STA355H1 - Assignment 1

1. (a) $\mathbb{E}(X_i) = 0\mathbb{P}(X_i = 0) + i\mathbb{P}(X_i = i) - i\mathbb{P}(X_i = -i) = (2i)^{-1} - (2i)^{-1} = 0$ and $\text{Var}(X_i) = \mathbb{E}(X_i^2) - \mathbb{E}(X_i)^2 = \mathbb{E}(X_i^2) = 0\mathbb{P}(X_i = 0) + i^2\mathbb{P}(X_i = i) + i^2\mathbb{P}(X_i = -i) = \frac{1}{2} + \frac{1}{2} = 1$.

   (b) Define $Y_i = \sqrt{n}X_i$ so that $\mathbb{E}(Y_i) = \sqrt{n}\mathbb{E}(X_i) = 0 < \infty$. Since the X_i's are independent, the Y_i's are independent as well. Then, by the WLLN, $\frac{1}{\sqrt{n}}\Sigma_{i=1}^n X_i = \frac{1}{n}\Sigma_{i=1}^n Y_i \xrightarrow{p} 0$.

   (c) See the attached code and plots, which show the sum's distribution with the corresponding normal distribution. I conclude that the normal approximation for this distribution is not entirely accurate, since the sum's distribution appears to have a larger kurtosis (higher peak, smaller tails) compared to the normal distribution, and this difference increases with n.

   (d) Notice that

$$\text{kurt}(X_i) = \frac{\mathbb{E}[(X_i - \mathbb{E}(X_i))^4]}{\sigma_i^4} = \frac{0^4\mathbb{P}(X_i = 0) + i^4\mathbb{P}(X_i = i) + i^4\mathbb{P}(X_i = -i)}{\text{Var}(X_i)^2} = \frac{i^2/2 + i^2/2}{1^2} = i^2.$$

   Then, by the proposition in the provided document "Skewness and Kurtosis and the Central Limit Theorem",

$$\begin{aligned}\text{kurt}(S_n) &= (\Sigma_{i=1}^n \sigma_i^2)^{-2}\Sigma_{i=1}^n \sigma_i^4(\text{kurt}(X_i) - 3) + 3 = (\Sigma_{i=1}^n 1)^{-2}\Sigma_{i=1}^n(\text{kurt}(X_i) - 3) + 3 \\ &= \frac{\Sigma_{i=1}^n i^2}{n^2} - \frac{\Sigma_{i=1}^n 3}{n^2} + 3 = \frac{n(n+1)(2n+1)}{6n^2} - \frac{3}{n} + 3 = \frac{2n^2 + 3n + 1 - 18 + 18n}{6n} \\ &= \frac{n}{3} + \frac{7}{2} - \frac{17}{6n}\end{aligned}$$

   where $\Sigma_{i=1}^n i^2 = n(n+1)(2n+1)/6$. Using this result, the kurtosis values calculated for $n = 10, 20, 50$, and 100 are $6.55, 10.03, 20.11$, and 36.81 respectively. In other words, as n increases, the kurtosis increases and diverges away from 3, which is the kurtosis of the normal distribution. Thus, the normal approximation is not entirely accurate, confirming my conclusions in (c).

2. (a) By the Taylor expansion,

$$g(X_n) - g(\theta) = g'(\theta)(X_n - \theta) + \frac{1}{2}g''(\theta)(X_n - \theta)^2 + r_n = \frac{1}{2}g''(\theta)(X_n - \theta)^2 + r_n$$

$$\implies \alpha_n^2(g(X_n) - g(\theta)) = \frac{1}{2}g''(\theta)(\alpha_n(X_n - \theta))^2 + \frac{r_n}{(X_n - \theta)^2}(\alpha_n(X_n - \theta))^2.$$

