Logan Hotz
Stochastic Calculus for Finance II
Shreve

Question 1

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a general probability space, and suppose a random variable X on this space is measurable with respect to the trivial σ-algebra $\mathcal{F}_0 = \{\varnothing, \Omega\}$. Show that X is not random (i.e., there is a constant c such that $X(\omega) = c$ for all $\omega \in \Omega$). Such a random variable is called *degenerate*.

Answer

A random variable X is said to be measurable with respect to a σ-algebra $\mathcal{G}$ if $\{\omega \in \Omega \,|\, X(\omega) \in B\} \subseteq \mathcal{G}$ for all Borel subsets of the range of X, B. Let X be a random variable measurable with respect to $\mathcal{F}_0 = \{\varnothing, \Omega\}$ and suppose there exists $\omega_1, \omega_2 \in \Omega$ for which $X(\omega_1) \neq X(\omega_2)$. Taking the Borel set $B_1 = \{X(\omega_1)\}$, it must be the case that $\{X \in B_1\}$ is not empty, and because X is measurable with respect to $\mathcal{F}_0$, we are left to conclude $\{X \in B_1\} = \Omega$. This is a contradiction, though, as $X(\omega_2) \notin B_1$ implies $\omega_2 \notin \{X \in B_1\} = \Omega$. Thus, for all $\omega \in \Omega$, it must be the case that $X(\omega) = c$ for some c in the image of X.

If one assumes X takes values in the reals and has an associated law, an alternative proof could look like the following.

A random variable X is said to be measurable with respect to a σ-algebra $\mathcal{G}$ if $\{\omega \in \Omega \,|\, X(\omega) \in B\} \subseteq \mathcal{G}$ for all Borel subsets B. For all $x \in \mathbb{R}$ (and therefore all Borel sets of the form $(-\infty, x)$), then, $\{X \leq x\} \in \mathcal{F}_0 \equiv \{\varnothing, \Omega\}$. Hence, $\mathbb{P}(X \leq x)$ takes one of two values:

$$
\begin{aligned}
\mathbb{P}(X \leq x) = 0 \quad &\text{if} \quad \{X \leq x\} = \{\omega \in \Omega \,|\, X(\omega) \leq x\} = \varnothing \\
\mathbb{P}(X \leq x) = 1 \quad &\text{if} \quad \{X \leq x\} = \{\omega \in \Omega \,|\, X(\omega) \leq x\} = \Omega
\end{aligned}
$$

Because $\lim_{x \to -\infty} \mathbb{P}(X \leq x) = 0$ and $\lim_{x \to \infty} \mathbb{P}(X \leq x) = 1$, there exists a $c \in \mathbb{R}$ for which $\mathbb{P}(X \leq c) = 1$ and $\mathbb{P}(X \leq x) = 0$ for all $x < c$. Thus

$$
\mathbb{P}(X = c) = \lim_{\varepsilon \to 0} \mathbb{P}(c - \varepsilon < X \leq c) = \mathbb{P}(X \leq c) - \lim_{\varepsilon \to 0} \mathbb{P}(X \leq c - \varepsilon) = 1 - 0 = 1.
$$

Question 2

Independence of random variables can be affectedby changes of measure. To illustrate this point, consider the space of two coin tosses $\Omega_2 = \{HH, HT, TH, TT\}$, and let stock prices be given by

$$\begin{array}{lll} S_0 = 4 & S_1(H) = 8 & S_2(HH) = 16 \\ & S_1(T) = 2 & S_2(HT) = S_2(TH) = 4 \\ & & S_2(TT) = 1. \end{array}$$

Consider two probability measures given by

$$\begin{array}{llll} \widetilde{\mathbb{P}}(HH) = 1/4 & \widetilde{\mathbb{P}}(HT) = 1/4 & \widetilde{\mathbb{P}}(TH) = 1/4 & \widetilde{\mathbb{P}}(TT) = 1/4 \\ \mathbb{P}(HH) = 4/9 & \mathbb{P}(HT) = 2/9 & \mathbb{P}(TH) = 2/9 & \mathbb{P}(TT) = 1/9. \end{array}$$

