Elliptic functions and exponential families

Gérard Letac, Université Paul Sabatier
Laboratoire de Statistique et Probabilités.
Roma, November 7th, 2005

Natural exponential families on $\mathbb{R}$. Let μ be a non Dirac measure on $\mathbb{R}$ with Laplace transform $L_\mu(\theta) = \int_{-\infty}^{\infty} e^{\theta x}\mu(dx) \leq \infty$. Assume that the interior $\Theta(\mu)$ of the interval $\{\theta \in \mathbb{R}; L_\mu(\theta) < \infty\}$ is not empty. Let us write $k_\mu = \log L_\mu$. Then the following family of probabilities

$$F = F(\mu) = \{P(\theta, \mu)\ ;\ \theta \in \Theta(\mu)\}$$

where

$$P(\theta, \mu)(dx) = e^{\theta x - k_\mu(\theta)}\mu(dx)$$

is called the natural exponential family (NEF) generated by μ. The function k_μ is strictly convex and $k'_\mu(\theta) = \int_{-\infty}^{\infty} xP(\theta, \mu)(dx)$. The set $k'_\mu(\Theta(\mu)) = M_F$ is called the domain of the means.

Denote by $\psi_\mu : M_F \to \Theta(\mu)$ the inverse function of k'_μ. Therefore $F(\mu)$ is parameterized by M_F as follows

$$m \mapsto P(\psi_\mu(m), \mu) = P(m, F).$$

The variance $V_F(m)$ of $P(m, F)$ is

$$V_F(m) = k''_\mu(\psi_\mu(m)) = \frac{1}{\psi'_\mu(m)}. \qquad (1)$$

The map $m \mapsto V_F(m)$ from M_F to $(0, \infty)$ is called the variance function. It characterizes F. The best way to interest a probabilist to variance functions is to recall that the explicit formula for large deviations : if $X_1, \dots, X_n$ are real iid rv with distribution $P(m, F)$ and if $S_n = X_1 + \cdots + X_n$ then

$$\lim_{n \to \infty} \frac{1}{n} \log \Pr\left(\frac{S_n}{n} \geq m + \epsilon\right) = -\int_m^{m+\epsilon} \frac{x - m}{V_F(x)} dx.$$

How to move from V_F to a generating measure μ? There are four steps :

- $d\theta = \psi'_\mu(m)dm = \frac{dm}{V_F(m)}$ leads by a quadrature to $\theta = \psi_\mu(m)$.
- An inversion of function gives $m = k'_\mu(\theta)$,
- A second quadrature gives $k_\mu = \log L_\mu$.
- A good dictionary of Laplace transforms leads to μ.

The Jorgensen set $\Lambda(\mu)$ is the set of $t > 0$ such that there exists μ_t with $\Theta(\mu_t) = \Theta(\mu)$ and $L_{\mu_t} = (L_\mu)^t$. It is an addititive semi-group containing $\mathbb{N}^*$. When $t \in \Lambda(\mu)$ we denote $F_t = F(\mu_t)$. Trivially we have $M_{F_t} = tM_F$ and $V_{F_t}(m) = tV_F(\frac{m}{t})$. The union $G = G(\mu) = \cup_{t\in\Lambda(\mu)} F(\mu_t)$ is called the exponential dispersion model generated by μ.

If F is a NEF and if $h(x) = ax+b$ (with $a \neq 0$) then $h(F)$ is a NEF with $M_{h(F)} = h(M_F)$ and

$$V_{h(F)}(m) = a^2 V_F(\frac{m-b}{a}) \qquad (2)$$

S. Bar-Lev, a statistician, has asked me a few months ago whether

$$V(m) = (1 + Cm^4)^{1/2}$$

is a variance function or not. The question is a very interesting one. Here is a list of reasons

- Carl Morris in 1982 has observed that variance functions which are second degree polynomials correspond to Meixner distributions and are completely known (normal, Poisson, binomial, negative-binomial, gamma and hyperbolic)
- One can dream of a complete classification of variance functions of the type $V(m) = \sqrt{Am^4 + Bm^2 + C}$. Experience shows that when $\lim_{m\to\infty} V(m)/m^2$ does exist, then quite often the distributions or their Laplace transforms are computable.
- The dual variance functions of $\sqrt{Am^4 + Bm^2 + C}$ are tractable NEF concentrated on positive integers.

