Exam in information theory 8.02.2024. Theoretical part

Please sign up this page with your first and last name.

Please mark the answers that you consider right. Note that there may be more than one or none. You need not justify your choices. You **may optionally** add an argument for the point 5d (on the reverse side), which will be graded as *bonus*.

1. Let w_n, for $n \in \mathbb{N}$, be a sequence of bit words, such that $w_n \in \{0,1\}^n$, and $C_U(w_n) \geq n$, for some fixed universal Turing machine U; thus each w_n is Kolmogorov random w.r..t.. U. Which of the properties below can hold for **infinitely many** words w_n ?

   (a) w_n belongs to some Hamming code $(2^m - 1, 2^m - m - 1)$, **no**

   (b) w_n belongs to some Huffman code (for some random variable), **yes**

   (c) w_n contains a subword 1^k, where $k \geq \sqrt{n}$, **no**

   (d) w_n is a binary representation (possibly with leading zeros) of a number $\lfloor \sqrt{k} \rfloor \cdot \lceil \log_2 m \rceil$, for some integers $k, m \geq 5$. **yes**

   Comment. By definition, a word of length $2^m - 1$ in a Hamming code $(2^m - 1, 2^m - m - 1)$ is generated from a word of length $2^m - m - 1$, so it cannot be random (except for a finite number). On the other hand, for *any* word, we can easily construct a Huffman code containing this word (with probabilities being powers of $\frac{1}{2}$). A word of the form $u1^k v$ with $|u| + k + |v| = n$ and $k \geq \sqrt{n}$, can be generated from $(|u|, k, uv)$ (with $|u|$ and k in binary), which using the standard encoding of tuples has the size $\leq n - \sqrt{n} + 4 \log n + 2 < n$, for sufficiently large n. For point 1d, recall from tutorials that only a finite number of prime numbers p can be random, and for analogous reasons, only a finite number of products $2 \cdot p$. Hence, almost all random numbers m admit a presentation required in this point, as they can be decomposed by $m = a \cdot b$ with $a, b \geq 3$, and clearly any $c \geq 3$ satisfies

   $$c = \sqrt{c^2} = \log 2^c, \text{ with } c^2, 2^c \geq 5.$$

2. Consider random variables A, B, C with some joint distribution. Suppose that $I(A;B|C) = I(A;B)$ and $I(A;C|B) = I(A;C)$. Which of the options below are consistent with this assumption (i.e., can happen)? Note that we ask about each option separately, not all of them together.

   (a) $I(B;C|A) \neq I(B;C)$, **no**

   (b) $I(B;C|A) < I(B;C)$, **no**

   (c) $I(B;C) = 0$, **yes**

   (d) $I(A;B|C) \neq I(A;C|B)$. **yes**

   Comment. The assumption implies that $R(A;B;C) = 0$, which excludes (a) and (b). But (c) and (d) are possible, for example if A and C are two independent coin tosses, and $B = A$.

3. Suppose that, for a channel Γ with matrix $\begin{pmatrix} p & q \\ q & p \end{pmatrix}$, there exists a sequence of codes $C_n \subseteq \{0,1\}^n$, such that $|C_n| \to \infty$, $R(C_n) \to C_\Gamma$ and $\Pr_E(\Delta_o, A_n) \to 0$ (for $n \to \infty$), where A is a random variable taking values in C with uniform distribution, and Δ_o is (attention!) the **ideal observer rule**. Then we can claim **for sure** that

   (a) $p = 0$ or $p = 1$, **no**

   (b) $p \neq \frac{1}{2}$, **yes**

   (c) $p > \frac{1}{2}$, **no**

   (d) such a sequence cannot exist. **no**

   Comment. The Shannon channel theorem has been stated at the lecture for $p > \frac{1}{2}$ and the maximal likelihood rule Δ, with an even stronger claim: for any $\forall \varepsilon, \delta > 0$, the respective inequalities hold for **almost all** n's. If $p < \frac{1}{2}$, it is easy to see that $\Pr_E(\Delta_o, A)$ coincides with $\Pr_E(\Delta, A)$ for the

dual channel $\begin{pmatrix} q & p \\ p & q \end{pmatrix}$ and the code $\bar{C} = \{\bar{w} : w \in C\}$. Hence we can use the Shannon theorem in this case as well. (In particular, we cannot claim **for sure** that $p > \frac{1}{2}$.) On the other hand, if $p = \frac{1}{2}$ then $\Pr_E(\Delta_o, A)$ is always $1 - \frac{1}{|C|}$.

