MATHEMATICAL COMMUNICATIONS
Math. Commun. **20**(2015), 219–228

Shannon's differential entropy asymptotic analysis in a Bayesian problem

Mark Kelbert[1,2] and Pavel Mozgunov[1,3,*]

[1] *Higher School of Economics, National Research University, 20 Myasnitskaya str., Moscow, 101 000, RF*
[2] *Department of Mathematics, Swansea University, Singleton Park, Swansea, SA2 8PP, UK*
[3] *Department of Mathematics and Statistics, Lancaster University, Lancaster, LA1 4YF, UK*

Received June 8, 2015; accepted August 22, 2015

Abstract. We consider a Bayesian problem of estimating of probability of success in a series of conditionally independent trials with binary outcomes. We study the asymptotic behaviour of the differential entropy for a posterior probability density function conditional on x successes after n conditionally independent trials, when $n \to \infty$. Three particular cases are studied: x is a proportion of n; $x \sim n^{\beta}$, where $0 < \beta < 1$; either x or $n-x$ is a constant. It is shown that after an appropriate normalization in the first and second case limiting distribution is Gaussian and the differential entropy of a standardized RV converges to the differential entropy of a the Gaussian distribution. In the third case, the limiting distribution in not Gaussian, but still the asymptotics of the differential entropy can be found explicitly.

AMS subject classifications: 94A17, 62B10, 62C10

Key words: differential entropy, Bayes' formula, Gaussian limit theorem

1. Introduction

Let U be a random variable (RV) uniformly distributed in interval $[0,1]$. Given a realization of this RV p, consider a sequence of conditionally independent identically distributed $(\xi_i, i = 1, 2, \ldots)$, where $\xi_i = 1$ with probability p and $\xi_i = 0$ with probability $1-p$. Let x_i, each 0 or 1, be an outcome in trial i. Denote $S_n = \xi_1 + \ldots + \xi_n$ and $x = \sum_{i=1}^{n} x_i$. Note that RVs (ξ_i) are positively correlated. Indeed, $\mathbb{P}(\xi_i = 1, \xi_j = 1) = \int_0^1 p^2 \mathrm{d}p = 1/3$ if $i \neq j$, but $\mathbb{P}(\xi_i = 1)\mathbb{P}(\xi_j = 1) = (\int_0^1 p \mathrm{d}p)^2 = 1/4$.

The probability that after n trials the exact sequence $(x_i, i = 1, ..., n)$ will appear equals:

$$\mathbb{P}(\xi_1 = x_1, ..., \xi_n = x_n) = \int_0^1 p^x (1-p)^{n-x} \mathrm{d}p = \frac{1}{(n+1)\binom{n}{x}}. \tag{1}$$

*Corresponding author. *Email addresses:* `mark.kelbert@gmail.com` (M. Kelbert), `pmozgunov@gmail.com` (P. Mozgunov)

This implies that the posterior probability density function (PDF) of the number of x successes after n trials is uniform:

$$\mathbb{P}(S_n = x) = \frac{1}{(n+1)}, x = 0, \dots, n.$$

It could be easily checked that the sufficient statistics for parameter p is S_n. Given the information that after n trials one observes x successes, the posterior PDF takes the form

$$f_n(p|\xi_1 = x_1, ..., \xi_n = x_n) = f_n(p|S_n = x) = (n+1)\binom{n}{x} p^x (1-p)^{n-x},\ 0 \le p \le 1. \quad (2)$$

Note that a conditional distribution given in (2) is a Beta-distribution $B(x+1, n-x+1)$. "It is known that a Beta-distribution is asymptotically normal with its mean and variance as x and $(n-x)$ tend to infinity, but this fact is lacking a handy reference"(see [3, p.1]). That is why we give the proof of this fact in two cases.

