Math 710 Final Exam

Austin Mohr

December 8, 2010

Problem 2

Let $\{X_n \mid n = 1, 2, \dots\}$ be a sequence of random variables with

$$P\{X_n = \pm n^3\} = \frac{1}{2n^2} \text{ and } P\{X_n = 0\} = 1 - \frac{1}{n^2}.$$

Proposition 1. *For the sequence described above,*

$$P\left\{\lim_{n\to\infty} X_n = 0\right\} = 1.$$

Proof. Let A_n be the event that X_n is nonzero. Formally,

$$A_n = \{\omega \in \Omega \mid X_n(\omega) \neq 0\}.$$

We see that

$$\begin{aligned} P(A_n) &= 1 - P\{X = 0\} \\ &= 1 - \left(1 - \frac{1}{n^2}\right) \\ &= \frac{1}{n^2}, \end{aligned}$$

and so

$$\begin{aligned} \sum_{n=1}^{\infty} P(A_n) &= \sum_{n=1}^{\infty} \frac{1}{n^2} \\ &< \infty. \end{aligned}$$

Thus, by Borel-Cantelli,

$$P([A_n \text{i.o}]) = P\left(\limsup_{n\to\infty} A_n\right) = 0.$$

By taking complements, we have

$$\begin{aligned} 1 &= P\left(\limsup_{n\to\infty} A_n^c\right) \\ &= P\left(\liminf_{n\to\infty} [X_n = 0]\right), \end{aligned}$$

and so

$$P\left(\lim_{n\to\infty} X_n = 0\right) = 1.$$

□

Proposition 2. *For the sequence described above,* $\lim_{n\to\infty} \mathbb{E}(X_n)$ *is either* $\pm\infty$ *or is undefined.*

Proof. For each n, $P\{X_n = \pm n^3\} = \frac{1}{2n^2}$. Hence,

$$P\{X_n^+ = n^3\} \geq \frac{1}{4n^2}$$

or

$$P\{X_n^- = n^3\} \geq \frac{1}{4n^2}$$

(possibly both). Suppose the former is true. It follows that

$$\begin{aligned}\mathbb{E}(X_n^+) &\geq n^3 P\{X_n^+ = n^3\} \\ &\geq n^3 \frac{1}{4n^2} \\ &= \frac{n}{4},\end{aligned}$$

and so

$$\begin{aligned}\lim_{n\to\infty} \mathbb{E}(X_n^+) &\geq \lim_{n\to\infty} \frac{n}{4} \\ &= \infty.\end{aligned}$$

Similarly, if $P\{X_n^- = n^3\} \geq \frac{1}{4n^2}$, then $\lim_{n\to\infty} \mathbb{E}(X_n^-) = \infty$. Therefore,

$$\lim_{n\to\infty} \mathbb{E}(X_n) = \begin{cases} \infty & \text{if } \lim_{n\to\infty} \mathbb{E}(X_n^+) = \infty \text{ and } \lim_{n\to\infty} \mathbb{E}(X_n^-) < \infty \\ -\infty & \text{if } \lim_{n\to\infty} \mathbb{E}(X_n^+) < \infty \text{ and } \lim_{n\to\infty} \mathbb{E}(X_n^-) = \infty \\ \text{undefined} & \text{if } \lim_{n\to\infty} \mathbb{E}(X_n^+) = \infty \text{ and } \lim_{n\to\infty} \mathbb{E}(X_n^-) < \infty. \end{cases}$$

□

Problem 3

Let $(\Omega, \mathfrak{F}, P)$ be a probability space.

Definition 1. *Two random variables* X *and* Y *are said to be independent provided that, for any* $A, B \in \mathfrak{B}(\mathfrak{R})$,

$$P[X^{-1}(A) \cdot Y^{-1}(B)] = P(X^{-1}(A)) \cdot P(Y^{-1}(B)).$$

Proposition 3. *Two random variables* X *and* Y *are independent if and only if, for every pair* f *and* g *of non-negative continuous functions on* $(\mathfrak{R}, \mathfrak{B}(\mathfrak{R}))$,

$$\mathbb{E}[f(X)g(Y)] = \mathbb{E}[f(X)]\mathbb{E}[g(Y)].$$

Proof. (⇒) Let f and g be any non-negative continuous functions on $(\mathfrak{R}, \mathfrak{B}(\mathfrak{R}))$. Since continuous functions between metric spaces are measurable (Resnick, 3.2.3), f and g are measurable. Since the composition of measurable functions is measurable (Resnick, 3.2.2), $f(X)$ and $g(Y)$ are measurable. Now, $\sigma(f(X)) \subseteq \sigma(X)$ (since $f \in \mathfrak{B}(\mathfrak{R})/\mathfrak{B}(\mathfrak{R})$) and $\sigma(g(Y)) \subset \sigma(Y)$ (since $g \in \mathfrak{B}(\mathfrak{R})/\mathfrak{B}(\mathfrak{R})$). Hence, $f(X)$ and $g(Y)$ are independent, since X and Y are independent.

