0.1 Conditional Expectation

Let X be a random variable defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Further let $\mathcal{G}$ be a sub-$\sigma-$algebra. If X is $\mathcal{G}$ measurable then the information in $\mathcal{G}$ is enough to determine the value of X. If X is independent of $\mathcal{G}$ then the information in $\mathcal{G}$ provides no help in determining X. In the intermediate case we can use the information in $\mathcal{G}$ to estimate X but not precisely evaluate X. The conditional expectation of X given $\mathcal{G}$ is such an estimate.

Conditional Expectation:- Conditional probability

$$\mathbb{P}(A/B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}, B \in \mathcal{F} \text{ with } \mathbb{P}(B) > 0.$$

Consider the random experiment tossing a coin three times. Then

$$\Omega = \Big\{HHH, HTH, HHT, HTT, THT, TTH, THH, TTT\Big\}.$$

$$X - \text{denotes the number of heads.}$$

$$\text{Range}(X) = \{0, 1, 2, 3\}.$$

$$\mathbb{P}(X = 0) = \frac{1}{8}, \mathbb{P}(X = 1) = \frac{3}{8} = \mathbb{P}(X = 2), \mathbb{P}(X = 3) = \frac{1}{8}.$$

$$\mathbb{E}[X] = \sum_{i=0}^{3} x_i \mathbb{P}(X = x_i) = 0 \cdot 1/8 + 1 \cdot 3/8 + 2 \cdot 3/8 + 3 \cdot 1/8 = \frac{3+6+3}{8} = 12/8 = 1 \cdot 5.$$

If the first toss is a head, then $X = 1, 2, 3$ and

$$\mathbb{E}[X|\text{first toss is a head}] = 1 \cdot 1/4 + 2 \cdot 1/2 + 3 \cdot 1/4 = 2.$$

Exercise:

$$\mathbb{E}[X|\text{first toss is a tail}] = 1.$$

Expectation changes as new information becomes available. If X is a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$, where Ω is finite, $\mathcal{F} = 2^{\Omega}$ (power set of Ω), $\mathbb{P}(\{\omega\}) > 0 \; \forall \omega \in \Omega$. For $A \in \mathcal{F}$, $0 < \mathbb{P}(A) < 1$, define

$$\mathbb{E}[X|A] := \sum_{\omega \in \Omega} X(\omega) \mathbb{P}(\{\omega\}/A),$$

where

$$\mathbb{P}(\omega/A) = \frac{\mathbb{P}(\{\omega\} \cap A)}{\mathbb{P}(A)} = \begin{cases} \frac{\mathbb{P}(\{\omega\})}{\mathbb{P}(A)} & \text{if } \omega \in A \\ 0 & \text{if } \omega \notin A. \end{cases}$$

Hence

$$\mathbb{E}[X|A] = \sum_{\omega \in A} X(\omega) \frac{\mathbb{P}(\{\omega\})}{\mathbb{P}(A)} = \frac{1}{\mathbb{P}(A)} \int_A X d\mathbb{P}.$$

similarly,

$$\mathbb{E}[X|A^c] = \sum_{\omega \in A^c} X(\omega) \frac{\mathbb{P}(\{\omega\})}{\mathbb{P}(A^c)} = \frac{1}{\mathbb{P}(A^c)} \int_{A^c} X d\mathbb{P}.$$

$$\sigma(A) = \{\phi, A, A^c, \Omega\}.$$

$$\mathbb{E}[X|\sigma(A)] = ?.$$

Here $\mathbb{E}[X|\sigma(A)]$:$\Omega \to \mathbb{R}$ be a random variable, if $\omega \in B \in \sigma(A)$ with $\mathbb{P}(B) > 0$, then

$$\mathbb{E}[X|\sigma(A)](\omega) = \frac{1}{\mathbb{P}(B)} \int_B X d\mathbb{P} = \begin{cases} \frac{1}{\mathbb{P}(A)} \int_A X d\mathbb{P} & \text{if } w \in A \\ \frac{1}{\mathbb{P}(A^c)} \int_{A^c} X d\mathbb{P} & \text{if } w \in A^c. \end{cases} \tag{1}$$

