EURASIAN MATHEMATICAL JOURNAL
ISSN 2077-9879
Volume 6, Number 2 (2015), 6 – 17

ASYMPTOTIC BEHAVIOUR OF THE WEIGHTED RENYI, TSALLIS AND FISHER ENTROPIES IN A BAYESIAN PROBLEM

M. Kelbert, P. Mozgunov

Communicated by E.L. Presman

Key words: weighted differential entropy, Bernoulli random variable, Renyi entropy, Tsallis entropy, Fisher information

AMS Mathematics Subject Classification: 94A17, 62B10, 62C10

Abstract. We consider the Bayesian problem of estimating the success probability in a series of conditionally independent trials with binary outcomes. We study the asymptotic behaviour of the weighted differential entropy for posterior probability density function conditional on x successes after n conditionally independent trials when $n \to \infty$. Suppose that one is interested to know whether the coin is approximately fair with a high precision and for large n is interested in the true frequency. In other words, the statistical decision is particularly sensitive in a small neighbourhood of the particular value $\gamma = 1/2$. For this aim the concept of the weighted differential entropy introduced in [1] is used when it is necessary to emphasize the frequency γ. It was found that the weight in suggested form does not change the asymptotic form of Shannon, Renyi, Tsallis and Fisher entropies, but changes the constants. The leading term in weighted Fisher Information is changed by some constant which depends on the distance between the true frequency and the value we want to emphasize.

1 Introduction

Let U be a random variable (RV) that uniformly distributed in interval $[0,1]$. Given a realization of this RV p, consider a sequence of conditionally independent identically distributed ξ_i 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 by $S_n = \xi_1 + \ldots + \xi_n$, by $\mathbf{x} = (x_i, i = 1, ..., n)$ and by $x = x(n) = \sum_{i=1}^n x_i$. 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 $\mathbf{x}$ will appear:

$$\mathbb{P}(\xi_1 = x_1, ..., \xi_n = x_n) = \int_0^1 p^x (1-p)^{n-x} \mathrm{d}p = \frac{1}{(n+1)\binom{n}{x}}. \tag{1.1}$$

The posteriori PDF given the information that after n throws we observe x heads takes the form

$$f^{(n)}(p) \equiv f_{p|S_n}(p|\xi_1 = x_1, ..., \xi_n = x_n) = (n+1)\binom{n}{x} p^x (1-p)^{n-x}. \tag{1.2}$$

Note that conditional distribution given in (1.2) is a Beta-distribution $B(x+1, n-x+1)$. The RV $Z^{(n)}$ with PDF (1.2) has the following conditional variance:

$$\mathbb{V}[Z^{(n)}|S_n = x] = \frac{(x+1)(n-x+1)}{(n+3)(n+2)^2}. \tag{1.3}$$

In our previous paper [8] Shannon's entropy of (1.2) was studied in three particular cases: $x = \lfloor \alpha n \rfloor$, $x \sim n^\beta$, where $0 < \alpha, \beta < 1$ and either x or $n - x$ is a constant. We have demonstrated that the limiting distributions when $n \to \infty$ in the cases 1 and 2 are Gaussian. However, the asymptotic normality does not imply automatically the limiting form of differential entropy. In general the problem of taking the limits under the sign of entropy is rather delicate and was extensively studied in literature, cf., e.g., [4, 6]. In stated problem, it was proved that in the first and second cases the differential entropies are asymptotically Gaussian with corresponding variances. In the third case the limiting distribution is not Gaussian, but still the asymptotics of the differential entropy can be found explicitly.

Consider the following statistical experiment with twofold goal: at the initial stage an experimenter mainly concerns whether the coin is approximately fair (i.e. $p \approx \frac{1}{2}$) with a high precision. As the size of a sample grows, he proceeds to estimate the true value of the parameter anyway. We want to quantify the differential entropy of this experiment taking into account its two sided objective. It seems that the quantitative measure of information gain of this experiment is provided by the concept of the weighted differential entropy [2, 1].

Let $\phi^{(n)} \equiv \phi^{(n)}(\alpha, \gamma, p)$ be a weight function that underlines the importance of some particular value γ ($\gamma = 1/2$ in the problem stated above). The goal of this work is to study the asymptotic behaviour of weighted Shannon's (1.4), Renyi's (1.5), Tsallis's (1.6) and Fisher's (1.7) differential entropies [3, 8] of RV $Z^{(n)}$ with PDF $f^{(n)}$ given in (1.2) and particular RV $Z^{(n)}_\alpha$ with PDF $f^{(n)}_\alpha$ given in (1.2) with $x = \lfloor \alpha n \rfloor$ where $0 < \alpha < 1$:

