A Gibbs Conditional Theorem under Extreme Deviation

M Biret, M Broniatowski, Z Cao

A GIBBS CONDITIONAL THEOREM UNDER EXTREME DEVIATION

MAEVA BIRET, MICHEL BRONIATOWSKI, AND ZANGSHEN CAO

Abstract. We explore some properties of the conditional distribution of an i.i.d. sample under large exceedances of its sum. Thresholds for the asymptotic independence of the summands are observed, in contrast with the classical case when the conditioning event is in the range of a large deviation. This paper is an extension to [6]. Tools include a new Edgeworth expansion adapted to specific triangular arrays where the rows are generated by tilted distribution with diverging parameters, together with some Abelian type results.

1. Introduction

Let $X_1^n := (X_1, .., X_n)$ be n independent unbounded real valued random variables and $S_1^n := X_1 + .. + X_n$ denote their sum. The purpose of this paper is to explore the limit distribution of the generic variable X_1 conditioned on extreme deviations (ED) pertaining to S_1^n. By extreme deviation we mean that S_1^n/n is supposed to take values which are going to infinity as n increases. Obviously such events are of infinitesimal probability. Our interest in this question stems from a first result which assesses that under appropriate conditions, when the sequence a_n is such that

$$\lim_{n\to\infty} a_n = \infty$$

then there exists a sequence ε_n for which $\lim_{n\to\infty} \varepsilon_n/a_n = 0$ such that

$$\lim_{n\to\infty} P\left(\cap_{i=1}^n \left(X_i \in (a_n - \varepsilon_n, a_n + \varepsilon_n)\right) \middle| S_1^n/n > a_n\right) = 1, \tag{1.1}$$

which is to say that when the empirical mean takes exceedingly large values, then all the summands share the same behavior. This result obviously requires a number of hypotheses, which we simply quote as "light tails" type. We refer to [6] for this result and the connection with earlier related works; arguments stating that such most unusual cases may be considered are presented in this latest paper, in relation with the Erdös-Rényi law of large numbers and the formation of high level aggregates in random sequences. Also these results have various applications in physics; see [12] and [17] where (1.1) is considered in the extreme deviation context, with applications to turbulence and fragmentation.

The above result is clearly to be put in relation with the so-called Gibbs conditional Principle which we recall briefly in its simplest form.

Consider the case when the sequence $a_n = a$ is constant with value larger than the expectation of X_1. Hence we consider the behavior of the summands when $(S_1^n/n > a)$, i.e. under a large deviation (LD) condition about the empirical mean. The asymptotic conditional distribution of X_1 given $(S_1^n/n > a)$ is the well known tilted distribution of P_X with parameter t associated to a. Let us introduce some

Key words and phrases. Gibbs conditional principle, extreme deviation.

notation to shed some light on this. The hypotheses to be stated now together with notation are kept throughout the entire paper. Without loss of generality it is assumed that the generic r.v. X_1 takes only non negative values.

It will be assumed that P_X , which is the distribution of X_1, has a density p with respect to the Lebesgue measure on $\mathbb{R}_+$. The fact that X_1 has a light tail is captured in the hypothesis that X_1 has a moment generating function

$$\Phi(t) := E[\exp tX_1],$$

which is finite in a non void neighborhood $\mathcal{N}$ of 0 (Cramer condition).

Defined on $\mathcal{N}$ are the following functions:

$$t \to m(t) := \frac{d}{dt} \log \Phi(t) \tag{1.2}$$

$$t \to s^2(t) := \frac{d}{dt} m(t) \tag{1.3}$$

$$t \to \mu_j(t) := \frac{d^j}{dt^j} \log \Phi(t),\ j \geq 3 \tag{1.4}$$

The function m measures the expectation of the r.v. $\mathcal{X}_t$ with density

$$\pi_t(x) := \frac{\exp tx}{\Phi(t)} p(x),$$

the tilted density with parameter t, the function s^2 measures its variance and μ_j measures its $j-$th centered moment. When Φ is steep, meaning that

$$\lim_{t \to t^{\pm}} m(t) = \infty \tag{1.5}$$

where $t^+ := ess\sup \mathcal{N}$ (resp. $t^- := ess\inf \mathcal{N}$) then m parametrizes the convex hull of the support of P_X; see [3] Theorem 9.2 for those properties. As a consequence of this fact, for all a in the support of P_X, it will be convenient to define

$$\pi^a = \pi_t$$

where a is the unique solution of the equation $m(t) = a$. The function

$$x \to V(x) := s^2 o m^{\leftarrow}(x) \tag{1.6}$$

is the "Variance function" of the model π_t , $t \in \mathcal{N}$ and it characterizes its heteroscedasticity property; see for instance [15], [2] and [14]. In (1.6) $m^{\leftarrow}$ designates the inverse (reciprocal) function of m.

The Gibbs conditional principle in the standard above setting can be stated as follows.

The distribution of X_1 given $(S_1^n/n > a)$ is asymptotically Π^a as n tends to infinity, where Π^a has density π^a; see [8]; we first state Gibbs principle where the conditioning event is a point condition $(S_1^n/n = a)$. The conditional distribution of X_1 given $(S_1^n/n = a)$ is a well defined distribution and Gibbs conditional principle states that it converges to Π^a as n tends to infinity; see [10]. Both convergences hold in total variation norm.

For all a (depending on n or not) denote p^a the density of the random vector X_1^k conditioned upon the local event $(S_1^n = na)$. The notation $p^a\left(X_1^k = x_1^k\right)$ is used to denote the value of the density p^a at point x_1^k. Similarly for a random variable or vector Z with density f we write $f(Z = z)$ for $f(z)$.

This article is organized as follows. Notation and hypotheses are stated in Section 2, along with some necessary facts from asymptotic analysis in the context of light tailed densities. Section 4 provides a local Gibbs conditional principle under ED, namely producing the pointwise approximation of the conditional density of X_1 conditionally on $(S_1^n/n = a_n)$ for sequences a_n which tend to infinity. We explore two rates of growth for the sequence a_n, which yield two different approximating distributions for the conditional law of X_1. The first one extends the classical approximation by the tilted one, substituting π^a by π^{a_n}. The second case, which corresponds to a faster growth of a_n, produces an approximation of a different kind. It may be possible to explore faster growth conditions than those considered here, leading to a wide class of approximating distributions; this would require some high order Edgeworth expansions for triangular arrays of variables, extending the corresponding result of order 3 presented in Section 3; we did not move further in this direction, in order to avoid technicalities.

For fixed k and fixed $a_n = a > E(X_1)$ it is known that the r.v's $X_1, \ldots, X_k$ are asymptotically independent given $(S_1^n/n = a_n)$; see [10] and [8] when the conditioning event is $(S_1^n/n > a_n)$. This statement is explored when a_n grows to infinity with n . It is shown that the asymptotic independence property holds for sequences a_n with moderate growth, and that it fails for sequences a_n with fast growth.

The local approximation of the density of X_1 conditionally on $(S_1^n/n = a_n)$ is further extended to typical paths under the conditional sampling scheme, which in turn provides the approximation in variation norm for the conditional distribution; the method used here follows closely the approach by [7]. Section 5 states similar results in the case when the conditioning event is $(S_1^n/n > a_n)$ .The relation between (1.1) and Gibbs conditional principle under exceedance is also discussed.

The main tools to be used come from asymptotic analysis and local limit theorems, developed in [11] and [4]; we also have borrowed a number of arguments from [16]. The basic Abelian type result which is used is stated in [5]. Unless specified, all proofs are differed to Section 6.

2. Notation and hypotheses

Thereafter we will use indifferently the notation $f(t) \underset{t\to\infty}{\sim} g(t)$ and $f(t) \underset{t\to\infty}{=} g(t)(1 + o(1))$ to specify that f and g are asymptotically equivalent functions.

The density p is assumed to be of the form

$$p(x) = \exp(-(g(x) - q(x))), \quad x \in \mathbb{R}_+. \tag{2.1}$$

The function q is assumed to be bounded, so that the asymptotic behavior of p is captured through the function g which is assumed to be is positive, convex, four times differentiable and which satisfies

$$\frac{g(x)}{x} \underset{x\to\infty}{\longrightarrow} \infty. \tag{2.2}$$

Define

$$h(x) := g'(x). \tag{2.3}$$

In the present context, due to (2.2), $t^+ = +\infty$.

Not all positive convex functions g's satisfying (2.2) are adapted to our purpose. We follow the line of Juszczak and Nagaev [16] to describe the assumed regularity conditions of h. See also [1] for somehow similar conditions.

We firstly assume that the function h, which is a positive function defined on $\mathbb{R}_+$, is either regularly or rapidly varying in a neighborhood of infinity; the function h is strictly monotone and, by (2.2), $h(x) \to \infty$ when $x \to \infty$.

$RV(\alpha)$ designates the class of regularly varying functions of index α defined on $\mathbb{R}_+$. The function

$$\psi(t) := h^{\leftarrow}(t)$$

designates the inverse of h. Hence ψ is strictly monotone for large t and $\psi(t) \to \infty$ when $t \to \infty$.

The two cases considered for h, the regularly varying case and the rapidly varying case, are described below.

The Regularly varying case. It will be assumed that h belongs to the subclass of $RV(\beta)$, $\beta > 0$, with

$$h(x) = x^{\beta} l(x),$$

where the Karamata form of the slowly varying function l takes the form

$$l(x) = c \exp \int_1^x \frac{\epsilon(u)}{u} du \tag{2.4}$$

for some positive c. We assume that $x \mapsto \epsilon(x)$ is twice differentiable and satisfies

$$\begin{cases} \epsilon(x) \underset{x\to\infty}{=} o(1), \\ x|\epsilon'(x)| \underset{x\to\infty}{=} O(1), \\ x^2|\epsilon^{(2)}(x)| \underset{x\to\infty}{=} O(1). \end{cases} \tag{2.5}$$

It will also be assumed that when $\beta \neq 1$, then

$$|h^{(2)}(x)| \in RV(\theta) \tag{2.6}$$

where θ is a real number such that $\theta \leq \beta - 2$.

When $\beta = 1$ and $|\epsilon'(t)| \in RV(\gamma)$ then it will be assumed that $|h^{(2)}(x)| \in RV(\gamma)$; taking $\epsilon'(t) = 0$ yields $\epsilon(t) = 0$, from which $h(x) = cx$, which is to say that p has Gaussian tail, when its support is $\mathbb{R}$, or $h(x) \sim cx$ when its support is $\mathbb{R}_+$.

