A refined saddle point approximation

THOMAS HÖGLUND

We consider high convolution powers

$$f^{n*}(x) = \sum_{x_1+\ldots+x_n=x} f(x_1)\ldots f(x_n)$$

of an arbitrary function $f\colon Z \to [0, \infty)$ with *finite* support $S = \{x \in Z | f(x) > 0\}$, and our aim is to obtain an approximation of $f^{n*}(x)$ which is sharp for *all* $x \in Z$.

It will be easy to see that our approximation below (properly interpreted) holds when S is a one point set. To avoid trivialities we exclude this case from the remainder of the text. Let $\underline{x}$ and $\bar{x}$ stand for the smallest and largest point of S, respectively, and let I denote the convex hull of S considered as a subset of R, $\underline{x} = \min S$, $\bar{x} = \max S$, $I = [\underline{x}, \bar{x}]$.

In order to be able to give the announced approximation we have to *require that the support of* f^{n*} *is a convex subset of* Z *for all* n *sufficiently large. This is the case if and only if both* $\underline{x} + 1$ *and* $\bar{x} - 1$ *belong to* S. For if $\bar{x} - 1$, say, does not belong to S then $n\bar{x} - 1 \notin \operatorname{supp} f^{n*}$ for all n, and hence the support of f^{n*} never becomes a convex subset of Z. (Note that the integer interval $[\underline{x}n, \bar{x}n]$ is the convex hull of $\operatorname{supp} f^{n*}$ considered as a subset of Z.) Conversely, if $\underline{x} + 1$ and $\bar{x} - 1$ belong to S then $\operatorname{supp} f^{n*}$ contains all points of form

$$k_1\underline{x} + k_2(\underline{x} + 1) + k_3(\bar{x} - 1) + k_4\bar{x} = n\underline{x} + (k_3 + k_4)(\bar{x} - \underline{x} - 1) + k_2 + k_4,$$

with $k_1, \ldots, k_4$ non-negative integers satisfying $k_1 + \ldots + k_4 = n$. That is, $\operatorname{supp} f^{n*}$ contains all points of form $n\underline{x} + h_1(\bar{x} - \underline{x} - 1) + h_2$, with $0 \le h_1 \le n$, $0 \le h_2 \le n$, and hence $\operatorname{supp} f^{n*} = [n\underline{x}, n\bar{x}]$ for all $n \ge \max(1, \bar{x} - \underline{x} - 2)$.

Let μ stand for the measure on Z which assigns the weight $f(x)$ to the point x,

$$\mu(E) = \sum_{x \in E} f(x).$$

The approximation will be formulated in terms of quantities which are naturally tied to the family of probability distributions whose densities relative to the measure μ consist of the closure of the exponential family $p_a(x) = e^{ax}/\varphi(a)$, $a \in R$. Here

$\varphi(a) = \sum_x e^{ax} f(x)$ is the Laplace transform of f. To complete the family we thus have to add the two probability distributions given by the densities

$$p_{-\infty}(x) = \lim_{a \to -\infty} p_a(x) = \delta_{\underline{x}x}/f(\underline{x}) \quad \text{and} \quad p_{+\infty}(x) = \lim_{a \to +\infty} p_a(x) = \delta_{\bar{x}x}/f(\bar{x})$$

(the Kronecker delta).

The above-mentioned quantities are among the following ones. The *meanvalue*

$$m_a = \int x p_a(x) d\mu(x) = \frac{d}{da} \log \varphi(a),$$

the *variance*

$$v_a = \int (x - m_a)^2 p_a(x) d\mu(x) = \frac{d^2}{da^2} \log \varphi(a),$$

and the *entropy*

$$H_a = - \int p_a(x) \log p_a(x) d\mu(x) = \log \varphi(a) - a m_a.$$

If we note that $m_{-\infty} = \underline{x}$, $m_{+\infty} = \bar{x}$ and $dm_a/da = v_a > 0$ for $a \in R$, we see that the meanvalue maps $\bar{R}$ onto I in a one to one manner. The fourth quantity we will need is the *maximum likelihood estimator*, $\hat{a} = m^{-1}$, which is a one to one mapping of I onto $\bar{R}$. Its name comes from the identity

$$p_{\hat{a}(x)}(x) = \max_{a \in \bar{R}} p_a(x).$$

Let us finally introduce the analytic function

$$\varrho(\lambda) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp [\lambda(e^{i\alpha} - 1 - i\alpha)] d\alpha.$$

