Martingale Theory
Problem set 2, with solutions
Conditional expectation

2.1 Prove directly from the definition the following basic properties of the conditional expectation. (Numbering corresponds to that in the handwritten lecture notes.)

- (1) $\mathbf{E}\big(X \mid \mathcal{T}\big) = \mathbf{E}\big(X\big)$, $\mathbf{E}\big(X \mid \mathcal{F}\big) = X$;
- (2) linearity;
- (3) positivity;
- (10) the Tower Rule;
- (11) the "take out what you know" rule;

SOLUTION:

□

2.2 Let $(\Omega, \mathcal{F}, \mu)$ be a probability space and $X, Y \in L^1(\Omega, \mathcal{F}, \mu)$. Prove that if $\mathbf{E}\big(X \mid \sigma(Y)\big) = Y$ and $\mathbf{E}\big(Y \mid \sigma(X)\big) = X$ then $X = Y$ a.s.

SOLUTION:
Let $c \in \mathbb{R}$ be fixed.

$$\mathbf{E}\big(X \mathbb{1}_{\{Y \le c\}}\big) = \mathbf{E}\big(\mathbb{1}_{\{Y \le c\}} \mathbf{E}\big(X \mid \sigma(Y)\big)\big) = \mathbf{E}\big(Y \mathbb{1}_{\{Y \le c\}}\big),$$
$$\mathbf{E}\big(Y \mathbb{1}_{\{X \le c\}}\big) = \mathbf{E}\big(\mathbb{1}_{\{X \le c\}} \mathbf{E}\big(Y \mid \sigma(X)\big)\big) = \mathbf{E}\big(X \mathbb{1}_{\{X \le c\}}\big).$$

Hence,

$$0 = \mathbf{E}\big((X - Y)\mathbb{1}_{\{Y \le c\}}\big) = \mathbf{E}\big((X - Y)\mathbb{1}_{\{X > c, Y \le c\}}\big) + \mathbf{E}\big((X - Y)\mathbb{1}_{\{X \le c, Y \le c\}}\big),$$
$$0 = \mathbf{E}\big((X - Y)\mathbb{1}_{\{X \le c\}}\big) = \mathbf{E}\big((X - Y)\mathbb{1}_{\{X \le c, Y > c\}}\big) + \mathbf{E}\big((X - Y)\mathbb{1}_{\{X \le c, Y \le c\}}\big),$$

and, equating the two

$$\mathbf{E}\big((X - Y)\mathbb{1}_{\{X > c, Y \le c\}}\big) = \mathbf{E}\big((X - Y)\mathbb{1}_{\{X \le c, Y > c\}}\big).$$

But, obviously

$$\mathbf{E}\big((X - Y)\mathbb{1}_{\{X > c, Y \le c\}}\big) \ge 0, \qquad \mathbf{E}\big((X - Y)\mathbb{1}_{\{X \le c, Y > c\}}\big) \le 0.$$

From the last two displays it follows that

$$\mathbf{E}\big((X-Y)\mathbb{1}_{\{X>c,Y\le c\}}\big)=0, \qquad \mathbf{E}\big((X-Y)\mathbb{1}_{\{X\le c,Y>c\}}\big).=0.$$

since this is true for any $c\in\mathbb{R}$, the assertion follows.

□

2.3 HW

Let $\Omega=\{-1,0,+1\}$, $\mathcal{F}=\mathcal{P}(\Omega)$ and $\mu(\{-1\})=\mu(\{0\})=\mu(\{+1\})=1/3$, and consider also the sub-σ-algebras

$$\mathcal{G}=\{\emptyset,\{-1\},\{0,+1\},\{-1,0,+1\}\}, \qquad \mathcal{H}=\{\emptyset,\{-1,0\},\{+1\},\{-1,0,+1\}\}.$$

Let $X:\Omega\to\mathbb{R}$, $X(\omega)=\omega$. Compute

$$\mathbf{E}\big(\mathbf{E}\big(X\mid\mathcal{G}\big)\mid\mathcal{H}\big) \quad\text{and}\quad \mathbf{E}\big(\mathbf{E}\big(X\mid\mathcal{H}\big)\mid\mathcal{G}\big).$$

SOLUTION:

$$\begin{aligned}
\mathbf{E}\big(X\mid\mathcal{G}\big)(-1,0,1)&=(-1,1/2,1/2),\\
\mathbf{E}\big(X\mid\mathcal{H}\big)(-1,0,1)&=(-1/2,-1/2,1),\\
\mathbf{E}\big(\mathbf{E}\big(X\mid\mathcal{G}\big)\mid\mathcal{H}\big)(-1,0,1)&=(-1/4,-1/4,1/2),\\
\mathbf{E}\big(\mathbf{E}\big(X\mid\mathcal{G}\big)\mid\mathcal{H}\big)(-1,0,1)&=(-1/2,1/4,1/4).
\end{aligned}$$

□

2.4 HW

Let X_j, $j=1,2,\dots$ i.i.d. random variables with the common distribution $\mathbf{P}\big(X_j=-1\big)=\mathbf{P}\big(X_j=+1\big)=1/2$ and $S_n:=X_1+\dots+X_n$. Compute the following conditional expectations:

$$\mathbf{E}\big(X_1\mid\sigma(S_n)\big), \qquad \mathbf{E}\big(S_n\mid\sigma(X_1)\big), \qquad \mathbf{E}\big(S_{n+m}^2\mid\sigma(S_n)\big).$$

SOLUTION:

□

$$\begin{aligned}
\mathbf{E}\big(X_1\mid\sigma(S_n)\big)&=\frac{S_n}{n},\\
\mathbf{E}\big(S_n\mid\sigma(X_1)\big)&=X_1,\\
\mathbf{E}\big(S_{n+m}^2\mid\sigma(S_n)\big)&=S_n^2+m.
\end{aligned}$$

2.5 HW

Let $(\Omega, \mathcal{F}, \mu)$ be a probability space and $\mathcal{G} \subset \mathcal{F}$ a sub-σ-algebra of $\mathcal{F}$. Let $X \in L^2(\Omega, \mathcal{F}, \mu)$ and $Y := \mathbf{E}(X \mid \mathcal{G})$. Prove that if $\mathbf{E}(X^2) = \mathbf{E}(Y^2)$ then $X = Y$ a.s. and thus, X is $\mathcal{G}$-measurable.

