CONDITIONAL EXPECTATION

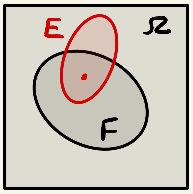
- Fix a prob. space (Ω, Σ, P) throughout.

1. Conditioning on an event

Def Let E, F be events. The conditional probability of E given F is

$$P(E|F) := \frac{P(E\cap F)}{P(F)}$$

- Interpetation:



$(F, \underbrace{\Sigma \cap F}_{= \{A\cap F \,:\, A \in \Sigma\}})$ is a new prob. space.

$E \mapsto P(E|F)$ is a new prob. measure on $(F, \Sigma\cap F)$ (check!)

- Ex:

	Cancer	No
Smoker	8	32
No	16	304

$$P(\text{Cancer} \,|\, \text{Smoke}) = \frac{8}{8+32} = 0.2$$

$$P(\text{Cancer} \,|\, \text{No smoke}) = \frac{16}{16+304} = 0.05$$

- Relation with independence:

$$E \perp\!\!\!\perp F \iff P(E|F) = P(E) \qquad \text{if } P(F) \neq 0.$$

- **Ex** Suppose you know that your friend has 2 children.
You saw one of them, and it was a girl.
What is the probability that the other child is also a girl?

$\Omega = \{GG, GB, BG, BB\}$ (older first), $P =$ uniform.

$F =$ "at least one child is a girl" $= \{GG, GB, BG\}$, $E =$ "both children are girls" $= \{GG\}$

$$\Rightarrow P(E|F) = \frac{1/4}{3/4} = \frac{1}{3}. \qquad ?! \quad \text{Why not } 1/2\,?$$

- Possibly add: ① Memoryless property of Exp, Geom? ② Bayes Formula?

Prop (Law of Total Probability) If $\Omega = F_1 \sqcup \dots \sqcup F_n$ ← n could be ∞ then

$$P(E) = \sum_i P(E|F_i)\cdot P(F_i) \quad \forall E \in \Sigma$$

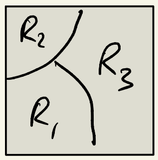
$$\ulcorner\; E = \bigsqcup_i (E \cap F_i) \Rightarrow P(E) = \sum P(E \cap F_i) \;\lrcorner$$

Bayes Formula

B.F. allows to swap the order of conditioning:

$$P(F|E) = \frac{P(E\cap F)}{P(E)} = \boxed{\frac{P(E|F)P(F)}{P(E)}} \underset{\uparrow \text{ L.T.P.}}{=} \frac{P(E|F)P(F)}{P(E|F)P(F)+P(E|F^c)P(F^c)}$$

Ex A hiker went missing in the wilderness.
She is one of the three regions with prob's 0.5, 0.3, 0.2.
Whenever anyone is lost in region 1,
the search is successful with prob. 0.4,
and for regions 2, 3 it is 0.7 and 0.8.
A search in region 1 was unsuccessful.
What is the prob. that the person is in region 1?

$S =$ 

$\ulcorner$ R_i = "the hiker is in region i",
U_i = "the search in region i is unsuccessful"

$$P(R_1|U_1) = \frac{P(U_1|R_1)P(R_1)}{P(U_1)} \underset{\nearrow \text{ L.T.P.}}{=} \frac{P(U_1|R_1)P(R_1)}{P(U_1|R_1)P(R_1)+P(U_1|R_2)P(R_2)+P(U_1|R_3)P(R_3)}$$

$$= \frac{(1-0.4)\,0.5}{(1-0.4)\,0.5 + 1\times 0.3 + 1\times 0.2} = 0.375 \;\lrcorner$$

- Similarly, $P(R_2|U_1) = \dfrac{P(U_1|R_2)P(R_2)}{\text{same}} = 0.375$

  $P(R_3|U_1) = \dfrac{P(U_1|R_3)P(R_3)}{\text{same}} = 0.25$

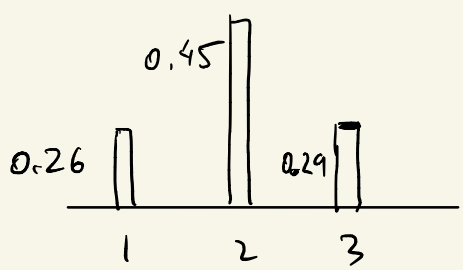
  $\Big\}$ Check!