Define the random variable

$$X = \begin{cases} 1 & \text{if } S_2 = 4 \\ 0 & \text{if } S_2 \neq 4 \end{cases}$$

(a) List all the sets in $\sigma(X)$.
(b) List all the sets in $\sigma(S_1)$.
(c) Show that $\sigma(X)$ and $\sigma(S_1)$ are independent under the probability measure $\widetilde{\mathbb{P}}$.
(d) Show that $\sigma(X)$ and $\sigma(S_1)$ are not independent under the probability measure $\mathbb{P}$.
(e) Under $\mathbb{P}$, we have $\mathbb{P}(S_1 = 8) = 2/3$ and $\mathbb{P}(S_1 = 2) = 1/3$. Explain intuitively why, if you are told that $X = 1$, you would want to revise your estimate of the distribution of S_1.

Answer

(a) The σ-algebra generated by X is $\sigma(X) = \big\{\varnothing, \Omega, \{HH, TT\}, \{HT, TH\}\big\}$.

(b) The σ-algebra generated by S_1 is $\sigma(S_1) = \big\{\varnothing, \Omega, \{HH, HT\}, \{TH, TT\}\big\}$.

(c) Note that it is sufficient to only consider the non-trivial events (i.e. those which are not $\varnothing$ nor Ω). For those non-trivial events in $\sigma(X)$, we have

$$\widetilde{\mathbb{P}}(\{HH, TT\}) = \widetilde{\mathbb{P}}(HH) + \widetilde{\mathbb{P}}(TT) = \tfrac{1}{4} + \tfrac{1}{4} = \tfrac{1}{2} \quad \text{and} \quad \widetilde{\mathbb{P}}(\{HT, TH\}) = \tfrac{1}{2}.$$

Similarly, both non-trivial events in $\sigma(S_1)$ also occur with probability $1/2$. Hence for those non-trivial $A \in \sigma(X), B \in \sigma(S_1)$, we have $\widetilde{\mathbb{P}}(A)\widetilde{\mathbb{P}}(B) = (1/2)(1/4) = 1/4$.

In the tables below we show the joint events and their probabilities, respectively:

$\{A \cap B\}$	$\{HH, HT\}$	$\{TH, TT\}$
$\{HH, TT\}$	HH	TT
$\{HT, HT\}$	HT	TH

$\widetilde{\mathbb{P}}(A \cap B)$	$\{HH, HT\}$	$\{TH, TT\}$
$\{HH, TT\}$	$1/4$	$1/4$
$\{HT, HT\}$	$1/4$	$1/4$

Since $\widetilde{\mathbb{P}}(A \cap B) = \widetilde{\mathbb{P}}(A)\widetilde{\mathbb{P}}(B)$ for all $A \in \sigma(A), B \in \sigma(S_1)$, those σ-algebras are independent under $\widetilde{\mathbb{P}}$.

(d) The probabilities of the non-trivial events in $\sigma(X)$ are

$$\mathbb{P}(\{HH, TT\}) = \mathbb{P}(HH) + \mathbb{P}(TT) = \tfrac{4}{9} + \tfrac{1}{9} = \tfrac{5}{9} \quad \text{and} \quad \mathbb{P}(\{HT, TH\}) = \tfrac{4}{9},$$

whereas those in $\sigma(S_1)$ are $\mathbb{P}(\{HH, HT\}) = 2/3$ and $\mathbb{P}(\{TH, TT\}) = 1/3$. Below we use the same table setup as in part (c) to show the probabilities of the joint events:

$\{A \cap B\}$	$\{HH, HT\}$	$\{TH, TT\}$
$\{HH, TT\}$	HH	TT
$\{HT, HT\}$	HT	TH

$\widetilde{\mathbb{P}}(A \cap B)$	$\{HH, HT\}$	$\{TH, TT\}$
$\{HH, TT\}$	$4/9$	$2/9$
$\{HT, HT\}$	$2/9$	$1/9$

Take, for example, $A = \{HH, TT\} \in \sigma(X)$ and $B = \{HH, HT\} \in \sigma(S_1)$. We have

$$\tfrac{4}{9} = \mathbb{P}(\{A \cap B\}) \neq \mathbb{P}(A)\mathbb{P}(B) = \left(\tfrac{5}{9}\right)\left(\tfrac{2}{3}\right) = \tfrac{10}{27}.$$

Thus $\sigma(X)$ and $\sigma(S_1)$ are not independent under $\mathbb{P}$.

