① σ-algebras : $\mathcal{F}$
② Conditional $E(X|\mathcal{F})$, $Var(X|\mathcal{F})$

Inform. definition:

$\sigma(X)$ – σ-algebra generated by random variable X – the list of all the events that can be stated in terms of X.

Problem 1.

e₁) X is measurable w.r.t $\sigma(X,Y)$.
e₂) Y is not measur. w.r.t $\sigma(X)$

	$X=0$	$X=1$
$Y=0$	0.1	0.2
$Y=1$	0.1	0.3
$Y=2$	0.2	0.1

a) $\sigma(X)$?
b) $\sigma(Y)$?
c) $\sigma(X+Y)$?
d) how many elements are in $\sigma(X,Y)$?

$\sigma(X)$: $\{X>\frac{1}{2}\} = \{X=1\} = \{X \geq \frac{1}{3}\}$

$\{X=0\} = \{X<\frac{1}{2}\} = \{X \text{ is even}\} =$
$= \{\sin X = 0\}$

2 trivial events:

$\{X>5\} = \emptyset$
$\{X>-1\} = \Omega$ (all the outcomes)

$\sigma(X) = \{\{X=0\}, \{X=1\}, \emptyset, \Omega\}$

4

	$X=0$	$X=1$
$Y=0$	0.1	0.2
$Y=1$	0.1	0.3
$Y=2$	0.2	0.1

$\sigma(Y) =$

$\{\underbrace{\emptyset, \Omega}_{\text{trivial}}, \{Y=0\}, \{Y=1\}, \{Y=2\}, \{Y \neq 0\}, \{Y\neq 1\}, \{Y \neq 2\}\}.$

$\text{card } \sigma(Y) = 8$

$\{Y > \frac{1}{2}\}, \{Y > 1.5\}, \{Y < 1.5\},$

$\{Y \neq 1\} = \{Y \text{ is even}\} = \{Y\cdot(Y-2)=0\} =$

$= \{\sin(Y\cdot\frac{\pi}{2}) = 0\} = \dots$

$\sigma(X+Y)$ Intuition: I know (I observe) $(X+Y)$

I know that some events have occurred and some events have not.

$\{X+Y=0\}, \{X+Y=1\}, \{X+Y=2\}$

$S = X+Y$

$\sigma(X+Y) = \{\emptyset, \Omega, \{S=0\}, \{S=1\}, \{S=2\},$
$\{S=3\}, \{S\in\{0,1\}\}, \{S\in\{0,2\}\}, \{S\in\{0,3\}\},$
$\{S\in\{1,2\}\}, \{S\in\{1,3\}\}, \{S\in\{2,3\}\},$
$\{S\neq 0\}, \{S\neq 1\}, \{S\neq 2\}, \{S\neq 3\}\}$

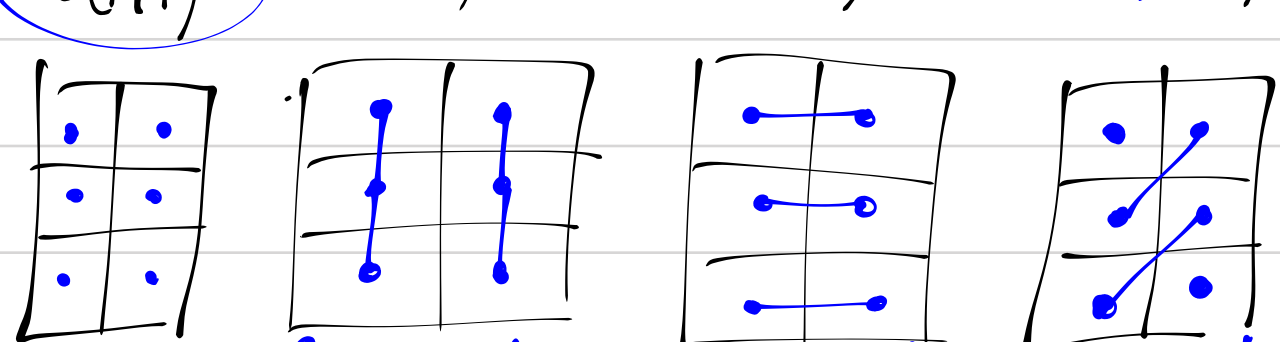
$\text{card } \sigma(X+Y) = 16$ events $\quad S\in\{0,1,2,3\}$

$\{S>1\} = \{S\in\{2,3\}\} = \{(S-2)\cdot(S-3)=0\}$

d) How many events are in $\sigma(X,Y)$?

