**Question 1**

1. FALSE. Consider independent uniformly random variables $X, Y$ on $\{0, 1\}$. Let $Z = X \oplus Y$.
2. TRUE.
3. TRUE. One can have $(0, 100, 101, 110, 1110, 1111)$
4. A weak converse states that there is no sequence of $(n, 2^{\lfloor nR \rfloor})$ codes with $R < H(X)$ such that we have vanishing error probability as $n \to \infty$ after decoding. A strong converse states that for any such sequence of codes, the error necessarily equals 1 in the limit $n \to \infty$.

**Question 2**

1. The entropy is 3.1291 bits.
2. A Huffman code for this source would be

$$\begin{aligned}
C(A) &= 00 \\
C(E) &= 10100 \\
C(I) &= 010 \\
C(O) &= 100 \\
C(U) &= 110 \\
C(G) &= 1010100 \\
C(K) &= 1010101 \\
C(P) &= 10110 \\
C(R) &= 011 \\
C(S) &= 101011 \\
C(T) &= 10111 \\
C(V) &= 111
\end{aligned}$$

3. A Shannon code would be

| x | P(x) | $\lceil -\log p(x) \rceil$ | Codeword |
|---|---|---|---|
| A | 0.23 | 3 | 000 |
| O | 0.13 | 3 | 001 |
| R | 0.12 | 4 | 0100 |
| V | 0.12 | 4 | 0101 |
| U | 0.12 | 4 | 0110 |
| I | 0.11 | 4 | 0111 |
| P | 0.05 | 5 | 10000 |
| T | 0.05 | 5 | 10001 |
| E | 0.04 | 5 | 10010 |
| S | 0.02 | 6 | 100110 |
| G | 0.005 | 8 | 10011100 |
| K | 0.005 | 8 | 10011101 |

**Question 3**

1. Huffman codes are prefix codes and hence

$$\mathbb{E}[l_k(X^k)] = \bar{l}_k(X^k) \le H(X^k) + 1 = kH(X) + 1. \tag{1}$$

2. Consider the random variable $\frac{1}{k} l_k(X^k)$. Since we know that

$$kH(X) \le \bar{l}_k(X^k) \le kH(X) + 1, \tag{2}$$

we have that the mean $\mu$ of $\frac{1}{k} l_k(X^k)$ is bounded as

$$H(X) \le \mu \le H(X) + \frac{1}{k}. \tag{3}$$

Then, by the weak law of large numbers we have

$$\lim_{m\to\infty} \Pr\left[\left|\frac{1}{mk}\sum_{i=1}^{m} l_k(X_i^k) - \mu\right| \ge \frac{1}{k}\right] = 0 \tag{4}$$

$$\lim_{m\to\infty} \Pr\left[\frac{1}{mk}\sum_{i=1}^{m} l_k(X_i^k) \ge H(X) + \frac{2}{k}\right] = 0 \tag{5}$$

3. The argument follows the proof of Theorem 2.18 in the notes.

Consider a prefix code for blocks of size $k = \lceil \frac{2}{\delta} \rceil + 1$. Now consider $m$ such blocks for any $m \in \mathbb{N}$. Let $n = mk$. By the previous part, we have that

$$\lim_{m\to\infty} \Pr\left[\sum_{i=1}^{m} l_k(X_i^k) \ge mkH(X) + 2m\right] = 0. \tag{6}$$

That is, the length of the encoded message $L_n$ for $m$ blocks of size $k = \lceil \frac{2}{\delta} \rceil + 1$ satisfies

$$\lim_{n\to\infty} \Pr\left[L_n \le nH(X) + \frac{2n}{k}\right] = 1 \tag{7}$$

$$\implies \lim_{n\to\infty} \Pr\left[L_n \le nH(X) + \frac{2n\delta}{2+\delta}\right] = 1. \tag{8}$$

For a rate $R = H(X) + \delta$, we have

$$\begin{aligned} n\left(H(X) + \frac{2\delta}{2+\delta}\right) &= n\left(R - \delta + \frac{2\delta}{2+\delta}\right) && (9)\\ &= n\left(R - \frac{\delta^2}{2+\delta}\right) && (10)\\ &\le \lfloor nR \rfloor. && (11) \end{aligned}$$

Hence a rate of $H(X) + \delta$ is achievable for any $\delta > 0$.

A possible encoding/decoding process is Huffman encoding described in Algorithm 2.1 of the notes for $x^k \in \mathcal{X}^k$ with probabilities $p_{x^k} = P_{X^k}(x^k)$.