The work done here was a little excessive, of course. We really just needed one joint event $A \cap B, A \in \sigma(X), B \in \sigma(S_1)$, to show that the two σ-algebras are not independent.

(e) Having been told $X = 1$, we know that one of the events in $\{TH, HT\}$ occured. These events are equally likely under $\mathbb{P}$, so with the information that it's a 50-50 chance the first coin flip is tails, one might want to shift more of the estimated mass of S_1's distribution to that outcome. This is all downstream of our conclusion in part (d): X and S_1 are not independent under $\mathbb{P}$, so information about one random variable gives us information about the other.

Question 3 - Rotating the Axes

Let X and Y be independent standard normal random variables. Let θ be a constant, and define the random variables

$$V = X\cos\theta + Y\sin\theta \quad \text{and} \quad W = -X\sin\theta + Y\cos\theta.$$

Show that V and W are independent standard normal random variables.

Answer

Note that $\cos\theta$ and $\sin\theta$ are simply constants, as θ is assumed to be one. Hence the mean of V is $\mathbb{E}V = \cos\theta\mathbb{E}X + \sin\theta\mathbb{E}Y = 0 + 0$ by the linearity of the expectations operator. The variance of V is $\text{Var}V = \cos^2\theta\text{Var}X + \sin^2\theta\text{Var}Y = (\sin^2\theta + \cos^2\theta)(1) = 1$. Finally, as V is a linear combination of independent normals, it is also a normally distributed random variable. For the same reasons, W is a mean-zero normal random variable with unit variance.

It remains to be shown that $\text{Cov}(V, W) = \mathbb{E}[VW] = 0$, which is a sufficient condition for the normal random variables V and W to be independent. We have

$$\begin{aligned} VW &= (Y^2 - X^2)\sin\theta\cos\theta + XY(\cos^2\theta - \sin^2\theta) \\ &\implies \mathbb{E}\big[VW\big] = (\text{Var}Y - \text{Var}X)\sin\theta\cos\theta + \mathbb{E}X\mathbb{E}Y(\cos^2\theta - \sin^2\theta) = 0, \end{aligned}$$

where the first equality in the second line uses the assumed independence of X and Y.

Question 4

In Example 2.2.8, X is a standard normal random variable and Z is an independent random variable satisfying

$$\mathbb{P}(Z = 1) = \mathbb{P}(Z = -1) = \tfrac{1}{2}.$$

We defined $Y = XZ$ and showed that Y is standard normal. We established that although X and Y are uncorrelated, they are not independent. In this exercise we use moment-generating functions to show that Y is standard normal and X and Y are not independent.

(i) Establish the joint moment-generating function formula

$$\mathbb{E}\big[e^{uX+vY}\big] = e^{\frac{1}{2}(u^2+v^2)}\frac{e^{uv}+e^{-uv}}{2}.$$

(ii) Use the formula above to show that $\mathbb{E}[e^{vY}] = e^{\frac{1}{2}v^2}$. This is the moment-generating function for a standard normal random variable, and thus Y must be a standard normal random variable.

(iii) Use the formula in (i) and Theorem 2.2.7(iv) to show that X and Y are not independent.

Answer

(i) From the Law of Total Expectation, we have $\mathbb{E}[e^{uX+vY}] = \mathbb{E}[e^{uX+vY} \mid Z = 1]\mathbb{P}(Z = 1) + \mathbb{E}[e^{uX+vY} \mid Z = -1]\mathbb{P}(Z = -1)$. Substituting $Y = XZ$ into these expressions, we have

$$\mathbb{E}\big[e^{uX+vY}\big] = \tfrac{1}{2}\mathbb{E}\big[e^{(u+v)X}\big] + \tfrac{1}{2}\mathbb{E}\big[e^{(u-v)X}\big].$$