Define now $Z_1 = f(X)$ and $Z_2 = g(Y)$. By the above, Z_1 and Z_2 are independent random variables. Thus, by Fubini's Theorem (as in Resnick, 5.9.2),

$$\mathbb{E}(Z_1 Z_2) = \mathbb{E}(Z_1)\mathbb{E}(Z_2),$$

but this is precisely

$$\mathbb{E}(f(X)g(Y)) = \mathbb{E}(f(X))\mathbb{E}(g(Y)).$$

(⇐) (Idea) Let a and b be real numbers and take $f = 1_{(0,a]}$ and $g = 1_{(0,b]}$. The support of $f(X)$ is $\{\omega \mid X(\omega) \leq a\}$ and the support of $g(Y)$ is $\{\omega \mid Y(\omega)\}$. Thus, the measures of the supports are $P(X \leq a)$ and $P(Y \leq b)$, respectively. Using the fact that

$$\mathbb{E}(f(X)g(Y)) = \mathbb{E}(f(X))\mathbb{E}(g(Y)),$$

I would like to derive that

$$P(X \leq a, Y \leq b) = P(X \leq a) \cdot P(Y \leq b).$$

Since a and b were arbitrary, we could conclude that X and Y are independent by the Factorization Criterion (Resnick, 4.2.1). Perhaps this may be accomplished by looking at the appropriate approximations of $f(X)$ and $g(Y)$ by simple functions (where the probabilty of the support becomes more evident in the computation). □

For each n, let X_n and Y_n be a pair of independent random variables and define

$$\lim_{n\to\infty} X_n = X \text{ and } \lim_{n\to\infty} Y_n = Y.$$

Proposition 4. *The functions X and Y are independent random variables.*

Proof. (Idea) We have, for each n and for all continuous, non-negative f and g,

$$\mathbb{E}(f(X_n)g(Y_n)) = \mathbb{E}(f(X_n))\mathbb{E}(g(Y_n)).$$

If we could switch limits with integrals, we would have

$$\begin{aligned} \lim_{n\to\infty} \mathbb{E}(f(X_n)g(Y_n)) &= \lim_{n\to\infty} \mathbb{E}(f(X_n))\mathbb{E}(g(Y_n)) \\ \mathbb{E}\left(\lim_{n\to\infty} f(X_n)g(Y_n)\right) &= \mathbb{E}\left(\lim_{n\to\infty} f(X_n)\right)\mathbb{E}\left(\lim_{n\to\infty} g(Y_n)\right) \\ \mathbb{E}(f(X)g(Y)) &= \mathbb{E}(f(X))\mathbb{E}(g(Y)), \end{aligned}$$

where the last step makes use of the continuity of f and g. Thus, appealing again to part b, we could conclude that X and Y are independent. I fail to see how to accomplish the interchange, however, as the X_n need not be monotone nor does there appear to be any bounding function. □

Problem 5

Suppose $\{p_k \mid k \geq 0\}$ is a probability mass function on $(\Omega = \{0, 1, 2, \dots\}, \mathfrak{P} = \mathfrak{P}(\Omega))$, where $\mathfrak{P}(\cdot)$ denotes the power set, so that $p_k \geq 0$ and $\sum_k p_k = 1$. Define for all $A \subset \Omega$,

$$P(A) = \sum_{k \in A} p_k.$$

Proposition 5. *The function P defined above is a probability measure on $(\Omega, \mathfrak{P})$.*

Proof. Since $p_k \geq 0$ for all k, $P(A) \geq 0$ for all $A \subset \Omega$.