The function $\varrho(\lambda), \lambda \geq 0$, *is strictly positive and satisfies*

$$\varrho(0) = 1 \geq \varrho(\lambda) = (2\pi\lambda)^{-1/2}(1 + O(1/\lambda)), \quad \text{as} \quad \lambda \to \infty.$$

This will be proved at the end of the paper.

It is clear that $f^{n*}(x) = 0$ when $x/n \notin I$, and that $f^{n*}(x) > 0$ when $x/n \in I$, provided $n \geq \max(1, \bar{x} - \underline{x} - 2)$. The local central limit theorem says

$$f^{n*}(x) = e^{nH_0}(2\pi n v_0)^{-\frac{1}{2}}(\exp[-\tfrac{1}{2}n(x/n - m_0)^2/v_0] + O(n^{-\frac{1}{2}}))$$

uniformly in $x \in Z$, and hence it tells us nothing more than $f^{n*}(x) = O(e^{nH_0}/n)$ except when x/n belongs to a subinterval of I of length $O(\sqrt{\log n/n})$ (centred around m_0). Richter [3] gave an approximation which holds when x/n belongs to a subinterval of length $o(1)$. But there are still better results. So-called saddle point approximations have long since been used in statistical mechanics. A rigorous result of that kind was given by Martin-Löf [2]:

$$f^{n*}(x) = \exp\,[nH_{\hat{a}(x/n)}](2\pi n v_{\hat{a}(x/n)})^{-\frac{1}{2}}(1 + O(1/n)),$$

uniformly in $x \in Z$ as x/n is within but stays away from the boundary of I. It is also clear that

$$f^{n*}(x) = f(x/n)^n = \exp\,[nH_{\hat{a}(x/n)}]$$

when x/n belongs to the boundary of I.

The remaining gap is filled in by *the refined saddle point approximation.*

THEOREM. *As* $n \to \infty$,

$$f^{n*}(x) = \exp\,[nH_{\hat{a}(x/n)}]\varrho(nv_{\hat{a}(x/n)})(1 + O(1/n))$$

uniformly for $x/n \in I$, *and* $f^{n*}(x) = 0$ *when* $x/n \notin I$.

It will follow from the details below that the statement above is still true if we replace the error $O(1/n)$ by $O(\min\,(nv^2_{\hat{a}(x/n)}, 1/n))$.

The proof starts with the identity

$$f^{n*}(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-(a+i\alpha)x}\varphi(a + i\alpha)^n d\alpha.$$

We put $a = \hat{a}(x/n)$ and conclude

$$f^{n*}(x) = \exp\,[nH_{\hat{a}(x/n)}]\,\frac{1}{2\pi}\int_{-\pi}^{\pi}\gamma_{\hat{a}(x/n)}(\alpha)^n d\alpha,$$

where

$$\gamma_a(\alpha) = \int_Z e^{i\alpha(x-m_a)}p_a(x)d\mu(x).$$

We have to show that

$$D = \left|\frac{1}{2\pi}\int_{-\pi}^{\pi}\gamma_a(\alpha)^n d\alpha - \varrho(nv_a)\right| = O(\varrho(nv_a)/n),$$

uniformly in a for $-\infty \leq a \leq \infty$ and a of the form $a = \hat{a}(x/n)$. The proof will be separated into two cases: $0 \leq a \leq \infty$ and $-\infty \leq a \leq 0$, and the second case will be omitted since it is quite analogous to the first one. The word constant (in formulas written Const.) will be used for numbers which do not depend on a, α, or n.

