1 Midterm

1. (a) $\mathbb{E}[|X|] = \frac{2}{\sqrt{2\pi}} \int_0^\infty x e^{-\frac{x^2}{2}} \mathrm{d}x = \frac{2}{\sqrt{2\pi}} \left(-e^{-\frac{x^2}{2}} \right) \Big|_0^\infty = -\frac{2}{\sqrt{2\pi}}(0-1) = \frac{2}{\sqrt{2\pi}}$

   (b) Let $Y = \max(Y_1, Y_2, Y_3, \ldots, Y_n)$, where $Y_i \mid 0 < i < n$ is the valuation of the ith bidder. We have that $F_Y(y) = \mathbb{P}(Y_i \le y)^n = y^n$, and $f_Y(y) = ny^{n-1}$ by differentiating. We then would like to evaluate the expected value of the profit when we win, $(v - Y)$. This works out to

$$\begin{aligned}
\int_{v=0}^1 \int_{y=0}^v (v-y)(ny^{n-1}) \mathrm{d}y \mathrm{d}v &= n \int_{v=0}^1 \left(\frac{vy^n}{n} - \frac{y^{n+1}}{n+1} \right) \Bigg|_{y=0}^v \mathrm{d}v \\
&= n \int_{v=0}^1 \left(\frac{v^{n+1}}{n} - \frac{v^{n+1}}{n+1} \right) \mathrm{d}v \\
&= n \left(\frac{v^{n+2}}{n(n+2)} - \frac{v^{n+2}}{(n+1)(n+2)} \right) \Bigg|_{v=0}^1 \\
&= n \left(\frac{v^{n+2}(n+1) - v^{n+2}(n)}{n(n+1)(n+2)} \right) \\
&= \frac{1}{(n+1)(n+2)}.
\end{aligned}$$

   (c) $\mathbb{P}(\textbf{failure}) \ge \mathbb{P}(\text{no degree 1 packets sent}) = \left(1 - \frac{1}{n}\right)^n$. $\lim_{n\to\infty} \left(1 - \frac{1}{n}\right)^n = e^{-1}$, so $\mathbb{P}(\textbf{failure}) \ge e^{-1}$ as $n \to \infty$.

   (d) Let $T = \sum_{i=1}^n T_i$, where T_i is an indicator variable indicating whether the ith throw is better than his previous $i-1$ throws. Then we have that

$$\begin{aligned}
\mathbb{E}[T] &= \sum_{i=1}^n \mathbb{E}[T_i] \\
&= \sum_{i=1}^n \mathbb{P}(T_i = 1) \\
&= \sum_{i=1}^n \frac{1}{i} \\
&\approx \ln n.
\end{aligned}$$

   (e) Let $R \sim \text{Poisson}(\lambda_r)$ denote the number of red balls, $B \sim \text{Poisson}(\lambda_b)$ denote the number of blue balls, and $N = (B+R) \sim \text{Poisson}(\lambda_r + \lambda_b)$ denote the number of balls in total. We have that

$$\begin{aligned}
p_{B|N}(b \mid n) &= \frac{p_{B,N}(b,n)}{p_N(n)} \\
&= \frac{p_R(n-b) \cdot p_B(b)}{p_N(n)} \\
&= \frac{\frac{\lambda_r^{(n-b)} e^{-\lambda_r} \cdot \lambda_b^b e^{-\lambda_b}}{(n-b)!b!}}{\frac{(\lambda_b+\lambda_r)^n e^{-(\lambda_b+\lambda_r)}}{n!}} \\
&= \binom{n}{b} \frac{\lambda_b^b \lambda_r^{(n-b)}}{(\lambda_b+\lambda_r)^n}
\end{aligned}$$

$$\sim \text{Binomial}\left(n, \frac{\lambda_b}{\lambda_b + \lambda_r}\right).$$

2. (a) Let $S_n = X + (n - X)(-1)$, where $X \sim \text{Binomial}\left(n, \frac{1}{2}\right)$, representing the number of $+1$ Y_is that are obtained. Then we have that

$$\begin{aligned}\mathbb{P}(|S_n| \geq t) &= \mathbb{P}(|X + (n - X)(-1)| \geq t) \\ &= \mathbb{P}(|2X - n| \geq t) \\ &= \mathbb{P}\left(\left|X - \frac{n}{2}\right| \geq \frac{t}{2}\right) \\ &\leq 2e^{-2\frac{\left(\frac{t}{2}\right)^2}{n}} = 2e^{-\frac{t^2}{2n}}.\end{aligned}$$

