

Recall: We consider the special case where X, Y have common density $f_{X,Y}(x,y)$:

$$P((X,Y) \in A) = \iint_A f_{X,Y}(x,y) dx dy.$$

We are interested in $E(X|Y)$ and define the conditional density

$$f_{X|Y}(x|y) = \begin{cases} \frac{f_{X,Y}(x,y)}{f_Y(y)} & \text{if } f_Y(y) \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

$$\text{where } f_Y(y) = \int_{\mathbb{R}} f_{X,Y}(x,y) dx.$$

$$\text{Now set } g(y) = \int_{\mathbb{R}} x f_{X|Y}(x|y) dx$$

("the expected value of X given $Y=y$ ").

We want to show that $g(y)$ satisfies the conditions of conditional expectations.

By inspection, g is $\sigma(Y)$ measurable and integrable.

It remains to check (3): Let $A \in \Sigma_2$, then

$$\int_{\{Y \in A\}} X dP \stackrel{\text{WTS}}{=} \int_{\{Y \in A\}} g(Y) dP$$

$$\int_{\Omega} X I_{\{Y \in A\}} dP$$

$$\int_{\Omega} g(Y) I_{\{Y \in A\}} dP$$

"

"

$$\int_{\mathbb{R} \times \mathbb{R}} x I_{\{Y \in A\}} f_{X,Y}(x,y) dx dy$$

$$\int_{\mathbb{R} \times \mathbb{R}} g(y) I_{\{Y \in A\}} f_{X,Y}(x,y) dx dy$$

$$\int_{\mathbb{R}} x \int_A f_{X,Y}(x,y) dy dx$$

" (*)

$$\int_{\mathbb{R}} \int_A g(y) f_{X,Y}(x,y) dy dx$$

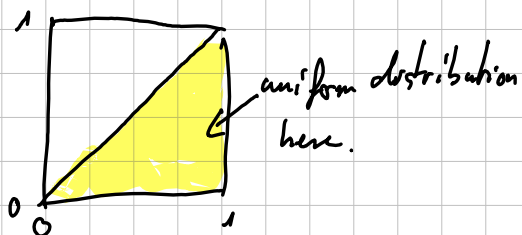
" (+)

By definition, $f_{X,Y}(x,y) = f_{X|Y}(x|y) f_Y(y)$ for Lebesgue almost every value. Then,

$$\begin{aligned}
 (*) &= \int_{\mathbb{R}} x \int_A f_{X|Y}(x|y) f_Y(y) dy dx \\
 &= \int_A f_Y(y) \int_{\mathbb{R}} x f_{X|Y}(x|y) dx dy \\
 &= \int_A f_Y(y) g(y) dy \\
 &= \int_A g(y) \int_{\mathbb{R}} f_{X,Y}(x,y) dx dy \\
 &= \int_{\mathbb{R}} \int_A g(y) f_{X,Y}(x,y) dy dx = (†)
 \end{aligned}$$

Which concludes the proof as A was arbitrary \square

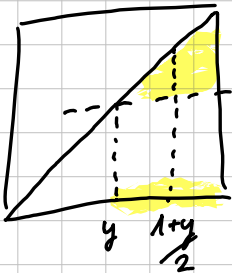
Example: Consider random variables X, Y on $[0,1] \times [0,1]$ with density $f(x,y) = 2 \mathbb{I}_{\{x \geq y\}}$



We get $f_Y(y) = \int_0^1 2 I_{\{x \geq y\}} dx = 2 \int_y^1 dx = 2(1-y)$

$$f_{X|Y}(x|y) = \frac{2 I_{\{x \geq y\}}}{2(1-y)} = \frac{I_{\{x \geq y\}}}{1-y}$$

$$g(y) = \int_0^1 x \frac{I_{\{x \geq y\}}}{1-y} dx = \frac{1}{1-y} \int_y^1 x dx = \frac{1}{1-y} \cdot \frac{1-y^2}{2} = \frac{1+y}{2} \Rightarrow E(X|Y=y) = \frac{1+y}{2}$$



Example: Let X_1, X_2, \dots, X_n be i.i.d. random variables. Let

$S_n = X_1 + \dots + X_n$. What is $E(X_1 | S_n)$?

By symmetry $E(X_1 | S_n) = E(X_2 | S_n) = \dots = E(X_n | S_n)$ a.s.

We must have

$$\int_A E(X_1 | S_n) dP + \dots + \int_A E(X_n | S_n) dP = \int_A X_1 dP + \dots + \int_A X_n dP = \int_A S_n dP$$

for all $A \in \sigma(S_n)$. Hence

$$n \int_A E(X_1 | S_n) dP = \int_A S_n dP \quad \text{and}$$

$$E(X_1 | S_n) = \frac{S_n}{n} \quad \text{a.s.}$$

Remark: We define cond. probabilities through cond. expectations:

$$P(A|G) = E(I_A | G) \quad (\text{Now a random variable!!})$$

This is uniquely determined (a.s.) and satisfies

$$P\left(\bigcup_{n \geq 1} A_n | G\right) = \sum_{n \geq 1} P(A_n | G) \quad \text{a.s.}$$

if A_1, A_2, \dots are disjoint.

