



Model Answers to Exercise 1 of September 20, 2017

<http://www.isi.ee.ethz.ch/teaching/courses/it1.html>

Problem 1

Expectation of a Chance Variable

- a) Note first that $\mathbb{E}[g(X)]$ is defined as a summation over all $x \in \mathcal{X}$ such that $P_X(x) > 0$. This implies that for those x the term $\frac{1}{P_X(x)}$ is finite and therefore we may write

$$\mathbb{E}\left[\frac{1}{P_X(X)}\right] = \sum_{x \in \text{supp}(P_X)} P_X(x) \frac{1}{P_X(x)} = \sum_{x \in \text{supp}(P_X)} 1 = |\text{supp}(P_X)| = L'.$$

- b) Simply by using the definition of $\mathbb{E}[g(X)]$ we obtain

$$\begin{aligned}\mathbb{E}[P_X(X)] &= \sum_{x \in \text{supp}(P_X)} P_X(x) \cdot P_X(x) = \sum_{x \in \text{supp}(P_X)} (P_X(x))^2, \\ \mathbb{E}[P_{X'}(X)] &= \sum_{x \in \text{supp}(P_X)} P_X(x) \cdot P_{X'}(x).\end{aligned}$$

(Note that we are interested in the expectation of a function of X . Within the expectation operator $\mathbb{E}[\cdot]$, the terms $P_X(\cdot)$ and $P_{X'}(\cdot)$ are just functions which do not have any connection to the chance variable X on their own.)

- c) Similarly, we have

$$\begin{aligned}\mathbb{E}[-\log P_X(X)] &= - \sum_{x \in \text{supp}(P_X)} P_X(x) \log P_X(x) = H(X), \\ \mathbb{E}[-\log P_{X'}(X)] &= - \sum_{x \in \text{supp}(P_X)} P_X(x) \log P_{X'}(x).\end{aligned}$$

Remarks:

- Note the difference between X (chance variable) and x (*value* of a chance variable).
- In the formula

$$\mathbb{E}[g(X)] = \sum_{x \in \text{supp}(P_X)} P_X(x) g(x),$$

the three symbols marked with an arrow must be the same, i.e., if we replace X by Z on the left-hand side of the equality, we must also change X to Z at both places on the right. On the right-hand side, x is just a dummy variable and could, therefore, be exchanged for any other variable name, e.g., ξ . However, it is common to take little x as a dummy variable for the chance variable X .

Problem 2***On the Expectation of a Discrete Random Variable***

Here we want to take advantage of the fact that the random variable T takes on only positive integer values. In this case, we can write

$$\begin{aligned} \sum_{v=1}^{\infty} \Pr[T \geq v] &= \underbrace{\Pr[T = 1] + \Pr[T = 2] + \dots}_{\Pr[T \geq 1]} \\ &\quad + \underbrace{\Pr[T = 2] + \Pr[T = 3] + \dots}_{\Pr[T \geq 2]} \\ &\quad + \underbrace{\Pr[T = 3] + \Pr[T = 4] + \dots}_{\Pr[T \geq 3]} \\ &\quad + \dots \end{aligned}$$

By rearranging the above terms, we obtain

$$\begin{aligned} \sum_{v=1}^{\infty} \Pr[T \geq v] &= \Pr[T = 1] + 2 \cdot \Pr[T = 2] + 3 \cdot \Pr[T = 3] + \dots \\ &= \sum_{t=1}^{\infty} t \Pr[T = t] \\ &= \mathbf{E}[T]. \end{aligned}$$

