

Grenander functionals and Cauchy's formula

Piet Groeneboom

*Delft University of Technology, Building 28, Van Mourik Broekmanweg 6, 2628 XE Delft, The Netherlands.
e-mail: P.Groeneboom@tudelft.nl*

Abstract: Let \hat{f}_n be the nonparametric maximum likelihood estimator of a decreasing density. Grenander characterized this in [6] as the left-continuous slope of the least concave majorant of the empirical distribution function. For a sample from the uniform distribution, the asymptotic distribution of the L_2 -distance of the Grenander estimator to the uniform density was derived in [12] by using a representation of the Grenander estimator in terms of conditioned Poisson and gamma random variables. This representation was also used in [8] to prove a central limit result of Sparre Andersen [19] on the number of jumps of the Grenander estimator. Here we extend this to the proof of the main result in [12] and also prove a similar asymptotic normality results for the entropy functional. In [12] the limit distribution of the sums of gamma and Poisson variables on which the conditioning was done did not have the right form, which is corrected here. Cauchy's formula and saddle point methods are the main tools in our development.

AMS 2000 subject classifications: Primary 62E20, Secondary 62G05.

Keywords and phrases: Grenander estimator, integral statistics, saddle points, Cauchy's formula.

1. Introduction

The Grenander estimator is the (nonparametric) maximum likelihood estimator (MLE) of a monotone decreasing density. It was introduced in [6], where it was proved that it is the left-continuous slope of the least concave majorant of the empirical distribution function. Some properties and limit results are discussed in [10] and also in [11] in the special issue on nonparametric inference under shape constraints of the journal *Statistical Science*. The Grenander estimator is shown in Figure 1 for a sample of size $n = 100$ from the uniform distribution on $[0, 1]$. It can be improved by using boundary penalties (in fact, the estimator is inconsistent at the boundary points 0 and 1), but this is not the concern of the present paper.

The Grenander estimator is a piecewise constant function with downward jumps at locations which correspond to the changes of slope ("kinks") of the least concave majorant of the empirical distribution function. Although the Grenander estimator is defined as the left-continuous slope of the empirical distribution function, we can make the Grenander estimator right-continuous by taking the limits on the right at its points of jump. This does not change the probability mass of the induced (absolutely continuous) probability distribution, which is absolutely continuous w.r.t. Lebesgue measure.

The number of jumps of the Grenander estimator is of order $\log n$ if the sample is from a uniform distribution (see Section 2), if the sample comes from a strictly decreasing smooth density like the exponential density, then the number of jumps is of order $cn^{1/3}$, for some constant $c > 0$. The limit behavior of the Grenander estimator for these situations is rather different. For a sample from the uniform distribution, we have for $t \in (0, 1)$:

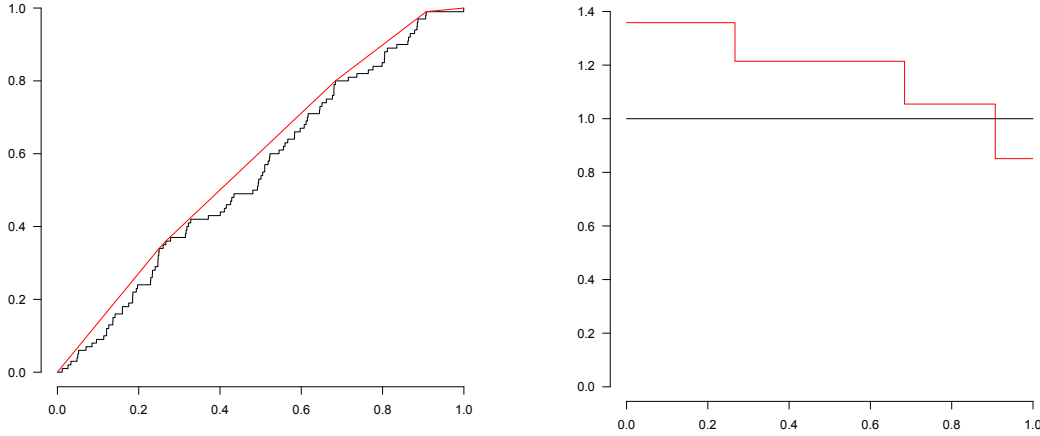
$$\sqrt{n} \left\{ \hat{f}_n(t) - 1 \right\} \xrightarrow{\mathcal{D}} S_t, \quad (1.1)$$

where S_t is the slope of the least concave majorant of the standard Brownian Bridge on $[0, 1]$, see Remark 2.2, p. 543 of [7]. The density of S_t is a function of the standard normal distribution function Φ and the standard normal density ϕ , see (3.11) in [9].

In contrast with the result (1.1), we get for a sample from a decreasing density f on $[0, \infty)$ at a point $t \in (0, \infty)$, where f is differentiable and $f'(t) < 0$, the following result, due to Prakasa Rao in [16]:

$$n^{1/3} |4f(t)f'(y)|^{-1/3} \left\{ \hat{f}_n(t) - f(t) \right\} \xrightarrow{\mathcal{D}} Z, \quad (1.2)$$

*This manuscript is dedicated to the memory of Ronald Pyke



(a) empirical distribution function (black) and its least concave majorant (red)

(b) Grenander estimator

Fig 1: Left: the empirical distribution function and its least concave majorant and right: the Grenander estimator, on the basis of a sample of size $n = 100$ from the uniform distribution on $[0, 1]$

where $Z = \operatorname{argmax}_t \{W(t) - t^2\}$, that is: Z is the (almost surely unique) location of the maximum of two-sided Brownian motion minus the parabola $y(t) = t^2$. For further details, see, e.g., [10] and [11].

Recently, integrated functionals of a monotone density were studied in [14]. Using the same notation as in [14] the following functionals were studied:

$$\mu(h, f) = \int_0^1 h(f(x)) dx,$$

where f is a nonincreasing function on \mathbb{R}_+ and h satisfies some regularity conditions. In the case that the underlying distribution is uniform, the following central limit result is proved:

Theorem 1.1 (Theorem 3.2 in [14]). *Let $h \in C^4([0, \infty))$ and let $h''(1) \neq 0$. Then:*

$$\frac{n \left\{ \mu(h, \hat{f}_n) - \mu(h, f) \right\} - \frac{1}{2} h''(1) \log n}{\sqrt{\frac{3}{4} h''(1)^2 \log n}} \xrightarrow{\mathcal{D}} N(0, 1), \quad (1.3)$$

where $N(0, 1)$ denotes the standard normal distribution.

We prove an analogous result for analytic functions h , defined on the positive open complex half plane. So our functions h have a lot more smoothness, but are on the other hand defined on the *open* complex half plane, which makes the result applicable to functions that are not covered by the conditions in [14]. We assume that h satisfies the following condition.

Condition 1.1. The function h is analytic on the complex half plane $\{z \in \mathbb{C} : \operatorname{Re}(z) > 0\}$, and satisfies the following conditions.

- (i) $h''(1) \neq 0$.
- (ii) For $t \in \mathbb{R}$:

$$|h(1 - it)| = O(t^2), \quad |t| \rightarrow \infty, \quad (1.4)$$

and

(iii) For $t \in \mathbb{R}$:

$$|h''(1 - it)| = O(1), \quad |t| \rightarrow \infty. \quad (1.5)$$

We now have the following result.

Theorem 1.2. *Let h satisfy Condition 1.1 and let \hat{f}_n be the Grenander estimator. Then*

$$\frac{n \left\{ \mu(h, \hat{f}_n) - \mu(h, f) \right\} - \frac{1}{2} h''(1) \log n}{\sqrt{\frac{3}{4} h''(1)^2 \log n}} \xrightarrow{\mathcal{D}} N(0, 1), \quad (1.6)$$

where $N(0, 1)$ denotes the standard normal distribution.