Consider an RV $Z^{(n)}$ on $[0;1]$ with a PDF (2). Note that $Z^{(n)}$ has the following expectation:

$$\mathbb{E}_x[Z^{(n)}] = \frac{x+1}{n+2}, \quad (3)$$

and the following variance:

$$\mathbb{V}_x[Z^{(n)}] = \frac{(x+1)(n-x+1)}{(n+3)(n+2)^2}. \quad (4)$$

Recall the definition of a differential entropy $h(f)$ of an RV Z with the PDF f:

$$h(f) = h_{diff}(f) = -\int_{\mathbb{R}} f(z)\log(f(z))\mathrm{d}z \quad (5)$$

with the convention $0\log 0 = 0$. When referring to the differential entropy of an RV Z we mean the entropy of its PDF f. Consider a linear transformation $X = b_1 Z + b_2$. Then [1, 7]:

$$h(g) = h(f) + \log b_1, \quad (6)$$

where g is a PDF of an RV X. Let $\bar{Z}$ be the standard Gaussian RV with PDF φ,

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Then the differential entropy of $\bar{Z}$ equals [7]:

$$h(\varphi) = \frac{1}{2}\log(2\pi e).$$

Recall the definition of the Kullback-Leibler divergence of g from f:

$$\mathbb{D}(f||g) = \int_{\mathbb{R}} f(x)\log\frac{f(x)}{g(x)}\mathrm{d}x. \quad (7)$$

Next, we consider a sequence of distributions f_n of the form (2) with $x = x(n)$ changing with the size of the sample. The goal of our work is to study the asymptotic behaviour of the differential entropy of the following RVs:

1. $Z_\alpha^{(n)}$ with PDF $f_\alpha^{(n)}$ given in (2) when $x = x(n) = \lfloor \alpha n \rfloor$, where $0 < \alpha < 1$ and $\lfloor a \rfloor$ is the integer part of a.

2. $Z_\beta^{(n)}$ with PDF $f_\beta^{(n)}$ given in (2) when $x = x(n) = \lfloor n^\beta \rfloor$, where $0 < \beta < 1$.

3. $Z_{c_1}^{(n)}$ with PDF $f_{c_1}^{(n)}$ given in (2) when $x = c_1$ and $Z_{n-c_2}^{(n)}$ with PDF $f_{n-c_2}^{(n)}$ given in (2) when $n - x(n) = c_2$, where c_1 and c_2 are some constants.

Generally, it does not play an important role whether $x(n)$ is integer or not. The analysis also holds for arbitrary positive values. In fact, the main steps of the analysis hold for complex values with positive real part although for complex values of x and $n - x$ the problem loses a direct probabilistic character. We stick to introduced notations in the context of the formulated problem.

2. Main results

Theorem 1. *Let* $\tilde{Z}_\alpha^{(n)} = n^{\frac{1}{2}}(\alpha(1-\alpha))^{-\frac{1}{2}}(Z_\alpha^{(n)} - \alpha)$ *be an RV with PDF* $\tilde{f}_\alpha^{(n)}$*. Let* $\bar{Z} \sim \mathcal{N}(0,1)$ *be the standard Gaussian RV. Then*

(a) $\tilde{Z}_\alpha^{(n)}$ *weakly converges to* $\bar{Z}$*:*

$$\tilde{Z}_\alpha^{(n)} \Rightarrow \bar{Z} \text{ as } n \to \infty,$$

(b) the differential entropy of $\tilde{Z}_\alpha^{(n)}$ *converges to the differential entropy of* $\bar{Z}$*:*

$$\lim_{n\to\infty} h(\tilde{f}_\alpha^{(n)}) = \frac{1}{2}\log(2\pi e),$$

(c) the Kullback-Leibler divergence of φ *from* $\tilde{f}_\alpha^{(n)}$ *tends to* 0 *as* $n \to \infty$*:*

$$\lim_{n\to\infty} \mathbb{D}(\tilde{f}_\alpha^{(n)}||\varphi) = 0.$$

Theorem 2. *Let* $\tilde{Z}_\beta^{(n)} = n^{1-\beta/2}(Z_\beta^{(n)} - n^{\beta-1})$ *be an RV with PDF* $\tilde{f}_\beta^{(n)}$ *and* $\bar{Z} \sim \mathcal{N}(0,1)$ *then*

