1 The Weak Law of Large Numbers

(a) $M_{\bar{X}_n}(s) = \mathbb{E}[e^{s\bar{X}_n}] = \mathbb{E}\left[\prod_{i=1}^n e^{\frac{sX_i}{n}}\right] = \mathbb{E}\left[\left(e^{\frac{sX_1}{n}}\right)^n\right] = \mathbb{E}\left[\left(e^{\frac{sX_1}{n}}\right)\right]^n = M_X(\frac{s}{n})^n$

(b)

$$\begin{aligned} M_X(s) = a + bs + o(s) &= \mathbb{E}[e^{sX}] \\ &= 1 + s\mathbb{E}[X] + o(s) \end{aligned}$$

Therefore, we have that $a = 1$ and $b = \mathbb{E}[X] = \mu$.

(c) $M_{\bar{X}_n}(s) = M_X(\frac{s}{n})^n = \left(a + b\left(\frac{s}{n}\right) + o\left(\frac{s}{n}\right)\right)^n = \left(1 + \mu\left(\frac{s}{n}\right) + o\left(\frac{s}{n}\right)\right)^n$.
As such, we have that $\lim_{n\to\infty} M_{\bar{X}_n}(s) = \lim_{n\to\infty}\left(1 + \frac{\mu s}{n} + o\left(\frac{s}{n}\right)\right)^n = \lim_{n\to\infty}\left(1 + \frac{\mu s}{n}\right)^n = e^{\mu s}$.

(d) $\lim_{n\to\infty} M_{\bar{X}_n}(s) = e^{\mu s} = \mathbb{E}[e^{\mu s}] = M_\mu(s)$, so $\bar{X}_n \xrightarrow{\text{d}} \mu$.

2 Huffman Questions

(a) Determine the probabilities of all possible outcomes of $X_1, X_2, \ldots, X_n$. On each iteration of the algorithm, perform the same procedure as is done in Huffman coding. In particular, each of the two lowest probability outcomes are combined into a larger subtree and readded to the queue. On each leaf of the tree the corresponding outcome sequence (e.g. $0\cdots 0$) will be stored. After the tree has been created, we can then proceed to ask questions such as whether the left subtree contains the specific sequence we have or not, and thus we can find the exact sequence and defective items.

(b) Similar to above we can ask something along the lines of which branch the desired sequence we have is in the final branch of the subtree. This will amount to asking essentially wehther the last item is defective, as we will be distinguishing $0\cdots 0$ and $0\cdots 1$ (since these are the two lowest probability outcomes).

3 Channel Capacity of the Binary Symmetric Channel Random Code

(a) From HW 4 3d, we have that $\mathbb{P}(|X - qn| \geq \epsilon n) \leq 2e^{-2n\epsilon^2}$ for $X \sim \text{Binomial}(n, q)$. Since $\sum_{i=1}^n y(i) \oplus x(i) \sim \text{Binomial}(n, p)$, we have that $\mathbb{P}(Y_w \notin \text{DecodeBox}(X_w))$
$= \mathbb{P}\left(\left|\sum_{i=1}^n y(i) \oplus x(i) - pn\right| \geq n\epsilon\right) \leq 2e^{-2n\epsilon^2}$.

(b) We can biject sequences $\in$ DecodeBox(0) to sequences $\in$ DecodeBox(x) such that starting from a base sequence X_w for DecodeBox(x) matches with the base sequence $(0, \cdots, 0)$ from DecodeBox(0). In particular, these are matched as any sequence $A \in$ DecodeBox(0) has a corresponding sequence $B \in$ DecodeBox(x) such that the indices of the values that are flipped with respect to $(0, \cdots, 0)$ for A are the same as that for B with respect to X_w. As such, $|\text{DecodeBox}(0)| = |\text{DecodeBox}(x)|\ \forall x \in \{0,1\}^n$.

(c) $\mathbb{P}(Y_w(i) = 1) = \frac{1}{2}(1-p) + \frac{1}{2}p = \frac{1}{2}$, so $Y_w(i) \sim \text{Bernoulli}\left(\frac{1}{2}\right)\ \forall i$.

(d) $\mathbb{P}(Y_w \in \text{DecodeBox}(X_u)) = \frac{|\text{DecodeBox}(0)|}{2^n}$ from part (b) and part (c).

(e) By the union bound, we have that $\mathbb{P}(\exists u \neq w : Y_w \in \text{DecodeBox}(X_u)) \leq 2^k \frac{|\text{DecodeBox}(0)|}{2^n}$.