The result is a corollary to Theorem 4.1 in Section 4, see the remark at the end of Section 4.

Examples of the application of Theorem 1.2 are:

Example 1.1. Let $h(z) = (z - 1)^2$. Then $h'(z) = 2(z - 1)$ and $h''(z) = 2$. The function h is obviously analytic on the positive complex half plane. Condition 1.1 is fulfilled, so we get

$$\frac{1}{\sqrt{3 \log n}} \left\{ n \int_0^1 \{ \hat{f}_n(t) - 1 \}^2 dt - \log n \right\} \xrightarrow{\mathcal{D}} N(0, 1), \quad n \rightarrow \infty. \quad (1.7)$$

This is the main result of [12]. Since in [14] Theorem 3.2 is deduced from this result, we also get Theorem 3.2 in [14] back from Theorem 1.2.

Example 1.2. Let $h(z) = z \log z$. Then $h'(z) = 1 + \log(z)$ and $h''(z) = 1/z$. The function h is again analytic on the positive complex half plane. Condition 1.1 is fulfilled, and we get:

$$\frac{1}{\sqrt{\frac{3}{4} \log n}} \left\{ n \int_0^1 \hat{f}_n(t) \log \hat{f}_n(t) dt - \frac{1}{2} \log n \right\} \xrightarrow{\mathcal{D}} N(0, 1), \quad n \rightarrow \infty. \quad (1.8)$$

For this example the conditions of Theorem 3.2 in [14] are not satisfied. The result follows from Theorem 1.2 and can be applied to the theory on a likelihood ratio test for monotonicity in [2].

To derive limit results for the uniform distribution, a special representation in terms of gamma and Poisson random variables was given in [12], with the purpose of proving a limit result for a two-sample rank statistic introduced in the dissertation of [1] and also for a test statistic in the combination of tests in [18]. We describe this representation now.

Let X_1, \dots, X_n be a sample from the uniform distribution, and let $0 = \xi_{n0} < \xi_{n1} < \dots < \xi_{nm} < \xi_{n,m+1} = 1$ be the locations of the jumps of the Grenander estimator \hat{f}_n for this sample, augmented with the points 0 and 1. Note that $[\xi_{n,0}, \xi_{n,1}]$, $(\xi_{n,1}, \xi_{n,2}]$, $(\xi_{n,2}, \xi_{n,3}]$, \dots , $(\xi_{n,m}, 1]$ are the successive intervals of constancy of the Grenander estimator if we take the estimator to be left-continuous.

Furthermore, let D_{ni} , J_{ni} and Q_{nj} be defined by:

$$\begin{aligned} D_{ni} &= \xi_{ni} - \xi_{n,i-1} & , i &= 1, \dots, m+1, \\ J_{ni} &= n \{ \mathbb{F}_n(\xi_{ni}) - \mathbb{F}_n(\xi_{n,i-1}) \} & , i &= 1, \dots, m+1, \\ Q_{nj} &= \# \{ i : J_{ni} = j \}, \end{aligned}$$

where \mathbb{F}_n is the empirical distribution function of the sample X_1, \dots, X_n , and where m is the number of jumps of the Grenander estimator.

Next, let $\{N_j : j \geq 1\}$ be independent Poisson random variables with $\mathbb{E}N_j = 1/j$, and let, for each i , $\{S_{ji}, i, j \geq 1\}$ be a collection of independent gamma random variables, independent of the N_j , where S_{ji} is Gamma($j, 1$) (the sum of j independent standard exponentials). We define:

$$S_n = \sum_{j=1}^n \sum_{i=1}^{N_j} S_{ji}, \quad T_n = \sum_{j=1}^n j N_j. \quad (1.9)$$

and

$$\mathbf{S}^{(n)} = (S_{11}, \dots, S_{1,N_1}, \dots, S_{n1}, \dots, S_{n,N_n}), \quad \mathbf{N}^{(n)} = (N_1, \dots, N_n).$$

Note that there are N_1 induced spacings ξ_{ni} (intervals of constancy of \hat{f}_n) of length 1, N_2 induced spacings ξ_{ni} (intervals of constancy of \hat{f}_n), consisting of two consecutive original spacings between locations of jumps of the least concave majorant, etc., where N_j can be zero.

We now have the following representation theorem:

Theorem 1.3 (Theorem 2.1 in [12]).

$$(nD_{n1}, \dots, nD_{n,m+1}; Q_{n1}, \dots, Q_{n,m}) \stackrel{\mathcal{D}}{=} (\mathbf{S}^{(n)}, \mathbf{N}^{(n)} \mid S_n = n, T_n = n)$$

Remark 1.1. For specificity, the random variables S_{ji} , $i = 1, \dots, N_j$, and D_{ji} , $i = 1, \dots, Q_{nj}$, are ordered in Theorem 2.1 in [12]. There is also a zero-step spacing introduced in [12], but this does not seem to be necessary.

Using this representation, we can reduce the proofs of the limit behavior of global functionals of the Grenander estimator to a theorem for gamma and Poisson random variables, under the condition $(T_n, S_n) = (n, n)$. For convenience in later proofs, we also use a further standardization of (S_n, T_n) :

$$V_n = n^{-1/2} \left\{ S_n - \sum_{j=1}^n j N_j \right\} = n^{-1/2} \{S_n - T_n\}, \quad W_n = \frac{T_n}{n}, \quad (1.10)$$

where S_n and T_n are defined by (1.9). A conditional central limit theorem for functionals of the Grenander estimator can then be proved under the condition:

$$V_n = 0, \quad W_n = 1. \quad (1.11)$$

The infinitely divisible limit distribution of the pair (V_n, W_n) was given in Lemma 3.1 of [12], but unfortunately Lemma 3.1 of [12] contains a rather silly error (the u^2 in the representation of the characteristic function should be u). The correct version of this result is given in Lemma 3.1 below, where also the origin of the error is explained. The proof in [8] does not use the result on the limit distribution of (V_n, W_n) , so is not influenced by the erroneous Lemma 3.1 in [12].

We use methods different from those in [12]. The conditional central limit theorem was proved in [12] using Le Cam's paper [13] (a paper apparently published without his permission, as became clear in a conversation of the author with him). In the present context, where we clearly have to deal with non-standard asymptotics, [13] does not seem to be the right tool to use. We replace this by a direct analysis of the characteristic function. The crucial tools here are Cauchy's formula and saddle point methods, using contour integration in the complex plane. To illustrate our method in a simple setting, we give a shortened version of the proof in [8] of Sparre Andersen's result [19] in Section 2.

2. Sparre Andersen's result

To illustrate our method in the simplest setting, we give a short version of the proof in [8] of the following remarkable result of Sparre Andersen in [19].

Theorem 2.1 (Sparre Andersen's result).

Let X_1, \dots, X_n be a sample from the Uniform(0, 1) distribution and let N_{jumps} be the number of jumps of the Grenander estimator for this sample. Then

$$\frac{N_{jumps} - \log n}{\sqrt{\log n}} \xrightarrow{\mathcal{D}} N(0, 1),$$

where $N(0, 1)$ is the standard normal distribution.