Remark 1. *Under (2.4) and (2.6), when $\beta \neq 1$ then $\theta = \beta - 2$. A sufficient condition for the last assumption (2.6) is that $|\epsilon'(t)| \in RV(\gamma)$, for some $\gamma < -1$. Also in this case when $\beta = 1$, then $\theta = \gamma$.*

Example 1. Weibull density. *Let p be a Weibull density with shape parameter $k > 1$ and scale parameter 1, namely*

$$\begin{aligned} p(x) &= kx^{k-1} \exp(-x^k), \qquad x \geq 0 \\ &= k \exp(-(x^k - (k-1)\log x)). \end{aligned}$$

Take $g(x) = x^k - (k-1)\log x$ and $q(x) = \log k$. Then it holds

$$h(x) = kx^{k-1} - \frac{k-1}{x} = x^{k-1}\left(k - \frac{k-1}{x^k}\right).$$

Set $l(x) = k - (k-1)/x^k, x \geq 1$, *which verifies*

$$l'(x) = \frac{k(k-1)}{x^{k+1}} = \frac{l(x)\epsilon(x)}{x}$$

with

$$\epsilon(x) = \frac{k(k-1)}{kx^k - (k-1)}.$$

The function $\epsilon(x)$ *satisfies the three conditions in (2.5).*

The Rapidly varying case. The function g in (2.1) is assumed to grow to infinity faster than any polynomial function. This is captured through a rapid growth condition on its derivative $h = g'$ which we assume to satisfy

$$h^{\leftarrow}(t) = \psi(t) \in RV(0).$$

Such a condition indeed characterizes the so called Rapid Variation class KR_∞ (see Theorem 2.4.7 in [4]), which is adequate for the description of the leading term of the function $\log p(x)$ in the upper tail; additional regularity conditions are defined assuming

$$\psi(t) = c \exp \int_1^t \frac{\epsilon(u)}{u} du, \tag{2.7}$$

for some positive c, and $t \mapsto \epsilon(t)$ is twice differentiable with

$$\begin{cases} \epsilon(t) \underset{t\to\infty}{=} o(1), \\ \frac{t\epsilon'(t)}{\epsilon(t)} \underset{t\to\infty}{\longrightarrow} 0, \\ \frac{t^2\epsilon^{(2)}(t)}{\epsilon(t)} \underset{t\to\infty}{\longrightarrow} 0. \end{cases} . \tag{2.8}$$

Example 2. A rapidly varying density. *Define* p *through*

$$p(x) = c \exp(-e^{x-1}), x \geq 0.$$

Then $g(x) = h(x) = e^{x-1}$ *and* $q(x) = 0$ *for all non negative* x. *It holds* $\psi(t) = \log t + 1$. *Since* $\psi'(t) = 1/t$, *let* $\epsilon(t) = 1/(\log t + 1)$ *so that* $\psi'(t) = \psi(t)\epsilon(t)/t$. *Then the three conditions in (2.8) are satisfied. Thus* $\psi(t) \in RV(0)$ *and* $h(x)$ *belongs to* KR_∞.

Denote by $\mathcal{R}$ the class of functions with either regular variation or with rapid variation defined above.

In the context of the present results the main argument which allows for the analysis of the extreme behavior of S_n/n lies in a description of the function $\left(\mu_3/s^3\right)(t)$ as t tends to infinity (Corollary 1 hereunder). This will be handled making use of a sharp Abelian result pertaining to the moment generating function of the r.v. X with density $p(x)$ defined in (2.1); the appropriate result (Theorem 3.1 in [5]) holds in the case when $p(x) = \exp(-g(x))$,with the condition $h \in \mathcal{R}$, hence when $q(x)$ is the null function; when $q(x)$ is constant, it can be incorporated in the function g. A straightforward extension holds when the bounded function q satisfies mild regularity conditions, making p nearly log-concave in its tail. We assume that

$$|q(x)| \in RV(\eta), \text{ for some } \eta < \theta - \tfrac{3\beta}{2} - \tfrac{3}{2} \text{ if } h \in RV(\beta) \tag{2.9}$$

and

$$|q(\psi(t))| \in RV(\eta), \text{ for some } \eta < -\tfrac{1}{2} \text{ if } h \text{ is rapidly varying.} \tag{2.10}$$

We will make use of the following extension of Theorem 3.1 in [5].

Theorem 1. *Let $p(x)$ be defined as in (2.1) and $h(x)$ belong to $\mathcal{R}$. Assume further (2.9) and (2.10). Denote by $m(t)$, $s^2(t)$ and $\mu_j(t)$ for $j = 3, 4, ...$ the functions defined in (1.2), (1.3) and (1.4). Then it holds*

$$\begin{aligned}
m(t) &\underset{t\to\infty}{=} \psi(t)(1+o(1)),\\
s^2(t) &\underset{t\to\infty}{=} \psi'(t)(1+o(1)),\\
\mu_3(t) &\underset{t\to\infty}{=} \psi^{(2)}(t)(1+o(1)),\\
\mu_j(t) &\underset{t\to\infty}{=} \begin{cases} M_j s^j(t)(1+o(1)), & \text{for even } j>3\\ \frac{(M_{j+3}-3jM_{j-1})\mu_3(t)s^{j-3}(t)}{6}(1+o(1)), & \text{for odd } j>3 \end{cases},
\end{aligned}$$

where M_i, $i > 0$, denotes the ith order moment of standard normal distribution.

Corollary 1. *Let $p(x)$ be defined as in (2.1) and $h(x) \in \mathcal{R}$. Assume further (2.9) and (2.10). Then as $t \to \infty$*

$$\frac{\mu_3(t)}{s^3(t)} \longrightarrow 0.$$

Proof. In the regularly varying case this follows from Corollaries 1 and 2 in [5], and in the rapidly varying case from Corollary 3 and Lemma 3 in [5], with the mentioned extension taking into account of the properties of the function q. □

Our results require an extension of the classical Edgeworth expansions to some specific triangular arrays; this is the scope of the following section.

3. Edgeworth expansion under extreme normalizing factors

With π^{a_n} defined through

$$\pi^{a_n}(x) = \frac{e^{tx}p(x)}{\Phi(t)},$$

and t determined by $m(t) = a_n$ together with $s^2 := s^2(t)$ define the normalized density of π^{a_n} by

$$\bar{\pi}^{a_n}(x) = s\pi^{a_n}(sx + a_n).$$

Denote the n-convolution of $\bar{\pi}^{a_n}(x)$ by $\bar{\pi}_n^{a_n}(x)$. Denote by ρ_n its normalized density

$$\rho_n(x) := \sqrt{n}\bar{\pi}_n^{a_n}(\sqrt{n}x).$$

The following result extends the local Edgeworth expansion of the distribution of normalized sums of i.i.d. r.v.'s to the present context, where the summands are generated under the density $\bar{\pi}^{a_n}$. Therefore the setting is that of a triangular array of row-wise independent summands; the fact that $a_n \to \infty$ makes the situation unusual. We mainly adapt Feller's proof (Chapiter 16, Theorem 2 [11]). However this variation on the classical Edgeworth expansion result requires some additional regularity assumption, which meets the requirements of Theorem 1, which are fulfilled in most models dealing with extremes and convolutions. Those are captured in cases when the density p is log-concave, or nearly log concave in the upper tail.

Similar conditions are considered in [6]. For notational simplicity let $s := s(t)$ and $\mu_3 := \mu_3(t)$.

Theorem 2. *Under the conditions stated in Corollary 1 , uniformly upon x it holds*

$$\rho_n(x) = \phi(x)\Big(1 + \frac{\mu_3}{6\sqrt{n}s^3}\big(x^3 - 3x\big)\Big) + o\Big(\frac{1}{\sqrt{n}}\Big).$$

where $\phi(x)$ is standard normal density.

4. Gibbs conditional principles under extreme events

4.1. **A local result.** Define t through

$$m := m(t) := a_n \tag{4.1}$$

and set

$$\begin{aligned} s &:= s(t) \\ \mu_3 &:= \mu_3(t) \end{aligned}$$

for brevity.

We consider two conditions pertaining to the growth of the sequence a_n to infinity. In the first case (moderate growth) we assume, making use of the Variance function defined in (1.6) that

$$\lim_{n\to\infty} \frac{a_n}{s\sqrt{n}} = \lim_{n\to\infty} \frac{a_n}{\sqrt{nV(a_n)}} = 0, \tag{4.2}$$

and in the second case (rapid growth) we consider sequences a_n which may grow faster to infinity, obeying

$$0 < \liminf_{n\to\infty} \frac{a_n}{\sqrt{nV(a_n)}} \leq \limsup_{n\to\infty} \frac{a_n}{\sqrt{nV(a_n)}} < \infty. \tag{4.3}$$

Remark 2. *In the Regularly varying case, i.e. when h belongs to the subclass of $RV(\beta)$, $\beta > 0$, then $V(x) = x^{1-\beta}\ l(x)$ for some slowly varying function l; see* [4] *and Lemma 4 in* [5].

Theorem 3. *When (4.2) holds then for all real number y_1,*

$$p^{a_n}(y_1) := p(X_1 = y_1 | S_1^n = na_n) = \pi^{a_n}(y_1)\left(1 + o\left(\frac{1}{\sqrt{n}}\right)\right), \tag{4.4}$$

The rate of growth defined through (4.2) is the limiting case when the natural extension of Gibbs conditional principle stated above holds. We state the following result.

Theorem 4. *Assume that the sequence a_n satisfies (4.3). Denote*

$$\alpha := t + \frac{\mu_3}{2(n-1)s^2}$$

and

$$\beta := (n-1)s^2.$$

Then for all real number y_1 ,

$$p(X_1 = y_1 | S_1^n = na_n) = g^{a_n}(y_1)(1 + o(1)) \tag{4.5}$$

where

$$g^{a_n}(y_1) := Cp(y_1)\mathfrak{n}\left(\alpha\beta + a_n, \beta, y_1\right);$$

in the above display $\mathfrak{n}\left(\mu, \sigma^2, x\right)$ *denotes the normal density with expectation* μ *and variance* σ^2 *evaluated at point* x*, and* C *is a normalizing constant.*

4.2. **On conditional independence under extreme events.** We now turn to the approximation of $p^{a_n}(y_1^k) := p^{a_n}(y_1, \ldots, y_k) = p^{a_n}(X_1^k = y_1^k \big| S_1^n = na_n)$. Denote $\mathbf{s}_i^j := y_i + \ldots + y_j$ for $i \leq j$ and $\mathbf{s}_1^0 := 0$.

We first consider the case when (4.2) holds. We then have

Proposition 1. *When (4.2) holds then for any fixed* k *and all* y_1^k*,*

$$p^{a_n}(y_1^k) = \prod_{i=1}^{k} \pi^{m_i}(y_i)\big(1 + o(1/\sqrt{n})\big)$$

where

$$m_i := m(t_i) := \frac{na_n - \mathbf{s}_1^{i-1}}{n - i + 1}. \tag{4.6}$$

We now explore the limit conditional independence of blocks of fixed length under extreme condition. As a consequence of the above Proposition 1 it holds

Theorem 5. *Under (4.2) it holds, for all* k *and all* y_1^k

$$p^{a_n}(y_1^k) = p(X_1^k = y_1^k | S_1^n = na_n) = \left(1 + o\left(\frac{1}{\sqrt{n}}\right)\right) \prod_{i=1}^{k} \pi^{a_n}(X_i = y_i). \tag{4.7}$$

Remark 3. *The above result shows that asymptotically the point condition* $(S_1^n = na_n)$ *leaves blocks of* k *of the* $X_i's$ *independent, with common density* π^{a_n}*. Obviously this property does not hold for large values of* k*, close to* n*. A similar statement holds in the LDP range, see* [10] *and* [8].