	$X=0$	$X=1$
$Y=0$	0.1	0.2
$Y=1$	0.1	0.3
$Y=2$	0.2	0.1

$\sigma(X,Y)$ $\quad$ $\sigma(X)$ $\quad$ $\sigma(Y)$ $\quad$ $\sigma(X+Y)$



$2^6 = 64$ $\quad$ $2^2 = 4$ events $\quad$ $2^3 = 8$ events $\quad$ $2^4 = 16$ events

$\mathcal{F} = \sigma(X,Y) = \{ \ldots$ big list of 64 events $\ldots \}$

$\{X=1\} \in \mathcal{F}$ $\{Y=0\} \in \mathcal{F}$

$\{X>Y\} \in \mathcal{F}$ $\{Y=2\cdot X\} \in \mathcal{F}$

...

$\{Y$ is divisible by $X\} \in \mathcal{F}$

....

$\sigma(X)$ formal definition | $\sigma(X,Y,Z)$

①. List (collection) of events. !!

②. I have enough information to calculate the value of X — X, Y, Z

$\forall t \quad \{X \le t\} \in$ list — $\forall t \quad \{X \le t\} \in$ list, $\{Y \le t\} \in$ list, $\{Z \le t\} \in$ list

I can compare X with any t

③. I can deduce! !!

If $A_1, A_2, A_3 \ldots\ldots \in$ list then

$A_1 \cup A_2 \cup A_3 \ldots\ldots \in$ list

$A_1 \cap A_2 \cap A_3 \ldots\ldots \in$ list

$\bar{A}_i = A_i^c \in$ list

$(A_1 \cup A_2)^c = A_1^c \cap A_2^c$

$A_i \Delta A_j, \; A_i \setminus A_j \in$ list

④ the list $\sigma(X)$ is the smallest list with properties ①–③. !!

why do we need these keys?

① $E(Y \mid \sigma(X)) \neq E(Y \mid \sigma(Y))$

forecast of Y if I know X. — forecast of Y if I know Y

② Rigorous foundations of prob-ty theory.

def. Random variable Y is measureable with respect to $\mathcal{G}$-algebra $\mathcal{F}$ if $\underline{\mathcal{G}(Y) \subseteq \mathcal{F}}$.

intuition $\mathcal{F}$ contains enough information to calculate the value of Y.

$\boxed{\begin{array}{l} \text{Q.A.} \quad A \subseteq B \\ \quad\quad\quad\quad \text{subset or equal} \\ A = \{a, b, c\} \quad B = \{a, b, c, d, e\} \end{array}}$

def event A and event B are independent
$$P(A \cap B) = P(A) \cdot P(B)$$

$$\boxed{\begin{array}{l} P(A|B) = P(A) \\ P(B|A) = P(B) \end{array}}$$

def list $\mathcal{F}$ and list $\mathcal{H}$ of events are independent if $\forall A \in \mathcal{F}$ and $\forall B \in \mathcal{H}$
$$P(A \cap B) = P(A) \cdot P(B).$$

Y - random variable $\quad\quad$ $\mathcal{G}(Y)$ - list of events.

def Random variables X and Y are independent if $\mathcal{G}(X)$ and $\mathcal{G}(Y)$ are independent.