Some technique shows that $C > 0$ and the question is : for which values of $t > 0$ can we claim that the function on $\mathbb{R}$ defined by

$$V_t(m) = t(1 + \frac{4m^4}{t^4})^{1/2}$$

is a variance function ? Who is μ_t?

We shall use for solving this problem the venerable theory of elliptic functions. May I mention that the best of the textbooks on the subject was written by Professor G. Sansone ?

More generally if $-2 \leq a < -1$ for which values of $t > 0$ can we claim that the function on $\mathbb{R}$

$$V_t(m) = t((1 + a\frac{m^2}{t^2})^2 + \frac{4m^2}{t^2})^{1/2}$$

is a variance function ? For answering we denote $k^2 = 1 + a = p - 1 \in [-1, 0)$ (The case $a = -2, k^2 = -1, p = 0$ is the preceeding case). We introduce

$$\begin{aligned} K &= \int_0^1 (1 - x^2)^{-1/2}(2 - p - x^2)^{-1/2}dx \\ K' &= \int_0^1 (1 - x^2)^{-1/2}(1 + (1 + p)x^2)^{-1/2}dx \end{aligned}$$

If $p = 0$ then

$$K = K' = B(1/4, 1/2) = \frac{\Gamma(1/4)^2}{2\sqrt{2\pi}}.$$

One now applies the four steps procedure to $t = 1$:

First quadrature

The minimum of $(1+am^2)^2+4m^2$ is reached on $m = 0$. Therefore one uses the change of variable $u^2 = (1 + am^2)^2 + 4m^2$ and the change of variable

$$u = (1 + \frac{1}{a})w^2 - \frac{1}{aw^2} = \frac{1}{a}(k^2w^2 - \frac{1}{w^2})$$

avec $0 < w < 1$.

$$m^2 = \frac{1}{a^2w^2}(1 - w^2)(1 - k^2w^2) \qquad (3)$$

One gets

$$\begin{aligned} d\theta &= \frac{1}{V(m)} \times dm \\ &= \frac{dw}{awm} \\ &= -\frac{dw}{\sqrt{(1 - w^2)(1 - k^2w^2)}} \end{aligned}$$

The first step is now performed

$$\theta = \int_0^m \frac{dx}{V_F(x)} = \int_{w(m)}^1 \frac{dw}{\sqrt{(1 - w^2)(1 - k^2w^2)}}$$

Note that $\Theta(\mu) = (-K', K')$.

Inversion and second quadrature

Denote $C(\theta) = w(m)$. One gets $C(0) = 1$ and the function C is defined on $[0, K']$. In $(0, K')$ it satisfies $C'(\theta) = -(1-C(\theta)^2)^{1/2}(1-k^2C(\theta)^2)^{1/2}$. Hence

$$\begin{aligned} m &= k'_{\mu_t}(\theta) = m(C(\theta)) \\ &= \frac{1}{|a|C(\theta)}(1 - C(\theta)^2)^{1/2}(1 - k^2C(\theta)^2)^{1/2} \\ &= \frac{C'(\theta)}{aC(\theta)}. \end{aligned}$$

Finally the Laplace transform of μ_t is

$$L_{\mu_t}(\theta) = \frac{1}{(C(\theta))^{t/|a|}}.$$

Other presentation of the Laplace transform

Now consider the restriction $c(s) = C(is)$ to the imaginary axis. It satisfies the differential equation $c'(s)^2 = (c(s)^2 - 1)(1 - k^2 c(s)^2)$ with the initial condition $c(0) = 1$. We now introduce $s \mapsto f(s) = -k^2 c^2(s)$ which satisfies

$$f'(s)^2 = 4(f(s) + 1) f(s)(f(s) + k^2).$$

Denoting $f(s) = -\frac{p}{3} + h(s)$ we get

$$h'(s)^2 = 4h(s)^3 - g_2 h(s) - g_3$$

with

$$g_2 = 4(1 - p + \frac{p^2}{3}) \text{ et } g_3 = -\frac{4p}{3}(1 - p + \frac{2p^2}{9}).$$

In other words h is a solution of the differential equation defining the Weierstrass function $\wp$ for g_2 and g_3. Note that in the case $p = 0$, the initial case, then $h'(s)^2 = 4h(s)^3 - 4h(s)$.

