Solutions to Problem Set 1

2.4.1 Let \( F(s,t) \equiv P1\{X \leq s, Y \leq t\} \), \( F_1(s) = F(s,\infty) \), and \( F_2(t) = F(\infty,t) \). For each \( \eta > 0 \), choose \(-\infty = s_1 < s_2 < \cdots < s_{k_1} = \infty \) and \(-\infty = t_1 < t_2 < \cdots < t_{k_2} = \infty\) such that \( F_1(s_j) - F_1(s_{j-1}) < \eta \), for all \( 1 < j \leq k_1 \), and \( F_2(t_l) - F_2(t_{l-1}) < \eta \), for all \( 1 < l \leq k_2 \). This can be done so that both \( k_1 \) and \( k_2 \) are \( 2 + 1/\eta \). Consider the set of brackets of the form \((l,u)\), where \( l(X,Y) = 1\{X \leq s_{j-1}, Y \leq t_{l-1}\} \) and \( u(X,Y) = 1\{X < s_j, Y < t_l\} \), for \( 1 < j \leq k_1 \) and \( 1 < l \leq k_2 \). The number of such brackets is bounded by \((1+1/\eta)^2\) by construction. Note that for these brackets, \( \|u - l\|_{P,r} < \eta^{1/r} \), for all \( 1 \leq r < \infty \). Hence, for \( \mathcal{F} = \{1\{X \leq s, Y \leq t\} : s,t \in \mathbb{R}\} \), we have \( N||\mathcal{F}, L_1(P) \leq (1+1/\epsilon)^2 \) and \( N||\mathcal{F}, L_2(P) \leq (1+1/\epsilon^2)\).

3.5.1 We need to make one more assumption that
\[
\int \hat{\ell}_\theta \hat{p}_\theta d\mu \to \int \hat{\ell}_\theta \hat{p}_\theta d\mu, \quad \text{as } \|\hat{\theta} - \theta\| \to 0. \tag{1}
\]

We begin the proof by first showing that
\[
H(\hat{\theta}, \theta) \equiv \int \left(p_\theta^{1/2} - p_{\hat{\theta}}^{1/2}\right)^2 d\mu \to 0, \quad \text{as } \|\hat{\theta} - \theta\| \to 0. \tag{2}
\]

By the mean value theorem, we have for some \( \hat{\theta}^* \) on the line segment between \( \theta \) and \( \hat{\theta} \) that
\[
H(\hat{\theta}, \theta) = \int \left(\hat{\ell}_{\hat{\theta}} - (\hat{\theta} - \theta)\hat{p}_{\hat{\theta}}^{1/2}\right)^2 d\mu \\
\leq \|\hat{\theta} - \theta\|^2 \int \|\hat{\ell}_{\hat{\theta}}\|^2 p_{\hat{\theta}} d\mu \\
\to 0,
\]
as \( \|\hat{\theta} - \theta\| \to 0 \) by (1). Hence (2) follows.

Next, we show that (2) implies
\[
\int |p_{\hat{\theta}} - p_\theta|d\mu \to 0, \quad \text{as } \|\hat{\theta} - \theta\| \to 0. \tag{3}
\]
This follows immediately from the fact that
\[ \int |p_\theta - p_\theta| d\mu = \int \left| p_\theta^{1/2} - p_\theta^{1/2} \right| \left( p_\theta^{1/2} + p_\theta^{1/2} \right) d\mu \]
\[ \leq \sqrt{\int \left( p_\theta^{1/2} - p_\theta^{1/2} \right)^2} \cdot \sqrt{\int \left( p_\theta^{1/2} + p_\theta^{1/2} \right)^2} d\mu \]
\[ \leq 2 \left[ H(\tilde{\theta}, \theta) \right]^{1/2}. \]
Now, fix \( h \in \mathbb{R}^k \), let the scalar sequence \( t \to 0 \), and set \( dP = p_\theta d\mu \) and \( dP_t = p_{\theta+th} d\mu \). We have by the mean value theorem that for some \( \tilde{\theta} = \theta + th \), the left-hand-side of expression (3.1) of page 37 of Kosorok (2008) is equal to
\[ \int \left( h' \dot{\ell}_{\theta} p_\theta^{1/2} - h' \dot{\ell}_{\theta} p_\theta^{1/2} \right)^2 d\mu = \int J_1 J_2 \left( h' \dot{\ell}_{\theta} p_\theta^{1/2} - h' \dot{\ell}_{\theta} p_\theta^{1/2} \right)^2 d\mu \]
\[ + \int (1 - J_1 J_2) \left( h' \dot{\ell}_{\theta} p_\theta^{1/2} - h' \dot{\ell}_{\theta} p_\theta^{1/2} \right)^2 d\mu \]
\[ \equiv A_t + B_t, \]
where \( J_1 = 1 \{ ||\dot{\ell}_{\theta}|| \leq k_1 \}, J_2 = 1 \{ ||\dot{\ell}_{\theta}|| \leq k_2 \}, \) and \( 0 < k_1, k_2 < \infty \) are scalars (to be chosen later).