4. A matrix of a channel Γ has dimension 8×8, but only 8 of its values is different from zero. Then the capacity C_Γ

   (a) can be an arbitrary real number in the set $[0, 3]$, **no**

   (b) $0 < C_\Gamma < 3$, **no**

   (c) can only assume a value from some **finite** set, **yes**

   (d) $C_\Gamma = 3$. **no**

   As the values in each row must sum up to 1, clearly all values are equal to 1 or 0. They can be arranged in finitely many of ways. We can have $C_\Gamma = 3$ (faithful channel) or $C_\Gamma = 0$ (bad channel with all 1's in one column), but also, e.g., $C_\Gamma = 1$ (if 1's occur in exactly two rows).

5. Three friends A, B, C decided to eat together **20** donuts[1] but some randomness enters in their feast. Let the random variables A, B, C take values in the set $\{0, 1, \ldots, 20\}$ and represent how many donuts each person eats; we assume that **all donuts** will be eaten. Then we can claim for sure that

   (a) $H(A, B, C) \leq \log 231$, **yes**

   (b) $H(A, B, C) = H(A + B, B + C, C + A)$, **yes**

   (c) $H(A|B + C) \leq I(A; B|C)$, **yes**

   (d) a disjunction holds: some pair of variables is dependent (i.e., $I(A; B) + I(B; C) + I(A; C) > 0$) or $H(A) = H(B) = H(C) = 0$. **yes**

   The number of possible partitions is $\begin{pmatrix} 20+2 \\ 2 \end{pmatrix} = 231$, hence $\log 231$ is the maximal entropy we can achieve. Clearly the mapping

   $$\{(a, b, c) \in \mathbb{N}^3 : a + b + c = 20\} \ni (a, b, c) \mapsto (a + b, b + c, a + c)$$

   is one-to-one, hence the entropy is preserved (point 5b). Further, $H(A|B + C) = 0$, since $A = 20 - B - C$ is a function of $B + C$ (point 5c).

   The claim in point 5d is also true. Note that if each variable assumes a single value, such that $A + B + C = 20$, we have a correct solution with

   $$I(A; B) = I(B; C) = I(A; C) = H(A) = H(B) = H(C) = 0.$$

   Now assume that variables A, B, C are pairwise independent. There are several ways to show that they must take single values; here we present an elementary solution. We will show that the following partitions must be assumed with non-zero probabilities

   $$\begin{matrix} \min(A) & \min(B) & \max(C) \\ \max(A) & \min(B) & \max(C) \end{matrix}$$

   implying that $\min(A) = \max(A)$, and similarly for other variables, by symmetry.

   Indeed, by the independence of A and B, we have $\Pr(A = \min(A) \wedge B = \min(B)) > 0$. We claim that the value of C must then be $\max(C)$. Suppose, for the contrary, that $\Pr(A = \min(A) \wedge B = \min(B) \wedge C = d)$, with some $d < \max(C)$. Then $\min(A) + \min(B) + d = 20$. But what about the values of A and B assumed along with $\max(C)$? If $\Pr(A = a \wedge B = b \wedge C = \max(C)) > 0$, we would have $a + b + \max(C) > 20$, a contradiction! Similarly, by the independence of A and C, we

[1] Donut or doughnut, in Polish: pączek, is a pastry that in Poland is traditionally eaten (in great quantities) on Fat Thursday (Tłusty Czwartek), which in 2024 has happened on 8 February.

have $\Pr(A = \max(A) \wedge C = \max(C)) > 0$. By an argument analogous to the above, we infer that in this case $B = \min(B)$. Thus we have

$$\min(A) + \min(B) + \max(C) = \max(A) + \min(B) + \max(C) = 20$$

hence $\min(A) = \min(B)$, as required.