isometry.*

The geometry does not distinguish metrics which are asymptotic to the cumulant metric. In terms of the random vector Z this means that we do not distinguish between the vectors Z and $\tilde{Z}$ if $\tilde{\kappa}''(\lambda) \sim \kappa''(\lambda)$ for $\lambda \to \partial$. Often it suffices for the metric d to be *equivalent* to the cumulant metric d_c in the sense that the quotient $d(p,p')/d_c(p,p')$ is bounded and bounded away from zero.

On the unit sphere the distance between two points is equal to the angle between two rays. On the surface of a rotund body there is an intrinsic Riemannian metric d_0 defined in terms of the curvature. It is invariant under affine transformations. Assume that D is rotund and contains the origin. Each point has the form $p = (1-t)b$ with $t \in (0,1)$ and $b \in \partial D$. The geometry in the second example above determines a metric on D. If $p_i = (1-t_i)b$ lie on the same ray then $d(p_1,p_2) = |\log(t_1/t_2)|$, if $p_i = (1-t)b_i$ lie on the same surface then $d(p_1,p_2) = d_0(b_1,b_2)/\sqrt{t}$. Similarly in the third example $d(t_1b, t_2b) = |\log(t_1/t_2)|$ and the distance between two points $p_i = tb_i \in t\partial\Lambda^o$ is $d_0(b_1,b_2)\sqrt{t}$ where d_0 is the intrinsic metric on the surface of Λ^o. This metric is equivalent to the metric with Λ^o replaced by the unit ball.

If D or Λ is a rotund body we may extend the natural conjugation between the boundary of the body and the boundary of the polar body. Thus if Λ is rotund and $\omega \in \partial\Lambda$ then $\omega^o = b \in \partial\Lambda^0$ satisfies $\omega b = 1$ and $\lambda b < 1$ for $\lambda \in \Lambda$. For $\lambda = (1-1/t)\omega$ with $0 < t < 1$ set $\lambda^o = tb$. A similar conjugation exists for the uniform distribution on D. This conjugation mimicks the Legendre transform.

Proposition 9.5 *Let D be a rotund body and d a metric equivalent to the metric described above. Then there exists a constant $M \geq 1$ so that*

$$d(p_0,p) \leq M(|\log(t/t_0)| + \sqrt{1 + |\lambda_0(p_0-p)|}) \qquad p, p_0 \in D.$$

Here $\lambda_0 = p_0^o$, $t, t_0 \in (0,1)$ satisfy $p = (1-t)b$ and $p_0 = (1-t_0)b_0$ with $b, b_0 \in \partial D$.

Corollary 9.6 *If $h : D \to (0,\infty)$ is flat then for any $\epsilon > 0$ there is a compact convex set $A \subset D$ and a constant $M \geq 1$ so that*

$$h(p)/h(p_0) \leq Me^{\epsilon|\log(t/t_0)|}e^{\epsilon|\lambda_0(p_0-p)|} \qquad p, p_0 \in D \setminus A.$$

Proposition 9.7 *Let Λ be a rotund body and d equivalent to the associated metric. There exists a constant $M \geq 1$ so that*

$$d(z,z_0) \leq M(|\log(r/r_0)| + 1 + \sqrt{r - \lambda_0(z)}) \qquad z, z_0 \in E.$$

Here r, r_0 is the value of the norm function of Λ^o in z, z_0 and $\lambda_0 = z_0^o$ is the conjugate of z_0.

Corollary 9.8 *If $h : E \to (0,\infty)$ is flat then for any $\epsilon > 0$ there is a compact set $A \subset E$ and a constant $M \geq 1$ so that*

$$h(z)/h(z_0) \leq M e^{\epsilon|\log(r/r_0)|} e^{\epsilon|r-\lambda_0(z)|} \qquad z, z_0 \in A^c.$$

These bounds make it possilbe to insert a flat function into the density of Z without affecting the weak convergence of $\alpha_\lambda^{-1}(Z^\lambda)$.

