STA347H1 - Assignment 1

1.11: Let $A_1, A_2, \ldots \subseteq \omega$ and $w \in \Omega$. Define $B_i^w = A_i$ if $w \in A_i$ and A_i^c if $w \notin A_i$. Notice that $\forall A_i$, w either $\in A_i$ or $\notin A_i$. Thus, $\{B_i^w\}_{i=1}^{\infty}$ are independent. By the definition of B_i^w, $\{w\} \subseteq B_1^w, B_2^w, \ldots$, meaning that $\forall n \in \mathbb{N}^+$, $\{w\} \subseteq \cap_{i=1}^n B_i^w$. Then, $\mathbb{P}(\{w\}) \leq \mathbb{P}(\cap_{i=1}^n B_i^w) = \Pi_{i=1}^n \mathbb{P}(B_i^w) = (\frac{1}{2})^n \to 0$ as $n \to \infty$, which contradicts the assumption that Ω is countable.

2.7:

- a) $F_X(x) = \mathbb{P}(YZ \leq x) = \mathbb{P}(YZ \leq x|Z=-1)\mathbb{P}(Z=-1) + \mathbb{P}(YZ \leq x|Z=1)\mathbb{P}(Z=1)$ by the law of total probability $= \frac{1}{2}(\mathbb{P}(-Y < x) + \mathbb{P}(Y < x)) = \frac{1}{2}(\mathbb{P}(Y > -x) + \mathbb{P}(Y < x)) = \mathbb{P}(Y < x) = F_Y(y)$ since $\mathbb{P}(Y > -x) = \mathbb{P}(Y < x)$ for the normal distribution. Differentiating both sides yields $f_X(x) = f_Y(x)$, so $X \sim \mathcal{N}(0,1)$.

  Revised: Instead of differentiating, notice that Theorem 2.2.3 implies X and Y have the same distribution.

- b) $\mathbb{P}(|X| = |Y|) = \mathbb{P}(\{X = Y\} \cup \{X = -Y\}) = \mathbb{P}(X = Y) + \mathbb{P}(X = -Y)$ since the sets are disjoint $= \mathbb{P}(YZ = Y) + \mathbb{P}(YZ = -Y) = \mathbb{P}(Z = 1) + \mathbb{P}(Z = -1) = \frac{1}{2} + \frac{1}{2} = 1$.

  Revised: $\{w \in \Omega : |X(w)| = |Y(w)|\} = \{w \in \Omega : |Z(w)Y(w)| = |Y(w)|\} = \{w \in \Omega : |Z(w)||Y(w)| = |Y(w)|\} = \{w \in \Omega : |Y(w)| = |Y(w)|\} = \Omega \implies \mathbb{P}(|X| = |Y|) = 1$.

- c) Suppose X and Y are independent. Then, for any $A \in \mathcal{B}(\mathbb{R})$ such that $\{0\} \notin A$ and $\mathbb{P}(Y \in A) > 0$, $\mathbb{P}(\{X \in A\} \cap \{Y \in A\}) = \mathbb{P}(\{X \in A\} \cap \{Y \in A\}|Z = 1)\mathbb{P}(Z = 1) + \mathbb{P}(\{X \in A\} \cap \{Y \in A\}|Z = -1)\mathbb{P}(Z = -1) = \frac{1}{2}\mathbb{P}(\{Y \in A\} \cap \{Y \in A\}) + \frac{1}{2}\mathbb{P}(\{-Y \in A\} \cap \{Y \in A\}) = \frac{1}{2}\mathbb{P}(Y \in A) + \frac{1}{2}\mathbb{P}(\varnothing) = \frac{1}{2}\mathbb{P}(Y \in A) = \mathbb{P}(X \in A)\mathbb{P}(Y \in A)$ by assumption. Since $\mathbb{P}(Y \in A) > 0$, this implies $\mathbb{P}(X \in A) = \frac{1}{2}$, but this is a contradiction since A is arbitrary.

2.8: Define $E_n = \{w \in \Omega : X(w) \geq \frac{1}{n}\} \nearrow E = \{w \in \Omega : X(w) > 0\}$, so that $\lim_{n\to\infty} \mathbb{P}(E_n) = \mathbb{P}(\lim_{n\to\infty} E_n) = \mathbb{P}(X > 0) > 0$ by the continuity of probability. Note that $\lim_{n\to\infty} \mathbb{P}(E_n) > 0$ means that $\exists N \in \mathbb{N}$ such that $\forall n \geq N$, $\mathbb{P}(\{X \geq \frac{1}{n}\}) > 0$. The result then follows if $\delta = \frac{1}{N}$.