We now turn to the case when a_n moves more quickly to infinity. With m_i defined as in (4.6), with

$$\begin{aligned} s_i^2 &: = s^2(t_i) \\ \mu_{3,i} &: = \mu_3(t_i) \end{aligned}$$

and following the same arguments as developed in the proof of Theorem 4 we state

Theorem 6. *Assume that (4.3) holds. Then for all fixed* k *and all* y_1^k *it holds*

$$p(X_1^k = y_1^k | S_1^n = na_n) = \prod_{i=1}^{k} g_i(y_i)\,(1 + o\,(1)) \tag{4.8}$$

where

$$g_i(y_i) := C_i p(y_i) \mathfrak{n}\left(\alpha_i \beta_i + a_n, \beta_i, y_i\right)$$

and

$$\alpha_i := t_i + \frac{\mu_{3,i}}{2(n-i)s_i^2} \tag{4.9}$$

$$\beta_i := (n-i)s_i^2. \tag{4.10}$$

Remark 4. *The following connection between Theorem 5 and Theorem 6 holds. Under (4.2), the above result (4.8) boils down to formula (4.7). Under (4.3) and when (4.2) does not hold, the approximations obtained in Lemma 4 in the Appendix do not hold, and the approximating density cannot be stated as a product of densities under which independence holds. In that case it follows that the conditional independence property under extreme events does not hold any longer.*

4.3. **Gibbs principle in variation norm.** We now turn to a stronger approximation of P^{a_n}, the distribution on $\mathbb{R}$ with density p^{a_n}.

Denote $d_V(Q,P)$ the total variation distance between two probability measures Q and P defined on the same space.

We put forward a principle, to be illustrated in this section. Consider two sequences of equivalent distributions F_n and G_n on $\mathbb{R}^k$ with densities f_n and g_n with respect to the Lebesgue measure. Let Y_n be a r.v. with distribution F_n and assume that

$$f_n(Y_n) = g_n(Y_n)\left(1 + o_{F_n}(\varepsilon_n)\right) \tag{4.11}$$

for some sequence ε_n which tends to 0 as n tends to infinity. Then the total variation distance between F_n and G_n tends to 0 as n tends to infinity; the reason is that (4.11) implies that reciprocally

$$g_n(Y_n) = f_n(Y_n)\left(1 + o_{G_n}(\varepsilon_n)\right), \tag{4.12}$$

hence with Y_n distributed under G_n, from which it readily follows that

$$\sup_{C \in \mathcal{B}(\mathbb{R}^k)} F_n(C) - G_n(C) \to 0.$$

Indeed it holds

Lemma 1. *Suppose that for some sequence ε_n which tends to 0 as n tends to infinity, (4.11) holds. Then (4.12) holds.*

Proof. Denote

$$A_{n,\varepsilon_n} := \left\{y : (1-\varepsilon_n) g_n(y) \le f_n(y) \le g_n(y)(1+\varepsilon_n)\right\}.$$

By (4.11) it holds

$$\lim_{n\to\infty} F_n(A_{n,\varepsilon_n}) = 1.$$

Write

$$G_n(A_{n,\varepsilon_n}) = \int \mathbf{1}_{A_{n,\varepsilon_n}}(y)\, \frac{g_n(y)}{f_n(y)} f_n(y) dy.$$

Since

$$G_n(A_{n,\varepsilon_n}) \ge (1-\varepsilon_n) F_n(A_{n,\varepsilon_n})$$

it follows that

$$\lim_{n\to\infty} G_n(A_{n,\varepsilon_n}) = 1,$$

which proves the claim. □

By Lemma 1 it holds for $\delta > 0$

$$\lim_{n\to\infty} F_n(E_n) = \lim_{n\to\infty} G_n(E_n) = 1 \tag{4.13}$$

where

$$E_n := \left\{ y \in \mathbb{R} : \left| \frac{f_n(y) - g_n(y)}{g_n(y)} \right| < \delta \right\}.$$

Then

$$\begin{aligned} & \sup_{C\in\mathcal{B}(\mathbb{R}^k)} |F_n(C\cap E_n) - G_n(C\cap E_n)| \\ \leq \quad & \delta \sup_{C\in\mathcal{B}(\mathbb{R}^k)} \int_{C\cap E_n} g_n(y)\,dy \leq \delta. \end{aligned}$$

By (4.13)

$$\sup_{C\in\mathcal{B}(\mathbb{R}^k)} |F_n(C\cap E_n) - G_n(C)| < \eta_n;$$

also

$$\sup_{C\in\mathcal{B}(\mathbb{R}^k)} |G_n(C\cap E_n) - G_n(C)| < \eta_n,$$

for the same sequence $\eta_n \to 0$; hence

$$\sup_{C\in\mathcal{B}(\mathbb{R}^k)} |F_n(C) - G_n(C)| < \delta + 2\eta_n$$

for all positive δ. We have proved the following result

Proposition 2. *Under the above assumption (4.13)*

$$\lim_{n\to\infty} d_V(F_n, G_n) = 0.$$

Consider Y_1 a r.v. with density p^{a_n}, and the random variable $p^{a_n}(Y_1) := p(X_1 = Y_1 | S_1^n = na_n)$. Denote P^{a_n} , Π^{a_n} and G^{a_n} the probability measures with respective densities p^{a_n}, π^{a_n} and g^{a_n}.

Now local results in Theorems 3 and 4 hold as far as y_1 in formulas (4.5) and (4.4) satisfy $y_1 = O(a_n)$; see Remarks 7 and 8 in in Section 6. Such is also the case for Y_1 when sampled under P^{a_n}, as stated below. Thus we may substitute the local result by the present statement, with the subsequent consequences.

Making use of Theorem 3 or Theorem 4 and substituting y_1 by Y in formula (6.11) in Section 6.2, defining a sequence $\epsilon_n := \sup\left(\frac{\mu^3}{s^3}\left(\frac{a_n}{s\sqrt{n}}\right)^3, \frac{1}{\sqrt{n}}\right)$ which tends to 0 under either (4.2) or (4.3) by Corollary 1 it then holds

Proposition 3. *(i) When (4.2) holds then*

$$p^{a_n}(Y_1) = \pi^{a_n}(Y_1)\,(1 + o_{P^{a_n}}(\epsilon_n))$$

where π^{a_n} is the tilted density at point a_n,

(ii)When (4.3) holds then, with t_n such that $m(t_n) = a_n$, $\alpha := \alpha_n$ and $\beta := \beta_n$

$$p^{a_n}(Y_1) = g^{a_n}(Y_1)(1 + o_{P^{a_n}}(\epsilon_n)).$$

Indeed it holds $Y_1 = O_{P^{a_n}}(a_n)$ since by Markov Inequality

$$P(Y_1 > u | S_1^n = na_n) \leq \frac{E(Y_1 | S_1^n = na_n)}{u} = \frac{a_n}{u}.$$

Since $Y_1 = O_{P^{a_n}}(a_n)$, substituting y_1 with Y_1 in Theorems 3 and 4 the proof of Proposition 3 is completed.

Making use of Lemma 1, Proposition 3 and Proposition 2 we obtain

Theorem 7. *Under (4.2), it holds*

$$\lim_{n\to\infty} d_V\left(P^{a_n}, \Pi^{a_n}\right) = 0$$

When (4.3) holds then

$$\lim_{n\to\infty} d_V\left(P^{a_n}, G^{a_n}\right) = 0$$

Remark 5. *This result is to be paralleled with Theorem 1.6 in Diaconis and Freedman* [10] *and Theorem 2.15 in Dembo and Zeitouni* [9] *which provide a rate for this convergence in the LDP range.*

4.4. **The asymptotic location of X_1 under the conditioned distribution.** This paragraph intends to provide some insight on the behavior of X_1 under the condition $(S_1^n = na_n)$. We concentrate on the growth condition (4.2) under which the natural extension of Gibbs conditional principle holds.

Let $\mathcal{X}_t$ be a r.v. with density π^{a_n} where $m(t) = a_n$ and a_n satisfies (4.2). Recall that $E\mathcal{X}_t = a_n$ and $Var\mathcal{X}_t = s^2$. The moment generating function of the normalized variable $(\mathcal{X}_t - a_n)/s$ satisfies

$$\log E[\exp(\lambda(\mathcal{X}_t - a_n)/s)] = -\lambda a_n/s + \log\Phi\left(t + \frac{\lambda}{s}\right) - \log\Phi(t).$$

A second order Taylor expansion in the above display yields

$$\log E[\exp(\lambda(\mathcal{X}_t - a_n)/s)] = \frac{\lambda^2}{2}\frac{s^2\left(t + \frac{\theta\lambda}{s}\right)}{s^2}$$

where $\theta = \theta(t,\lambda) \in (0,1)$. The proof of the following Lemma is deferred to Section 6. It holds

Lemma 2. *Under the above hypotheses and notation, for any compact set K,*

$$\lim_{n\to\infty}\sup_{u\in K}\frac{s^2\left(t + \frac{u}{s}\right)}{s^2} = 1.$$

Applying the above Lemma it follows that the normalized r.v's $(\mathcal{X}_t - a_n)/s$ converge to a standard normal variable $N(0,1)$ in distribution, as $n \to \infty$. This amounts to say that

$$\mathcal{X}_t = a_n + sN(0,1) + o_{\Pi^{a_n}}(1). \tag{4.14}$$

which implies that $\mathcal{X}_t$ concentrates around a_n with rate s. Due to Theorem 7 the same holds for X_1 under $(S_1^n = na_n)$. The behavior of the function s determines the asymptotic order of magnitude of X_1 around a_n. When $h \in RV(\beta)$,making use of Theorem 1 it is readily seen that when $\beta > 1$ then $s(t) \to 0$ as $t \to \infty$; when $\beta < 1$ then $s(t) \to \infty$; when $\beta = 1$ then $s(t)$ may be bounded away from 0 and infinity (for example, $s(t)$ is a constant in the Gaussian tail case), but may also tend to 0 or infinity as $t \to \infty$. When $h \in KR_\infty$ then $s(t) \to 0$ as $t \to \infty$.

5. ED UNDER EXCEEDANCE

5.1. **Approximation of the conditional distribution under exceedance.** We now consider conditioning events of the form $A_n = \{S_1^n > na_n\}$. We prove that when (4.2) holds then the variation distance between the distribution of X_1 given A_n and Π^{a_n} (resp. G^{a_n}, when (4.3) holds) tends to 0, extending the classical result valid for fixed $a_n = a$.