Proof. Let, for the sample X_1, \dots, X_n , U_n be defined by:

$$U_n = \frac{\sum_{j=1}^n N_j - \log n}{\sqrt{\log n}},$$

and let T_n be defined as in (1.9). Using (part of) the representation, introduced in Section 1, we only have to prove that $(U_n | T_n = n)$ tends in law to a standard normal distribution. To this end we consider the conditional characteristic function

$$\mathbb{E} \{ e^{isU_n} \mid T_n = n \}.$$

Lemma 3.2 of [12] implies:

$$\mathbb{P} \{ T_n = n \} = \exp \left\{ - \sum_{j=1}^n \frac{1}{j} \right\}.$$

Hence we get, by Fourier inversion and using the notation $b_n = \sqrt{\log n}$,

$$\begin{aligned} \mathbb{E} \{ e^{isU_n} \mid T_n = n \} &= \exp \left\{ \sum_{j=1}^n \frac{1}{j} \right\} \frac{1}{2\pi} \int_{u=-\pi}^{\pi} \mathbb{E} e^{isU_n + iuT_n - inu} du \\ &= \exp \left\{ \sum_{j=1}^n \frac{1}{j} \right\} \frac{1}{2\pi} \int_{u=-\pi}^{\pi} \exp \left\{ -isb_n + \sum_{j=1}^n \frac{1}{j} \left(e^{iju + is/b_n} - 1 \right) - inu \right\} du \\ &= e^{-isb_n} \frac{1}{2\pi} \int_{u=-\pi}^{\pi} \exp \left\{ e^{is/b_n} \sum_{j=1}^n \frac{1}{j} e^{iju} - inu \right\} du. \end{aligned}$$

Denoting the contour $u \mapsto e^{iu}$, $u \in [-\pi, \pi)$, by C , we write this in the form

$$e^{-isb_n} \frac{1}{2\pi i} \int_C \exp \left\{ \delta_n \sum_{j=1}^n \frac{z^j}{j} \right\} z^{-n-1} dz, \quad (2.1)$$

where δ_n is given by:

$$\delta_n = e^{is/b_n}.$$

The expression in (2.1), multiplied by e^{isb_n} , is by Cauchy's formula equal to the coefficient of z^n in the power series around $z = 0$ of the function:

$$z \mapsto \exp \left\{ \delta_n \sum_{j=1}^n \frac{z^j}{j} \right\}.$$

Comparing this with the power series of the function $z \mapsto (1 - z)^{-\delta_n}$, we see that the coefficient of z^n is the same in both series. This coefficient is:

$$(-1)^n \binom{-\delta_n}{n} = \prod \left(1 + \frac{\delta_n - 1}{j} \right) = \exp \left\{ \sum_{j=1}^n \log \left(1 + \frac{\delta_n - 1}{j} \right) \right\}.$$

So we get:

$$\begin{aligned} \mathbb{E} \{ e^{isU_n} \mid T_n = n \} &= \exp \left\{ -isb_n + \sum_{j=1}^n \log \left(1 + \frac{\delta_n - 1}{j} \right) \right\} \\ &= \exp \left\{ -isb_n + \sum_{j=1}^n \log \left(1 + \frac{e^{is/b_n} - 1}{j} \right) \right\} \\ &= \exp \left\{ -isb_n + \sum_{j=1}^n \frac{e^{is/b_n} - 1}{j} + o(1) \right\} = \exp \left\{ -\frac{s^2}{2b_n^2} \sum_{j=1}^n \frac{1}{j} + o(1) \right\} \\ &= \exp \left\{ -\frac{1}{2}s^2 + o(1) \right\}. \end{aligned}$$

□

3. The limit distribution of the conditioning variables (V_n, W_n)

Let the pair (V_n, W_n) be defined by (1.10). We prove the following lemma, which corrects Lemma 3.1 in [12].

Lemma 3.1. *The pair (V_n, W_n) converges in distribution to (V, W) , where (V, W) has the infinitely divisible characteristic function*

$$\phi_{(V,W)}(t, u) = \exp \left\{ \int_0^1 \frac{e^{-\left(\frac{1}{2}t^2 - iu\right)y} - 1}{y} dy \right\}.$$

Proof. We have:

$$\begin{aligned} \mathbb{E} \exp \{itV_n + iuW_n\} &= \exp \left\{ \sum_{j=1}^n \frac{e^{iu j/n} \phi_{S_{j1}-j}(tn^{-1/2}) - 1}{j} \right\} \\ &= \exp \left\{ \sum_{j=1}^n \frac{e^{ij(u/n - tn^{-1/2})} (1 - itn^{-1/2})^{-j} - 1}{j} \right\}, \end{aligned} \quad (3.1)$$

where $\phi_{S_{j1}-j}$ is the characteristic function of the centered gamma variable $S_{j1} - j$, see (3.9) of [12].

Writing $y_{j,n} = j/n$, and noting that for $y \in (0, 1)$:

$$\frac{e^{iny(u/n - tn^{-1/2})} (1 - itn^{-1/2})^{-ny} - 1}{y} = \frac{e^{iyu - \frac{1}{2}t^2y} - 1}{y} - \frac{1}{3}ie^{-\frac{1}{2}t^2y + iuy}t^3n^{-1/2} + O(n^{-1}),$$

and that the limit of the expression on the left for $y \downarrow 0$ is equal to:

$$i(-t\sqrt{n} + u) - n \log(1 - it/\sqrt{n}) = iu - \frac{1}{2}t^2 + O(n^{-1/2}), \quad n \rightarrow \infty,$$

we can write the exponent in the form

$$\sum_{j=1}^n \frac{e^{-\frac{1}{2}t^2y_{j,n} + iuy_{j,n}} - 1}{ny_{j,n}} + O(n^{-1/2}),$$

(it is here that the mistake was made in [12], in the formula after (3.9) on p. 333), so we get a Riemann sum converging to the integral

$$\int_0^1 \frac{e^{-\frac{1}{2}t^2y + iuy} - 1}{y} dy.$$

The infinite divisibility of the limit distribution is shown below. □

Remark 3.1. In [12] first the limit distribution of U_n is computed, using moment conditions (going up to the 8th moments). Next the limit distribution of (V_n, W_n) is computed and it is stated that this distribution is infinite divisible and has no normal component, implying that therefore the limit (V_n, W_n) has to be independent of the limit of U_n .

The s^2u^2 in the exponent of the characteristic function of the limit distribution of (V_n, W_n) in Lemma 3.1 of [12] should be s^2u . The incorrect u^2 arose on p. 333 of [12], where the limit of the characteristic function of the rescaled gamma random variable $(S_j - j)/\sqrt{n}$ was given by $\exp\{-s^2u^2/2\}$ instead of $\exp\{-s^2u/2\}$. This also invalidates the ensuing remarks on p. 333 of [12]. We correct these remarks below.

The distribution of (V, W) is infinitely divisible, as we now show. A general characterization of infinitely divisible distributions in \mathbb{R}^d is given in [17] and given below for convenience.

Theorem 3.1 (Theorem 1.3 in [17], Lévy-Khintchine representation). *If the distribution μ is infinitely divisible, then its characteristic function $\hat{\mu}(\mathbf{s}) = \int_{\mathbb{R}^d} \exp\{i\langle \mathbf{s}, \mathbf{z} \rangle\} d\mu(\mathbf{z})$ is given by:*

$$\hat{\mu}(\mathbf{z}) = \exp \left\{ -\frac{1}{2} \mathbf{s}^T \mathbf{A} \mathbf{s} + \int_{\mathbb{R}^d} \left(e^{i\langle \mathbf{s}, \mathbf{z} \rangle} - 1 - i\langle \mathbf{s}, \mathbf{z} \rangle 1_{\{\|\mathbf{z}\| \leq 1\}}(\mathbf{z}) \right) d\nu(\mathbf{z}) + i\langle \boldsymbol{\delta}, \mathbf{s} \rangle \right\}, \quad (3.2)$$

where \mathbf{A} is a symmetric nonnegative-definite $d \times d$ matrix, $\|\cdot\|$ is the Euclidean norm, ν is a measure on \mathbb{R}^d satisfying $\nu(\{0\}) = 0$, $\int_{\mathbb{R}^d} (\|\mathbf{z}\|^2 \wedge 1) d\nu(\mathbf{z}) < \infty$, and where $\boldsymbol{\delta} \in \mathbb{R}^d$. The representation (3.2) by \mathbf{A} , ν and $\boldsymbol{\delta}$ is unique. Conversely, for any choice of \mathbf{A} , ν and $\boldsymbol{\delta}$ satisfying the conditions above, there exists an infinite divisible distribution μ having characteristic function (3.2).

