Probability, solutions to HW13

1. By Proposition 3.8.6 with \( a_k = 1/n \) for \( k = 1, 2, ..., n \)

\[
\text{Var}[\bar{X}] = \frac{\sigma^2}{n} + \frac{1}{n^2} \sum_{i \neq j} \text{Cov}[X_i, X_j] < \frac{\sigma^2}{n}
\]

since \( \text{Cov}[X_i, X_j] < 0 \) for all \( i \neq j \). Let \( c = \epsilon/\sqrt{\text{Var}[\bar{X}]} \) and apply Cheb to \( \bar{X} \) which has mean \( \mu \) and variance \( \text{Var}[\bar{X}] \):

\[
P(|\bar{X} - \mu| > \epsilon) = P \left(|\bar{X} - \mu| > c\sqrt{\text{Var}[\bar{X}]}\right) \leq \frac{\text{Var}[\bar{X}]}{\epsilon^2} < \frac{\sigma^2}{n\epsilon^2} \rightarrow 0
\]
as \( n \rightarrow \infty \). Note that this result shows that the Law of Large Numbers may hold for random variables that are not independent.

3. Let \( Y_k = g(X_k) \) and consider the sequence of random variables \( Y_1, Y_2, ... \) and its sample mean \( \bar{Y} \). By LLN, \( \bar{Y} \overset{p}{\rightarrow} \mu \) as \( n \rightarrow \infty \) where \( \mu \) is the mean of the \( Y_k \), that is

\[
\mu = E[Y_k] = E[g(X_k)] = \int g(x)f_{X}(x)dx
\]

In particular, if the \( X_k \) are \text{unif}[0, 1], we get

\[
\mu = \int_0^1 g(x)dx
\]

so in conclusion

\[
\frac{1}{n} \sum_{k=1}^{n} g(X_k) \overset{p}{\rightarrow} \int_0^1 g(x)dx
\]
as \( n \rightarrow \infty \). We can use this result to approximate the integral as

\[
\int_0^1 g(x)dx \approx \frac{1}{n} \sum_{k=1}^{n} g(X_k)
\]

where we simulate \( n \) observations of the \( X_k \). This is an alternative way to integrate via simulation compared to what we did in class. The method there requires the integrand to be bounded which is not necessary here.
4. Let $Y_k = 1/X_k$ and note that $H_n = 1/\bar{Y}$. First apply LLN to the $Y_k$ to get $\bar{Y} \xrightarrow{P} \mu$ where

$$\mu = E[Y_k] = E \left[ \frac{1}{X_k} \right] = \int_0^1 \frac{1}{x} 3x^2 dx = \frac{3}{2}$$

and by Corollary 4.2.3, $H_n \xrightarrow{P} 2/3$.

9. In a Poisson process with rate $\lambda$, the time $T$ between two consecutive events has an exponential distribution with parameter $\lambda$. The mean and variance of $T$ are $1/\lambda$ and $1/\lambda^2$, respectively. The $n$th event comes after a time $S = \sum_{k=1}^n T_k$

and hence CLT gives that

$$S \xrightarrow{d} N \left( \frac{n}{\lambda}, \frac{n}{\lambda^2} \right)$$

where $T_1, T_2, \ldots$ are i.i.d. $\exp(\lambda)$. We have $n = 10^6$ and $\lambda = 10^4$ and get

$$S \xrightarrow{d} N(100, 0.01)$$

and hence

$$P(S \leq 100.2) = \Phi \left( \frac{100.2 - 100}{\sqrt{0.01}} \right) = \Phi(2) = 0.98$$

2(a) Since $f(x) = 1/(b - a)$ for $a \leq x \leq b$, we get

$$M_X(t) = E[e^{tX}] = \int_a^b e^{tx} \frac{1}{b - a} dx = \frac{e^{tb} - e^{ta}}{b - a}$$

(b) By (a) with $a = 0, b = 1$ and Proposition 3.11.9, the sum $X + Y$ has mgf

$$M_{X+Y}(t) = M_X(t)M_Y(t) = (e^t - 1)(e^t - 1)$$

whereas the unif[0, 2] distribution has mgf

$$M(t) = \frac{e^{2t} - 1}{2}$$
by uniqueness of the mgf, we conclude that $X + Y$ is not uniform on $[0, 2]$.

3. For a single household the mean is $\mu = 0.1 + 1.0.6 + 2.0.3 = 1.2$. Further, $E[X^2] = 0^2 \cdot 0.1 + 1^2 \cdot 0.6 + 2^2 \cdot 0.3 = 1.8$ which gives $\sigma^2 = 1.8 - 1.2^2 = 0.36$ and standard deviation $\sigma = \sqrt{0.36} = 0.6$. The total number of cars $S$ is then the sum of 200 random variables, each with $\mu = 1.2$ and $\sigma = 0.6$ and by the central limit theorem, $S_n$ is approximately $N(200 \cdot 1.2, 200 \cdot 0.36) = N(240, 72)$.

(a) $P(S \leq 250) = \Phi\left(\frac{250 - 240}{\sqrt{72}}\right) = \Phi(1.18) = 0.88$

(b) Let $n$ be the number of parking spaces. We must then find $n$ such that $P(S \leq n) = 0.95$ which gives

$$P(S \leq n) \approx \Phi\left(\frac{n - 240}{\sqrt{72}}\right) = 0.95$$

which gives $(n - 240)/\sqrt{72} = 1.64$ which gives $n \approx 253.9$ so we need at least 254 spaces.