Solutions for Problem Set One

**Problem 1.** (1) (i) We have
\[ \mathbb{E}[X \mathbb{E}[Y | \mathcal{G}]] = \mathbb{E}[\mathbb{E}[X \mathbb{E}[Y | \mathcal{G}]] = \mathbb{E}[\mathbb{E}[X | \mathcal{G}] \cdot \mathbb{E}[Y | \mathcal{G}]]. \]
Similarly for \( \mathbb{E}[Y \mathbb{E}[X | \mathcal{G}]] \).

(ii) We say that a bounded measurable function satisfying property \( P \) if
\[ \mathbb{E}[f(X, Y) | \mathcal{G}] = \mathbb{E}[f(x, Y)]_{x=X}. \]
Let \( \mathcal{E} = \{ E \in \mathcal{B}(\mathbb{R}^2) : \mathbf{1}_E \text{ satisfies property } P \} \). Then \( \mathcal{E} \) is a monotone class containing the \( \pi \)-system \( \mathcal{C} \triangleq \{ A \times B : A, B \in \mathcal{B}(\mathbb{R}^1) \} \). By the monotone class theorem in measure theory, we conclude that \( \mathcal{E} = \mathcal{B}(\mathbb{R}^2) \). In other words, \( \mathbf{1}_E \) satisfies property \( P \) for every \( E \in \mathcal{B}(\mathbb{R}^2) \).

Note that the property \( P \) is linear in \( f \). By writing \( f = f^+ - f^- \), we only need to consider the case when \( f \) is bounded and non-negative. But then there exists a sequence \( f_n \) of simple functions on \( \mathbb{R}^2 \) such that \( 0 \leq f_n \uparrow f \). We know that each \( f_n \) satisfies property \( P \). By the monotone convergence theorem for both conditional and unconditional expectations, we conclude that \( f \) satisfies property \( P \).

(iii) Since both sides are \( \sigma(\mathcal{G}, \mathcal{H}) \)-measurable, it suffices to show that
\[ \int_E X d\mathbb{P} = \int_E \mathbb{E}[X | \mathcal{G}] d\mathbb{P}, \forall E \in \sigma(\mathcal{G}, \mathcal{H}). \quad (1) \]
Let \( \mathcal{E} = \{ E \in \sigma(\mathcal{G}, \mathcal{H}) : \text{ equation (1) holds} \} \), and let \( \mathcal{C} = \{ A \cap B : A \in \mathcal{G}, B \in \mathcal{H} \} \).

Apparentvly, \( \mathcal{C} \) is a \( \pi \)-system. For any \( A \in \mathcal{G}, B \in \mathcal{H} \), we have
\[ \mathbb{E}[X \mathbf{1}_A \mathbf{1}_B] = \mathbb{E}[X \mathbf{1}_A] \mathbb{P}(B) = \mathbb{E}[\mathbb{E}[X | \mathcal{G}] \mathbf{1}_A] \mathbb{P}(B) = \mathbb{E}[\mathbb{E}[X | \mathcal{G}] \mathbf{1}_A \mathbf{1}_B]. \]
Therefore, \( \mathcal{C} \subseteq \mathcal{E} \). Moreover, it is easy to see that \( \mathcal{E} \) is a monotone class. By the monotone class theorem, we conclude that \( \sigma(\mathcal{G}, \mathcal{H}) = \mathcal{E} \).

(2) By assumption, we know that for every \( r \in \mathbb{R}^1 \),
\[ \mathbb{E} [(X - Y) \mathbf{1}_{\{X \leq r\}}] = \mathbb{E} [(X - Y) \mathbf{1}_{\{Y \leq r\}}] = 0. \]
Therefore,
\[
\mathbb{E} [(X - Y) 1_{\{X \leq r, Y > r\}}] + \mathbb{E} [(X - Y) 1_{\{X \leq r, Y \leq r\}}] = 0,
\]
\[
\mathbb{E} [(X - Y) 1_{\{X > r, Y \leq r\}}] + \mathbb{E} [(X - Y) 1_{\{X \leq r, Y \leq r\}}] = 0.
\]
It follows that
\[
\mathbb{E} [(X - Y) 1_{\{X > r, Y \leq r\}}] + \mathbb{E} [(Y - X) 1_{\{X \leq r, Y > r\}}] = 0.
\]
But the integrand inside each of the above expectations is non-negative. Therefore,
\[
(X - Y) 1_{\{X > r, Y \leq r\}} = (Y - X) 1_{\{X \leq r, Y > r\}} = 0 \text{ a.s.}
\]
This implies that
\[
\mathbb{P}(X > r, Y \leq r) = \mathbb{P}(X \leq r, Y > r) = 0.
\]
And this is true for all \( r \in \mathbb{R}^1 \). The result then follows from the fact that
\[
\{X \neq Y\} \subseteq \{X > Y\} \bigcup \{X < Y\} \subseteq \bigcup_{n \in \mathbb{Z}} \left( \{X > n \geq Y\} \bigcup \{Y > n \geq X\} \right).
\]