(a) $\tilde{Z}_\beta^{(n)}$ *weakly converges to* $\bar{Z}$*:*

$$\tilde{Z}_\beta^{(n)} \Rightarrow \bar{Z} \text{ as } n \to \infty,$$

(b) the differential entropy of $\tilde{Z}_\beta^{(n)}$ *converges to the differential entropy of* $\bar{Z}$*:*

$$\lim_{n\to\infty} h(\tilde{f}_\beta^{(n)}) = \frac{1}{2}\log(2\pi e),$$

(c) the Kullback-Leibler divergence of φ *from* $\tilde{f}_\beta^{(n)}$ *tends to* 0 *as* $n \to \infty$*:*

$$\lim_{n\to\infty} \mathbb{D}(\tilde{f}_\beta^{(n)}||\varphi) = 0.$$

Theorem 3. *Let $\tilde{Z}^{(n)}_{c_1} = nZ^{(n)}_{c_1}$ be an RV with PDF $\tilde{f}^{(n)}_{c_1}$ and $\tilde{Z}^{(n)}_{n-c_2} = nZ^{(n)}_{n-c_2}$ an RV with PDF $\tilde{f}^{(n)}_{n-c_2}$. Let $H_k = 1 + \frac{1}{2} + \ldots + \frac{1}{k}$ denote the partial sum of harmonic series and γ the Euler-Mascheroni constant. Then*

$$\text{(a)}\ \lim_{n\to\infty} h(\tilde{f}^{(n)}_{c_1}) = c_1 + \sum_{i=0}^{c_1-1} \log(c_1 - i) - c_1(H_{c_1} - \gamma) + 1,$$

$$\text{(b)}\ \lim_{n\to\infty} h(\tilde{f}^{(n)}_{n-c_2}) = c_2 + \sum_{i=0}^{c_2-1} \log(c_2 - i) - c_2(H_{c_2} - \gamma) + 1.$$

3. Proof of Theorem 1

Proof. *(a)* Let $x = x(n) = \lfloor \alpha n \rfloor$, where $0 < \alpha < 1$, and consider a RV

$$\tilde{Z}^{(n)}_{\alpha} = n^{\frac{1}{2}}(\alpha(1-\alpha))^{-\frac{1}{2}}(Z^{(n)}_{\alpha} - \alpha).$$

We proceed by the method of characteristic functions, and establish that

$$\phi(t) = \mathbb{E}[e^{it\tilde{Z}^{(n)}_{\alpha}}] \to e^{-t^2/2} \tag{8}$$

for all $t \in \mathbb{R}$. Indeed,

$$\begin{aligned}\phi(t) =& \int_0^1 e^{it\frac{(p-\alpha)\sqrt{n}}{\sqrt{\alpha(1-\alpha)}}} f^{(n)}_{\alpha}(p)\mathrm{d}p \\ =&(n+1)\binom{n}{x} e^{it\frac{(-\alpha)\sqrt{n}}{\sqrt{\alpha(1-\alpha)}}} \int_0^1 e^{it\frac{p\sqrt{n}}{\sqrt{\alpha(1-\alpha)}}} p^x(1-p)^{n-x}\mathrm{d}p\end{aligned}$$

and consider the integral:

$$I(t,\alpha,n) = \int_0^1 e^{n(it\frac{p}{\sqrt{\alpha(1-\alpha)n}} + \alpha\log p + (1-\alpha)\log(1-p))}\mathrm{d}p. \tag{9}$$

Denote $g(p) = it\frac{p}{\sqrt{\alpha(1-\alpha)n}} + \alpha\log p + (1-\alpha)\log(1-p)$. The integrand in (9) has a narrow sharp peak, and the integral is completely dominated by the maximum of $\mathrm{Re}[g(p)]$ when $n \to \infty$. For fixed values of t, α and $n \to \infty$, it can be studied by the saddle point method [4, Theorem 1.3, p.170]:

$$I(t,\alpha,n) \simeq e^{ng(p^*)}\sqrt{\frac{2\pi}{-ng''(p^*)}}\left(1 + O\left(\frac{1}{n}\right)\right). \tag{10}$$

Find the point of maximum of $\mathrm{Re}[g(p)]$ and deform the initial contour $[0,1]$ into the steepest descent contour through the saddle point:

$$p^* = \alpha + it\frac{\sqrt{(1-\alpha)\alpha}}{\sqrt{n}} + O\left(\frac{1}{n}\right).$$

So, $\phi(t)$ takes the form:

$$\phi(t) = e^{-t^2}(n+1)\binom{n}{x}(p^*)^x(1-p^*)^{n-x}\sqrt{\frac{2\pi}{-ng''(p^*)}} + O\left(\frac{1}{n}\right).$$

Here and below $x = \lfloor \alpha n \rfloor$. Next, by Stirling's formula:

$$(n+1)\binom{n}{x} \simeq (n+1)\frac{n^n}{x^x(n-x)^{(n-x)}}\sqrt{\frac{n}{2\pi x(n-x)}}.$$

