STA347 Probability I
Assignment #3

Due: Novmember 26, 2018 before class starts

1. Solve Problem 4.7.74

Solution. Let N_i be bhe number of organisms in the ith generation. Also assume N_1 is constant. Let $N_{1,j}$ be the number of offsprings of organism j in the first generation. Then $N_2 = N_{1,1} + \cdots + N_{1,N_1}$ and number the second generation from 1 to N_2. Again let $N_{2,j}$ be the number of offsprings of organism j in the second generation. Finally the number of organisms in the third generation is $N_3 = N_{2,1} + \cdots + N_{2,N_2}$. All $N_{1,j}$s and $N_{2,j}$s are independent and identially distributed and have mean μ and variance σ^2. The mean and variance of N_3 can be computed using the law of total expectation, that is,

$$\mathbb{E}[N_3] = \mathbb{E}[\mathbb{E}(N_3 \mid N_2)] = \mathbb{E}[\mathbb{E}(N_{2,1} + \cdots + N_{2,N_2} \mid N_2)] = \mathbb{E}[\mu N_2] = \mu\mathbb{E}[\mathbb{E}(N_{1,1} + \cdots + N_{1,N_1} \mid N_1)] = \mu\mathbb{E}[\mu N_1] = \mu^2 N_1$$

and

$$\begin{aligned}\mathbb{V}\mathrm{ar}(N_3) &= \mathbb{E}[\mathbb{V}\mathrm{ar}(N_3 \mid N_2)] + \mathbb{V}\mathrm{ar}[\mathbb{E}(N_3 \mid N_2)] = \mathbb{E}[\mathbb{V}\mathrm{ar}(N_{2,1} + \cdots + N_{2,N_2} \mid N_2)] + \mathbb{V}\mathrm{ar}[\mathbb{E}(N_{2,1} + \cdots + N_{2,N_2} \mid N_2)] \\ &= \mathbb{E}[\sigma^2 N_2] + \mathbb{V}\mathrm{ar}(\mu N_2) = \sigma^2\mathbb{E}[N_2] + \mu^2\{\mathbb{E}[\mathbb{V}\mathrm{ar}(N_2 \mid N_1)] + \mathbb{V}\mathrm{ar}[\mathbb{E}(N_2 \mid N_1)]\} = \sigma^2\mu N_1 + \mu^2\{\mathbb{E}[\sigma^2 N_1] + \mathbb{V}\mathrm{ar}(\mu N_1)\} \\ &= \sigma^2\mu N_1 + \mu^2\sigma^2 N_1.\end{aligned}$$

2. Solve Problem 4.7.49

Solution. (a) $\mathbb{E}(Z) = \mathbb{E}[\alpha X + (1-\alpha)Y] = \alpha\mathbb{E}(X) + (1-\alpha)\mathbb{E}(Y) = \alpha\mu + (1-\alpha)\mu = \mu$.
(b) The variance of Z is

$$\begin{aligned}\mathbb{V}\mathrm{ar}(Z) &= \mathbb{V}\mathrm{ar}(\alpha X + (1-\alpha)Y) = \mathbb{V}\mathrm{ar}(\alpha X) + \mathbb{V}\mathrm{ar}((1-\alpha)Y) = \alpha^2\mathbb{V}\mathrm{ar}(X) + (1-\alpha)^2\mathbb{V}\mathrm{ar}(Y) = \alpha^2\sigma_X^2 + (1-\alpha)^2\sigma_Y^2 \\ &= \alpha^2(\sigma_X^2 + \sigma_Y^2) - 2\alpha\sigma_Y^2 + \sigma_Y^2 = (\sigma_X^2 + \sigma_Y^2)(\alpha - \sigma_Y^2/(\sigma_X^2 + \sigma_Y^2))^2 + \sigma_Y^2 - (\sigma_Y^2)^2/(\sigma_X^2 + \sigma_Y^2)\end{aligned}$$

hence, $\mathbb{V}\mathrm{ar}(Z)$ is minimized at $\alpha = \sigma_Y^2/(\sigma_X^2 + \sigma_Y^2)$.
(c) The variance of Z when $\alpha = 1/2$ is $(\sigma_X^2 + \sigma_Y^2)/4$ which is less than the minimum of σ_X^2 or σ_Y^2) if and only if $\sigma_X^2 + \sigma_Y^2 < 4\min(\sigma_X^2, \sigma_Y^2)$ which is equivalent to $\max(\sigma_X^2, \sigma_Y^2) < 3\min(\sigma_X^2, \sigma_Y^2)$ or $1/3 < \sigma_X^2/\sigma_Y^2 < 3$.