10 Uniform distribution on the unit ball

Let $Z = (X, Y)$ be uniformly distributed on the open disk

$$D = \{(x,y) \mid (x-1)^2 + y^2 < 1\}.$$

The density is $f = 1_D/\pi$. Let $\lambda = (0, -\tau)$ with $\tau > 0$. Then Z^λ has density $f^\lambda(x,y) = c(\lambda)e^{-\tau y}1_D(x,y)$. For τ large the mass concentrates in a thin region close to the origin. So we blow up the disk into a long vertical ellipse E_τ centered in $(0,\tau)$ and passing through the origin. Then the random vector has density $c(\lambda)e^{-y}1_{E_\tau}$. In order to obtain a limit for $\tau \to \infty$ we have to ensure that the curvature in the origin does not blow up. This can be achieved by also expanding the ellipse in the horizontal direction, by a factor $\sqrt{\tau}$. The limit of these ellipses for $\tau \to \infty$ is the open parabola $Q = \{y > x^2/2\}$. The limit vector has density $g(x,y) = e^{-y}/\sqrt{2\pi}$.

The disk and its boundary are invariant under rotations. There is a corresponding one parameter group of skew translations $(x,y) \mapsto (x+t, y+xt+t^2/2)$ which leave Q and its boundary invariant. In addition there is the group of expansions $(x,y) \mapsto (rx, r^2y)$ which were used to transform the circle into a parabola. These also leave the limit set Q invariant. They add a multiplicative constant to Lebesgue measure which drops out when we normalise. Lebesgue measure on Q is quasi-invariant under a two-dimensional group $\mathcal{G}$ in $\mathcal{A}$. It generates a stable exponential family.

Now start with a unit ball in $\mathbf{R}^d \times \mathbf{R}$ centered in $(0,1)$. The procedure above yields the open convex set

$$Q = \{v > 0\} \qquad v = y - x^T x/2, \quad (x,y) \in \mathbf{R}^d \times \mathbf{R} \tag{10.1}$$

which is invariant under the $d+1$ dimensional group $\mathcal{G}$ of affine transformations generated by

$$\alpha_p \sim \begin{pmatrix} 1 & 0^T & 0 \\ p & I & 0 \\ p^T p/2 & p^T & 1 \end{pmatrix} \qquad \beta_r \sim \begin{pmatrix} 1 & 0^T & 0 \\ 0 & rI & 0 \\ 0 & 0^T & 1 \end{pmatrix} \qquad p \in \mathbf{R}^d,\ r > 0.$$

The skew translations α_p leave Lebesgue measure invariant, the diagonal linear transformations β_r only add an innocuous factor.

So let μ be Lebesgue measure restricted to Q. The measure μ_λ has density $e^{\lambda z}$ (by definition) and is finite if and only if λ lies in $\Lambda = \mathbf{R}^d \times (0, \infty)$. The Casalis isomorphism maps α_p into α_p^* and β_r into β_r^* where

$$\alpha_p^* \sim \begin{pmatrix} 1 & p^T & 0 \\ 0 & I & 0 \\ 0 & p^T & 1 \end{pmatrix} \qquad \beta_r^* \sim \begin{pmatrix} 1 & 0^T & 0 \\ 0 & rI & 0 \\ 0 & 0^T & 1 \end{pmatrix} \qquad p \in \mathbf{R}^d,\ r > 0.$$

This is the basic construction.

Now start with the uniform distribution on the surface of the unit ball. This distribution generates the Von Mises-Fisher exponential family. The arguments above yield a stable exponential family on the paraboloid $\partial Q = \{y = x^T x/2\}$. It is generated by the measure μ which is the image of Lebesgue measure on the horizontal hyperplane $\mathbf{R}^d \times \{0\}$ under the map $x \mapsto (x, x^T x/2)$. The measure μ_λ is finite for $\lambda \in \mathbf{R}^d \times (-\infty, 0)$. For $\lambda = (0, -1)$ it projects onto a multiple of the standard normal distribution on the horizontal plane $\{y = 0\}$.

Theorem 10.1 *Let Q and v be defined by (10.1). For $s > 0$ let W_s have density*

$$h_s(x, y) = w^{s-1} e^{-y} 1_Q(x, y)/(2\pi)^{d/2}\Gamma(s)$$

and for $s = 0$ let W_0 be the vector $(X_0, X_0^T X_0/2)$ where X_0 has a standard normal distribution on $\mathbf{R}^d$. Each vector W_s, $s \geq 0$, generates a stable exponential family W_s^θ, $\theta \in \Theta = \mathbf{R}^d \times (-\infty, -1)$.