We have, by definition of the probability mass function,

$$\begin{aligned} P(\Omega) &= \sum_{k \in \Omega} p_k \\ &= 1. \end{aligned}$$

Let $\{A_n\}$ be a countable sequence of disjoint events and let $A = \bigcup \{A_n\}$. It follows that

$$\begin{aligned} P\left(\bigcup_{n=1}^{\infty} A_n\right) &= P(A) \\ &= \sum_{k \in A} p_k \\ &= \sum_{n=1}^{\infty} \sum_{k \in A_n} p_k \qquad \text{(since the } A_n \text{ are disjoint)} \\ &= \sum_{n=1}^{\infty} P(A_n). \end{aligned}$$

□

Define the generating function $\Psi : ([0,1], \mathfrak{B}[0,1]) \to (\mathfrak{R}, \mathfrak{B})$ via

$$\Psi(s) = \sum_{k=0}^{\infty} p_k s^k.$$

Proposition 6. *The function Ψ defined above satisfies*

$$\Psi'(s) \equiv \frac{d}{ds} \Psi(s) = \sum_{k=1}^{\infty} k p_k s^{k-1}$$

for $0 \leq s \leq 1$.

Proof. Define the function $X_n = \sum_{k=1}^{n} kp_k s^{k-1}$. Observe that $0 \leq X_n \uparrow X$, and so by the Monotone Convergence Theorem

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{E}(X_n) &= \mathbb{E}\left(\lim_{n\to\infty} X_n\right) \\
\lim_{n\to\infty} \mathbb{E}\left(\sum_{k=1}^{n} kp_k s^{k-1}\right) &= \mathbb{E}\left(\lim_{n\to\infty} \sum_{k=1}^{n} kp_k s^{k-1}\right) \\
\lim_{n\to\infty} \sum_{k=0}^{n} p_k s^k &= \mathbb{E}\left(\lim_{n\to\infty} \sum_{k=1}^{n} kp_k s^{k-1}\right) \\
\Psi(s) &= \mathbb{E}\left(\Psi'(s)\right).
\end{aligned}$$

□

Proposition 7. *If X has probability measure P, then $\mathbb{E}(X) = \lim_{s\uparrow 1} \Psi'(s)$.*

Proof. We have,

$$\begin{aligned}
\mathbb{E}(X) &= \int_\Omega X(\omega) dP \\
&= \sum_{k=0}^{\infty} kP(X = k) \\
&= \sum_{k=1}^{\infty} kp_k \\
&= \lim_{s\uparrow 1} \Psi'(s).
\end{aligned}$$

□

Problem 6

Let $X_1, X_2, \ldots, X_n \in L_2(P)$ be random variables defined on a probability space $(\Omega, \mathfrak{F}, P)$. For each $i, j \in \{1, 2, \ldots, n\}$, define the covariances

$$\sigma_{ij} = \mathbb{C}(X_i, X_j) = \mathbb{E}\{[X_i - \mu_i][X_j - \mu_j]\},$$

where

$$\mu_i = \mathbb{E}(X_i) \text{ and } \sigma_i^2 = \sigma_{ii} = \mathbb{V}(X_i) = \mathbb{E}[(X_i - \mu_i)^2].$$

Lemma 1. *For any random variable X and real numbers a and b,*

$$\mathbb{V}(aX + b) = a^2 \mathbb{V}(X).$$

Proof. It follows from the linearity of expectation that,

$$\begin{aligned}\mathbb{V}(aX+b) &= \mathbb{E}[(aX+b-\mathbb{E}(aX+b))^2] \\ &= \mathbb{E}[(aX+b-a\mathbb{E}(X)-b)^2] \\ &= \mathbb{E}[(a(X-\mathbb{E}(X)))^2] \\ &= a^2\mathbb{E}[(X-\mathbb{E}(X))^2] \\ &= a^2\mathbb{V}(X).\end{aligned}$$

□

Lemma 2. *For any random variables X and Y,*

$$\mathbb{V}(X+Y) = \mathbb{V}(X) + 2\mathbb{C}(X,Y) + \mathbb{V}(Y).$$

Proof. It follows from the linearity of expectation that,

$$\begin{aligned}\mathbb{V}(X+Y) &= \mathbb{E}[(X+Y-\mathbb{E}(X)-\mathbb{E}(Y))^2] \\ &= \mathbb{E}[((X-\mathbb{E}(X))+(Y-\mathbb{E}(Y)))^2] \\ &= \mathbb{E}[(X-\mathbb{E}(X))^2 + 2(X-\mathbb{E}(X))(Y-\mathbb{E}(Y)) + (Y-\mathbb{E}(Y))^2] \\ &= \mathbb{E}[(X-\mathbb{E}(X))^2] + \mathbb{E}[2(X-\mathbb{E}(X))(Y-\mathbb{E}(Y))] + \mathbb{E}[(Y-\mathbb{E}(Y))^2] \\ &= \mathbb{V}(X) + 2\mathbb{C}(X,Y) + \mathbb{V}(Y).\end{aligned}$$

□

Proposition 8. *For all i and j,*

$$\sigma_{ij} \le |\sigma_{ij}| \le \sigma_i\sigma_j.$$

Moreover, $|\sigma_{ij}| = \sigma_i\sigma_j$ if and only if, for some α and β, we have $P\{X_j = \alpha + \beta X_i\} = 1$.