- Thus, the <u>prior</u> probabilities $P(R_1), P(R_2), P(R_3)$ got updated to the <u>posterior</u> probabilities $P(R_1|U_1), P(R_2|U_1), P(R_3|U_1)$.

PRIOR (initial degree of belief) — POSTERIOR (degree of belief after incorporating new info U_1)

Ex Another search in reg. 1 is unsuccessful

$\Rightarrow$ prob's are further updated to



(Check!)

Remark Bayesian model of <u>learning</u>

(text $\to$ image generative AI, spam filtering, etc.)

Def Let X be a r.v. and F be an event.
The <u>conditional expectation</u> of X given F is

$$\mathbb{E}[X|F] := \frac{\mathbb{E}[X \mathbb{1}_F]}{\mathbb{P}(F)}$$

- <u>Ex</u>: X = lifetime, F: smoker $\Rightarrow$ $\mathbb{E}[X|F]$ = expected lifetime of a smoker.
- This def. is more general than the previous one, if we define

$$\mathbb{P}(E|F) := \mathbb{E}[\mathbb{1}_E | F].$$

So we will work with <u>conditional expectation</u> from now on.

More generally:

<u>2. Conditioning on a σ-algebra</u>

Def Let X be a r.v, and $\mathcal{F} \subset \Sigma$ be a σ-algebra.
The <u>conditional expectation of X given $\mathcal{F}$</u> is a r.v. Y satisfying

(i) Y is $\mathcal{F}$-measurable ← i.e. $\{Y \in B\} \in \mathcal{F}$ $\forall B \in \mathcal{B}$

(ii) $\mathbb{E}[Y \mathbb{1}_F] = \mathbb{E}[X \mathbb{1}_F]$ $\forall F \in \mathcal{F}$

Notation: $Y = \mathbb{E}[X|\mathcal{F}]$.

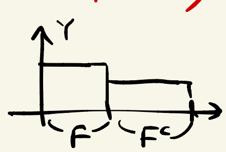
- NB: $\mathbb{E}[X|\mathcal{F}]$ is a <u>random variable</u>

<u>Examples</u>:

(a) Let $F \in \Sigma$ be ~~an~~ event, (e.g. "smoker") $\mathcal{F} := \sigma(F) = \{\emptyset, F, F^c, \Omega\}$

$$\Rightarrow \mathbb{E}[X|\mathcal{F}] \overset{Y :=}{=} \begin{cases} \mathbb{E}[X|F] & \text{if } F \text{ occurs} & \leftarrow \text{life expectancy of smokers} \\ \mathbb{E}[X|F^c] & \text{if } F \text{ does not occur} & \leftarrow \text{life expectancy of nonsmokers} \end{cases}$$

<u>Proof</u> (i) Y is constant on F and F^c



$\Rightarrow \forall B \in \mathcal{B}$, $\{Y \in B\} = Y^{-1}(B)$ is $\emptyset$, F, F^c or Ω $\Rightarrow \in \mathcal{F}$.