Clearly,
\[ A_t \leq \int J_1 J_2 \left[ h' (\dot{\ell}_{\theta} - \dot{\ell}_{\theta}) \right]^2 p_\theta d\mu + \int J_1 J_2 \left( h' \dot{\ell}_{\theta} \right)^2 \left( p_\theta^{1/2} - p_\theta^{1/2} \right)^2 d\mu \]
\[ \leq ||h||^2 \int ||\dot{\ell}_{\theta} - \dot{\ell}_{\theta}||^2 p_\theta d\mu + ||h||^2 k_2^2 H(\tilde{\theta}, \theta) \]
\[ \to 0, \]
where, in the second-to-last line, the first term goes to zero by assumption and the second term goes to zero by (2).

To facilitate proof that \( B_t \to 0 \), we first argue that
\[ \int (1 - J_1 J_2) \dot{\ell}_{\theta} \dot{\ell}_{\theta} p_\theta d\mu = \int (1 - J_1 J_2) \dot{\ell}_{\theta} \dot{\ell}_{\theta} p_\theta d\mu + o(1). \]
By applying (3) followed by bounded convergence,
\[
\int J_1 J_2 \dot{\ell}_\theta \dot{\ell}_\theta p_\theta d\mu = \int J_1 J_2 \dot{\ell}_\theta \dot{\ell}_\theta p_\theta d\mu + o(1)
\]
\[
= \int J_1 J_2 \dot{\ell}_\theta \dot{\ell}_\theta p_\theta d\mu + o(1).
\]
Combining this with assumption (1), (4) follows. Hence
\[
B_t \leq 4 \int (1 - J_1 J_2) (h' \dot{\ell}_\theta)^2 p_\theta d\mu + o(1)
\]
\[
\leq 4 \int (1 - J_1) (h' \dot{\ell}_\theta)^2 p_\theta d\mu + 4 \int J_1 (1 - J_2) (h' \dot{\ell}_\theta)^2 p_\theta d\mu + o(1)
\]
\[
\leq 4 \|h\|^2 \int 1\{\|\dot{\ell}_\theta\| > k_1\} \|\dot{\ell}_\theta\|^2 p_\theta d\mu + 4 \|h\|^2 k_1^2 P(\|\dot{\ell}_\theta\| > k_2) + o(1)
\]
\[
\equiv 4 \|h\|^2 C_1(k_1) + 4 \|h\|^2 k_1^2 C_2(k_2) + o(1).
\]
Since $k_1$ and $k_2$ are arbitrary, we can, for any $\epsilon > 0$, first choose $k_1 \geq 1$ large enough so that $C_1(k_1) \leq \epsilon/(8\|h\|^2)$ and then pick $k_2$ large enough so that $C_2(k_2) \leq \epsilon/(8\|h\|^2 k_1^2)$, and thus $B_t \leq \epsilon + o(1)$. Hence $B_t \to 0$ since $\epsilon$ was also arbitrary. Thus $A_t + B_t \to 0$, and the desired result follows. □