Local Limit Theorem

State of Affairs

Owen Mireles

August 26, 2020

In Appendix A of [1], R. Hough stated the following:
For $k > 1$, recall that we define the measure on $\mathbb{Z}^k$,

$$\nu_k = \frac{1}{2k+1}\left(\delta_0 + \sum_{j=1}^{k}(\delta_{e_j} + \delta_{e_{-j}})\right)$$

and that we write

$$\eta_k(\sigma, x) = \frac{1}{(2\pi\sigma^2)^{k/2}}\exp\left(-\frac{||x||_2^2}{2\sigma^2}\right)$$

for the density of the centered standard normal distribution on $\mathbb{R}^k$. In this appendix, we prove

Lemma 1. *Let $n, k(n) \geq 1$ with $k^2 = o(n)$ for large n. As $n \to \infty$ we have*

$$\left\|\nu_k^{*n} * 1_{\left[-\frac{1}{2},\frac{1}{2}\right)^k} - \eta_k\left(\sqrt{\frac{2n}{2k+1}}, \cdot\right)\right\|_{\mathrm{TV}(\mathbb{R}^k)} = o(1).$$

Actually, a stronger estimate is proven, which is a local limit theorem on $\mathbb{R}^k$.

Lemma 2. *Let $n, k(n) \geq 1$ with $k^2 = o(n)$ for large n. Uniformly for $\alpha \in \mathbb{Z}^k$ such that $||\alpha||_2^2 \leq \frac{2kn}{2k+1} + \frac{n\log n}{\sqrt{k}}$, and $||\alpha||_4^4 \ll \frac{n^2}{k}\left(1 + \frac{\log n}{\sqrt{k}}\right)$, as $n \to \infty$,*

$$\nu_k^{*n}(\alpha) = \{1 + o(1)\}\eta_k\left(\sqrt{\frac{2n}{2k+1}}, \alpha\right).$$

The goal of an ongoing research project is to prove this result for other, more general measures on $\mathbb{Z}^k$. Understanding the proof of these two lemmas seems to be a natural starting point on how to achieve this. We do a quick review of the steps taken.

The easier part is showing that Lemma 2 implies Lemma 1.

Lemma 2 implies Lemma 1. Both $\nu_k^{*n} * 1_{\left[-\frac{1}{2},\frac{1}{2}\right)^k}$ and $\eta_k\left(\sqrt{\frac{2n}{2k+1}}, \cdot\right)$ are measures that concentrate at the origin. So it suffices to estimate the difference

$$\nu_k^{*n} * 1_{\left[-\frac{1}{2},\frac{1}{2}\right)^k}(x) - \eta_k\left(\sqrt{\frac{2n}{2k+1}}, x\right)$$

for $\|x\|_2^2 \leq \frac{2kn}{2k+1} + O\left(\right)$.

For $x \in \mathbb{Z}^k$ satisfying this upper bound and for $y \in \left[-\frac{1}{2}, \frac{1}{2}\right)^k$, it is not hard to see that

$$\eta_k\left(\sqrt{\frac{2n}{2k+1}}, x+y\right) = (1+o(1))\eta_k\left(\sqrt{\frac{2n}{2k+1}}, x\right)\exp\left(-\frac{(2k+1)x\cdot y}{2n}\right).$$

Thus,

$$\begin{aligned}\int_{\left[-\frac{1}{2},\frac{1}{2}\right)^k} &\left|\eta_k\left(\sqrt{\frac{2n}{2k+1}}, x+y\right) - \eta_k\left(\sqrt{\frac{2n}{2k+1}}, x\right)\right| dy \\ &= \eta_k\left(\sqrt{\frac{2n}{2k+1}}, x\right)\left(o(1) + (1+o(1))\int_{\left[-\frac{1}{2},\frac{1}{2}\right)^k}\left|\exp\left(-\frac{(2k+1)x\cdot y}{2n}\right) - 1\right| dy\right).\end{aligned}$$

For $\|x\|_2^2 \ll \frac{2kn}{2k+1} + \frac{n\log n}{\sqrt{k}}$, using concentration of the measures and Azuma's inequality, it can be shown that

$$\int_{\left[-\frac{1}{2},\frac{1}{2}\right)^k}\left|\exp\left(-\frac{(2k+1)x\cdot y}{2n}\right) - 1\right| dy = o(1).$$

Let

$$\mathscr{B} = \left\{x \in \mathbb{Z}^k : \|x\|_2^2 \leq \frac{2kn}{2k+1} + \frac{n\log n}{\sqrt{k}}, \|x\|_4^4 \leq C\frac{n^2}{k}\left(1 + \frac{\log n}{\sqrt{k}}\right)\right\}.$$

Using the above estimates we get

$$\begin{aligned}\sum_{x\in\mathscr{B}}\int_{\left[-\frac{1}{2},\frac{1}{2}\right)^k} &\left|\eta_k\left(\sqrt{\frac{2n}{2k+1}}, x+y\right) - \eta_k\left(\sqrt{\frac{2n}{2k+1}}, x\right)\right| dy \\ &\leq o(1)\sum_{x\in\mathscr{B}}\eta_k\left(\sqrt{\frac{2n}{2k+1}}, x\right) \\ &= o(1)\sum_{x\in\mathscr{B}}(1+o(1))\nu_k^{*n}(x) = o(1).\end{aligned}$$

The last line follows from Lemma 2.