Denote $p^{A_n}(y) = p(X_1 = y | S_1^n > na_n)$ and P^{A_n} the corresponding probability measure.

Theorem 8. *When h belongs to $\mathcal{R}$ and a_n satisfies (4.2) it holds*

$$\lim_{n\to\infty} d_V\left(P^{A_n}, \Pi^{a_n}\right) = 0;$$

when (4.3) holds then

$$\lim_{n\to\infty} d_V\left(P^{A_n}, G^{a_n}\right) = 0.$$

Remark 6. *A detailed proof of Theorem 8 under condition (4.2) is provided in the Appendix; we omit the proof under condition (4.3) for sake of brevity.*

5.2. **Location property** (1.1) **vs Gibbs conditional result under exceedance.** Let us consider the relation between (1.1) and the Gibbs conditional principle under exceedance in the case when (4.2) holds and h belongs to $RV(\beta)$ for some positive β. By (4.14) For all ε_n such that $s(t_n) = o(\varepsilon_n)$ as $n \to \infty$, where $m(t_n) = a_n$, it holds

$$P^{A_n}\left(X_1 \in (a_n - \varepsilon_n, a_n + \varepsilon_n)\right) \to 1, \tag{5.1}$$

following from Theorem 8 and (4.14). When $\beta < 1$ then $s(t_n) \to \infty$ and there exists ε_n which satisfies jointly $s(t_n) = o(\varepsilon_n)$ and $\lim_{n\to\infty} \varepsilon_n/a_n = 0$. When $\beta > 1$, then $s(t_n) \to 0$ and we also may find $\varepsilon_n \to 0$ satisfying (5.1). The case when $\beta = 1$ also yields ((5.1)) for adequate sequence $\varepsilon_n \to \infty$. Therefore Theorem 8 implies a marginal form of (1.1).

However for $h \in RV(\beta)$ and when condition (2.9) holds, and considering sequences a_n with moderate growth (4.2) then property (1.1) and Gibbs conditional principle under exceedance cannot coexist. Indeed standard calculation together with the results stated in Remark 2 yields

$$\frac{a_n}{s\sqrt{n}} = \frac{a_n^{\frac{\beta+1}{2}}}{\sqrt{n}}\left(1 + o(1)\right) \tag{5.2}$$

which should tend to 0 in order to fulfill the Gibbs conditional result under exceedance. Now (1.1) holds with some sequence ε_n satisfying

$$\lim_{n\to\infty} \frac{\varepsilon_n}{a_n} = 0$$

when

$$\lim\inf_{n\to\infty} \frac{a_n}{n^\delta} > 0 \tag{5.3}$$

for some $\delta > 1/(\beta+1)$. Inserting this latest requirement in (5.2) yields inconsistency. Therefore Gibbs conditional result with approximating distribution Π^{a_n} does not yield (1.1) when a_n moves to infinity too slowly but only (5.1).

Consider now the more complex case when h is a function rapidly growing at infinity. Assume that the function ψ' belongs to $RV(-1)$ (recall that $\psi := h^{\longleftarrow}$ is a slowly varying function at $+\infty$). Consider Condition (4.2) firstly. Making use of Theorem 1 in [6] , this latest condition is equivalent to

$$a_n h(a_n)/\sqrt{n} \to 0 \tag{5.4}$$

as n tends to infinity, which certainly holds for sequences $a_n \to \infty$ with slow increase; by Theorem 1 in [6], $m(t)$ is slowly varying at infinity and is asymptotically equivalent to $h^{\longleftarrow}$, and (6.20),(6.21) and (6.22) in Section 6 hold for some suitable

sequence η_n. Therefore with the above assumptions pertaining to the model, Gibbs conditional result holds. We now turn to (1.1) in this case. Conditions (10) and (11) in [6] should be fulfilled. Now (11) in [6] holds when $1/g$ is a self neglecting function at infinity. Recall that (10) in [6] writes

$$na_n \frac{\log g(a_n)}{h(a_n)} = O(1). \tag{5.5}$$

With h in KR_∞ it holds that $x \to x\frac{\log g(x)}{h(x)}$ tends to 0 as x tends to infinity. Therefore we may tune a_n with slow increase such that (5.4) holds. Since (5.5) is fulfilled for slow a_n we may make (5.4) and (5.5) match.

Let h belong to $RV(\beta)$ and a_n with controlled rapid growth captured by condition (4.3). Then making use of Theorem 1

$$\liminf_{n\to\infty} \frac{a_n}{n^{1/(1+\beta)}} > 0 \text{ and } \limsup_{n\to\infty} \frac{a_n}{n^{1/(1+\beta)}} < \infty. \tag{5.6}$$

Assume that for some $0 < \eta < 1$ it holds

$$\liminf_{n\to\infty} \frac{a_n}{n^{\eta}} > 0 \text{ and } \limsup_{n\to\infty} \frac{a_n}{n^{\eta}} < \infty$$

then whenever $\beta > (1-\eta)/\eta$ both (5.3) and (5.6) hold. Condition (6.20) in Section 6 is fulfilled so that (1.1) and the Gibbs conditional principle under exceedance simultaneously hold, with conditional distribution approximated by G^{a_n}.

As in Section 4.4 the behavior of the sequence ε_n relies on the function s at infinity

6. Appendix

6.1. **Proof of Theorem 2.** We state a preliminary Lemma, whose role is to provide some information on the characteristic function of the normalized random variable $(\mathcal{X}_t - m(t))/s(t)$ with density $\widetilde{\pi}_t$ defined by

$$\widetilde{\pi}_t(x) := \frac{s(t)\exp t\,(s(t)x + m(t))\,p(s(t)x + m(t))}{\phi(t)} \tag{6.1}$$

as $t \to \infty$. The density p satisfies the hypotheses in Section 2. Denote $\varphi^{a_n}(u) := \int e^{iux}\widetilde{\pi}_t(x)dx$ the characteristic function of $(\mathcal{X}_t - m(t))/s(t)$. Then

Lemma 3. *Assume that there exists c_1, c_2 both positive such that for all t*

$$\widetilde{\pi}_t(x) > c_1 \text{ for } |x| < c_2 \tag{6.2}$$

then under the hypotheses stated in Section 2, for any $c > 0$ there exists $\rho < 1$ such that

$$|\varphi^{a_n}(u)| \le \rho \tag{6.3}$$

for $|u| > c$ and all a_n.

Proof. See [13], p150. □

We now turn to the Proof of Theorem 2.

We denote $\widetilde{\pi}_t$ the normalized conjugate density of $p(x)$ and ρ_n is the normalized $n-$fold convolution of $\widetilde{\pi}_t$. We consider the triangular array whose $n-$th row consists in n i.i.d. copies of a r.v. with standardized density $\widetilde{\pi}_t$ and the sum of the row, divided by $\sqrt{n}$, has density ρ_n. The standard Gaussian density is denoted ϕ. The c.f. of $\widetilde{\pi}_t$ is denoted φ^{a_n} so that the c.f. of ρ_n is $\left(\varphi^{a_n}(.)\right)^n$, and $m(t) = a_n$.

Step 1: Let

$$G(x) := \rho_n(x) - \phi(x) - \frac{\mu_3}{6\sqrt{n}s_n^3}\left(x^3 - 3x\right)\phi(x),$$

for which we provide an upper bound.

It holds

$$\phi'''(x) = -(x^3 - 3x)\phi(x),$$

which gives

$$(x^3 - 3x)\phi(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}(i\tau)^3 e^{-i\tau x}e^{-\frac{1}{2}\tau^2}d\tau. \tag{6.4}$$

By Fourier inversion

$$\rho_n(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\tau x}\left(\varphi^{a_n}(\tau/\sqrt{n})\right)^n d\tau. \tag{6.5}$$

Using (6.4) and (6.5), we have

$$|G(x)| \leq \frac{1}{2\pi}\int_{-\infty}^{\infty}\left|\left(\varphi^{a_n}(\tau/\sqrt{n})\right)^n - e^{-\frac{1}{2}\tau^2} - \frac{\mu_3}{6\sqrt{n}s^3}(i\tau)^3 e^{-\frac{1}{2}\tau^2}\right|d\tau.$$

Step 2: In this step, we show that for large n, the characteristic function φ^{a_n} satisfies

$$\int |\varphi^{a_n}(\tau)|^2 d\tau < \infty.$$

Indeed by Parseval identity

$$\int |\varphi^{a_n}(\tau)|^2 d\tau = 2\pi\int(\widetilde{\pi}_t(x))^2 dx \leq 2\pi\sup_{x\in\mathbb{R}}\widetilde{\pi}_t(x) < \infty$$

where we used Theorem 5.4 of Nagaev [16] which states that $\widetilde{\pi}_t(x)$ satisfies

$$\lim_{a_n\to\infty}\sup_{x\in\mathbb{R}}|\widetilde{\pi}_t(x) - \varphi(x)| = 0.$$

Step 3: In this step, we complete the proof by showing that when $n \to \infty$

$$\int_{-\infty}^{\infty}\left|\left(\varphi^{a_n}(\tau/\sqrt{n})\right)^n - e^{-\frac{1}{2}\tau^2} - \frac{\mu_3}{6\sqrt{n}s^3}(i\tau)^3 e^{-\frac{1}{2}\tau^2}\right|d\tau = o\Big(\frac{1}{\sqrt{n}}\Big). \tag{6.6}$$

The LHS in (6.6) is splited on $|\tau| > \omega\sqrt{n}$ and on $|\tau| \leq \omega\sqrt{n}$. It holds

$$\begin{aligned}
&\sqrt{n}\int_{|\tau|>\omega\sqrt{n}}\left|\left(\varphi^{a_n}(\tau/\sqrt{n})\right)^n - e^{-\frac{1}{2}\tau^2} - \frac{\mu_3}{6\sqrt{n}s^3}(i\tau)^3 e^{-\frac{1}{2}\tau^2}\right|d\tau \\
&\leq \sqrt{n}\int_{|\tau|>\omega\sqrt{n}}\left|\left(\varphi^{a_n}(\tau/\sqrt{n})\right)\right|^n d\tau + \sqrt{n}\int_{|\tau|>\omega\sqrt{n}}\left|e^{-\frac{1}{2}\tau^2} + \frac{\mu_3}{6\sqrt{n}s^3}(i\tau)^3 e^{-\frac{1}{2}\tau^2}\right|d\tau \\
&\leq \sqrt{n}\rho^{n-2}\int_{|\tau|>\omega\sqrt{n}}\left|\left(\varphi^{a_n}(\tau/\sqrt{n})\right)\right|^2 d\tau + \sqrt{n}\int_{|\tau|>\omega\sqrt{n}} e^{-\frac{1}{2}\tau^2}\Big(1 + \Big|\frac{\mu_3\tau^3}{6\sqrt{n}s^3}\Big|\Big)d\tau \\
&=: A + B.
\end{aligned} \tag{6.7}$$

where we used Lemma 3 from the second line to the third one. Now

$$A = \exp\left(\frac{1}{2}\log n + (n-2)\log\rho + \log\int_{|\tau|>\omega\sqrt{n}}\left(\varphi^{a_n}(\tau/\sqrt{n})\right)^2 d\tau\right) \longrightarrow 0.$$