SOLUTION:
Note first that Y and $X - Y$ are orthogonal:

$$\begin{aligned}
\mathbf{E}\big(Y(X-Y)\big) &= \mathbf{E}\big(\mathbf{E}(X \mid \mathcal{G})(X - \mathbf{E}(X \mid \mathcal{G}))\big) \\
&= \mathbf{E}\big(\mathbf{E}\big(\mathbf{E}(X \mid \mathcal{G})(X - \mathbf{E}(X \mid \mathcal{G})) \mid \mathcal{G}\big)\big) \\
&= \mathbf{E}\big(\mathbf{E}(X \mid \mathcal{G}) \underbrace{\mathbf{E}\big((X - \mathbf{E}(X \mid \mathcal{G})) \mid \mathcal{G}\big)}_{=0}\big) \\
&= 0.
\end{aligned}$$

Hence

$$\mathbf{E}(Y^2) = \mathbf{E}(X^2) = \mathbf{E}(Y^2) + \mathbf{E}\big((X-Y)^2\big),$$

and

$$\mathbf{E}\big((X-Y)^2\big) = 0.$$

Thus $X = Y$ almost surely. □

2.6 Bonus

[Change of conditional expectation]
Let ν and μ be two probability measures on $(\Omega, \mathcal{F})$, with $\nu \ll \mu$, and Radon-Nikodym derivative $\frac{d\nu}{d\mu}(\omega) = \varrho(\omega)$. Let $\mathcal{G} \subseteq \mathcal{F}$ a sub-σ-algebra. Show that, for any $\mathcal{F}$-measurable random variable X, we have

$$\mathbf{E}_\nu(X \mid \mathcal{G}) = \frac{\mathbf{E}_\mu(\varrho X \mid \mathcal{G})}{\mathbf{E}_\mu(\varrho \mid \mathcal{G})} \tag{1}$$

(a) First prove (1) for discrete probability space, by applying the elementary notion of conditional probability and conditional expectation.
(a) Then prove (1) for general case, by applying basic properties of the (general, measure theoretic notion of) conditional expectation.

SOLUTION:
(a) Assume that the σ-algebra $\mathcal{G}$ is discretely generated by the partition H_k, $k = 1, 2, \dots$, where $\mu(H_k) > 0$ for all k. Then

$$\mathbf{E}_\nu(X \mid H_k) = \frac{\int_{H_k} X d\nu}{\int_{H_k} d\nu} = \frac{\int_{H_k} X \varrho d\mu}{\int_{H_k} \varrho d\mu} = \frac{\int_{H_k} X \varrho d\mu}{\mu(H_k)} \cdot \frac{\mu(H_k)}{\int_{H_k} \varrho d\mu} = \frac{\mathbf{E}_\mu(\varrho X \mid H_k)}{\mathbf{E}_\mu(\varrho \mid H_k)}.$$

(b) Let Z be $\mathcal{G}$-measurable and bounded random variable. Then

$$\mathbf{E}_\nu\big(Z\mathbf{E}_\nu(X \mid \mathcal{G})\big) = \mathbf{E}_\nu\big(\mathbf{E}_\nu(ZX \mid \mathcal{G})\big) = \mathbf{E}_\nu(ZX) = \mathbf{E}_\mu(\varrho Z X).$$

On the other hand:

$$\mathbf{E}_\nu\big(Z\frac{\mathbf{E}_\mu(\varrho X \mid \mathcal{G})}{\mathbf{E}_\mu(\varrho \mid \mathcal{G})}\big) = \mathbf{E}_\mu\big(\varrho Z\frac{\mathbf{E}_\mu(\varrho X \mid \mathcal{G})}{\mathbf{E}_\mu(\varrho \mid \mathcal{G})}\big) = \mathbf{E}_\mu\big(\varrho\frac{\mathbf{E}_\mu(\varrho ZX \mid \mathcal{G})}{\mathbf{E}_\mu(\varrho \mid \mathcal{G})}\big)$$
$$= \mathbf{E}_\mu\big(\mathbf{E}_\mu(\varrho \mid \mathcal{G})\frac{\mathbf{E}_\mu(\varrho ZX \mid \mathcal{G})}{\mathbf{E}_\mu(\varrho \mid \mathcal{G})}\big) = \mathbf{E}_\mu(\varrho ZX).$$

We have proved that for any $\mathcal{G}$-measurable and bounded random variables Z,

$$\mathbf{E}_\nu\big(Z\mathbf{E}_\nu(X \mid \mathcal{G})\big) = \mathbf{E}_\nu\big(Z\frac{\mathbf{E}_\mu(\varrho X \mid \mathcal{G})}{\mathbf{E}_\mu(\varrho \mid \mathcal{G})}\big),$$

which implies (6).

□