The following example is a variant of *Example 2.18 The Two-Envelopes Paradox* from the text book. Reading the that example before reading the following discussion might be a good idea.

Example 1. Alice observes an integer valued random variable X with the following Geometric probability mass function:

$$p_{\mathsf{X}}(k) = \begin{cases} p(1-p)^{k-1} & k \in \{1,2,\dots\} \\ 0 & k \in \{0,-1,\dots\} \end{cases}. \tag{1}$$

If $\mathsf{X} = k$, then Alice puts m^k liras in the envelope Y and m^{k+1} liras in the envelope Z with probability $1/2$ and m^{k+1} liras in the envelope Y and m^k liras in the envelope Z with probability $1/2$, where m is an integer strictly greater than 1, i.e., $m \in \{2,3,\dots\}$. You know the values of parameters p and m. You open the envelope Y and observe that it has y liras in it. You can either keep y liras or switch envelopes and get the money in the envelope Z instead. If you want to maximize your expected earnings from this game, for which values of y will you switch envelopes? Is there a value of p and m for which you will always switch envelopes?

Solution. With a slight abuse of notation let use denote the amount money —in liras— in the envelope Y by Y and in the envelope Z by Z. The expected amount of money we get by switching to the envelope Z is $\mathbf{E}[\mathsf{Z}|\,\mathsf{Y} = y]$. Thus in order to maximize our earnings we need to switch whenever $\mathbf{E}[\mathsf{Z}|\,\mathsf{Y} = y] > y$.

Remark 1. Whenever $\mathbf{E}[\mathsf{Z}|\,\mathsf{Y} = y] = y$ holds, switching to the envelope Z will not change the expected earnings. Thus if we switch to the envelope Z whenever $\mathbf{E}[\mathsf{Z}|\,\mathsf{Y} = y] \ge y$, the expected gain will be the same. Thus our rule for switching to the envelope Z can be $\mathbf{E}[\mathsf{Z}|\,\mathsf{Y} = y] > y$ or $\mathbf{E}[\mathsf{Z}|\,\mathsf{Y} = y] \ge y$.

In order to determine the values of y satisfying $\mathbf{E}[\mathsf{Z}|\,\mathsf{Y} = y] > y$, we need to calculate the expected value of Z given Y and hence the conditional PMF of Z given Y. Let us start with determining the conditional joint PMF of Y and Z given X, described in the question.

$$p_{\mathsf{Y},\mathsf{Z}|\mathsf{X}}(y,z|k) = \begin{cases} 1/2 & y = m^k, z = m^{k+1} \\ 1/2 & y = m^{k+1}, z = m^k \\ 0 & \text{else} \end{cases} \tag{2}$$

Recall that $p_{\mathsf{Y},\mathsf{Z},\mathsf{X}}(y,z,k) = p_{\mathsf{X}}(k)\, p_{\mathsf{Y},\mathsf{Z}|\mathsf{X}}(y,z|k)$ and $p_{\mathsf{Y},\mathsf{Z}}(y,z) = \sum_k p_{\mathsf{Y},\mathsf{Z},\mathsf{X}}(y,z,k)$. Thus using (1) and (2) we get

$$\begin{aligned} p_{\mathsf{Y},\mathsf{Z}}(y,z) &= \sum\nolimits_k p_{\mathsf{X}}(k)\, p_{\mathsf{Y},\mathsf{Z}|\mathsf{X}}(y,z|k) \\ &= \begin{cases} \frac{p(1-p)^{k-1}}{2} & y = m^k, z = m^{k+1} \text{ for some } k \in \{1,2,\dots\} \\ \frac{p(1-p)^{k-1}}{2} & y = m^{k+1}, z = m^k \text{ for some } k \in \{1,2,\dots\} \\ 0 & \text{else} \end{cases} \end{aligned} \tag{3}$$

Since $p_{\mathsf{Y}}(y) = \sum_z p_{\mathsf{Y},\mathsf{Z}}(y,z)$, (3) implies

$$p_{\mathsf{Y}}(y) = \begin{cases} \frac{p}{2} & y = 1 \\ \frac{p(1-p)^{k-1} + p(1-p)^{k-2}}{2} & y = m^k \text{ for some } k \in \{2,3,\dots\} \\ 0 & \text{else} \end{cases} \tag{4}$$