By Corollary 1 when $n \to \infty$

$$\begin{aligned} B &\le \sqrt{n}\int_{|\tau|>\omega\sqrt{n}} e^{-\frac{1}{2}\tau^2}|\tau|^3 d\tau = \sqrt{n}\int_{|\tau|>\omega\sqrt{n}} \exp\Big\{-\frac{1}{2}\tau^2 + 3\log|\tau|\Big\}d\tau \\ &= 2\sqrt{n}\exp\big(-\omega^2 n/2 + o(\omega^2 n/2)\big) \longrightarrow 0, \end{aligned}$$

where the second equality holds from, for example, Chapiter 4 of [4]. Summing up, when $n \to \infty$

$$\int_{|\tau|>\omega\sqrt{n}} \Big|\big(\varphi^{a_n}(\tau/\sqrt{n})\big)^n - e^{-\frac{1}{2}\tau^2} - \frac{\mu_3}{6\sqrt{n}s^3}(i\tau)^3 e^{-\frac{1}{2}\tau^2}\Big| d\tau = o\Big(\frac{1}{\sqrt{n}}\Big).$$

If $|\tau| \le \omega\sqrt{n}$, it holds

$$\begin{aligned} &\int_{|\tau|\le\omega\sqrt{n}} \Big|\big(\varphi^{a_n}(\tau/\sqrt{n})\big)^n - e^{-\frac{1}{2}\tau^2} - \frac{\mu_3}{6\sqrt{n}s^3}(i\tau)^3 e^{-\frac{1}{2}\tau^2}\Big| d\tau \\ &= \int_{|\tau|\le\omega\sqrt{n}} e^{-\frac{1}{2}\tau^2}\Big|\exp\Big\{n\log\varphi^{a_n}(\tau/\sqrt{n}) + \frac{1}{2}\tau^2\Big\} - 1 - \frac{\mu_3}{6\sqrt{n}s^3}(i\tau)^3\Big| d\tau. \end{aligned} \tag{6.8}$$

We make use of the Inequality

$$|e^\alpha - 1 - \beta| = |(e^\alpha - e^\beta) + (e^\beta - 1 - \beta)| \le (|\alpha - \beta| + \frac{1}{2}\beta^2)e^\gamma,$$

where $\gamma \ge \max(|\alpha|, |\beta|)$ in the integrand in the last display. Denote

$$\gamma(\tau) = \log\varphi^{a_n}(\tau) + \frac{1}{2}\tau^2.$$

Since $\gamma'(0) = \gamma''(0) = 0$, the third order Taylor expansion of $\gamma(\tau)$ at $\tau = 0$ yields

$$\gamma(\tau) = \gamma(0) + \gamma'(0)\tau + \frac{1}{2}\gamma''(0)\tau^2 + \frac{1}{6}\gamma'''(\xi)\tau^3 = \frac{1}{6}\gamma'''(\xi)\tau^3,$$

where $0 < \xi < \tau$. Hence it holds

$$\Big|\gamma(\tau) - \frac{\mu_3}{6s^3}(i\tau)^3\Big| = \Big|\gamma'''(\xi) - \frac{\mu_3}{s_n^3}i^3\Big|\frac{\tau^3}{6}.$$

Here γ''' is continuous; also as $n \to \infty$, making use of Corollary 1 $|\gamma'''(0)| \to 0$. Thus for any small $\eta > 0$, we can choose ω small enough such that $|\gamma'''(\xi)| < \eta$ for $|\tau| < \omega$. Making use of Corollary 1 again, for n large enough

$$\Big|\gamma(\tau) - \frac{\mu_3}{6s^3}(i\tau)^3\Big| \le \Big(|\gamma'''(\xi)| + \eta\Big)\frac{|\tau|^3}{6} < \eta\tau^3. \tag{6.9}$$

Choose ω small enough, such that for n large enough it holds for $|\tau| < \omega$

$$\max\Big(\Big|\frac{\mu_3}{6s^3}(i\tau)^3\Big|, |\gamma(\tau)|\Big) \le \frac{1}{4}\tau^2.$$

Replacing τ by $\tau/\sqrt{n}$, it holds for $|\tau| < \omega\sqrt{n}$, and using (6.9)

$$\begin{aligned} &\Big|n\log\varphi^{a_n}(\tau/\sqrt{n}) + \frac{1}{2}\tau^2 - \frac{\mu_3}{6\sqrt{n}s^3}(i\tau)^3\Big| \\ &= n\Big|\gamma\Big(\frac{\tau}{\sqrt{n}}\Big) - \frac{\mu_3}{6s^3}\Big(\frac{i\tau}{\sqrt{n}}\Big)^3\Big| < \frac{\eta|\tau|^3}{\sqrt{n}}. \end{aligned}$$

In a similar way, it also holds for $|\tau| < \omega\sqrt{n}$

$$\begin{aligned}&\max\Big(\Big|n\log\varphi^{a_n}(\tau/\sqrt{n})+\frac{1}{2}\tau^2\Big|,\Big|\frac{\mu_3}{6\sqrt{n}s^3}(i\tau)^3\Big|\Big)\\&=n\max\Big(\Big|\gamma\Big(\frac{\tau}{\sqrt{n}}\Big)\Big|,\Big|\frac{\mu_3}{6s^3}\Big(\frac{i\tau}{\sqrt{n}}\Big)^3\Big|\Big)\le\frac{1}{4}\tau^2.\end{aligned}$$

Turn to the integrand in (6.8). We then for $|\tau| < \omega\sqrt{n}$

$$\begin{aligned}&\Big|\exp\Big\{n\log\varphi^{a_n}(\tau/\sqrt{n})+\frac{1}{2}\tau^2\Big\}-1-\frac{\mu_3}{6\sqrt{n}s^3}(i\tau)^3\Big|\\&\le\Big(\Big|n\log\varphi^{a_n}(\tau/\sqrt{n})+\frac{1}{2}\tau^2-\frac{\mu_3}{6\sqrt{n}s^3}(i\tau)^3\Big|+\frac{1}{2}\Big|\frac{\mu_3}{6\sqrt{n}s^3}(i\tau)^3\Big|^2\Big)\\&\qquad\times\exp\Big[\max\Big(\Big|n\log\varphi^{a_n}(\tau/\sqrt{n})+\frac{1}{2}\tau^2\Big|,\Big|\frac{\mu_3}{6\sqrt{n}s^3}(i\tau)^3\Big|\Big)\Big]\\&\le\Big(\frac{\eta|\tau|^3}{\sqrt{n}}+\frac{1}{2}\Big|\frac{\mu_3}{6\sqrt{n}s^3}(i\tau)^3\Big|^2\Big)\exp\Big(\frac{\tau^2}{4}\Big)\\&=\Big(\frac{|\tau|^3}{\sqrt{n}}+\frac{\mu_3^2\tau^6}{72ns^6}\Big)\exp\Big(\frac{\tau^2}{4}\Big).\end{aligned}$$

Use this upper bound to obtain

$$\begin{aligned}&\int_{|\tau|\le\omega\sqrt{n}}\Big|\big(\varphi^{a_n}(\tau/\sqrt{n})\big)^n-e^{-\frac{1}{2}\tau^2}-\frac{\mu_3}{6\sqrt{n}s^3}(i\tau)^3e^{-\frac{1}{2}\tau^2}\Big|d\tau\\&\le\int_{|\tau|\le\omega\sqrt{n}}\exp\Big(-\frac{\tau^2}{4}\Big)\Big(\frac{\eta|\tau|^3}{\sqrt{n}}+\frac{\mu_3^2\tau^6}{72ns^6}\Big)d\tau\\&=\frac{\eta}{\sqrt{n}}\int_{|\tau|\le\omega\sqrt{n}}\exp\Big(-\frac{\tau^2}{4}\Big)|\tau|^3d\tau+\frac{\mu_3^2}{72ns^6}\int_{|\tau|\le\omega\sqrt{n}}\exp\Big(-\frac{\tau^2}{4}\Big)\tau^6d\tau,\end{aligned}$$

and Corollary 1, which yields

$$\int_{|\tau|\le\omega\sqrt{n}}\Big|\big(\varphi^{a_n}(\tau/\sqrt{n})\big)^n-e^{-\frac{1}{2}\tau^2}-\frac{\mu_3}{6\sqrt{n}s^3}(i\tau)^3e^{-\frac{1}{2}\tau^2}\Big|d\tau=o\Big(\frac{1}{\sqrt{n}}\Big).$$

This gives (6.6), and therefore we obtain

$$\Big|\bar{\pi}_n^{a_n}(x)-\phi(x)-\frac{\mu_3}{6\sqrt{n}s^3}(x^3-3x)\phi(x)\Big|=o\Big(\frac{1}{\sqrt{n}}\Big),$$

which concludes the proof.

6.2. **Proof of Theorem 3.** It is well known and easily checked that the conditional density $p(X_1^k = y_1^k | S_1^n = na_n)$ is invariant under any i.i.d sampling scheme in the family of densities π^α as α belongs to $Im(X_1)$ (commonly called tilting change of measure). Namely

$$p(X_1^k = y_1^k | S_1^n = na_n) = \pi^\alpha(X_1^k = y_1^k | S_1^n = na_n)$$

where on the LHS the X_i's are sampled i.i.d. under p and on the RHS they are sampled i.i.d. under π^α.