$$h^\phi(f^{(n)}_\alpha) = -\int_{\mathbb{R}} \phi^{(n)} f^{(n)}_\alpha \log f^{(n)}_\alpha \mathrm{d}p, \tag{1.4}$$

$$H^\phi_\nu(f^{(n)}_\alpha) = \frac{1}{1-\nu} \log \int_{\mathbb{R}} \phi^{(n)} \left(f^{(n)}_\alpha\right)^\nu \mathrm{d}p, \tag{1.5}$$

$$S^\phi_q(f^{(n)}_\alpha) = \frac{1}{q-1} \left(1 - \int_{\mathbb{R}} \phi^{(n)} \left(f^{(n)}_\alpha\right)^q \mathrm{d}p\right), \tag{1.6}$$

$$I^\phi(\alpha) = \mathbb{E}\left(\phi^{(n)}(Z^{(n)}_\alpha) \left(\frac{\partial}{\partial \alpha} \log f(Z^{(n)}_\alpha)\right)^2 \middle| \alpha\right) \tag{1.7}$$

where $q, \nu \geq 0$ and $q, \nu \neq 1$.When the weight function is uniform ($\phi \equiv 1$) we will omit the superscript ϕ. The following special cases are considered:

1. $\phi^{(n)} \equiv 1$,
2. $\phi^{(n)}$ depends on both n and p.

We assume that $\phi^{(n)}(x) \geq 0$ for all x. Choosing the weight function we adopt the following normalization rule:

$$\int_{\mathbb{R}} \phi^{(n)} f_\alpha^{(n)} \mathrm{d}p = 1. \tag{1.8}$$

It can be easily checked that if a weight function $\phi^{(n)}$ satisfies (1.8) then the Renyi weighted entropy (1.5) and Tsallis weighted entropy (1.6) tend to Shannon's weighted entropy as $\nu \to 1$ and $q \to 1$ correspondingly.

In this paper we consider the weight function of the following form:

$$\phi^{(n)}(p) = \Lambda^{(n)}(\alpha, \gamma) p^{\gamma\sqrt{n}} (1-p)^{(1-\gamma)\sqrt{n}} \tag{1.9}$$

where $\Lambda^{(n)}(\alpha, \gamma, p)$ is found from normalizing condition (1.8) and is given explicitly in (3.1). This weight function is selected as a model example with a twofold goal to emphasize a particular value γ for moderate n while preserving the estimate to be asymptotically unbiased

$$\lim_{n\to\infty} \int_0^1 p\phi^{(n)} f^{(n)} \mathrm{d}p = \alpha.$$

2 Main results

Proposition 2.1. *For the weighted Shannon differential entropy of RV $Z_\alpha^{(n)}$ with PDF $f_\alpha^{(n)}$ and weight function $\phi^{(n)}$ given in (1.9) the following limit exists*

$$\lim_{n\to\infty} \left(h^\phi(f_\alpha^{(n)}) - \frac{1}{2}\log\left(\frac{2\pi e \alpha(1-\alpha)}{n}\right)\right) = \frac{(\alpha-\gamma)^2}{2\alpha(1-\alpha)}. \tag{2.1}$$

If $\alpha = \gamma$ then

$$\lim_{n\to\infty} \left(h^\phi(f_\alpha^{(n)}) - h(f_\alpha^{(n)})\right) = 0 \tag{2.2}$$

where $h(f_\alpha^{(n)})$ is the standard ($\phi \equiv 1$) Shannon's differential entropy.

Theorem 2.1. *Let $Z^{(n)}$ be a RV with PDF $f^{(n)}$ given in (1.2), $Z_\alpha^{(n)}$ be a RV with PDF $f_\alpha^{(n)}$ given in (1.2) with $x = \lfloor \alpha n \rfloor$, $0 < \alpha < 1$ and $H_\nu(f^{(n)})$ be the weighted Renyi differential entropy given in (1.5).*

(a) *When $\phi^{(n)} \equiv 1$ and both x and $n - x$ tend to infinity as $n \to \infty$ the following limit holds*

$$\lim_{n\to\infty} \left(H_\nu(f^{(n)}) - \frac{1}{2}\log\frac{2\pi x(n-x)}{n^3}\right) = -\frac{\log(\nu)}{2(1-\nu)}, \tag{2.3}$$

and for any fixed n

$$\lim_{\nu\to 1} \left(H_\nu(f^{(n)}) - h(f^{(n)})\right) = 0. \tag{2.4}$$

(b) *When the weight function $\phi^{(n)}$ is given in (1.9) the following limit for the Renyi weighted entropy of $f_\alpha^{(n)}$ holds*

$$\lim_{n\to\infty} \left(H_\nu^\phi(f_\alpha^{(n)}) - \frac{1}{2}\log\frac{2\pi\alpha(1-\alpha)}{n}\right) = -\frac{\log(\nu)}{2(1-\nu)} + \frac{(\alpha-\gamma)^2}{2\alpha(1-\alpha)\nu}, \tag{2.5}$$

and for any fixed n

$$\lim_{\nu\to 1}\left(H_\nu^\phi(f_\alpha^{(n)}) - h^\phi(f_\alpha^{(n)})\right) = 0. \tag{2.6}$$