The moment-generating function of a standard normal random variable X is $\mathbb{E}[e^{sX}] = e^{\frac{1}{2}s^2}$. Evaluating this at $s = u \pm v$, we have the desired result:

$$\mathbb{E}\big[e^{uX+vY}\big] = \tfrac{1}{2}e^{\frac{1}{2}(u+v)^2} + \tfrac{1}{2}e^{\frac{1}{2}(u-v)^2} = e^{\frac{1}{2}(u^2+v^2)}\frac{e^{uv}+e^{-uv}}{2}.$$

(ii) Evaluate the joint moment-generating function from (i) at $u = 0$:

$$\mathbb{E}\big[e^{vY}\big] = \mathbb{E}\big[e^{(0)X+vY}\big] = e^{\frac{1}{2}((0)^2+v^2)}\frac{e^{(0)v}+e^{-(0)v}}{2} = e^{\frac{1}{2}v^2}.$$

(iii) Theorem 2.2.7 states, among other things, that random variables X and Y are independent if and only if their moment-generating functions factor as $\mathbb{E}[e^{uX+vY}] = \mathbb{E}[e^{uX}]\mathbb{E}[e^{vY}]$. Factoring a portion of $\mathbb{E}[e^{uX+vY}]$ to resemble $\mathbb{E}[e^{vY}]$, it is clear the remainder is *not* $\mathbb{E}[e^{uX}]$:

$$\mathbb{E}[e^{uX+vY}] = e^{\frac{1}{2}v^2}e^{\frac{1}{2}u^2}\frac{e^{uv}+e^{-uv}}{2} = \mathbb{E}[e^{vY}]\underbrace{e^{\frac{1}{2}u^2}\frac{e^{uv}+e^{-uv}}{2}}_{\neq\mathbb{E}[e^{uX}]}.$$

Thus, we conclude X and Y are not independent.

Question 5

Let (X, Y) be a pair of random variables with joint density function

$$f_{X,Y}(x, y) = \begin{cases} \frac{2|x|+y}{\sqrt{2\pi}} \exp\left\{-\frac{(2|x|+y)^2}{2}\right\} & \text{if } y \geq -|x| \\ 0 & \text{if } y < -|x|. \end{cases}$$

Show that X and Y are standard normal random variables and that they are uncorrelated but not independent.

Answer

We recover the density of X by integrating out Y. First, consider

$$f_X(x) = \int f_{X,Y}(x, y)\,\mathrm{d}y = \int_{-|x|}^{\infty} \frac{2|x|+y}{\sqrt{2\pi}} \exp\left\{-\frac{(2|x|+y)^2}{2}\right\} \mathrm{d}y = \frac{1}{\sqrt{2\pi}} \int_{|x|}^{\infty} ue^{-\frac{u^2}{2}}\,\mathrm{d}u.$$

The last equality above follows from the substitution $u = 2|x|+y$. This integral evaluates to

$$f_X(x) = -\frac{1}{\sqrt{2\pi}} e^{-v}\Big|_{\frac{1}{2}|x|^2}^{\infty} = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2},$$

which is the density of a standard normal random variable.

The process for recovering the marginal density of Y is the same, although a little more involved. Let $\mathbf{1}_A$ be the indicator function for the statement A, so that

$$f_Y(y) = \int f_{X,Y}(x, y)\,\mathrm{d}x = \int \mathbf{1}_{y \leq |x|} \frac{2|x|+y}{\sqrt{2\pi}} \exp\left\{-\frac{(2|x|+y)^2}{2}\right\} \mathrm{d}x.$$

We split this integral into two parts, based on the sign of y relative to zero:

$$f_Y(y) = \int_{-|x|}$$

Question 6

Consider a probability space Ω with four elements, which we call $a, b, c,$ and d (i.e. $\Omega = \{a, b, c, d\}$). The σ-algebra $\mathcal{F}$ is the collection of all subsets of Ω; i.e. $\mathcal{F} = 2^{\Omega}$. We define a probability measure $\mathbb{P}$ by specifying that

$$\mathbb{P}(a) = \tfrac{1}{6}, \quad \mathbb{P}(b) = \tfrac{1}{3}, \quad \mathbb{P}(c) = \tfrac{1}{4}, \quad \mathbb{P}(d) = \tfrac{1}{4},$$

and, as usual, the probability of every other set in $\mathcal{F}$ is the sum of the probabilities of the elements in the set, e.g. $\mathbb{P}(a, b, c) = \mathbb{P}(a) + \mathbb{P}(b) + \mathbb{P}(c) = 3/4$.