So, the straightforward computation yields:

$$\begin{aligned}(p^*)^x(1-p^*)^{n-x} \simeq & \alpha^x(1-\alpha)^{(n-x)}e^{it\sqrt{(1-\alpha)\alpha n}+\frac{(1-\alpha)t^2}{2}-it\sqrt{(1-\alpha)\alpha n}+\frac{\alpha t^2}{2}} \\ = & e^{\frac{t^2}{2}}\left(\frac{x}{n}\right)^x\left(\frac{n-x}{n}\right)^{n-x}.\end{aligned}$$

It can be checked that the next term in the asymptotics of $\log p^*$ (as well as $\log(1-p^*)$) is decaying to 0 after multiplication by αn and $(1-\alpha)n$, respectively.

We have for $t \in \mathbb{R}$

$$\begin{aligned}\phi(t) \simeq & e^{-t^2}\frac{(n+1)n^n}{x^x(n-x)^{(n-x)}}\sqrt{\frac{n}{2\pi x(n-x)}}e^{\frac{t^2}{2}}\left(\frac{x}{n}\right)^x\left(\frac{n-x}{n}\right)^{n-x}\sqrt{\frac{2\pi x(n-x)}{n^3}} \\ \simeq & e^{-\frac{t^2}{2}}\end{aligned}$$

This fact establishes the pointwise convergence of characteristic function to its Gaussian limit and it completes the proof of part (a).

(b) Write the differential entropy in the form:

$$h(f_\alpha^{(n)}) = -\left(\log\left[(n+1)\binom{n}{x}\right] + (n+1)\binom{n}{x}xI_1 + (n+1)\binom{n}{x}(n-x)I_2\right), \quad (11)$$

where

$$I_1 = \int_0^1 p^x(1-p)^{n-x}\log p\,\mathrm{d}p, \quad (12)$$

$$I_2 = \int_0^1 p^x(1-p)^{n-x}\log(1-p)\,\mathrm{d}p. \quad (13)$$

Integrals I_1 and I_2 can be computed explicitly by reducing to the standard integral

$$\int_0^1 x^{\mu-1}(1-x^r)^{\nu-1}\log x\,\mathrm{d}x = \frac{1}{r^2}B\left(\frac{\mu}{r},\nu\right)\left(\psi\left(\frac{\mu}{r}\right) - \psi\left(\frac{\mu}{r}+\nu\right)\right), \quad (14)$$

where $\psi(x)$ is the digamma function, and $B(x,y)$ is the Beta-function [5, #4.253.1] and in considering case $r \equiv 1, \mu - 1 \equiv x, \nu - 1 \equiv n - x$.

For integral I_1, we get:

$$U_1 = (n+1)\binom{n}{x}xI_1 = -x(\psi(n+2) - \psi(x+1)).$$

Similarly, for the second integral I_2, we obtain:

$$U_2 = (n+1)\binom{n}{x}(n-x)I_2 = -(n-x)(\psi(n+2) - \psi(n-x+1)).$$

After summation of these two integrals and by using the asymptotics for the digamma function [5, #8.362.2], we obtain:

$$U_1 + U_2 = x\log x - n\log n + (n-x)\log(n-x) - \frac{1}{2} + O\left(\frac{1}{n}\right).$$

Next, we apply Stirling's formula to the first term in (11):

$$\begin{aligned} U_0 =& \log\left[(n+1)\binom{n}{x}\right] \\ =& n\log n - x\log x - (n-x)\log(n-x) \\ &+ \frac{1}{2}\log n - \frac{1}{2}\log\alpha - \frac{1}{2}\log(1-\alpha) - \log(\sqrt{2\pi}) + O\left(\frac{1}{n}\right). \end{aligned}$$

Here, as before, $x = \lfloor \alpha n \rfloor$. So, we obtain the following asymptotics of the differential entropy:

$$\lim_{n\to\infty}\left[h(f_\alpha^{(n)}) - \frac{1}{2}\log\frac{2\pi e[\alpha(1-\alpha)]}{n}\right] = 0. \tag{15}$$

Due to (6), the differential entropy of RV $\tilde{Z}_\alpha^{(n)}$ has the form:

$$\lim_{n\to\infty}\left[h(\tilde{f}_\alpha^{(n)})\right] = \frac{1}{2}\log(2\pi e). \tag{16}$$