   By the continuous mapping theorem, $(\alpha_n(X_n - \theta))^2 \xrightarrow{d} Z^2$ since $f(x) = x^2$ is continuous. Thus, $\frac{1}{2}g''(\theta)(\alpha_n(X_n - \theta))^2 \xrightarrow{d} \frac{1}{2}g''(\theta)Z^2$. Also, assuming that $\frac{r_n}{(X_n - \theta)^2} \xrightarrow{p} 0$, $\frac{r_n}{(X_n - \theta)^2}(\alpha_n(X_n - \theta))^2 \xrightarrow{d} 0$ by Slutsky's Theorem, and since 0 is a constant, this term actually $\xrightarrow{p} 0$. As a result,

$$\alpha_n^2(g(X_n) - g(\theta)) \xrightarrow{d} \frac{1}{2}g''(\theta)Z^2 + 0 = \frac{1}{2}g''(\theta)Z^2$$

   by Slutsky's Theorem.

   (b) First, $Z/\sigma \sim \mathcal{N}(0, 1)$. Since the sum of n squares of standard normal r.v's follows the χ_n^2 distribution, $(Z/\sigma)^2 \sim \chi_1^2$ and so $Z^2 \sim \sigma^2\chi_1^2$. Thus, $\alpha_n^2(g(X_n) - g(\theta)) \xrightarrow{d} \frac{1}{2}g''(\theta)Z^2$ or $\frac{\sigma^2}{2}g''(\theta)\chi_1^2$.

(c) Extending (a),

$$\begin{aligned}
g(X_n) - g(\theta) &= g'(\theta)(X_n - \theta) + \frac{1}{2}g''(\theta)(X_n-\theta)^2 + \ldots + \frac{1}{k!}g^{(k)}(\theta)(X_n-\theta)^k + r_n \\
&= \frac{1}{k!}g^{(k)}(\theta)(X_n-\theta)^k + r_n \\
\implies \alpha_n^k(g(X_n) - g(\theta)) &= \frac{1}{k!}g^{(k)}(\theta)(\alpha_n(X_n-\theta))^k + \frac{r_n}{(X_n-\theta)^k}(\alpha_n(X_n-\theta))^k.
\end{aligned}$$

Similar to in (a), $(\alpha_n(X_n-\theta))^k \xrightarrow{d} Z^k$ and so $\frac{1}{k!}g^{(k)}(\theta)(\alpha_n(X_n-\theta))^k \xrightarrow{d} \frac{1}{k!}g^{(k)}(\theta)Z^k$. Moreover, assuming that $\frac{r_n}{(X_n-\theta)^k} \xrightarrow{p} 0$, $\frac{r_n}{(X_n-\theta)^k}(\alpha_n(X_n-\theta))^k \xrightarrow{p} 0$ by Slutsky's Theorem. Thus,

$$\alpha_n^k(g(X_n) - g(\theta)) \xrightarrow{d} \frac{1}{k!}g^{(k)}(\theta)Z^k + 0 = \frac{1}{k!}g^{(k)}(\theta)Z^k$$

by Slutsky's Theorem.

3. By the triangle inequality, $|X_i - \bar{X}_n| = |X_i - \mu + \mu - \bar{X}_n| \le |X_i - \mu| + |\bar{X}_n - \mu|$ and $|X_i - \mu| = |X_i - \bar{X}_n + \bar{X}_n - \mu| \le |X_i - \bar{X}_n| + |\bar{X}_n - \mu|$. Then,

$$\begin{aligned}
&|X_i - \mu| - |\bar{X}_n - \mu| \le |X_i - \bar{X}_n| \le |X_i - \mu| + |\bar{X}_n - \mu| \\
\implies &\frac{1}{n}\Sigma_{i=1}^n(|X_i - \mu| - |\bar{X}_n - \mu|) = \frac{1}{n}\Sigma_{i=1}^n|X_i - \mu| - |\bar{X}_n - \mu| \le \frac{1}{n}\Sigma_{i=1}^n|X_i - \bar{X}_n| \\
\le &\frac{1}{n}\Sigma_{i=1}^n(|X_i - \mu| + |\bar{X}_n - \mu|) = \frac{1}{n}\Sigma_{i=1}^n|X_i - \mu| + |\bar{X}_n - \mu|
\end{aligned}$$

Since the X_i's are independent, by the WLLN, $\frac{1}{n}\Sigma_{i=1}^n|X_i - \mu| \xrightarrow{p} \mathbb{E}|X_i - \mu|$ and $\bar{X}_n \xrightarrow{p} \mu$. Thus, both the lower and upper bounds $\xrightarrow{p} \mathbb{E}|X_i - \mu|$, and the squeeze theorem for convergence in probability (proved below as the lemma) then implies $\frac{1}{n}\Sigma_{i=1}^n|X_i - \bar{X}_n| \xrightarrow{p} \mathbb{E}|X_i - \mu|$.