$\{X^2 + X^3 > 5\}$ $\quad\quad$ $\{Y \text{ is divisible by } 7\}$

def (in books):

$\mathcal{F}$ is σ-algebra if.

$$\boxed{\begin{array}{ll} ① & \emptyset \in \mathcal{F} \\ ② & \text{If } A \in \mathcal{F} \text{ then } A^c \in \mathcal{F} \\ ③ & \text{If } A_1, A_2 \ldots\ldots \in \mathcal{F} \text{ then } \\ & \qquad A_1 \cup A_2 \cup A_3 \cup \ldots \in \mathcal{F}. \end{array}}$$

$\Omega = \mathbb{R}$ $\quad X \sim N(7; 42)$

[def] $\mathcal{B}$ - Borel σ-algebra $= \sigma(X)$

= the smallest σ-algebra that contains all subsets of real line of the form $(-\infty; t]$.

$\to$ $E(X) = \text{const}$ $\qquad$ X - random variable

$\to$ $E(X \mid A) = \text{const}$ $\qquad$ A - event

$\to$ $\boxed{E(X \mid \mathcal{F})} = \boxed{\text{Random variable}}$ $\qquad$ $\mathcal{F}$ - σ-algebra.

$$\boxed{E(X) \begin{cases} E(X) = \sum_x x \cdot P(X = x) & \text{if } X \text{ is discrete} \\ E(X) = \int_{-\infty}^{\infty} x \cdot f(x)\, dx & \text{if } X \text{ has pdf } f(x) \end{cases}}$$

$$\boxed{E(X \mid A) \begin{cases} \sum x \cdot P(X = x \mid A) \\ \int_{-\infty}^{\infty} x \cdot f(x \mid A) \cdot dx \end{cases}}$$

[Problem]

	$X=0$	$X=1$	$X=2$	
$P(X=x)$	0.2	0.3	0.5	$\sum_x P(X=x) = 1$

$E(X \mid X > 0) \overset{?}{=}$

$P(X=x \mid A)$	0	$\frac{0.3}{0.8}$	$\frac{0.5}{0.8}$	$\sum_x P(X=x \mid A) = 1$

$A = \{X > 0\}$

$= 0 \cdot 0 + 1 \cdot \frac{0.3}{0.8} + 2 \cdot \frac{0.5}{0.8}$

$E(X|\mathcal{F})$

X - random variable
$\mathcal{F}$ - σ-algebra.

intuition behind

① the „best" point forecast of X given info in $\mathcal{F}$.

Problem

discrete case

X=0	X=1	X=2	X=3
Y=0	Y=7	Y=5	Y=5
0,1	0,4	0.2	0,3

$E(X|\sigma(Y)) \leftarrow$ best forecast of X if we know Y.

$Y=0 \Rightarrow E(X|\sigma(Y)) = 0$

$Y=7 \Rightarrow E(X|\sigma(Y)) = 1$

$Y=5 \Rightarrow E(X|\sigma(Y)) = f. = 2.6$

„best" - minimizes expected squared error of forecast.

	X=0	X=1	X=2	X=3
	Y=0	Y=7	Y=5	Y=5
$p(w)$	0,1.	0,4	0,2	0,3
$p(w\|Y=5)$	0	0	$\frac{0,2}{0,5}$	$\frac{0,3}{0,5}$

$$\min_f \boxed{(2-f)^2 \cdot \frac{0,2}{0,5} + (3-f)^2 \cdot \frac{0,3}{0,5}}$$

$$-2\cdot(2-f)\cdot\frac{0,2}{0,5} - 2\cdot(3-f)\cdot\frac{0,3}{0,5} = 0$$

$$\boxed{f = 2\cdot\frac{0,2}{0,5} + 3\cdot\frac{0,3}{0,5}} = E(X|Y=5)$$

$= 2\cdot 0,4 + 3\cdot 0,6 = 0.8 + 1.8 = 2.6$

$$E(X \mid \sigma(Y)) = \begin{cases} 0, & \text{if } Y = 0 \\ 1 & \text{if } Y = 7 \\ 2.6 & \text{if } Y = 5 \end{cases}$$ ← random variable

$$E(X \mid \sigma(X)) = X$$

① def $E(X)$ . . . . [using Lebesgue integral]

② def $E(X \mid \mathcal{F})$ - - - - (!)

