Information Theory **DCC/ICEx/UFMG**
Prof. Mário S. Alvim **2017/02**

SOLUTION OF HOMEWORK

Discrete Probability - Part 2/2
(Based on slide-set)

Necessary reading for this assignment:

- *Slide-set of Lecture 01 - Discrete Probability*:
  - *Bayes' Theorem*
  - *Expected Value and Variance*

Note: The exercises are labeled according to their level of difficulty: `[Easy]`, `[Medium]` or `[Hard]`. This labeling, however, is subjective: different people may disagree on the perceived level of difficulty of any given exercise. Don't be discouraged when facing a hard exercise, you may find a solution that is simpler than the one the instructor had in mind!

Review questions.

1. (Rosen Review Question 7-6) `[Easy]`

   (a) Define the expected value of a random variable X.

   Instructor's solution: The expected value of a random variable X in a sample space S is

   $$E(X) = \sum_{s \in S} p(s)X(s) = \sum_{r \in X(S)} p(X = r) \cdot r.$$

   (b) What is the expected value of the random variable X that assigns to a roll of two dice the larger number that appears on the two dice?

   Instructor's solution: The sample space of rolling two dice can be represented by the following table, where the entry (i, j) is the result where the first die (d_1) comes up i and the second die (d_2) comes up j. Each entry (i, j) in the table contains the value of the random variable $X((i, j)) = \max(i, j)$.

d_1/d_2	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	2	3	4	5	6
3	3	3	3	4	5	6
4	4	4	4	4	5	6
5	5	5	5	5	5	6
6	6	6	6	6	6	6

Because each of the 36 outcomes is equally likely, $p((i,j)) = 1/36$ for every pair in the table, and the expectation of X can be calculated as

$$\begin{aligned}
E(XY) =& \sum_{i=1}^{6}\sum_{j=1}^{6} p((i,j)) \cdot \max(i,j) \\
=& \frac{1}{36}(1 + 3\cdot 2 + 5\cdot 3 + 7\cdot 4 + 9\cdot 5 + 11\cdot 6) \\
=& \frac{161}{36} \\
\approx& 4.47
\end{aligned}$$

2. (Rosen Review Question 7-8) `[Easy]`

   (a) What is meant by a "*Bernoulli trial*"?

   Instructor's solution: A Bernoulli trial is an expirement with two possible outcomes (often generally referred to as "success" and "failure").

   (b) What is the probability of k successes in n independent Bernoulli trials?

   Instructor's solution: If p is the probability of success in one Bernoulli trial, the probability of k successes in n independent Bernoulli trials is $C(n,k)p^k(1-p)^{n-k}$.

   (c) What is the expected value of the number of successes in n independent Bernoulli trials?

   Instructor's solution: The expected value of the number of successes in n independent Bernoulli trials is np, where p is the probability of success in one Bernoulli trial.

3. (Rosen Review Question 7-9) `[Easy]`

   (a) What does the linearity of expectations of random variables mean?

   Instructor's solution: The linearity of expecations of random variables means that expecatation is a linear function, in the sense that: (i) if X_i are random variables ($1 \le i \le n$) then $E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i)$; and (ii) if X is a random variable and a and b are real numbers, then $E(aX+b) = aE(X) + b$.

4. (Rosen Review Question 7-11) `[Easy]` State Bayes' theorem and use it to find $p(F \mid E)$ if $p(E \mid F) = 1/3$, $p(E \mid \overline{F}) = 1/4$, and $p(F) = 2/3$, where E and F are events from a sample space S.

Instructor's solution: Baye's Theorem says that if E and F are events in a sample space S, then

$$p(F \mid E) = \frac{p(F)p(E \mid F)}{p(E)} = \frac{p(F)p(E \mid F)}{p(F)p(E \mid F) + p(\overline{F})p(E \mid \overline{F})}.$$

Noting that $p(\overline{F}) = 1 - p(F) = 1/3$, we can substitute the given values in the above equation to obtain

$$\begin{aligned} p(F \mid E) &= \frac{p(F)p(E \mid F)}{(p(F)p(E \mid F) + p(\overline{F})p(E \mid \overline{F}))} \\ &= \frac{2/3 \cdot 1/3}{2/3 \cdot 1/3 + 1/3 \cdot 1/4} \\ &= \frac{8}{11}. \end{aligned}$$

5. (Rosen Review Question 7-13) `[Easy]`

   (a) What is the variance of a random variable?