Proposition 2.2. *For any continuous random variable X with PDF f and for any non-negative weight function $\phi(x)$ which satisfies condition (1.8) and such that*

$$\int_{\mathbb{R}} \phi(x) f(x)^\nu |\log(f(x))| \mathrm{d}x < \infty,$$

the weighted Renyi differential entropy $H_\nu^\phi(f)$ is a non-increasing function of ν and

$$\frac{\partial}{\partial\nu} H_\nu^\phi(f) = -\frac{1}{(1-\nu)^2}\int_{\mathbb{R}} z(x)\log\frac{z(x)}{\phi(x)f(x)}dx \tag{2.7}$$

where

$$z(x) = \frac{\phi(x)(f(x))^\nu}{\int_{\mathbb{R}}\phi(x)(f(x))^\nu \mathrm{d}x}.$$

Similarly, the Tsallis weighted entropy $S_q^\phi(f)$ given in (1.6) is a non-increasing function of q.

Theorem 2.2. *Let $Z^{(n)}$ be a RV with PDF $f^{(n)}$ given in (1.2), $Z_\alpha^{(n)}$ be a RV with PDF $f_\alpha^{(n)}$ given in (1.2) with $x = \lfloor \alpha n \rfloor$, $0 < \alpha < 1$ and $S_q^\phi(f^{(n)})$ be the weighted Tsallis differential entropy given in (1.6).*

(a) *When both x and $n - x$ tend to infinity as $n \to \infty$ and $\phi^{(n)}(p) \equiv 1$,*

$$\lim_{n\to\infty}\left(S_q(f^{(n)}) - \frac{1}{q-1}\left(1 - \frac{1}{\sqrt{q}}\left(\frac{2\pi x(n-x)}{n^3}\right)^{\frac{1-q}{2}}\right)\right) = 0 \tag{2.8}$$

and for any fixed n

$$\lim_{q\to 1}\left(S_q(f^{(n)}) - h(f^{(n)})\right) = 0. \tag{2.9}$$

(b) *When the weight function $\phi^{(n)}$ is given in (1.9) the following limit for the Tsallis weighted entropy of $f_\alpha^{(n)}$ holds*

$$\lim_{n\to\infty}\left(S_q^\phi(f_\alpha^{(n)}) - \frac{1}{q-1}\left(1 - \frac{1}{\sqrt{q}}\left(\frac{2\pi\alpha(1-\alpha)}{n}\right)^{\frac{1-q}{2}} \exp\left(\frac{(\alpha-\gamma)^2(1-q)}{2\alpha(1-\alpha)q}\right)\right)\right) = 0 \tag{2.10}$$

and for any fixed n

$$\lim_{q\to 1}\left(S_q^\phi(f_\alpha^{(n)}) - h^\phi(f_\alpha^{(n)})\right) = 0. \tag{2.11}$$

Remark. It can be seen by Theorem 2.1 and Theorem 2.2 that for large n standard Renyi's entropy and standard Tsallis's entropy (for $\phi \equiv 1$) "behaves" like respective entropies of the Gaussian RV with the variance $\frac{x(n-x)}{n^3}$.

Theorem 2.3. *Let $Z_\alpha^{(n)}$ be a RV with PDF $f_\alpha^{(n)}$ given in (1.2) with $x = \lfloor \alpha n \rfloor$, $0 < \alpha < 1$ and $I(f_\alpha^{(n)})$ be the weighted Fisher information given in (1.7).*

(a) *When $\phi^{(n)} \equiv 1$*

$$\lim_{n\to\infty}\left[I(f_\alpha^{(n)}) - \left(\frac{1}{\alpha(1-\alpha)}\right)n\right] = -\frac{2\alpha^2-2\alpha+1}{2\alpha^2(1-\alpha)^2}. \tag{2.12}$$

(b) *When $\phi^{(n)}$ is given in (1.9)*

$$\lim_{n\to\infty}\left[I^\phi(f_\alpha^{(n)}) - \left(\frac{1}{\alpha(1-\alpha)} + \frac{(\alpha-\gamma)^2}{(1-\alpha)^2\alpha^2}\right)n - B(\alpha,\gamma)\sqrt{n}\right] = C(\alpha,\gamma) \tag{2.13}$$

where $B(\alpha,\gamma)$ and $C(\alpha,\gamma)$ are constants which depend only on α and γ and are given in (3.25) and (3.26) respectively .