Another way to see this is the following:

$$\begin{aligned} \mathbf{E}[T] &= \sum_{v=1}^{\infty} v \Pr[T = v] \\ &= \sum_{v=1}^{\infty} v (\Pr[T \geq v] - \Pr[T \geq v + 1]) \\ &= \sum_{v=1}^{\infty} v \Pr[T \geq v] - \sum_{v=2}^{\infty} (v - 1) \Pr[T \geq v] \\ &= \Pr[T \geq 1] + \sum_{v=2}^{\infty} (v - (v - 1)) \Pr[T \geq v] \\ &= \Pr[T \geq 1] + \sum_{v=2}^{\infty} \Pr[T \geq v] \\ &= \sum_{v=1}^{\infty} \Pr[T \geq v]. \end{aligned}$$

Problem 3***Statistical Independence***

Recall that k discrete random variables Y_1, \dots, Y_k are statistically independent if and only if

$$\Pr[Y_1 = y_1, \dots, Y_k = y_k] = \prod_{i=1}^k \Pr[Y_i = y_i] \quad \forall y_1, \dots, y_k.$$

Observe that the parity check on IID Bernoulli(1/2) random variables is again a Bernoulli(1/2) random variable:

$$\Pr[Z = z] = \Pr[X_1 \oplus X_2 \oplus \dots \oplus X_n = z]$$

$$\begin{aligned}
& \stackrel{(i)}{=} \sum_{x_1 \in \{0,1\}} \Pr[X_1 \oplus X_2 \oplus \cdots \oplus X_n = z | X_1 = x_1] \cdot \Pr[X_1 = x_1] \\
& = \sum_{x_1 \in \{0,1\}} \Pr[X_2 \oplus \cdots \oplus X_n = z \oplus x_1 | X_1 = x_1] \cdot \Pr[X_1 = x_1] \\
& \stackrel{(ii)}{=} \sum_{x_1 \in \{0,1\}} \Pr[X_2 \oplus \cdots \oplus X_n = z \oplus x_1] \cdot \Pr[X_1 = x_1] \\
& = \sum_{x_1 \in \{0,1\}} \Pr[X_2 \oplus \cdots \oplus X_n = z \oplus x_1] \cdot \frac{1}{2} \\
& \stackrel{(iii)}{=} \frac{1}{2},
\end{aligned}$$

where (i) follows from the law of total probability; (ii) holds because $X_2 \oplus \cdots \oplus X_n$ is independent of X_1 ; and (iii) holds because $\Pr[X_2 \oplus \cdots \oplus X_n = 0] + \Pr[X_2 \oplus \cdots \oplus X_n = 1] = 1$.

a) For any choice of $z, x_1 \in \{0, 1\}$, we have

$$\begin{aligned}
\Pr[Z = z, X_1 = x_1] &= \Pr[X_1 \oplus X_2 \oplus \cdots \oplus X_n = z, X_1 = x_1] \\
&= \Pr[X_2 \oplus \cdots \oplus X_n = z \oplus x_1, X_1 = x_1] \\
&\stackrel{(i)}{=} \Pr[X_2 \oplus \cdots \oplus X_n = z \oplus x_1] \cdot \Pr[X_1 = x_1] \\
&\stackrel{(ii)}{=} \frac{1}{2} \cdot \Pr[X_1 = x_1] \\
&\stackrel{(iii)}{=} \Pr[Z = z] \cdot \Pr[X_1 = x_1],
\end{aligned}$$

where (i) holds because $X_2 \oplus \cdots \oplus X_n$ is independent of X_1 ; (ii) and (iii) hold because the parity check on IID Bernoulli(1/2) random variables is again a Bernoulli(1/2) random variable. We conclude that Z and X_1 are statistically *independent*.

