Question 1

Recall question 1 from assignment 2 (we rename the variables): If $F(y) = P(Y \leq y)$ is continuous, then $X = F(Y)$ has a uniform distribution on [0,1].

Let $C(x)$ be the CDF of the Cantor distribution, which is continuous and non-decreasing. Define on $[0,1]$ the function $F(x) = \frac{x+C(x)}{2}$. Note F is continuous and strictly increasing; also, $F(1) = 1, F(0) = 0$. Hence it has inverse $F^{-1} : [0,1] \to [0,1]$, which is also continuous and strictly increasing.

$F(x)$ satisfies the required properties as a CDF as stated in class. Let Y be a random variable with CDF F. Y does not have a density, since the Cantor distribution does not have a density.

Finally, by the assignment question, $X = F(Y)$ has a uniform distribution on $[0,1]$. Now $Y = F^{-1}(X)$, as required.

Problem 2. Suppose $X_1, X_2, \ldots$ are iid random variables with distribution

$$\Pr[X_k - 2^{2j}] = 2^{-j}$$

for $j = 1, 2, \ldots$.

(i) Let N_n be the number of X_k's such that $X_k = 2^{2n}$. Prove that for every $m_1 < \infty$ there exists $p_1 > 0$ and $n_1 < \infty$ such that for every $n \geq n_1$, $\Pr[N_n \geq m_1] \geq p_1$.

We know $N_n = \sum_{i=1}^{2^n} \mathbb{1}_{X_i = 2^{2n}} = \sum_{i=1}^{2^n} \text{Bernoulli}(2^{-n}) = \text{Binomial}(2^n, 2^{-n})$. Thus,

$$\begin{aligned}
\Pr[N_n \geq m_1] &= \Pr[\text{Binomial}(2^n, 2^{-n}) \geq m_1] \\
&\geq \binom{2^n}{m_1} \left(\frac{1}{2^n}\right)^{m_1} \\
&\geq \left(\frac{2^n}{m_1}\right)^{m_1} \left(\frac{1}{2^n}\right)^{m_1} \\
&= m_1^{-m_1}
\end{aligned}$$

The first inequality follows from the fact that seeing $\geq m_1$ heads from 2^n coin flips implies that some subset of coins of size m_1 consists of all heads. The second inequality is that $\binom{n}{k} \geq (n/k)^k$. Thus, we can choose $p_1 = m_1^{-m_1}$ and $n_1 = 1$.

(ii) Let M_n be the number of X_k's such that $X_k > 2^{2n}$. Prove that there exists $p_2 > 0$ and $n_2 < \infty$ such that for every $n \geq n_2$, $\Pr[M_n = 0] \geq p_2$.

We know $M_n = \sum_{i=1}^{2^n} \mathbb{1}_{X_i > 2^{2n}} = \sum_{i=1}^{2^n} \text{Bernoulli}(2^{-n-1} + 2^{-n-2} + \cdots) = \text{Binomial}(2^n, 2^{-n})$. Thus,

$$\begin{aligned} \Pr[M_n = 0] &= \Pr[\text{Binomial}(2^n, 2^{-n}) = 0] \\ &= (1 - 2^{-n})^{2^n} \\ &\to e^{-1} \end{aligned}$$

Because this probability converges to $\frac{1}{e}$, we choose n_2 sufficiently large so that $p_2 > \frac{1}{2e}$.

(iii) Let $S_n^* = \sum_{k=1}^{2^n} X_k \mathbb{1}_{X_k < 2^{2n}}$. Prove that for some $a < \infty$ there exists $p_3 < 1$ and $n_3 < \infty$ such that for every $n \geq n_3$, $\Pr[S_n^* \geq a 2^{2n}] \leq p_3$.

It is sufficient to show that $\mathrm{E}[S_n^*] \leq 2^{2n}$ by Markov's inequality: $\Pr[S_n^* \geq a2^{2n}] \leq \Pr[S_n^* \geq a\,\mathrm{E}[S_n^*]] \leq \frac{1}{a}$.

$$\begin{aligned} \mathrm{E}[S_n^*] &= \sum_{k=1}^{2^n} \mathrm{E}\left[X_k \mathbb{1}_{X_k < 2^{2n}}\right] \\ &= \sum_{k=1}^{2^n} \sum_{\ell=1}^{n-1} 2^{-\ell} 2^{2\ell} \\ &\leq \sum_{k=1}^{2^n} 2^n = 2^{2n} \end{aligned}$$

(iv) Let $S_n = \sum_{k=1}^{2^n} X_k$. Prove that there is no sequence of real numbers μ_n such that $S_n/\mu_n \to 1$ in distribution.

It suffices to show that, even for large n, a large range of options for S_n are still possible. The following computation holds for all n sufficiently large.