Proof. For all real numbers x, we have $x \le |x|$, so certainly $\sigma_{ij} \le |\sigma_{ij}|$.

For the second inequality, let t be a real variable. It follows from the lemmas that

$$\begin{aligned}0 &\le \mathbb{V}[tX_i + X_j] \\ &= \mathbb{V}(tX_i) + 2\mathbb{C}(X_i, X_j) + \mathbb{V}(X_j) \\ &= \sigma_i^2 t^2 + 2\sigma_{ij}t + \sigma_j^2.\end{aligned}$$

Viewing this as a non-negative quadratic in t, we have that

$$0 \ge 4\sigma_{ij}^2 - 4\sigma_i^2\sigma_j^2,$$

and so

$$|\sigma_{ij}| \le \sigma_i\sigma_j.$$

For the remaining claim, observe that

$$\begin{aligned}|\sigma_{ij}| = \sigma_i\sigma_j &\Leftrightarrow \sigma_{ij}^2 = \sigma_i^2\sigma_j^2 \\ &\Leftrightarrow 0 = 4\sigma_{ij}^2 - 4\sigma_i^2\sigma_j^2.\end{aligned}$$

Hence, $\mathbb{V}[tX_i + X_j]$ has a unique real root t_0. Now, the variance of a random variable is equal to 0 if and only if it is constant with probability one. That is,

$$P\{t_0X_i + X_j = \alpha\} = 1,$$

or equivalently

$$P\{X_j = \alpha - t_0X_i\} = 1.$$

□

Proposition 9. *For real constants α_i and β_i, $i = 1, 2, \ldots, n$,*

$$\mathbb{C}\left\{\sum_{i=1}^n \alpha_i X_i, \sum_{j=1}^n \beta_j X_j\right\} = \sum_{i=1}^n\sum_{j=1}^n \alpha_i\beta_j\sigma_{ij}.$$

Proof. Applying linearity of expectation, we have

$$\begin{aligned}\mathbb{C}\left\{\sum_{i=1}^n \alpha_i X_i, \sum_{j=1}^n \beta_j X_j\right\} &= \mathbb{E}\left\{\left(\sum_{i=1}^n \alpha_i X_i\right)\left(\sum_{j=1}^n \beta_j X_j\right)\right\} - \mathbb{E}\left\{\sum_{i=1}^n \alpha_i X_i\right\}\mathbb{E}\left\{\sum_{j=1}^n \beta_j X_j\right\} \\ &= \mathbb{E}\left\{\sum_{i=1}^n\sum_{j=1}^n \alpha_i\beta_j X_iX_j\right\} - \mathbb{E}\left\{\sum_{i=1}^n \alpha_i X_i\right\}\mathbb{E}\left\{\sum_{j=1}^n \beta_j X_j\right\} \\ &= \sum_{i=1}^n\sum_{j=1}^n \alpha_i\beta_j\mathbb{E}(X_iX_j) - \left\{\sum_{i=1}^n \alpha_i\mathbb{E}(X_i)\right\}\left\{\sum_{j=1}^n \beta_j\mathbb{E}(X_j)\right\} \\ &= \sum_{i=1}^n\sum_{j=1}^n \alpha_i\beta_j\mathbb{E}(X_iX_j) - \sum_{i=1}^n\sum_{j=1}^n \alpha_i\beta_j\mathbb{E}(X_i)\mathbb{E}(X_j) \\ &= \sum_{i=1}^n\sum_{j=1}^n \alpha_i\beta_j\mathbb{C}(X_i, X_j) \\ &= \sum_{i=1}^n\sum_{j=1}^n \alpha_i\beta_j\sigma_{ij}.\end{aligned}$$

□

Proposition 10. *For real constants α_i, $i = 1, 2, \ldots, n$,*

$$\mathbb{V}\left\{\sum_{i=1}^n \alpha_i X_i\right\} = \sum_{i=1}^n \alpha_i^2\sigma_i^2 + 2\sum_{1\le i<j\le n} \alpha_i\alpha_j\sigma_{ij}.$$

