Problem Q1:

(a) Let the steady state probability of the Markov chain be \((q_A, 1 - q_A)\), then

\[
q_A = q_A(1 - \alpha) + (1 - q_A)\beta \quad \Rightarrow \quad q_A = \frac{\beta}{\alpha + \beta}.
\]

Therefore \(H(S_n) = H\left(\frac{\alpha}{\alpha + \beta}\right)\).

(b) Since the Markov chain is in steady state,

\[
H[X_n|S_{n-1}] = \sum_S H[X_n|s_{n-1}]q(s_{n-1}) = \frac{\beta}{\alpha + \beta}H(\alpha) + \frac{\alpha}{\alpha + \beta}H(\beta).
\]

(c) The LZ method asymptotically achieves \(\bar{L} \approx \frac{1}{2}H[XY|S]\):

\[
\bar{L} \approx \frac{1}{2}H[XY|S] = \frac{1}{2} \left( H[X|S] + H[Y|XS] \right) = \frac{1}{2} \left( H[X|S] + H[Y|X] \right).
\]

The last equality holds here because given \(X\), the distribution of \(Y\) does not depend on the state \(S\). Therefore

\[
\bar{L} \approx \frac{1}{2} \left\{ \frac{\beta}{\alpha + \beta}H(\alpha) + \frac{\alpha}{\alpha + \beta}H(\beta) + H(\gamma) \right\}
\]

(d) \(\bar{L} \approx \frac{1}{2}(H[X|S] + H[Y|S])\). Since conditioning reduces entropy on average, \(H[Y|S] \geq H[Y|XS]\). Therefore the expected number of bits per source symbol is larger than that from part (c).

Problem Q2:
(a) Applying the Lloyd-Max with the initial representation points \( a_1 = \frac{1}{2} \) and \( a_2 = \frac{3}{2} \), we see that only one iteration is required and the quantizing regions are \( R_1 = [0, 1] \), \( R_2 = (1, 2] \).

Alternatively, we can find the minimum by first noticing that the endpoint \( b_1 \) must lie within \([1, 2]\) to minimize the MSE, and \( a_2 \) must be center of the interval \((b_1, 2]\). Also, \( b_1 \) is equal distance from \( a_1 \) and \( a_2 \). As shown in the figure on the right, \( a_1 \) is the conditional mean of \( R_1 \), therefore

\[
\int_0^{\frac{1}{4}} u f_U(u) du = 2 - 3x \quad \Rightarrow \quad \int_0^{\frac{1}{4}} \frac{1}{4} u du + \int_{\frac{1}{4}}^{\frac{3}{4}} \frac{3}{4} u du = \frac{(2 - 3x)^2}{2}
\]

\[
\Rightarrow \quad x = \frac{1}{2}
\]

\[
MSE = \frac{1}{4} \times \frac{1}{12} + \frac{3}{4} \times \frac{1}{12} = \frac{1}{12}
\]

\[
H(V) = H_B(\frac{1}{4}) = -\frac{1}{4} \log \frac{1}{4} - \frac{3}{4} \log \frac{3}{4} = 2 - \frac{3}{4} \log 3
\]

(b) Since the pdf of \( u \) is uniform within \( R_1 \), the representation points are uniformly spaced: \( a_{11} = \frac{1}{4}, a_{12} = \frac{3}{4} \). The point separating the quantization regions is \( b_{11} = \frac{1}{2} \).

(c) Similarly to part (b), the representation points are uniformly spaced, with \( a_{21} = \frac{5}{4}, a_{22} = \frac{7}{4}, b_{12} = \frac{3}{2} \). \( V_2(U) \) leads to lower overall MSE. \( U \) lies within \( R_2 \) with higher probability, but the conditional mean squared error are the same, so further quantization here reduces the overall MSE by a larger amount. \( V_2(U) \) leads to higher entropy, since the resulting probability mass function is closer to the uniform distribution.

(d) Similarly to part (a), apply the Lloyd-Max algorithm with starting points \( a_1 = \frac{1}{4}, a_2 = \frac{3}{4}, a_3 = \frac{5}{4}, a_4 = \frac{7}{4} \). The resulting quantization regions are intervals all of length \( \frac{1}{2} \).

\[
MSE = 2 \times \frac{1}{8} \times \left( \frac{1}{2} \right)^2 \times \frac{1}{12} + 2 \times \frac{3}{8} \times \left( \frac{1}{2} \right)^2 \times \frac{1}{12} = \frac{1}{48}
\]

(e) \( V'(U) \) achieves lower MSE. Denote the quantization output from \( V'(U) \) by \( \{0, 10, 11\} \) and quantization output from \( V''(U) \) by \( \{00, 01, 10, 11\} \). When the first bit is dropped, the receiver sees either \( \{?, 0, ?, 1\} \) or \( \{0, ?, 0, ?, 1\} \). For \( V'(U) \), MSE at the receiver is the same as that at the transmitter, since the three codewords can still be distinguished. This MSE is lower than that of the 1 bit quantizer from part (a). For \( V''(U) \), MSE at the receiver is at least that of the 1 bit quantizer, depending on how the representation points are labeled.

**Problem Q3:**

(a) \( \hat{u}_n(f) = 2 \text{sinc}(2f) - 2^{1-n} \text{sinc}(2^{1-n}f) \), \( \hat{u}(f) = \lim_{n \to \infty} \hat{u}_n(f) = 2 \text{sinc}(2f) \).
(b) No. From the plots of $u_n(t)$, we see that in the limit, a discontinuity exists at $t = 0$.

(c) The sampling theorem does not apply to $u(t)$ because of its discontinuity. The sampling theorem does apply to $\hat{u}(f)$ because it is continuous, $L_2$, and time-limited to $[-1, 1]$.

(d) To apply Gram-Schmidt orthonormalization, first note for all $n > 1$, $u_n(t)$ is $u_{n-1}(t)$ combined with rectangles of width $2^{-n}$ and height 1, all of which are non-overlapping:

$$u_n(t) = u_{n-1}(t) + \text{rect}\left(\frac{t - \frac{3}{2^n+1}}{2^{-n}}\right) + \text{rect}\left(\frac{t + \frac{3}{2^n+1}}{2^{-n}}\right).$$

Therefore $\phi_1(t) = u_1(t)$, $\phi_2(t) = \sqrt{2}\text{rect}\left(\frac{t + \frac{2}{4}}{4}\right) + \sqrt{2}\text{rect}\left(\frac{t - \frac{2}{4}}{4}\right)$, \ldots,

$$\phi_n(t) = \sqrt{2^{n-1}} \left\{ \text{rect}\left(\frac{t - \frac{3}{2^n+1}}{2^{-n}}\right) + \text{rect}\left(\frac{t + \frac{3}{2^n+1}}{2^{-n}}\right) \right\}$$

(e) No. $\phi_n(t)$ is piecewise constant and equal to 0 at $t = 0$, so continuous $L_2$ functions on $[-1, 1]$ can not equal to linear combinations of $\{\phi_n, 1 \leq n < \infty\}$. Also, $\phi_n(t)$ are symmetric about $t = 0$, hence non-symmetric functions cannot be expressed as their linear combinations.