(c) By the definition of the the Kullback-Leibler divergence:

$$\begin{aligned} \mathbb{D}(\tilde{f}_\alpha^{(n)}||\varphi) =& - h(\tilde{f}_\alpha^{(n)}) - \int_0^1 \tilde{f}_\alpha^{(n)}(p)\log\varphi(p)\mathrm{d}p \\ =& - \frac{1}{2}\log(2\pi e) + \frac{1}{2}\log(2\pi) + \frac{1}{2}\int_0^1 p^2\tilde{f}_\alpha^{(n)}\mathrm{d}p + O\left(\frac{1}{n}\right) = O\left(\frac{1}{n}\right), \end{aligned}$$

$\int_0^1 p^2\tilde{f}_\alpha^{(n)}\mathrm{d}p = 1 + O\left(\frac{1}{n}\right)$ is the second moment of $\tilde{Z}_\alpha^{(n)}$. This completes the proof. □

4. Proof of Theorem 2

Proof. *(a)* Let $x = x(n) = \lfloor n^\beta \rfloor$, where $0 < \beta < 1$, and consider $\tilde{Z}_\beta^{(n)}$ such that

$$\tilde{Z}_\beta^{(n)} = n^{1-\beta/2}(Z_\beta^{(n)} - n^{\beta-1}).$$

In this case, it is more convenient to proceed by the method of moments. We use the following classical result. Let f_n be a sequence of distribution functions with finite moments $\mu_k(n)$. Let $\mu_k(n)$ tends to ν_k for each k as $n \to \infty$, where ν_k are moments of distribution f and the distribution f is uniquely defined by its moments. Then f_n weakly converges to f as $n \to \infty$ [9].

Consider RV $\tilde{Z}_\beta^{(n)} = n^{1-\beta/2}(Z_\beta^{(n)} - n^{\beta-1})$, where $Z_\beta^{(n)}$ has PDF (2) when $x = \lfloor n^\beta \rfloor$, and compute all moments of $\tilde{Z}_\beta^{(n)}$. First, $\mathbb{E}(\tilde{Z}_\beta^{(n)}) \to 0$ as $n \to \infty$ because $\mathbb{E}(Z_\beta^{(n)}) = n^{\beta-1} + O\left(\frac{1}{n}\right)$. Next, we check that $\mathbb{E}\left[\left(\tilde{Z}_\beta^{(n)}\right)^2\right] = n^{2(1-\beta/2)}\mathbb{E}(Z_\beta^{(n)} - n^{\beta-1})^2 \to 1$ as $n \to \infty$. Compute central moments for any $k > 1$:

$$\mathbb{E}\left[\left(\tilde{Z}_\beta^{(n)}\right)^k\right] = n^{k-\frac{\beta k}{2}}(1-n^{1-\beta})^{-k}(1-n^{\beta-1})^k \ {}_2F_1[-k, n^\beta+1; n+2; n^{1-\beta}], \quad (17)$$

where ${}_2F_1[-k, n^\beta+1; n+2; n^{1-\beta}]$ is the hypergeometric function, which is in this case the polynomial:

$${}_2F_1[-k, n^\beta+1; n+2; n^{1-\beta}] = \sum_{i=0}^{k}(-1)^i \binom{k}{i} \frac{(n^\beta+1)_i}{(n+2)_i} n^{i(1-\beta)},$$

where $(q)_n$ is the rising Pochhammer symbol. For $n > 0$,

$$(q)_n = q(q+1)\dots(q+n-1)$$

and $(q)_0 = 1$.

Consider the asymptotics of terms separately:

$$n^{k-\frac{\beta k}{2}}(1-n^{1-\beta})^{-k}(1-n^{\beta-1})^k \simeq O(n^{\frac{k\beta}{2}})$$

and

$${}_2F_1[-k, n^\beta+1; n+2; n^{1-\beta}] \simeq O(n^{-[0.5+0.5k]\beta}), \quad (18)$$

where $\lfloor k \rfloor$ is the integer part of k. For k odd

$$n^{k(1-\beta/2)}\mathbb{E}(Z_\beta^{(n)} - n^{\beta-1})^k = O(n^{\frac{k\beta}{2}})O(n^{-[0.5+0.5k]\beta}) \simeq O(n^{-\beta/2}) \to 0 \quad (19)$$

as $n \to \infty$. For k even

$$n^{k(1-\beta/2)}\mathbb{E}(Z_\beta^{(n)} - n^{\beta-1})^k = O(n^{\frac{k\beta}{2}})O(n^{-[0.5+0.5k]\beta}) = O(1). \quad (20)$$

