Question 1. (5 points)

Agent receives rewards as follows:

$$r_k = \begin{cases} 0; \text{ if } y_k = 0 \\ 1; \text{ if } y_k = 1 \text{ and } k = 1 \\ 2r_{k-1} \text{with probability} \frac{1}{2}; \text{ if } y_k = 1 \text{ and } k > 1 \\ \frac{1}{2}r_{k-1} \text{with probability} \frac{1}{2}; \text{ if } y_k = 1 \text{ and } k > 1 \end{cases} \tag{1}$$

Determine $U^{y_{\leq 3}}$ of $y_1 = y_2 = y_3 = 1$.

Answer:

There are four possible reward sequences $r_{\leq 3}$ each with probability $\frac{1}{4}$:

$$\begin{gathered} 1, \tfrac{1}{2}, \tfrac{1}{4} \\ 1, \tfrac{1}{2}, 1 \\ 1, 2, 1 \\ 1, 2, 4 \end{gathered}$$

Their respective sums are $\frac{7}{4}, \frac{10}{4}, \frac{16}{4}, \frac{28}{4}$. The mean value is thus $\frac{1}{4}(\frac{7}{4} + \frac{10}{4} + \frac{16}{4} + \frac{28}{4}) = \frac{61}{16}$

Question 2. (10 points)

Agent receives rewards as follows:

$$r_k = \begin{cases} 0; \text{ if } y_k = 0 \\ 1; \text{ if } y_k = 1 \text{ and } k = 1 \\ r_{k-1} + 1 \text{ with probability} \frac{1}{2}; \text{ if } y_k = 1 \text{ and } k > 1 \\ r_{k-1} - 1 \text{ with probability} \frac{1}{2}; \text{ if } y_k = 1 \text{ and } k > 1 \end{cases} \tag{2}$$

Determine $U^{y_{\leq\infty}}$ of $y_{\leq\infty} = 1, 1, \ldots$ for $\gamma = \frac{1}{2}$.

Answer:

Consider the conditional expected value of reward r_k in a reward sequence $r_{\leq m}$ of length $m \geq k$

$$\mathbb{E}\left(r_k | y_{\leq\infty} = 1, 1, \ldots\right) = \sum_{r_{\leq m}} P(r_{\leq m} | y_{\leq\infty} = 1, 1, \ldots) r_k$$

Since the condition $y_{\leq\infty} = 1, 1, \ldots$ is fixed by assumption, we drop it from the probability and expectation expressions. Due to the second case in (2), we have $\mathbb{E}\left(r_1\right) = 1$, and from the last two cases we have $\mathbb{E}\left(r_k\right) = \mathbb{E}\left(r_{k-1}\right)$ for $k \geq 2$. So by induction

$$\mathbb{E}\left(r_k\right) = \sum_{r_{\leq m}} P(r_{\leq m}) r_k = 1 \tag{3}$$

for all $k \in \mathbb{N}$. The utility can now be computed as follows:

$$U^{y_{\leq\infty}} = \lim_{m\to\infty} \sum_{r_{\leq m}} \left(P(r_{\leq m}) \sum_{k=1}^{m} r_k \gamma^k \right) = \lim_{m\to\infty} \sum_{k=1}^{m} \gamma^k \underbrace{\sum_{r_{\leq m}} P(r_{\leq m}) r_k}_{=1 \text{ from } (3)} = \lim_{m\to\infty} \sum_{k=1}^{m} \left(\frac{1}{2}\right)^k = 1$$