b) For any choice of $z, x_1, \dots, x_{n-1} \in \{0, 1\}$, we have

$$\begin{aligned}
& \Pr[Z = z, X_1 = x_1, \dots, X_{n-1} = x_{n-1}] \\
&= \Pr[X_1 = x_1, X_2 = x_2, \dots, X_{n-1} = x_{n-1}, X_n = z \oplus x_1 \oplus \cdots \oplus x_{n-1}] \\
&\stackrel{(i)}{=} \Pr[X_1 = x_1] \cdot \Pr[X_2 = x_2] \cdots \Pr[X_{n-1} = x_{n-1}] \cdot \Pr[X_n = z \oplus x_1 \oplus \cdots \oplus x_{n-1}] \\
&= \left(\frac{1}{2}\right)^n \\
&\stackrel{(ii)}{=} \Pr[Z = z] \cdot \Pr[X_1 = x_1] \cdots \Pr[X_{n-1} = x_{n-1}],
\end{aligned}$$

where (i) holds because X_1, \dots, X_n are independent and (ii) holds since Z is a Bernoulli(1/2) random variable. We conclude that Z, X_1, \dots, X_{n-1} are statistically *independent*.

c) The probability $\Pr[Z = 1, X_1 = 0, X_2 = 0, \dots, X_n = 0]$ is obviously 0, however

$$\Pr[Z = 1] \cdot \Pr[X_1 = 0] \cdot \Pr[X_2 = 0] \cdots \Pr[X_n = 0] = \left(\frac{1}{2}\right)^{n+1} \neq 0.$$

Therefore, Z, X_1, \dots, X_n are statistically *dependent*.

d) Let $\Pr[X_i = 1] = p$ and $\Pr[X_i = 0] = 1 - p$. With $n = 2$, we find

$$\begin{aligned}
\Pr[Z = 0] &= (1 - p)^2 + p^2 &= 1 - 2p + 2p^2, \\
\Pr[Z = 1] &= p(1 - p) + p(1 - p) &= 2p - 2p^2.
\end{aligned}$$

We want to find all values of p for which Z and X_1 are statistically independent. First, look at the case when $Z = 0$ and $X_1 = 0$. Then

$$\begin{aligned}\Pr[Z = 0, X_1 = 0] &= \Pr[X_1 = 0, X_2 = 0] \\ &= \Pr[X_1 = 0] \cdot \Pr[X_2 = 0] \\ &= (1 - p)^2,\end{aligned}\tag{1}$$

whereas

$$\Pr[Z = 0] \cdot \Pr[X_1 = 0] = (1 - 2p + 2p^2)(1 - p).\tag{2}$$

For Z and X_1 to be independent, (1) and (2) must be equal, which leads to the following condition on p :

$$\begin{aligned}(1 - p)^2 &\stackrel{!}{=} (1 - 2p + 2p^2)(1 - p) \\ (1 - p)^2 - (1 - 2p + 2p^2)(1 - p) &= 0 \\ (1 - p)(1 - p - 1 + 2p - 2p^2) &= 0 \\ (1 - p)(p - 2p^2) &= 0 \\ p(1 - p)(1 - 2p) &= 0.\end{aligned}$$

Obviously, this last equation has three solutions: $p = 0$, $p = 1$, and $p = \frac{1}{2}$. The first two solutions correspond to the trivial cases where the random variables X_1 , X_2 , and Z have fixed values and therefore are statistically independent. The third solution is the case from Part a). As we cannot find any $p \notin \{0, 1/2, 1\}$ for which (1) equals (2), there is no need to investigate the other cases ($\{Z = 0, X_1 = 1\}$, $\{Z = 1, X_1 = 0\}$ or $\{Z = 1, X_1 = 1\}$), and we conclude that for $p \notin \{0, 1/2, 1\}$, Z and X_1 are statistically *dependent*.

Problem 4

Markov's Inequality and Chebyshev's Inequality

- a) We first consider the case where X is a discrete random variable. We take the definition of the expectation and split the summation into two parts:

$$\mathbb{E}[X] = \sum_x x P_X(x) = \underbrace{\sum_{x: x < \delta} x P_X(x)}_{\geq 0} + \sum_{x: x \geq \delta} x P_X(x).$$

As X is nonnegative, the first part of the summation is also nonnegative and we can write

$$\mathbb{E}[X] \geq \sum_{x: x \geq \delta} x P_X(x) \geq \sum_{x: x \geq \delta} \delta P_X(x) = \delta \underbrace{\sum_{x: x \geq \delta} P_X(x)}_{=\Pr[X \geq \delta]}.\tag{3}$$