Let

$$c_k(a) = \int_Z (x - m_a)^k p_a(x) d\mu(x).$$

Then $c_2(a) = v_a$ and

$$\gamma_a(\alpha) = 1 + \sum_{k=2}^{\infty} \frac{(i\alpha)^k}{k!} c_k(a).$$

LEMMA 1. *$|c_k(a) - (-1)^k e^{-a} f(\bar{x} - 1)/f(\bar{x})| \leq C^k e^{-2a}$, $k = 2, 3, \ldots$ for some constant $C > 1$ and all $a \geq 0$.*

This and the lemma below will be proved at the end of the paper. Put

$$r_a(\alpha) = \sum_{k=3}^{\infty} \frac{(i\alpha)^k}{k!} (c_k(a) - (-1)^k v_a).$$

It follows from the lemma that

$$|c_k(a) - (-1)^k v_a| \leq 2C^k e^{-2a} \leq \text{Const. } C^k v_a^2,$$

and hence also that $|r_a(\alpha)| \leq \text{Const. } v_a^2 |\alpha|^3$, for $|\alpha| \leq \pi$. This motivates the decomposition

$$\gamma_a(\alpha) = 1 + v_a(e^{-i\alpha} - 1 + i\alpha) + r_a(\alpha).$$

Note that $r_a(\alpha)$ is negligible compared to the other terms not only when v_a is small but also otherwise provided $|\alpha|$ is small. Small α give the main contribution to $\int_{-\pi}^{\pi} \gamma_a(\alpha)^n d\alpha$ in the latter case.

Introduce the abbreviations

$$\gamma = \gamma_a(\alpha), \quad q = v_a(e^{-i\alpha} - 1 + i\alpha), \quad r = r_a(\alpha), \quad s = \sum_{k=4}^{\infty} \frac{(i\alpha)^k}{k!} (c_k(a) - (-1)^k v_a),$$

and $M = \max(|\gamma|, |e^{\gamma-1}|, |e^{q+|r|}|, e^{-v_a\alpha^2/2})$.

Then $\gamma = 1 + q + r$ and

$$\gamma^n - e^{nq} = (\gamma^n - e^{n(\gamma-1)}) + (e^{n(\gamma-1)} - e^{nq}(1 + nr))$$
$$+ nr(e^{nq} - e^{-nv_a\alpha^2/2}) + nse^{-nv_a\alpha^2/2} + n(r - s)e^{-nv_a\alpha^2/2}.$$

(This complicated decomposition is unnecessary if we are content with the error $O(n^{-\frac{1}{2}})$ instead of $O(1/n)$.) The inequalities

$$|x^n - y^n| \leq n|x - y| (\max(|x|, |y|))^{n-1} \quad \text{and} \quad |e^x - 1 - x| \leq |x|^2 e^{|x|}/2$$

together with the familiar estimate $|\gamma - 1| \leq v_a\alpha^2/2$ yields

$$|\gamma^n - e^{n(\gamma-1)}|$$
$$\leq nM^{n-1}|\gamma - e^{\gamma-1}| \leq nM^{n-1}|\gamma - 1|^2 e^{\gamma-1}/2 \leq \text{Const. } M^n(nv_a\alpha^2)^2/n \quad \textit{for} \quad |\alpha| \leq \pi.$$

(Note that $\sup_a v_a < \infty$, and hence also $1/M \leq \exp v_a\alpha^2/2 \leq \text{Const.}$) In a similar way we obtain

$$|e^{n(\gamma-1)} - e^{nq}(1 + nr)| \leq \text{Const } M^n(nv_a\alpha^2)^3/n$$

and

$$|nr(e^{nq} - \exp(- nv_a\alpha^2/2))| \leq \text{Const. } M^n(nv_a\alpha^2)^3/n \quad \textit{for} \quad |\alpha| \leq \pi.$$

(Here we used the inequality $|e^{i\alpha} - 1 - i\alpha - (i\alpha)^2/2| \leq |\alpha|^3/6$.) Finally

$$|ns \exp(- nv_a\alpha^2/2)| \leq \text{Const. } M^n(nv_a\alpha^2)^2/n.$$

(The estimate $|s| \leq \text{Const. } v_a^2|\alpha|^4$ is a consequence of lemma 1.)