3 Proofs

The normalizing constant $\Lambda^{(n)}(\gamma)$ in the weight function (1.9) can be found from condition (1.8). We obtain that

$$\Lambda^{(n)}(\gamma) = \frac{\Gamma(x+1)\Gamma(n-x+1)\Gamma(n+2+\sqrt{n})}{\Gamma(x+\gamma\sqrt{n}+1)\Gamma(n-x+1+\sqrt{n}-\gamma\sqrt{n})\Gamma(n+2)} = \frac{\mathbb{B}(x+1,n-x+1)}{\mathbb{B}(x+\gamma\sqrt{n}+1,n-x+\sqrt{n}-\gamma\sqrt{n}+1)} \tag{3.1}$$

where $\Gamma(x)$ is the Gamma function and $\mathbb{B}(x,y)$ is the Beta function. We denote by $\psi^{(0)}(x) = \psi(x)$ and by $\psi^{(1)}(x)$ the digamma function and its first derivative respectively

$$\psi^{(j)}(x) = \frac{\mathrm{d}^{\mathrm{j}+1}}{\mathrm{d}x^{j+1}}\log\Gamma(x). \tag{3.2}$$

In further calculations we will need the asymptotics of digamma functions in two particular cases $j = 0$ and $j = 1$ only

$$\psi(x) = \log(x) - \frac{1}{2x} + O\left(\frac{1}{x^2}\right) \text{ as } x\to\infty,$$

$$\psi^{(1)}(x) = \frac{1}{x} + \frac{1}{2x^2} + O\left(\frac{1}{x^3}\right) \text{ as } x\to\infty.$$

Recall also the Stirling formula [5]:

$$n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n\left(1+\frac{1}{12n}+O\left(\frac{1}{n^2}\right)\right) \text{ as } n\to\infty. \tag{3.3}$$

3.1 Proposition 2.1

The Shannon differential entropy of PDF $f_\alpha^{(n)}$ given in (1.2) with the weight function $\phi^{(n)}$ given in (1.9) takes the form:

$$h^\phi(f_\alpha^{(n)}) = \log\left[(n+1)\binom{n}{x}\right] + x\int_0^1 \log(p)\phi^{(n)} f_\alpha^{(n)}\mathrm{d}p + (n-x)\int_0^1 \log(1-p)\phi^{(n)} f_\alpha^{(n)}\mathrm{d}p.$$

The integrals can be computed explicitly [5] (4.253.1):

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

Applying this formula, we get

$$\int_0^1 \log(p)\phi^{(n)} f_\alpha^{(n)}\mathrm{d}p = \psi(x+z+1)-\psi(n+\sqrt{n}+2)$$

and

$$\int_0^1 \log(1-p)\phi^{(n)} f_\alpha^{(n)}\mathrm{d}p = \psi(n-x+\sqrt{n}-z+1)-\psi(n+\sqrt{n}+2)$$

where $z=\gamma\sqrt{n}$.

Applying Stirling's formula (3.3) and using the asymptotics for the digamma function we have that

$$h^\phi(f_\alpha^{(n)}) = \frac{1}{2}\log\frac{2\pi e[\alpha(1-\alpha)]}{n} + \frac{(\alpha-\gamma)^2}{2\alpha(1-\alpha)} + O\left(\frac{1}{\sqrt{n}}\right). \tag{3.4}$$

The leading term in (3.4) is the Shannon differential entropy of Gaussian RV with the weight function $\phi^{(n)}\equiv 1$. Moreover, note that leading term of the asymptotics for the weighted differential entropy exceeds that for the classical differential entropy studied in [8]. The difference tends to zero as $\gamma\to\alpha$.

3.2 Theorem 2.1

(a) In this case $\phi^{(n)}(p)\equiv 1$ the Renyi entropy has the form

$$(1-\nu)H_\nu(f^{(n)}) = \nu\log\left[(n+1)\binom{n}{x}\right] + \log\left[\int_0^1 p^{\nu x}(1-p)^{\nu(n-x)}\mathrm{d}p\right].$$

Consider the integral:

$$\int_0^1 p^{\nu x}(1-p)^{\nu(n-x)}\mathrm{d}p = \mathbb{B}(\nu x+1,\nu(n-x)+1) = \frac{\Gamma(\nu x+1)\Gamma(\nu(n-x)+1)}{\Gamma(\nu n+2)}.$$

Applying Stirling's formula, we obtain that

$$(1-\nu)H_\nu(f^{(n)}) = \frac{1-\nu}{2}\log\left(\frac{2\pi x(n-x)}{n^3}\right) - \frac{1}{2}\log(\nu) + O\left(\frac{1}{n}\right). \tag{3.5}$$

So, we have that

$$H_\nu(f^{(n)}) = \frac{1}{2}\log\left(\frac{2\pi x(n-x)}{n^3}\right) - \frac{\log(\nu)}{2(1-\nu)} + O\left(\frac{1}{n}\right). \tag{3.6}$$

Note that the leading terms in (3.6) looks like the Renyi differential entropy of the Gaussian RV with variance $\frac{x(n-x)}{n^3}$.