What can one say about the domains of attraction? Note that the examples above give a global result which depends only on the local behaviour of the original distribution of Z.

Let D be a convex bounded open set in $\mathbf{R}^{d+1}$ whose boundary can locally be described by a C^2 function with a positive definite second derivative. Let U be uniformly distributed on D. Then U lies in the domain of attraction of W_1. If U is uniformly distributed on the surface ∂D it lies in the domain of attraction of W_0. These two results remain true if we replace the uniform distribution by a density which is continuous on the closure of D and positive on the boundary.

Now suppose D is the open unit ball. Let r denote the distance from the centre and set $t = 1 - r$ on D. The random vector Z with density $f_s \propto t^{s-1}$ lies in the domain of attraction of W_s for $s > 0$. Again we may multiply this density by any continuous function on the closed ball which is positive on the boundary sphere. Let A be a compact convex subset of D. Let Z' have a probability distribution which agrees with the distribution of Z on $D \setminus A$. Then Z' lies in the domain of attraction of Z_s. See Lemma 6.4.

What happens if $s = 0$? The function $g = (\log(1-r))^2/(1-r)$ is integrable over the unit ball D. Let the vector Z have density $\propto g1_C$. Then Z lies in the domain of attraction of W_0.

Let us concentrate now on the simple case $s = 1$. For which densities h on the ball D will the corresponding vector Z lie in the domain of attraction of W_1? This turns out to be determined by the geometry on D. The associated metric is akin to the hyperbolic metric. Take $d = 1$ and use polar coordinates (r, θ) on the open disk $D = \{x^2 + y^2 < 1\}$. Assume f is strictly positive on D and write $f = e^{-\varphi}$. If φ is C^1 and

$$(1-r)\frac{\partial\varphi}{\partial r} \to 0 \qquad \sqrt{1-r}\frac{\partial\varphi}{\partial\theta} \to 0 \qquad r \to 1-0 \tag{10.2}$$

then f is flat. See Section 9 for details.

Flat functions are less flat than they seem. It is possible to construct a C^1 function f on the open disk C_0 which is flat in the sense of (10.2) but which also has the following property: For each constant $c \in \mathbf{R}$ there exists a dense subset $S_c \subset [0, 2\pi]$ so that

$$\lim_{r \to 1-0} f(r, \theta) = c \qquad \theta \in S_c.$$

10.1 Rotund bodies

Definition 10.2 *A* body *is a bounded convex set which contains an interior point. Assume the body B contains the origin as interior point. There is a unique function $r = r_B$ on E, the* norm function *of B with the properties: $r(0) = 0$, $\{r = 1\} = \partial B$, and $r(cx) = cr(x)$ for $c > 0$ and $x \in E$. It is sometimes more convenient to work with the* tent function *on B. This is the function $t = 1 - r$. A tent function may have its top in an interior point $p \neq 0$.*

The convex sets C and D (or their boundaries) osculate *in a common boundary point b if the tangent planes in b coincide, and if the functions ψ_C and ψ_D which describe the boundaries around b have the same second derivative in that point.*

Given affine coordinates $(x, y) \in \mathbf{R}^d \times \mathbf{R}$ on E let $U_0 \subset \{y = 0\}$ denote the vertical projection of D. There exists a continuous function $\psi_0 : U_0 \to \mathbf{R}$ so that $(u, \psi_0(u)) \in \partial D$ for all $u \in U$.

In fact there exist two such functions. We choose the convex function. It describes the lower boundary of D.

The following are equivalent for any $k \geq 1$:

1) ∂D is a C^k manifold;

2) The norm function r is C^k on $E \setminus \{0\}$;

3) The functions $\psi_0 : U_0 \to \mathbf{R}$ above are C^k.