If $G = \sigma(B)$ is generated by an event B ,

then $P(A|G)$ is a random variable Y with

$$Y(\omega) = \begin{cases} a & \omega \in B \\ b & \omega \in B^c \end{cases} \quad (G\text{-measurability}).$$

$$\text{We get } a P(B) = \int_B Y dP = \int_B I_A dP = \int I_{A \cap B} dP = P(A \cap B)$$

$$\Rightarrow a = \frac{P(A \cap B)}{P(B)} = P(A|B) \quad \text{and similarly,}$$

$$b = \frac{P(A \cap B)}{P(A)} = P(B|A).$$

Independent Random Variables

If X_1, X_2, \dots, X_n are independent, then
$$\mathbb{E}(h(X_1, X_2, \dots, X_n) | X_1) = g(X_1)$$
with $g(x) = \mathbb{E}(h(x, X_2, \dots, X_n))$.

This follows from Fubini's Theorem:

Since $\sigma(X_1)$ is generated by $\{X \in A\}$, $A \in \mathcal{B}(\mathbb{R})$
we only need to consider $\int_{\{X \in A\}} dP$.

$$\begin{aligned} & \int_{\{X_1 \in A\}} h(X_1, X_2, \dots, X_n) dP \\ &= \int_{X_1 \in A} \int_{\mathbb{R}^{n-1}} h(X_1, X_2, \dots, X_n) d(\mu_1 \times \mu_2 \times \dots \times \mu_n) \quad \begin{array}{l} \text{law} \\ \downarrow \end{array} \\ & \quad \quad \quad \text{(by independence)} \\ &= \int_{\mathbb{R}^n} \int_{\{X_1 \in A\}} h(X_1, \dots, X_n) d(\mu_2 \times \mu_3 \times \dots \times \mu_n \times \mu_1) \quad \text{(Fubini)} \\ &= \int_{\mathbb{R}} \int_{\{X_1 \in A\}} \int_{\mathbb{R}^{n-1}} h(X_1, \dots, X_n) d(\mu_2 \times \dots \times \mu_n) d\mu_1 \\ &= \int_{\{X_1 \in A\}} g(X_1) d\mu_1. \quad \text{So } g(X_1) \text{ satisfies (3).} \end{aligned}$$

Example: Let X_1, \dots, X_n be independent and write $\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$. What is $E(\bar{X} | X_1)$?

Take $h(x_1, \dots, x_n) = \frac{x_1 + \dots + x_n}{n}$, then,

$$g(x) = E\left(\frac{x + X_2 + \dots + X_n}{n}\right) = \frac{x + E(X_2) + \dots + E(X_n)}{n}$$

$$\text{and } E(\bar{X} | X_1) = \frac{X_1 + E(X_2) + \dots + E(X_n)}{n}.$$

Martingales

Stochastic processes & filtrations

A (discrete) stochastic process is a sequence X_0, X_1, X_2, \dots of random variables.

A filtration is a sequence of σ -algebras

$$\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \dots \subseteq \mathcal{F}_n \subseteq \dots \subseteq \mathcal{F}$$

We write $\mathcal{F}_\infty = \sigma\left(\bigcup_{n=0}^{\infty} \mathcal{F}_n\right) \subseteq \mathcal{F}$.

The process X_0, X_1, \dots is said to be adapted to the filtration (\mathcal{F}_n) if X_n is \mathcal{F}_n measurable.

A ^{super-}martingale is a sequence X_0, X_1, X_2, \dots
_{sub-}

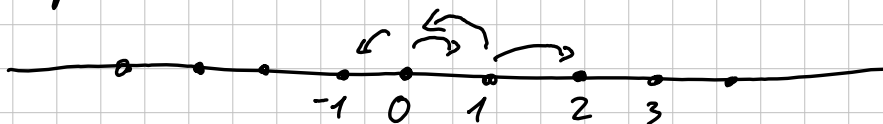
adapted to the filtration $\tilde{\mathcal{F}}_0, \tilde{\mathcal{F}}_1, \dots$ such that

$$\mathbb{E}(X_n | \tilde{\mathcal{F}}_{n-1}) \stackrel{\leq}{=} X_{n-1}.$$

Equivalently, in terms of increments,

$$\mathbb{E}(X_n - X_{n-1} | \tilde{\mathcal{F}}_{n-1}) = \mathbb{E}(X_n | \tilde{\mathcal{F}}_{n-1}) - X_{n-1} \stackrel{\leq}{=} 0$$