   Instructor's solution: The variance of a random variable is a measure of how much, in average, its values spread around the expected value. The variance of a random variable X in a sample space S is given by $V(X) = \sum_{s \in S} p(s)[X(s) - E(X)]^2 = E(X^2) - [E(X)]^2$.

6. (Rosen Review Question 7-14) `[Easy]`

   (a) What is the variance of the sum of n independent random variables?

   Instructor's solution: The variance of n independent random variables is given by $V(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} V(X_i)$.

Exercises.

7. (Rosen 7.3-3) `[Medium]` Suppose that Frida selects a ball by first picking one of two boxes at random and then selecting a ball from this box at random. The first box contains two white balls and three blue balls, and the second box contains four white balls and one blue ball. What is the probability that Frida picked a ball from the first box if she has selected a blue ball?

   Instructor's solution: Let's use box_1 to represent the first box and box_2 to represent the second box.

   We know that each box is selected with equal probability, so $p(box_1) = p(box_2) = 1/2$.

   Also, let's use *white* and *blue* to represent the color of one ball.

   We know that $p(white \mid box_1) = 2/5$ and that $p(blue \mid box_1) = 3/5$.

   Also, we know that $p(white \mid box_2) = 4/5$ and that $p(blue \mid box_2) = 1/5$.

   We want to calculate $p(box_1 \mid blue)$, and for that we can use Bayes' Theorem as follows.

$$\begin{aligned} p(box_1 \mid blue) &= \frac{p(box_1)p(blue \mid box_1)}{p(blue)} \\ &= \frac{p(box_1)p(blue \mid box_1)}{p(box_1)p(blue \mid box_1) + p(box_2)p(blue \mid box_2)} \\ &= \frac{1/2 \cdot 3/5}{1/2 \cdot 3/5 + 1/2 \cdot 1/5} \\ &= \frac{3/10}{2/5} \\ &= \frac{3}{4}. \end{aligned}$$

8. (Rosen 7.3-9) `[Medium]` Suppose that 8% of the patients tested in a clinic are infected with HIV. Furthermore, suppose that when a blood test for HIV is given, 98% of the patients infected with HIV test positive and that 3% of the patients not infected with HIV test positive. What is the probability that

   (a) a patient testing positive for HIV with this test is infected with it?

   Instructor's solution: First let us formally model the problem, and calculate all the values needed to apply Bayes' theorem.
   Let I be the event that a patient is infected, and $\overline{I}$ be the event a patient is not infected.
   The prior probability is that $p(I) = 0.08$, hence $p(\overline{I}) = 0.92$.
   Let P be the event that the HIV test comes up positive, and $\overline{P}$ the event that HIV test comes up negative.
   We know that $p(P \mid I) = 0.98$, and hence $p(\overline{P} \mid I) = 0.02$. Also, we know that $p(P \mid \overline{I}) = 0.03$, and hence $p(\overline{P} \mid \overline{I}) = 0.97$.
   Now we calculate the probability of a patient testing positive and negative for HIV.

   $$p(P) = p(I)p(P \mid I) + p(\overline{I})p(P \mid \overline{I}) = 0.008 \cdot 0.98 + 0.92 \cdot 0.03 = 0.106$$
   $$p(\overline{P}) = p(I)p(\overline{P} \mid I) + p(\overline{I})p(\overline{P} \mid \overline{I}) = 0.008 \cdot 0.02 + 0.92 \cdot 0.97 = 0.894$$

   Now we are ready to apply Bayes' theorem to calculate the a patient testing positive for HIV when this patient is actually infected with the virus as follows.

   $$p(I \mid P) = \frac{p(I)p(P \mid I)}{p(P)} = \frac{0.08 \cdot 0.98}{0.106} \approx 0.740$$

   (b) a patient testing positive for HIV with this test is not infected with it?

   Instructor's solution: Using the results from the previous item, we can calculate the following.

   $$p(\overline{I} \mid P) = \frac{p(\overline{I})p(P \mid \overline{I})}{p(P)} = \frac{0.92 \cdot 0.03}{0.106} \approx 0.260$$

   (c) a patient testing negative for HIV with this test is infected with it?