Definition:- Let $(\Omega, \mathcal{F}, \mathbb{P})$ denote a probability space and let $\mathcal{G}$ be a σ-algebra on Ω generated by a countable partition $(G_n)_{n=1}^{\infty}$ of Ω. Suppose $\mathcal{G} \subset \mathcal{F}$ and $\mathbb{P}(G_n) > 0$ for all n. If X is an integrable random variable on $(\Omega, \mathcal{F}, \mathbb{P})$, define

$$\mathbb{E}[X|\mathcal{G}](\omega) = \frac{1}{\mathbb{P}(G_n)} \int_{G_n} X d\mathbb{P} \text{ for all } n \text{ and all } \omega \in G_n. \tag{2}$$

We call $\mathbb{E}[X|\mathcal{G}]$ the conditional expectation of X given $\mathcal{G}$. If $\mathcal{G}$ is generated by a random variable Y on $(\Omega, \mathcal{F}, \mathbb{P})$, we also write $\mathbb{E}[X|Y]$ in place of $\mathbb{E}[X|\sigma(Y)]$.

Note that $\omega \in \Omega \implies \omega \in G_n$ for some n

$$-\mathbb{E}[X|\mathcal{G}] : \Omega \to \mathbb{R} \text{ mapping.}$$

This mapping is constant on each G_n.

$-$Hence $\mathbb{E}[X|\mathcal{G}]$ is $\mathcal{G}-$ measurable.

$-\mathbb{E}[X|\mathcal{G}]$ is a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$.

1. $\mathbb{E}[X|\mathcal{G}]$ is integrable.

   If $\omega \in G_n$, then

$$\mathbb{E}[X|\mathcal{G}](\omega) = \frac{1}{\mathbb{P}(G_n)} \int_{G_n} X d\mathbb{P}.$$

   So,

$$|\mathbb{E}[X|\mathcal{G}](\omega)| \leq \frac{1}{\mathbb{P}(G_n)} \int_{G_n} |X| d\mathbb{P}$$

   and

$$\begin{aligned} \int_\Omega |\mathbb{E}[X|\mathcal{G}]| d\mathbb{P} &= \sum_{n=1}^{\infty} \int_{G_n} |\mathbb{E}[X|\mathcal{G}]| d\mathbb{P} \\ &= \sum_{n=1}^{\infty} |\mathbb{E}[X|\mathcal{G}]| \mathbb{P}(G_n) \\ &\leq \sum_{n=1}^{\infty} \int_{G_n} |X| d\mathbb{P} = \int_\Omega |X| d\mathbb{P} = \mathbb{E}[|X|]. \end{aligned}$$

   Therefore $\mathbb{E}[X|\mathcal{G}]$ is also integrable.

2. $$\int_A \mathbb{E}[X|\mathcal{G}] d\mathbb{P} = \int_A X d\mathbb{P} \; \forall A \in \mathcal{G}.$$

Let $\omega \in G_n$. Since $\mathbb{E}[X|\mathcal{G}](\omega)$ is constant on each G_n, therefore

$$\begin{aligned}\int_{G_n} \mathbb{E}[X|\mathcal{G}]d\mathbb{P} &= \mathbb{E}[X|\mathcal{G}](\omega)\int_{G_n} d\mathbb{P} \\ &= \left(\frac{1}{\mathbb{P}(G_n)}\int_{G_n} Xd\mathbb{P}\right)\mathbb{P}(G_n) \\ &= \int_{G_n} Xd\mathbb{P}.\end{aligned}$$