We see that every even central moment tends to a constant which is the coefficient in front of term $n^{-[0.5+0.5k]\beta}$ in the hypergeometric function. For k even, we have:

$$n^{k(1-\beta/2)}\mathbb{E}(Z_\beta^{(n)} - n^{\beta-1})^k \to (k-1)!!. \quad (21)$$

These imply that a RV $\tilde{Z}_\beta^{(n)}$ weakly converges to the standard Gaussian RV.

(b) Write the differential entropy in the form:

$$\begin{aligned} h(f_\beta^{(n)}) &= -\left(\log\left[(n+1)\binom{n}{x}\right] + (n+1)\binom{n}{x}xI_1 + (n+1)\binom{n}{x}(n-x)I_2\right) \\ &= -(U_0 + U_1 + U_2), \end{aligned} \quad (22)$$

where I_1 and I_2 are defined in (12) and (13) and can be computed explicitly by (14).

As before, we apply Stirling's formula for U_0:

$$U_0 = n\log n - x\log x - (n-x)\log(n-x) + \log n + \frac{1}{2}(-\log n^\beta - \log(1-n^{\beta-1})) - \frac{1}{2}\log(2\pi) + O\left(\frac{1}{n^\beta}\right).$$

As far as $0 < \beta < 1$, the remainder tends to 0 as $n \to \infty$. Note that the rate of decaying depends on parameter β, contrary to remainder in Theorem 1. Now $U_1 + U_2$ can be computed as follows:

$$U_1 + U_2 = x\log x - n\log n + (n-x)\log(n-x) - \frac{1}{2} + O\left(\frac{1}{n^\beta}\right).$$

So, we proved that

$$\lim_{n\to\infty}\left[h(f_\beta^{(n)}) - \frac{1}{2}\log\frac{2\pi e(1-n^{\beta-1})}{n^{2-\beta}}\right] = 0.$$

Due to (6), the differential entropy of RV $\tilde{Z}_\beta^{(n)}$ has the form:

$$\lim_{n\to\infty} h(\tilde{f}_\beta^{(n)}) = \frac{1}{2}\log(2\pi e).$$

(c) Similarly, by the definition of the Kullback-Leibler divergence:

$$\begin{aligned}\mathbb{D}(\tilde{f}_\beta^{(n)}||\varphi) &= -h(\tilde{f}_\beta^{(n)}) - \int_0^1 \tilde{f}_\beta^{(n)}(p)\log\varphi(p)\mathrm{d}p \\ &= -\frac{1}{2}\log(2\pi e) + \frac{1}{2}\log(2\pi) + \frac{1}{2}\int_0^1 p^2\tilde{f}_\beta^{(n)}\mathrm{d}p + O\left(\frac{1}{n^\beta}\right) = O\left(\frac{1}{n^\beta}\right),\end{aligned}$$

$\int_0^1 p^2\tilde{f}_\beta^{(n)}\mathrm{d}p = 1 + O\left(\frac{1}{n^\beta}\right)$ is the second moment of $\tilde{Z}_\beta^{(n)}$.

□

5. Proof of Theorem 3

Proof. *(a)* Let $x = c_1$, where c_1 is some integer constant. Consider the differential entropy $h(f_{c_1}^{(n)}) = -(U_0 + U_1 + U_2)$, where U_0, U_1 and U_2 defined in (22). Applying Stirling's formula to U_0:

$$U_0 = \log n - \log(x!) + x\log n + O\left(\frac{1}{n}\right).$$

Next, we compute $U_1 + U_2$ via formula (14) as before. The only difference will be in the asymptotics of digamma functions [5, #8.365.3, #8.365.4], because of $x = c_1$, where c_1 is constant:

$\psi(n-x+1) \simeq \log n + \frac{1/2-x}{2n}$, and $\psi(x+1) = H_x - \gamma$, here H_x is the partial sum of harmonic series and γ stands for the Euler-Mascheroni constant. Using that $x = c_1$:

$$\lim_{n\to\infty}\left[h(f_{c_1}^{(n)}) + \log n\right] = c_1 + \sum_{i=0}^{c_1-1}\log(c_1-i) - c_1(H_{c_1}-\gamma) + 1.$$