③ def $E(X \mid Y) := E(X \mid \sigma(Y))$

motivation

u_t is indep. of $\mathcal{F}_{t-1}$

AR(1) $\quad y_t = 0{,}7\, y_{t-1} + u_t \quad$ y_t - stationary

$E(y_{t+2} \mid \mathcal{F}_t)$

$\mathcal{F}_t = \sigma(y_t, y_{t-1}, y_{t-2}, y_{t-3}, \ldots)$

$$E(y_{t+2} \mid \boxed{\mathcal{F}_t}) = E(0{,}7\, y_{t+1} + u_{t+2} \mid \mathcal{F}_t) =$$

$$= 0{,}7 \cdot E(y_{t+1} \mid \mathcal{F}_t) = 0{,}7 \cdot (0.7 \cdot y_t)$$

10 minutes break !!

Case 2 (X,Y) has joint pdf $f(x,y)$

Step 1. $f(x|y) = \frac{f(x,y)}{f(y)}$

step 2. $E(X|Y=y) = \int_{-\infty}^{\infty} x \cdot f(x|y)\,dx = a(y)$

Step 3. $E(X|\sigma(Y)) = a(Y)$

General case. $\hat{X} = E(X|\mathcal{F})$

X – Random variable

$\mathcal{F}$ – σ-algebra

① $\hat{X}$ – random variable

② Info in $\mathcal{F}$ is enough to calculate $\hat{X}$

$\boxed{\sigma(\hat{X}) \subseteq \mathcal{F}}$

③ „Best" point forecast

(3.1) $\boxed{E(\hat{X}) = E(X)}$

„on average the forecast is right"

$E(X-\hat{X})=0$

(3.2) You can't predict the error of the forecast.

$\mathrm{Cov}(\underbrace{X-\hat{X}}, R) = 0$ if $\sigma(R) \subseteq \mathcal{F}$

error of the forecast

$X-\hat{X}$ = is the forecasting error.

$\sigma(R) \subseteq \mathcal{F} \iff$ I know the value of R if I know info in $\mathcal{F}$.

3.1.
$$E(X-\hat{X})=0$$

3.2.
for any random variable R such that $\sigma(R) \subseteq \mathcal{F}$

$$\boxed{Cov(X-\hat{X}, R)=0.}$$

$$1_A = \begin{cases} 1 & \text{if } A \text{ is TRUE} \\ 0 & \text{if } A \text{ is FALSE} \end{cases}$$ „Bernoulli"

let's take $R = 1_A$

$$0 = Cov(X-\hat{X}, 1_A) = E((X-\hat{X})\cdot 1_A) - \underbrace{E(X-\hat{X})}_{\overset{\|}{0}\,[3.1]}\cdot \underbrace{E(1_A)}_{P(A)} =$$
$$= E(X\cdot 1_A) - E(\hat{X}\cdot 1_A) = 0$$

in books: $E(X|\mathcal{F}) = \hat{X}$ such that.

② $\hat{X}$ is measureable wrt $\mathcal{F}$
$\sigma(\hat{X}) \subset \mathcal{F}$

③ $E(X\cdot 1_A) = E(\hat{X}\cdot 1_A)$ for all $A \in \mathcal{F}$

you can't predict the error of the forecast

$\Downarrow$

$\boxed{E(X|\mathcal{F})}$ def. $\Rightarrow$

const $\to$ $Var(X) = E(X^2) - (E(X))^2$

const $\to$ $P(A) = E(1_A)$

$Var(X|\mathcal{F}) = E(X^2|\mathcal{F}) - (E(X|\mathcal{F}))^2 \leftarrow$ random variable

$P(A|\mathcal{F}) = E(1_A|\mathcal{F}) \leftarrow$ random variable

Properties of $E(X|\mathcal{F})$

[may be proven from general definition]

① linearity

α, β - constant
X, Y - random variables

$$E(\alpha X + \beta Y | \mathcal{F}) = \alpha \cdot E(X|\mathcal{F}) + \beta \cdot E(Y|\mathcal{F})$$

② Extreme cases.