(b) Let $X \sim \text{Bernoulli}\left(\frac{1}{k}\right)$, so that $\mathbb{P}(X = 1 = \frac{k}{k} = k\mathbb{E}[X]) = \frac{1}{k}$, so $\mathbb{P}(X \geq k\mathbb{E}[X]) = \frac{1}{k}$.

3. (a) $H(U) = -\sum_u p_u \log_2 p_u(u) = -\sum_{u=1}^{n} \frac{1}{n} \log_2 \frac{1}{n} = -\log_2 \frac{1}{n} = \log_2 n$

(b)

$$\begin{aligned}H(X,Y) &= -\sum_x \sum_y p_{X,Y}(x,y) \log_2 p_{X,Y}(x,y) \\ &= -\sum_x \sum_y p_X(x) p_Y(y) \log_2 p_X(x) p_Y(y) \\ &= -\sum_x \sum_y p_X(x) p_Y(y) \left(\log_2 p_X(x) + \log_2 p_Y(y)\right) \\ &= -\sum_x \sum_y p_X(x) p_Y(y) \log_2 p_X(x) - \sum_x \sum_y p_X(x) p_Y(y) \log_2 p_Y(y) \\ &= -\sum_x p_X(x) \log_2 p_X(x) \sum_y p_Y(y) - \sum_x p_X(x) \sum_y p_Y(y) \log_2 p_Y(y) \\ &= H(X) + H(Y)\end{aligned}$$

4. (a) Let $A = -\ln Y$. Then we have that $F_A(a) = \mathbb{P}(Y \geq e^{-a}) = \int_{e^{-a}}^{1} \mathrm{d}y = 1 - e^{-a}$, which is the CDF of an exponential random variable with $\lambda = 1$. As such, we have that $A = -\ln Y \sim \text{Exponential}(1)$.

(b) We have $Z = \ln \frac{X}{Y} = \ln X - \ln Y$, so by the convolution theorem, we have that $M_Z(s) = M_B(s) \cdot M_A(s)$, where $B = \ln X$ and $A = -\ln Y$. We get that

$$\begin{aligned}M_Z(s) &= M_B(s) \cdot M_A(s) \\ &= \mathbb{E}[e^{s \ln X}] \cdot \frac{1}{1-s} \\ &= \int_0^1 e^{s \ln X} \mathrm{d}x \left(\frac{1}{1-s}\right) \\ &= \int_0^1 x^s \mathrm{d}x \left(\frac{1}{1-s}\right) \\ &= \left.\frac{x^{s+1}}{s+1}\right|_0^1 \left(\frac{1}{1-s}\right)\end{aligned}$$

$$= \frac{1}{s+1} \cdot \frac{1}{1-s}$$
$$= \frac{1}{1-s^2}.$$

(c) $\mathbb{E}[X] = \frac{d}{ds}\left(\frac{1}{1-s^2}\right)\Big|_{s=0} = -(1-s^2)^{-2}(-2s)|_{s=0} = \frac{2s}{(1-s^2)^2}\Big|_{s=0} = 0$
$\mathbb{E}[X^2] = \frac{d}{ds}\left(\frac{2s}{(1-s^2)^2}\right)\Big|_{s=0} = \frac{(1-s)^2(2)-2s(2(1-s^2)(-2s))}{(1-s^2)^4}\Big|_{s=0} = 2$
$\text{var}(Z) = \mathbb{E}[X^2] - \mathbb{E}[X] = 2 - 0 = 2$