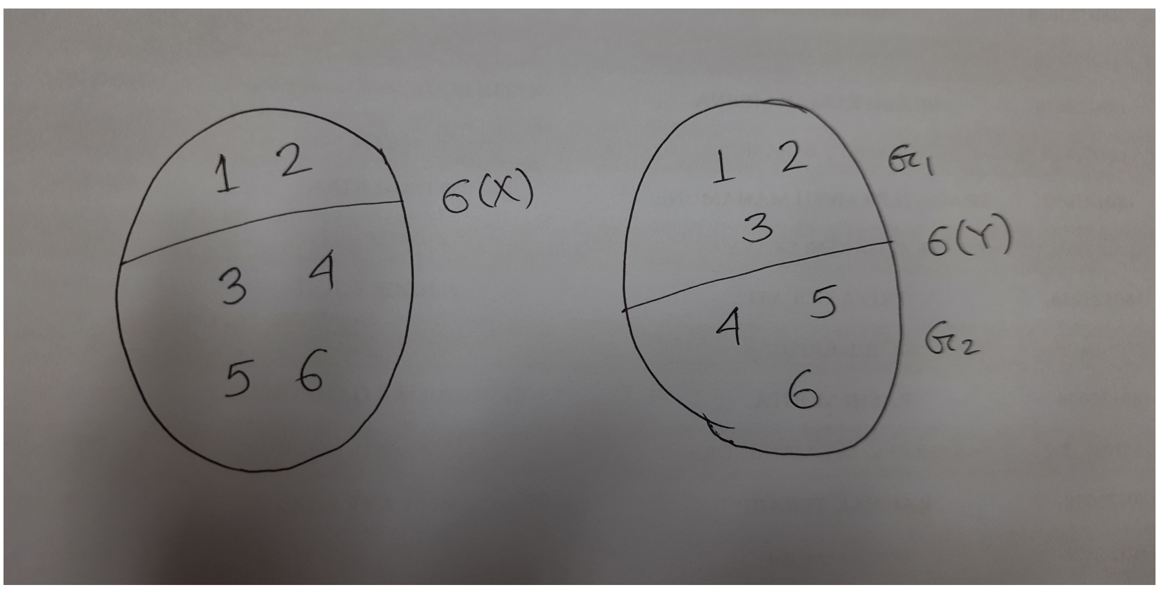
If $A \in \mathcal{G}$ then $A = \cup_{n\in M} G_n$ for some $M \subseteq \mathbb{N}$. Hence

$$\begin{aligned}\int_A \mathbb{E}[X|\mathcal{G}](\omega)d\mathbb{P} &= \int_{\cup_{n\in M}G_n} \mathbb{E}[X|\mathcal{G}]d\mathbb{P} = \sum_{n\in M}\int_{G_n} \mathbb{E}[X|\mathcal{G}]d\mathbb{P} \\ &= \sum_{n\in M}\int_{G_n} Xd\mathbb{P} = \int_{\cup_{n\in M}G_n} Xd\mathbb{P} = \int_A Xd\mathbb{P}.\end{aligned}$$

If $A = \Omega$, then $\mathbb{E}[\mathbb{E}[X|\mathcal{G}](\omega)] = \int_\Omega \mathbb{E}[X|\mathcal{G}]d\mathbb{P} = \int_\Omega Xd\mathbb{P} = \mathbb{E}[X]$.

Example 1: Let $\Omega = \{1,2,3,4,5,6\}$. $\mathcal{F} = 2^\Omega$, $\mathbb{P}(\{1\}) = \mathbb{P}(\{2\}) = 1/16$. $\mathbb{P}(\{3\}) = \mathbb{P}(\{4\}) = 1/4$ and $\mathbb{P}(\{5\}) = \mathbb{P}(\{6\}) = 3/16$. Suppose $X(1) = X(2) = 2$ and $X(3) = X(4) = X(5) = X(6) = 8$ and $Y = 4\mathbb{I}_{\{1,2,3\}} + 6\mathbb{I}_{\{4,5,6\}}$, calculate $\mathbb{E}[X|Y]$.