Proof. By definition, $\mathbb{V}(X) = \mathbb{C}(X, X)$ for any random variable X. Thus, by the previous proposition,

$$\begin{aligned}
\mathbb{V}\left\{\sum_{i=1}^{n} \alpha_i X_i\right\} &= \mathbb{C}\left\{\sum_{i=1}^{n} \alpha_i X_i, \sum_{i=1}^{n} \alpha_i X_i\right\} \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \alpha_j \mathbb{C}(X_i, X_j) \\
&= \sum_{i=1}^{n} \alpha_i^2 \mathbb{C}(X_i, X_i) + 2 \sum_{1 \le i < j \le n} \alpha_i \alpha_j \mathbb{C}(X_i, X_j) \\
&= \sum_{i=1}^{n} \alpha_i^2 \mathbb{V}(X_i) + 2 \sum_{1 \le i < j \le n} \alpha_i \alpha_j \mathbb{C}(X_i, X_j) \\
&= \sum_{i=1}^{n} \alpha_i^2 \sigma_i^2 + 2 \sum_{1 \le i < j \le n} \alpha_i \alpha_j \sigma_{ij}.
\end{aligned}$$

□

Proposition 11. *Let $X_1, \ldots, X_n$ be independent random variables. Furthermore, suppose that the constants α_i are restricted to belong to $[0, 1]$ and must satisfy $\sum_{i=1}^{n} \alpha_i = 1$. Finally, letting $s = \sum_{i=1}^{n} \sigma_i^{-1}$,*

$$\mathbb{V}\left\{\sum_{i=1}^{n} \alpha_i X_i\right\} \ge ns^{-2}$$

with equality if and only if $\alpha_i = (\sigma_i s)^{-1}$ for all i.

Proof. Since the X_i are independent, $\sigma_{ij} = 0$ for all $i \ne j$ (Resnick, 5.9.2). Thus, the expression for the variance reduces to

$$\mathbb{V}\left\{\sum_{i=1}^{n} \alpha_i X_i\right\} = \sum_{i=1}^{n} (\alpha_i \sigma_i)^2.$$

Now, since we require $\sum_{i=1}^{n} \alpha_i = 1$, the variance will be minimized precisely when all the terms in its summation are equal. To that end, tentatively set $\alpha_i = \sigma_i^{-1}$ for all i. Thus, $\alpha_i \sigma_i = 1$ for all i, and so all the terms are equal, as desired. Now,

$$\begin{aligned}
\sum_{i=1}^{n} \alpha_i &= \sum_{i=1}^{n} \sigma_i^{-1} \\
&= s.
\end{aligned}$$

To ensure that the α_i indeed sum to 1, we scale them all by a factor of s^{-1}. Thus, we take instead $\alpha_i = (\sigma_i s)^{-1}$ for all i. Using these values, we have

$$\begin{aligned}\mathbb{V}\left\{\sum_{i=1}^{n} \alpha_i X_i\right\} &= \sum_{i=1}^{n} (\alpha_i \sigma_i)^2 \\ &= \sum_{i=1}^{n} ((\sigma_i s)^{-1} \sigma_i)^2 \\ &= \sum_{i=1}^{n} s^{-2} \\ &= n s^{-2}.\end{aligned}$$

□

Problem 8

For $i = 1, 2$, let $(\Omega_i, \mathfrak{B}_i, P_i)$ be probability spaces. Define $\Omega = \Omega_1 \times \Omega_2$ and $\mathfrak{B} = \mathfrak{B}_1 \otimes \mathfrak{B}_2 = \sigma(RECTS)$, where

$$RECTS = \{B_1 \times B_2 \mid B_1 \in \mathfrak{B}_1, B_2 \in \mathfrak{B}_2\}.$$

Let $P = P_1 \times P_2$ be the product probability measure so that, for $B_1 \times B_2 \in RECTS$, we have $P(B_1 \times B_2) = P_1(B_1)P_2(B_2)$. Define the class of subsets

$$\mathfrak{C} = \left\{B \subset \Omega \mid \int_{\Omega} 1_B(\omega_1, \omega_2) dP(\omega_1, \omega_2) = \int_{\Omega_1} Y(\omega_1) dP_1(\omega_1)\right\},$$

where $Y(\omega_1) = \int_{\Omega_2} 1_B(\omega_1, \omega_2) dP_2(\omega_2)$.