(d) Let $C = -\ln X$, so that $Z = A - C$, where $A = -\ln Y$ from part (a), such that $C \sim$ Exponential(1) and $A \sim$ Exponential(1). For the case where $Z > 0$ and $A > C$, we have that

$$f_Z(z) = \int_0^\infty f_A(x+z) f_C(x)\mathrm{d}x$$
$$= \int_0^\infty e^{-x-z}e^{-x}\mathrm{d}x$$
$$= \int_0^\infty e^{-2x-z}\mathrm{d}x$$
$$= \frac{e^{-2x-z}}{-2}\Big|_0^\infty$$
$$= -\frac{1}{2}\left(0 - e^{-z}\right)$$
$$= \frac{e^{-z}}{2}.$$

For the case where $Z < 0$, and $A < C$, we have that

$$f_Z(z) = \int_{-z}^\infty f_A(x+z) f_C(x)\mathrm{d}x$$
$$= \int_{-z}^\infty e^{-x-z}e^{-x}\mathrm{d}x$$
$$= \int_{-z}^\infty e^{-2x-z}\mathrm{d}x$$
$$= -\frac{e^{-2x-z}}{2}\Big|_{-z}^\infty$$
$$= -\frac{1}{2}\left(0 - e^{2z-z}\right)$$
$$= \frac{e^{z}}{2}.$$

5. (a) $(\pi + 4) f_{X,Y} = 1 \implies f_{X,Y} = \frac{1}{\pi+4}$

(b) For $-2 \le y \le 0$, we have that $f_Y = \int_{-(y+2)}^{y+2} \frac{1}{\pi+4}\mathrm{d}x = \frac{x}{\pi+4}\Big|_{-(y+2)}^{y+2} = \frac{2y+4}{\pi+4}$.

For $0 \le y \le 1$, we have that $f_Y = \int_{-\sqrt{1-y^2}+1}^{\sqrt{1-y^2}+1} \frac{2}{\pi+4} = \frac{4\sqrt{1-y^2}}{\pi+4}$.

(c) $\mathbb{E}[X \mid Y] = 0$, by symmetry along the y-axis. $\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X \mid Y]] = 0$.
$\mathbb{E}[XY] = \int_y \mathbb{E}[XY \mid Y] f_Y(y)\mathrm{d}y = 0$.
Thus, we have that $\text{cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = 0$.

6. (a) Let $X \sim$ Exponential(a) and $Y \sim$ Exponential(b). We can determine $\mathbb{P}(X < Y)$ as follows

$$\begin{aligned}
\mathbb{P}(X < Y) &= \int_{y=0}^{\infty} \int_{x=0}^{y} ae^{-ax} be^{-by} \mathrm{d}x\mathrm{d}y \\
&= \int_{y=0}^{\infty} be^{-by} \left(1 - e^{-ay}\right) \mathrm{d}y \\
&= 1 - \int_{y=0}^{\infty} be^{-(a+b)y} \mathrm{d}y \\
&= 1 - \frac{b}{a+b} \\
&= \frac{a}{a+b}
\end{aligned}$$

(b) True. Let A_1 denote the event that the Alice finishes with her customer first then Bob finishes with his customer before Alice finishes again. Let B_1 denote the same event, but with Bob finishing before Alice first. We are looking for $\mathbb{P}(A_1 \cup B_1) = \mathbb{P}(A_1) + \mathbb{P}(B_1) = \frac{2ab}{(a+b)^2} = \frac{2ab}{a^2+2ab+b^2}$. To get $\mathbb{P}(A_1 \cup B_1) = \frac{2ab}{a^2+2ab+b^2} < \frac{1}{2}$, we need $a^2 + b^2 - 2ab > 0$, so that $(a-b)^2 > 0$ or $(b-a)^2 > 0$. As a result, we find that $a > b$ or $b > a$, naturally implying that $a \neq b$ for $\mathbb{P}(A_1 \cup B_1) < \frac{1}{2}$.