Solution: Note that $\mathbb{P}(\{1,2,3\}) = 3/8$, if $\omega \in G_1$ i.e., $\omega = 1$ or 2 or 3.



$$\mathbb{E}[X|Y](\omega) = \frac{1}{\mathbb{P}(\{1,2,3\})}\int_{\{1,2,3\}} Xd\mathbb{P} = 8/3[2\cdot 1/16 + 2\cdot 1/16 + 8\cdot 1/4] = 6.$$

Similarly, $\mathbb{E}[X|Y](\omega) = 8$ if $\omega \in G_2 = \{4,5,6\}$.

$$\mathbb{E}[X|Y](\omega) = 6\cdot\mathbb{I}_{\{1,2,3\}} + 8\cdot\mathbb{I}_{\{4,5,6\}}.$$

Example 2:- $\Omega = [0,1]$, $\mathcal{F} = \mathcal{B}([0,1])$.

$$\mathbb{P} - \text{Lebesgue measure on } [0,1].$$

$$X(\omega) = 2\omega^2 \text{ and } Y(\omega) = \begin{cases} 1 \text{ if } \omega \in [0, 1/3] \\ 2 \text{ if } \omega \in (1/3, 2/3) \\ 0 \text{ if } \omega \in [2/3, 1]. \end{cases} \tag{3}$$

Solution: Clearly, Y is a discrete random variable with the possible values 0,1,2. The corresponding events are $\{Y = 1\} = [0, 1/3]$, $\{Y = 2\} = (1/3, 2/3)$, $\{Y = 0\} = [2/3, 1]$. for $\omega \in [0, 1/3]$

$$\mathbb{E}[X|Y](\omega) = \mathbb{E}[X|[0, 1/3]] = \frac{1}{\frac{1}{3}} \int_0^{1/3} X(\omega) d\mathbb{P}(\omega) = 3 \int_0^{1/3} 2x^2 dx = 2/27.$$

For $\omega \in (1/3, 2/3)$

$$\mathbb{E}[X|Y](\omega) = 3 \int_{1/3}^{2/3} 2x^2 dx = 2[8/27 - 1/27] = 14/27.$$

Similarly, for $\omega \in [2/3, 1]$,

$$\mathbb{E}[X|Y](\omega) = 3 \int_{2/3}^{1} 2x^2 dx = 2[1 - 8/27] = 38/27.$$

So,

$$\mathbb{E}[X|Y](\omega) = \begin{cases} 2/27 \text{ if } \omega \in [0, 1/3] \\ 14/27 \text{ if } \omega \in (1/3, 2/3) \\ 38/27 \text{ if } \omega \in [2/3, 1]. \end{cases} \tag{4}$$

-If G is generated by the countable partition $\{G_n\}_{n=1}^{\infty}$ with $\mathbb{P}(G_n) > 0 \; \forall n$ and X is an integrable random variable on $(\Omega, \mathcal{F}, \mathbb{P})$, then for $\omega \in G$, we have

$$\begin{aligned} \mathbb{E}[X|\mathcal{G}](\omega) &= \text{Expected value of } X \text{ given that } G_n \text{ has occured} \\ &= \text{Expected value of } X \text{ with respect to the probability measure } \mathbb{P}(\cdot|G_n) \\ &= \text{Average value of } X \text{ over } G_n \\ &= \text{Average value of } \mathbb{E}[X|\mathcal{G}] \text{ over } G_n. \end{aligned}$$

When $\mathcal{G}$ is countable generated, we can interpret $\mathbb{E}[X|\mathcal{G}]$ pointwise. But this is not possible in the general case.

Definition 0.1. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\mathcal{G}$ be a sub-$\sigma-$algebra. Let X be a random variable which is either non-negative or integrable. The conditional expectation of X given $\mathcal{G}$ denoted by $\mathbb{E}[X|\mathcal{G}]$ is any random variable that satisfies*
i) (Measurability Property) $\mathbb{E}[X|\mathcal{G}]$ is $\mathcal{G}$ measurable.
ii) (Partial Averaging Property) $\int_A \mathbb{E}[X|\mathcal{G}] d\mathbb{P} = \int_A X d\mathbb{P}$, for all $A \in \mathcal{G}$.