LEMMA 2. *There are positive constants* ε *and* δ *such that* $M \leq \exp(- \varepsilon\alpha^2 v_a)$ *for all* $- \infty \leq a \leq \infty$ *and* $- \pi \leq \alpha \leq \pi$ *satisfying* $v_a|\alpha| \leq \delta$.

We use the fact $\int_{-\pi}^{\pi} (r - s) \exp(- nv_a\alpha^2/2)d\alpha = 0$, make the substitution $\alpha \to - \alpha$ in the integral representation of $\varrho(\lambda)$, and split the domains of integration in the expression D into two parts: $|\alpha| \leq \min(\pi, \delta/v_a)$ and $\delta/v_a < |\alpha| \leq \pi$. The result is

$$D \leq \text{Const. } g(nv_a)/n + \int\limits_{\delta/v_a \leq |\alpha| \leq \pi} (|\gamma_a(\alpha)|^n + |\exp(e^{i\alpha} - 1 - i\alpha)|^{nv_a})d\alpha,$$

where

$$g(\lambda) = \int\limits_{|\alpha| \leq \pi} e^{-\varepsilon\lambda\alpha^2}[(\lambda\alpha^2)^2 + (\lambda\alpha^2)^3]d\alpha.$$

It is clear that $g(\lambda) \leq \text{Const. } (\lambda^2 + \lambda^3), \lambda \geq 0$, and the substitution $\alpha \to \alpha\lambda^{-\frac{1}{2}}$ shows that $g(\lambda) \leq \text{Const. } \lambda^{-\frac{1}{2}}$. Hence $g(\lambda) \leq \text{Const.} \min(1, \lambda^2)\varrho(\lambda)$ for all $\lambda \geq 0$.

The second term in the sum dominating D is non-zero only if the domain of integration is non-empty, i.e. only when $v_a > \delta/\pi$. According to lemma 1 this implies that a belongs to a compact subset, K, of R. We also know that $\eta = \inf_a \delta/v_a > 0$, and it is wellknown that $|\gamma_a(\alpha)| < 1$ for all $a \in R$ and $0 < |\alpha| < \pi$. (See lemma 3, p. 475 of Feller [1].) It is easy to see that the function $(a, \alpha) \to \gamma_a(\alpha)$ is continuous,

$$\left(|\gamma_a(\alpha) - \gamma_b(\beta)| \leq \int |p_a(x) - p_b(x)|d\mu(x) + |\alpha - \beta|\right),$$

and hence also

$$\sup_{\substack{a \in K \\ \eta \leq |\alpha| \leq \pi}} |\gamma_a(\alpha)| < 1.$$

The estimate

$$|e^q| = e^{\operatorname{Re} q} = \exp(-v_a(1-\cos\alpha)) \le e^{-\delta(1-\cos\alpha)/\pi}$$

(valid when $\delta/v_a < |\alpha| \le \pi$) finally shows that there is a $\zeta > 0$ such that the integral in question is dominated by

$$\text{Const. } e^{-\zeta n} \le \text{Const. } e^{-\zeta' n v_a} \le \text{Const.} \min(1, (nv_a)^2)\varrho(nv_a)/n, \quad \textit{where} \quad \zeta' < \zeta/\max_a v_a.$$

(Remember that $v_a \ge \delta/\pi$.)