(c) The wait time is $Z = \min(X, Y)$. We find the CCDF of Z as $\mathbb{P}(Z > z) = \mathbb{P}(X > z) \cdot \mathbb{P}(Y > z) = e^{-az}e^{-bz} = e^{-(a+b)z}$, where X and Y are from part (a). Thus, $Z \sim$ Exponential$(a+b)$.

2 Confidence Interval Comparisons

(a) From Chebyshev's inequality, we have $\mathbb{P}(|\hat{p} - p| \geq \epsilon) \leq \frac{\text{var}(\hat{p})}{\epsilon^2}$, where $\text{var}(\hat{p}) = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}$, so that $\mathbb{P}(|\hat{p} - p| \geq \epsilon) \leq \frac{p(1-p)}{n\epsilon^2}$.

(i) Solving for n where $\delta = \frac{p(1-p)}{n\epsilon^2}$, we have that, $n = \frac{p(1-p)}{\epsilon^2\delta}$. Plugging in the values $\epsilon = .05, \delta = .1$, we have $n = \frac{p(1-p)}{(.05)^2(.1)} = 4000p(1-p)$, such that for the worst case of $p = \frac{1}{2}$, we have $n = 1000$. Using $\epsilon = .1, \delta = .1$, we get $n = \frac{p(1-p)}{(.1)^3} = 1000p(1-p)$, and a worst case $n = 250$. Doubling ϵ divides n by 4.

(ii) Reusing the equation for n from part (i), we can plug in $\epsilon = .1, \delta = .05$ to get $n = \frac{p(1-p)}{(.1)^2(.05)} = 2000p(1-p)$, getting $n = 500$ in the worst case. The value of n for $\epsilon = .1, \delta = .1$ is the same as in part (i). Doubling δ halves n.

(b) Since $\mathbb{E}[\hat{p}] = \frac{np}{n} = p$, the mean of $\frac{\hat{p}-p}{p}$ is already standardized to 0. The variance must still be standardized to use the CLT. From part (a), $\text{var}(\hat{p}) = \frac{p(1-p)}{n}$, so standardize by $Z_n = \frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}} = \frac{\sqrt{n}(\hat{p}-p)}{\sqrt{p(1-p)}}$. Therefore, by the CLT, we have $\mathbb{P}\left(\left|\frac{\hat{p}-p}{p}\right| \leq .05\right) = \mathbb{P}\left(\left|\frac{\sqrt{n}(\hat{p}-p)}{\sqrt{p(1-p)}}\right| \leq \frac{\sqrt{np}.05}{\sqrt{1-p}}\right)$, and $\lim_{n\to\infty} \mathbb{P}(|Z_n| \leq \frac{\sqrt{np}.05}{\sqrt{1-p}}) = \phi(\frac{\sqrt{np}.05}{\sqrt{1-p}}) - \phi(-\frac{\sqrt{np}.05}{\sqrt{1-p}})$. Setting this equal to .95, we have that $\frac{\sqrt{np}.05}{\sqrt{1-p}} = 1.96$, so that $n = \frac{(1.96)^2(1-p)}{(.05)^2 p}$. In the worst case, we have $p = .4$, which gives $n = \frac{1536.64 \cdot .6}{.4} \approx 2305$.

3 Convergence in Probability

(a) $\mathbb{P}(|Y_n| \geq \epsilon) = \mathbb{P}(|(X_n)^n| \geq \epsilon) = \mathbb{P}(|X_n| \geq \epsilon^{\frac{1}{n}}) = 2\mathbb{P}(X_n \geq \epsilon^{\frac{1}{n}}) = 1 - \epsilon^{\frac{1}{n}}$.
$\lim_{n\to\infty} \mathbb{P}(|Y_n| \geq \epsilon) = \lim_{n\to\infty} 1 - e^{\frac{1}{n}} = 0$, so $Y_n \xrightarrow{p} 0$.