→ I know X

$$E(X|\mathcal{F}) = X \quad \text{if} \quad \sigma(X) \subseteq \mathcal{F}$$

→ I know „nothing"

$$\mathcal{F} = \{\phi, \Omega\}$$

(False) (True)

$$E(X|\mathcal{F}) = E(X)$$

→ "My info is irrelevant"

If X and $\mathcal{F}$ are independent then

$$E(X|\mathcal{F}) = E(X)$$

③ „Tower property." „Adam's law"

$$E(E(X|\mathcal{F})) = E(X)$$

Proof: Find this in the definition. !!

④ If $\mathcal{F} \subseteq \mathcal{H}$ are σ-algebras then

$$E(E(X|\mathcal{F})|\mathcal{H}) = E(X|\mathcal{F}) = E(E(X|\mathcal{H})|\mathcal{F})$$

$$E(E(X|\mathcal{F})|\mathcal{H}) = E(X|\mathcal{F}) = E(E(X|\mathcal{H})|\mathcal{F})$$

→ Fedor knows less

→ Hercule Poirot knows more

$\mathcal{F} \subseteq \mathcal{H}$

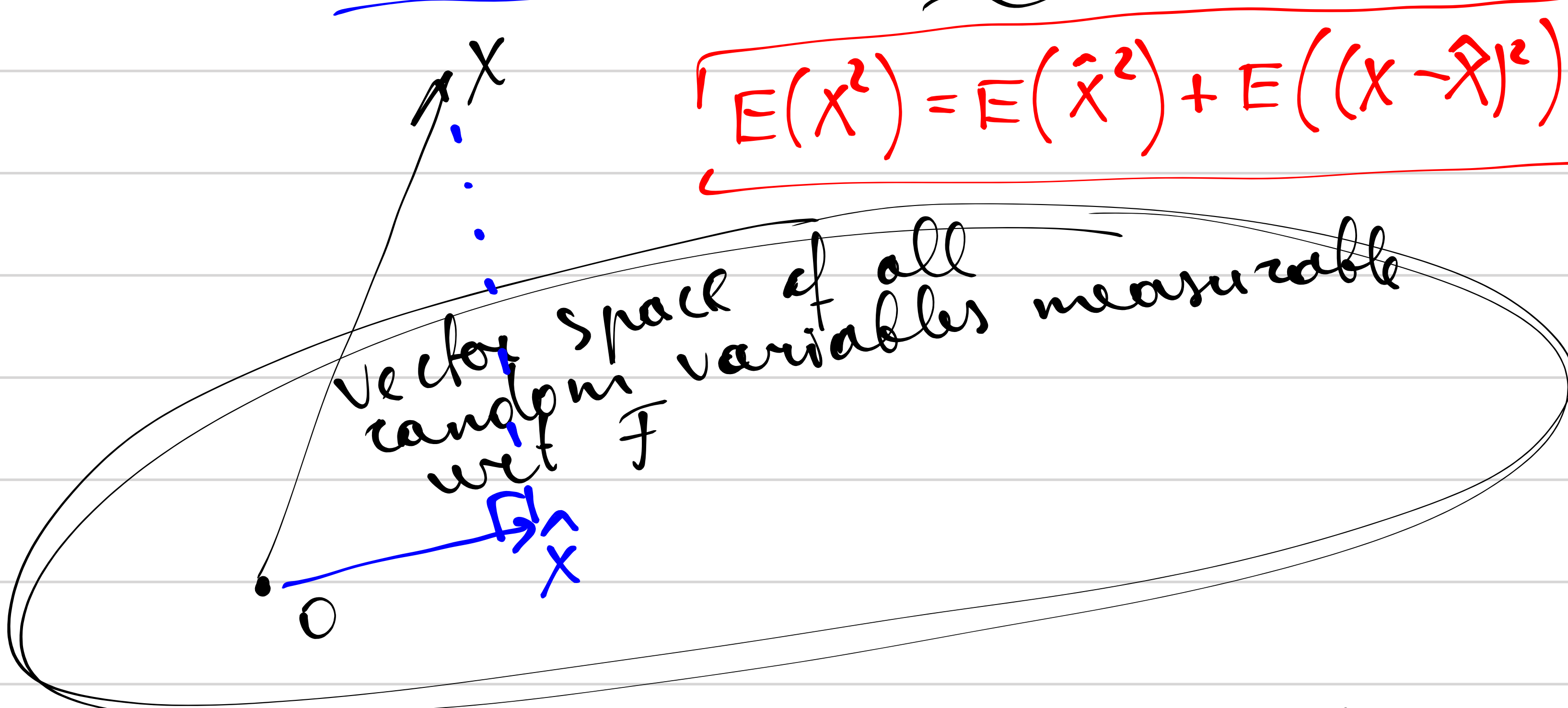
Hercule forecasts the forecast of X made by Fedor.

