ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

School of Computer and Communication Sciences

Handout 3 — Information Theory and Coding

Solutions to homework 1 — Sep. 23, 2019

Problem 1. Note that $E_0 = E_1 \cup E_2 \cup E_3$.

(a) (1) For disjoint events, $P(E_0) = P(E_1) + P(E_2) + P(E_3)$, so $P(E_0) = 3/4$.

(2) For independent events, $1 - P(E_0)$ is the probability that none of the events occur, which is the product of the probabilities that each one doesn't occur. Thus $1 - P(E_0) = (3/4)^3$ and $P(E_0) = 37/64$.

(3) If $E_1 = E_2 = E_3$, then $E_0 = E_1$ and $P(E_0) = 1/4$.

(b) (1) From the Venn diagram in Fig. 1, $P(E_0)$ is clearly maximized when the events are disjoint, so $\max P(E_0) = 3/4$.

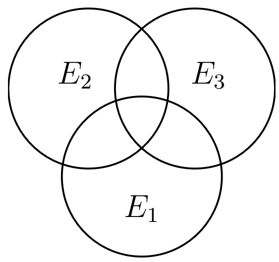

Figure 1: Venn Diagram for problem 1 (b)(1)

(2) The intersection of each pair of sets has probability $1/16$. As seen in Fig. 2, $P(E_0)$ is maximized if all these pairwise intersections are identical, in which case $P(E_0) = 3(1/4 - 1/16) + 1/16 = 5/8$. One can also use the formula $P(E_0) = P(E_1)+P(E_2)+P(E_3)-P(E_1\cap E_2)-P(E_1\cap E_3)-P(E_2\cap E_3)+P(E_1\cap E_2\cap E_3)$, and notice that all the terms except the last is fixed by the problem, and the last term cannot be made more than $\min_{i,j} P(E_i \cap E_j) = 1/16$.

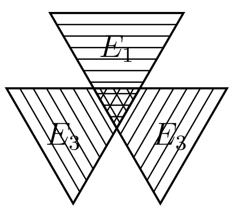

Figure 2: Venn Diagram for problem 1 (b)(2)

(c) Same considerations as in (b)(2) yields the upper bound $P(E_0) \leq 3p - 2p^2$ As $P(E_0) = 1$, we find that $p \geq 1/2$.

Problem 2. (a) Since the die is fair, the probability of a toss being 6 is $1/6$. Then, $P(N_1 = k)$ is simply the probability that the child does not observe 6 for the first $k-1$ tosses and observes 6 at k^{th} toss. Hence, $P(N_1 = k) = (5/6)^{k-1}1/6$,

(b) $E[N_1] = \sum_{k=1}^{\infty} P(N_1 = k)k = 1/6 \sum_{k=1}^{\infty}(5/6)^{k-1}k = 6^2.1/6 = 6$. Here, we used the hint $\sum_{k=1}^{\infty} x^{k-1}k = 1/(1-x)^2$.

(c) The only way $\tilde{N} = k, k \geq m$ is when (i) k^{th} toss is a 6 and (ii) in the previous $k-1$ tosses exactly $m-1$ 6's and $k-m$ non-6's are observed. There are $\binom{k-1}{m-1}$ distinct ways for (ii) to happen each with probability $(5/6)^{k-m}(1/6)^m$. Consequently, $P(\tilde{N} = k) = \binom{k-1}{m-1}(5/6)^{k-m}(1/6)^m$

To find $E[\tilde{N}]$, consider new random variables $N_i, i \in \{1, 2, \dots, m\}$ which denotes the number of tosses after the $i-1^{th}$ 6 is observed until the i^{th} 6 occurs. Since $\tilde{N} = N_1 + N_2 + \ldots + N_m$, and N_i's are independent and identically distributed, we have $E[\tilde{N}] = mE[N_1] = 6m$.