Using Bayes formula, it thus holds

$$\begin{aligned} p(X_1 = y_1 | S_1^n = na_n) &= \pi^m(X_1 = y_1 | S_1^n = na_n) \\ &= \pi^m(X_1 = y_1) \frac{\pi^m(S_2^n = na_n - y_1)}{\pi^m(S_1^n = na_n)} \\ &= \frac{\sqrt{n}}{\sqrt{n-1}} \pi^m(X_1 = y_1) \frac{\widetilde{\pi_{n-1}}(\frac{m-y_1}{s\sqrt{n-1}})}{\widetilde{\pi_n}(0)}, \end{aligned} \tag{6.10}$$

where $\widetilde{\pi_{n-1}}$ is the normalized density of S_2^n under i.i.d. sampling with density π^{a_n}; correspondingly, $\widetilde{\pi_n}$ is the normalized density of S_1^n under the same sampling. Note that a r.v. with density π^{an} has expectation m and variance s^2. Perform a third-order Edgeworth expansion of $\widetilde{\pi_{n-1}}(z)$, using Theorem 2. Setting

$$z := \frac{m - y_1}{s\sqrt{n-1}}$$

it follows that

$$\widetilde{\pi_{n-1}}(z) = \phi(z)\Big(1 + \frac{\mu_3}{6s^3\sqrt{n-1}}(z^3 - 3z)\Big) + o\Big(\frac{1}{\sqrt{n}}\Big),$$

The approximation of $\widetilde{\pi_n}(0)$ is

$$\widetilde{\pi_n}(0) = \phi(0)\Big(1 + o\big(\frac{1}{\sqrt{n}}\big)\Big).$$

Hence (6.10) becomes

$$\begin{aligned} &p(X_1 = y_1 | S_1^n = na_n) \\ &= \frac{\sqrt{n}}{\sqrt{n-1}} \pi^m(X_1 = y_1) \frac{\phi(z)}{\phi(0)} \Big[1 + \frac{\mu_3}{6s^3\sqrt{n-1}}(z^3 - 3z) + o\Big(\frac{1}{\sqrt{n}}\Big)\Big] \\ &= \frac{\sqrt{2\pi n}}{\sqrt{n-1}} \pi^m(X = y_1) \phi(z)\big(1 + R_n + o(1/\sqrt{n})\big), \end{aligned} \tag{6.11}$$

where

$$R_n = \frac{\mu_3}{6s^3\sqrt{n-1}}(z^3 - 3z).$$

Under condition (4.2), by Corollary 1, $\mu_3/s^3 \to 0$. This yields

$$R_n = o\big(1/\sqrt{n}\big), \tag{6.12}$$

which gives

$$p(X_1 = y_1 | S_1^n = na_n) = \pi^m(X = y_1)\big(1 + o(1/\sqrt{n})\big)$$

as claimed.

Remark 7. *The rate given in (6.12) holds also when $y_1 = O(a_n)$, due to (4.2).*

6.3. **Proof of Theorem 4.** In contrast with the above proof of Theorem 3, the second summand in (6.11) does not tend to 0 any longer and contributes to the approximating density. Standard development then yields the result. When (4.2) holds instead of (4.3) then standard expansions in (4.5) provide $g^{a_n}(y_1) \sim \pi^{a_n}(y_1)$ for all y_1 as n tends to infinity.

Remark 8. *The same fact as quoted in Remark 7 holds under (4.3).*

6.4. **Proof of Proposition 1.** Denote

$$z_i := \frac{m_i - y_{i+1}}{s_i\sqrt{n-i-1}}$$

where

$$s_i^2 := s^2(t_i).$$

We first state a Lemma pertaining to the order of magnitude of z_i. The proof of this Lemma is in the next Subsection

Lemma 4. *Assume that $h(x) \in \mathcal{R}$. Let t_i be defined by (4.6). Assume that $a_n \to \infty$ as $n \to \infty$ and that (4.2) holds. Then as $n \to \infty$*

$$\lim_{n\to\infty} \sup_{0\le i\le k-1} z_i = 0, \qquad \text{and} \qquad \sup_{0\le i\le k-1} z_i^2 = o\left(\frac{1}{\sqrt{n}}\right).$$

We turn to the proof of Proposition 1.

It holds by Bayes formula,

$$p^{a_n}(y_1^k) = \prod_{i=0}^{k-1} p(X_{i+1} = y_{i+1} | S_{i+1}^n = na_n - \mathbf{s}_1^i). \tag{6.13}$$

Making use of the same arguments as in the proof of Theorem 3 for all i we get

$$p(X_{i+1} = y_{i+1} | S_{i+1}^n = na_n - \mathbf{s}_1^i) == \frac{\sqrt{n-i}}{\sqrt{n-i-1}} \pi^{m_i}(X_{i+1} = y_{i+1}) \big(1 - z_i^2/2 + o(z_i^2)\big) \left(1 + o(1/\sqrt{n})\right)\big),$$

Making use of Lemma 4 in each of the factors in (6.13) completes the proof.

6.5. **Proof of Lemma 4.** By Theorem 1, when $n \to \infty$, it holds

$$z_i \sim m_i/(s_i\sqrt{n}) \sim \frac{\psi(t_i)}{\sqrt{n\psi^{'}(t_i)}}$$

as $n \to \infty$. Since $m_i \sim m_k$ as $n \to \infty$, it holds $m_i \sim \psi(t_k)$; hence $\psi(t_i) \sim \psi(t_k)$.

Case 1: if $h(x) \in RV(\beta)$

$$h^{'}(x) = x^{\beta-1} l_0(x) \left(\beta + \epsilon(x)\right).$$

Set $x = \psi(t)$; we get

$$h^{'}\left(\psi(t)\right) = \psi(t)^{\beta-1} l_0\left(\psi(t)\right) \left(\beta + \epsilon\left(\psi(t)\right)\right).$$

Since $\psi^{'}(t) = 1/h^{'}\left(\psi(t)\right)$ we obtain

$$\frac{\psi'(t_i)}{\psi'(t_k)} = \frac{h^{'}\left(\psi(t_k)\right)}{h^{'}\left(\psi(t_i)\right)} = \frac{\left(\psi(t_k)\right)^{\beta-1} l_0\left(\psi(t_k)\right) \left(\beta + \epsilon\left(\psi(t_k)\right)\right)}{\left(\psi(t_i)\right)^{\beta-1} l_0\left(\psi(t_i)\right) \left(\beta + \epsilon\left(\psi(t_i)\right)\right)} \longrightarrow 1,$$

by the slowly varying property of l_0. Thus

$$\psi'(t_i) \sim \psi'(t_k),$$

which yields

$$z_i \sim \frac{\psi(t_k)}{\sqrt{n\psi^{'}(t_k)}}.$$

Hence we have under condition (4.2)

$$z_i^2 \sim \frac{\psi(t_k)^2}{n\psi^{'}(t_k)} = \frac{\psi(t_k)^2}{\sqrt{n}\psi^{'}(t_k)} \frac{1}{\sqrt{n}} = o\left(\frac{1}{\sqrt{n}}\right).$$

Case 2: Assume that $h(x) \in KR_\infty$. Since m is increasing and $m(t_k) \geq m(t_i)$, it holds $t_i \leq t_k$. The function $t \to \psi^{'}(t)$ is decreasing, since

$$\psi^{''}(t) = -\frac{\psi(t)}{t^2}\epsilon(t)\,(1+o(1)) < 0 \qquad as \quad t \to \infty.$$

Therefore as $n \to \infty$, it holds $\psi'(t_i) \geq \psi'(t_k) > 0$, which yields

$$z_i \sim \frac{\psi(t_i)}{\sqrt{n\psi'(t_i)}} \leq \frac{2\psi(t_k)}{\sqrt{n\psi'(t_k)}}.$$

Hence

$$z_i^2 \leq \frac{4\psi(t_k)^2}{n\psi'(t_k)} = \frac{4\psi(t_k)^2}{\sqrt{n}\psi'(t_k)}\frac{1}{\sqrt{n}} = o\left(\frac{1}{\sqrt{n}}\right),$$

where the last step holds from condition (4.2).

This closes the proof of the Lemma.

6.6. **Proof of Lemma 2. Case 1:** if $h(t) \in RV(\beta)$. By Theorem 1, $s^2 \sim \psi'(t)$ with $\psi(t) \sim t^{1/\beta}l_1(t)$, where l is some slowly varying function. Since $\psi'(t) = 1/h^{'}\big(\psi(t)\big)$,

$$\begin{aligned}\frac{1}{s^2} &\sim h^{'}\big(\psi(t)\big) = \psi(t)^{\beta-1}l_0\big(\psi(t)\big)\big(\beta+\epsilon\big(\psi(t)\big)\big)\\ &\sim \beta t^{1-1/\beta}l_1(t)^{\beta-1}l_0\big(\psi(t)\big) = o(t),\end{aligned}$$

where $l_0 \in RV(0)$. This implies for any $u \in K$ it holds $u/s = o(\sqrt{t})$,which using (2.5) yields

$$\begin{aligned}\frac{s^2\,(t+u/s)}{s^2} \sim \frac{\psi'(t+u/s)}{\psi'(t)} &= \frac{\psi(t)^{\beta-1}l_0\big(\psi(t)\big)\big(\beta+\epsilon\big(\psi(t)\big)\big)}{\big(\psi(t+u/s)\big)^{\beta-1}l_0\big(\psi(t+u/s)\big)\big(\beta+\epsilon\big(\psi(t+u/s)\big)\big)}\\ \sim \frac{\psi(t)^{\beta-1}}{\psi(t+u/s)^{\beta-1}} &\sim \frac{t^{1-1/\beta}l_1(t)^{\beta-1}}{(t+u/s)^{1-1/\beta}l_1(t+u/s)^{\beta-1}} \longrightarrow 1.\end{aligned}$$

Case 2: if $h(t) \in KR_\infty$, then

$$\frac{1}{st} \sim \frac{1}{t\sqrt{\psi'(t)}} = \sqrt{\frac{1}{t\psi(t)\epsilon(t)}} \longrightarrow 0,$$

making use of condition (2.8). Hence for any $u \in K$, we get as $n \to \infty$

$$\frac{u}{s} = o(t).$$

Thus using the slowly varying property of $\psi(t)$ we have

$$\begin{aligned}\frac{s^2\,(t+u/s)}{s^2} \sim \frac{\psi'(t+u/s)}{\psi'(t)} &= \frac{\psi(t+u/s)\epsilon(t+u/s)}{t+u/s}\frac{t}{\psi(t)\epsilon(t)}\\ &\sim \frac{\epsilon(t+u/s)}{\epsilon(t)} = \frac{\epsilon(t)+O\big(\epsilon'(t)u/s\big)}{\epsilon(t)} \longrightarrow 1,\end{aligned} \tag{6.14}$$

where we used a Taylor expansion in the second line, and where the last step holds from condition (2.8). This completes the proof.