In the present situation we can take \mathbf{A} the 2×2 matrix with zeroes, $\boldsymbol{\delta} = (0, 0)^T$ and define ν by the density

$$\frac{\partial^2 \nu(v, w)}{\partial v \partial w} = \phi(v/\sqrt{w}) w^{-3/2} 1_{(0,1)}(w),$$

where ϕ is the standard normal density. With these choices of \mathbf{A} , $\boldsymbol{\delta}$ and ν we get, using the notation $\mathbf{s} = (t, u)^T$ and $\mathbf{z} = (v, w)^T$,

$$\begin{aligned} & \exp \left\{ -\frac{1}{2} \mathbf{s}^T \mathbf{A} \mathbf{s} + \int_{\mathbb{R}^2} \left(e^{i\langle \mathbf{s}, \mathbf{z} \rangle} - 1 - i\langle \mathbf{s}, \mathbf{z} \rangle 1_{\{\|\mathbf{z}\| \leq 1\}}(\mathbf{z}) \right) d\nu(\mathbf{z}) + i\langle \boldsymbol{\delta}, \mathbf{s} \rangle \right\} \\ &= \exp \left\{ \int_{y=0}^1 \frac{e^{-\frac{1}{2}t^2y + iuy} - 1}{y} dy \right\}. \end{aligned}$$

We note that the computer package Mathematica evaluates the characteristic function for (V, W) in the following way:

$$\exp \left\{ \int_{y=0}^1 \frac{e^{-\frac{1}{2}t^2y + iuy} - 1}{y} dy \right\} = \exp \left\{ -\gamma - \Gamma(0, \frac{1}{2}t^2 - iu) - \log(\frac{1}{2}t^2 - iu) \right\}, \quad (3.3)$$

where γ is Euler's gamma and $\Gamma(0, \frac{1}{2}t^2 - iu)$ is the complementary incomplete gamma function, defined by:

$$\Gamma(0, \frac{1}{2}t^2 - iu) = \exp \left\{ -\frac{1}{2}t^2 + iu \right\} \int_0^\infty \exp \left\{ -(\frac{1}{2}t^2 - iu)x \right\} (1+x)^{-1} dx, \quad t \neq 0.$$

see, e.g., (2.01) on p. 109 of [15].

4. Central limit theorem for $\int h(\hat{f}_n(x)) dx$

In this section, we use the notation

$$b_n = \sqrt{\frac{3}{4} h''(1)^2 \log n}, \quad c_n = \sqrt{n}. \quad (4.1)$$

Using the conditioning of Section 1, the statistic $\int h(\hat{f}_n(x)) dx$ has the following representation:

$$\int h(\hat{f}_n(x)) dx = n^{-1} \sum_{j=1}^n \sum_{i=1}^{N_j} h\left(\frac{j}{S_{ji}}\right) S_{ji},$$

where the Poisson random variables N_j and the gamma random variables S_{ji} are defined as in (1.9), and where we condition on $(V_n, W_n) = (0, 1)$, where (V_n, W_n) is defined by (1.10). We define

$$U_n = \frac{1}{\sqrt{\frac{3}{4} h''(1)^2 \log n}} \left\{ \sum_{j=1}^n \sum_{i=1}^{N_j} \left[\left(h\left(\frac{j}{S_{ji}}\right) - h(1) \right) S_{ji} + h'(1) (S_{ji} - j) \right] - \frac{1}{2} h''(1) \log n \right\}. \quad (4.2)$$

Remark 4.1. The terms $h'(1)(S_{ji} - j)$ are present in U_n as variance reducing terms and give, after the summation over i and j , a zero contribution to U_n if $(V_n, W_n) = (0, 1)$. Also note that

$$h(1) \sum_{j=1}^n \sum_{i=1}^{N_j} S_{ji} = nh(1) = n\mu(h, f),$$

if the condition $(V_n, W_n) = (0, 1)$ is satisfied.

We assume that the function h satisfies condition 1.1 and first consider the conditional density of V_n , given $W_n = 1$.

Lemma 4.1. *The conditional density of V_n , given $W_n = 1$, is the density of a centered and standardized Gamma($n, 1$) variable:*

$$f_{V_n|W_n=1}(x) = \Gamma(n)^{-1} n^{n-1/2} e^{-n-x\sqrt{n}} (1 + x/\sqrt{n})^{n-1}, \quad x \in \mathbb{R}.$$

The density $f_{V_n|W_n=1}(x)$ converges uniformly to the standard normal density, as $n \rightarrow \infty$.

Proof. By Fourier inversion, the conditional characteristic function is given by:

$$\begin{aligned} & \frac{1}{\mathbb{P}\{W_n = 1\}} \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathbb{E} \exp \{itV_n + niuW_n - niu\} du \\ &= \frac{1}{\mathbb{P}\{W_n = 1\}} \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp \left\{ \sum_{j=1}^n \frac{e^{ij(u-t/c_n)} (1 - it/c_n)^{-j} - 1}{j} - niu \right\} du, \end{aligned}$$

where $c_n = \sqrt{n}$, see (3.1). Denoting the contour $w \mapsto e^{iw}$, $w \in [-\pi, \pi)$, by C , we write this in the form

$$\begin{aligned} & P\{W_n = 1\}^{-1} \exp \left\{ - \sum_{j=1}^n \frac{1}{j} \right\} \frac{1}{2\pi i} \int_C \exp \left\{ \sum_{j=1}^n \frac{(\beta_n(t)z)^j}{j} \right\} z^{-n-1} dz \\ &= \frac{1}{2\pi i} \int_C \exp \left\{ \sum_{j=1}^n \frac{(\beta_n(t)z)^j}{j} \right\} z^{-n-1} dz, \end{aligned} \tag{4.3}$$

where $\beta_n(t)$ is given by:

$$\beta_n(t) = e^{-\frac{it}{c_n}} \left(1 - \frac{it}{c_n} \right)^{-1},$$

and where we use:

$$\mathbb{P}\{W_n = 1\} = \exp \left\{ - \sum_{j=1}^n \frac{1}{j} \right\},$$

by Lemma 3.2 of [12]. An application of Cauchy's formula yields:

$$\frac{1}{2\pi i} \int_C \exp \left\{ \sum_{j=1}^n \frac{(\beta_n(t)z)^j}{j} \right\} z^{-n-1} dz = \beta_n(t)^n,$$

where we use that the coefficient of z^n in the power series around $z = 0$ of the function $z \mapsto \exp\{\sum_{j=1}^n (\beta_n(t)z)^j/j\}$ is the same as the coefficient of z^n in the power series of the function $z \mapsto (1 - \beta_n(t)z)^{-n}$.