(d) Using Bayes' Rule, we have

$$P(\text{Fair} \mid \ N = k) = \frac{P(N = k \mid \text{Fair})P(\text{Fair})}{P(N = k \mid \text{Loaded})P(\text{Loaded}) + P(N = k \mid \text{Fair})P(\text{Fair})}$$

$$= \frac{(5/6)^{k-1}1/6}{(5/6)^{k-1}1/6 + (1-1/6^5)^{k-1}1/6^5}$$

The statement $P(\text{Fair} \mid \ N = k) < P(\text{Loaded} \mid \ N = k)$ is equivalent to

$$(5/6)^{k-1}1/6 < (1-1/6^5)^{k-1}1/6^5$$
$$(k-1)\ln(6/5) + \ln 6 > 5\ln 6 + (k-1)\ln(6^5/6^5 - 1)$$
$$(k-1)\ln(\frac{6(6^5-1)}{5.6^5}) + \ln 6 > 5\ln 6$$
$$k > 4\ln 6/(\ln(6(6^5-1)) - \ln(5.6^5)) + 1 \approx 40.3$$

- *An alternative way to find $P(\tilde{N} = k)$:*

  Recalling that $\tilde{N} = N_1 + N_2 + \ldots + N_m$, and N_i's are i.i.d, the distribution of $\tilde{N}$ is the m-fold convolution of the distribution of N_1. To find the m-fold convolution, we can take the easier z-transform approach. (For convenience, let $p = 1/6$ and $q = 5/6$)

  Define the z-transform of P_{N_1} as $\psi_{N_1}(z) = E[z^{-N_1}] = \sum_{k=1}^{\infty} P(N_1 = k)z^{-k} = \sum_{k=1}^{\infty} pq^{k-1}z^{-k}$

$$= \frac{pz^{-1}}{1-qz^{-1}}$$

  As $\tilde{N} = N_1 + \cdots + N_m$, the z-transform of $\tilde{N}$ will be

$$\psi_{\tilde{N}}(z) = E[z^{-(N_1+N_2+\ldots+N_m)}] = E[z^{-N_1}]E[z^{-N_2}]\ldots E[z^{-N_1m}] = (\psi_{N_1}(z))^m \tag{1}$$

$$= \left(\frac{pz^{-1}}{1-qz^{-1}}\right)^m = p^m z^{-m}\frac{1}{(1-qz^{-1})^m}$$

  From geometric series, we know that $\sum_{k=0}^{\infty} r^k = 1/1-r$. Taking the derivative of both sides with respect to r, $m-1$ times, one can obtain

$$\sum_{k=m-1}^{\infty} \frac{k!}{(k-m+1)!} r^{k-m+1} = \sum_{k=0}^{\infty} \frac{(k+m-1)!}{k!} r^k = (m-1)!\frac{1}{(1-r)^m}$$

Thus,

$$\sum_{k=0}^{\infty} \binom{k+m-1}{m-1} r^k = \frac{1}{(1-r)^m}$$

Here, if we substitute r with qz^{-1}, we get

$$\sum_{k=0}^{\infty} \binom{k+m-1}{m-1} (qz^{-1})^k = \frac{1}{(1-qz^{-1})^m}$$

and substituting in (1), we obtain

$$\psi_{\tilde{N}}(z) = \sum_{k=0}^{\infty} \binom{k+m-1}{m-1} q^k z^{-(m+k)} p^m = \sum_{k=m}^{\infty} \binom{k-1}{m-1} q^{k-m} z^{-k} p^m$$

Since by definition, $\psi_{\tilde{N}}(z) = \sum_{k=m}^{\infty} P(\tilde{N}=k) z^{-k}$, it can be seen that
$P(\tilde{N}=k) = \binom{k-1}{m-1} q^{k-m} p^m, \forall k \geq m$

Problem 3. Since A, B, C, D form a Markov chain their probability distribution is given by

$$p(a)p(b|a)p(c|b)p(d|c) \tag{2}$$

(a) Yes: Summing (2) over d shows that A, B, C have the probability distribution $p(a)p(b|a)p(c|b)$.

(b) Yes: The reverse of a Markov chain is also a Markov chain. Applying this to A, B, C, D and using part (a) we get that D, C, B is a Markov chain. Reversing again we get the desired result.

(c) Yes: Since A, B, C, D is a Markov chain, given C, D is independent of B, and thus $p(d|c) = p(d|(b,c))$. So (2) can be written as

$$p(a,(b,c),d) = p(a)p((b,c)|a)p(d|(b,c)).$$

Problem 4. No. Take for example $A = D$ and let A be independent of the pair (B, C). Then both A, B, C and B, C, A (same as B, C, D) are Markov chains. But A, B, C, D is not: A is not independent of D when B and C are given.

Problem 5.