(b) We have $\text{var}(Y_n) = \text{var}(X_1)^n = \left(\frac{2^2}{12}\right)^n = \left(\frac{1}{3}\right)^n$. By Chebyshev's inequality, $\mathbb{P}(|Y_n - 0| \geq \epsilon) \leq \frac{\left(\frac{1}{3}\right)^n}{\epsilon^2}$, so that $\lim_{n\to\infty} \mathbb{P}(|Y_n| \geq \epsilon) = \lim_{n\to\infty} \frac{1}{\epsilon^2 3^n} = 0$, therefore $Y_n \xrightarrow{p} 0$.

(c) $\mathbb{P}(|Y_n - 1| \geq \epsilon) = \mathbb{P}(Y_n \geq \epsilon + 1) + \mathbb{P}(Y_n \leq -\epsilon + 1) = \mathbb{P}(Y_n \leq -\epsilon + 1)$, where $\mathbb{P}(Y_n \geq \epsilon + 1) = 0$, since Y_n and X_n are upper bounded by 1. Working this out, we get $\mathbb{P}(Y_n \leq -\epsilon + 1) = \left(\frac{-\epsilon+2}{2}\right)^n$, and since $\epsilon \in (0, 2)$ for nontrivial bounds, we have $\lim_{n\to\infty} \mathbb{P}(|Y_n - 1| \geq \epsilon) = \lim_{n\to\infty} \left(\frac{-\epsilon+2}{2}\right)^n = 0$, so $Y_n \xrightarrow{p} 1$.

(d) By the WLLN, we have $\lim_{n\to\infty} \mathbb{P}(\left|Y_n - \frac{1}{3}\right| \geq \epsilon) = 0$, where $\mathbb{E}[X_1^2] = \text{var}(X_1) = \frac{1}{3}$, as $\mathbb{E}[X_1] = 0$. As such, $Y_n \xrightarrow{p} \frac{1}{3}$.

4 Almost Sure Convergence

(a) Yes. Since X_n oscillates between two values infinitely often, it cannot converge a.s.

(b) Yes. $\lim_{n\to\infty} \mathbb{P}(X_n = \frac{1}{y+\frac{1}{n}} \mid Y = y) = \mathbb{P}(X_n = \frac{1}{y} \mid Y = y) = 1$, so $X_n \xrightarrow{a.s.} \frac{1}{Y}$, where $Y \neq 0$. However, since $\mathbb{P}(Y = 0) = 0$, the convergence holds.

(c) No. X_n oscillates infinitely often, between 0 and powers of 2.

(d) Yes, $X_n \xrightarrow{p} 0$, since $\lim_{n\to\infty} \mathbb{P}(|X_n| \geq \epsilon) = \lim_{k\to\infty} \frac{1}{2^k} = 0$.
$\mathbb{E}[X] = 0$, while $\mathbb{E}[X_n] = 1$.

5 Compression of a Random Source

(a) We have $-\frac{1}{n}\log_2 p(X_1, \ldots, X_n) = \frac{1}{n}\left(\log_2\left(\frac{1}{p(x_1)}\right) + \log_2\left(\frac{1}{p(x_2)}\right) + \cdots + \log_2\left(\frac{1}{p(x_n)}\right)\right)$, and $\mathbb{E}\left[\log_2 \frac{1}{p(X_1)}\right] = H(X_1)$. Thus, from the SLLN, we get that $-\frac{1}{n}\log_2 p(X_1, \ldots, X_n) \xrightarrow{a.s.} H(X_1)$.

(b)

$$\begin{aligned}
\mathbb{P}((X_1, \ldots, X_n) \in A_e^{(n)}) &= \mathbb{P}\left(2^{-n(H(X_1)+\epsilon)} \leq p(x_1, \ldots, x_n) \leq 2^{-n(H(X_1)-\epsilon}\right) \\
&= \mathbb{P}(-n(H(X_1) + \epsilon) \leq \log_2 p(x_1, \ldots, x_n) \leq -n(H(X_1) - \epsilon)) \\
&= \mathbb{P}(H(X_1) + \epsilon \geq -\frac{1}{n}\log_2 p(x_1, \ldots, x_n) \geq H(X_1) - \epsilon) \\
&= \mathbb{P}\left(\left|-\frac{1}{n}\log_2 p(x_1, \ldots, x_n) - H(X_1)\right| \leq \epsilon\right)
\end{aligned}$$

By the WLLN, we have $\lim_{n\to\infty} \mathbb{P}\left(\left|-\frac{1}{n}\log_2 p(x_1, \ldots, x_n) - H(X_1)\right| \geq \epsilon\right) = 0$, so that for n sufficiently large, $\mathbb{P}((X_1, \ldots, X_n) \in A_\epsilon^{(n)}) > 1 - \epsilon$.