Proposition 12. *The class $RECTS$ is a subset of the class $\mathfrak{C}$.*

Proof. Let $B = B_1 \times B_2$ belong to $RECTS$. We have

$$\begin{aligned}\int_{\Omega} 1_B(\omega_1, \omega_2) dP(\omega_1, \omega_2) &= \int_B dP(\omega_1, \omega_2) \\ &= P(B).\end{aligned}$$

At the same time, we have

$$\begin{aligned}
\int_{\Omega_1}\int_{\Omega_2} 1_B(\omega_1,\omega_2)dP_2(\omega_2)dP_1(\omega_1) &= \int_{\Omega_1}\int_{\Omega_2} 1_{B_1}(\omega_1)1_{B_2}(\omega_2)dP_2(\omega_2)dP_1(\omega_1) \\
&= \int_{\Omega_1}\int_{B_2} 1_{B_1}(\omega_1)dP_2(\omega_2)dP_1(\omega_1) \\
&= \int_{\Omega_1} 1_{B_1}(\omega_1)P_2(B_2)dP_1(\omega_1) \\
&= \int_{B_1} P_2(B_2)dP_1(\omega_1) \\
&= P_1(B_1)P_2(B_2) \\
&= P(B).
\end{aligned}$$

Hence, for all $B \in RECTS$,

$$\int_{\Omega} 1_B(\omega_1,\omega_2)dP(\omega_1,\omega_2) = \int_{\Omega_1} Y(\omega_1)dP_1(\omega_1),$$

and so $RECTS \subseteq \mathfrak{C}$. □

Proposition 13. *The class* $\mathfrak{C}$ *is a* λ*-system.*

Proof. We have immediately that $\Omega = \Omega_1 \times \Omega_2 \in RECTS \subseteq \mathfrak{C}$, so $\Omega \in \mathfrak{C}$.

Next, let $B \in \mathfrak{C}$. We have

$$\begin{aligned}
\int_{\Omega} 1_{B^c}(\omega_1,\omega_2)dP(\omega_1,\omega_2) &= \int_{B^c} dP(\omega_1,\omega_2) \\
&= P(B^c) \\
&= 1 - P(B) \\
&= 1 - \int_{\Omega_1}\int_{\Omega_2} 1_{B_{\omega_1}}(\omega_2)dP_1(\omega_1) \qquad \text{(since } B \in \mathfrak{C}\text{)} \\
&= \int_{\Omega_1}\int_{\Omega_2} 1 - 1_{B_{\omega_1}}(\omega_2)dP_1(\omega_1) \\
&= \int_{\Omega_1}\int_{\Omega_2} 1_{(B_{\omega_1})^c}(\omega_2)dP_1(\omega_1) \\
&= \int_{\Omega_1}\int_{\Omega_2} 1_{(B^c)_{\omega_1}}(\omega_2)dP_1(\omega_1) \\
&= \int_{\Omega_1}\int_{\Omega_2} 1_{B^c}(\omega_1,\omega_2)dP_2(\omega_2)dP_1(\omega_1).
\end{aligned}$$

Finally, let $\{B_n \mid n = 1, 2, \ldots\}$ be a collection of disjoint elements of $\mathfrak{C}$. We have

$$\begin{aligned}
\int_\Omega 1_{\sum_{n=1}^\infty A_n} dP(\omega_1,\omega_2) &= \int_{\sum_{n=1}^\infty A_n} dP(\omega_1,\omega_2) \\
&= P\left(\sum_{n=1}^\infty A_n\right) \\
&= \sum_{n=1}^\infty P(A_n) \\
&= \sum_{n=1}^\infty \int_{\Omega_1}\int_{\Omega_2} 1_{(A_n)_{\omega_1}}(\omega_2) dP_1(\omega_1) && \text{(since } A_n \in \mathfrak{C} \text{ for all } n\text{)} \\
&= \int_{\Omega_1}\int_{\Omega_2} \sum_{n=1}^\infty 1_{(A_n)_{\omega_1}}(\omega_2) dP_1(\omega_1) && \text{(by MCT)} \\
&= \int_{\Omega_1}\int_{\Omega_2} 1_{(\sum_{n=1}^\infty A_n)_{\omega_1}}(\omega_2) dP_1(\omega_1) \\
&= \int_{\Omega_1}\int_{\Omega_2} 1_{\sum_{n=1}^\infty A_n}(\omega_1,\omega_2) dP_2(\omega_2) dP_1(\omega_1).
\end{aligned}$$

Therefore, $\mathfrak{C}$ is a λ-system. □

Proposition 14. *For every $B \in \mathfrak{B}$,*

$$\int_\Omega 1_B(\omega_1,\omega_2) dP(\omega_1,\omega_2) = \int_{\Omega_1}\left\{\int_{\Omega_2} 1_B(\omega_1,\omega_2) dP_2(\omega_2)\right\} dP_1(\omega_1).$$

Proof. We have shown $RECTS \subseteq \mathfrak{C}$ and that $\mathfrak{C}$ is a λ-system. If we can show also that $RECTS$ is a π-system, then Dynkin's Theorem gives $\mathfrak{B} = \sigma(RECTS) \subset \mathfrak{C}$, from which the conclusion follows.

