Smooth Rényi Entropy and Applications

Renato Renner [1]
Department of Computer Science
ETH Zürich, Switzerland
e-mail renner@inf.ethz.ch

Stefan Wolf [2]
Département d'Informatique et R.O.
Université de Montréal, Canada
e-mail: wolf@iro.umontreal.ca

Abstract — **We introduce a new entropy measure, called *smooth Rényi entropy*. The measure characterizes fundamental properties of a random variable Z, such as the amount of uniform randomness that can be extracted from Z or the minimum length of an encoding of Z.**

I. Definition and Properties

For a probability distribution P and $\varepsilon \geq 0$, let $\mathcal{B}^\varepsilon(P) := \{Q : \delta(P,Q) \leq \varepsilon\}$ be the set of probability distributions which are *ε-close* to P, with respect to the variational distance δ.[3]

Definition I.1 *Let P be a probability distribution with range $\mathcal{Z}$, let $\alpha \in [0,\infty]$, and let $\varepsilon \geq 0$. The ε-*smooth Rényi entropy (of order α) of P *is*[4]

$$H_\alpha^\varepsilon(P) := \frac{1}{1-\alpha} \inf_{Q \in \mathcal{B}^\varepsilon(P)} \log_2 \left(\sum_{z \in \mathcal{Z}} Q(z)^\alpha \right) .$$

For a random variable Z with probability distribution P_Z, $H_\alpha^\varepsilon(P_Z)$ is also denoted as $H_\alpha^\varepsilon(Z)$.

The smooth Rényi entropy H_α^ε inherits many of its properties from conventional Rényi entropy H_α as introduced in [4]. E.g., for any $\varepsilon \geq 0$,

$$H_\alpha^\varepsilon(Z) \geq H_\beta^\varepsilon(Z) \tag{1}$$

if $\alpha < \beta \leq 1$ or $1 < \alpha \leq \beta$. Moreover, for the case of a great number of i.i.d. random variables, the smooth Rényi entropy (of any order α) approaches the Shannon entropy.

Lemma I.2 *Let $Z^n := (Z_1, \ldots, Z_n)$ be an n-tuple of independent random variables Z_i distributed according to P_Z. Then, for any $\alpha \in [0,\infty]$,*

$$\lim_{\varepsilon \to 0} \lim_{n \to \infty} \frac{H_\alpha^\varepsilon(Z^n)}{n} = H(Z) .$$

The following lemma, together with (1), implies that the smooth Rényi entropy $H_\alpha^\varepsilon(Z)$ of a random variable Z, for any order $\alpha \in [0,\infty]$, is—up to an additive constant—determined by $H_0^\varepsilon(Z)$ and $H_\infty^\varepsilon(Z)$. We will see in Section II that these two entropy measures also characterize many natural properties of Z (e.g., the amount of extractable randomness or the encoding length). This yields a new interpretation of the Shannon entropy $H(Z)$ which is the common value of these (natural) quantities, for the special case where Z consists of many independent repetitions (cf. Lemma I.2).

Lemma I.3 *Let Z be a random variable and let $\varepsilon > 0$. Then,*

$$H_\infty^\varepsilon(Z) \geq H_\alpha(Z) - \frac{1}{\alpha-1}\log(1/\varepsilon) \qquad \text{for } \alpha > 1$$

$$\text{and} \qquad H_0^\varepsilon(Z) \leq H_\alpha(Z) + \frac{1}{1-\alpha}\log(1/\varepsilon) \qquad \text{for } \alpha < 1 .$$

II. Applications

A fundamental property of a random variable Z is the amount of (almost) uniform randomness that can be extracted from Z (see, e.g. [2] and [1]) by application of a randomly chosen function F (called *extractor* [3]).[5] For a set $\mathcal{P}$ of probability distributions, the *ε-extractable uniform randomness* of $\mathcal{P}$ is defined as the amount of randomness that can be extracted from a random variable Z with any probability distribution $P_Z \in \mathcal{P}$, where the actual distribution P_Z does not have to be known. Formally,[6]

$$H_{\mathrm{ext}}^\varepsilon(\mathcal{P}) := \max_U \big(\log_2 |U|\big) ,$$

where the maximum ranges over all uniform random variables U such that there exists a random function F satisfying the following: For any random variable Z with $P_Z \in \mathcal{P}$, the pair $(F(Z), F)$ is ε-close to the pair (U, F).