3. Solve Problem 5.4.21

Solution. (a) the expectation of $\hat{I}$ is

$$\mathbb{E}(\hat{I}) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\Big[\frac{f(X_i)}{g(X_i)}\Big] = \frac{1}{n}\cdot n \cdot \int_a^b \frac{f(x)}{g(x)}\cdot g(x)\,dx = \int_a^b f(x)\,dx = I(f).$$

(b) The variance of $\hat{I}$ is

$$\mathbb{V}\mathrm{ar}(\hat{I}) = \frac{1}{n^2}\sum_{i=1}^{n}\mathbb{V}\mathrm{ar}\Big[\frac{f(X_i)}{g(X_i)}\Big] = \frac{1}{n^2}\cdot n\cdot\mathbb{V}\mathrm{ar}\Big[\frac{f(X_1)}{g(X_1)}\Big].$$

The second moment of $f(X_1)/g(X_1)$ is $I(f)$ is

$$\mathbb{E}\Big[\frac{f(X_1)^2}{g(X_1)^2}\Big] = \int_a^b \frac{f(x)^2}{g(x)^2}\cdot g(x)\,dx = \int_a^b \frac{f(x)^2}{g(x)}\,dx.$$

Hence the variance of $\hat{I}$ is

$$\mathbb{V}\mathrm{ar}(\hat{I}) = \frac{1}{n}\mathbb{V}\mathrm{ar}(\frac{f(X_1)}{g(X_1)}) = \frac{1}{n}\Big(\int_a^b \frac{f(x)^2}{g(x)}\,dx - I(f)^2\Big).$$

(c) The considered function is $f(x) = (2\pi)^{-1/2}\exp(-x^2/2)1(0 < x < 1)$. Let $X_1, X_2, \ldots$ be an i.i.d. sequence from uniform$(0,1)$ having density $g(x) = 1(0 < x < 1)$. Then

$$\hat{I}(f) = \frac{1}{n}\sum_{i=1}^{n}\frac{f(X_i)}{g(X_i)} = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{\sqrt{2\pi}}e^{-X_i^2/2}.$$

Hence Example A used uniform distribution as g.

The variance comparison is equivalent to comparison of the integral

$$\int_0^1 \frac{f(x)^2}{g(x)}\,dx = \mathbb{E}\Big[\Big(\frac{f(X)}{g(X)}\Big)^2\Big]$$

as seen in part (b). Using Cauchy-Schwartz' inequality,

$$(I(|f|))^2 = \Big[\mathbb{E}\Big(\frac{|f(X)|}{g(X)}\cdot 1\Big)\Big]^2 \le \mathbb{E}\Big[\Big(\frac{f(X)}{g(X)}\Big)^2\Big]\mathbb{E}[1^2] = \mathbb{E}\Big[\Big(\frac{f(X)}{g(X)}\Big)^2\Big].$$

The equality holds if and only if $|f(X)|/g(X) = c\cdot 1 = c$, that is, $g(x) \propto |f(x)|$, or $g(x) = |f(x)|/I(|f|)$.

4. Let $f_n(x)$ be the density of X_n. Show that $X_n \longrightarrow X$ in distribution if $f_n(x) \to f(x)$ where $f(x)$ is the density of X. [Hint: Use dominated convergence theorem in integral, that is, if $g_n(x) \to g(x), |g_n(x)| \le h(x), \int h(x)\,dx < \infty$, then $\int |g_n(x) - g(x)|\,dx \to 0$.]

Solution. The problem is called the Lehmann-Scheffe theorem. The following proof is very rigorous and mathematical.