Taking the limit as $\nu \to 1$ and applying L'Hopital's rule we get that

$$H_{\nu\to 1}(f^{(n)}) = \lim_{\nu\to 1} H_\nu(f^{(n)}) = \frac{1}{2}\log\left(\frac{2e\pi x(n-x)}{n^3}\right) + O\left(\frac{1}{n}\right). \tag{3.7}$$

For example, when $x = \lfloor \alpha n \rfloor$, $0 < \alpha < 1$ the Renyi entropy:

$$H_{\nu\to 1}(f^{(n)}) = \frac{1}{2}\log\frac{2\pi e[\alpha(1-\alpha)]}{n} + O\left(\frac{1}{n}\right)$$

where the leading term is Shannon's differential entropy of the Gaussian RV with the corresponding variance.

(b) When $\phi^{(n)}$ is given in (1.9) and $x = \lfloor \alpha n \rfloor$, the weighted Renyi differential entropy of PDF $f_\alpha^{(n)}$ takes the following form

$$H_\nu^\phi(f_\alpha^{(n)}) = \frac{1}{1-\nu}\log\int_0^1 \phi^{(n)}\left(f_\alpha^{(n)}\right)^\nu \mathrm{d}p,$$

$$\int_0^1 \phi^{(n)}\left(f_\alpha^{(n)}\right)^\nu \mathrm{d}p \equiv U_1 U_2 U_3 \text{ where}$$

$$U_1 = \frac{\Gamma(\nu x + \gamma\sqrt{n} + 1)\Gamma(\nu(n-x) + (1-\gamma)\sqrt{n} + 1)}{\Gamma(\nu n + \sqrt{n} + 2)},$$

$$U_2 = \left(\frac{\Gamma(n+2)}{\Gamma(x+1)\Gamma(n-x+1)}\right)^{\nu-1},$$

$$U_3 = \frac{\Gamma(n+\sqrt{n}+2)}{\Gamma(x+z+1)\Gamma(n-x+\sqrt{n}-z+1)}$$

where $z = \gamma\sqrt{n}$ as before.

Applying Stirling's formula for each term and taking all parts together, we obtain that

$$H_\nu^\phi(f_\alpha^{(n)}) = \frac{1}{2}\log\frac{2\pi\alpha(1-\alpha)}{n} - \frac{\log(\nu)}{2(1-\nu)} + \frac{(\alpha-\gamma)^2}{2\alpha(1-\alpha)\nu} + O\left(\frac{1}{\sqrt{n}}\right). \tag{3.8}$$

Taking the limit when $\nu \to 1$ and applying L'Hopital's rule we get that

$$H_{\nu\to 1}^\phi(f_\alpha^{(n)}) = \lim_{\nu\to 1} H_\nu^\phi(f_\alpha^{(n)}) = \frac{1}{2}\log\frac{2\pi e[\alpha(1-\alpha)]}{n} + \frac{(\alpha-\gamma)^2}{2\alpha(1-\alpha)} + O\left(\frac{1}{\sqrt{n}}\right). \tag{3.9}$$

So, for any fixed n the weighted Renyi differential entropy tends to Shannon's weighted differential entropy as $\nu \to 1$.

3.3 Proposition 2.2

We need to show that

$$\frac{\partial}{\partial\nu}H_\nu^\phi(f) \le 0,$$

$$\frac{\partial}{\partial\nu}H_\nu^\phi(f) = \frac{\log\int_{\mathbb{R}}\phi(x)(f(x))^\nu \mathrm{d}x}{(1-\nu)^2} + \frac{\int_{\mathbb{R}}\phi(x)(f(x))^\nu \log(f(x))\mathrm{d}x}{(1-\nu)\int_{\mathbb{R}}\phi(x)(f(x))^\nu \mathrm{d}x} = I_1 + I_2. \quad (3.10)$$

Denote

$$z(x) = \frac{\phi(x)(f(x))^\nu}{\int_{\mathbb{R}}\phi(x)(f(x))^\nu \mathrm{d}x}. \quad (3.11)$$