If $\mathcal{G}$ is the σ-algebra generated by another random variable Y (i.e., $\mathcal{G} = \sigma(Y)$), then we write $\mathbb{E}[X|Y]$ instead of $\mathbb{E}[X|\sigma(Y)]$.

Uniqueness of conditional expectation: Suppose there are two random variables Y and Z which satisfy the above two properties. Since both Y and Z are $\mathcal{G}$ measurable the set $A = \{Y - Z > 0\} \in \mathcal{G}$. But by the second property,

$$\int_A Y d\mathbb{P} = \int_A Z d\mathbb{P} = \int_A X d\mathbb{P}.$$

Thus $\int_A (Y - Z) d\mathbb{P} = 0$, implies $\mathbb{P}(A) = 0$. Similarly, we can show that if $B = \{Z - Y > 0\}$, then $\mathbb{P}(B) = 0$. Thus we get $\mathbb{P}(Y = Z) = 1$.

An Important Consequence: Suppose X is integrable then, putting $A = \Omega$ in the partial averaging property we get

$$\mathbb{E}(X) = \mathbb{E}(\mathbb{E}[X|\mathcal{G}])\,.$$

Problem Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\mathcal{G}$ be a σ-algebra with $\mathcal{G} \subseteq \mathcal{F}$. If ξ is a $\mathcal{G}-$ measurable r.v. and $\int_B \xi d\mathbb{P} = 0\ \forall B \in \mathcal{G}$, then $\xi = 0$ a.s.

Proof. $0 \leq \varepsilon\mathbb{P}(\xi \geq \varepsilon) = \int_{\{\xi \geq \varepsilon\}} \varepsilon d\mathbb{P} \leq \int_{\{\xi \geq \varepsilon\}} \xi d\mathbb{P} = 0$, this imples $\mathbb{P}\{\xi \geq \varepsilon\} = 0$. Similarly, $\mathbb{P}\{\xi \leq -\varepsilon\} = 0$. Therefore $\mathbb{P}\{-\varepsilon < \xi < \varepsilon\} = 1$. Set $A_n = \{-1/n < \xi < 1/n\}$. See $A_n \downarrow$. $\{\xi = 0\} = \cap_{n=1}^{\infty} A_n$. So, $\mathbb{P}(\xi = 0) = \lim_{n\to\infty} \mathbb{P}(A_n) = 1$. $\square$

Theorem 0.2. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\mathcal{G}$ be a sub-$\sigma-$algebra.*

- *[1] (Linearity) Let X and Y be two random variables and let c_1 and c_2 be two constants, then $\mathbb{E}[c_1X + c_2Y|\mathcal{G}] = c_1\mathbb{E}[X|\mathcal{G}] + c_2\mathbb{E}[Y|\mathcal{G}]$.*
- *[2] (Taking out what is known) If X is $\mathcal{G}$ measurable then $\mathbb{E}[XY|\mathcal{G}] = X\mathbb{E}[Y|\mathcal{G}]$.*
- *[3] (Tower Property) If $\mathcal{H}$ is a sub-$\sigma-$algebra such that $\mathcal{H} \subset \mathcal{G}$, then $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}]$.*
- *[4] (Independence) If X is independent of $\mathcal{G}$ then $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}(X)$*
- *[5] (Conditional Jensen's Inequality) If φ is a convex function then $\varphi(\mathbb{E}[X|\mathcal{G}]) \leq \mathbb{E}[\varphi(X)|\mathcal{G}]$.*

