Problem 1

**Expectation of a Chance Variable**

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

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

b) By simply using the definition of \( E[g(X)] \) we obtain

\[
E[P_X(X)] = \sum_{x \in \text{supp}(P_X)} P_X(x) \cdot P_X(x) = \sum_{x \in \text{supp}(P_X)} (P_X(x))^2,
\]

\[
E[P_{X'}(X)] = \sum_{x \in \text{supp}(P_X)} P_X(x) \cdot P_{X'}(x).
\]

(Note that within the expectation operator \( E[\cdot] \) the terms \( P_X(\cdot) \) and \( P_{X'}(\cdot) \) are just names of functions; by themselves they have no connection to the chance variable \( X \). Here we are interested in the expectation of a function of \( X \).)

c) Similarly to Part a) we have

\[
E[-\log P_X(X)] = - \sum_{x \in \text{supp}(P_X)} P_X(x) \log P_X(x) = H(X),
\]

\[
E[-\log P_{X'}(X)] = - \sum_{x \in \text{supp}(P_X)} P_X(x) \log P_{X'}(x).
\]

**Remarks:**

- Note the difference between \( X \) (chance variable) and \( x \) (value of a chance variable).
- In the formula
  \[
  E[g(X)] = \sum_{x \in \text{supp}(P_X)} P_X(x) g(x),
  \]

the three symbols marked with an arrow must be the same, i.e., if on the left-hand side of the equality we replace \( X \) by \( Z \), we also must 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., \( \xi \). However, it is common to take little \( x \) as a dummy variable for the chance variable \( X \).
Problem 2

Two random variables $X$ and $Y$ are statistically independent if

$$P_{X,Y}(x,y) = P_X(x)P_Y(y), \quad \forall x \in \mathcal{X}, y \in \mathcal{Y}.$$  

It can be shown by induction that $\Pr[Z = 0] = \Pr[Z = 1] = 1/2$.

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

$$\Pr[Z = z, X_1 = x_1] = \Pr[X_1 \oplus X_2 \oplus \cdots \oplus X_n = z, X_1 = x_1]$$

$$= \Pr[X_1 = x_1] \cdot \Pr[X_1 \oplus X_2 \oplus \cdots \oplus X_n = z | X_1 = x_1]$$

$$= \Pr[X_1 = x_1] \cdot \Pr[X_2 \oplus X_3 \oplus \cdots \oplus X_n = z \oplus x_1 | X_1 = x_1]$$

$$= \Pr[X_1 = x_1] \cdot \Pr[X_2 \oplus X_3 \oplus \cdots \oplus X_n = z \oplus x_1]$$

where in the second equality we used the independence of $X_1, \ldots, X_n$. We can conclude that $Z$ and $X_1$ are statistically independent.

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

$$\Pr[Z = z, X_1 = x_1, \ldots, X_{n-1} = x_{n-1}]$$

$$= \Pr[X_1 = x_1, X_2 = \ldots, X_{n-1} = x_{n-1}, X_n = z \oplus x_1 \oplus \cdots \oplus x_{n-1}]$$

$$= \Pr[X_1 = x_1] \Pr[X_2 = x_2] \cdots \Pr[X_{n-1} = x_{n-1}] \Pr[X_n = z \oplus x_1 \cdots \oplus x_{n-1}]$$

$$= \left( \frac{1}{2} \right)^n$$

$$= \Pr[Z = z] \cdot \Pr[X_1 = x_1] \cdots \Pr[X_{n-1} = x_{n-1}],$$

where in the second equality we used the independence of $X_1, \ldots, X_n$. We can conclude that $Z, X_1, \ldots, X_{n-1}$ are statistically independent.

c) The probability $\Pr[Z = 1, X_1 = 0, X_2 = 0, \ldots, X_n = 0]$ 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, $Z, X_1, \ldots, X_n$ are statistically dependent.

d) Let $\Pr[X_i = 0] = p$ and $\Pr[X_i = 1] = 1 - p$. Then, under the assumption that $n = 2$, we find

$$\Pr[Z = 0] = p^2 + (1 - p)^2 = 1 - 2p + 2p^2$$

$$\Pr[Z = 1] = p(1 - p) + p(1 - p) = 2p - 2p^2.$$

We want to find all values of $p$ for which $Z$ and $X_1$ are statistically independent. Firstly, look at the case when $Z = 0$ and $X_1 = 1$. Then

$$\Pr[Z = 0, X_1 = 1] = \Pr[X_1 = 1, X_2 = 1]$$

$$= \Pr[X_1 = 1] \cdot \Pr[X_2 = 1]$$

$$= (1 - p)^2$$

whereas

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

© Amos Lapidoth, 2014/2015
If \( X_1 \) and \( Z \) are independent, then (1) and (2) must be equal. This leads to the following condition on \( p \) for independent \( X_1 \) and \( Z \):

\[
(1 - p)^2 \frac{1}{(1 - 2p + 2p^2)(1 - p)}
\]

\[
(1 - p)^2 - (1 - 2p + 2p^2)(1 - p) = 0
\]

\[
(1 - p)(1 - p - 1 + 2p - 2p^2) = 0
\]

\[
(1 - p)(p - 2p^2) = 0
\]

\[
p(1 - p)(1 - 2p) = 0.
\]

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 (\( \{Z = 0, X_1 = 0\}, \{Z = 1, X_1 = 1\} \) or \( \{Z = 1, X_1 = 0\} \)), and we can conclude that for \( p \notin \{0, 1/2, 1\} \), \( Z \) and \( X_1 \) are statistically dependent.

