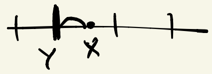
Conditioning on a random variable

Def Let X, Y be r.v's. $\mathbb{E}[X|Y] := \mathbb{E}[X|\underset{\text{r.v.}}{\sigma(Y)}]$ $= h(Y)$, function of Y.

allows to condition on events of measure 0!

- **Ex** $X \sim \mathrm{Unif}[0,1]$ $Y := [nX]$ (integer part). $\mathbb{E}[X|Y] = Y + \frac{1}{2}$ (check!)

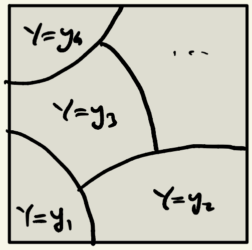


- Q: Suppose we want to use Y to predict X, how should we choose $h: \mathbb{R} \to \mathbb{R}$ so as to minimize $\mathbb{E}(X - h(Y))^2$?

  Ans: $h(Y) = \mathbb{E}[X|Y]$ is the best predictor (Cor p. 150)

- Formula for discrete distributions? ← SKIP

Let (X, Y) take on finitely many or countable # of values,



joint pmf: $p(x,y) := \mathbb{P}\{X = x, Y = y\}$.

$\sigma(Y) = \sigma\big(\{Y = y\} : y \text{ ranges over all values } Y \text{ takes}\big) \to \Omega$

$\mathbb{E}[X|Y] = \mathbb{E}[X|\sigma(Y)] = \mathbb{E}[X|Y=y]$ if $\{Y=y\}$ occurs (Ex (B) p. 144)

$$= \frac{\mathbb{E}[X \mathbb{1}_{\{Y=y\}}]}{\mathbb{P}\{Y=y\}} \quad \text{(def p. 143)}$$

↖ takes value x with prob. $p(x,y)$ $\forall x$

$$= \frac{\sum_x x\, p(x,y)}{\sum_{x'} p(x',y)} = \sum_x x\, p(x|y) \quad \text{where} \quad p(x|y) = \frac{p(x,y)}{\sum_{x'} p(x',y)}$$

$$\Rightarrow \boxed{\mathbb{E}[X|Y] = \sum_x x\, p(x|Y)}$$

HW: Prove

p. 57 2019

- Similarly for continuous distributions:

Prop If (X, Y) have joint pdf $f(x,y)$, then

$$\mathbb{E}[X|Y] = \int_{-\infty}^{\infty} x\, f(x|Y)\, dx \quad \text{where} \quad f(x|y) := \frac{f(x,y)}{\int_{-\infty}^{\infty} f(x',y)\, dx'}$$

Example 1 (Random sums)

Let $X_1, X_2, \dots$ be iid r.v's with mean μ, variance σ^2

Let $N \in \mathbb{N}$ be an indep. r.v. with finite variance.

Compute mean and variance of $S_N = \sum_{i=1}^{N} X_i$

- $\forall$ fixed $n \in \mathbb{N}$, $X_1, \dots, X_n$ are independent $\Rightarrow$

$$\mathbb{E}[S_N \mid N=n] = \mathbb{E}\Big[\sum_1^n X_i\Big] = \mu n. \qquad (*)$$

$$\underbrace{\operatorname{Var}(S_N \mid N=n)}_{\mathbb{E}[S_N^2 \mid N=n] - (\underbrace{\mathbb{E}[S_N \mid N=n]}_{\mu n})^2} = n \cdot \operatorname{Var}(X_1 \mid N=n) \underset{\text{indep}}{=} n \cdot \operatorname{Var}(X_1) = n\sigma^2.$$

$$\Rightarrow \mathbb{E}[S_N^2 \mid N=n] = n\sigma^2 + \mu^2 n^2 \qquad (**)$$

- LTP & $(*)$ $\Rightarrow$ $\mathbb{E}S_N = \mathbb{E}[\mu N] = \mu\,\mathbb{E}N.$ natural.

LTP & $(**)$ $\Rightarrow$ $\mathbb{E}[S_N^2] = \mathbb{E}[\sigma^2 N + \mu^2 N^2] = \sigma^2\,\mathbb{E}N + \mu^2\,\mathbb{E}[N^2]$

$$\Rightarrow \operatorname{Var}(S_N) = \mathbb{E}[S_N^2] - (\mathbb{E}S_N)^2 = \sigma^2\,\mathbb{E}N + \underline{\mu^2\,\mathbb{E}[N^2] - \mu^2(\mathbb{E}N)^2}$$

$$= \sigma^2\,\mathbb{E}N + \mu^2 \operatorname{Var}(N)$$

- This is a partial case of the law of total variance:

$$\operatorname{Var}(X) = \mathbb{E}[\operatorname{Var}(X \mid Y)] + \operatorname{Var}(\mathbb{E}[X \mid Y])$$

HW

Ex. 2 (An Exchange Paradox)

There are 2 closed envelops,
one containing twice as much money as the other.
You are asked to choose one of the envelops for yourself.
You pick an envelop at random, open it, observe the content.
You are given the option to exchange it for the other envelop.
Should you exchange or not?

$Y :=$ amount in the chosen envelop
$X :=$ amount in the other envelop

$\mathbb{E}[X] = \mathbb{E}[\,\mathbb{E}[X|Y]\,]$ (law of total expectation p. 149)

Given Y, X can be $2Y$ or $Y/2$, each with probability $1/2$ (*)

$\Rightarrow \mathbb{E}[X|Y] = 2Y \cdot \frac{1}{2} + \frac{Y}{2} \cdot \frac{1}{2} = 1.25\,Y$

$\Rightarrow \boxed{\mathbb{E}[X] = 1.25\,\mathbb{E}[Y]}$

$\Rightarrow$ you should always exchange! (and get 25% more money)

• Resolution of paradox (?)

Symmetry argument (*) is only justified if $\exists\ \infty$ resources.
In reality, if you observe a large amount Y,
you should conclude that the other envelop contains $Y/2$
more likely than $2Y$.