This can be rewritten

$$h'^2(s) = 4(h(s) - e_1)(h(s) - e_2)(h(s) - e_3)$$

with $e_1 = 1 - \frac{2p}{3} > e_2 = \frac{p}{3} > e_3 = -1 + \frac{p}{3}$ and discriminant

$$\Delta = g_2^3 - 27g_3^2 = [4(1-p)(2-p)]^2.$$

Now the periods of $\wp$ are $2K$ and $2iK'$

$$\begin{tikzcd}
2iK' \arrow[r, dash] \arrow[d, dash] & K + 2iK' \arrow[r, dash] \arrow[d, dash] & 2K + 2iK' \arrow[d, dash] \\
iK' \arrow[r, dash] \arrow[d, dash] & K + iK' \arrow[r, dash] \arrow[d, dash] & 2K + iK' \arrow[d, dash] \\
0 \arrow[r, dash] & K \arrow[r, dash] & 2K
\end{tikzcd}$$

Remember : in general

$$\wp(K) = e_1,\ \wp(K + iK') = e_2,\ \wp(iK') = e_3.$$

We certainly cannot have $h(s) = \wp(s)$ since 0 is a pole of $\wp$. However $h(s) = \wp(s + C)$ for some constant C.

Since the variance function is symmetric there exists a symmetric measure which generates it. Thus the characteristic function

$$s \mapsto \frac{1}{f(s)^{t/'2|a|)}}$$

is real. We have to take C such that $c(0) = 1$ or $f(0) = 1 - p$ or $h(0) = 1 - \frac{2p}{3}$ or $\wp(C) = e_1$. Hence $C = K$. As a result we have

$$L_{\mu_t}(is) = \left(\frac{1-p}{\wp(s+K) - \frac{p}{3}}\right)^{\frac{t}{2|a|}}$$

Dictionary : Finding the measure μ_t

The periods of $s \mapsto \wp(s)$ are $2K$ and $2iK'$. Its poles are on $2aK + 2ibK'$ with $(a,b) \in \mathbb{Z}^2$. Thus the characteristic function

$$s \mapsto \left(\frac{1-p}{f(s)}\right)^{t/2|a|}$$

has zeros on the odd multiples of K. A more important fact is that it has period $2K$. This implies that it concentrated on the multiples of π/K :

$$\sum_{\nu \in \mathbb{Z}} p_\nu(t) e^{i\nu\frac{\pi}{K}} = p_0(t) + 2\sum_{\nu=1}^{\infty} p_\nu(t) \cos\frac{\pi\nu}{K}.$$

The last task is to show that we have a probability. For this we have to compute $p_\nu(t)$ in order to decide whether they are positive or not. With this aim we compute the Fourier series of $-\frac{t}{2|a|}\log f(s)$

We now pass from Weierstrass to Jacobi. Let $q = e^{-\pi\frac{K'}{K}}$ (and thus $q = e^{-\pi}$ si p=0). We introduce the functions

$$\begin{aligned}\vartheta_1(z) &= i\sum_{\nu\in\mathbb{Z}}(-1)^\nu q^{(\nu-\frac{1}{2})^2}e^{(2\nu-1)\pi iz}\\ \vartheta_3(z) &= \sum_{\nu\in\mathbb{Z}} q^{\nu^2}e^{2\nu\pi iz}\end{aligned}$$

$$f(2Kz)^{1/2} = \sqrt{\wp(2Kz+K)-p/3} = \frac{\vartheta_3(z+\frac{1}{2})}{C\vartheta_1(z+\frac{1}{2})}$$

where C is some constant. We obtain

$$-\frac{1}{2}\log f(s) = \log C + \log\vartheta_1(\frac{s}{2K}+\frac{1}{2}) - \log\vartheta_3(\frac{s}{2K}+\frac{1}{2}).$$

We take derivatives :

$$-\frac{1}{2}(\log f(s))' = \frac{1}{2K}\frac{\vartheta_1'(\frac{s}{2K}+\frac{1}{2})}{\vartheta_1(\frac{s}{2K}+\frac{1}{2})} - \frac{1}{2K}\frac{\vartheta_3'(\frac{s}{2K}+\frac{1}{2})}{\vartheta_3(\frac{s}{2K}+\frac{1}{2})}.$$