For the case where X is a continuous nonnegative random variable with probability density $f_X(x)$, we have

$$\mathbb{E}[X] \geq \int_{\delta}^{\infty} x f_X(x) dx \geq \delta \int_{\delta}^{\infty} f_X(x) dx = \delta \cdot \Pr[X \geq \delta].\tag{4}$$

It follows from (3) and (4) that

$$\Pr[X \geq \delta] \leq \frac{\mathbb{E}[X]}{\delta}.$$

An example of a random variable X that achieves equality is the binary random variable with $P_X(0) = p$ and $P_X(1) = 1 - p$ for any $p \in [0, 1]$. Let $\delta = 1$ and thus

$$\Pr[X \geq \delta] = \Pr[X \geq 1] = 1 - p,$$

and

$$\frac{\mathbf{E}[X]}{\delta} = \frac{1 - p}{1} = 1 - p.$$

b) Since $X = (Y - \mu)^2$ is nonnegative, we can apply *Markov's Inequality*, which yields

$$\begin{aligned} \Pr[|Y - \mu| \geq \epsilon] &= \Pr[|Y - \mu|^2 \geq \epsilon^2] \\ &= \Pr[(Y - \mu)^2 \geq \epsilon^2] \\ &= \Pr[X \geq \epsilon^2] \\ &\stackrel{\text{a)}}{\leq} \frac{\mathbf{E}[X]}{\epsilon^2} \\ &= \frac{\mathbf{E}[(Y - \mu)^2]}{\epsilon^2} = \frac{\sigma^2}{\epsilon^2}. \end{aligned}$$

c) We know that Z_1, Z_2, \dots, Z_n are IID and that

$$\begin{aligned} \mathbf{E}[Z_k] &= \mu, \\ \text{Var}(Z_k) &= \mathbf{E}[(Z_k - \mu)^2] = \sigma^2. \end{aligned}$$

Therefore, we can calculate the mean and variance of \bar{Z}_n :

$$\mathbf{E}[\bar{Z}_n] = \mathbf{E}\left[\frac{1}{n} \sum_{k=1}^n Z_k\right] = \frac{1}{n} \sum_{k=1}^n \mathbf{E}[Z_k] = \mu$$

and

$$\begin{aligned} \text{Var}(\bar{Z}_n) &= \text{Var}\left(\frac{1}{n} \sum_{k=1}^n Z_k\right) \\ &\stackrel{\text{(i)}}{=} \frac{1}{n^2} \text{Var}\left(\sum_{k=1}^n Z_k\right) \\ &\stackrel{\text{(ii)}}{=} \frac{1}{n^2} \underbrace{\sum_{k=1}^n \text{Var}(Z_k)}_{=n\sigma^2} \\ &= \frac{\sigma^2}{n}, \end{aligned}$$

where (i) holds because $\text{Var}(\alpha X) = \alpha^2 \text{Var}(X)$ holds for all constants α and for all random variables X and (ii) holds because the variance of a sum of *independent* random variables is the sum of their variances (in fact, it is easy to show that for all finite-variance random variables X_1 and X_2 , $\text{Var}(X_1 + X_2) = \text{Var}(X_1) + \text{Var}(X_2) + 2 \text{Cov}(X_1, X_2)$ holds).

Now we can apply *Chebyshev's Inequality* for \bar{Z}_n to obtain

$$\begin{aligned} \Pr[|\bar{Z}_n - \mu| \geq \epsilon] &= \Pr[|\bar{Z}_n - \mathbf{E}[\bar{Z}_n]| \geq \epsilon] \\ &\stackrel{\text{b)}}{\leq} \frac{\text{Var}(\bar{Z}_n)}{\epsilon^2} \\ &= \frac{\sigma^2}{n\epsilon^2}. \end{aligned}$$