Hence

$$\begin{aligned} \mathbb{E} \left\{ e^{itV_n} \mid W_n = 1 \right\} &= \frac{1}{\mathbb{P}\{W_n = 1\}} \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathbb{E} \exp \{itV_n + niuW_n - niu\} du \\ &= \beta_n(t)^n = \exp \{ -nit/c_n \} \left(1 - \frac{it}{c_n} \right)^{-n}. \end{aligned}$$

This is just the characteristic function of the sum of n standardized exponential variables, and hence its density tends uniformly to the standard normal density by Theorem 2 on p. 516 of [5]. \square

So, in particular, we get:

$$\mathbb{P}\{W_n = 1\}f_{V_n|W_n=1}(0) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\sum_{j=1}^n 1/j\right\}(1 + o(1)) = \frac{e^{-\gamma}}{n\sqrt{2\pi}}(1 + o(1)), \quad n \rightarrow \infty,$$

where γ is Euler's gamma. The characteristic function of $(U_n|V_n = 0, W_n = 1)$ is therefore given by

$$\begin{aligned} & \frac{1}{4\pi^2} \int_{t=-\infty}^{\infty} \int_{-\pi}^{\pi} \mathbb{E} e^{isU_n + itV_n + iuW_n - inu} du dt / (\mathbb{P}\{W_n = 1\}f_{V_n|W_n=1}(0)) \\ & \sim \frac{ne^\gamma}{(2\pi)^{3/2}} \int_{t=-\infty}^{\infty} \int_{-\pi}^{\pi} \mathbb{E} e^{isU_n + itV_n + iuW_n - inu} du dt. \end{aligned} \quad (4.4)$$

We now consider the characteristic function ϕ_{nj} , defined by:

$$\phi_{nj}(s, t) = E \exp \left\{ is \frac{\{h(j/S_{j1}) - h(1)\} S_{j1} + h'(1)(S_{j1} - j)}{b_n} + it \frac{S_{j1} - j}{c_n} \right\}. \quad (4.5)$$

which involves the components of the random variables U_n and V_n . Its asymptotic behavior is determined using a saddle point method.

Lemma 4.2. *The characteristic function (4.5) satisfies*

$$\phi_{nj}(s, t) \sim \frac{1}{j} \exp \left\{ -\frac{jt^2}{2n} + O\left(\frac{j}{nb_n}\right) \right\} \left\{ 1 + \frac{ish''(1)}{2b_n} - \frac{3s^2h''(1)^2}{8b_n^2} \right\}, \quad j \rightarrow \infty. \quad (4.6)$$

uniformly for t in a bounded interval.

Proof. After a change of variables, $\phi_{nj}(s, t)$ can be written:

$$\begin{aligned} & \frac{j^j}{\Gamma(j)} \int_{x=0}^{\infty} \exp \left[j \left\{ is \frac{\{h(1/x) - h(1)\}x + h'(1)(x-1)}{b_n} + \frac{it(x-1)}{c_n} - x + \log x \right\} \right] x^{-1} dx \\ & = \frac{j^j}{\Gamma(j)} \int_{x=0}^{\infty} \exp \{ j f_{n,s,t}(x) \} x^{-1} dx, \end{aligned}$$

where $f_{n,s,t}$ is defined on the right half plane by:

$$f_{n,s,t}(z) = is \frac{\{h(1/z) - h(1)\}z + h'(1)(z-1)}{b_n} + \frac{it(z-1)}{c_n} - z + \log z, \quad \text{Re}(z) > 0, \quad (4.7)$$

The derivative of $f_{n,s,t}$ is given by

$$f'_{n,s,t}(z) = \frac{is\{h(1/z) - h(1)\}}{b_n} - \frac{ish'(1/z)}{b_n z} + \frac{ish'(1)}{b_n} + \frac{it}{c_n} - 1 + \frac{1}{z}. \quad (4.8)$$

A saddle point is given by the equation

$$f'_{n,s,t}(z) = 0.$$

Multiplying both sides of this equation with z , we get the equation

$$z = g_n(z), \quad g_n(z) \stackrel{\text{def}}{=} \left(1 - \frac{it}{c_n}\right)^{-1} \left\{ 1 + \frac{is\{h(1/z) - h(1)\}z}{b_n} - \frac{ish'(1/z)}{b_n} + \frac{ish'(1)z}{b_n} \right\} \quad (4.9)$$

This equation has, for sufficiently large n , a unique solution in a neighborhood of $z_0(t) = \left(1 - \frac{it}{c_n}\right)^{-1}$, as is clear from the following properties.

(i)

$$g'_n(z) = \left(1 - \frac{it}{c_n}\right)^{-1} b_n^{-1} \left\{ ish(1) + ish(1/z) + ish'(1) - \frac{ish'(1/z)}{z} + \frac{ish''(1/z)}{z^2} \right\} \rightarrow 0, \quad n \rightarrow \infty.$$

(ii) For $z_0(t) = \left(1 - \frac{it}{c_n}\right)^{-1}$ we have:

$$g_n(z_0(t)) - z_0(t) = -\frac{is \{h(1) - h(1 - it/c_n) - h'(1)/c_n + (1 - it/c_n)h'(1 - it/c_n)\}}{b_n(1 - it/c_n)^2} \rightarrow 0, \quad n \rightarrow \infty,$$

see, e.g., (3.3), p. 56 of [4]. So the saddle point is given by a solution of the equation (4.9) and can be found by the simple iteration $z_{k+1} = g_n(z_k)$, starting at $z_0(t)$. We can, however, also take $z_0(t)$ itself instead of the real saddle point for the asymptotic expansion, since this gives us the same terms in the expansion we need.

The value of $f_{n,s,t}$ at $z_0(t)$ has the following expansion:

$$f_{n,s,t}(z_0(t)) = -1 - \frac{t^2 + ist^2 h''(1)/(2b_n) + o(b_n^{-1})}{2n} = -1 - \frac{t^2}{2n} + O\left(\frac{1}{nb_n}\right). \quad (4.10)$$

Furthermore,

$$f''_{n,s,t}(z) = \frac{-b_n z + ish''(1/z)}{b_n z^3}.$$

so

$$f''_{n,s,t}(z_0(t)) = -1 + \frac{ish''(1)}{b_n} + O(c_n^{-1}).$$

It follows that we get:

$$\frac{\alpha_n}{|f''_{n,s,t}(z_0(t))|^{1/2}} \approx 1 + \frac{ish''(1)}{2b_n} - \frac{3h''(1)^2 s^2}{8b_n^2} + O(c_n^{-1}) = 1 + \frac{ish''(1)}{2b_n} - \frac{s^2}{2 \log n} + O(c_n^{-1}),$$

where $\alpha_n = \exp\{ish''(1)/(2b_n)\}$ is a complex number with absolute value 1, corresponding to the argument of the main axis of the saddle point (note that the argument of this axis is $\frac{1}{2}\pi - \frac{1}{2}\arg f''_{n,s,t}(z_0(t))$, see [3], p. 84).

Evaluating the integrand at $z_0(t)$, and applying Stirling's formula on $\Gamma(j)$, we obtain the following asymptotic representation:

$$\begin{aligned} \phi_{nj}(s, t) &= \frac{j^j}{\Gamma(j)} \int_{x=0}^{\infty} \exp\{j f_{n,s,t}(x)\} x^{-1} dx \sim \frac{\alpha_n \sqrt{2\pi} e^{j\{f_{n,s,t}(z_0(t))+1\}}}{\sqrt{2\pi j} \sqrt{j |f''_{n,s,t}(z_0(t))|}} \\ &\sim \frac{1}{j} \exp\left\{-\frac{jt^2}{2n} + O\left(\frac{j}{nb_n}\right)\right\} \left\{1 + \frac{ish''(1)}{2b_n} - \frac{3s^2 h''(1)^2}{8b_n^2}\right\}, \quad j \rightarrow \infty. \end{aligned} \quad (4.11)$$

see, e.g., [3], (5.10.3) on p. 92 for the first asymptotic equivalence. The second asymptotic equivalence holds uniformly for t in a bounded interval. Note that this corresponds to changing the path of integration for x to a path in the complex plane, going through the saddle point. \square

We can now prove the following property of the characteristic function of (U_n, V_n, W_n) . This is “almost” the Fourier inversion for (V_n, W_n) , but we still have to extend the inversion for V_n to the whole real line. Cauchy's formula is an essential ingredient of the proof of Lemma 4.3.