One now use the formulas of classical analysis for ϑ_j'/ϑ_j. This gives

$$\frac{\vartheta_1'(z)}{\vartheta_1(z)} = \pi\frac{\cos\pi z}{\sin\pi z} + 4\pi\sum_{\nu=1}^{\infty}\frac{q^{2\nu}}{1-q^{2\nu}}\sin 2\nu\pi z$$
$$\frac{\vartheta_3'(z)}{\vartheta_3(z)} = 4\pi\sum_{\nu=1}^{\infty}\frac{(-1)^\nu q^{\nu}}{1-q^{2\nu}}\sin 2\nu\pi z$$

In such an expression we replace z by $\frac{s}{2K}+\frac{1}{2}$ and we get

$$\frac{\vartheta_1'(\frac{\pi s}{2K}+\frac{\pi}{2})}{\vartheta_1(\frac{\pi s}{2K}+\frac{\pi}{2})} = -\pi\frac{\sin\frac{\pi s}{2K}}{\cos\frac{\pi s}{2K}} + 4\pi\sum_{\nu=1}^{\infty}\frac{(-1)^\nu q^{2\nu}}{1-q^{2\nu}}\sin\frac{\pi s}{K}$$
$$\frac{\vartheta_3'(\frac{\pi s}{2K}+\frac{\pi}{2})}{\vartheta_3(\frac{\pi s}{2K}+\frac{\pi}{2})} = 4\pi\sum_{\nu=1}^{\infty}\frac{q^{\nu}}{1-q^{2\nu}}\sin\frac{\pi s}{K}$$

Finally $-\frac{K}{\pi}(\log f(s))'$ is equal to

$$-\frac{\sin\frac{\pi s}{2K}}{\cos\frac{\pi s}{2K}} + 4\sum_{\nu=1}^{\infty} \frac{(-1)^\nu q^{2\nu} - q^\nu}{1-q^{2\nu}} \sin\frac{\pi s}{K},$$

and up to some additive constant we get $-\frac{1}{2}\log f(s)$ as the following sum

$$\log\cos\frac{\pi s}{2K} + \sum_{\nu\in\mathbb{Z}\setminus\{0\}} \frac{q^{|\nu|} - (-1)^\nu q^{2|\nu|}}{1-q^{2|\nu|}} \exp \nu i\pi\frac{s}{K}.$$

Thus the characteristic function $\frac{1}{c(s)^{t/|a|}}$ is (up to a multiplicative function)

$$(\cos\frac{\pi s}{2K})^{t/|a|} \exp[\frac{t}{|a|} \sum_{\nu\in\mathbb{Z}\setminus\{0\}} \frac{q^{|\nu|} - (-1)^\nu q^{2|\nu|}}{1-q^{2|\nu|}} \exp \nu i\pi\frac{s}{K}]$$

The job is not over : one has to find the values of t such that it is a characteristic function.

This is the case for $t = |a|$ and all the positive multiples of $|a|$. The NEF $G_{|a|}$ with variance function

$$((|a| - m^2)^2 + 4m^2)^{1/2}$$

is generated by a measure $\mu_{|a|}$ which is the convolution of a Bernoulli distribution

$$\frac{1}{2}(\delta_{-\frac{\pi}{2K}} + \delta_{\frac{\pi}{2K}})$$

by an infinitely divisible distribution $\alpha_{|a|}$ concentrated on $\frac{\pi}{K}\mathbb{Z}$. For any positive integer ν we denote

$$c_\nu = c_{-\nu} = \frac{q^\nu - (-1)^\nu q^{2\nu}}{1 - q^{2\nu}} > 0.$$

The Laplace transform of α_t is

$$\int_{-\infty}^{\infty} e^{\theta x} \alpha_t(dx) = \exp\left(\frac{t}{|a|} \sum_{\nu \in \mathbb{Z}\setminus\{0\}} c_\nu (e^{\frac{\nu\pi\theta}{K}} - 1)\right).$$

Since a Bernoulli factor is present, the characteristic function has zeros and cannot be associated to an infinitely divisible distribution. The Jorgensen set of $\mu_{|a|}$ is unknown.