Fix $\epsilon > 0$. There exist $M_1, M_2 > 0$ such that

$$P(|X| > M_1) < \epsilon/7 \text{ and } \int 1(|x| \le M_1)\max(0, f(x) - M_2)\,dx < \epsilon/7.$$

Note that $f(x) = M_2 + \max(0, f(x) - M_2)$ if $f(x) \ge M_2$. The second part is the amount of $f(x)$ which is greater than M_2. Also it is easy to see that $f(x) = \min(M_2, f(x)) + \max(0, f(x) - M_2)$. The convergence $\min(M_2, f_n(x))1(|x| \le M_1) \to \min(M_2, f(x))1(|x| \le M_1)$ and integrability imply $\int 1(|x| \le M_1)|\min(M_2, f_n(x)) - \min(M_2, f_n(x))|\,dx \to 0$ by the dominated convergence theorem. Hence there exists $M_3 > 0$ such that

$$\int 1(|x| \le M_1)|\min(M_2, f_n(x)) - \min(M_2, f_n(x))|\,dx < \epsilon/7$$

for any $n \ge M_3$. Then for any $n \ge M_3$,

$$\begin{aligned}
P(|X_n| > M_1) + \int 1(|x| \le M_1)\max(0, f_n(x) - M_2)\,dx &= 1 - \int 1(|x| \le M_1)\min(M_2, f_n(x))\,dx \\
&\le 1 - \int 1(|x| \le M_1)\min(M_2, f(x))\,dx + \int 1(|x| \le M_1)|\min(M_2, f_n(x) - \min(M_2, f(x))|\,dx \\
&\le P(|X| > M_1) + \int 1(|x| \le M_1)\max(0, f(x) - M_2)\,dx + \frac{\epsilon}{7} \le \frac{3\epsilon}{7}.
\end{aligned}$$

For any z, if $n \geq M_3$, then

$$|\mathrm{cdf}_{X_n}(z) - \mathrm{cdf}_X(z)| = |P(X_n \leq z) - P(X \leq z)|$$
$$\leq P(|X_n| > M_1) + \int 1(|x| \leq M_1) \max(0, f_n(x) - M_2)\, dx + P(|X| > M_1)$$
$$+ \int 1(|x| \leq M_1) \max(0, f(x) - M_2)\, dx + \left| \int 1(|x| \leq M_1, x \leq z)(\min(M_2, f_n(x)) - \min(M_2, f(x)))\, dx \right|$$
$$\leq \frac{3\epsilon}{7} + \frac{\epsilon}{7} + \frac{\epsilon}{7} + \frac{\epsilon}{7} < \epsilon.$$

Therefore $\mathrm{cdf}_{X_n}(z) \to \mathrm{cdf}_X(x)$ and $X_n \xrightarrow{d} X$ as $n \to \infty$.

5. Let $X_1, X_2, \ldots$ be a random sample from uniform$(\theta - 1, \theta + 1)$ where $\theta \in \mathbb{R}$.
(a) Show both $Y_n = \min(X_1, \ldots, X_n) + 1$ and $Z_n = \max(X_1, \ldots, X_n) - 1$ converges to θ in probability.
(b) Show that $n(Y_n - Z_n)$ converges to W in distribution where W is a non-degerate random variable.

Solution. (a) Note that $Y_n > \theta$ and $Z_n < \theta$ since $-1 < X_i - \theta < 1$ for any $i = 1, \ldots, n$. Fix any $\epsilon > 0$. For convenience assume $\epsilon < 2$. $P(|Y_n - \theta| > \epsilon) = P(Y_n - \theta > \epsilon) = P(Y_n > \theta + \epsilon) = P(\min(X_1, \ldots, X_n) + 1 > \theta + \epsilon) = P(X_i > \theta - 1 + \epsilon, i = 1, \ldots, n) = [P(X_1 > \theta - 1 + \epsilon)]^n = (1 - (\theta - 1 + \epsilon - (\theta - 1))/(\theta + 1 - (\theta - 1)))^n = (1 - \epsilon/2)^n \to 0$. Similarly, $P(|Z_n - \theta| > \epsilon) = P(X_i > \theta + 1 - \epsilon, i = 1, \ldots, n) = (1 - \epsilon/2)^n \to 0$.
(b) The joint cumulative distribution function of Y_n and Z_n is, for $\theta < y < \theta + 2, \theta - 2 < z < \theta, y - z < 2$,

$$\begin{aligned}
\mathrm{cdf}_{Y_n, Z_n}(y, z) &= P(Y_n \leq y, Z_n \leq z) = P(Z_n \leq z) - P(Y_n > y, Z_n \leq z) \\
&= P(\max(X_1, \ldots, X_n) - 1 \leq z) - P(\min(X_1, \ldots, X_n) + 1 > y, \max(X_1, \ldots, X_n) - 1 \leq z) \\
&= P(X_1 \leq 1 + z)^n - P(y - 1 < X_1 \leq z + 1)^n = (1 + (z - \theta)/2)^n - (1 + (z - y)/2)^n.
\end{aligned}$$

and the joint probability density function is

$$\mathrm{pdf}_{Y_n, Z_n}(y, z) = \frac{\partial^2}{\partial y \partial z} \mathrm{cdf}_{Y_n, Z_n}(y, z) = n(n-1)(1 + (z - y)/2)^{n-2}/4.$$