To finish the proof, let $B_1 \times B_2$ and $B_1' \times B_2'$ belong to $RECTS$. It follows immediately that

$$(B_1 \times B_2) \cap (B_1' \times B_2') = (B_1 \cap B_1') \times (B_2 \cap B_2').$$

Since $\mathfrak{B}_1$ and $\mathfrak{B}_2$ are closed under intersections, $B_1 \cap B_1' \in \mathfrak{B}_1$ and $B_2 \cap B_2' \in \mathfrak{B}_2$, and so $(B_1 \cap B_1') \times (B_2 \cap B_2') \in RECTS$. □

To establish the more general result where 1_B in part c is replaced with any $\mathfrak{B}$-measurable positive random variable X, we first establish the result for simple functions of the form $X_n = \sum_{i=1}^n a_i 1_{B_i}$, where $B_i \in \mathfrak{B}$ for all i. The result for simple functions follows readily from the linearity of the integral. Since each X_n is positive, we can take a sequence $X_n \uparrow X$. By hypothesis,

$$\int_\Omega X_n(\omega_1,\omega_2) dP(\omega_1,\omega_2) = \int_{\Omega_1}\left\{\int_{\Omega_2} X_n(\omega_1,\omega_2) dP_2(\omega_2)\right\} dP_1(\omega_1)$$

for all n. Applying the Monotone Convergence Theorem, we can conclude that

$$\int_\Omega X_n(\omega_1,\omega_2)dP(\omega_1,\omega_2) \uparrow \int_\Omega X(\omega_1,\omega_2)dP(\omega_1,\omega_2)$$

and

$$\int_{\Omega_1}\left\{\int_{\Omega_2} X_n(\omega_1,\omega_2)dP_2(\omega_2)\right\}dP_1(\omega_1) \uparrow \int_{\Omega_1}\left\{\int_{\Omega_2} X(\omega_1,\omega_2)dP_2(\omega_2)\right\}dP_1(\omega_1),$$

from which it follows that

$$\int_\Omega X(\omega_1,\omega_2)dP(\omega_1,\omega_2) = \int_{\Omega_1}\left\{\int_{\Omega_2} X(\omega_1,\omega_2)dP_2(\omega_2)\right\}dP_1(\omega_1).$$

Problem 10

Suppose that X and Y are independent random variables and let $h : \mathfrak{R}^2 \to [0,\infty)$ be a measurable function such that $\mathbb{E}\{h^2(X,Y)\} < \infty$. Define

$$g(x) = \mathbb{E}\{h(x,Y)\} \text{ and } k(x) = \mathbb{V}\{h(x,Y)\}.$$

Proposition 15. *The functions g and k are both measureable on $\mathfrak{R} \to \mathfrak{R}$.*

Proof. Define $\hat{h}(x,\omega) = h(x,Y(\omega))$. Since h and Y are measurable, $\hat{h}$ is measurable, as it is defined by the composition of two measurable functions. Hence, we can take a collection $\{\hat{h}_n\}$ of simple functions with $\hat{h}_n \uparrow \hat{h}$. Define now $g_n(x) = \int_\Omega \hat{h}_n(x,\omega)dP(\omega)$ for $n = 1, 2, \ldots$. By the Monotone Convergence Theorem, $g_n \uparrow g$. To conclue that g is measurable, it remains to show that each g_n is simple. To that end, observe that

$$\begin{aligned}
g_n(x) &= \int_\Omega \hat{h}_n(x,\omega)dP(\omega) \\
&= \int_\Omega \sum_{j=1}^k a_j 1_{A_j}(x,\omega)dP(\omega) && \text{(constants } a_j \text{ and } \{A_j\} \text{ a partition of } \mathfrak{R}) \\
&= \sum_{j=1}^k \int_\Omega a_j 1_{A_j}(x,\omega)dP(\omega) && \text{(by MCT)} \\
&= \sum_{j=1}^k \int_\Omega a_j 1_{A_j}(x) 1_{A_j}(\omega)dP(\omega) \\
&= \sum_{j=1}^k a_j P(A_j) 1_{A_j}(x),
\end{aligned}$$

and so g_n is simple.