Proof: Proof of [1]: $c_1\mathbb{E}[X|\mathcal{G}] + c_2\mathbb{E}[Y|\mathcal{G}]$ is $\mathcal{G}$-measurable, since $\mathbb{E}[X|\mathcal{G}]$ and $\mathbb{E}[Y|\mathcal{G}]$ are $\mathcal{G}$-measurable. Let $A \in \mathcal{G}$, then

$$\begin{aligned}
&\int_A (c_1\mathbb{E}[X|\mathcal{G}](\omega) + c_2\mathbb{E}[Y|\mathcal{G}](\omega))d\mathbb{P}(\omega) \\
&= c_1 \int_A \mathbb{E}[X|\mathcal{G}](\omega)d\mathbb{P}(\omega) + c_2 \int_A \mathbb{E}[Y|\mathcal{G}](\omega)d\mathbb{P}(\omega) \\
&= c_1 \int_A X(\omega)d\mathbb{P}(\omega) + c_2 \int_A Y(\omega)d\mathbb{P}(\omega) \\
&= \int_A (c_1X(\omega) + c_2Y(\omega))d\mathbb{P}(\omega).
\end{aligned}$$

Proof of [2]: Measurability is trivial. Suppose that $X = 1_B$ where $B \in \mathcal{G}$. Then for any $A \in \mathcal{G}$ we have

$$\begin{aligned}
&\int_A X\mathbb{E}[Y|\mathcal{G}]d\mathbb{P} = \int_A 1_B\mathbb{E}[Y|\mathcal{G}]d\mathbb{P} \\
&= \int_{A\cap B} \mathbb{E}[Y|\mathcal{G}]d\mathbb{P} = \int_{A\cap B} Y d\mathbb{P} = \int_A XY d\mathbb{P}\,.
\end{aligned}$$

Now use the standard machinery.
Proof of [3]: Again measurability is trivial. Now suppose $A \in \mathcal{H} \subset \mathcal{G}$. Then

$$\begin{aligned}
&\int_A \mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}]d\mathbb{P} = \int_A \mathbb{E}[X|\mathcal{G}]d\mathbb{P} \\
&= \int_A X d\mathbb{P} = \int_A \mathbb{E}[X|\mathcal{H}]d\mathbb{P}\,.
\end{aligned}$$

Hence the result.

Proof of [4]: $\mathbb{E}[X]$ is a constant and so is measurable w.r.t any σ-algebra. For partial averaging property first assume that $X = \mathbb{I}_B$, where B is independent of $\mathcal{G}$, then

$$\begin{aligned}\int_A \mathbb{E}[X]d\mathbb{P} &= \int_A \mathbb{P}(B)d\mathbb{P} = \mathbb{P}(B)\mathbb{P}(A) = \mathbb{P}(A\cap B)\\ &= \int_A \mathbb{I}_B d\mathbb{P} = \int_A X d\mathbb{P} = \int_A \mathbb{E}[X|\mathcal{G}]d\mathbb{P}.\end{aligned}$$

Now use the standard machinery.

Proof of [5]: Same as Jensen's inequality with conditional expectation replacing expectation. □

Exercise: If $X \geq 0$, then show that $\mathbb{P}(\mathbb{E}[X|\mathcal{G}] \geq 0) = 1$.

Theorem 0.3. $\mathbb{E}[X|\mathcal{G}]$ *is the best estimate (in terms of mean square error) of* X *given* $\mathcal{G}$*, i.e.,*

$$\mathbb{E}(X - \mathbb{E}[X|\mathcal{G}])^2 \leq \mathbb{E}(X-Y)^2\,,$$

for any Y*,* $\mathcal{G}$ *measurable.*

Proof:

$$\begin{aligned}\mathbb{E}[(X-Y)^2] &= \mathbb{E}[\mathbb{E}(X-Y)^2|\mathcal{G}]\\ &= \mathbb{E}[\mathbb{E}(X-\mathbb{E}[X|\mathcal{G}]+\mathbb{E}[X|\mathcal{G}]-Y)^2|\mathcal{G}]\\ &= \mathbb{E}[\mathbb{E}(X-\mathbb{E}[X|\mathcal{G}])^2|\mathcal{G}] + \mathbb{E}[(\mathbb{E}[X|\mathcal{G}]-Y)^2|\mathcal{G}] + 2\mathbb{E}[\mathbb{E}([X|\mathcal{G}]-Y)(X-\mathbb{E}[X|\mathcal{G}))|\mathcal{G}]]\\ &\geq \mathbb{E}[X-\mathbb{E}[X|\mathcal{G}]]^2.\end{aligned}$$

Fact: Let X be a random variable defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let Y be another random variable which is $\sigma(X)$ measurable. Then there exists a Borel measurable function g such that $Y = g(X)$.