6.7. **Proof of Theorem 5.** Making use of

$$p(X_1^k = y_1^k | S_1^n = na_n) = \prod_{i=0}^{k-1} p(X_{i+1} = y_{i+1} | S_{i+1}^n = na_n - \mathbf{s}_1^i),$$

and using the tilted density π^{a_n} instead of π^{m_i} it holds

(6.15)

$$p(X_{i+1} = y_{i+1} | S_{i+1}^n = na_n - \mathbf{s}_1^i) = \frac{\sqrt{n-i}}{\sqrt{n-i-1}} \pi^{a_n}(X_{i+1} = y_{i+1}) \frac{\widetilde{\pi_{n-i-1}}(\frac{(i+1)a_n - \mathbf{s}_1^{i+1}}{s\sqrt{n-i-1}})}{\widetilde{\pi_{n-i}}\left(\frac{ia_n - \mathbf{s}_1^i}{s\sqrt{n-i}}\right)},$$

where $\widetilde{\pi_{n-i-1}}$ is the normalized density of S_{i+2}^n under i.i.d. sampling with π^{a_n}. Correspondingly, denote $\widetilde{\pi_{n-i}}$ the normalized density of S_{i+1}^n under the same sampling. Write

$$z_i = \frac{ia_n - \mathbf{s}_1^{i-1}}{s\sqrt{n-i+1}}.$$

By Theorem 2 a third-order Edgeworth expansion yields

$$\widetilde{\pi_{n-i-1}}(z_i) = \phi(z_i)\left(1 + R_n^i\right) + o\left(\frac{1}{\sqrt{n}}\right),$$

where

$$R_n^i = \frac{\mu_3}{6s^3\sqrt{n-i-1}}(z_i^3 - 3z_i).$$

Accordingly

$$\widetilde{\pi_{n-i}}(z_{i-1}) = \phi(z_{i-1})\left(1 + R_n^{i-1}\right) + o\left(\frac{1}{\sqrt{n}}\right).$$

When $a_n \to \infty$, using Theorem 1, it holds

$$\begin{aligned} \sup_{0\le i\le k-1} z_i^2 \sim \frac{(i+1)^2 a_n^2}{s^2 n} &\le \frac{2k^2 a_n^2}{s^2 n} = \frac{2k^2 (m(t))^2}{s^2 n} \\ &\sim \frac{2k^2(\psi(t))^2}{\psi'(t)n} = \frac{2k^2(\psi(t))^2}{\sqrt{n}\psi'(t)}\frac{1}{\sqrt{n}} = o\left(\frac{1}{\sqrt{n}}\right), \end{aligned} \tag{6.16}$$

where the last step holds under condition (4.2). Hence $z_i \to 0$ for $0 \le i \le k-1$ as $a_n \to \infty$. By Corollary 1

$$R_n^i = o\left(1/\sqrt{n}\right) \text{ and } R_n^{i-1} = o\left(1/\sqrt{n}\right).$$

We thus get

$$p(X_{i+1} = y_{i+1} | S_{i+1}^n = na_n - \mathbf{s}_1^i) = \frac{\sqrt{n-i}}{\sqrt{n-i-1}} \pi^{a_n}(X_{i+1} = y_{i+1}) \frac{\phi(z_i)}{\phi(z_{i-1})}\left(1 + o(1/\sqrt{n})\right)$$

$$= \frac{\sqrt{n-i}}{\sqrt{n-i-1}} \pi^{a_n}(X_{i+1} = y_{i+1})\left(1 - (z_i^2 - z_{i-1}^2)/2 + o(z_i^2 - z_{i-1}^2)\right)\left(1 + o(1/\sqrt{n})\right),$$

where we used a Taylor expansion in the second equality. Using (6.16), we have as $a_n \to \infty$

$$|z_i^2 - z_{i-1}^2| = o(1/\sqrt{n}),$$

which yields

$$p(X_1^k = y_1^k | S_1^n = na_n) = \left(1 + o\left(\frac{1}{\sqrt{n}}\right)\right) \prod_{i=0}^{k-1} \pi^{a_n}(X_{i+1} = y_{i+1}).$$

This completes the proof.

6.8. **Proof of Theorem 8.** We first establish some asymptotics for the distribution of S_1^n. It holds, denoting

$$I(x) := xm^{\leftarrow}(x) - \log \Phi\left(m^{\leftarrow}(x)\right)$$

and recall that the Variance function $V(x)$ has been defined in (1.6).

Lemma 5. *Set $m(t) = a_n$. Suppose that $a_n \to \infty$ as $n \to \infty$. Then*

$$P(S_1^n > na_n) = \frac{\exp(-nI(a_n))}{\sqrt{2\pi}\sqrt{n}m^{\leftarrow}(a_n)V(a_n)}\left(1 + o\left(\frac{1}{\sqrt{n}}\right)\right). \tag{6.17}$$

Let further t_τ be defined by $m(t_\tau) = \tau$ with $\tau \geq a_n$. Then uniformly upon τ

$$p(S_1^n = n\tau) = \frac{\exp(-nI(\tau))}{\sqrt{2\pi}\sqrt{n}V(\tau)}\left(1 + o\left(\frac{1}{\sqrt{n}}\right)\right). \tag{6.18}$$

Proof. For a density $p(x)$ defined in as in (2.1), we show that $g(x)$ is a convex function when x is large. If $h(x) \in R_\beta$, for x large

$$g^{''}(x) = h^{'}(x) = \frac{h(x)}{x}\left(\beta + \epsilon(x)\right) > 0.$$

If $h(x) \in KR_\infty$, set $x := \psi(v)$. Then

$$g^{''}(x) = h'(x) = \frac{1}{\psi'(v)} = \frac{v}{\psi(v)\epsilon(v)} > 0,$$

where the inequality holds since $\epsilon(v) > 0$ when v is large enough. Hence $g(x)$ is convex for large x. Therefore, the density $p(x)$ with $h(x) \in \mathcal{R}$ satisfies the conditions of Theorem 6.2.1 in [13]. Denote by p_n the density of $\bar{X} = (X_1 + \ldots + X_n)/n$. We obtain from formula (2.2.6) of [13], using a third order Edgeworth expansion

$$P(S_1^n \geq na_n) = \frac{\Phi(t)^n \exp(-nta_n)}{\sqrt{n}ts(t)}\left(B_0(\lambda_n)\right) + O\left(\frac{\mu_3(t)}{6\sqrt{n}s^3(t)}B_3(\lambda_n)\right),$$

where $\lambda_n = \sqrt{n}ts(t)$; $B_0(\lambda_n)$ and $B_3(\lambda_n)$ are defined by

$$B_0(\lambda_n) = \frac{1}{\sqrt{2\pi}}\left(1 - \frac{1}{\lambda_n^2} + o\left(\frac{1}{\lambda_n^2}\right)\right), \qquad B_3(\lambda_n) \sim -\frac{3}{\sqrt{2\pi}\lambda_n}.$$

We show that as $a_n \to \infty$

$$\frac{1}{\lambda_n^2} = o\left(\frac{1}{n}\right). \tag{6.19}$$

Since $n/\lambda_n^2 = 1/(t^2s^2(t))$, (6.19) is equivalent to

$$t^2s^2(t) \longrightarrow \infty.$$

By Theorem 1, $m(t) \sim \psi(t)$ and $s^2(t) \sim \psi'(t)$; combined with $m(t) = a_n$, it holds $t \sim h(a_n)l_1(a_n)$, where l_1 is some slowly varying function. If $h \in RV(\beta)$, notice that

$$\psi'(t) = \frac{1}{h'(\psi(t))} = \frac{\psi(t)}{h\left(\psi(t)\right)\left(\beta + \epsilon(\psi(t))\right)} \sim \frac{a_n}{h(a_n)\left(\beta + \epsilon(\psi(t))\right)};$$

hence

$$t^2s^2(t) \sim h(a_n)^2 l_1(a_n)^2 \frac{a_n}{h(a_n)\left(\beta + \epsilon(\psi(t))\right)} = \frac{a_n h(a_n) l_1(a_n)^2}{\beta + \epsilon(\psi(t_n))} \longrightarrow \infty.$$

If $h \in KR_\infty$, then

$$t^2 s^2(t) \sim t^2 \frac{\psi(t)\epsilon(t)}{t} = t\psi(t)\epsilon(t) \longrightarrow \infty;$$

we have proved (6.19) and, summing up, that

$$B_0(\lambda_n) = \frac{1}{\sqrt{2\pi}} \left(1 + o\left(\frac{1}{n}\right)\right).$$

By (6.19), λ_n goes to ∞ as $a_n \to \infty$; this implies further that $B_3(\lambda_n) \to 0$. On the other hand, by Corollary 1 it holds $\mu_3/s^3 \to 0$. Hence we obtain

$$P(S_1^n \geq na_n) = \frac{\Phi(t)^n \exp(-nta_n)}{\sqrt{2\pi n} t s(t)} \left(1 + o\left(\frac{1}{\sqrt{n}}\right)\right),$$

which gives (6.17). By Jensen's Theorem 6.2.1 ([13]) and formula (2.2.4) in [13] it follows that uniformly in τ

$$p(S_1^n/n = \tau) = \frac{\sqrt{n}\Phi(t_\tau)^n \exp(-nt_\tau\tau)}{\sqrt{2\pi} s(t_\tau)} \left(1 + o\left(\frac{1}{\sqrt{n}}\right)\right),$$

which, together with $p(S_1^n = n\tau) = (1/n)p(S_1^n/n = \tau)$, gives (6.18). □

We now turn to the pointwise approximation of $p^{A_n}(y_1)$. Let η_n be a positive sequence and denote

$$\eta_n/a_n \longrightarrow 0 \tag{6.20}$$

$$nm^{\leftarrow}(a_n)\eta_n \longrightarrow \infty \tag{6.21}$$

and

$$\eta^2/V(a_n) \to 0. \tag{6.22}$$

It holds

$$\begin{aligned}
p^{A_n}(y_1) &= \int_{a_n}^{\infty} p(X_1 = y_1 | S_1^n = n\tau) p(S_1^n = n\tau | S_1^n \geq na_n) d\tau \\
&= \frac{p(X_1 = y_1)}{P(S_1^n \geq na_n)} \int_{a_n}^{\infty} p(S_2^n = n\tau - y_1) d\tau \\
&= \left(1 + \frac{P_2}{P_1}\right) \frac{p(X_1 = y_1)}{P(S_1^n \geq na_n)} \int_{a_n}^{a_n+\eta_n} p(S_2^n = n\tau - y_1) d\tau \\
&= \left(1 + \frac{P_2}{P_1}\right) \int_{a_n}^{a_n+\eta_n} p(X_1 = y_1 | S_1^n = n\tau) p(S_1^n = n\tau | S_1^n \geq na_n) d\tau \qquad (6.23)
\end{aligned}$$

where the second equality is obtained by Bayes formula, and

$$P_1 = \int_{a_n}^{a_n+\eta_n} p(S_2^n = n\tau - y_1) d\tau,$$

$$P_2 = \int_{a_n+\eta_n}^{\infty} p(S_2^n = n\tau - y_1) d\tau.$$

Lemma 6. *Whenever h belongs to $\mathcal{R}$, with η_n satisfying (6.20) and (6.21) it holds*

$$\lim_{n\to\infty} \frac{P_2}{P_1} = 0.$$

Proof. Indeed

$$P_2 = \frac{1}{n} P\left(S_2^n \geq n(a_n + \eta_n) - y_1\right) = \frac{1}{n} P\left(S_2^n \geq (n-1)c_n\right),$$
$$P_1 + P_2 = \frac{1}{n} P\left(S_2^n \geq na_n - y_1\right) = \frac{1}{n} P\left(S_2^n \geq (n-1)d_n\right),$$

where $c_n = \left(n(a_n + \eta_n) - y_1\right)/(n-1)$ and $d_n = (na_n - y_1)/(n-1)$. Denote $t_{c_n} = m^{\leftarrow}(c_n)$ and $t_{d_n} = m^{\leftarrow}(d_n)$. Using Lemma 5, it holds

$$\frac{P_2}{P_1 + P_2} = \left(1 + o\left(\frac{1}{\sqrt{n}}\right)\right) \frac{t_{d_n} s(t_{d_n})}{t_{c_n} s(t_{c_n})} \exp\left(-(n-1)\left(I(c_n) - I(d_n)\right)\right).$$

