HOMEWORK 5 SOLUTIONS

Problem 1

a) We can show this immediately using Corollary 6.4 from Chapter 1 of Durret.

b)Let $X = c$ and for any $\epsilon > 0$ define the open interval $A_\epsilon = (c - \epsilon, c + \epsilon)$. The boundary $\partial A_\epsilon = \{c - \epsilon, c + \epsilon\}$, therefore $\mathbb{P}\{X \in \partial A_\epsilon\} = 0$. By the Portmanteau Theorem

$$\begin{aligned} 1 = \mathbb{P}\{X \in A_\epsilon\} &= \lim_{n\to\infty} \mathbb{P}\{X_n \in A_\epsilon\} \\ &= \lim_{n\to\infty} \mathbb{P}\{|X_n - c| < \epsilon\} \end{aligned}$$

Problem 2: Observe that f_n is always nonnegative. To show that is a density function we need to compute the integral:

$$\int_0^1 1 - \cos(2\pi n x)\, dx = \left[x - \frac{\sin(2\pi n x)}{2\pi n}\right]_{x=0}^1 = 1.$$

For $x \in [0, 1]$, define

$$F_n(x) = x - \frac{\sin(2\pi n x)}{2\pi n}.$$

Observe that for all $x \in [0, 1]$,

$$\lim_{n\to\infty} F_n(x) = F(x) = x \tag{1}$$

which is the distribution function of a uniform [0,1] r.v.

We actually proved that for any point of continuity of $F(x)$ the convergence (1) holds, therefore the μ_n converge weakly to the Lebesque measure.

Problem 3: Let $\sigma^2 = Var(X_1)$. Classical CLT implies that $\dfrac{S_n}{\sigma\sqrt{n}} \Longrightarrow X \sim N(0, 1)$. Let $M > 0$. The set $(\sigma M, +\infty)$ is open and so by the Portmanteau Theorem,

$$\liminf_{n\to\infty} \mathbb{P}\left\{\frac{S_n}{\sigma\sqrt{n}} > M\right\} \geq \mathbb{P}\{X > M\} = 1 - \Phi(M) > 0. \tag{2}$$

Hence there exists an $\epsilon > 0$ and $N = N(\epsilon)$ such that for all $n > N$,

$$\mathbb{P}\left\{\frac{S_n}{\sqrt{n}} > \sigma M\right\} = \epsilon > 0.$$

An easy lemma from measure theory shows that if $\mathbb{P}\{A_n\} \geq \epsilon$ for all n, then $\mathbb{P}\{A_n \quad i.o.\} \geq \epsilon$. (Why?). Therefore

$$\mathbb{P}\left\{\frac{S_n}{\sqrt{n}} > \sigma M \quad i.o.\right\} \geq \epsilon > 0. \tag{3}$$

To finish the proof, observe that by Kolmogorov's (or Hewitt-Savage) 0-1 law we have

$$\mathbb{P}\left\{\limsup_{n\to\infty}\frac{S_n}{\sqrt{n}} > \sigma M\right\} \in \{0,1\}. \tag{4}$$

and since

$$\mathbb{P}\left\{\limsup_{n\to\infty}\frac{S_n}{\sqrt{n}} > \sigma M\right\} = \mathbb{P}\left\{\frac{S_n}{\sqrt{n}} > \sigma M \quad i.o.\right\} > 0, \tag{5}$$

it has to be that $\limsup_{n\to\infty}\frac{S_n}{\sqrt{n}} > \sigma M$a.s. Since M was arbitrary take M to infinity.

Problem 4: Define $X_{n,j}$ to be the number of trials required to collect j coupons ($j \geq 1$), if we already collected $j-1$ coupons. Then we have that

(1) $\tau_k^n = \sum_{j=1}^{k} X_{n,j}$.

(2) $X_{n,1} \equiv 1, \mathbb{P}\{X_{n,j} = l\} = \left(1 - \frac{j-1}{n}\right)\left(\frac{j-1}{n}\right)^{l-1}$ for $l \geq 1$

(3) $\mathbb{E}X_{n,j} = \frac{n}{n-j+1}$

(4) $Var(X_{n,j}) = \frac{j-1}{n}\left(1 - \frac{j-1}{n}\right)^{-2}$.