Proof of lemma 1. We point out the fact that

$$1 \ge p_a(x) = e^{-a(\bar{x}-x)}f(x)/f(\bar{x})(1 + O(e^{-a}))$$

as $a \to \infty$. It is clear that $m_a = \bar{x} + O(e^{-a})$. Also

$$|c_k(a) - (-1)^k e^{-a} f(\bar{x}-1)/f(\bar{x})| \le |(\bar{x} - m_a)^k p_a(\bar{x})| +$$

$$+ |(\bar{x} - 1 - m_a)^k p_a(\bar{x}-1) - (-1)^k e^{-a} f(\bar{x}-1)/f(\bar{x})| + |\sum_{x=\underline{x}}^{\bar{x}-2} (x - m_a)^k p_a(x)|.$$

The first expression to the right is dominated by $(\text{Const. } e^{-a})^k$, the third by $(\text{Const.})^k e^{-2a}$, and the second by

$$|(\bar{x} - 1 - m_a)^k - (-1)^k| p_a(\bar{x}-1) + \text{Const. } e^{-2a} \le$$

$$\le (\text{Const.})^k\, |\bar{x} - m_a| e^{-a} + \text{Const. } e^{-2a} \le (\text{Const.})^k\, e^{-2a}.$$

The lemma follows.

Proof of lemma 2. We have

$$\operatorname{Re}(\gamma_a(\alpha) - 1) = \operatorname{Re} v_a(e^{-i\alpha} - 1 + i\alpha) + \operatorname{Re} r_a(\alpha) \le -v_a(1-\cos\alpha) + |r_a(\alpha)| \le$$

$$\le -v_a\alpha^2 2/\pi^2 + \text{Const. } v_a^2|\alpha|^3 \le -v_a\alpha^2/10,$$

provided $|\alpha| \le \pi$ and $v_a|\alpha|$ is sufficiently small. Here we have used the inequality $1 - \cos\alpha \ge \alpha^2 2/\pi^2$, valid for $|\alpha| \le \pi$. These estimates take care of $|e^{\gamma-1}|$ and $|e^{q+|r|}|$.

In order to estimate $|\gamma_a(\alpha)|$ we note that $|\log(1+z) - z| \le |z|^2$ for $|z| \le \frac{1}{2}$. But $|\gamma_a(\alpha) - 1| \le v_a\alpha^2/2 \le \frac{1}{2}$ for $|\alpha| \le \pi$ and $v_a|\alpha| \le 1/\pi$. Hence

$$|\gamma_a(\alpha)| = |e^{\log(1+\gamma-1)}| \le |\exp(\gamma - 1 + v_a^2\alpha^4/4)| \le$$

$$\exp(-v_a\alpha^2/10 + v_a^2\alpha^4/4) \le \exp(-v_a\alpha^2/20),$$

provided $|\alpha| \le \pi$ and $v_a|\alpha|$ is sufficiently small.

The function $\varrho(\lambda)$. It is clear that $|\varrho(\lambda)| \le \varrho(0) = 1$ for $\lambda \ge 0$. In order to show that $\varrho(\lambda) > 0$ for $\lambda > 0$, we note that

$$\varrho(\lambda) = e^{-\lambda}\sum_{k=0}^{\infty} \frac{\lambda^k}{k!}\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i\alpha(k-\lambda)}d\alpha.$$

Hence

$$\varrho(\lambda) = e^{-\lambda} \frac{\lambda^\lambda}{\lambda!} > 0$$

for $\lambda = 0, 1, 2, \ldots,$ and

$$\varrho(\lambda) = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \frac{\sin \pi(\lambda - k)}{\pi(\lambda - k)}$$

otherwise. But $\sin \pi(\lambda - k) = (-1)^k \sin \pi\lambda$, and $\operatorname{sgn}(\sin \pi\lambda) = (-1)^{[\lambda]}$ for $\lambda \in R - Z$. It therefore follows from lemma 3 that $\varrho(\lambda) > 0$ also for $0 < \lambda \notin Z$.