Using the convexity of the function I and a first order Taylor expansion it holds

$$\exp\left(-(n-1)I(c_n) - I(d_n)\right) \leq \exp -n\eta_n m^{\leftarrow}(a_n)$$

which tends to 0 by (6.21). We now show that

$$\frac{t_{d_n} s(t_{d_n})}{t_{c_n} s(t_{c_n})} \longrightarrow 1.$$

By (6.20), $c_n/d_n \to 1$ as $a_n \to \infty$. If $h \in RV(\beta)$, it holds

$$\left(\frac{t_{d_n} s(t_{d_n})}{t_{c_n} s(t_{c_n})}\right)^2 \sim \left(\frac{d_n h(d_n)}{\beta + \epsilon\left(\psi(d_n)\right)}\right)^2 \left(\frac{\beta + \epsilon\left(\psi(c_n)\right)}{c_n h(c_n)}\right)^2 \sim \left(\frac{h(d_n)}{h(c_n)}\right)^2 \longrightarrow 1.$$

If $h \in KR_\infty$,

$$t^2 s^2(t) \sim t\psi(t)\epsilon(t),$$

hence

$$\left(\frac{t_{d_n} s(t_{d_n})}{t_{c_n} s(t_{c_n})}\right)^2 \sim \frac{d_n \psi(d_n)\epsilon(d_n)}{c_n \psi(c_n)\epsilon(c_n)} \sim \frac{\epsilon(d_n)}{\epsilon(c_n)} = \frac{\epsilon\left(c_n - n\eta_n/(n-1)\right)}{\epsilon(c_n)} \longrightarrow 1,$$

where the last step holds by using the same argument as in the second line of (6.14). We obtain

$$\frac{P_2}{P_1} = o\left(1\right). \tag{6.24}$$

□

Now turning back to (6.23) under (4.2) and making use of Lemma 6

$$\begin{aligned} p^{A_n}(y_1) &\sim \int_{a_n}^{a_n+\eta_n} \left(\frac{p(X_1 = y_1 | S_1^n = n\tau) - \pi^\tau(y_1)}{\pi^\tau(y_1)}\right) p(S_1^n = n\tau | S_1^n \geq na_n) \pi^\tau(y_1) d\tau \\ &= (1 + o(1))\left(1 + o\left(1/\sqrt{n}\right)\right) \pi^{a_n^*}(y_1) = (1 + r_n)\, \pi^{a_n^*}(y_1) \end{aligned}$$

with $\lim_{n\to\infty} r_n = 0$ is independent on y_1. The second equality follows from (4.2) making use of Theorem 3; we also used that fact that that the mapping $\tau \to \pi^\tau(y_1)$ is decreasing for $\tau > y_1$, hence for n large enough; also $a_n^* := a_n + \theta_n \eta_n$ with $0 < \theta_n < 1$ and we made use of the version of the mean value theorem for product of integrands.

We now prove that when Y is drawn under p^{A_n}, then $Y = O_{P^{A_n}}(a_n)$.

By linearity of the expectation and integration by parts, $E\left(Y|\, S_1^n > na_n\right) = a_n + \frac{1}{P\left(S_1^n/n > a_n\right)} \int_{a_n}^{\infty} P\left(S_1^n/n > u\right) du$; now by Markov Inequality and convexity of I,

$$\begin{aligned}\int_{a_n}^{\infty} P\left(S_1^n/n > u\right) du &\leq \int_{a_n}^{\infty} \exp\left(-nI(u)\right) du \\ &\leq \left(\exp nI(a_n)\right) \int_{a_n}^{\infty} \exp\left(-n\left(I(u) - I(a_n)\right)\right) du \\ &\leq \left(\exp nI(a_n)\right) / nI'(a_n).\end{aligned}$$

Making use of (6.17) to handle $P\left(S_1^n/n > a_n\right)$ it results that $E\left(Y|\, S_1^n > na_n\right) = a_n + (1 + o(1/\sqrt{n}))s(t)/\sqrt{n}$; now $s(t)/\sqrt{n} = o(a_n)$ whenever $h \in RV(\beta)$ for some $\beta > 0$ or when $h \in KR_\infty$, for any sequence a_n bounded away from 0. Therefore $E\left(Y|\, S_1^n > na_n\right) = a_n + o(a_n)$ and $Y = O_{P^{A_n}}(a_n)$ follows by Markov Inequality.

It follows that, similarly as in Theorem 7, $\lim_{n\to\infty} d_V\left(P^{A_n}, \Pi^{a_n^*}\right) = 0$. Now by Kemperman Inequality, $d_V\left(\Pi^{a_n}, \Pi^{a_n^*}\right) \leq \sqrt{2K\left(\Pi^{a_n^*}, \Pi^{a_n}\right)}$.

We prove that $K\left(\Pi^{a_n^*}, \Pi^{a_n}\right) = (t^* - t)\, a_n + \log\Phi\left(t^*\right) - \log\Phi(t)$ tends to 0 as n tends to infinity, where $m(t^*) = a_n^*$. By Taylor expansion $K\left(\Pi^{a_n^*}, \Pi^{a_n}\right) = (t^* - t)\left(m(t^{**}) - a_n\right)$ for some t^{**} such that $m(t^{**})$ belongs to $(a_n, a_n + \eta_n)$. Therefore, through a second Taylor expansion

$$K\left(\Pi^{a_n^*}, \Pi^{a_n}\right) \leq \frac{\eta_n^2}{\inf_{\alpha\in(a_n, a_n+\eta_n)} V(\alpha)}.$$

Now $V(\alpha) = s^2(m^{\leftarrow}(\alpha)) = \psi'\left(m^{\leftarrow}(\alpha)\right))o(1))$.

In the case when h belongs to $RV(\beta)$ for some $\beta > 0$ it holds denoting $\alpha := a_n + \theta_n\eta_n$ for some θ_n in $(0, 1)$, under (6.20)

$$V(\alpha) \sim \psi'\left(\psi^{-1}(a_n\left(1 + \theta_n\eta_n\right)\right) \sim \psi'\left(\psi^{-1}(a_n)\right)$$

Therefore under (6.20) and (6.22) $K\left(\Pi^{a_n^*}, \Pi^{a_n}\right)$ tends to 0.

In the case when h belongs to KR_∞, making use of Condition (2.8) proves that ψ' is ultimately decreasing, which in turn proves that $\inf_{\alpha\in(a_n, a_n+\eta_n)} V(\alpha)$ can be substituted by $V(a_n)$, Therefore under (6.22), $K\left(\Pi^{a_n^*}, \Pi^{a_n}\right)$ tends to 0.

All conditions (6.20), (6.21) and (6.22) are fulfilled by adequate sequence η_n when h belongs to $\mathcal{R}$, hence $\lim_{n\to\infty} d_V\left(P^{A_n}, \Pi^{a_n}\right) \leq \lim_{n\to\infty} d_V\left(P^{A_n}, \Pi^{a_n^*}\right) + d_V\left(\Pi^{a_n}, \Pi^{a_n^*}\right) = 0$ as sought.

References

[1] Balkema, A. A.; Klüppelberg, C.; Resnick, S. I. Densities with Gaussian tails. Proc. London Math. Soc. (3) 66 (1993), no. 3, 568–588.

[2] Bar-Lev, S K.; Bshouty, D; Enis, P On polynomial variance functions. Probab. Theory Related Fields 94 (1992), no. 1, 69–82.

[3] Barndorff-Nielsen, O., "Information and exponential families in statistical theory," Wiley Series in Probability and Mathematical Statistics. Chichester: John Wiley & Sons, (1978).

[4] Bingham, N.H., Goldie, C.M., Teugels, J.L., "Regular Variation," Cambridge University Press, Cambridge, (1987).

[5] Biret ,M., Broniatowski, M.,Cao, Z, A sharp Abelian theorem for the Laplace transform. Mathematical statistics and limit theorems, 67–92, Springer, Cham, 2015.
[6] Broniatowski, M., Cao, Z. , Light tails: all summands are large when the empirical mean is large. Extremes 17 (2014), no. 2, 305–336.
[7] Broniatowski, M., Caron, V. Long runs under a conditional limit distribution. Ann. Appl. Probab. 24 (2014), no. 6, 2246–2296.
[8] Csiszár, I. Sanov property, generalized I-projection and a conditional limit theorem. Ann. Probab. 12 (1984), no. 3, 768–793.
[9] Dembo, A. and Zeitouni, O., Refinements of the Gibbs conditioning principle, Probab. Theory Related Fields 104 1¨C14, (1996).
[10] Diaconis, P., Freedman, D., Conditional Limit Theorems for Exponential Families and Finite Versions of de Finetti's Theorem, Journal of Theoretical Probability, Vol. 1, No. 4, (1988).
[11] Feller, W. "An introduction to probability theory and its applications," Vol. 2, second edition, John Wiley and Sons Inc., New York, (1971).
[12] Frisch, U., Sornette, D. Extreme Deviations and Applications. Journal de Physique I, EDP Sciences, 1997, 7 (9), pp.1155-1171.
[13] Jensen, J. L. "Saddlepoint approximations," Oxford Statistical Science Series, 16. Oxford Science Publications. Oxford University Press, New York, (1995).
[14] Jørgensen, B; Martínez, J R. Tauber theory for infinitely divisible variance functions. Bernoulli 3 (1997), no. 2, 213–224.
[15] Letac, G.; Mora, M. Natural real exponential families with cubic variance functions. Ann. Statist. 18 (1990), no. 1, 1–37.
[16] Juszczak, D., Nagaev, A. V., Local large deviation theorem for sums of i.i.d. random vectors when the Cramer condition holds in the whole space, Probability and Mathematical Statistics, Vol. 24, (2004), pp. 297-320.
[17] Sornette. D; Critical Phenomena in Natural Sciences; Chaos, Fractals, Selforganization and Disorder; concepts and tools. Ed Springer, Berlin Heidelberg ISBN-13 978-3-540-30882-9, (2006)

Safran, Villaroche, France
E-mail address: maeva.biret@safrangroup.com

Current address: LPSM, Sorbonne Université, Paris, France
E-mail address: michel.broniatowski@sorbonne-universite.fr
URL: https://www.lpsm.paris/pageperso/broniatowski/

LPSM, Sorbonne Université, Paris, France