Lemma: For random variables X_n, Y_n, Z_n with $X_n, Z_n \xrightarrow{p} \mu \in \mathbb{R}$, $X_n \le Y_n \le Z_n \implies Y_n \xrightarrow{p} \mu$. Proof: $X_n \le Y_n \le Z_n \implies X_n - \mu \le Y_n - \mu \le Z_n - \mu \implies \{|Y_n - \mu| \ge \varepsilon\} \subseteq \{|X_n - \mu| \ge \varepsilon\} \cup \{|Z_n - \mu| \ge \varepsilon\} \implies \lim_{n\to\infty} \mathbb{P}(|Y_n - \mu| \ge \varepsilon) \le \lim_{n\to\infty} \mathbb{P}(|X_n - \mu| \ge \varepsilon) + \lim_{n\to\infty} \mathbb{P}(|Z_n - \mu| \ge \varepsilon) = 0$.

4. (a) Define $Y = \lambda X$. Note that

$$\begin{aligned}
\text{skew}(Y) &= \frac{\mathbb{E}[(\lambda X - \mathbb{E}(\lambda X))^3]}{(\text{Var}(\lambda X))^{3/2}} = \frac{\lambda^3\mathbb{E}[(X - \mathbb{E}(X))]^3}{\lambda^3(\text{Var}(X))^{3/2}} = \text{skew}(X) \text{ and} \\
\text{kurt}(Y) &= \frac{\mathbb{E}[(\lambda X - \mathbb{E}(\lambda X))^4]}{(\text{Var}(\lambda X))^2} = \frac{\lambda^4\mathbb{E}[(X - \mathbb{E}(X))]^4}{\lambda^4(\text{Var}(X))^2} = \text{kurt}(X).
\end{aligned}$$

Thus, skewness and kurtosis does not depend on λ. Also, since $g(x) = \lambda x$ is monotonic, by the change of variables formula,

$$f_Y(y) = \frac{f_X(g^{-1}(y))}{g'(g^{-1}(y))} = \frac{\lambda^\alpha}{\lambda\Gamma(\alpha)}(\frac{y}{\lambda})^{\alpha-1}e^{-\lambda y/\lambda} = \frac{y^{\alpha-1}e^{-y}}{\Gamma(\alpha)},$$

so $Y \sim \text{Gamma}(\alpha, 1)$ and

$$\mathbb{E}(Y^k) = \int_0^\infty f_Y(y)dy = \frac{1}{\Gamma(\alpha)}\int_0^\infty y^k y^{\alpha-1}e^{-y}dy = \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)}$$

for $k \in \mathbb{N}^+$ by the definition of the gamma function. Using this identity and the fact that $\Gamma(k+1) = k! = k\Gamma(k)$, we have