Lemma 4.3. *Let U_n be defined by (4.2) and let h satisfy condition 1.1. Then, for each $M > 0$:*

$$\frac{n e^\gamma}{(2\pi)^{3/2}} \int_{t=-M}^M \int_{-\pi}^{\pi} \mathbb{E} e^{isU_n + itV_n + niuW_n - inu} du dt \rightarrow \frac{e^{-\frac{1}{2}s^2}}{\sqrt{2\pi}} \int_{t=-M}^M e^{-\frac{1}{2}t^2} dt, \quad n \rightarrow \infty,$$

where γ is Euler's gamma.

Proof. Let ϕ_{nj} be defined by (4.5). Since, by (4.1) and (4.2),

$$\begin{aligned} U_n &= \frac{1}{\sqrt{\frac{3}{4}h''(1)^2 \log n}} \left\{ \sum_{j=1}^n \sum_{i=1}^{N_j} \left[\left(h\left(\frac{j}{S_{ji}}\right) - h(1) \right) S_{ji} + h'(1)(S_{ji} - j) \right] - \frac{1}{2}h''(1) \log n \right\} \\ &= b_n^{-1} \left\{ \sum_{j=1}^n \sum_{i=1}^{N_j} \left[\left(h\left(\frac{j}{S_{ji}}\right) - h(1) \right) S_{ji} + h'(1)(S_{ji} - j) \right] - \frac{1}{2}h''(1) \log n \right\}, \end{aligned}$$

where

$$b_n = \sqrt{\frac{3}{4}h''(1)^2 \log n},$$

we get, evaluating the probabilities for the Poisson random variables N_j in the third line:

$$\begin{aligned} \mathbb{E} \exp \{isU_n + itV_n + iuW_n\} &= \exp \left\{ -\frac{1}{2}ish''(1)(\log n)/b_n \right\} \mathbb{E} \left\{ \prod_{j=1}^n \{\phi_{nj}(s, t)\}^{N_j} e^{iuN_j/n} \right\} \\ &= \exp \left\{ -\frac{1}{2}ish''(1)(\log n)/b_n + \sum_{j=1}^n \frac{1}{j} \left\{ \phi_{nj}(s, t) e^{iju/n} - 1 \right\} \right\} \\ &= \exp \left[-\frac{1}{2}ish''(1)(\log n)/b_n + \sum_{j=1}^n \frac{1}{j} \left\{ \exp \left(-\frac{jt^2}{2n} + \frac{iju}{n} \right) - 1 \right\} \right. \\ &\quad \left. + \left\{ \frac{ish''(1)}{2b_n} - \frac{3s^2h''(1)^2}{8b_n^2} \right\} \sum_{j=1}^n \frac{1}{j} \exp \left(-\frac{jt^2}{2n} + \frac{iju}{n} \right) + o(1) \right]. \quad (4.12) \end{aligned}$$

As in the proof of Lemma 4.1 we consider the contour $w \mapsto e^{iw}$, $w \in [-\pi, \pi)$, and denote this contour by C . So, integrating (4.12) w.r.t. u and changing variables we get:

$$\frac{1}{2\pi} \int_{u=-\pi}^{\pi} \mathbb{E} \exp \{isU_n + itV_n + inuW_n - inu\} du \quad (4.13)$$

$$= \exp \left\{ -\frac{1}{2}ish''(1)(\log n)/b_n - \sum_{j=1}^n \frac{1}{j} \right\} \frac{1}{2\pi i} \int_C \exp \left\{ \sum_{j=1}^n \frac{\delta_n(\beta_n(t)z)^j}{j} \right\} z^{-n-1} dz, \quad (4.14)$$

where

$$\beta_n(t) \sim \exp \left(-\frac{t^2}{2n} \right), \quad (4.15)$$

and

$$\delta_n(t) = \frac{\alpha_n}{\sqrt{|f''_{n,s,t}(z_0(t))|}} = 1 + \frac{ish''(1)}{2b_n} - \frac{3s^2h''(1)^2}{8b_n^2} + O(b_n^{-3}), \quad (4.16)$$

and where α_n is a complex number of absolute value 1, corresponding to the angle of the main axis through the (approximate) saddle point $z_0(t)$.

So we get, by Cauchy's formula, for $t \in [-M, M]$ and M arbitrarily large,

$$\begin{aligned}
& \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathbb{E} \exp \{isU_n + itV_n + inuW_n - niu\} du \\
& \sim \exp \left\{ -\frac{1}{2} ish''(1)(\log n)/b_n - \sum_{j=1}^n \frac{1}{j} \right\} (-1)^n \binom{-\delta_n(t)}{n} \beta_n(t)^n \\
& = \exp \left\{ -\frac{1}{2} ish''(1)(\log n)/b_n - \sum_{j=1}^n \frac{1}{j} \right\} \prod_{j=1}^n \left(1 + \frac{\delta_n(t) - 1}{j} \right) \beta_n(t)^n \\
& \sim \exp \left\{ -\frac{1}{2} ish''(1)(\log n)/b_n - \sum_{j=1}^n \frac{1}{j} \right\} \exp \left\{ -\frac{1}{2} t^2 + \sum_{j=1}^n \log \left(1 + \frac{\delta_n(t) - 1}{j} \right) \right\} \\
& = \exp \left\{ -\frac{1}{2} ish''(1)(\log n)/b_n - \log n - \gamma \right\} \exp \left\{ -\frac{1}{2} t^2 + \{\delta_n(t) - 1\} \sum_{j=1}^n \frac{1}{j} + o(1) \right\} \\
& = \frac{1}{n} \exp \left\{ -\frac{1}{2} ish''(1)(\log n)/b_n - \gamma \right\} \exp \left\{ -\frac{1}{2} t^2 + \left\{ \frac{ish''(1)}{2b_n} - \frac{3s^2 h''(1)^2}{8b_n^2} \right\} \sum_{j=1}^n \frac{1}{j} + o(1) \right\}. \quad (4.17)
\end{aligned}$$

Thus:

$$\begin{aligned}
& \frac{ne^\gamma}{(2\pi)^{3/2}} \int_{t=-M}^M \int_{-\pi}^{\pi} \mathbb{E} e^{isU_n + itV_n + niuW_n - inu} du dt \\
& = \frac{1}{\sqrt{2\pi}} \int_{t=-M}^M \exp \left\{ -\frac{1}{2} t^2 - \frac{3s^2 h''(1)^2}{8b_n^2} \sum_{j=1}^n \frac{1}{j} + o(1) \right\} dt \\
& = \exp \left\{ -\frac{1}{2} s^2 + o(1) \right\} \frac{1}{\sqrt{2\pi}} \int_{t=-M}^M \exp \left\{ -\frac{1}{2} t^2 \right\} dt.
\end{aligned}$$

□

We still have to prove that the remaining part of integral w.r.t. the integration variable t can be made arbitrarily small by choosing M large. To this end we split the remaining region into two regions: $A_1 = \{t \in \mathbb{R} : M < |t| \leq \delta n^{1/2}\}$ and $A_2 = \{t \in \mathbb{R} : |t| > \delta n^{1/2}\}$. This split-up is familiar from inversion theorems for densities, see, e.g., the proof of Theorem 2 on p. 516 of [5]. We start with the region $A_1 = \{t \in \mathbb{R} : M < |t| \leq \delta n^{1/2}\}$.