   Instructor's solution: Using the results from the previous item, we can calculate the following.

   $$p(I \mid \overline{P}) = \frac{p(I)p(\overline{P} \mid I)}{p(\overline{P})} = \frac{0.08 \cdot 0.02}{0.894} \approx 0.002$$

   (d) a patient testing negative for HIV with this test is not infected with it?

   Instructor's solution: Using the results from the previous item, we can calculate the following.

   $$p(\overline{I} \mid \overline{P}) = \frac{p(\overline{I})p(\overline{P} \mid \overline{I})}{p(\overline{P})} = \frac{0.92 \cdot 0.97}{0.894} \approx 0.998$$

9. (Rosen 7.4-3) `[Easy]` What is the expected number of times a 6 appears when a fair die is rolled 10 times?

Instructor's solution: Let X_i, with $1 \leq i \leq 10$, be the random variable modeling the number of times a number 6 occurs in the i^{th} roll of the die.

We know that $E(X_i) = \frac{1}{6}$ for $1 \leq i \leq 10$.

We are interested on the expected number of times a 6 appears in the random variable $E(X_1 + X_2 + \ldots + X_{10})$. By the linearity of expectations we have:

$$
\begin{aligned}
E(X_1 + X_2 + \ldots + X_n) =& \; E(X_1) + E(X_2) + \ldots + E(X_{10}) \\
=& \; 10 \cdot \frac{1}{6} \\
=& \; \frac{5}{3}
\end{aligned}
$$

10. (Rosen 7.4-7) `[Easy]` The final exam of a discrete mathematics course consists of 50 true/false questions, each worth two points, and 25 multiple-choice questions, each worth four points. The probability that Linda answers a true/false question correctly is 0.9, and the probability that she answers a multiple-choice question correctly is 0.8. What is her expected score on the final?

Instructor's solution: Let i range over $1, \ldots, 50$ and let X_i be the random variable modeling Linda's score when answering the i^{th} yes/no question, that is, X_i takes value 2 if the i^{th} yes/no question is answered correctly, and value 0 otherwise.

We know that the probability of Linda answering a yes/no question correctly is 0.9, hence $P(X_i = 2) = 0.9$ and $P(X_i = 0) = 0.1$.

The expected value of each X_i can be calculated as

$$E(X_i) = p(X_i = 2) \cdot 2 + p(X_i = 0) \cdot 0 = 0.9 \cdot 2 + 0.1 \cdot 0 = 1.8$$

Let j range over $1, \ldots, 25$ and let Y_i be the random variable modeling Linda's score when answering the j^{th} multiple-choice question, that is, Y_j takes value 4 if the j^{th} yes/no question is answered correctly, and value 0 otherwise.

We know that the probability of Linda answering a multiple-choice question correctly is 0.8, hence $P(Y_j = 4) = 0.8$ and $P(Y_j = 0) = 0.2$.

The expected value of each Y_j can be calculated as

$$E(Y_j) = p(Y_j = 4) \cdot 4 + p(Y_j = 0) \cdot 0 = 0.8 \cdot 4 + 0.2 \cdot 0 = 3.2$$

Let S be the random variable representing Linda's final score, that is $S = \sum_i^{50} X_i + \sum_{j=1}^{25} Y_j$. The expectation of S can be calculated as follows.

$$
\begin{aligned}
E(S) =& E\left(\sum_i^{50} X_i + \sum_{j=1}^{25} Y_j\right) \\
=& \sum_i^{50} E(X_i) + \sum_{j=1}^{25} E(Y_j) \qquad \text{(by the linearity of expecations)} \\
=& \sum_i^{50} 1.8 + \sum_{j=1}^{25} 3.2 \qquad \text{(as we calculated above)} \\
=& 50 \cdot 1.8 + 25 \cdot 3.2 \\
=& 170.
\end{aligned}
$$

11. (Rosen 7.4-19) `[Hard]` Let X be the number appearing on the first die when two fair dice are rolled and let Y be the sum of the numbers appearing on the two dice. Show that $E(X)E(Y) \neq E(XY)$.