Note that $z(x) \ge 0$ for any x and

$$\int_{\mathbb{R}} z(x)\mathrm{d}x = 1.$$

Denote $Q = \log\int_{\mathbb{R}}\phi(x)(f(x))^\nu \mathrm{d}x$. Using the substitution (3.11)

$$Q = \log(\phi(x)) + \nu\log(f(x)) - \log(z(x)) \quad (3.12)$$

we obtain

$$I_2 = \frac{1}{1-\nu}\int_{\mathbb{R}} z(x)\log(f(x))\mathrm{d}x,$$

$$I_1 + I_2 = \frac{1}{(1-\nu)^2}\left(\log\int_{\mathbb{R}}\phi(x)(f(x))^\nu \mathrm{d}x + (1-\nu)\int_{\mathbb{R}} z(x)\log(f(x))\mathrm{d}x\right).$$

By substitution $\log(f(x))$ using (3.12) we obtain that

$$-\frac{\partial}{\partial\nu}H_\nu^\phi(f) = \frac{1}{(1-\nu)^2}\int_{\mathbb{R}} z(x)\log\left(\frac{z(x)}{\phi(x)f(x)}\right)\mathrm{d}x = \frac{1}{(1-\nu)^2}\mathbb{D}_{KL}(z||\phi f). \quad (3.13)$$

Here $\mathbb{D}_{KL}(z||\phi f)$ is the Kullback-Leibler divergence between z and ϕf which is always non-negative [3, 7]. Due to conditions $\phi(x)f(x) \ge 0$ and (1.8), $\phi(x)f(x)$ is itself a PDF:

$$\int_{\mathbb{R}}\phi(x)f(x)\mathrm{d}x = 1.$$

Similarly, one can show that the Tsallis weighted differential entropy given in (1.6) is a non-increasing function of q. So, the result follows.

3.4 Theorem 2.2.

(a) When $\phi^{(n)} \equiv 1$, the Tsallis entropy has the form

$$S_q(f^{(n)}) = \frac{1}{q-1}\left(1 - \int_0^1 (f(p))^q \mathrm{d}p\right) = \frac{1}{q-1}\left(1 - \int_0^1 \left((n+1)\binom{n}{x}p^x(1-p)^{n-x}\right)^q \mathrm{d}p\right).$$

Let us denote the integral above by V_0. Compute its asymptotics using (3.3), (3.5) and the Taylor expansion for exponential function we get

$$V_0 \equiv \int_0^1 \left(f^{(n)}\right)^q \mathrm{d}p = e^{\log \int_0^1 \left(f^{(n)}\right)^q \mathrm{d}p} = \frac{1}{\sqrt{q}} \left(\frac{2\pi x(n-x)}{n^3}\right)^{\frac{1-q}{2}} \left(1 + O\left(\frac{1}{n}\right)\right).$$

By straightforward computation we obtain that

$$S_q(f^{(n)}) = \frac{1}{q-1}\left(1 - \left(\frac{1}{\sqrt{q}}\frac{2\pi x(n-x)}{n^3}\right)^{\frac{1-q}{2}} \left(1 + O\left(\frac{1}{n}\right)\right)\right). \qquad (3.14)$$

Note that $V_0 \to 1$ as $q \to 1$, applying L'Hospital's rule we get that

$$\lim_{q\to 1} S_q(f^{(n)}) = S_{q\to 1}(f^{(n)}) = \frac{1}{2}\log\left(\frac{2e\pi x(n-x)}{n^3}\right) + O\left(\frac{1}{n}\right). \qquad (3.15)$$

The leading term in the expression above is nothing else but Shannon's differential entropy of the Gaussian RV.

(b) When $\phi^{(n)}$ is given in (1.9) the Tsallis entropy of PDF $f_\alpha^{(n)}$ has the form

$$S_q^\phi(f_\alpha^{(n)}) = \frac{1}{q-1}\left(1 - \int_0^1 \phi^{(n)} \left(f_\alpha^{(n)}\right)^q \mathrm{d}p\right).$$

Using that $x = \lfloor \alpha n \rfloor$, (3.3), (3.8) and the Taylor expansion for exponential function, we obtain that

$$V_1 \equiv \int_0^1 \phi^{(n)} \left(f_\alpha^{(n)}\right)^q \mathrm{d}p = e^{\log\left[\int_0^1 \phi^{(n)} \left(f_\alpha^{(n)}\right)^q \mathrm{d}p\right]} =$$

$$= \frac{1}{\sqrt{q}}\left(\frac{2\pi\alpha(1-\alpha)}{n}\right)^{\frac{1-q}{2}} \exp\left(\frac{(\alpha-\gamma)^2}{2\alpha(1-\alpha)}\left(\frac{1}{q}-1\right)\right)\left(1 + O\left(\frac{1}{\sqrt{n}}\right)\right).$$

