Spring 2017

STAT 364 - Information Theory

Homework 3

Due: Feb 20 5pm, 2017 in Jason's mailbox

`Prof. Yihong Wu`

1 Exercises

1. (Total variation and coupling.) Let P and Q be probability measures on some discrete alphabet. A joint distribution P_{XY} is called a *coupling* of P and Q if $P_X = P$ and $P_Y = Q$. Recall the total variation distance we encountered in previous homework:

$$\mathrm{TV}(P,Q) \triangleq \sup_E \left(P[E] - Q[E]\right) ,$$

where the supremum is over all event E. In fact we also have the following dual representation of the total variation

$$\mathrm{TV}(P,Q) = \min_{P_{XY}} \{\mathbb{P}\left[X \neq Y\right] : P_X = P, P_Y = Q\}. \tag{1}$$

Let us prove this fact in the discrete case following the steps below (the method work in general as well):

a. Show that for any coupling P_{XY}, $\mathrm{TV}(P,Q) \leq \mathbb{P}\left[X \neq Y\right]$.

b. Prove that $\mathrm{TV}(P,Q) = \frac{1}{2}\sum |P(x) - Q(x)| = 1 - \sum_x P(x) \wedge Q(x)$.

c. Let $t = \mathrm{TV}(P,Q)$. Assume that $0 < t < 1$. Define three probability measures: $R = \frac{P\wedge Q}{1-t}$, $P' = \frac{P - P\wedge Q}{t}$ and $Q' = \frac{Q - P\wedge Q}{t}$. Construct a coupling P_{XY} as follows:

1) Generate $B \sim \mathrm{Bernoulli}(t)$.
2) If $B = 0$, draw $Z \sim R$ and set $X = Y = Z$.
3) If $B = 1$, draw $X \sim P'$ and $Y \sim Q'$ independently.

Verify that this P_{XY} is a coupling of P and Q.

d. Conclude that (1) holds (also verify the case where TV = 0 or 1).

2. (Continuity of entropy on finite alphabet.) In Exercise we learned that entropy is continuous on on finite alphabet. Now let us study how continuous it is with respect to the total variation. Prove

$$|H(P) - H(Q)| \leq h(\mathrm{TV}(P,Q)) + \mathrm{TV}(P,Q)\log(|\mathcal{X}| - 1)$$

for any P and Q supported on $\mathcal{X}$.

Hint: Use Fano's inequaility and the previous exercise.

3. (Information radius v.s. diameter.) Let $\{P_{Y|X=x} : x \in \mathcal{X}\}$ be a set of distributions. Prove that

$$\sup_{P_X} I(X;Y) \triangleq C \leq \sup_{x,x' \in \mathcal{X}} D(P_{Y|X=x} \| P_{Y|X=x'})$$

Comment: This is the information-theoretic version of "radius $\leq$ diameter".

4. Prove that if $X_1, \ldots, X_n$ are independent then

$$I(X^n; W) \geq \sum_{i=1}^{n} I(X_i; W)\,.$$

What is the necessary and sufficient condition for equality?

5. The Hewitt-Savage[1] 0-1 law states that certain symmetric events have no randomness. Let $\{X_i\}_{i\geq 1}$ be a sequence be iid random variables. Let E be an event determined by this sequence. We say E is exchangeable if it is invariant under permutation of finitely many indices in the sequence of $\{X_i\}$'s, e.g., the occurance of E is unchanged if we permute the values of (X_1, X_4, X_7), etc.

   Let's prove the Hewitt-Savage 0-1 law information-theoretically in the following steps:

   (a) (Warm-up) Verify that $E = \{\sum_{i\geq 1} X_i \text{ converges}\}$ and $E = \{\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} X_i = \mathbb{E}[X_1]\}$ are exchangeable events.

   (b) Let E be an exchangeable event and $W = \mathbf{1}_E$ is its indicator random variable. Show that for any k, $I(W; X_1, \ldots, X_k) = 0$. (Hint: apply the previous exercise, show that for arbitrary n, $nI(W; X_1, \ldots, X_k) \leq 1$ `bit`.)

   (c) Conclude that $I(W; X_1, \ldots, X_k) = 0$ for any k.

   (d) Since E is determined by the sequence $\{X_i\}_{i\geq 1}$, we have (proof not required)

   $$H(W) = I(W; X_1, \ldots) = \lim_{k\to\infty} I(W; X_1, \ldots, X_k) = 0.$$

   Conclude that E has no randomness, i.e., $P(E) = 0$ or $P(E) = 1$.

   (e) (Application to random walk) Often after the application of Hewitt-Savage, further efforts are needed to determine whether the probability is 0 or 1. As an example, consider X_i's are iid ± 1 and $S_n = \sum_{i=1}^{n} X_i$ denotes the symmetric random walk. Verify that the event $E = \{S_n = 0 \text{ finitely often}\}$ is exchangeable. Now show that $P(E) = 0$.

   (Hint: consider $E^+ = \{S_n > 0 \text{ eventually}\}$ and E^- similarly. Apply Hewitt-Savage to them and invoke symmetry.)

2 Optional reading

1. Read [1, Chapter 5]

References

[1] T. Cover and J. Thomas, *Elements of Information Theory,* Second Edition, Wiley, 2006

[1] A portrait of Jimmy Savage can be found in the department library.