Method I: The density of $V_n = Y_n - Z_n$ is, for $0 < v < 2$,

$$\mathrm{pdf}_{V_n}(v) = \int \mathrm{pdf}_{Y_n, Z_n}(y, y - v)\, dy = \int_0^v n(n-1)(1 + (y - v - y)/2)^{n-2}/4\, dy = n(n-1)v(1 - v/2)^{n-2}/4$$

and the density of $W_n = n(Y_n - Z_n) = nV_n$ is

$$\mathrm{pdf}_{W_n}(w) = \mathrm{pdf}_{V_n}(w/n)/n = (n-1)(w/n)(1 - w/(2n))^{n-2}/4 \approx we^{-w/2}/4 \sim \mathrm{gamma}(2, 1/2).$$

Method II: The probability

$$P(n(Y_n - Z_n) \leq w) = P(Y_n - Z_n \leq w/n) = \int_\theta^{\theta+2} \int_{\theta-2}^{\theta} 1(y - z \leq w/n) \cdot n(n-1)(1 + (z - y)/2)^{n-2}/4\, dz\, dy$$

apply change of variables $(y, z) \mapsto (y', z')$ where $y' = y - \theta, z' = z - \theta$.

$$\begin{aligned}
&= \int_0^2 \int_{-2}^0 1(y - z \leq w/n) \cdot n(n-1)(1 + (z - y)/2)^{n-2}/4\, dz\, dy \\
&= \int_0^{w/n} \int_{y - w/n}^0 n(n-1)(1 + (z - y)/2)^{n-2}/4\, dz\, dy \\
&= \int_0^{w/n} n(1 + (z - y)/2)^{n-1}/2|_{y-w/n}^0\, dy \\
&= \int_0^{w/n} n(1 - y/2)^{n-1}/2 - n(1 - w/n/2)^{n-1}/2\, dy \\
&= -(1 - y/2)^n|_0^{w/n} - n(1 - w/n/2)^{n-1}/2(w/n - 0) \\
&= -(1 - w/n/2)^n + 1 - w(1 - w/n/2)^{n-1}/2
\end{aligned}$$

from which the density of $W_n = n(Y_n - Z_n)$ is

$$\begin{aligned}
\mathrm{pdf}_{W_n}(w) &= \frac{d}{dw}\mathrm{cdf}_{W_n}(w) = -n(1-w/(2n))^{n-1}/(-2n) - (1-w/(2n))^{n-1}/2 - w(n-1)(1-w/(2n))^{n-2}/(-4n) \\
&= \frac{n-1}{4n}w\Big(1-\frac{w}{2n}\Big)^{n-2} \approx \frac{w}{4}e^{-w(n-2)/(2n)} \approx \frac{w}{4}e^{-w/2} \sim \mathrm{gamma}(2, 1/2).
\end{aligned}$$

6. [optional] If you solve this problem correctly, the score obtained will be used only to make up deductions acrued in problems 1-5. Prove the following rigorously. It is a bit tedius but provable using only definitions and theorems in the lecture note.
(a) If $X_n \xrightarrow{p} X$ and $\mathbb{E}(|X_n|) \to \mathbb{E}(|X|) < \infty$, then $\mathbb{E}(|X_n - X|) \to 0$.
(b) Consider random variables $X, X_1, X_2, \ldots, Y, Y_1, Y_2, \ldots$ satisfying $X_n \xrightarrow{p} X, Y_n \xrightarrow{p} Y, |X_n| \le Y_n, \mathbb{E}(Y_n) \to \mathbb{E}(Y)$. Prove that $\mathbb{E}[|X_n - X|] \to 0$ as $n \to \infty$.
(c) Let $Z_1, Z_2, \ldots$ be a sequence of i.i.d. random variables satisfying $\mathbb{E}(Z_1) = \mu$ where $|\mu| < \infty$. Prove that the sample mean $\bar{Z}_n = (Z_1 + \cdots + Z_n)/n$ converges to μ in L^1 sense.