So, we have the following form for the weighted Tsallis differential entropy:

$$S_q^\phi(f_\alpha^{(n)}) = \frac{1}{q-1}\left(1 - \frac{1}{\sqrt{q}}\left(\frac{2\pi\alpha(1-\alpha)}{n}\right)^{\frac{1-q}{2}} e^{\frac{(\alpha-\gamma)^2}{2\alpha(1-\alpha)}\left(\frac{1}{q}-1\right)}\left(1 + O\left(\frac{1}{\sqrt{n}}\right)\right)\right).$$

Note that $V_1 \to 1$ as $q \to 1$. Applying L'Hospital's rule we get that

$$S_{q\to 1}^\phi(f_\alpha^{(n)}) \equiv \lim_{q\to 1} S_q^\phi(f_\alpha^{(n)}) = \frac{1}{2}\log\frac{2\pi e[\alpha(1-\alpha)]}{n} + \frac{(\alpha-\gamma)^2}{2\alpha(1-\alpha)} + O\left(\frac{1}{\sqrt{n}}\right). \qquad (3.16)$$

Then the weighted Tsallis entropy tends to weighted Shannon's differential entropy as $q \to 1$.

3.5 Theorem 2.3

(a) The Fisher information in the case $\phi^{(n)} \equiv 1$ and PDF $f_\alpha^{(n)}$ takes the form:

$$I(\alpha) = \mathbb{E}\left(\left(\frac{\partial}{\partial\alpha}\log f_\alpha^{(n)}(Z_\alpha^{(n)})\right)^2 \Big|\alpha\right) = \int_0^1 \left(\frac{\partial}{\partial\alpha}\log f_\alpha^{(n)}\right)^2 f_\alpha^{(n)}\mathrm{d}p,$$

$$\frac{\partial}{\partial\alpha}\log f_\alpha^{(n)} = n\log(p) - n\log(1-p) + n\psi(n-x+1) - n\psi(x+1). \tag{3.17}$$

Denote

$$W_0 = \frac{\Gamma(n-x+1)\Gamma(x+1)}{\Gamma(n+2)}.$$

In order to compute the expectation we will need the following formulas [5] (4.261.17):

$$\begin{aligned}&\int_0^1 \log^2(p)p^x(1-p)^{n-x}\mathrm{d}p\\ &= W_0(\psi(n+2)-\psi(x+1))^2 - \psi^{(1)}(n+2) + \psi^{(1)}(x+1),\end{aligned} \tag{3.18}$$

$$\begin{aligned}&\int_0^1 \log^2(1-p)p^x(1-p)^{n-x}\mathrm{d}p\\ &= W_0(\psi(n+2)-\psi(n-x+1))^2 - \psi^{(1)}(n+2) + \psi^{(1)}(n-x+1),\end{aligned} \tag{3.19}$$

$$\begin{aligned}&\int_0^1 \log(p)\log(1-p)p^x(1-p)^{n-x}\mathrm{d}p\\ &= W_0(\psi(n+2)-\psi(n-x+1)(\psi(n+2)-\psi(x+1)) - \psi^{(1)}(n+2).\end{aligned} \tag{3.20}$$

So, we have that

$$I(\alpha) = n^2(\psi^{(1)}(x+1) + \psi^{(1)}(n-x+1)). \tag{3.21}$$

Using the asymptotics for the digamma function we get

$$I(\alpha) = \frac{1}{\alpha(1-\alpha)}n - \frac{1}{2}\frac{2\alpha^2-2\alpha+1}{\alpha^2(1-\alpha)^2} + O\left(\frac{1}{n}\right). \tag{3.22}$$

Remark. When $x = \lfloor \alpha n \rfloor$

$$\int_0^1 p f_\alpha^{(n)}\mathrm{d}p = \alpha + b_n(\alpha)$$

where $b_n(\alpha)$ is a bias

$$b_n(\alpha) \simeq \frac{1-2\alpha}{n}.$$

Note that $\frac{\partial}{\partial\alpha}b_n(\alpha) \to 0$ as $n \to \infty$. So, our estimate is asymptotically unbiased. Also note that the leading term in Theorem 2.3 has the same form as in the classical problem of estimating p in a series of n binary trials $I(p) = \frac{n}{p(1-p)}$.