Consequence: $\mathbb{E}[Y|X] \doteq \mathbb{E}[Y|\sigma(X)] = g(X)$ for some g Borel measurable.

Example: Suppose X and Y have joint density $f_{X,Y}(\cdot)$. Further assume that the marginal $f_X(\cdot)$ is strictly positive for all x. Now define the function

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

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

Again measurability is trivial. Take $A \in \sigma(X)$. Then there exists B such that $A = \{\omega : X(\omega) \in B\}$. Then

$$\begin{aligned}\int_A g(X)d\mathbb{P} &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} 1_B(x)g(x)f_{X,Y}(x,y)dydx\\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} 1_B(x)\frac{\int_{-\infty}^{\infty} t f_{X,Y}(x,t)dt}{f_X(x)} f_{X,Y}(x,y)dydx\\ &= \int_{-\infty}^{\infty} 1_B(x)\frac{\int_{-\infty}^{\infty} t f_{X,Y}(x,t)dt}{f_X(x)}\big(\int_{-\infty}^{\infty} f_{X,Y}(x,y)dy\big)dx\\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} 1_B(x)t f_{X,Y}(x,t)dtdx\\ &= \int_{\{X\in B\}} Y d\mathbb{P} = \int_A Y d\mathbb{P}\,.\end{aligned}$$

Thus we are done.

Lemma 0.4. *(Independence Lemma) Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\mathcal{G}$ be a sub-$\sigma-$algebra. Suppose that the random variables $X_1, X_2, \ldots, X_k$ are $\mathcal{G}$ measurable and $Y_1, \ldots, Y_l$ are independent of $\mathcal{G}$. Then for a Borel measurable function $f : \mathbb{R}^{k+l} \to \mathbb{R}$ define*

$$g(x) = \mathbb{E}(f(x_1, \ldots, x_k, Y_1, \ldots, Y_l)) \,.$$

Then

$$\mathbb{E}[f(X_1, \ldots, X_k, Y_1, \ldots, Y_l)|\mathcal{G}] = g(X_1, \ldots, X_k) \,.$$

Definition 0.5. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let T be a fixed positive number and let $\mathcal{F}_t, 0 \le t \le T$ be a filtration of sub-$\sigma-$algebras of $\mathcal{F}$. Consider an adapted stochastic process $M(t), 0 \le t \le T$.*

1. *If $\mathbb{E}[M(t)|\mathcal{F}_s] = M(s)$ for all $0 \le s \le t \le T$ then we say that the process is a martingale.*

2. *If $\mathbb{E}[M(t)|\mathcal{F}_s] \ge M(s)$ for all $0 \le s \le t \le T$ then we say that the process is a sub-martingale.*

3. *If $\mathbb{E}[M(t)|\mathcal{F}_s] \le M(s)$ for all $0 \le s \le t \le T$ then we say that the process is a super-martingale.*

Thus if $M(t)$ is a martingale then $\mathbb{E}(M(t)) = \mathbb{E}(\mathbb{E}[M(t)|\mathcal{F}_0]) = \mathbb{E}(M(0))$ for all $t \ge 0$.

Definition 0.6. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let T be a fixed positive number and let $\mathcal{F}_t, 0 \le t \le T$ be a filtration of sub-$\sigma-$algebras of $\mathcal{F}$. Consider an adapted stochastic process $X(t), 0 \le t \le T$. If for all $0 \le s \le t \le T$ and for every non-negative Borel measurable function f, there exists another Borel measurable function g such that*

$$\mathbb{E}[f(X(t))|\mathcal{F}_s] = g(X(s)) \,,$$

then we say that X is a Markov process.