# Information Theory II

Spring Semester 2017

Prof. Dr. A. Lapidoth / Prof. Dr. S. Moser

Signal and Information Processing Laboratory
Institut für Signal- und Informationsverarbeitung

## Exercise 13 of June 1, 2017

http://www.isi.ee.ethz.ch/teaching/courses/it2.html

---

### Problem 1 — *Optimal Guessing*

Let $\rho > 0$, and let $X$ and $Y$ be distributed according to the joint PMF $P_{XY}$:

| $P_{XY}(x,y)$ | $y=0$ | $y=1$ |
|---|---|---|
| $x=0$ | 0.1 | 0.2 |
| $x=1$ | 0.3 | 0.4 |

a) Compute $\mathsf{E}[\mathsf{G}^*(X)^\rho]$.

b) Compute $\mathsf{E}[\mathsf{G}^*(X|Y)^\rho]$. What do you observe?

c) Let $X_1, X_2, \ldots, X_n$ be IID $P_X$. Compute $\lim\limits_{n\to\infty} \frac{1}{n} \log \mathsf{E}[\mathsf{G}^*(X^n)^\rho]$.

### Problem 2 — *Rényi Entropy*

The Rényi entropy of order $\alpha$ of a discrete chance variable $X$ taking values in $\mathcal{X}$ according to the PMF $P_X$ is defined for $\alpha > 0$ and $\alpha \neq 1$ as

$$H_\alpha(X) \triangleq \frac{1}{1-\alpha} \log \sum_{x\in\mathcal{X}} P_X(x)^\alpha.$$

a) Show that $H_\alpha(X) \geq 0$ with equality if and only if $X$ is deterministic.

*Hint: Treat the cases $\alpha \in (0,1)$ and $\alpha > 1$ separately.*

b) Show that the Rényi entropy approaches Shannon entropy as $\alpha$ approaches unity, i.e.,

$$\lim_{\alpha\to 1} H_\alpha(X) = H(X).$$

### Problem 3 — *Rényi Divergence*

The Rényi divergence of order $\alpha$ between two PMFs $P$ and $Q$ is defined for $\alpha > 0$ and $\alpha \neq 1$ as

$$D_\alpha(P\|Q) \triangleq \frac{1}{\alpha-1} \log \sum_{x\in\mathcal{X}} P(x)^\alpha Q(x)^{1-\alpha}$$

with the convention that for $\alpha > 1$, we read $P(x)^\alpha Q(x)^{1-\alpha}$ as $\frac{P(x)^\alpha}{Q(x)^{\alpha-1}}$ and say that $\frac{0}{0} = 0$ and $\frac{p}{0} = \infty$ for $p > 0$.

a) Show that the Rényi divergence approaches relative entropy as $\alpha$ approaches unity, i.e.,

$$\lim_{\alpha\to 1} D_\alpha(P\|Q) = D(P\|Q).$$

b) Let $U$ be the uniform distribution over $\mathcal{X}$. What is $D_\alpha(P\|U)$?

## **Problem 4** ***Guessing with Chosen Side Information***

Let $X_1, X_2, \ldots, X_n$ be IID $P_X$, and let $\mathsf{R} \geq 0$. Given the side information $Y_n = f_n(X_n)$, where $f_n$ is a function from $\mathcal{X}^n$ to $\{1, \ldots, 2^{n\mathsf{R}}\}$, the decoder has to guess the sequence $X^n$.

a) Show that if $\mathsf{R} < H_{\frac{1}{1+\rho}}(X)$, then for every sequence of functions $(f_1, f_2, \ldots)$,

$$\lim_{n\to\infty} \mathsf{E}\big[\mathsf{G}^*(X^n|f_n(X^n))^\rho\big] = \infty.$$

b) Show that if $\mathsf{R} > H_{\frac{1}{1+\rho}}(X)$, then there exists a sequence of functions $(f_1, f_2, \ldots)$ for which

$$\lim_{n\to\infty} \mathsf{E}\big[\mathsf{G}^*(X^n|f_n(X^n))^\rho\big] = 1.$$

*Hint: Partition every type class $P_{\mathbf{x}}$ into $\left\lceil \frac{2^{n\mathsf{R}}}{2(n+1)^{|\mathcal{X}|}} \right\rceil$ subsets. For $n$ large enough, the total number of these sets does not exceed $2^{n\mathsf{R}}$. To bound the number of guesses, you can use the inequality $\lceil \xi \rceil^\rho \leq 1 + 2^\rho \xi^\rho$, which holds for all $\rho > 0$ and $\xi \geq 0$.*