Statistics 521, Problem Set 4 Solution

Wellner; 10/25/2019

Reminder: Make-up lecture: Friday, 25 October, 8:30 - 9:20, Low 101.
Reminder: Midterm exam: Friday, November 1.

1. PfS, Exercise 2.3.4, page 32: (a) Suppose that $\mu(\Omega) < \infty$ and g is continuous a.e. μ_X (that is, g is continuous except perhaps on a set of μ_X measure 0). Then $X_n \rightarrow_\mu X$ implies that $g(X_n) \rightarrow_\mu g(X)$.
(b) Let g be uniformly continuous on the real line. Then $X_n \rightarrow_\mu X$ implies that $g(X_n) \rightarrow_\mu g(X)$. (Here $\mu(\Omega) = \infty$ is allowed.)

Solution: (a) Let $\{n'\}$ be a subsequence. We want to show that for some subsequence $\{n''\}$ it follows that $g(X_{n''}) \rightarrow_{a.e.} g(X)$. Then by (15) of Theorem 2.3.1 it follows that $g(X_n) \rightarrow_\mu g(X)$. But since $X_n \rightarrow_\mu X$ we know, by (15) of Theorem 2.3.1, that there is a further subsequence $\{n''\}$ such that $X_{n''} \rightarrow_{a.e.} X$. For this subsequence we have $g(X_{n''}) \rightarrow_{a.e.} g(X)$ (by restricting in addition to the set $[X \in C_g]$ with $\mu([X \in C_g^c]) = \mu_X(C_g^c) = 0$ for which g is continuous). Thus we conclude that $g(X_n) \rightarrow_\mu g(X)$.
(b) Let $\epsilon > 0$. Since g is uniformly continuous there is a $\delta = \delta_\epsilon$ such that $|y - x| < \delta_\epsilon$ implies $|g(y) - g(x)| < \epsilon$. Since $X_n \rightarrow_\mu X$, for every $\gamma > 0$ there exists an $N = N_{\epsilon,\gamma}$ such that

$$\mu([|X_n - X| \geq \delta_\epsilon]) < \gamma, \qquad \text{for all } n \geq N_{\epsilon,\gamma}.$$

Then we have

$$\begin{aligned}
&\mu([|g(X_n) - g(X)| \geq \epsilon]) \\
&\quad = \mu([|g(X_n) - g(X)| \geq \epsilon] \cap [|X_n - X| \geq \delta_\epsilon]) \\
&\qquad\qquad + \mu([|g(X_n) - g(X)| \geq \epsilon] \cap [|X_n - X| < \delta_\epsilon]) \\
&\quad \leq \mu([|X_n - X| \geq \delta_\epsilon]) + \mu(\emptyset) \\
&\quad \leq \gamma + 0 = \gamma \qquad \text{for } n \geq N_{\epsilon,\gamma}.
\end{aligned}$$

Thus $\mu([|g(X_n) - g(X)| \geq \epsilon]) \rightarrow 0$ as $n \rightarrow \infty$; i.e. $g(X_n) \rightarrow_\mu g(X)$.

2. PfS, Exercise 3.2.1, page 42: Show that $X \geq 0$ and $\int X d\mu = 0$ implies $\mu([X > 0]) = 0$.

   Solution: Let $\epsilon > 0$. Then $X \geq \epsilon 1_{[X \geq \epsilon]}$, and hence

$$0 = \int X d\mu \geq \epsilon \mu([X \geq \epsilon])$$

   Since $[X > 0] = \cup_{n=1}^{\infty} [X \geq 1/n] = \lim [X \geq 1/n]$, we find that $\mu([X > 0]) = \lim_n \mu([X \geq 1/n]) = \lim_n 0 = 0$.

3. PfS, Exercise 3.2.2, page 42: Show that

$$\int_A X d\mu = \begin{cases} = 0, \\ \geq 0, \end{cases} \qquad \text{for all } A \in \mathcal{A} \text{ implies } X = \begin{cases} = 0 & \text{a.e.,} \\ \geq 0 & \text{a.e.} \end{cases}$$

   Solution: Suppose first that $\int_A X d\mu = 0$ for all $A \in \mathcal{A}$. Then with $A = [X^+ \geq 0]$ we have $0 = \int X 1_{[X^+ \geq 0]} d\mu = \int Y d\mu$ where $Y \equiv X 1_{[X^+ \geq 0]} = X^+ \geq 0$. Then by the previous exercise, $0 = \mu([Y > 0]) = \mu([X^+ > 0])$; i.e $X^+ = 0$ a.e. Similarly, choosing $A = [X^- \geq 0]$ yields $X^- = 0$ a.e.; combining the two results gives $X = X^+ - X^- = 0$ a.e.
   Now suppose that $\int_A X d\mu \geq 0$ for all $A \in \mathcal{A}$. Taking $A = [X < 0] = [X^- > 0]$ yields $0 \leq \int_A X d\mu = \int -X^- d\mu \leq 0$ since $X^- \geq 0$. Thus $\int X^- d\mu = 0$. By problem 2 this implies $\mu([X < 0]) = \mu([X^- > 0]) = 0$. Hence $X \geq 0$ a.e. μ.