We next define two random variables, X and Y, by the formulas

$$\begin{aligned} &X(a) = 1, \quad X(b) = 1, \quad &X(c) = -1, \quad X(d) = -1, \\ &Y(a) = 1, \quad Y(b) = -1, \quad &Y(c) = 1, \quad Y(d) = -1. \end{aligned}$$

We then define $Z = X + Y$.

(i) List the sets in $\sigma(X)$.
(ii) Determine $\mathbb{E}[Y \mid X]$ (i.e. specify the values of this random variable for a, b, c, and d). Verify the partial-averaging property is satisfied.
(iii) Determine $\mathbb{E}[Z \mid X]$. Again verify the partial-averaging property.
(iv) Compute $\mathbb{E}[Z \mid X] - \mathbb{E}[Y \mid X]$. Citing the appropriate properties of conditional expectation from Theorem 2.3.2, explain why you get X.

Answer

(i) The σ-algebra generated by X is $\sigma(X) = \big\{\varnothing, \Omega, \{a, b\}, \{c, d\}\big\}$.

(ii) Given the σ-algebra specified in (i), evaluating the expectation of Y conditional on a or b is equivalent to $\mathbb{E}[Y \mid X = 1]$; it is the same for $\{c, d\}$ and $\mathbb{E}[Y \mid X = -1]$. Hence,

$$\mathbb{E}\big[Y \mid X = 1\big] = \tfrac{1}{\mathbb{P}(X=1)}\left[(1)\mathbb{P}(a) + (-1)\mathbb{P}(b)\right] = \tfrac{1}{1/2}\left[\tfrac{1}{6} - \tfrac{1}{3}\right] = -\tfrac{1}{3}$$

and

$$\mathbb{E}\big[Y \mid X = -1\big] = \tfrac{1}{\mathbb{P}(X=-1)}\left[(1)\mathbb{P}(c) + (-1)\mathbb{P}(d)\right] = \tfrac{1}{1/2}\left[\tfrac{1}{4} - \tfrac{1}{4}\right] = 0.$$

The law of total expectation implies

$$\begin{aligned}\mathbb{E}[Y] &= \mathbb{E}[Y \mid X = 1]\mathbb{P}(X = 1) + \mathbb{E}[Y \mid X = -1]\mathbb{P}(X = -1)\\ &= \left(-\tfrac{1}{3}\right)\left(\tfrac{1}{2}\right) + (0)\left(\tfrac{1}{2}\right) = -\tfrac{1}{2}.\end{aligned}$$

On the other hand, by direct evaluation of $\mathbb{E}[Y]$ we have

$$\mathbb{E}[Y] = (1)\mathbb{P}(a) + (-1)\mathbb{P}(b) + (1)\mathbb{P}(c) + (-1)\mathbb{P}(d) = (1)\tfrac{1}{6} + (-1)\tfrac{1}{3} = -\tfrac{1}{2}.$$

(iii) From the linearity of the expectations operator and the fact that X is measurable with respect to $\sigma(X)$, we have

$$\mathbb{E}\big[Z \mid X\big] = \mathbb{E}\big[Y + X \,||, X\big] = \mathbb{E}\big[Y \mid X\big] + \mathbb{E}\big[X \mid X\big] = \mathbb{E}\big[Y \mid X\big] + X.$$

Hence,

$$\mathbb{E}\big[Z \mid X = 1\big] = \mathbb{E}\big[Y \mid X = 1\big] + (1) = \tfrac{2}{3}$$

and

$$\mathbb{E}\big[Z \mid X = -1\big] = \mathbb{E}\big[Y \mid X = -1\big] + (-1) = -1.$$

By the law of total expectation and direct computation, respectively, we have

UNFINISHED

Question 7

Let Y be an integrable random variable on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathcal{G}$ eb a sub-σ-algebra of $\mathcal{F}$. Based on the information in $\mathcal{G}$, we can form the estimate $\mathbb{E}[Y|\mathcal{G}]$ of Y and define the error of the estimation $\mathrm{Err} = Y - \mathbb{E}[Y|\mathcal{G}]$. This is a random variable with expectation zero and some variance $\mathrm{Var}(\mathrm{Err})$. Let X be some other $\mathcal{G}$-measurable random variable, which we can regard as another estimate of Y. Show that

$$\mathrm{Var}(\mathrm{Err}) \leq \mathrm{Var}(Y - X).$$

In other words, the estimate $\mathbb{E}[Y|\mathcal{G}]$ minimizes the variance of the error among all estimates based on the information in $\mathcal{G}$.