Lemma 4.4. *Let h satisfy condition 1.1. Then there exists for each $\varepsilon > 0$ an $M > 0$ and $\delta > 0$ such that*

$$\left| \int_{t: M < |t| \leq \delta n^{1/2}} \int_{u=-\pi}^{\pi} \mathbb{E} \exp \{isU_n + itV_n + inuW_n - niu\} du dt \right| < \varepsilon,$$

Proof. We consider again the expansion of the function $f_{n,s,t}$ defined by (4.7) at the point $z_0(t) = (1 - it/c_n)^{-1}$. We get:

$$f_{n,t,u}(z_0(t)) = -1 - \frac{t^2}{2n} - \frac{ist^2 h''(\theta)}{2nb_n},$$

where θ is a point on the line segment between 1 and $1/(1 - it/c_n)$. Likewise

$$f''_{n,t,u}(z_0(t)) = -1 + \frac{ish''(1)}{b_n} + \frac{2i\theta'}{c_n},$$

where θ' is a point on the line segment between 1 and $1/(1 - it/c_n)$.

So we have a local expansion

$$\phi_{nj}(s, t) \sim \frac{1}{j} \exp \left\{ -\frac{jt^2}{2n} + O \left(\frac{jt^2}{nb_n} \right) \right\} \left\{ 1 + \frac{ish''(1)}{2b_n} - \frac{3s^2 h''(1)^2}{8b_n^2} + O(c_n^{-1}) \right\}, \quad j \rightarrow \infty. \quad (4.18)$$

as in (4.11). This means that we can follow the same steps as in the proof of Lemma 4.3 and that we can choose M and $\delta > 0$ in such a way that

$$\begin{aligned} & \left| \frac{ne^\gamma}{(2\pi)^{3/2}} \int_{M < |t| \leq \delta n^{1/2}} \int_{-\pi}^{\pi} \mathbb{E} e^{isU_n + itV_n + niuW_n - inu} du dt \right| \\ & \leq \exp \left\{ -\frac{1}{2}s^2 + o(1) \right\} \frac{1}{\sqrt{2\pi}} \int_{M < |t| \leq \delta n^{1/2}} \exp \left\{ -\frac{1}{2}t^2 + \frac{1}{4}t^2 \right\} dt \\ & = \exp \left\{ -\frac{1}{2}s^2 + o(1) \right\} \frac{1}{\sqrt{2\pi}} \int_{M < |t| \leq \delta n^{1/2}} \exp \left\{ -\frac{1}{4}t^2 \right\} dt < \varepsilon. \end{aligned}$$

□

The following lemma deals with the region $A_2 = \{t \in \mathbb{R} : |t| > \delta n^{1/2}\}$.

Lemma 4.5. *Let h satisfy condition 1.1. Then, for each $\delta > 0$:*

$$\int_{|t| > \delta c_n} \int_{u=-\pi}^{\pi} \mathbb{E} \exp \{ isU_n + itV_n + inuW_n - niu \} du dt \longrightarrow 0, \quad n \rightarrow \infty.$$

Proof. We consider the characteristic function:

$$\bar{\phi}_{nj}(s, t) = E \exp \left\{ is \frac{\{h(j/S_{j1}) - h(1)\} S_{j1} + h'(1)(S_{j1} - j)}{b_n} + it(S_{j1} - j) \right\},$$

so we replace c_n by 1 in (4.5). This means that, for the saddle point analysis, the constant c_n is replaced by 1 in the function (4.7). So we now define

$$\bar{f}_{n,s,t}(z) = is \frac{\{h(1/z) - h(1)\}z + h'(1)(z - 1)}{b_n} + it(z - 1) - z + \log z, \quad \operatorname{Re}(z) > 0, \quad (4.19)$$

The saddle point equation (4.9) now turns into

$$z = g_n(z), \quad g_n(z) \stackrel{\text{def}}{=} (1 - it)^{-1} \left\{ 1 + \frac{is\{h(1/z) - h(1)\}z}{b_n} - \frac{ish'(1/z)}{b_n} + \frac{ish'(1)z}{b_n} \right\} \quad (4.20)$$

and has a unique solution in a neighborhood of $(1 - it)^{-1}$ for the same reasons as before.

We define:

$$z_0(t) = 1/(1 - it).$$

Then:

$$\begin{aligned} \bar{f}_{n,s,t}(z_0(t)) &= -\frac{t^2 + 1}{1 - it} - \log(1 - it) - is \frac{h(1) - h(1 - it) + sth'(1)}{b_n(1 - it)} \\ &= -\frac{t^2 + 1}{1 - it} + O(b_n^{-1}|t|), \end{aligned}$$

and

$$\begin{aligned} \bar{f}_{n,s,t}'(z_0(t)) &= (t + i)^2 - \frac{s(t + i)h''(1 - it)}{b_n} \\ &= (t + i)^2 + O(b_n^{-1}|t|), \end{aligned}$$

uniformly in $|t| > \delta$, using Condition 1.1. This implies that, uniformly for $|t| > \delta$,

$$\begin{aligned} \phi_{nj}(s, t) &= \frac{j^j}{\Gamma(j)} \int_{x=0}^{\infty} \exp\{j f_{n,s,t}(x)\} x^{-1} dx \sim \frac{\alpha_n(t) e^{j \{\bar{f}_{n,s,t}(z_0(t)) + 1\}}}{j \sqrt{|\bar{f}_{n,s,t}''(z_0(t))|}} \\ &= \frac{\alpha_n(t) \exp\{-j(it + \log(1 - it) + O(b_n^{-1}))\}}{j |(t + i)^2 + O(b_n^{-1}|t|)|^{1/2}}, \quad j \rightarrow \infty, \end{aligned} \quad (4.21)$$

where $\alpha_n(t)$ is a complex number with absolute value 1 and argument $\frac{1}{2}\pi - \frac{1}{2} \arg \bar{f}_{n,s,t}''(z_0(t))$.

So we find, using Cauchy's formula again, as in the proof of Lemma 4.3, and using Condition 1.1,

$$\begin{aligned} &\frac{1}{2\pi} \int_{-\pi}^{\pi} \mathbb{E} \exp\{isU_n + itc_n V_n + inuW_n - niu\} du \\ &\sim \exp\left\{-\frac{1}{2}ish''(1)(\log n)/b_n - \sum_{j=1}^n \frac{1}{j}\right\} (-1)^n \binom{-\delta_n(t)}{n} \beta_n(t)^n \\ &= \exp\left\{-\frac{1}{2}ish''(1)(\log n)/b_n - \sum_{j=1}^n \frac{1}{j}\right\} \prod_{j=1}^n \left(1 + \frac{\delta_n(t) - 1}{j}\right) \beta_n(t)^n, \end{aligned}$$

where

$$\beta_n(t) = \exp\{-(it + \log(1 - it) + O(b_n^{-1}))\},$$

and

$$\delta_n(t) = \frac{\alpha_n}{|(t + i)^2 + O(b_n^{-1}|t|)|^{1/2}}.$$

Hence

$$\begin{aligned} &\exp\left\{-\frac{1}{2}ish''(1)(\log n)/b_n - \sum_{j=1}^n \frac{1}{j}\right\} \prod_{j=1}^n \left(1 + \frac{\delta_n(t) - 1}{j}\right) \beta_n(t)^n \\ &= \frac{1}{n} \exp\left\{-\frac{1}{2}ish''(1)(\log n)/b_n - \gamma\right\} \\ &\quad \cdot \exp\left\{-n\{it + \log(1 - it) + O(b_n^{-1})\} + \frac{\alpha_n \sum_{j=1}^n 1/j}{|(t + i)^2 + O(b_n^{-1}|t|)|^{1/2}}\right\}, \end{aligned}$$

