Homework 2

EE 668 — Information Theory

Name: Ali Zafari **Student Number:** 800350381 **Fall 2022**

Problem 2.1

((a)) According to **central limit theorem**, M_n will follow a normal distribution with mean and variance of the data distribution divided by the sample sequence size, i.e., $Exp(1)$:

$$M_n \sim \mathcal{N}(E[X], \frac{Var[X]}{n})$$
$$M_n \sim \mathcal{N}(1, \frac{1}{n})$$

We can define A_n being defined as $A_n \triangleq \frac{M_n - 1}{1/\sqrt{n}}$:

$$A_n \sim \mathcal{N}(0, 1)$$

Hence, the probability can be written in terms of Q-function of a Standard Normal(Φ):

$$P[M_n > 2] = P[A_n > \sqrt{n}]$$
$$= \Phi(\sqrt{n})$$

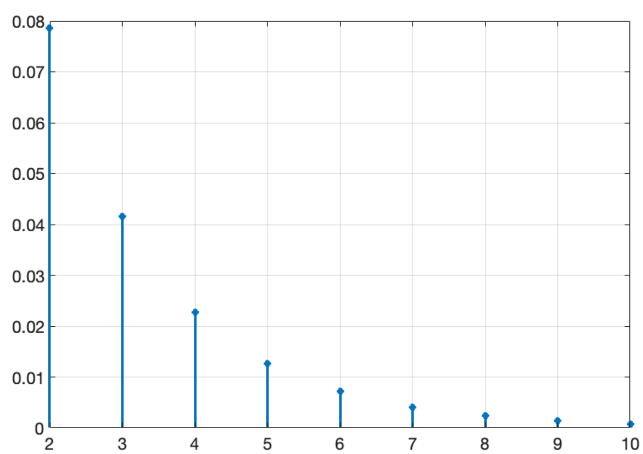

Figure 1: $P(M_n > 2)$

((b)) To develop the Chernoff bound on the probability distribution of the sample mean, we recall (from EE 513):

$$P[M_n > \gamma] \leq e^{-\max\limits_{s\geq 0}\{s\gamma - \phi_{M_n}(s)\}}$$

where $\phi_{M_n}(s)$ is the logarithm of the moment generating function (log-MGF).

As $M_n \sim \mathcal{N}(1, \frac{1}{n})$, its MGF function would be:

$$\begin{aligned}
\text{(MGF)} \qquad \Phi_{M_n}(s) &= e^{\mu s + \frac{\sigma^2_{M_n}}{2}s^2} \\
&= e^{s + \frac{s^2}{2n}} \\
\text{(log-MGF)} \qquad \Rightarrow \phi_{M_n}(s) &= s + \frac{s^2}{2n}
\end{aligned}$$

To find the value of s which maximized the exponent, we set its derivative equal to zero:

$$\frac{d}{ds}(s(\gamma - 1) - \frac{s^2}{2n})\Big|_{s^*} = \gamma - 1 - \frac{s}{n}\Big|_{s^*} = 0 \Rightarrow s^* = n$$

Finally the Chernoff bound will be as follows, by having $\gamma = 2$:

$$\begin{aligned}
P[M_n > 2] &\leq e^{-\{n\times 2 - n - \frac{n}{2}\}} \\
P[M_n > 2] &\leq e^{-\frac{n}{2}}
\end{aligned}$$

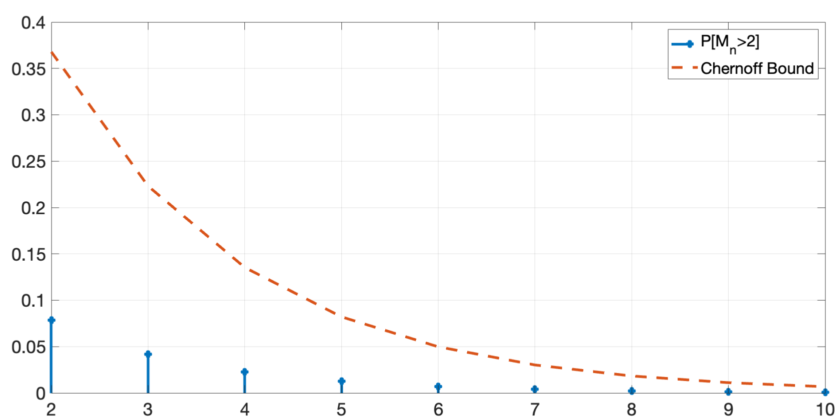

Figure 2: Chernoff Bound vs. $P(M_n > 2)$

Problem 2.2

((a)) Let's define A_n as the product of the drawn samples:

$$A_n = (X_1 X_2 \dots X_n)^{\frac{1}{n}}$$

then

$$\log A_n = \frac{1}{n} \log(X_1 X_2 \dots X_n) = \frac{1}{n} \sum_{i=1}^{n} \log X_i$$

Now we can infer that $\log A_n$ is the sample mean of the random variable Y being defined as $Y \triangleq \log X$.