Smooth Rényi entropy quantifies the amount of extractable uniform randomness, up to some small additive constant.

Theorem II.1 *For any set $\mathcal{P}$ of probability distributions with range $\mathcal{Z}$ and $\varepsilon, \varepsilon_1, \varepsilon_2 \in \mathbb{R}^+$ with $\varepsilon_1 + \varepsilon_2 = \varepsilon$,*

$$\min_{P \in \mathcal{P}} \big(H_\infty^{\varepsilon_1}(P)\big) - 2\log(1/\varepsilon_2) \leq H_{\mathrm{ext}}^\varepsilon(\mathcal{P}) \leq \min_{P \in \mathcal{P}} \big(H_\infty^\varepsilon(P)\big) .$$

In particular, it follows from Lemma I.3 that the conventional Rényi entropy $\min_{P \in \mathcal{P}}(H_\alpha(P))$, for any $\alpha > 1$, is a lower bound for $H_{\mathrm{ext}}^\varepsilon(\mathcal{P})$ (up to some additive constant).

Another basic property of a random variable Z is the minimum length to which one can compress Z. The *ε-encoding length* $H_{\mathrm{enc}}^\varepsilon(\mathcal{P})$ is defined as the number of bits needed for encoding a random variable Z distributed according to any $P_Z \in \mathcal{P}$—where the encoding is independent of P_Z—such that Z can be recovered with probability at least $1-\varepsilon$. Then, similarly as in Theorem II.1, the smooth Rényi entropy of order 0, $\max_{P \in \mathcal{P}}(H_0^\varepsilon(P))$, can be shown equal to $H_{\mathrm{enc}}^\varepsilon(\mathcal{P})$, up to some small additive constant. Thus, again by Lemma I.3, the (conventional) Rényi entropy of order α, for any $\alpha < 1$, yields an upper bound for the encoding length.

References

[1] C. Cachin Smooth entropy and Rényi entropy. In *Adv. in Cryptology — EUROCRYPT '97*, vol. 1233 of *Lecture Notes in Computer Science*, pp. 193–208. Springer-Verlag, 1997.

[2] R. Impagliazzo, L. A. Levin, and M. Luby. Pseudo-random generation from one-way functions. In *Proc. of the Twenty-First Annual ACM Symp. on Theory of Computing*, pp. 12–24, 1989.

[3] N. Nisan and D. Zuckerman. Randomness is linear in space. *Journal of Computer and System Sciences*, 52:43–52, 1996.

[4] A. Rényi. On measures of entropy and information. In *Proc. of the 4th Berkeley Symp. on Math. Statistics and Prob.*, vol. 1, pp. 547–561. Univ. of Calif. Press, 1961.

[1] Supported by SNF No. 20-66716.01.

[2] Supported by Canada's NSERC.

[3] $\delta(P,Q) := (\sum_z |P(z) - Q(z)|)/2$.

[4] If $\alpha = 0$ or $\alpha = \infty$, $H_\alpha^\varepsilon(P)$ is defined by the continuous extension, $H_\alpha^\varepsilon(P) := \lim_{\beta \to \alpha} H_\beta^\varepsilon(P)$. For $\alpha = 1$, set $H_1^\varepsilon(P) := H(P)$.

[5] In a cryptographic context, randomness extraction is also known as *privacy amplification*.

[6] $|U|$ denotes the size of the range of U.