implying

$$\begin{aligned} &\left| \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathbb{E} \exp\{isU_n + itc_n V_n + inuW_n - niu\} du \right| \\ &\leq \frac{1}{n} \exp\left\{-n\left\{\frac{1}{2} \log(1 + t^2) + O(b_n^{-1})\right\} + \frac{\sum_{j=1}^n 1/j}{|(t + i)^2 + O(b_n^{-1}|t|)|^{1/2}}\right\} \\ &\leq |1 + t^2|^{-n/4}, \end{aligned}$$

for large n , uniformly for $|t| > \delta$, using Condition 1.1. It now follows that

$$\begin{aligned} &\left| \frac{1}{2\pi} \int_{|t| > \delta c_n} \int_{u=-\pi}^{\pi} \mathbb{E} \exp\{isU_n + itV_n + inuW_n - niu\} du dt \right| \\ &\left| \frac{1}{2\pi} \int_{|t| > \delta} \int_{u=-\pi}^{\pi} \mathbb{E} \exp\{isU_n + itc_n V_n + inuW_n - niu\} du dt \right| \\ &\leq \int_{|t| > \delta} (1 + t^2)^{-n/4} dt \longrightarrow 0, \quad n \rightarrow \infty. \end{aligned}$$

□

This leads to the main result of this section.

Theorem 4.1. *Let h satisfy Condition 1.1. Then $U_n | (V_n, W_n) = (0, 1)$ converges in law to a standard normal distribution.*

Proof. The preceding lemma's imply

$$\frac{ne^\gamma}{(2\pi)^{3/2}} \int_{t=-\infty}^{\infty} \int_{-\pi}^{\pi} \mathbb{E} e^{isU_n + itV_n + niuW_n - inu} du dt \longrightarrow \frac{e^{-\frac{1}{2}s^2}}{\sqrt{2\pi}} \int_{t=-\infty}^{\infty} e^{-\frac{1}{2}t^2} dt = e^{-\frac{1}{2}s^2}, \quad n \rightarrow \infty,$$

where γ is Euler's gamma. Hence

$$\mathbb{E} \{ e^{isU_n} \mid (V_n, W_n) = (0, 1) \} \sim \frac{ne^\gamma}{(2\pi)^{3/2}} \int_{t=-\infty}^{\infty} \int_{-\pi}^{\pi} \mathbb{E} e^{isU_n + itV_n + niuW_n - inu} du dt \longrightarrow e^{-\frac{1}{2}s^2}, \quad n \rightarrow \infty.$$

□

Theorem 1.2 now follows from the conditional representation from Section 1, definition (4.2) of U_n , Remark 4.1 and Theorem 4.1.

5. Conclusion

We derived a general theorem (Theorem 1.2) for integrals of the Grenander estimator when the distribution is uniform from a representation in terms of Poisson and gamma random variables in [12]. The result implies the main result of [12] and gives also the limit behavior of the entropy functional. We corrected the limit distribution of the conditioning variables given in Lemma 3.1 of [12] in Section 3. The methods used are rather different from the methods in [12], where a result in [13] was used.

Our main result was inspired by [14] who derived a similar result from [12] under different conditions. The main tools are Cauchy's formula and the saddle point method for integrals of analytic functions of a complex variable. A simple version of the approach is given in Section 2 to illustrate the method without the complications of the saddle point method.

References

- [1] K. Behnen. *Güteeigenschaften von Rangtests unter Bindungen*. Habilitationsschrift, University of Freiburg, 1974.
- [2] K. C. G. Chan, C-F Tang, and S.C.P Yam. Likelihood ratio test for monotonicity of density. Submitted, 2018.
- [3] N. G. de Bruijn. *Asymptotic methods in analysis*. Dover Publications, Inc., New York, third edition, 1981. ISBN 0-486-64221-6.
- [4] Jean Dieudonné. *Calcul infinitésimal*. Hermann, Paris, 1968.
- [5] William Feller. *An introduction to probability theory and its applications. Vol. II*. Second edition. John Wiley & Sons, Inc., New York-London-Sydney, 1971.
- [6] Ulf Grenander. On the theory of mortality measurement. II. *Skand. Aktuarietidskr.*, 39:125–153 (1957), 1956.
- [7] P. Groeneboom. Estimating a monotone density. In *Proceedings of the Berkeley conference in honor of Jerzy Neyman and Jack Kiefer, Vol. II (Berkeley, Calif., 1983)*, Wadsworth Statist./Probab. Ser., pages 539–555, Belmont, CA, 1985. Wadsworth.
- [8] P. Groeneboom and H. P. Lopuhaä. Isotonic estimators of monotone densities and distribution functions: basic facts. *Statist. Neerlandica*, 47:175–183, 1993. ISSN 0039-0402. . URL <http://dx.doi.org/10.1111/j.1467-9574.1993.tb01415.x>.
- [9] Piet Groeneboom. The concave majorant of Brownian motion. *Ann. Probab.*, 11:1016–1027, 1983. ISSN 0091-1798. URL [http://links.jstor.org/sici?sici=0091-1798\(198311\)11:4<1016:TCMOBM>2.0.CO;2-7&origin=MSN](http://links.jstor.org/sici?sici=0091-1798(198311)11:4<1016:TCMOBM>2.0.CO;2-7&origin=MSN).
- [10] Piet Groeneboom and Geurt Jongbloed. *Nonparametric Estimation under Shape Constraints*. Cambridge Univ. Press, Cambridge, 2014.
- [11] Piet Groeneboom and Geurt Jongbloed. Some Developments in the Theory of Shape Constrained Inference. *Statist. Sci.*, 33(4):473–492, 2018. ISSN 0883-4237. . URL <https://doi.org/10.1214/18-STS657>.
- [12] Piet Groeneboom and Ronald Pyke. Asymptotic normality of statistics based on the convex minorants of empirical distribution functions. *Ann. Probab.*, 11:328–345, 1983. ISSN 0091-1798. URL [http://links.jstor.org/sici?sici=0091-1798\(198305\)11:2<328:ANOSBO>2.0.CO;2-Q&origin=MSN](http://links.jstor.org/sici?sici=0091-1798(198305)11:2<328:ANOSBO>2.0.CO;2-Q&origin=MSN).
- [13] Lucien Le Cam. Un théorème sur la division d'un intervalle par des points pris au hasard. *Publ. Inst. Statist. Univ. Paris*, 7(3/4):7–16, 1958.
- [14] Rajarshi Mukherjee and Bodhisattva Sen. Estimation of integrated functionals of a monotone density. Submitted, 2019. URL <https://arxiv.org/abs/1808.07915>.

- [15] F. W. J. Olver. *Asymptotics and special functions*. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1974. Computer Science and Applied Mathematics.
- [16] B.L.S. Prakasa Rao. Estimation of a unimodal density. *Sankhya Ser. A*, 31:23–36, 1969. ISSN 0581-572X.
- [17] Ken-iti Sato. Basic results on Lévy processes. In *Lévy processes*, pages 3–37. Birkhäuser Boston, Boston, MA, 2001.
- [18] F.-W. Scholz. Combining independent P -values. In *A Festschrift for Erich L. Lehmann*, Wadsworth Statist./Probab. Ser., pages 379–394. Wadsworth, Belmont, Calif., 1983.
- [19] E. Sparre Andersen. On the fluctuations of sums of random variables. II. *Math. Scand.*, 2:195–223, 1954. ISSN 0025-5521.