4. PfS, Exercise 3.2.4, page 43. Let $Y \equiv g(X)$ in the context of Theorem 3.2.6 (the "Theorem of the unconscious statistician"). Show that the second equality holds in:

$$\int_{X^{-1}(g^{-1}(B))} g(X(\omega)) d\mu(\omega) = \int_{g^{-1}(B)} g(x) d\mu_X(x) = \int_B y d\mu_Y(y) \quad \text{for } B \in \overline{\mathcal{B}}$$

   where μ_Y is the induced measure of Y on $(\overline{R}, \overline{\mathcal{B}})$.

   Solution:
   First proof: This can be viewed as the identity proved in the first equality with an appropriate identification of terms. Thus $(\Omega, \mathcal{A})$ is replaced by $(\Omega', \mathcal{A}')$, $(\Omega', \mathcal{A}')$ is replaced by $(\overline{R}, \overline{\mathcal{B}})$, g is replaced by the

identity function $h : (\Omega', \mathcal{A}') \mapsto (\overline{R}, \overline{\mathcal{B}})$ given by $h(v) = v$, and X is replaced by $g : (\Omega', \mathcal{A}') \mapsto (\overline{R}, \overline{\mathcal{B}})$, and μ and μ_X are replaced by μ_X and μ_Y respectively. Thus we think of the identity

$$\int g(X(\omega))d\mu(\omega) = \int g(x)d\mu_X(x)$$

as being replaced by

$$\int h(g(x))d\mu_X(x) = \int h(y)d\mu_{g(X)}(y) = \int h(y)d\mu_Y(y) \tag{1}$$

or, when $h(v) = v$, this yields

$$\int g(x)d\mu_X(x) = \int yd\mu_Y(y)$$

where $Y \equiv g(X)$. With this set of identifications the the slightly more general identity (1) follows from the first equality, and the desired second equality is given by the special case $h(v) = v$.

Second proof: (via the "standard machine"). With the inclusion of an additional function h as indicated in the first proof above, we want to show that

$$\int h(g(x))d\mu_X(x) = \int h(y)d\mu_Y(y) \tag{2}$$

for measurable functions h and g from $(\mathbb{R}, \mathcal{B})$ to $(\mathbb{R}, \mathcal{B})$. Then the claimed identity is just the special case $h(y) = y$.

Case 1: (2) holds when $h = 1_B$ for a Borel set B:

$$\begin{aligned} \int 1_B(g(x))d\mu_X(x) = \int 1_{g^{-1}(B)}d\mu_X &= \mu_X(g^{-1}(B)) \\ &= \mu_Y(B) = \int 1_B(y)d\mu_Y(y). \end{aligned}$$

Case 2: (2) holds when $h(y) = \sum_{i=1}^n c_i 1_{B_i}(y)$ where $\sum_1^n B_i = \mathbb{R}$ with

$B_i \in \mathcal{B}$ and all $c_i \geq 0$. To see this, note that

$$\begin{aligned}
\int h(g(x))d\mu_X(x) &= \int \sum_1^n c_i 1_{B_i}(g(x))d\mu_X(x) \\
&= \sum_1^n c_i \int 1_{B_i}(g(x))d\mu_X(x) \text{ by linearity of the integral} \\
&= \sum_1^n c_i \int 1_{B_i}(y)d\mu_Y(y) \text{ by case 1} \\
&= \int \sum_1^n c_i 1_{B_i}(y)d\mu(y) = \int h(y)d\mu_Y(y) \text{ by linearity again.}
\end{aligned}$$

Case 3: $h \geq 0$. Let $h_n \geq 0$ be simple functions with $h_n \nearrow h$. Then

$$\begin{aligned}
\int h(g(x))d\mu_X(x) &= \lim_n \int h_n(g(x))d\mu_X(x) \text{ by the MCT theorem} \\
&= \lim_n \int h_n(y)d\mu_Y(y) \text{ by case 2} \\
&= \int h(y)d\mu_Y(y) \text{ by the MCT again.}
\end{aligned}$$