Fedor forecasts the forecast of X made by Hercule

Fedor forecasts X.

Geometry of $E(X|\mathcal{F})$

$$\boxed{E(X^2) = E(\hat{X}^2) + E((X-\hat{X})^2)}$$



R is measurable wrt $\mathcal{F}$

↳ $\sigma(R) \subseteq \mathcal{F}$

$\hat{X}$ the random variable that you can calculate given your info in $\mathcal{F}$ the closest to X

minimizes $E[(X-\hat{X})^2]$

→ Reincarnation of Pythagorean theorem

$$E(X^2) = E(\hat{X}^2) + E((X-\hat{X})^2)$$

$$E(X^2) - (E(X))^2 = E(\hat{X}^2) - (E(\hat{X}))^2 + E((X-\hat{X})^2)$$

$- (E(X))^2$

$- (E(\hat{X}))^2$

$$\mathrm{Var}(X) = \mathrm{Var}(\hat{X}) + E((X-\hat{X})^2)$$

$$\mathrm{Var}(X) = \mathrm{Var}(E(X|\mathcal{F})) + E(\mathrm{Var}(X|\mathcal{F}))$$

"EVE's law" / Pythagorean theorem

$$\boxed{\mathrm{Var}(X) = E(\mathrm{Var}(X|\mathcal{F})) + \mathrm{Var}(E(X|\mathcal{F}))}$$

Properties of $\mathrm{Var}(X|\mathcal{F})$:

1. $\mathrm{Var}(\alpha X + \beta | \mathcal{F}) = \alpha^2 \cdot \mathrm{Var}(X|\mathcal{F})$
2. extreme cases.
   - → I know X
     $\mathrm{Var}(X|\mathcal{F}) = 0$ if $\sigma(X) \subseteq \mathcal{F}$.
   - → I know nothing
     $\mathrm{Var}(X|\mathcal{F}) = \mathrm{Var}(X)$ if $\mathcal{F} = \{\emptyset, \Omega\}$
   - → My info is irrelevant
     $\mathrm{Var}(X|\mathcal{F}) = \mathrm{Var}(X)$ if X and $\mathcal{F}$ are independent.
3. Take out known values with square

$$\mathrm{Var}(X \cdot Y | \mathcal{F}) = X^2 \cdot \mathrm{Var}(Y|\mathcal{F}) \text{ if } \sigma(X) \subseteq \mathcal{F}.$$

$\boxed{\text{Look the past exams collection.}}$

One / two problems in every exam

$\boxed{\text{Ex.}}$

I throw a dice (standard) until I obtain 6.

$\boxed{X}$ - the result of the first throw

$\boxed{Y}$ - the number of throws.

$E(X|Y)$ $\quad [E(X|\sigma(Y))]$ $\quad$?

$Var(Y|X)$ $\quad [Var(Y|\sigma(X))]$ $\quad$?