2.9: Let $A_1, A_2 \in \mathcal{B}(\mathbb{R})$ and define $B_1 = \{x \in \mathbb{R} : f(x) \in A_1\} \in \mathcal{B}(\mathbb{R})$ and $B_2 = \{y \in \mathbb{R} : g(y) \in A_2\} \in \mathcal{B}(\mathbb{R})$. Then, $\mathbb{P}(f(X) \in A_1 \cap g(Y) \in A_2) = \mathbb{P}(X \in \{x \in \mathbb{R} : f(x) \in A_1\} \cap Y \in \{y \in \mathbb{R} : g(y) \in A_2\}\} = \mathbb{P}(X \in B_1 \cap Y \in B_2) = \mathbb{P}(X \in B_1)\mathbb{P}(Y \in B_2)$ by the assumption of independence $= \mathbb{P}(f(X) \in A_1)\mathbb{P}(f(Y) \in A_2)$.

2.11: X and Y are random variables since they map from Ω to $\{0,1\} \subseteq \mathbb{R}$. Also, notice that $\mathbb{P}(\{w \in \Omega : X(w) = 1\} \cap \{w \in \Omega : Y(w) = 1\}) = \mathbb{P}(\{w \in \Omega : \mathbb{I}(w \in A) = 1\} \cap \{w \in \Omega : \mathbb{I}(w \in B) = 1\}) = \mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$ by assumption $= \mathbb{P}(\{w \in \Omega : \mathbb{I}(w \in A) = 1\})\mathbb{P}(\{w \in \Omega : \mathbb{I}(w \in B) = 1\}) = \mathbb{P}(\{w \in \Omega : X(w) = 1\})\mathbb{P}(\{w \in \Omega : Y(w) = 1\})$. By proposition 1.2.2, this equality also holds for the cases of $X = 0$ and $Y = 1$; $X = 1$ and $Y = 0$; and $X = Y = 0$. Thus, X and Y are independent.

3.3: Notice that $A_n \cap B_n \subseteq A_n \implies \cup_{k=n}^{\infty}(A_k \cap B_k) \subseteq \cup_{k=n}^{\infty} A_k \implies \cap_{n=1}^{\infty} \cup_{k=n}^{\infty} (A_k \cap B_k) \subseteq \cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k \implies \limsup_{n\to\infty}(A_n \cap B_n) \subseteq \limsup_{n\to\infty} A_n$. Similarly, $\limsup_{n\to\infty}(A_n \cap B_n) \subseteq \limsup_{n\to\infty} B_n$. Thus, $\limsup_{n\to\infty}(A_n \cap B_n) \subseteq (\limsup_{n\to\infty} A_n) \cap (\limsup_{n\to\infty} B_n)$. When $A_n = B_n$, $\limsup_{n\to\infty}(A_n \cap B_n) = \limsup_{n\to\infty} A_n = (\limsup_{n\to\infty} A_n) \cap (\limsup_{n\to\infty} B_n)$. When $\{A_i\}_i = \{B_j\}_j = \varnothing$ for odd i and even j and $\{B_j\}_j = \{A_k\}_k = \Omega$ for odd j and even k, $\limsup_{n\to\infty}(A_n \cap B_n) = \varnothing \subset \Omega = \Omega \cap \Omega = (\limsup_{n\to\infty} A_n) \cap (\limsup_{n\to\infty} B_n)$.

3.6: For $t \in [0,1]$, construct the sequence $X_1 = \mathbb{I}(t \in [0,1]), X_2 = \mathbb{I}(t \in [0, \frac{1}{2}]), X_3 = \mathbb{I}(t \in [\frac{1}{2}, 1]), X_4 = \mathbb{I}(t \in [0, \frac{1}{3}]), X_5 = \mathbb{I}(t \in [\frac{1}{3}, \frac{2}{3}])$, and so on. The chance that $X_n = 1$ decreases as $n \to \infty$, so $\{X_n\}_n$ converges in probability; however, $X_n = 1$ for infinity many n, so $\{X_n\}_n$ does not converge almost surely.

3.7: $X_n \to X$ a.s. $\iff \mathbb{P}(\lim_{n\to\infty} X_n = X) = \mathbb{P}(\lim_{n\to\infty}(X_n - X) = 0) = 1 \iff (X_n - X) \to 0$ a.s. $X_n \xrightarrow{p} X \iff \forall\varepsilon > 0, \lim_{n\to\infty} \mathbb{P}(|X_n - X| \le \varepsilon) = \lim_{n\to\infty} \mathbb{P}(|(X_n - X) - 0| \le \varepsilon) = 1 \iff (X_n - X) \xrightarrow{p} 0$.