Define

$$T_n^* = \sum_{k=1}^{2^n} X_k \mathbb{1}_{X_k \leq 2^{2n}}$$

So T_n^* is S_n^* where we also include X_k's that are 2^{2n}. We can use a near-identical argument to (iii) to show that $\mathrm{E}[T_n^*] \leq 2^{2n+1}$, which means that, by Markov's inequality

$$\Pr[T_n^* < a2^{2n+1}] \geq 1 - \frac{1}{a}$$

From part (ii) we know that $M_n = 0$ with probability at least p_2; pick a such that $\frac{1}{a} \leq p_2$, so that $\Pr[M_n = 0 \text{ and } T_n^* < a2^{2n+1}] > 0$ by union bound. Call this probability p_4. Since $S_n = T_n^*$ if $M_n = 0$,

$$\Pr[S_n \leq (2a)2^{2n}] \geq \Pr[M_n = 0, T_n^* < (2a)2^{2n}] = p_4$$

for a constant p_4.

Further, using (i),

$$\Pr[S_n \geq (4a)2^{2n}] \geq \Pr[N_n \geq 4a] \geq p_1$$

for some constant p_1.

Thus, for any sequence μ_n, for n sufficiently large, $S_n/\mu_n \leq (2a)2^{2n}/\mu_n$ and $S_n/\mu_n \geq (4a)2^{2n}/\mu_n$ each occur with constant probability. Because they differ by a factor of two, $\Pr[|S_n/\mu_n - 1| \geq \frac{1}{3}]$ is always positive (since at least one of the above events satisfies this criterion for any μ_n). This implies that S_n/μ_n does not converge in distribution.

Problem 3. For a random variable X, the Laplace transform of X is

$$L_X(t) = \mathrm{E}(e^{-tX}).$$

(i) Prove that, if X and Y are independent, then $L_{X+Y}(t) = L_X(t)L_Y(t)$, provided the Laplace transforms exist.

First, we note that if X and Y are independent, so are $f(X)$ and $g(Y)$ for $f, g : \mathbb{R} \to \mathbb{R}$ measurable:

$$\begin{aligned}
\Pr[f(X) < a, g(Y) < b] &= \Pr[X \in f^{-1}(\infty, a), Y \in g^{-1}(\infty, b)] \\
&= \Pr[X \in f^{-1}(\infty, a)] \Pr[Y \in g^{-1}(\infty, b)] \\
&= \Pr[f(X) < a] \Pr[g(Y) < b]
\end{aligned}$$

Thus, e^{-tX} and e^{-tY}, and the statement follows:

$$L_{X+Y}(t) = \mathrm{E}(e^{-tX}e^{-tY}) = \mathrm{E}(e^{-tX})\,\mathrm{E}(e^{-tY}) = L_X(t)L_Y(t)$$

(ii) Compute $L_{S_n}(t)$, where S_n is Binomial$(n, 1/2)$.

We know $S_n = \sum_{i=1}^n C_i$, where the C_i's are iid uniformly chosen from $\{0, 1\}$. Using (i),

$$L_{S_n}(t) = \Big(L_{C_1}(t)\Big)^n = \left(\frac{1}{2} + \frac{1}{2}e^{-t}\right)^n$$

(iii) Compute $L_Z(t)$, where Z is standard normal.

Using the Radon-Nikodym derivative of Z with respect to Lebesgue measure, we get that

$$\begin{aligned}
L_Z(t) &= \int e^{-tx}\left(\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\right)dx \\
&= \int \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x+t)^2}{2} + \frac{t^2}{2}\right)dx \\
&= e^{t^2/2}\int \frac{1}{\sqrt{2\pi}}e^{-y^2/2}dy \qquad (y = x+t) \\
&= e^{t^2/2}
\end{aligned}$$

(iv) Prove that for some constant c and every real t, $L_{cR_n}(t) \to L_Z(t)$, where $R_n = (S_n - n/2)/\sqrt{n}$.

We can treat R_n as the sum of n coin flips, normalized by $c/\sqrt{n}$, and a constant $-c\sqrt{n}/2$. Thus,

$$\begin{aligned} L_{cR_n}(t) &= \left(\frac{1}{2} + \frac{1}{2}e^{-ct/\sqrt{n}}\right)^n e^{ct\sqrt{n}/2} \\ &= \left(\frac{\exp\left(\frac{-ct}{2\sqrt{n}}\right) + \exp\left(\frac{ct}{2\sqrt{n}}\right)}{2}\right)^n \end{aligned}$$

We use the Taylor series approximation:

$$\begin{aligned} &= \left(1 + \frac{1}{2!}\frac{c^2t^2}{4n} + \frac{1}{4!}\frac{c^4t^4}{16n^2} + \cdots\right)^n \\ &\approx \exp\left(\frac{c^2t^2}{8} + \frac{c^4t^4}{384n} + \cdots\right) \\ &\to \exp\left(\frac{c^2t^2}{8}\right) \end{aligned}$$

Where we use that $(1 + \frac{x}{n})^n$ conveges to e^x and that the lower order terms of the above expression vanish as n grows large.

When $c = 2$, this is $L_Z(t)$.