Due to (6), it can be written in the following form:

$$\lim_{n\to\infty} h(\tilde{f}_{c_1}^{(n)}) = c_1 + \sum_{i=0}^{c_1-1}\log(c_1-i) - c_1(H_{c_1}-\gamma) + 1.$$

(b) Let $n - x(n) = c_2$, where c_2 is some integer constant. In a similar way, we compute $h(f_{n-c_2}^{(n)})$, where $n - x = c_2$ and c_2 is a constant. The asymptotics of the digamma function is given as follows [5, #8.365.4]:

$$\psi(n-x+1) = H_{c_2} - \gamma \text{ where } x = n - c_2,$$

and the final result for the differential entropy:

$$h(f_{n-c_2}^{(n)}) = -\log n + c_2 - c_2(H_{c_2}-\gamma) + \sum_{i=0}^{c_2-1}\log(c_2-i) + 1 + O\left(\frac{1}{n}\right).$$

In terms of standardized RV $\tilde{Z}_{n-c_2}^{(n)}$, due to (6) we obtain:

$$\lim_{n\to\infty} h(\tilde{f}_{n-c_2}^{(n)}) = c_2 + \sum_{i=0}^{c_2-1}\log(c_2-i) - c_2(H_{c_2}-\gamma) + 1.$$

□

6. Conclusion

We demonstrated that the limiting distributions of the standardized RV $\tilde{Z}^{(n)}$ when $n \to \infty$ in cases 1 and 2 are Gaussian. However, the asymptotic normality does not imply automatically the limiting form of the differential entropy. In general, the problem of taking the limits under the sign of entropy is rather delicate and was extensively studied in the literature, cf., i.e., [2, 6]. In the third case, the limiting distribution is not Gaussian, but still the asymptotics of the differential entropy can be found explicitly.

For the Bayesian problem studied here, the explicit asymptotic expansions of the Shannon, Renyi, Tsallis and Fisher entropies are presented in [8]. The considered problem seems important because of natural extensions of this topic. The so-called weighted differential entropy is the most recent one (see [10, 11, 12] and the reference therein). The weighted version of these entropies is defined in [8] and the explicit asymptotic expansions for the stated Bayesian problem are obtained.

Acknowledgement

We thank the anonymous referees for useful comments and remarks.

The article was prepared within the framework of a subsidy granted to the HSE by the Government of the Russian Federation for the implementation of the Global Competitiveness Program.

References

[1] T. M. Cover, J. M. Thomas, *Elements of Information Theory*, John Wiley & Sons, New York, 2006.

[2] R. L. Dobrushin, *Passing to the limit under the sign of the information and entropy*, Theory Prob.Appl. **5** (1960), 29–37.

[3] R. Dudley, *The Delta-Method and Asymptotics of some Estimators*, Lecture notes on "Topics in Statistics: Nonparametrics and Robustness", MIT, 2005.

[4] M. V. Fedoruk, *Saddle Point Method*, Nauka, Moscow, 1977.

[5] I. S. Gradshteyn, I. M. Ryzhik, *Table of Integrals, Series, and Product*, Elsevier, New York, 2007.

[6] M. Kelbert, Yu. Suhov, *Continuity of mutual entropy in the large signal-to-noise ratio limit*, in: *Stochastic Analysis*, (D. Crisan, Ed.), Springer, Berlin, 2010, 281–299.

[7] M. Kelbert, Yu. Suhov, *Information Theory and Coding by Example*, Cambridge University Press, Cambridge, 2013.

[8] M. Kelbert, P. Mozgunov, *Asymptotic behaviour of weighted differential entropies in a Bayesian problem*, 2015, arXiv: 1504.01612.

[9] Yu. V. Prokhorov, *Moments, method of (in probability theory)*, Encyclopedia of Mathematics, 2001, `http://www.encyclopediaofmath`.

[10] Yu. Suhov, S. Yasaei Sekeh, *Simple inequalities for weighted entropies*, 2014, arXiv: 1409.4102.

[11] Yu. Suhov, S. Yasaei Sekeh, M. Kelbert, *Entropy-power inequality for weighted entropy*, 2015, arXiv: 1502.02188.

[12] Yu. Suhov, S. Yasaei Sekeh, I. Stuhl, *Weighted Gaussian entropy and determinant inequalities*, 2015, arXiv: 1505.01753.