By using the strong law of large numbers:

$$\overline{Y}_n \xrightarrow{Prob.=1} \mu_Y \qquad \text{when } n \to +\infty$$

so we only need to calculate μ_Y:

$$\begin{aligned}
\mu_Y &= E_{y \sim f_Y(y)}[Y] \\
&= E_{x \sim f_X(x)}[\log X] \\
&= \frac{1}{2} \log 1 + \frac{1}{4} \log 2 + \frac{1}{4} \log 3 \\
&= \frac{1}{4} \log 6
\end{aligned}$$

Therefore:

$$\overline{Y}_n = \log A_n \xrightarrow{Prob.=1} \log 6^{\frac{1}{4}} \qquad \text{when } n \to +\infty$$

$$A_n = (X_1 X_2 \dots X_n)^{\frac{1}{n}} \xrightarrow{Prob.=1} 1.5650 \qquad \text{when } n \to +\infty$$

((b)) Entropy of X:

$$\begin{aligned}
H(X) &= -\sum p \log p \\
&= -\left(\frac{1}{2} \log \frac{1}{2} + \frac{1}{4} \log \frac{1}{4} + \frac{1}{4} \log \frac{1}{4}\right) \\
&= 1.5
\end{aligned}$$

Problem 2.3

((a))

$$\lim_{n\to+\infty}\left[-\frac{1}{n}\log q(X_1, X_2, \ldots, X_n)\right] = \lim_{n\to+\infty}\left[-\frac{1}{n}\sum_{i=1}^{n}\log q(X_i)\right]$$

Now we can infer that the above limit is on the sample mean of the random variable Y being defined as $Y \triangleq -\log q(X_i)$.

By using the strong law of large numbers:

$$\overline{Y}_n \xrightarrow{Prob.=1} \mu_Y \qquad \text{when } n \to +\infty$$

so we only need to calculate μ_Y:

$$\begin{aligned}
\mu_Y &= E_{y\sim f_Y(y)}[y] \\
&= E_{x\sim f_X(x)}[-\log q(x_i)] \\
&= -\sum_{i=1}^{n} p(x_i)\log q(x_i) \\
&= -\sum_{i=1}^{n} p(x_i)\log[q(x_i)\times \frac{p(x_i)}{p(x_i)}] \\
&= \sum_{i=1}^{n} p(x_i)\log\frac{p(x_i)}{q(x_i)} - \sum_{i=1}^{n} p(x_i)\log p(x_i) \\
&= D(p||q) + H(p)
\end{aligned}$$

((b))

$$\lim_{n\to+\infty}\left[\frac{1}{n}\log\frac{q(X_1, X_2, \ldots, X_n)}{p(X_1, X_2, \ldots, X_n)}\right] = \lim_{n\to+\infty}\left[\frac{1}{n}\sum_{i=1}^{n}\log\frac{q(X_i)}{p(X_i)}\right]$$

Now we can infer that the above limit is on the sample mean of the random variable Y being defined as $Y \triangleq \log\frac{q(X)}{p(X)}$.

By using the strong law of large numbers:

$$\overline{Y}_n \xrightarrow{Prob.=1} \mu_Y \qquad \text{when } n \to +\infty$$

so we only need to calculate μ_Y:

$$\begin{aligned}\mu_Y &= E_{y \sim f_Y(y)}[y] \\ &= E_{x \sim f_X(x)}[\log \frac{q(x)}{p(x)}] \\ &= \sum_{i=1}^{n} p(x_i) \log \frac{q(x_i)}{p(x_i)} \\ &= -D(p||q)\end{aligned}$$

Problem 2.4

$$\log l = \frac{1}{n} \ln V_n = \frac{1}{n} \sum_{i=1}^{n} \ln X_i$$

Now we can infer that $\ln l$ is the sample mean of the random variable Y being defined as $Y \triangleq \ln X$.

By using the strong law of large numbers:

$$\overline{Y}_n \xrightarrow{Prob.=1} \mu_Y \qquad \text{when } n \to +\infty$$

so we only need to calculate μ_Y:

$$\begin{aligned}\mu_Y &= E_{y \sim f_Y(y)}[Y] \\ &= E_{x \sim f_X(x)}[\ln X] \\ &= \int_0^1 \ln x dx \\ &= x(\ln x - 1)\Big|_0^1 \\ &= -1\end{aligned}$$

Hence:

$$\overline{Y}_n = \ln l \xrightarrow{Prob.=1} -1 \qquad \text{when } n \to +\infty$$

$$l \xrightarrow{Prob.=1} e^{-1} = 0.3679 \qquad \text{when } n \to +\infty$$