**Problem 2.** (1) For \( \lambda > 0 \), we have
\[
|\mathbb{E}[X|\mathcal{G}_i]| 1_{\{|\mathbb{E}[X|\mathcal{G}_i]| > \lambda\}} \leq \mathbb{E}[|X|]\mathbb{1}_{\{|\mathbb{E}[X|\mathcal{G}_i]| > \lambda\}}.
\]
Therefore, by taking expectations on both sides, we obtain that
\[
\mathbb{E} \left[ |\mathbb{E}[X|\mathcal{G}_i]| 1_{\{|\mathbb{E}[X|\mathcal{G}_i]| > \lambda\}} \right] \leq \mathbb{E}[|X|] 1_{\{|\mathbb{E}[X|\mathcal{G}_i]| > \lambda\}}.
\]
But
\[
\mathbb{E}[|X|] 1_{\{|\mathbb{E}[X|\mathcal{G}_i]| > \lambda\}} = \mathbb{E}[|X|] 1_{\{|\mathbb{E}[X|\mathcal{G}_i]| > \lambda\}}; |X| > \sqrt{\lambda} + \mathbb{E}[|X|] 1_{\{|\mathbb{E}[X|\mathcal{G}_i]| > \lambda\}}; |X| \leq \sqrt{\lambda}
\]
\[
\leq \mathbb{E}|X|; |X| > \sqrt{\lambda} + \sqrt{\lambda} \cdot \frac{1}{\lambda} \mathbb{E}[|X|]|\mathcal{G}_i]
\]
\[
= \mathbb{E}|X|; |X| > \sqrt{\lambda} + \frac{1}{\sqrt{\lambda}} \mathbb{E}[|X|],
\]
which goes to zero uniformly in \( i \in \mathcal{I} \) as \( \lambda \to \infty \) since \( X \) is integrable. Therefore, \( \{\mathbb{E}[X|\mathcal{G}_i]; i \in \mathcal{I}\} \) is uniformly integrable.

(2) Let \( M = \sup_{t \in \mathcal{T}} \mathbb{E}[\varphi(|X_t|)] \). For \( \varepsilon > 0 \), let \( R = M/\varepsilon \). Then there exists some \( \Lambda > 0 \), such that for any \( x > \Lambda \), we have \( \varphi(x)/x > R \). Therefore, for \( \lambda > \Lambda \), we have
\[
\mathbb{E}[|X_t| 1_{\{|X_t| > \lambda\}}] \leq \frac{1}{R} \mathbb{E}[\varphi(|X_t|)] \leq \frac{M}{R} = \varepsilon, \, \forall t \in \mathcal{T}.
\]
Consequently, \( \{X_t; t \in \mathcal{T}\} \) is uniformly integrable.
Problem 3. (1) $\mathbb{P}(X_n > \alpha \log n) = e^{-\alpha \log n} = 1/n^\alpha$. Therefore, by the Borel-Cantelli lemma, we have
\[
\mathbb{P}(X_n > \alpha \log n \text{ for infinitely many } n) = \begin{cases} 
0, & \alpha > 1; \\
1, & 0 < \alpha \leq 1. 
\end{cases}
\]
Therefore, by the Borel-Cantelli lemma, we have
\[
\mathbb{P}(X_n > \alpha \log n \text{ for infinitely many } n) = \begin{cases} 
0, & \alpha > 1; \\
1, & 0 < \alpha \leq 1. 
\end{cases}
\]

(2) Let $A_\alpha = \{X_n > \alpha \log n \text{ for infinitely many } n\}$. Since $\mathbb{P}(A_1) = 1$, we know that $L \geq 1$ almost surely. Moreover,
\[
\{L > 1\} \subseteq \bigcup_{k=1}^{\infty} \left\{ L > 1 + \frac{1}{k} \right\} \subseteq \bigcup_{k=1}^{\infty} A_{1+\frac{1}{k}}.
\]
It follows that $\mathbb{P}(L > 1) = 0$. Therefore, $L = 1$ almost surely.