3.8: Notice that $\forall\varepsilon > 0$, $\{w \in \Omega : |X_n(w) - a| \ge \varepsilon\} = \{w \in \Omega : |X_n(w) - a_n + a_n - a| \ge \varepsilon\} \subseteq \{w \in \Omega : |X_n(w) - a_n| + |a_n - a| \ge \varepsilon\}$ by the triangle inequality $\subseteq \{w \in \Omega : |X_n(w) - a_n| \ge \varepsilon\} \cup \{|a_n - a| \ge \varepsilon\}$. Thus, by lemma 1.1.11, $\lim_{n\to\infty} \mathbb{P}(|X_n - a| \ge \varepsilon) \le \lim_{n\to\infty} \left[\mathbb{P}(|X_n - a_n| \ge \varepsilon) + \mathbb{P}(|a_n - a| \ge \varepsilon)\right] = \lim_{n\to\infty} \mathbb{P}(|X_n - a_n| \ge \varepsilon) + \lim_{n\to\infty} \mathbb{P}(|a_n - a| \ge \varepsilon) = 0$ by assumption.

3.9: By Theorem 3.2.7, there exists a subsequence $\{X_{n_k}\}_{n_k}$ such that $X_{n_k} \to X$ a.s. Notice that $\forall n \in \mathbb{N}^+, \exists k \in \mathbb{N}^+$ such that $n_k \le n \le n_{k+1}$. By monotonicity, $X_{n_k} \le X_n \le X_{n_{k+1}} \le X \implies |X_n - X| \le |X_{n_k} - X|$. Thus, $\forall\varepsilon > 0, \exists m \in \mathbb{N}$ such that $n \ge m \implies |X_{n_k} - X| \le \varepsilon \implies |X_n - X| \le \varepsilon$, so $X_n \to X$ a.s.

4.3:

- Bernoulli: $\mathbb{E}(X) = 0p_X(0) + 1p_X(1) = 0(1-p) + 1(p) = 1$. $\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = (0^2 p_X(0) + 1^2 p_X(1)) - p^2 = p - p^2 = p(1-p)$.
- Binomial: Since $X \sim \Sigma_{i=1}^n X_i$ where each $X_i \sim$ Bernoulli(p) and is independent, $\mathbb{E}(X) = \Sigma_{i=1}^n \mathbb{E}(X_i)$ since the sum is finite $= np$ and $\mathrm{Var}(X) = \Sigma_{i=1}^n \mathrm{Var}(X_i)$ by independence $= np(1-p)$.
- Poisson: $\mathbb{E}(X) = \Sigma_{x=0}^\infty x \frac{\lambda^x e^{-\lambda}}{x!} = \lambda e^{-\lambda} \Sigma_{x=1}^\infty \frac{\lambda^{x-1}}{(x-1)!} = \lambda e^{-\lambda} \Sigma_{x=0}^\infty \frac{\lambda^x}{x!} = \lambda e^{-\lambda} e^\lambda$ by the exponential power expansion $= \lambda$. $\mathbb{E}(X(X-1)) = \Sigma_{x=0}^\infty x(x-1) \frac{\lambda^x e^{-\lambda}}{x!} = \lambda^2 e^{-\lambda} \Sigma_{x=2}^\infty \frac{\lambda^{x-2}}{(x-2)!} = \lambda^2 e^{-\lambda} \Sigma_{x=0}^\infty \frac{\lambda^x}{x!} = \lambda^2 e^{-\lambda} e^\lambda$ by the exponential power expansion $= \lambda^2$. $\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \mathbb{E}(X^2 - X) + \mathbb{E}(X) - \mathbb{E}(X)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$.

4.4:

- Bernoulli: $\mathbb{E}(e^{\lambda X}) = e^0 p_X(0) + e^\lambda p_X(1) = 1 - p + e^\lambda p$.
- Binomial: Since $X \sim \Sigma_{i=1}^n X_i$ where each $X_i \sim$ Bernoulli(p) and is independent, by Lemma 4.4.4, $\mathbb{E}(e^{\lambda X}) = \Pi_{i=1}^n \mathbb{E}(e^{\lambda X_i}) = (1 - p + e^\lambda p)^n$.
- Poisson: $\mathbb{E}(e^{tX}) = \Sigma_{x=0}^\infty e^{tx} \frac{\lambda^x e^{-\lambda}}{x!} = e^{-\lambda} \Sigma_{x=0}^\infty \frac{(\lambda e^t)^x}{x!} = e^{-\lambda} \exp(\lambda e^t)$ by the exponential power expansion $= \exp(\lambda(e^t - 1))$.