For $k(x)$, we have

$$\begin{aligned} k(x) &= \mathbb{V}(\hat{h}) \\ &= \mathbb{E}(\hat{h}^2) - \mathbb{E}(\hat{h})^2 \\ &= \mathbb{E}(\hat{h}^2) - g^2. \end{aligned}$$

Now, since $\hat{h}$ is measureable, $\hat{h}^2$ is measurable. Following the same argument as above, we find that $\mathbb{E}(\hat{h}^2)$ is measurable, and so k is measurable. □

Proposition 16. *For g and h as defined above,*

$$\mathbb{E}\{g(X)\} = \mathbb{E}\{h(X,Y)\}.$$

Proof. Suppose X is a random variable on Ω_1 with probability measure P_1 and Y is a random variable on Ω_2 with probability measure P_2. Finally, let P be the probability measure on $\Omega = \Omega_1 \times \Omega_2$ induced by P_1 and P_2. In order to make use of Fubini's Theorem later in the proof, we must establish first that $P = P_1 \times P_2$. To that end, observe that for any measurable sets $A \subset \Omega_1$ and $B \subset \Omega_2$,

$$\begin{aligned} P(A \times B) &= \int_\Omega 1_{A\times B} dP \\ &= \int_\Omega 1_A \cdot 1_B dP \\ &= \int_\Omega 1_A dP \cdot \int_\Omega 1_B dP && \text{(since } X \text{ and } Y \text{ are independent)} \\ &= \int_{\Omega_1} 1_A(\omega_1) dP_1(\omega_1) \cdot \int_{\Omega_2} 1_B(\omega_2) dP_2(\omega_2) \\ &= P_1(A) \cdot P_2(B). \end{aligned}$$

Now,

$$\begin{aligned} \mathbb{E}(g(X)) &= \int_{\Omega_1} g(X(\omega_1)) dP_1(\omega_1) \\ &= \int_{\Omega_1} \int_{\Omega_2} h(X(\omega_1), Y(\omega_2)) dP_2(\omega_2) dP_1(\omega_1) \\ &= \int_\Omega h(X(\omega_1), Y(\omega_2)) dP(\omega_1, \omega_2) && \text{(by Fubini's Theorem)} \\ &= \mathbb{E}(h(X,Y)). \end{aligned}$$

□

Proposition 17. *For g, h, and k as defined above,*

$$\mathbb{V}\{g(X)\} + \mathbb{E}\{k(X)\} = \mathbb{V}\{h(X,Y)\}.$$

Proof. We have

$$
\begin{aligned}
&\mathbb{V}(g(X)) + \mathbb{E}(k(X)) \\
=&\int_{\Omega_1} g(X(\omega_1))^2 dP_1(\omega_1) - \left[\int_{\Omega_1} g(X(\omega_1)) dP_1(\omega_1)\right]^2 + \int_{\Omega_1} k(X(\omega_1)) dP_1(\omega_1) \\
=&\int_{\Omega_1} g(X(\omega_1))^2 + k(X(\omega_1)) dP_1(\omega_1) - \left[\int_{\Omega_1} g(X(\omega_1)) dP_1(\omega_1)\right]^2 \\
=&\int_{\Omega_1} \left(\int_{\Omega_2} h(X(\omega_1), Y(\omega_2)) dP_2(\omega_2)\right)^2 + \int_{\Omega_2} h(X(\omega_1), Y(\omega_2))^2 dP_2(\omega_2) \\
&\quad - \left(\int_{\Omega_2} h(X(\omega_1), Y(\omega_2)) dP_2(\omega_2)\right)^2 dP_1(\omega_1) - \left[\int_{\Omega_1}\int_{\Omega_2} h(X(\omega_1), Y(\omega_2)) dP_2(\omega_2) dP_1(\omega_1)\right]^2 \\
=&\int_{\Omega_1}\int_{\Omega_2} h(X(\omega_1), Y(\omega_2))^2 dP_2(\omega_2) dP_1(\omega_1) - \left[\int_{\Omega_1}\int_{\Omega_2} h(X(\omega_1), Y(\omega_2)) dP_2(\omega_2) dP_1(\omega_1)\right]^2 \\
=&\mathbb{E}(h(X, Y)^2) - \mathbb{E}(h(X, Y))^2 \\
=&\mathbb{V}(h(X, Y)).
\end{aligned}
$$

□