(3) For each $x \in \mathbb{R}^1$, we have
\[
\mathbb{P}(M_n \leq x) = \mathbb{P}\left( \max_{1 \leq i \leq n} X_i \leq x + \log n \right) = (1 - e^{-x - \log n})^n,
\]
provided that $x + \log n > 0$. Therefore,
\[
\lim_{n \to \infty} \mathbb{P}(M_n \leq x) = e^{-e^{-x}}, \ \forall x \in \mathbb{R}^1.
\]
Apparently, the function $F(x) \triangleq e^{-e^{-x}}$ defines a continuous distribution function on $\mathbb{R}^1$. Therefore, $M_n$ converges weakly to $F$.

Problem 4. (1) $\implies$ (2). Suppose that $\mathbb{P}_n$ converges weakly to $\mathbb{P}$. According to Theorem 1.7, we know that $\mathbb{P}_n(A) \to \mathbb{P}(A)$ for every $A \in \mathcal{B}(\mathbb{R}^1)$ satisfying $\mathbb{P}(\partial A) = 0$. In particular, let $x$ be a continuity point of $F$ and let $A = (-\infty, x]$. Then $\mathbb{P}(\partial A) = dF(\{x\}) = 0$. Therefore,
\[
F_n(x) = \mathbb{P}_n(A) \to \mathbb{P}(A) = F(x).
\]

(2) $\implies$ (1). Suppose that $F_n$ converges in distribution to $F$. Let $C_F$ be the set of continuity points of $F$. Since $C_F^c$ is at most countable, we conclude that $C_F$ is dense in $\mathbb{R}^1$.

Let $\varphi \in C_b(\mathbb{R}^1)$. Given $\varepsilon > 0$, let $a, b \in C_F$ be such that $a < 0 < b$ and
\[
F(a) < \varepsilon, \ 1 - F(b) < \varepsilon.
\]
Then there exists $N \geq 1$, such that for any $n > N$,
\[
|F_n(a) - F(a)| < \varepsilon, |F_n(b) - F(b)| < \varepsilon.
\]

3
It follows that

\[ F_n(a) < 2\varepsilon, \quad 1 - F_n(b) < 2\varepsilon, \quad \forall n > N. \]

Therefore,

\[
\left| \int_{\mathbb{R}} \varphi \left( dF_n - dF \right) \right| \\
\leq \left| \int_{(a,b]} \varphi \left( dF_n - dF \right) \right| + \| \varphi \|_{\infty} (dF_n((a,b]) + dF((a,b])) \\
\leq \left| \int_{(a,b]} \varphi \left( dF_n - dF \right) \right| + 6\| \varphi \|_{\infty} \varepsilon \tag{2}
\]

for every \( n > N. \)

Since \( \varphi \) is uniformly continuous on \([a,b]\), there exists \( \delta > 0 \), such that whenever \( x, y \in [a,b] \) with \( |x - y| < \delta \), we have \( |\varphi(x) - \varphi(y)| < \varepsilon \). Choose a finite partition \( \mathcal{P} : a = x_0 < x_1 < \cdots < x_k = b \) of \([a,b]\), such that \( x_0, x_1, \cdots, x_k \in C_F \) and \( |x_i - x_{i-1}| < \delta \) for each \( i \). Define a step function \( \psi \) by taking \( \psi(x) = \varphi(x_i-1) \) for \( x \in [x_{i-1}, x_i] \). It follows that

\[
\sup_{x \in [a,b]} |\varphi(x) - \psi(x)| \leq \varepsilon.
\]

Therefore,

\[
\left| \int_{[a,b]} \varphi \left( dF_n - dF \right) \right| \\
\leq 2 \sup_{x \in [a,b]} |\varphi(x) - \psi(x)| + \left| \int_{[a,b]} \psi \left( dF_n - dF \right) \right| \\
\leq 2\varepsilon + \sum_i |\varphi(x_{i-1})| \cdot ((F_n(x_i) - F(x_i)) - (F_n(x_{i-1}) - F(x_{i-1}))). \tag{3}
\]

Note that the partition \( \mathcal{P} \) we chose before does not depend on \( n \).