$$\begin{aligned}
\mathbb{E}(Y) &= \frac{\Gamma(\alpha+1)}{\Gamma(\alpha)} = \frac{\alpha\Gamma(\alpha)}{\Gamma(\alpha)} = \alpha \\
\mathbb{E}(Y^2) &= \frac{\Gamma(\alpha+2)}{\Gamma(\alpha)} = \frac{(\alpha+1)\Gamma(\alpha+1)}{\Gamma(\alpha)} = (\alpha+1)\mathbb{E}(Y) = \alpha(\alpha+1) \\
\mathbb{E}(Y^3) &= \frac{\Gamma(\alpha+3)}{\Gamma(\alpha)} = \frac{(\alpha+2)\Gamma(\alpha+2)}{\Gamma(\alpha)} = (\alpha+2)\mathbb{E}(Y^2) = \alpha(\alpha+1)(\alpha+2) \\
\mathbb{E}(Y^4) &= \frac{\Gamma(\alpha+4)}{\Gamma(\alpha)} = \frac{(\alpha+3)\Gamma(\alpha+3)}{\Gamma(\alpha)} = (\alpha+3)\mathbb{E}(Y^3) = \alpha(\alpha+1)(\alpha+2)(\alpha+3) \\
\mathrm{Var}(Y) &= \mathbb{E}(Y^2) - \mathbb{E}(Y)^2 = \alpha(\alpha+1) - \alpha^2 = \alpha \\
\mathrm{skew}(X) = \mathrm{skew}(Y) &= \frac{\mathbb{E}[(Y-\alpha)^3]}{\alpha^{3/2}} = \frac{\mathbb{E}(Y^3 - 3\alpha Y^2 + 3\alpha^2 Y - \alpha^3)}{\alpha^{3/2}} \\
&= \frac{(\alpha+1)(\alpha+2) - 3\alpha(\alpha+1) + 3\alpha^2 - \alpha^2}{\alpha^{1/2}} = \frac{2}{\alpha^{1/2}} \\
\mathrm{kurt}(X) = \mathrm{kurt}(Y) &= \frac{\mathbb{E}[(Y-\alpha)^4]}{\alpha^2} = \frac{\mathbb{E}(Y^4 - 4\alpha Y^3 + 6\alpha^2 Y^2 - 4\alpha^3 Y + \alpha^4)}{\alpha^2} \\
&= \frac{(\alpha+1)(\alpha+2)(\alpha+3) - 4\alpha(\alpha+1)(\alpha+2) + 6\alpha^2(\alpha+1) - 4\alpha^3 + \alpha^3}{\alpha} \\
&= \frac{3\alpha+6}{\alpha} = 3 + \frac{6}{\alpha}.
\end{aligned}$$

As $\alpha \to \infty$, $\mathrm{skew}(X) \to 0$ and $\mathrm{kurt}(X) \to 3$, which are the skewness and kurtosis of a normal distribution.

(b) Note that

$$\begin{aligned}
\mathbb{E}(S_n^3) &= \mathbb{E}[(X_1 + \ldots + X_n)(X_1 + \ldots + X_n)(X_1 + \ldots + X_n)] \\
&= \Sigma_{i=1}^n \mathbb{E}(X_i^3) + \Sigma_{i \neq j} \mathbb{E}(X_i^2)\mathbb{E}(X_j) + \Sigma_{i \neq j \neq k} \mathbb{E}(X_i)\mathbb{E}(X_j)\mathbb{E}(X_k) \\
&= \Sigma_{i=1}^n \mathbb{E}(X_i^3) = \Sigma_{i=1}^n \sigma_i^3 \mathrm{skew}(X_i)
\end{aligned}$$

since the X_i's are independent, $\mathbb{E}(X_i) = 0$, and $\mathrm{skew}(X_i) = \mathbb{E}(X_i^3)/\sigma_i^3$. Also, $\mathbb{E}(S_n) = \Sigma_{i=1}^n \mathbb{E}(X_i) = 0$ and $\mathrm{Var}(S_n) = \mathrm{Var}(\Sigma_{i=1}^n X_i) = \Sigma_{i=1}^n \mathrm{Var}(X_i)$ by independence. Hence, by definition,

$$\mathrm{skew}(S_n) = \frac{\mathbb{E}[(S_n - \mathbb{E}(S_n))^3]}{(\mathrm{Var}(S_n))^{3/2}} = \frac{\Sigma_{i=1}^n \mathbb{E}(X_i^3)}{(\Sigma_{i=1}^n \mathrm{Var}(X_i))^{3/2}} = \frac{\Sigma_{i=1}^n \sigma_i^3 \mathrm{skew}(X_i)}{(\Sigma_{i=1}^n \sigma_i^2)^{3/2}}.$$