Idea 1: $\boxed{X, Y \text{ are discrete}}$

$W_1, W_2, W_3 \ldots$ - throw results.

$W_1 = X$

$E(X|Y)$

$Y=1 \Rightarrow E(X|Y) \stackrel{?}{=} 6$!!

$Y=2 \Rightarrow \boxed{\text{not } 6}\ \boxed{6}$ $\quad E(X|Y) = \frac{1+2+\ldots+5}{5} = 3$

$Y=3 \Rightarrow \boxed{\text{not } 6}\ \boxed{\text{not } 6}\ \boxed{6}$ $\quad E(X|Y) =$

$Y=4$

$\vdots$

$$E(X|Y) = \begin{cases} 6 & \text{if } Y=1 \\ 3 & \text{if } Y>1 \end{cases}$$!!

$$E(X|Y) = 6 \cdot 1_{Y=1} + 3 \cdot 1_{Y>1} =$$!!

$$= 6 \cdot 1_{Y=1} + 3 \cdot (1 - 1_{Y=1}) =$$

$$= 3 + 3 \cdot 1_{Y=1}$$

$\boxed{Y=3}$ $\quad \boxed{1,2,3,4,5\ \cancel{6}}\ \boxed{1,2,3,4,5\ \cancel{6}}\ \boxed{6}$

$\| \\ X$

$$E(X|Y=3) = 1 \cdot \frac{1}{5} + 2 \cdot \frac{1}{5} + 3 \cdot \frac{1}{5} + 4 \cdot \frac{1}{5} + 5 \cdot \frac{1}{5} = 3$$

$$\frac{5+1}{2}$$

$Var(Y|X)$

$X=6$ $\quad Var(Y|X=6)=0$

$X=6 \Rightarrow Y=1$

$X=5$ conditional distribution of Y

$p(Y=y\|X=5)$	$Y=1$	$Y=2$	$Y=3$	....
	0	$\frac{1}{6}$	$\frac{5}{6}\cdot\frac{1}{6}$	

$Var(Y|X=5) =$ → boring way $E(Y^2|X=5)-(E(Y|X=5))^2$

→ recognize geometric distribution „+1"

geom.

$G=1$	$G=2$	$G=3$	$G=4$	
$\frac{1}{6}$	$\frac{5}{6}\cdot\frac{1}{6}$	$\left(\frac{5}{6}\right)^2\cdot\frac{1}{6}$	....	....

$Var(G) =$ [look at wiki] $= E(G^2)-(E(G))^2$

! [On the exam you can use a cheat sh A4]

Moment Generating function.

$$E(\exp(G\cdot t)) = \frac{1}{6}\cdot\exp(t)+\frac{5}{6}\cdot\frac{1}{6}\cdot\exp(2t)+\left(\frac{5}{6}\right)^2\cdot\frac{1}{6}\exp(3t) + \dots$$

$$= \boxed{\text{"}\frac{b_1}{1-q}\text{"}} = \frac{\frac{1}{6}\cdot\exp(t)}{1-\frac{5}{6}\cdot\exp(t)} = \frac{1}{6\exp(-t)-5}$$

$$E(\exp(Gt)) = \frac{1}{6\exp(-t)-5}$$

$\frac{d}{dt}$ ↓

$$E(G\cdot\exp(Gt)) = \frac{-1}{(6\exp(-t)-5)^2}\cdot 6\cdot\exp(-t)\cdot(-1)$$

$t=0$ $\quad E(G)=6$

$$E\left(G\cdot\exp(Gt)\right) = \frac{6\exp(-t)}{(6\exp(-t)-5)^2}$$

$$\Big\downarrow \frac{d}{dt}$$

$$E\left(G^2\exp(Gt)\right) = \left[\;\ldots\ldots\;\right]$$

$$E(G^2) = \left[\qquad\right]\Big|_{t=0}$$

$$\boxed{\mathrm{Var}(Y|X) = \begin{cases} 0 & \text{if } X = 6 \\ \mathrm{Var}(G) & \text{if } X < 6 \end{cases}}$$