Since

$$\sum_{x\in\mathscr{B}}\int_{\left[-\frac{1}{2},\frac{1}{2}\right)^k}\eta_k\left(\sqrt{\frac{2n}{2k+1}}, x+y\right) dy = 1 + o(1)$$

and

$$\sum_{x\in\mathscr{B}}\nu_k^{*n}(x) = 1 + o(1),$$

both quantities can be compared, so

$$\left\|\nu_k^{*n} * 1_{\left[-\frac{1}{2},\frac{1}{2}\right)^k} - \eta_k\left(\sqrt{\frac{2n}{2k+1}}, \cdot\right)\right\|_{\mathrm{TV}(\mathbb{R}^k)}$$

$$= \sum_{x\in\mathbb{Z}^k} \int_{\left[-\frac{1}{2},\frac{1}{2}\right)^k} \left| \eta_k\left(\sqrt{\frac{2n}{2k+1}}, x+y\right) - \eta_k\left(\sqrt{\frac{2n}{2k+1}}, x\right)\right| dy = o(1).$$

The last equality comes from splitting $\mathbb{Z}^k$ into $\mathscr{B}$ and $\mathscr{B}^c$ and summing over each $x \in \mathscr{B}^c$ separately.

□

Remark 1. *From the proof of this, it is safe to assume that if an analogous to Lemma 2 can be proven for more general measures, then Lemma 1 will follow easily.*

Lemma 2. Lemma 2 is the result of a long and detailed computation of the Saddle-Point method.

Let f be the probability generating function of ν_k, that is,

$$f(z_1,\ldots,z_k) = \frac{1}{2k+1}(1 + z_1 + z_1^{-1} + \ldots + z_k + z_k^{-1}) = \mathrm{E}(z_1^{X_1}\cdots z_k^{X_k}),$$

where $X = (X_1,\ldots,X_k)$ has the law of ν_k. We have $\nu_k(\alpha) = C_\alpha[f]$, where

$$C_\alpha[f] = \left(\frac{1}{2\pi i}\right)^k \int_{|z_1|=R_1} \cdots \int_{|z_k|=R_k} \frac{f(z_1,\ldots,z_k)}{z_1^{\alpha_1}\cdots z_k^{\alpha_k}} \frac{dz_1}{z_1} \cdots \frac{dz_k}{z_k}.$$

This is true for positive R_j because of Cauchy's integral formula.

The probability generating function associated to ν_k^{*n} is f^n. By symmetry, we may assume $\alpha \geq 0$ coordinate-wise. Thus,

$$\begin{aligned}\nu_k^{*n}(\alpha) &= \left(\frac{1}{2\pi i}\right)^k \int_{|z_1|=R_1} \cdots \int_{|z_k|=R_k} \frac{f(z_1,\ldots,z_k)^n}{z_1^{\alpha_1}\cdots z_k^{\alpha_k}} \frac{dz_1}{z_1} \cdots \frac{dz_k}{z_k} \\ &= \frac{1}{R_1^{\alpha_1}\cdots R_k^{\alpha_k}} \int_{(\mathbb{R}/\mathbb{Z})^k} f_0(\theta_1,\ldots,\theta_k)^n e(-\alpha\cdot\theta)d\theta,\end{aligned}$$

where $e(y) = \exp(2\pi i y)$ and

$$f_0(\theta_1,\ldots,\theta_k) = f(R_1 e(\theta_1),\ldots,R_k e(\theta_k)).$$

Let $\mathscr{D}_{sm}$ be the domain

$$\mathscr{D}_{sm} = \left\{\theta \in (\mathbb{R}/\mathbb{Z})^k : \|\theta\|_\infty \leq \frac{1}{12}\right\}.$$

In $\mathscr{D}_{sm}$, it is possible to define $F(\theta) = \log \frac{f_0(\theta_1,\ldots,\theta_k)^n e(-\alpha\cdot\theta)}{R_1^{\alpha_1}\cdots R_k^{\alpha_k}}$.

To apply the Saddle-Point method, we want to find the R_j that solve the stationary-phase equation for $F(\theta)$.

In [1], the proof is bolstered by some lemmas which roughly state:

1. Find partial derivatives of $F(\theta)$ up to a large enough order.
2. Find a solution for the stationary phase equation; find the leading terms of the R_j. Express $f_0(0)$ with the correct leading terms and order.
3. Express the correct order of the partial derivatives of F.

4. Find a good enough estimate for $F(\theta) - F(0)$.
5. Bound $f_0(\theta)/f_0(0)$.

Items 1-4 are used to estimate the integral in $\mathscr{D}_{sm}$, while Item 5 is used in $\mathscr{D}_{sm}^c$. Upon inspection of [1], we find that Item 2 of the previous list is the most crucial part of the lemma.

Finding the derivatives is important, but can be done symbolically regardless of which probability generating function we get.

However, for Item 2, we see that there is a strong dependence on the fact that, if $0 \leq a \leq b$ and $R_j \geq 1$, then solving for R_{j_1} in $R_{j_1} + R_{j_1}^{-1} = a$ and for R_{j_2} in $R_{j_2} + R_{j_2}^{-1} = b$ yields $R_{j_1} \leq R_{j_2}$.

□

References

[1] Robert Hough. *Mixing and cut-off in cycle walks.* Electronic Journal of Probability, 22:149, 2017.