Let $b \in \partial D$ be a boundary point of D. Assume there is a unique supporting plane T_b to D in b, the tangent plane. We regard T_b as a hyperplane in E. It becomes a vector space T_b^0 by declaring b to be the origin. So a point $a \in T_b \subset E$ corresponds to the vector $w = a - b \in T_b^0$. Choose new coordinates $(x, y) \in \mathbf{R}^d \times \mathbf{R}$ on E so that $b = (0, 0)$ and $D \subset \{y > 0\}$. Then T_b is the horizontal plane $\{y = 0\}$.

Let $U_b \subset T_b$ be the vertical projection of D. Set $W_b = U_b - b \subset T_b^0$, and let $\psi_b : W_b \to \mathbf{R}$ describe the lower boundary of D. Then $\psi_b(0) = 0$ and $\psi_b'(0) = 0$. Suppose $\psi_b''(0)$ exists and is positive definite. Then there are two euclidean structures on T_b^0, the euclidean structure on $T_b \subset E$ inherited from E, and the intrinsic euclidean structure determined by the symmetric positive definite bilinear form $\psi_b''(0)$. We endow T with the intrinsic structure. One may choose affine coordinates on E so that the unit sphere in E centered in $(0, 1)$ osculates ∂D in b. Then the two structures coincide in b.

Definition 10.3 *The body D is* rotund *if D is open, ∂D is a C^2 manifold and if the* curvature *$\psi_b''(0)$ is positive definite for each $b \in \partial D$.*

Let D be a rotund body, $p \in D$. For each boundary point b there exists a unique ellipsoid $E_b(p)$ centered in p which osculates D in b. These ellipsoids vary continuously with b.

Let T denote the tangent bundle to ∂D. We may identify T with the set of pairs $\{(b, w) \mid b \in \partial D, b + w \in T_b \subset E\}$. The set $W = \{(b, w) \mid b \in \partial D, w \in W_b \subset T_b^0\}$ is open in T. Define $\psi : W \to \mathbf{R}$ by $\psi(b, w) = \psi_b(w)$, and similarly define the partial derivatives $\psi_w(b, w) = \psi_b'(w)$ and $\psi_{ww}(b, w) = \psi_b''(w)$.

Proposition 10.4 *The functions ψ and ψ_w are C^1 on W.*

The set $A_\delta = \{(b, w) \mid \psi_b(0)[w, w] \leq \delta^2\}$ is a compact subset of T which is contained in W for some $\delta > 0$. Uniform continuity of ψ on A_δ implies

Lemma 10.5 *Let $b_n \in \partial D$ and $w_n \in T^0_{b_n}$. Suppose $b_n \to b$ and $w_n \to w$. Let $\tau_n \to \infty$. Then*

$$\tau_n \psi_{b_n}(w_n/\sqrt{\tau_n}) \to w^T w/2.$$

10.2 Weak convergence for strictly smooth convex bodies

Proposition 10.6 *Let t be the tent function on the rotund body D in $d+1$-dimensional space E for some $d \geq 1$. For $s > 0$ the density $t^{s-1}1_D$ lies in the domain of attraction of the stable exponential family generated by the density*

$$g = v^{s-1}e^{-y}1_Q \qquad Q = \{v > 0\},\ v = y = x^2/2.$$

The normalised densities converge uniformly on compact subsets of Q.

Proof The proof consists of two steps: We first prove vague convergence, then weak convergence. The domain Λ of the mgf of μ is the whole space E^T, hence is open. Take $\lambda \neq 0$ and let $b \in \partial D$ be the point where $\tau = \lambda b$ is maximal. Then $\{\lambda = \tau\}$ is the tangent plane to ∂D in b. We assume that the center of gravity of D is the origin.

Define $\alpha_\lambda = \sigma_b B_\tau$ where σ_b is an initial transformation to bring the convex body into the right position. Choose σ_β so that $D_b = \sigma_b^{-1}(D) \subset \mathbf{R}^d \times \mathbf{R}$ lies in the upper half space $\{y > 0\}$, with center of gravity $(0,1)$, and so that $\sigma_b^{-1}(b) = (0,0)$. Then the horizontal coordinate plane $\{y = 0\}$ is the tangent plane and the function e^λ on D transforms into the function $e^\tau e^{-\tau y}$ on D_b. We also choose σ_b so that the unit sphere centered in $(0,1)$ osculated D_b in $(0,0)$.