Résumé

Theorem. Let $0 \leq p < 1$ and

$$K = \int_0^1 (1-x^2)^{-1/2}(2-p-x^2)^{-1/2}dx$$
$$K' = \int_0^1 (1-x^2)^{-1/2}(1+(1+p)x^2)^{-1/2}dx.$$

There exists a natural exponential family G_t with domain of the means $\mathbb{R}$ and variance function

$$m \mapsto t\sqrt{1+2p\frac{m^2}{t^2}+(2-p)^2\frac{m^4}{t^4}}$$

when t is a multiple of $2-p$. It is concentrated on $\frac{\pi}{2K}\mathbb{Z}$. The family G_{2-p} is generated by a probability measure

$$\mu_{2-p} = \frac{1}{2}(\delta_{-\frac{\pi}{2K}} + \delta_{\frac{\pi}{2K}}) * \alpha_{2-p}$$

where the infinitely divisible distribution α_{2-p} is concentrated on $\frac{\pi}{K}\mathbb{Z}$. We denote $q = e^{-\pi K'/K}$ and for a positive integer ν we denote

$$c_\nu = c_{-\nu} = \frac{q^\nu - (-1)^\nu q^{2\nu}}{1-q^{2\nu}} > 0.$$

Then the Laplace transform of α_t is

$$\int_{-\infty}^{\infty} e^{\theta x}\alpha_t(dx) = \exp\left(\frac{t}{2-p}\sum_{\nu\in\mathbb{Z}\setminus\{0\}} c_\nu\left(e^{\frac{\nu\pi\theta}{K}}-1\right)\right).$$

Finally the characteristic function of μ_{4-2p} is $\frac{1}{\wp(s+K)-\frac{p}{3}}$ where $\wp$ is the elliptic Weierstrass function satisfying

$$\wp'^2(s) = 4\left(\wp(s)-1+\frac{2p}{3}\right)\left(\wp(s)-\frac{p}{3}\right)\left(\wp(s)+1-\frac{p}{3}\right)$$

which is doubly periodic with primitive periods $2K$ and $2iK'$. In particular it has zeros and G_t cannot be infinitely divisible.

$$\begin{array}{ccccccccccccccccccc}
 & -3K/2 & & & & -K/2 & & & & K/2 & & & & 3K/2 & & & & 5K/2 & \\
- & \bullet & - & \circ & - & \bullet & - & \circ & - & \bullet & - & \circ & - & \bullet & - & \circ & - & \bullet & -
\end{array}$$

Supplementary material : the dual of G_t.

It is worthwhile to recall what is the dual of a real NEF. Given a variance function V defined on the domain of the means M we denote $M^+ = M \cap (0, \infty)$ and $M_1^+ = \{m\,;\, 1/m \in M^+\}$. We say that the NEF F_1 is the dual of F if its domain of the means M_1 satisfies $M_1^+ = M_1 \cap (0, \infty)$ and if its variance function V_1 satisfies $V_1(m) = m^3 V(1/m)$ for $m \in M_1^+$. We do not comment on the challenging problem of the probabilistic interpretation of the duality by stopping times and Zolotarev formula .

The next theorem describes the dual of the NEF G_{2-p}

Theorem 2. Let $p \in [0,1)$. Let $x > 0$. The NEF F_x with domain of the means $(0,\infty)$ and variance function

$$V_{F_x}(m) = m[1 + 2p\frac{m^2}{x^2} + \frac{(2-p)^2 m^4}{x^4}]^{1/2}$$

is generated by a positive measure on $\mathbb{N}$ with generating function $f_x(z)$ given by

$$\exp[x\sqrt{\frac{2}{2-p}}\int_0^z (1+qw^2)^{1/2}(1-w^4)^{-1/2}dw]. \tag{4}$$

where $q = p/(2-p)^2$.