Instructor's solution: If the die if fair, the expectation of X is

$$E(X) = \frac{1}{6} \cdot 1 + \frac{1}{6} \cdot 2 + \frac{1}{6} \cdot 3 + \frac{1}{6} \cdot 4 + \frac{1}{6} \cdot 5 + \frac{1}{6} \cdot 6 = \frac{7}{2}.$$

The random variable Y can be written as $Y = X_1 + X_2$, where X_1 is the result of the first die, and X_2 is the result of the second die. Hence, the expectation of Y can be written as

$$\begin{aligned}
E(Y) =& E(X_1 + X_2) \\
=& E(X_1) + E(X_2) && \text{(by the linearity of expecations)} \\
=& \frac{7}{2} + \frac{7}{2} && \text{(because the expectation of each fair die is 7/2)} \\
=& 7
\end{aligned}$$

Now let's calculate the expectation of XY. The sample space of rolling two dice can be represented by the following table, where the entry (i, j) is the result where the first die (d_1) comes up i and the second die (d_2) comes up j. We know that $X((i, j)) = i$ and $Y((i, j)) = i + j$, hence XY takes the value $i \cdot (i + j)$ for each pair (i, j).

d_1/d_2	1	2	3	4	5	6
1	2	3	4	5	6	7
2	6	8	20	12	14	16
3	12	15	18	21	24	27
4	20	24	28	32	36	40
5	30	35	40	45	50	55
6	12	48	54	60	66	72

Because each of the 36 outcomes is equally likely, $p((i, j)) = 1/36$ for every pair in the table, and the expectation of XY can be calculated as

$$\begin{aligned}
E(XY) =& \sum_{i=1}^{6} \sum_{j=1}^{6} p((i, j)) \cdot i \cdot (i + j) \\
=& \frac{1}{36} \sum_{i=1}^{6} \sum_{j=1}^{6} i \cdot (i + j) \\
=& \frac{329}{12}.
\end{aligned}$$

To finish the question, we have that $E(X)E(Y) = 7/2 \cdot 7 = 49/2 = 294/12$, which is different from $E(XY) = 319/12$. (This shows that X and Y are not independent.)

12. (Rosen 7.4-27) What is the variance of the number of heads that come up when a fair coin is flipped 10 times?

Instructor's solution: Let X_i, with $1 \leq x \leq 10$ be the random variable representing the number of heads in the i^{th} flip of a coin. That is, $X_i(t) = 1$ if t is heads and $X_i(t) = 0$ it t is tails. Because the coin is fair, we have that $p(X_i = 1) = \frac{1}{2}$ and $p(X_i = 0) = \frac{1}{2}$.

We are interested in the random variable $X = \sum_{i=1}^{10} X_i$ representing the number of heads when a fair coin is flipped 10 times.

We want the variance of X, which can be calculated as

$$\begin{aligned}
V(X) &= V\left(\sum_{i=1}^{10} X_i\right) \\
&= \sum_{i=1}^{10} V(X_i) && \text{(since every } X_i \text{ is independent of each other).}
\end{aligned}$$

Now we have to calculate the variance $V(X_i)$, which for every $1 \leq i \leq 10$ is

$$\begin{aligned}
V(X_i) =& E(X_i^2) - [E(X_i)]^2 \\
=& p(X_i = 1) \cdot (1)^2 + p(X_i = 0) \cdot (0)^2 - [p(X_i = 1) \cdot 1 + p(X_i = 0) \cdot 0]^2 \\
=& \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 0 - \left[\frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 0\right]^2 \\
=& \frac{1}{2} - \left[\frac{1}{2}\right]^2 \\
=& \frac{1}{4}.
\end{aligned}$$

Finally, we can replace the value of $V(X_i)$ in the formula for $V(X)$ and derive

$$\begin{aligned}
V(X) &= V\left(\sum_{i=1}^{10} X_i\right) \\
&= \sum_{i=1}^{10} V(X_i) && \text{(because every } X_i \text{ is independent of each other)} \\
&= \sum_{i=1}^{10} \frac{1}{4} && \text{(as we calculated above)} \\
&= \frac{10}{4} \\
&= \frac{5}{2}.
\end{aligned}$$

Hence, the variance of the number of heads that come up when a fair coin is flipped 10 times is $5/2$.