(c) From part (b), have that

$$\begin{aligned} 1 &= \sum_{(x_1,\ldots,x_n)\in\mathcal{X}^n} p((x_1,\ldots,x_n)) \\ &\geq \sum_{(x_1,\ldots,x_n)\in A_\epsilon^{(n)}} p((x_1,\ldots,x_n)) \geq \left|A_\epsilon^{(n)}\right| 2^{-n(H(X_1)+\epsilon)} \end{aligned}$$

So that $\left|A_\epsilon^{(n)}\right| \leq 2^{n(H(X_1)+\epsilon)}$. In addition, $(1-\epsilon) \leq \mathbb{P}\left((X_1,\ldots,X_n) \in A_\epsilon^{(n)}\right) \leq \left|A_\epsilon^{(n)}\right| 2^{-n(H(X_1)-\epsilon)}$, so $\left|A_\epsilon^{(n)}\right| \geq (1-\epsilon)2^{n(H(X_1)-\epsilon)}$.

(d) $\mathbb{P}((X_1,\ldots,X_n) \in B_n) = \mathbb{P}\left((X_1,\ldots X_n) \in \left(B_n \cap A_\epsilon^{(n)}\right) \cup \left(B_n \setminus A_\epsilon^{(n)}\right)\right)$. From part (b), we get that this is equivalent to $\mathbb{P}\left((X_1,\ldots,X_n) \in A_e^{(n-x)} \cup \left(B_n \setminus A_e^{(n)}\right)\right) = \mathbb{P}\left((X_1,\ldots,X_n) \in A_e^{(n)}\right)$ for sufficiently large n. However, from part (c), $(1-\epsilon)2^{n(H(X_1)-\epsilon)} \leq \left|A_\epsilon^{(n)}\right| \leq 2^{n(H(X_1)+\epsilon)}$, such that $\left|A_\epsilon^{(n)}\right| \geq |B_n|$, so that $\mathbb{P}\left((X_1,\ldots,X_n) \in B_n\right) \to 0$ as $n \to \infty$.

(e) We have

$$\begin{aligned} \mathbb{E}\left[L_n\right] &= \sum_{\mathcal{X}^n} p((x_1,\ldots,x_n)) \\ &\leq \sum_{(x_1,\ldots,x_n)\in A_\epsilon^{(n)}} (1+nH(X_1)) + \sum_{(x_1,\ldots,x_n)\notin A_\epsilon^{(n)}} (1+n\lceil\log_2|\mathcal{X}|\rceil) \\ &\leq 1+nH(X_1) + (1+n\lceil\log_2|\mathcal{X}|\rceil)\left(1-\mathbb{P}\left((x_1,\ldots,x_n)\in A_\epsilon^{(n)}\right)\right) \\ &\leq 1+nH(X_1)+(1+n\lceil\log_2|\mathcal{X}|\rceil)\epsilon \end{aligned}$$

So we get that $\frac{\mathbb{E}[L_n]}{n} \leq \frac{1}{n}\left(1+nH(X_1)+(1+n\lceil\log_2|\mathcal{X}|\rceil)\epsilon\right)$. Taking the limit as $n\to\infty$, we have $\frac{\mathbb{E}[L_n]}{n} \leq H(X_1)+\epsilon'$, where $\epsilon' = \epsilon\lceil\log_2|\mathcal{X}|\rceil$, and thus the number of bits per symbol can be made arbitrarily close to the entropy.

6 [Bonus] Balls and Bins: Poisson Convergence

(a)

(b)

(c)

(d) (i)

(ii)

(iii)

(e)