LEMMA 3. *Put*

$$S(\lambda) = \sum_{k=0}^{\infty} \frac{(-\lambda)^k}{k!} \frac{1}{\lambda - k}, \quad \lambda \in R - Z.$$

Then $S(\lambda) > 0$ *if* $[\lambda]$ *is even and non-negative, and* $S(\lambda) < 0$ *otherwise.*

Proof. Observe that

$$S(\lambda) = \sum_{k=0}^{\infty} \left(\frac{(-\lambda)^{2k}}{(2k)!} \frac{1}{\lambda - 2k} + \frac{(-\lambda)^{2k+1}}{(2k+1)!} \frac{1}{\lambda - 2k - 1} \right) =$$
$$= - \sum_{k=0}^{\infty} \frac{\lambda^{2k}}{(2k)!} \left(\frac{1}{2k+1} + \frac{1}{(\lambda - 2k)(\lambda - 2k - 1)} \right),$$

and that $(\lambda - 2k)(\lambda - 2k - 1) > 0$ if and only if $\lambda - 2k < 0$ or $\lambda - 2k > 1$. Hence $S(\lambda) < 0$ if $(\lambda - 2k)(\lambda - 2k - 1) > 0$ for all $k = 0, 1, 2, \ldots,$ i.e. if $\lambda < 0$ or $[\lambda]$ is odd.

Similarly

$$S(\lambda) = \frac{1}{\lambda} + \sum_{k=0}^{\infty} \left(\frac{(-\lambda)^{2k+1}}{(2k+1)!} \frac{1}{\lambda - 2k - 1} + \frac{(-\lambda)^{2k+2}}{(2k+2)!} \frac{1}{\lambda - 2k - 2} \right) =$$
$$= \frac{1}{\lambda} + \sum_{k=0}^{\infty} \frac{\lambda^{2k+1}}{(2k+1)!} \left(\frac{1}{2k+2} + \frac{1}{(\lambda - 2k - 1)(\lambda - 2k - 2)} \right),$$

and hence $S(\lambda) > 0$ if $[\lambda]$ is even and non-negative. The lemma follows.

Since

$$(2\pi\lambda)^{-\frac{1}{2}} = \frac{1}{2\pi} \int_R e^{-\frac{1}{2}\lambda\alpha^2} (1 + \lambda(i\alpha)^3/6) d\alpha$$

it remains to show that the difference between this integral and the integral defining $\varrho(\lambda)$ is $O(\lambda^{-3/2})$ as $\lambda \to \infty$. It should be clear that the contribution to these integrals from values of α satisfying $|\alpha| > 1$ is less than $O(\lambda^{-3/2})$. Put $s_k(\alpha) = e^{i\alpha} - \sum_{j=0}^{k-1} (i\alpha)^j/j!$. Then $|s_k(\alpha)| \leq |\alpha|^k/k!$, and

$$|\varrho(\lambda) - (2\pi\lambda)^{-\frac{1}{2}}| = \left| \frac{1}{2\pi} \int_{|\alpha|<1} e^{-\frac{1}{2}\lambda\alpha^2}(e^{\lambda s_3(\alpha)} - 1 - \lambda(i\alpha)^3/6)d\alpha \right| + O(\lambda^{-3/2}).$$

The inequality $|e^x - 1 - y| \leq |e^x - 1 - x| + |x - y| \leq |x|^2 e^{|x|}/2 + |x - y|$, applied to $x = \lambda s_3(\alpha)$ and $y = \lambda(i\alpha)^3/6$, shows that the integrand to the right is dominated by

$$[(\lambda\alpha^2)^3 e^{-\lambda\alpha^2/3} + (\lambda\alpha^2)^2 e^{-\lambda\alpha^2/2}]/\lambda$$

provided $|\alpha| < 1$. Hence $\varrho(\lambda) = (2\pi\lambda)^{-\frac{1}{2}}(1 + O(1/\lambda))$.

References

1. Feller, W., *An introduction to probability theory and its applications*, Vol. 2. Wiley, New York, 1966.
2. Martin-Löf, P., *Statistiska modeller, anteckningar från seminarier läsåret 1969—70*, utarbetade av R. Sundberg (in Swedish). University of Stockholm, Department of Mathematical statistics.
3. Richter, W., Local limit theorems for large deviations. *Theor. Probability Appl.*, 2 (1957), 206—220.

Received September 3, 1973

Thomas Höglund
Institut for Matematisk Statistik
H. C. Ørsted Institut
Universitetsparken 5
DK-2100 København
Danmark