On the other hand $p_{\mathsf{Z}|\mathsf{Y}}(z|y) = \frac{p_{\mathsf{Y},\mathsf{Z}}(y,z)}{p_{\mathsf{Y}}(y)}$ for all y satisfying $p_{\mathsf{Y}}(y) > 0$. Thus (4) implies

$$\begin{aligned} p_{\mathsf{Z}|\mathsf{Y}}(z|y) &= \begin{cases} 1 & y = 1, z = m \\ \frac{(1-p)^{k-1}}{(1-p)^{k-1} + (1-p)^{k-2}} & y = m^k, z = m^{k+1} \text{ for some } k \in \{2,3,\dots\} \\ \frac{(1-p)^{k-2}}{(1-p)^{k-1} + (1-p)^{k-2}} & y = m^k, z = m^{k-1} \text{ for some } k \in \{2,3,\dots\} \\ 0 & \text{else} \end{cases} \\ &= \begin{cases} 1 & y = 1, z = m \\ \frac{1-p}{2-p} & y = m^k, z = m^{k+1} \text{ for some } k \in \{2,3,\dots\} \\ \frac{1}{2-p} & y = m^k, z = m^{k-1} \text{ for some } k \in \{2,3,\dots\} \\ 0 & \text{else} \end{cases}. \end{aligned} \tag{5}$$

Recall that $\mathbf{E}[\mathsf{Z}|\,\mathsf{Y} = y] = \sum_z z p_{\mathsf{Z}|\mathsf{Y}}(z|y)$. Thus (5) implies

$$\begin{aligned} \mathbf{E}[\mathsf{Z}|\,\mathsf{Y} = y] &= \begin{cases} m & y = 1 \\ ym\frac{1-p}{2-p} + \frac{y}{m}\frac{1}{2-p} & y = m^k \text{ for some } k \in \{2,3,\dots\} \end{cases} \\ &= \begin{cases} m & y = 1 \\ y\frac{(1-p)m^2+1}{(2-p)m} & y = m^k \text{ for some } k \in \{2,3,\dots\} \end{cases} \end{aligned} \tag{6}$$

- If $y = 1$ we switch to the envelope Z, because $\mathbf{E}[\mathsf{Z}|\,\mathsf{Y} = 1] = m > 1$ by (6). In fact we know that it has exactly m liras with probability 1 by (5).
- If $y = m^k$ for some $k \in \{2, 3, \ldots\}$, then $\mathbf{E}[\mathsf{Z}|\,\mathsf{Y} = y] > y$ holds if and only if $\frac{(1-p)m^2+1}{(2-p)m} > 1$ by (6). On the other hand, $\frac{(1-p)m^2+1}{(2-p)m} > 1$ holds if and only if $(1-p)m^2 - (2-p)m + 1 > 0$. Note that

$$(1-p)m^2 - (2-p)m + 1 = ((1-p)m - 1)(m - 1).$$

Thus $m = 1$ and $m = \frac{1}{1-p}$ are the roots of the equation $(1-p)m^2-(2-p)m+1 = 0$. Thus $(1-p)m^2-(2-p)m+1 > 0$ if and only if $m > \frac{1}{1-p}$ because $m > 1$ by the hypothesis. Thus
  - If $m > \frac{1}{1-p}$, then $\mathbf{E}\left[\mathsf{Z}|\,\mathsf{Y} = \mathsf{m}^{\mathsf{k}}\right] > m^k$ for all $k \in \{2, 3, \ldots\}$. Thus we need to switch to the envelope Z for all $k \in \{2, 3, \ldots\}$.
  - If $m = \frac{1}{1-p}$, then $\mathbf{E}\left[\mathsf{Z}|\,\mathsf{Y} = \mathsf{m}^{\mathsf{k}}\right] = m^k$ for all $k \in \{2, 3, \ldots\}$. Thus we may switch to the envelope Z or retain the envelope Y for all $k \in \{2, 3, \ldots\}$.
  - If $m < \frac{1}{1-p}$, then $\mathbf{E}\left[\mathsf{Z}|\,\mathsf{Y} = \mathsf{m}^{\mathsf{k}}\right] < m^k$ for all $k \in \{2, 3, \ldots\}$. Thus we need to retain the envelope Y for all $k \in \{2, 3, \ldots\}$.

Thus a strategy that maximize the expected earnings is given as follows for different values of the parameters m and p:

- If $m > \frac{1}{1-p}$, then we will always (i.e. for all values of y satisfying $p_{\mathsf{Y}}(y) > 0$) switch to the envelope Z.
- If $1 < m \leq \frac{1}{1-p}$, then we will switch to the envelope Z if $y = 1$ and retain the envelope Y else.

Remark 2. It is easy to that $\mathbf{E}[\mathsf{Y}|\,\mathsf{X}] = \mathbf{E}[\mathsf{Z}|\,\mathsf{X}]$ by the description of the example or by (2). Thus as a result of the law of iterated expectations, we also have $\mathbf{E}[\mathsf{Y}] = \mathbf{E}[\mathsf{Z}]$. Thus one might be tempted to doubt the conclusion that "If $m > \frac{1}{1-p}$, then we will always switch to the envelope Z." Our conclusions, however, is sound; the subtlety here ia that if $m > \frac{1}{1-p}$ then both $\mathbf{E}[\mathsf{Y}]$ and $\mathbf{E}[\mathsf{Z}]$ are infinite.