(f) We have $-n\epsilon + pn \leq \sum_{i=1}^n y(i) \leq n\epsilon + pn$. In addition, since $|-\frac{1}{n}\log_2 p(X_1, \ldots X_n) - H(X_1)| \leq \epsilon'$, we have that

$$
\begin{aligned}
|-(\frac{1}{n}\sum_{i=1}^n y(i)\log_2 p + (n - \sum_{i=1}^n y(i))\log_2(1-p)) - H(p)| &\leq \\
|-(\frac{1}{n}\sum_{i=1}^n y(i)\log_2 p + (n - \sum_{i=1}^n y(i))\log_2(1-p)) - H(p)| & \\
\leq |-(\frac{1}{n}(n\epsilon + pn)\log_2 p + (n - n\epsilon + pn)\log_2(1-p)) - H(p)| & \\
\leq |-((\epsilon + p)\log_2 p + (1 - \epsilon + p)\log_2(1-p)) - H(p)| & \\
\leq \epsilon'. &
\end{aligned}
$$

So that $|\text{DecodeBox}(0) \leq 2^{n(H(p)+\epsilon')}$ where $\epsilon' = -((\epsilon+p)\log_2 p + (1-\epsilon+p)\log_2(1-p)) - H(p)$.

(g) $\mathbb{P}(\exists u \neq w : Y_w \in \text{DecodeBox}(X_u)) \leq 2^k \frac{|\text{DecodeBox}(0)|}{2^n} \leq 2^k 2^{n(H(p)+\epsilon'-1)}$, so that we would need $k < -n(H(p)-1) = n(1-H(p))$, leading to $C = 1 - H(p)$.

4 Number of Parameters

(a) 2^{n+1} parameters are necessary to characterize all the possible combinations of values. However, due to the normalization property, the last parameter will be completely determined by the first $2^{n+1}-1$, such that only $2^{n+1}-1$ parameters are necessary to parameterize the joint.

(b) Since random variables are independent and binary, only $n+1$ parameters are necessary to parameterize the joint.

(c) 3 parameters are necessary. In particular, by the Markov property, $\mathbb{P}(X_0 = x_0, X_1 = x_1, \ldots, X_n = x_n) = \pi_0 P_{0,1} \cdots P_{n-1,n}$. Since there are only two states and the rows in P must sum to 1, only 3 parameters are needed in total. One comes from parameterizing the initial distribution π_0, and the others come from the probability of transitioning from one state to the other.

(d) This violates the time homogeneity property of Markov chains. In particular, let $\mathbb{P}(Z_n) = \frac{1}{n+1}$. Then $\mathbb{P}(Z_n = a \mid Z_{n-1} = b) = \frac{1}{n+1}$ and $\mathbb{P}(Z_m = a \mid Z_{m-1} = b) = \frac{1}{m+1}$, which are not equal for $n \neq m$.

5 Backwards Markov Property

We have that

$$
\begin{aligned}
\mathbb{P}(X_k = i_0 \mid X_{k+1} = i_1, \ldots, X_{k+m} = i_m) &= \frac{\mathbb{P}(X_k = i_0, X_{k+1} = i_1, \ldots, X_{k+m} = i_m)}{\mathbb{P}(X_{k+1} = i_1, \ldots, X_{k+m} = i_m)} \\
&= \frac{\mathbb{P}(X_k = i_0, X_{k+1} = i_1, \ldots, X_{k+m} = i_m)}{\mathbb{P}(X_{k+1} = i_1, \ldots, X_{k+m} = i_m)}
\end{aligned}
$$

$$= \frac{\mathbb{P}(X_{k+m} = i_{k+m} \mid X_k = i_0, X_{k+1} = i_1, \dots, X_{k+m-1} = i_{m-1}) \cdot \mathbb{P}(X_k = i_0, X_{k+1} = i_1, \dots, X_{k+m-1} = i_{m-1})}{\mathbb{P}(X_{k+m} = i_{k+m} \mid X_{k+1} = i_1, \dots, X_{k+m-1} = i_{m-1}) \cdot \mathbb{P}(X_{k+1} = i_1, \dots, X_{k+m-1} = i_{m-1})}$$
$$= \frac{\mathbb{P}(X_k = i_0, X_{k+1} = i_1, \dots, X_{k+m-1} = i_{m-1})}{\mathbb{P}(X_{k+1} = i_1, \dots, X_{k+m-1} = i_{m-1})}.$$

Continuing in this manner, we get that $\mathbb{P}(X_k = i_0 \mid X_{k+1} = i_1, \dots, X_{k+m} = i_m) = \frac{\mathbb{P}(X_k = i_0, X_{k+1} = i_{k+1})}{\mathbb{P}(X_{k+1} = i_{k+1})} = \mathbb{P}(X_k = i_0 \mid X_{k+1} = i_1)$.

6 [Bonus] The CLT Implies the WLLN

(a)

(b)