Answer

Define the constant $\mu = \mathbb{E}[Y - X)$, and note that the variance of $Y - X$ is then written as $\mathrm{Var}(Y - X) = \mathbb{E}\big[(Y - X - \mu)^2\big]$. Judiciously inserting a $\mathbb{E}[Y|\mathcal{G}] - \mathbb{E}[Y|\mathcal{G}]$ term, we further have

$$\mathbb{E}\big[(Y - X - \mu)^2\big] = \mathbb{E}\left[\big((Y - \mathbb{E}[Y|\mathcal{G}]) + (\mathbb{E}[Y|\mathcal{G}] - X - \mu)\big)^2\right].$$

Multiplying this out, we have

$$\begin{aligned}\mathrm{Var}(Y - X) &= \mathbb{E}\big[(Y - X - \mu)^2\big] \\ &= \mathbb{E}\big[(\mathrm{Err} - 0)^2\big] + \mathbb{E}\big[(\mathbb{E}[Y|\mathcal{G}] - X - \mu)^2\big] + 2\mathbb{E}\big[\mathrm{Err}(\mathbb{E}[Y|\mathcal{G}] - X - \mu)\big]\end{aligned}$$

Since the Err random variable has zero mean, the first term on the right-hand side above is $\mathrm{Var}(\mathrm{Err})$. The second term is weakly greater than zero, and we will show below that the last term is exactly zero.

Towards that end, we condition an inner expectation operator on $\mathcal{G}$ to arrive at

$$\begin{aligned}\mathbb{E}\big[\mathrm{Err}(\mathbb{E}[Y|\mathcal{G}] - X - \mu)\big] &= \mathbb{E}\left[\mathbb{E}\big[\mathrm{Err}(\mathbb{E}[Y|\mathcal{G}] - X - \mu)|\mathcal{G}\big]\right] \\ &= \mathbb{E}\left[\mathbb{E}\big[\mathrm{Err}|\mathcal{G}\big](\mathbb{E}[Y|\mathcal{G}] - X - \mu)\right].\end{aligned}$$

The final equality follows from the $\mathcal{G}$-measurability of X, the fact that $\mathbb{E}[Y|\mathcal{G}]$ is known conditioned on $\mathcal{G}$, and that μ is a constant. Turning to $\mathbb{E}[\mathrm{Err}|\mathcal{G}]$, we immediately have that

$$\mathbb{E}\big[\mathrm{Err}|\mathcal{G}\big] = \mathbb{E}\big[Y - \mathbb{E}[Y|\mathcal{G}]\,|\,\mathcal{G}\big] = \mathbb{E}[Y|\mathcal{G}] - \mathbb{E}[Y|\mathcal{G}] = 0.$$

Thus, the cross-term $\mathbb{E}[\mathrm{Err}(\mathbb{E}[Y|\mathcal{G}] - X - \mu)]$ evaluates to zero. This implies that

$$\mathrm{Var}(Y - X) = \mathrm{Var}(\mathrm{Err}) + \underbrace{\mathbb{E}\big[(\mathbb{E}[Y|\mathcal{G}] - X - \mu)^2\big]}_{\geq 0} + 0.$$

Because the second term is weakly greater than zero, we can conclude that

$$\mathrm{Var}(\mathrm{Err}) \geq \mathrm{Var}(Y - X),$$

and we are done.

Question 8

Let X and Y be integrable random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then $Y = Y_1 + Y_2$, where $Y_1 = \mathbb{E}[Y|X]$ is $\sigma(X)$-measurable and $Y_2 = Y - \mathbb{E}[Y|X]$. Show that Y_2 and X are uncorrelated. More generally, show that Y_2 is uncorrelated with every $\sigma(X)$-measurable random variable.