Now treat D_b as we did the unit ball. Blow it up by a factor τ in the vertical direction and by the factor $\sqrt{\tau}$ in the horizontal directions. Set

$$Q_\lambda = B_\tau^{-1}(D_b) = \alpha_\lambda^{-1}(D) \qquad B_t^{-1}(x,y) = (\sqrt{\tau}x, \tau y).$$

Then $\alpha_\lambda(0,1) = p = (1 - 1/\tau)b$ and the lower boundary of Q_λ is given by the function

$$\psi_p(x) = \tau\psi_b(x/\sqrt{\tau}) \to x^2/2 \qquad \text{uniformly in } b \in \partial D. \tag{10.3}$$

The function $e^{-\tau y}$ on D_b transforms into the function e^{-y} on Q_λ, and the tent function t on D into the tent function t_λ on Q_λ. Since the center of gravity of Q_λ is $(0,\tau)$ the limit relation (10.3) gives

$$\tau t_\lambda(x,y) \to y - x^2/2 \qquad \tau \to \infty,\ (x,y) \in Q.$$

So the densities, properly normalised, converge to $g(x,y) = (y - x^T x/2)^{s-1} e^{-y} 1_Q$. It remains to prove weak convergence

$$\int_{Q_\lambda} g_\lambda dxdy \to \int_Q gdxdy \qquad \lambda \to \partial. \tag{10.4}$$

For $\delta > 0$ let $D_b(\delta)$ be the polar cap $D_b \cap \{y \leq \delta\}$ of the set D_b around the point $(0,0)$. Then

$$P\{X^\lambda \in D_b(\delta)\} \to 1 \qquad \lambda \to \partial$$

uniformly in $b \in \partial D$ by Lemma 6.3. So it suffices to take the integral (10.4) over the set

$$Q_\lambda(\delta) = \{\psi_p(x) < y \leq \delta\tau\}.$$

Instead of the functions g_λ we shall consider the functions

$$h_\lambda(x,y) = (y - \psi_p(x))^{s-1} e^{-y} 1_{Q_\lambda(\delta)}(x,y).$$

Note that by Lemma 10.5 we may choose $\delta > 0$ so small that

$$g_\lambda(x,y) \leq 2^{s-1} h_\lambda(x,y) \qquad (x,y) \in Q_\lambda(\delta), \lambda \in E^T \setminus \{0\}.$$

Tightness of the measures $h_\lambda dxdy$ implies tightness of the measures $g_\lambda dxdy$. We claim that

$$h_\lambda \to g1_Q \text{ a. e. and in } \mathbf{L}^1.$$

The a.e. convergence was proved above. The $\mathbf{L}^1$ convergence is a simple computation:

$$\begin{aligned}
\int_{Q_\lambda(\delta)} h_\lambda dxdy &= \int_{\mathbf{R}^d} \int_{\psi_p(x)}^{\delta\tau} (y - \psi_{b,\tau}(x))^{s-1} e^{-y} dxdy \\
&= \int_{\mathbf{R}^d} e^{-\psi_p(x)} \int_0^{(\delta\tau\psi_p(x))+} e^{-v} dxdv \\
&\to (2\pi)^{d/2}\Gamma(s) = \int_Q (y - x^T x/2)^{s-1} e^{-y} dxdy \qquad \lambda \to \ partial.
\end{aligned}$$

The limit relation follows from (10.3) and convexity of ψ_b. ¶

Corollary 10.7 *The mgf $K(\lambda) = \mu_\lambda(E)$ satisfies the asymptotic equality*

$$K(\lambda) \sim (\Gamma(s)\Gamma((d+3)/2)|E_b|/\sqrt{\pi})/\tau^{d/2+s} \qquad lambda \to \partial.$$

Here $|E_b|$ is the volume of the ellipsoid osculating D in the boundary point b and $\{\lambda = \tau\}$ is the tangent plane in b.

We may insert a flat function h into the density without harming the weak convergence, see Section 9. It is also possible to replace Lebesgue measure by any measure μ which is asymptotically Lebesgue. This can be done by defining suitable partitions and showing that the change in the integral is managable if one redefines the measure of the atoms of the partition without changing the total mass of the atom.