Case 4: h is measurable and either $\int h(g(x))^+ d\mu_X(x) < \infty$ or $\int h(g(x))^- d\mu_X(x) < \infty$. Since $h = h^+ - h^-$ we note that

$$h(g(x))^+ = h^+(g(x)) \quad \text{and} \quad h(g(x))^- = h^-(g(x)).$$

Then

$$\begin{aligned}
\int h(g(x))d\mu_X(x) &= \int h(g(x))^+ d\mu_X(x) - \int h(g(x))^- d\mu_X(x) \\
&= \int h^+(g(x))d\mu_X(x) - \int h^-(g(x))d\mu_X(x) \\
&= \int h^+(y)d\mu_Y(y) - \int h^-(y)d\mu_Y(y) \text{ by case 3} \\
&= \int h(y)d\mu_Y(y) \text{ by linearity and } h = h^+ - h^-.
\end{aligned}$$

This completes the proof of the identity (2).

5. (i) PfS, Exercise 3.3, page 45, part (a).
(ii) Suppose that μ is Lebesgue measure on the unit interval $[0,1]$ and that $(a,b) = (0,1)$ in Exercise 3.3. If $X(t,\omega) = 1_{[\omega\le t]}$, then for each t, $(\partial/\partial t)X(t,\omega) = 0$ almost everywhere. But $\int X(t,\omega)d\mu(\omega)$ does not differentiate to 0. Why is this not a contradiction?

Solution: (i) We write $h(t) \equiv \int_\Omega X(t,\omega)d\mu(\omega$. Then by linearity of the integral and the fundamental theorem of calculus it follows that

$$\begin{aligned} h(t+\epsilon) - h(t) &= \int_\Omega \left(X(t+\epsilon,\omega) - X(t,\omega)\right) d\mu(\omega) \\ &= \int_\Omega \left(\int_t^{t+\epsilon} \frac{\partial}{\partial s} X(s,\omega)ds\right) d\mu(\omega), \end{aligned}$$

and hence that

$$\frac{h(t+\epsilon) - h(t)}{\epsilon} = \int_\Omega \left(\frac{1}{\epsilon}\int_t^{t+\epsilon} \frac{\partial}{\partial s} X(s,\omega)ds\right) d\mu(\omega)$$

where the integrand $\epsilon^{-1}\int_t^{t+\epsilon} \frac{\partial}{\partial s}X(s,\omega)ds$ satisfies

$$\begin{aligned} \left|\epsilon^{-1}\int_t^{t+\epsilon} \frac{\partial}{\partial s} X(s,\omega)ds\right| &\le \frac{1}{\epsilon}\int_t^{t+\epsilon} \left|\frac{\partial}{\partial s}X(s,\omega)\right| ds \\ &\le \frac{1}{\epsilon}\int_t^{t+\epsilon} Y(\omega)ds \quad \text{if} \quad [t,t+\epsilon] \subset (a,b) \\ &= Y(\omega) \in \mathcal{L}_1, \end{aligned}$$

and where

$$\lim_{\epsilon\downarrow 0} \frac{1}{\epsilon}\int_t^{t+\epsilon} \frac{\partial}{\partial s} X(s,\omega)ds = \lim_{\epsilon\downarrow 0} \frac{1}{\epsilon}(X(t+\epsilon,\omega) - X(t,\omega)) = \frac{\partial}{\partial t}X(t,\omega)$$

for a.e. ω. Thus it follows from the dominated convergence theorem that

$$\lim_{\epsilon\downarrow 0} \frac{h(t+\epsilon) - h(t)}{\epsilon} = \int_\Omega \frac{\partial}{\partial t} X(t,\omega)d\mu(\omega).$$

By using one-sided derivatives, the same argument works at the end-points a and b of the interval $[a,b]$. Thus the claimed equality holds for all $t \in [a,b]$.

(ii) For $X(t,\omega) = 1_{[\omega \le t]}$ and μ Lebesgue measure on $[0,1]$ we have

$$\int_{(0,1)} X(t,\omega) d\omega = \int_{(0,1)} 1_{[\omega \le t]} d\omega = t,$$

so

$$\begin{aligned} \frac{\partial}{\partial t} \int_{(0,1)} X(t,\omega) d\omega &= \frac{\partial}{\partial t} t = 1 \\ &\neq 0 = \int_{(0,1)} \frac{\partial}{\partial t} X(t,\omega) d\omega. \end{aligned}$$

This does not contradict (i) above because the derivative $(\partial/\partial t) 1_{[\omega \le t]}$ does not exist at $t = \omega$, and hence (i) does not apply.