**Problem 3**

**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

\[
\sum_{v=1}^{\infty} \Pr[T \geq v] = \underbrace{\Pr[T = 1]}_{\Pr[T \geq 1]} + \underbrace{\Pr[T = 2]}_{\Pr[T \geq 2]} + \underbrace{\Pr[T = 3]}_{\Pr[T \geq 3]} + \cdots
\]

Then, by rearranging the above terms we obtain

\[
\sum_{v=1}^{\infty} \Pr[T \geq v] = \Pr[T = 1] + 2 \cdot \Pr[T = 2] + 3 \cdot \Pr[T = 3] + \cdots
\]

\[
= \sum_{t=1}^{\infty} t \Pr[T = t]
\]

\[
= \mathbb{E}[T].
\]

Another way to see this is as follows:

\[
\mathbb{E}[T] = \sum_{v=1}^{\infty} v \Pr[T = v]
\]

\[
= \sum_{v=1}^{\infty} v (\Pr[T \geq v] - \Pr[T \geq v + 1])
\]

\[
= \sum_{v=1}^{\infty} v \Pr[T \geq v] - \sum_{v=2}^{\infty} (v - 1) \Pr[T \geq v]
\]

© Amos Lapidoth, 2014/2015
\[ \begin{align*}
&= \Pr[T \geq 1] + \sum_{v=2}^{\infty} (v - (v - 1)) \Pr[T \geq v] \\
&= \Pr[T \geq 1] + \sum_{v=2}^{\infty} \Pr[T \geq v] \\
&= \sum_{v=1}^{\infty} \Pr[T \geq v].
\end{align*} \]

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:

\[ \begin{align*}
\mathbb{E}[X] &= \sum_{x} x P_X(x) = \underbrace{\sum_{x < \delta} x P_X(x)}_{\geq 0} + \underbrace{\sum_{x \geq \delta} x P_X(x)}_{\geq 0}.
\end{align*} \]

As \( X \) is nonnegative, the first part of the summation is also nonnegative and thus we can write

\[ \mathbb{E}[X] \geq \sum_{x \geq \delta} x P_X(x) \geq \sum_{x \geq \delta} \delta P_X(x) = \delta \sum_{x \geq \delta} P_X(x). \tag{3} \]

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

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

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

\[ \Pr[|Y - \mu| \geq \varepsilon] = \Pr[|Y - \mu|^2 \geq \varepsilon^2] = \Pr[(Y - \mu)^2 \geq \varepsilon^2] = \Pr[X \geq \varepsilon^2] \leq \frac{\mathbb{E}[X]}{\varepsilon^2} = \frac{\mathbb{E}[(Y - \mu)^2]}{\varepsilon^2} = \frac{\sigma^2}{\varepsilon^2}. \]
c) We know that $Z_1, Z_2, \ldots, Z_n$ are IID and that

$$E[Z_k] = \mu,$$

$$\text{Var}(Z_k) = E[(Z_k - \mu)^2] = \sigma^2.$$ 

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

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

and

$$\text{Var}(\bar{Z}_n) = E\left[(\bar{Z}_n - \mu)^2\right]$$

$$= E\left[\left\{\frac{1}{n} \sum_{k=1}^{n} Z_k - \mu\right\}^2\right]$$

$$= E\left[\left\{\frac{1}{n} \sum_{k=1}^{n} (Z_k - \mu)\right\}^2\right]$$

$$= \frac{1}{n^2} E\left[\sum_{k=1}^{n} \sum_{\ell=1}^{n} (Z_k - \mu)(Z_\ell - \mu)\right]$$

$$= \frac{1}{n^2} \sum_{k=1}^{n} \sum_{\ell=1}^{n} E[(Z_k - \mu)(Z_\ell - \mu)]$$

$$= \frac{1}{n^2} \sum_{k=1}^{n} \text{Var}(Z_k) + \frac{1}{n^2} \sum_{k=1}^{n} \sum_{\ell=1, \ell \neq k}^{n} E[(Z_k - \mu)] E[(Z_\ell - \mu)]$$

$$= \frac{1}{n^2} \sum_{k=1}^{n} \sigma^2$$

$$= \frac{\sigma^2}{n}.$$ 

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

$$\Pr[|\bar{Z}_n - \mu| \geq \epsilon] = \Pr[|\bar{Z}_n - E[\bar{Z}_n]| \geq \epsilon]$$

$$\leq \frac{\text{Var}(\bar{Z}_n)}{\epsilon^2}$$

$$= \frac{\sigma^2}{n\epsilon^2}.$$