By substituting (3) into (2) and letting \( n \to \infty \), we arrive at

\[
\limsup_{n \to \infty} \left| \int_{\mathbb{R}} \varphi dF_n - \int_{\mathbb{R}} \varphi dF \right| \leq (2 + 6\| \varphi \|_{\infty})\varepsilon.
\]

Since \( \varepsilon \) is arbitrary, we conclude that \( \int_{\mathbb{R}} \varphi dF_n \to \int_{\mathbb{R}} \varphi dF \) as \( n \to \infty \). Therefore, \( \mathbb{P}_n \) converges weakly to \( \mathbb{P} \).
Problem 5. (1) Necessity. Suppose that \( \{\mathbb{P}_n\} \) is tight. Then there exists \( M > 0 \), such that
\[
\mathbb{P}_n([-M, M]) \geq \frac{3}{4}, \quad \forall n \geq 1.
\]
It follows that \( |\mu_n| \leq M \) for all \( n \). Indeed, if this is not the case, suppose for instance that \( \mu_n > M \) for some \( n \). Then
\[
\frac{1}{2} \leq \mathbb{P}_n(\mu_n, \infty) \leq \mathbb{P}_n((M, \infty)) < \frac{1}{4},
\]
which is a contradiction. In addition, we have
\[
\frac{3}{4} \leq \mathbb{P}_n([-M, M]) = \frac{1}{\sqrt{2\pi} \sigma_n} \int_{-M}^{M} e^{-\frac{(x-\mu_n)^2}{2\sigma_n^2}} dx
\]
\[
= \frac{1}{\sqrt{2\pi} \sigma_n} \int_{-M-\mu_n}^{M-\mu_n} e^{-\frac{x^2}{2}} dx \leq \frac{1}{\sqrt{2\pi} \sigma_n} \int_{-2M}^{2M} e^{-\frac{x^2}{2}} dx.
\]
This implies that \( \sigma_n \) is bounded. Indeed, if \( \sigma_n \uparrow \infty \) along a subsequence, then the right hand side of (4) goes to zero along this subsequence, which is a contradiction.

Sufficiency. Suppose that \( |\mu_n| \leq M_1, \sigma_n \leq M_1 \) for some \( M_1 > 0 \). Then for any \( M > M_1 \), we have
\[
\mathbb{P}_n([-M, M]) = \frac{1}{\sqrt{2\pi} \sigma_n} \int_{-M-\mu_n}^{M-\mu_n} e^{-\frac{x^2}{2}} dx
\]
\[
\geq \frac{1}{\sqrt{2\pi} \sigma_n} \int_{-M_1-M_1}^{M_1-M_1} e^{-\frac{x^2}{2}} dx
\]
\[
\geq \frac{1}{\sqrt{2\pi} \sigma_n} \int_{-M_1}^{M_1} e^{-\frac{x^2}{2}} dx.
\]
Since the right hand side of (5) converges to 1 as \( M \to \infty \), we conclude that
\[
\lim_{M \to \infty} \inf_{n \geq 1} \mathbb{P}_n([-M, M]) = 1.
\]
In other words, \( \{\mathbb{P}_n\} \) is tight.

(2) Sufficiency. Suppose that \( \mu_n \to \mu \) and \( \sigma_n^2 \to \sigma^2 \). Then
\[
e^{\mu_n t - \frac{1}{2} \sigma_n^2 t} \to e^{\mu t - \frac{1}{2} \sigma^2 t}
\]
for every $t \in \mathbb{R}^1$ as $n \to \infty$. Therefore, $\mathbb{P}_n$ converges weakly to $\mathcal{N}(\mu, \sigma^2)$.

Necessity. Suppose that $\{\mathbb{P}_n\}$ is weakly convergent. From the first part we already know that $\{\mu_n\}$ and $\{\sigma^2_n\}$ are both bounded. Assume that $\mu$ and $\mu'$ are two limit points of $\mu_n$. We may further assume without loss of generality that $\mu_{n_k} \to \mu, \sigma^2_{n_k} \to \sigma^2$, and $\mu_{n'_l} \to \mu', \sigma^2_{n'_l} \to \sigma'^2$ along two subsequences $n_k$ and $n'_l$. By the sufficiency part and the uniqueness of weak limits, we know that $\mathcal{N}(\mu, \sigma^2) = \mathcal{N}(\mu', \sigma'^2)$, and hence $\mu = \mu'$ and $\sigma^2 = \sigma'^2$. Therefore, $\mu_n$ converges to some $\mu \in \mathbb{R}^1$. Similarly, we conclude that $\sigma^2_n$ has exactly one limit point, which means that it converges to some $\sigma^2 \geq 0$. 