(a)

$$\begin{aligned}
E[X+Y] &= \sum_{x,y} (x+y) P_{XY}(x,y) \\
&= \sum_{x,y} x P_{XY}(x,y) + \sum_{x,y} y P_{XY}(x,y) \\
&= \sum_{x} x P_X(x) + \sum_{y} y P_Y(y) \\
&= E[X] + E[Y].
\end{aligned}$$

Note that independence is not necessary here and that the argument extends to non-discrete variables if the expectation exists.

(b)

$$\begin{aligned} E[XY] &= \sum_{x,y} xyP_{XY}(x,y) \\ &= \sum_{x,y} xyP_X(x)P_Y(y) \\ &= \sum_x xP_X(x) \sum_y yP_Y(y) \\ &= E[X]\,E[Y]. \end{aligned}$$

Note that the statistical independence was used on the second line. Let X and Y take on only the values ± 1 and 0. An example of uncorrelated but dependent variables is

$$P_{XY}(1,0) = P_{XY}(0,1) = P_{XY}(-1,0) = P_{XY}(0,-1) = \frac{1}{4}.$$

An example of correlated and dependent variables is

$$P_{XY}(1,1) = P_{XY}(-1,-1) = \frac{1}{2}.$$

(c) Using (a), we have

$$\begin{aligned} \sigma^2_{X+Y} &= E\big[(X - E[X] + Y - E[Y])^2\big] \\ &= E[(X - E[X])^2] + 2E[(X - E[X])(Y - E[Y])] + E[(Y - E[Y])^2]. \end{aligned}$$

The middle term, from (a), is $2(E[XY] - E[X]E[Y])$. For uncorrelated variables that is zero, leaving us with $\sigma^2_{X+Y} = \sigma^2_X + \sigma^2_Y$.

PROBLEM 6. We solve the problem for a general vehicle with n wheels.

(a) Out of $n!$ possible orderings $(n-1)!$ has the tyre 1 in its original place. Thus tyre 1 is installed in its original position with probability $1/n$.

(b) All tyres end up in their original position in only 1 of the $n!$ orders. Thus the probability of this event is $1/n!$.

(c) Let X_i be the indicator random variable that tyre i is installed in its original position, so that the number of tyres installed in their original positions is $N = \sum_{i=1}^n X_i$. By (a), $E[X_i] = 1/n$. By the linearity of expectation, $E[N] = n(1/n) = 1$. Note that the linearity of the expectation holds even if the X_i's are not independent (as it is in this case).

(e) Let A_i be the event that the ith tyre remains in its original position. Then, the event we are interested in is the complement of the event $\bigcup_i A_i$ and thus has probability $1 - \Pr(\bigcup_i A_i)$. Furthermore, by the inclusion/exclusion formula,

$$\Pr(\bigcup_i A_i) = \sum_i \Pr(A_i) - \sum_{i_1<i_2} \Pr(A_{i_1} \cap A_{i_2}) + \sum_{i_1<i_2<i_3} P(A_{i_1} \cap A_{i_2} \cap A_{i_3}) - \dots.$$

The jth sum above consists of $\binom{n}{j}$ terms, each term having the value $P(A_1 \cap \cdots \cap A_j)$. Note that this is the probability of the event that tyres 1 through j have remained in their original positions, and equals $(n-j)!/n!$. Consequently,

$$\Pr\Big(\bigcup_i A_i\Big) = \sum_{j=1}^n (-1)^{j-1} \binom{n}{j} \frac{(n-j)!}{n!} = \sum_{j=1}^n (-1)^{j-1} 1/j!,$$

and the event that no tyre remains in its original position has probability

$$1 - \Pr\Big(\bigcup_i A_i\Big) = \sum_{j=0}^{n} \frac{(-1)^j}{j!}.$$

(For the case $n = 4$, the value is $3/8$.)

Problem 7.

(a) Let A_i denote the event that $X_i \neq X$. The event that X does not appear in the inventory is thus

$$A = A_1 \cap \dots A_n.$$

Note that the events $A_1, \dots, A_n$ are not independent—because they involve the common random variable X. However, they become independent when conditioned on the value of X, with $P(A_i|X = x) = 1 - p(x)$. Thus,

$$P(A|X = x) = (1 - p(x))^n.$$

Consequently $P(A) = \sum_x p(x)(1 - p(x))^n$..

(b) With p the uniform distribution on n items, the above value for $P(A)$ equals $(1 - 1/n)^n$.

(c) For n large, $(1 - 1/n)^n$ approaches $1/e \approx 37\%$.