Theorem 10.8 *Let D be a rotund body in $d+1$-dimensional euclidean space with $d \geq 1$. Let Z be a random vector with distribution*

$$d\pi = t^{s-1} h d\mu \tag{10.5}$$

where t is a tent function on D, s a positive parameter, h a flat function and μ a finite measure which is asymptotically Lebesgue. Then Z lies in the domain of attraction of the stable exponential family generated by the random vector $W_s$in Theorem 10.1. Conversely any random vector Z with convex domain D which is attracted to W_s has a distribution of the form (10.5).

10.3 Spherically symmetric distributions

Let π be a probability measure on the unit disk which is concentrated on a sequence of concentric circles with radii $r_n \to 1$. Assume that π is uniformly distributed over each of these circles. Can one choose the sequence r_n so that π is in the domain of attraction of W_1?

Let D be the unit ball in $\mathbf{R}^d \times \mathbf{R}$ centered in $q = (0,1)$. Let t denote the tent function on D with top in q and $\omega = \omega(p)$ the angle in q between the point $p \in D$ and $(0,0)$. Then $y - 1 = (s-1)\cos|\omega|$ and for $|\omega| < \pi/4$ and $\tau > 0$

$$e^{-\tau t} e^{-\tau|\omega|^2/2} \leq e^{-\tau y} \leq e^{-\tau t/2} e^{-|\omega|^2/4}.$$

Let $\mu = \sigma \times \mu_0$ be a product measure on $S^d \times (0,1]$ where σ is the uniform distribution. The map $(\omega, s) \to (\sqrt{\tau}\omega, \tau s)$ maps μ into a product measure μ_τ on $\sqrt{\tau} S^d \times (0,\tau]$. Now suppose $\rho_\tau = \mu_\tau / c(\tau) \to \rho$ vaguely. Then

$$e^{-\xi^2/2} e^{-y} d\rho_\tau(\xi, y) \to e^{-x^2/2} e^{-y} d\rho(x,y) \quad \text{vaguely.}$$

Suppose that

$$\int e^{-\xi^2/4} e^{-y/2} d\rho_\tau(\xi, y) \to \int e^{-x^2/4} e^{-y} d\rho(x,y).$$

Then

$$e^{-y\cos(\xi/\tau)} e^{-(1-\cos\xi/\tau)} d\rho_\tau(\xi, y) \to e^{-y} e^{-x^2/2} d\rho(x,y)$$

weakly for $\tau \to \infty$. The latter is equivalent to convergence $\alpha_\lambda^{-1}(Z^\lambda) \Rightarrow W$ for $\lambda = (0, -\tau)$ and $W = f(V)$ where the vector V has distribution $ce^{-y}e^{-x^2/2}d\rho$ for some $c > 0$ and $f(x,y) = (x, y + x^2/2)$ maps the half space $y > 0$ onto the paraboloid Q.

Convergence $\alpha_\lambda^{-1}(Z) \Rightarrow W$ for a vector Z on the unit ball with the symmetric probability distribution corresponding to μ is equivalent to convergence of the univariate exponential family generated by the probability measure μ_0 on $(0,1]$. This is equivalent to regular variation of the distribution function $M_0(t) = \mu_0(0,t]$ for $t \to 0$. See BKR [1999b].

Theorem 10.9 *Let the random vector Z have the open unit ball D as convex domain. Suppose that Z is spherically symmetric. Define*

$$M(t) = P\{\|X\| \geq 1 - t\} \qquad t \in (0,1].$$

Let $s > 0$. The vector Z lies in the domain of attraction of W_s in Theorem 10.1 if and only if M varies regularly in $t = 0$ with exponent s.

11 The Laplace distribution

The function

$$f = e^{-r}/c \qquad c = 2\pi^{d/2}\Gamma(d)/\Gamma(d/2),\ r = \|x\|_2 \tag{11.1}$$

is a probability density on $\mathbf{R}^{d+1}$.

Proposition 11.1 *For $s > -(1 + d/2)$ the function*

$$g_s(x,y) = y^s e^{-(y + x^T x/2y)}/2\pi^{d-1}\Gamma(s + d/2 + 1)$$

is the density of a vector W_s on $D = \mathbf{R}^d \times (0,\infty)$ which generates a stable exponential family.