Solution. (a) Fix $\varepsilon > 0$. There exists $M, N > 0$ such that $\mathbb{E}[|X|1(|X| > M)] < \varepsilon/5$ and $|\mathbb{E}(|X_n|) - \mathbb{E}(|X|)| < \varepsilon/5$ for all $n \ge N$. Let $g(x) = \max(\min(x, M), -M)$ so that $g(x) = x$ for $-M \le x \le M$ and $|x - g(x)| = |x| - g(|x|)$. Then $g(X_n)$ is bounded and $g(X_n) \xrightarrow{p} g(X)$ by the continuous mapping theorem. The dominated convergence theorem implies $\mathbb{E}[|g(X_n) - g(X)|] \to 0$, hence, there exists $N_2 \ge N$ such that $\mathbb{E}[|g(X_n) - g(X)|] \le \epsilon/5$ for all $n \ge N_2$ and

$$\begin{aligned}
\mathbb{E}[|X_n - X|] &\le \mathbb{E}[|X_n - g(X_n)|] + \mathbb{E}[|g(X_n) - g(X)|] + \mathbb{E}[|g(X) - X|] \\
&= \mathbb{E}(|X_n|) - \mathbb{E}[|g(X_n)|] + \mathbb{E}[|g(X_n) - g(X)|] + \mathbb{E}(|X|) - \mathbb{E}[|g(X)|] \\
&\le \mathbb{E}(|X_n|) + \mathbb{E}(|X|) - 2\mathbb{E}[|g(X)|] + 2\mathbb{E}[|g(X_n) - g(X)|] \\
&\le \mathbb{E}(|X|) + \epsilon/5 + \mathbb{E}(|X|) - 2\mathbb{E}(|X|) + 2\epsilon/5 + 2\epsilon/5 \le \epsilon.
\end{aligned}$$

Then $\limsup_{n\to\infty} \mathbb{E}[|X_n - X|] \le \epsilon$. The arbitrariness of $\epsilon > 0$ implies $\limsup_{n\to\infty} \mathbb{E}[|X_n - X|] = 0$.
(b) Fix any $\epsilon > 0$. There exists $\delta > 0$ such that $\mathbb{E}[Y 1_A] < \epsilon/5$ for any A with $P(A) < \delta$. Also there exists $M, N_1 > 0$ such that $\mathbb{E}(|Y_n - Y|) < \epsilon/5, P(|X_n - X| > \epsilon/5) < \epsilon/5$ for $n \ge N_1$. Then

$$\mathbb{E}(|X_n - X|) = \mathbb{E}(|X_n - X| \cdot 1(|X_n - X| \le \epsilon/5)) + \mathbb{E}(|X_n - X| \cdot 1(|X_n - X| > \epsilon/5))$$

the boundedness $|X_n| \le Y_n, |X| \le Y$ imply

$$\le \frac{\epsilon}{5} + \mathbb{E}[(Y_n + Y)1(|X_n - X| > \frac{\epsilon}{5})]$$

From $Y_n = Y_n - Y + Y \le Y + |Y_n - Y|$

$$\begin{aligned}
&\le \frac{\epsilon}{5} + \mathbb{E}[2Y1(|X_n - X| > \frac{\epsilon}{5})] + \mathbb{E}[|Y_n - Y|1(|X_n - X| > \frac{\epsilon}{5})] \\
&\le \frac{\epsilon}{5} + \frac{2\epsilon}{5} + \mathbb{E}[|Y_n - Y|] \le \frac{3\epsilon}{5} + \frac{\epsilon}{5} < \epsilon.
\end{aligned}$$

Which implies $\mathbb{E}(|X_n - X|) \to 0$.
(c) Let $X_n = (Z_1 + \cdots + Z_n)/n, Y_n = (|Z_1| + \cdots + |Z_n|)/n$. Then the weak law of large numbers implies $X_n \xrightarrow{p} \mathbb{E}(Z_1) = \mu$ and $Y_n \xrightarrow{p} \mathbb{E}(|Z_1|)$. The expectation of Y_n is $\mathbb{E}(Y_n) = n\mathbb{E}(|X_1|)/n = \mathbb{E}(|X_1|) = \mathbb{E}(Y)$. Also by the definition, $|X_n| = |Z_1 + \cdots + Z_n|/n \le (|Z_1| + \cdots + |Z_n|)/n = Y_n$. All conditoins in part (b) is satisfied. Thus $\mathbb{E}(|\bar{Z}_n - \mu|) = \mathbb{E}(|X_n - \mu|) \to 0$.