Proof. It is convenient to denote $c = p/(2-p)$ and to observe that

$$0 \le c < 1 \ , \qquad \sqrt{1-c^2} = \frac{2\sqrt{1-p}}{2-p} \ ,$$

$$\frac{c \pm 1}{\sqrt{1-c^2}} = \pm \frac{1}{\sqrt{1-p}}.$$

We use successively the change of variables $u = (2-p)m^2/x^2$ and $u = \sqrt{1-c^2}\,\sinh v - c$.

$$\begin{aligned} d\theta &= \frac{dm}{V(m)} = \frac{4mdm}{4m^2\sqrt{1+2p\frac{m^2}{x^2}+\frac{(2-p)^2m^4}{x^4}}} \\ &= \frac{du}{u\sqrt{1+2cu+u^2}} = \frac{dv}{2\sqrt{1-c^2}\,\sinh v - 2c} \\ &= \frac{e^v dv}{\sqrt{1-c^2}e^{2v} - 2ce^v - \sqrt{1-c^2}} \\ &= \frac{1}{2}\left[\frac{1}{e^v - \frac{1}{\sqrt{1-p}}} - \frac{1}{e^v + \frac{1}{\sqrt{1-p}}}\right] e^v dv \end{aligned}$$

Denoting $z = e^\theta$ we get

$$\begin{aligned}
z^2 &= \frac{e^v - \frac{1}{\sqrt{1-p}}}{e^v + \frac{1}{\sqrt{1-p}}} \\
e^v &= \frac{1}{\sqrt{1-p}} \frac{1+z^2}{1-z^2} \\
e^{-v} &= \sqrt{1-p}\frac{1-z^2}{1+z^2} \\
\sqrt{1-c^2}\,\mathrm{sinh}\, v &= \frac{4z^2 + p(1-z^2)^2}{(2-p)(1-z^4)} \\
u &= \frac{2z^2}{1-z^4}(1+qz^2) \\
m^2 &= \frac{2}{2-p}\frac{z^2}{1-z^4}(1+qz^2).
\end{aligned}$$

Thus

$$m = k'_\mu(\theta) = x\sqrt{\frac{2}{2-p}}\frac{(1+qe^{2\theta})^{1/2}}{(1-e^{4\theta})^{1/2}}e^\theta$$

and this leads to the result 4.

It remains to prove that the Taylor expansion of $z \mapsto f_x(z)$ defined by 4 has positive coefficients. For this it is enough to prove that the argument of the exponential

$$z \mapsto \int_0^z (1+qw^2)^{1/2}(1-w^4)^{-1/2}dw$$

has positive coefficients. It is enough to prove that $z \mapsto (1+qz^2)^{1/2}(1-z^4)^{-1/2}$ has positive coefficients. It is enough to prove that $z \mapsto (1+qz)^{1/2}(1-z^2)^{-1/2}$ has positive coefficients. It is enough to prove that

$$z \mapsto \log[(1+qz)(1-z^2)^{-1}] = \sum_{n=1}^{\infty} a_n z^n$$

has positive coefficients. But this very last point is easy to check since $0 \leq q < 1$ and since a_n is computable : for odd n then $a_n = q^n/n > 0$ and for even $n = 2p$ we have

$$a_n = \frac{1}{p} - \frac{q^{2p}}{2p} > 0.$$

The theorem is proved.

The case $p=0$ for F_x

For $p=0$ it is possible to be more specific :
Theorem 3. Let $x>0$. The NEF F_x with domain of the means $(0,\infty)$ and variance function

$$V_{F_x}(m)=m(1+\frac{4m^4}{x^4})^{1/2}$$

is generated by a positive measure on $\mathbb{N}$ which is $\nu_x(dt)=\sum_{n=0}^{\infty}\frac{p_n(x)}{n!}\delta_n(dt)$ with generating function

$$f_x(z)=\sum_{n=0}^{\infty}\frac{p_n(x)}{n!}z^n=e^{x\int_0^z\frac{dw}{(1-w^4)^{1/2}}}. \qquad (5)$$

which satisfies

$$(1-z^4)f_x''(z)-2z^3f_x'(z)-x^2f_x(z)=0.$$

The total mass of ν_x is $\exp(x\frac{1}{4}B(\frac{1}{2},\frac{1}{4}))$. The polynomials p_n are given by $p_n(x)=x^n$ for $n=0,1,2,3,4$, $p_5(x)=x^5+12x$ and for $n\geq 2$

$$p_{n+2}(x)=x^2p_n(x)+n(n-1)^2(n-2)p_{n-2}(x).$$

Proof. For the beginning, do $p = 0$ in Theorem 2. The trick to obtain the differential equation for f_x is to write $(1 - z^4)^{1/2} f_x' = x f_x$, then to differentiate with respect to z and then to multiply both sides of the result by $(1 - z^4)^{1/2}$. The remainder is standard calculus on power series.