Therefore, $S_n = \sum_{j=1}^{\lfloor n\theta \rfloor} X_{n,j}$. We compute

$$\begin{aligned}\mathbb{E}S_n &= n\sum_{j=1}^{\lfloor n\theta\rfloor}\frac{1}{n-j+1} = n\sum_{j=n-\lfloor n\theta\rfloor+1}^{n}\frac{1}{j}\\ &\sim n(\log n - \log(n - \lfloor n\theta\rfloor + 1) \sim n\log\frac{1}{1-\theta}.\end{aligned}$$

$$\begin{aligned}Var(S_n) &= n\sum_{j=1}^{\lfloor n\theta\rfloor}\frac{j-1}{(n-j+1)^2} = n\sum_{j=n-\lfloor n\theta\rfloor+1}^{n}\frac{n-j}{j^2}\\ &= \sum_{j=n-\lfloor n\theta\rfloor+1}^{n}\frac{1-j/n}{(j/n)^2} \sim n\int_{1-\theta}^{1}\frac{1-x}{x^2}\,dx\\ &\sim n\left(\frac{\theta}{1-\theta} - \log\frac{1}{1-\theta}\right).\end{aligned}$$

We would like to apply Lindeberg - Feller on the sequence $\frac{S_n - \mathbb{E}S_n}{\sqrt{n}}$. We check the two conditions:

(i) $\sum_{j=1}^{\lfloor n\theta\rfloor}\frac{Var(X_{n,j})}{n} \longrightarrow \frac{\theta}{1-\theta} - \log\frac{1}{1-\theta} = \sigma^2 > 0$ (why?)

(ii) This requires to show $\lim_{n\to\infty} n^{-1} \sum_{j=1}^{\lfloor n\theta \rfloor} \mathbb{E}\left((X_{n,j} - \mathbb{E}X_{n,j})^2 1\{|X_{n,j} - \mathbb{E}X_{n,j}| \geq \varepsilon\sqrt{n}\}\right) = 0.$

To obtain this, let us show that for the $X_{n,j}$'s we care about, all moments are bounded. Note that

$$\mathbb{P}\{X_{n,j} \geq l\} = \sum_{k=l}^{\infty} \left(1 - \frac{j-1}{n}\right)\left(\frac{j-1}{n}\right)^{k-1} = \left(\frac{j-1}{n}\right)^{l-1} \leq \theta^{l-1}$$

Hence

$$\mathbb{E}(X_{n,j}^p) = \int_0^{+\infty} \mathbb{P}\{X_{n,j} > t\} p t^{p-1}\, dt \leq \int_0^{+\infty} \theta^{\lceil t \rceil - 1} p t^{p-1}\, dt$$
$$= \theta^{-1} \int_0^{\infty} e^{t \log\theta} p t^{p-1}\, dt = C(p, \theta) < +\infty$$

since $0 < \theta < 1$. Consequently

$$\begin{aligned}\mathbb{E}\left((X_{n,j} - \mathbb{E}X_{n,j})^2 1\{|X_{n,j} - \mathbb{E}X_{n,j}| \geq \varepsilon\sqrt{n}\}\right) &\leq \left(\mathbb{E}|X_{n,j} - \mathbb{E}X_{n,j}|^4\right)^{1/2} \mathbb{P}\{|X_{n,j} - \mathbb{E}X_{n,j}| \geq \varepsilon\sqrt{n}\}^{1/2} \\ &\leq C \frac{1}{\varepsilon\sqrt{n}} \mathbb{E}|X_{n,j} - \mathbb{E}X_{n,j}| \\ &\leq \frac{C}{\varepsilon\sqrt{n}}.\end{aligned}$$

Now for (ii):

$$\lim_{n\to\infty} n^{-1} \sum_{j=1}^{\lfloor n\theta \rfloor} \mathbb{E}\left((X_{n,j} - \mathbb{E}X_{n,j})^2 1\{|X_{n,j} - \mathbb{E}X_{n,j}| \geq \varepsilon\sqrt{n}\}\right) \leq \lim_{n\to\infty} n^{-1} \lfloor n\theta \rfloor \frac{C}{\varepsilon\sqrt{n}} = 0.$$

We conclude that $\frac{S_n - \mathbb{E}S_n}{\sqrt{n}} \xrightarrow{w} N(0, \sigma^2)$.