(b) The weighted Fisher Information in the case of the weight (1.9) and PDF $f_\alpha^{(n)}$ takes the following form

$$I^\phi(\alpha) = \mathbb{E}\left(\phi^{(n)}\left(\frac{\partial}{\partial\alpha}\log f_\alpha^{(n)}\right)^2 \Big| \alpha\right) = \int_0^1 \phi^{(n)}\left(\frac{\partial}{\partial\alpha}\log f_\alpha^{(n)}\right)^2 f_\alpha^{(n)} \mathrm{d}p.$$

All the integrals can be found exactly similarly to integrals (3.18)-(3.20):

$$I^\phi(\alpha) = n^2\left(\psi^{(1)}(x+z+1) + \psi^{(1)}(n-x+1+\sqrt{n}-z)\right)$$

$$+n^2\left[(\psi(x+z+1)-\psi(x+1)) + \left(\psi(n-x+1+\sqrt{n}-z)-\psi(n-x+1)\right)\right]^2.$$

Using the asymptotics for the digamma function we get

$$I^\phi(\alpha) = A(\alpha,\gamma)n + B(\alpha,\gamma)\sqrt{n} + C(\alpha,\gamma) + O\left(\frac{1}{\sqrt{n}}\right), \tag{3.23}$$

where

$$A(\alpha,\gamma) = \frac{1}{\alpha(1-\alpha)} + \frac{(\alpha-\gamma)^2}{(1-\alpha)^2\alpha^2}, \tag{3.24}$$

$$B(\alpha,\gamma) = \frac{2\alpha\gamma-\gamma-\alpha^2}{(1-\alpha)^2\alpha^2} + \frac{(\alpha-\gamma)^2}{(1-\alpha)^3\alpha^3}(\alpha(2\gamma-1)-\gamma), \tag{3.25}$$

$$\begin{aligned} C(\alpha,\gamma) = {} & \frac{\alpha-2\alpha^4-2\gamma^2+6\alpha\gamma^3+\alpha^3(2+4\gamma)-3\alpha(1+\gamma^2)}{-2(1-\alpha)^3\alpha^3)} \\ & + \frac{\alpha^4(-31-44\gamma+72\gamma^2-56\gamma^3+28\gamma^4+36\alpha-12\alpha^2)}{12(1-\alpha)^4\alpha^4} \\ & + \frac{6\alpha^2(\gamma^2-2\gamma^3+12\gamma^4-1)-4\gamma^3(11\gamma-44\alpha\gamma-6+3\gamma^2-6\gamma^3+14\gamma^4)}{12(1-\alpha)^4\alpha^4}. \end{aligned} \tag{3.26}$$

An impact of the weight function of form (1.9) results in appearance of the term of order $\sqrt{n}$, but the leading order, n, remains the same. However, the coefficient at n is larger by $\frac{(\alpha-\gamma)^2}{(1-\alpha)^2\alpha^2}$. Evidently, when the frequency of special interest is equal to the true frequency the leading term is the same as in the Fisher information with constant weight. Also note that the rate depends on the distance between γ and α and as $\gamma \to \alpha$ only the leading term remains.

Acknowledgements

We would like to thank the anonymous referees for useful comments and remarks. The study (reaserch grant №14-05-0007) was supported by the Academic Fund Program of thet the National Research University Higher School of Economics in 2014-2015.

References

[1] M. Belis, S. Guiasu, *A Quantitative and qualitative measure of information in cybernetic systems,* IEEE Trans. Inf. Th. 14 (1968), 593-594.

[2] A. Clim, *Weighted entropy with application*, Analele Universitatii Bucuresti Matematica, Anul LVII, 2008, 223-231.

[3] T.M. Cover, J.M. Thomas, *Elements of information theory*, NY: Basic Books, 2006.

[4] R.L. Dobrushin, *Passing to the limit under the sign of the information and entropy*, Theory Prob. Appl. 1960, 29-37.

[5] I.S. Gradshteyn, I.M. Ryzhik, *Table of integrals, series, and product*, Elsevier, 2007.

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

[7] M. Kelbert, Yu. Suhov, *Information theory and coding by example*, Cambridge: Cambridge University Press, 2013

[8] M. Kelbert, P. Mozgunov, *Shannon's differential entropy asymptotic analysis in a Bayesian problem*, Mathematical Communications, 20 (2015), no. 2, 219–228.

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

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

[11] M. Kelbert, Yu. Suhov, I. Stuhl, S. Yasaei Sekeh, *Basic inequalities for weighted entropies*, arXiv:1510.02184, 2015

Mark Kelbert, Pavel Mozgunov
International Laboratory of Stochastic Analysis and Its Applications
National Research University Higher School of Economics
31 Shabolovka St.,
Moscow, Russia
E-mails: mark.kelbert@gmail.com, pmozgunov@gmail.com

Mark Kelbert
Department of Mathematics
Swansea University
Singleton Park,
Swansea, UK
E-mail: m.kelbert@swansea.ac.uk

Received: 21.04.2015