Example: Standard random walk



Let Y_1, Y_2, \dots be independent random variables with

$$P(Y_i = 1) = P(Y_i = -1) = \frac{1}{2}$$

and set $X_0 = 0, X_n = Y_1 + Y_2 + \dots + Y_n$

With $\tilde{\mathcal{F}}_n = \sigma(Y_1, Y_2, \dots, Y_n)$, the sequence X_0, X_1, \dots is adapted and

$$\mathbb{E}(X_n | \tilde{\mathcal{F}}_{n-1}) = \mathbb{E}(X_{n-1} + Y_n | \tilde{\mathcal{F}}_{n-1})$$

$$= \mathbb{E}(X_{n-1} | \tilde{\mathcal{F}}_{n-1}) + \mathbb{E}(Y_n | \tilde{\mathcal{F}}_{n-1}) = X_{n-1} + \mathbb{E}(Y_n) = X_{n-1}.$$

This also works for any other Y_i with $\mathbb{E}(Y_i) = 0$.

Example: Let Y_1, Y_2, \dots be independent random variables with $\mathbb{E}(Y_i) = 1$.

Set $X_n = X_0 \prod_{i=1}^n Y_i$. This is a martingale

since $\mathbb{E}(X_n | \tilde{\mathcal{F}}_{n-1}) = \mathbb{E}(X_{n-1} \cdot Y_n | \tilde{\mathcal{F}}_{n-1})$ (wrt $\tilde{\mathcal{F}}_n = \sigma(Y_1, \dots, Y_n)$)

$$= X_{n-1} \mathbb{E}(Y_n | \tilde{\mathcal{F}}_{n-1}) = X_{n-1} \mathbb{E}(Y_n) = X_{n-1}.$$

Example: Let X be a fixed $\tilde{\mathcal{F}}$ -meas. random variable. Fix a filtration $(\tilde{\mathcal{F}}_n)$ and set $X_n = \mathbb{E}(X | \tilde{\mathcal{F}}_n)$. This is a martingale.

$$\begin{aligned} \mathbb{E}(X_n | \tilde{\mathcal{F}}_{n-1}) &= \mathbb{E}(\mathbb{E}(X | \tilde{\mathcal{F}}_n) | \tilde{\mathcal{F}}_{n-1}) \\ &= \mathbb{E}(X | \tilde{\mathcal{F}}_{n-1}) = X_{n-1} \end{aligned}$$

Remark: Let $m < n$. For every martingale, we have

$$\begin{aligned} \mathbb{E}(X_n | \tilde{\mathcal{F}}_m) &= \mathbb{E}(\mathbb{E}(\dots \mathbb{E}(\underbrace{\mathbb{E}(X_n | \tilde{\mathcal{F}}_{n-1})}_{= X_{n-1}} | \tilde{\mathcal{F}}_{n-2}) \dots | \tilde{\mathcal{F}}_m) | \tilde{\mathcal{F}}_m) \\ &= \mathbb{E}(\mathbb{E}(\dots \underbrace{\mathbb{E}(X_{n-1} | \tilde{\mathcal{F}}_{n-2})}_{X_{n-2}} \dots | \tilde{\mathcal{F}}_{m+1}) | \tilde{\mathcal{F}}_m) \\ &\quad \vdots \\ &= \mathbb{E}(X_{m+1} | \tilde{\mathcal{F}}_m) = X_m. \end{aligned}$$

Definition A pre-visible process is a sequence C_1, C_2, \dots of random variables, such that C_n is \mathcal{F}_{n-1} -measurable for all n .

Let C_n be a pre-visible process. The martingale transform of X by C is

$$(C \cdot X)_n = \sum_{k=1}^n C_k (X_k - X_{k-1}).$$

In particular, if $C_k = 1$ for all k , then $(C \cdot X)_n = X_n - X_0$.

Proposition: If C is a bounded pre-visible process with $|C_n(\omega)| \leq K$ for all n and $\omega \in \Omega$, then

$(C \cdot X)_n$ is a martingale if X_n is.

If C is also non-negative, then $(C \cdot X)_n$ is a sub-/super martingale whenever X_n is.

Proof: We have

$$\begin{aligned} & E((C \cdot X)_n - (C \cdot X)_{n-1} \mid \mathcal{F}_{n-1}) \\ &= E(C_n(X_n - X_{n-1}) \mid \mathcal{F}_{n-1}) \\ &= C_n(E(X_n \mid \mathcal{F}_{n-1}) - E(X_{n-1} \mid \mathcal{F}_{n-1})) \\ &= C_n(E(X_n \mid \mathcal{F}_{n-1}) - X_{n-1}) \end{aligned}$$

$$\begin{cases} = 0 & \text{if } X_n \text{ is a martingale} \\ \geq 0 & \text{if } C_n \geq 0 \text{ and } X_n \text{ is a submartingale} \\ \leq 0 & \text{if } C_n \geq 0 \text{ and } X_n \text{ is a supermartingale} \end{cases}$$

□