Question 9

Let X be a random variable.

(i) Given an example of a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a random variable X defined on this probability space, and a function f so that the σ-algebra generated by $f(X)$ is not the trivial σ-algebra $\{\varnothing, \Omega\}$ but is strictly smaller than the σ-algebra generated by X.

(ii) Can the σ-algebra generated by $f(X)$ ever be strictly larger than the σ-algebra generated by X?

Question 10

Let X and Y be random variables (on some unspecified probability space $(\Omega, \mathcal{F}, \mathbb{P})$), assume they have a joint density $f_{X,Y}(x, y)$, and assume $\mathbb{E}[|Y|] < \infty$. In particular, for every Borel subset C of $\mathbb{R}^2$, we have

$$\mathbb{P}\big((X, Y) \in C\big) = \int_C f_{X,Y}(x, y) dx dy.$$

In elementary probability, one learns to compute $\mathbb{E}[Y|X = x]$, which is a *nonrandom* function of the *dummy variable* x, by the formula

$$\mathbb{E}[Y \mid X = x] = \int_{-\infty}^{\infty} y f_{Y|X}(y|x) \, \mathrm{d}y,$$

where $f_{Y|X}(y|x)$ is the *conditional density* defined by

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x, y)}{f_X(x)}.$$

The denominator in this expression, $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, \eta) \, \mathrm{d}\eta$, is the *marginal density* of X, and we must assume it is strictly positive for every x. We introduce the symbol $g(x)$ for the function $\mathbb{E}[Y|X = x]$ defined above; i.e.,

$$g(x) = \int_{-\infty}^{\infty} y f_{Y|X}(y|x) \, \mathrm{d}y = \int_{-\infty}^{\infty} \frac{y f_{X,Y}(x, y)}{f_X(x)} \, \mathrm{d}y.$$

In measure-theoretic probability, conditional expectation is a *random variable* $\mathbb{E}[Y|X]$. This exercise is to show that when there is a joint density for (X, Y), this random variable can be obtained by substituting the random variable X in place of the dummy variable x in the function $g(x)$. In other words, this exercise is to show that

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

(We introduced the symbol $g(x)$ in order to avoid the mathematically confusing expression $\mathbb{E}[Y|X=X]$.)

Since $g(X)$ is obviously $\sigma(X)$-measurable, to verify that $\mathbb{E}[Y|X]=g(X)$, we need only check that the partial-averaging property is satisfied. For every Borel-measurable function h mapping $\mathbb{R}$ to $\mathbb{R}$ and satisfying $\mathbb{E}[|h(X)|]<\infty$, we have

$$\mathbb{E}[h(X)] = \int_{-\infty}^{\infty} h(x) f_X(x)\,\mathrm{d}x.$$

This is Theorem 1.5.2 in Chapter 1. Similarly, if h is a function of both x and y, then

$$\mathbb{E}[h(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y) f_{X,Y}(x,y)\,\mathrm{d}x\mathrm{d}y$$

whenever (X,Y) has a joint density $f_{X,Y}(x,y)$. You may use both expressions, $\mathbb{E}[h(X)]$ and $\mathbb{E}[h(X,Y)]$, in your solution to this problem.

Let A be a set in $\sigma(X)$. By the definition of $\sigma(X)$, there is a Borel subset B of $\mathbb{R}$ such that $A=\{\omega\in\Omega\,|\,X(\omega)\in B\}$ or, more simply, $A=\{X\in B\}$. Show the partial-averaging property

$$\int_A g(X)\,\mathrm{d}\mathbb{P} = \int_A Y\,\mathrm{d}\mathbb{P}.$$

Question 11

(i) Let X be a random variable on a probability space $(\Omega,\mathcal{F},\mathbb{P})$, and let W be a nonnegative $\sigma(X)$-measurable random variable. Show there exists a function g such that $W=g(X)$.

(ii) Let X be a random variable on a probability space $(\Omega,\mathcal{F},\mathbb{P})$, and let Y be a nonnegative random variable on this space. We do not assume that X and Y have a joint density. Nonetheless, show there is a function g such that $\mathbb{E}[Y|X]=g(X)$.