4.7: The MGF of the normal distribution is $M_X(\lambda) = \exp(\mu\lambda + \frac{1}{2}\sigma^2\lambda^2)$ for some $X \sim \mathcal{N}(\mu, \sigma^2)$. By Lemma 4.4.4, for $S = \Sigma_{i=1}^n X_i$, $M_S(\lambda) = \Pi_{i=1}^n M_{X_i}(\lambda) = \Pi_{i=1}^n \exp(\mu\lambda + \frac{1}{2}\sigma^2\lambda^2) = \exp(\Sigma_{i=1}^n(\mu\lambda + \frac{1}{2}\sigma^2\lambda^2)) = \exp(n\mu\lambda + \frac{n}{2}\sigma^2\lambda^2))$. Thus, $S \sim \mathcal{N}(n\mu, n\sigma^2)$.

Revised: $\exp(n\mu\lambda + \frac{n}{2}\sigma^2\lambda^2))$ is the MGF of some $Y \sim \mathcal{N}(n\mu, n\sigma^2)$. Additionally, $\forall\lambda, M_S(\lambda) < \infty$. By Theorem 4.4.3, S and Y have the same distribution.

4.10: Notice that $X \ge x\mathbb{I}(X \ge x)$ $\forall x > 0$, so $0 = \mathbb{E}(X) \ge \mathbb{E}(x\mathbb{I}(X \ge x)) = x\mathbb{E}(\mathbb{I}(X \ge x)) = x\mathbb{P}(X \ge x)$ by the properties of expectation, implying $\mathbb{P}(X \ge x) = 0$. Then, $\mathbb{P}(X > 0) = \mathbb{P}(\cup_{n=1}^\infty \{X \ge \frac{1}{n}\}) \le \Sigma_{n=1}^\infty \mathbb{P}(X \ge \frac{1}{n}) = 0$, so $\mathbb{P}(X = 0) = \mathbb{P}(X \ge 0) - \mathbb{P}(X > 0) = 1$. Alternatively, suppose that $\mathbb{P}(X > 0) > 0 \iff \exists\varepsilon > 0, \delta > 0$ s.t. $\mathbb{P}(X \ge \delta) = \varepsilon$. Notice that $\mathbb{E}(X) = \mathbb{E}(X\mathbb{I}(X < \delta)) + \mathbb{E}(X\mathbb{I}(X \ge \delta)) \ge \mathbb{E}(X\mathbb{I}(X \ge \delta)) \ge \mathbb{E}(\delta\mathbb{I}(X \ge \delta)) = \delta\mathbb{P}(X \ge \delta) = \delta\varepsilon > 0$, a contradiction.

4.14: Define $Y = \Sigma_{i=1}^N X_i$. By the law of total expectation, $\mathbb{E}(Y) = \mathbb{E}(\mathbb{E}(Y|N)) = \Sigma_{n=1}^\infty \mathbb{E}(Y|N = n)\mathbb{P}(N = n) = \Sigma_{n=1}^\infty (\Sigma_{i=1}^n \mathbb{E}(X_i|N = n))\mathbb{P}(N = n)$ since the inner sum is finite $= \Sigma_{n=1}^\infty (\Sigma_{i=1}^n \mathbb{E}(X_i))\mathbb{P}(N = n)$ by independence $= \Sigma_{n=1}^\infty n\mu\mathbb{P}(N = n) = \mu\Sigma_{n=1}^\infty n\mathbb{P}(N = n) = \mu\mathbb{E}(N) = \mu m$. Notice that $\mathbb{E}_N(\mathrm{Var}(Y|N)) = \mathbb{E}_N(\Sigma_{i=1}^N \mathrm{Var}(X_i))$ by independence $= \mathbb{E}_N(N\mathrm{Var}(X_i)) = \mathrm{Var}(X_i)\mathbb{E}_N(N) = \sigma^2 m$ and $\mathrm{Var}_N(\mathbb{E}(Y|N)) = \mathrm{Var}_N(\Sigma_{i=1}^N \mathbb{E}(X_i|N))$ since the inner sum is finite $= \mathrm{Var}_N(\Sigma_{i=1}^N \mathbb{E}(X_i))$ by independence $= \mathrm{Var}_N(N\mathbb{E}(X_i)) = \mathbb{E}(X_i)^2\mathrm{Var}(N) = \mu^2 v$. The result follows by the law of total variance.

4.18: Define $Y_t = t\mathbb{I}(X > t)$ $\forall t > 0$. Notice that $\mathbb{I}(X > t) = 0$ for large enough $t \implies \mathbb{P}(\lim_{t\to\infty} t\mathbb{I}(X > t) = 0 < X) = 1$. Thus, $Y_n \to 0$ a.s. and $|Y_n| \leq X$ a.s., so by the DCT, $0 = \lim_{t\to\infty} \mathbb{E}(Y_t) = \lim_{t\to\infty} t\mathbb{P}(X > t)$.