Proposition 11.2 *The density f in (11.1) lies in the domain of attraction of the density g_0.*

More generally one can show

Theorem 11.3 *Let Λ be a rotund body in $\mathbf{R}^{d+1}$ which contains the origin and let r be the norm function of the polar set Λ^o. Let h be a flat function for the geometry associated with this norm function. Let $s + d/2 + 1 > 0$. Then the vector Z with density $f = he^{-r}r^s$ is attracted to the vector W_s in Proposition 11.1.*

Proof It is not hard to see that vague convergence holds.

The norm function r is C^2 outside the origin and can be lifted to the tangent bundle T of the set Λ^o. Set $\varphi(b,w) = r(b+w) - 1$ for $b+w \in T_b$, $b \in \partial\Lambda^o$. Then continuity of φ_{ww} together with $\varphi_{ww}(b,0) \equiv I$ imply

$$\frac{\varphi(b,w)}{w^T w} \to \frac{1}{2} \qquad \frac{r_b(x,1)-1}{x^T x} \to \frac{1}{2} \qquad \frac{r_b(tx,t^2 y)}{t^2} \to \frac{x^T x}{2y}$$

for $w \to 0$, $x \to 0$, $t \to \infty$, all uniformly in $b \in \partial\Lambda^o$. We can now choose $\delta_0 > 0$ independent of the boundary point b so that in coordinates $(x,y) \in \mathbf{R}^d \times \mathbf{R}$ such that $b = (0,1)$ and Λ^o osculates the unit sphere in b one has the inclusion $\Lambda^o \subset E_2$ on $\{y \geq 1-\delta\}$ where $E_2 = \{x^T x/2 + y^2 \leq 1\}$. Then there exist δ and δ_1 in $(0,1)$ so that

$$\|x\| \geq \delta y > 0 \ \Rightarrow \ r(z) - y \geq \delta_1 \|z\| \qquad z = (x,y).$$

These bounds ensure that convergence of the densities holds in $\mathbf{L}^1$. ¶

For spherically symmetric distributions of the form $e^{-\|z\|}d\mu(z)$ one has convergence in (1.1) if and only if the function $M(r) = \mu\{\|z\| \leq r\}$ varies regularly. The proof is as for Theorem 10.9.

Nagaev and Zagraiev [1999] have treated the theory of the multivariate Laplace distribution and its generalisations from a different angle. What happens if λ approaches a fixed point on the boundary of Λ?

References

A.A. Balkema, C. Klüppelberg and S.I. Resnick (1999) Domains of attraction for exponential families and asymptotic behaviour of Laplace transforms. Submitted.

A.A. Balkema, C. Klüppelberg and S.I. Resnick (1999) Limit laws for exponential families.

A.A. Balkema and Yong-Cheng Qi (1998) Strictly stable laws for multivariate residual lifetimes. Proceedings of the 18th Seminar on Stability Problems for Stochastic Models, Part I (Hajdúszoboszló, 1997). J. Math. Sci. (New York) **92** 3873-3880.

O.E. Barndorff-Nielsen, P. Blæsild and P.S. Eriksen (1989) *Decomposition and Invariance of Measures, and Statistical transformation Models.* L.N. Statistics **58**. Springer, Berlin.

O.E. Barndorff-Nielsen and C. Klüppelberg (1999) Tail exactness of multivariate saddlepoint approximations. *Scand. J. Stat.* **26** 253-264.

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

J. Faraut and A. Korányi (1994) *Analysis on symmetric cones.* Clarendon Press, Oxford.

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

A. Nagaev and A. Zaigraev (1999) Abelian theorems for a class of probability distributions in $\mathbb{R}^d$ and their application. Preprint.

August A. Balkema
Department of Mathematics
University of Amsterdam
NL-1018TV Amsterdam
Holland
guus@wins.uva.nl

Claudia Klüppelberg
Center of Mathematical Sciences
Munich University of Technology
D-80290 Munich
Germany
cklu@ma.tum.de
www.ma.tum.de/stat/

Sidney I. Resnick
ORIE
Cornell University
Ithaca, NY 14853-7501
USA
sid@orie.cornell.edu