(ii) $Y\mathbb{1}_F = \begin{cases} \mathbb{E}[X|F] & \text{if } F \text{ occurs} \\ 0 & \text{if not} \end{cases}$

$$\Rightarrow \mathbb{E}[Y\mathbb{1}_F] = \underbrace{\mathbb{E}[X|F]}_{\overset{\text{def}}{=} \frac{\mathbb{E}[X \mathbb{1}_F]}{\mathbb{P}(F)}} \cdot \mathbb{P}(F) = \mathbb{E}[X \mathbb{1}_F].$$

Similarly for $F^c, \emptyset, \Omega$. □

(b) More generally, consider a partition

$$\Omega = F_1 \sqcup \cdots \sqcup F_n$$

and let $\mathcal{F} := \sigma(F_1, \dots, F_n)$. Then

$$\mathbb{E}[X|\mathcal{F}] = \mathbb{E}[X|F_i] \text{ if } F_i \text{ occurs.} \qquad \text{(Check!)}$$

e.g. $F_1 = \{\text{person's age} \in [0,10)\}$, $F_2 = \{\text{age} \in [10,20)\}$, ..., $F_{10} = \{\text{age} \in [90,100)\}$

X = person's height.

$\Rightarrow$ $\mathbb{E}[X|\mathcal{F}]$ takes on 10 values = ave hight in each group

(c) Same example, in the analytic form:

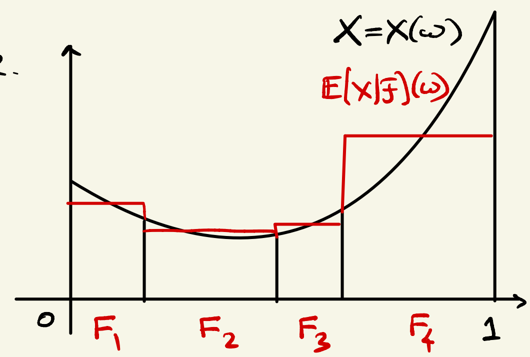
$\Omega = [0,1]$, $\Sigma = \mathcal{B}(\mathbb{R})$, P = Lebesgue measure.

$X : [0,1] \to \mathbb{R}$ an integrable function.

$$\mathbb{E}X = \int_0^1 X(\omega)\, d\omega$$

$$\mathbb{E}[X|F_i] = \frac{1}{|F_i|} \int_{F_i} X(\omega)\, d\omega$$

$$\mathbb{E}[X|\mathcal{F}](t) = \frac{1}{|F_i|} \int_{F_i} X(\omega)\, d\omega \quad \text{if } t \in F_i$$



Remarks

(a) For the coarsest σ-algebra $\mathcal{F} = \{\emptyset, \Omega\}$, $\mathbb{E}[X|\mathcal{F}] = \mathbb{E}(X)$

For the finest σ-algebra $\mathcal{F} = \Sigma$, $\mathbb{E}[X|\Sigma] = X$

Intermediate $\mathcal{F}$ $\Rightarrow$ interpolates between $\mathbb{E}[X]$ and X.

(b) $\mathcal{F}$ encodes available information (e.g. smoking habits, age, ...)

$\mathbb{E}[X|\mathcal{F}]$ encodes the best prediction of X given that info

No info $\Rightarrow$ $\mathbb{E}[X]$ All info $\Rightarrow$ X (exact)

(c) General def. of $\mathbb{E}[X|\mathcal{F}]$ allows us to condition on events of probability = 0 (e.g. ave height of a 20 y.o. person)

3. Existence, Uniqueness

THM $\forall$ integrable r.v. X and $\forall$ σ-algebra $\mathcal{F} \subset \Sigma$,

(i) $\mathbb{E}[X|\mathcal{F}]$ exists

(ii) and is unique: if Y, Y' are both cond. exp's of X given $\mathcal{F}$, then $Y = Y$ a.s.

Proof of Existence is based on:

THM (Radon-Nikodym Thm)

Let $(\Omega, \mathcal{F})$ be a measurable space. (Ω = countable union of mble sets with finite measures)

Let μ, ν be two σ-finite measures on (S, Σ)

Assume $\nu \ll \mu$ (absolutely continuous) ($\mu(F)=0 \Rightarrow \nu(F)=0 \; \forall F \in \mathcal{F}$)

Then $\exists$ $\mathcal{F}$-measurable function $f: S \to [0, \infty)$

s.t. $\nu(F) = \int_F f \, d\mu \quad \forall \; F \in \mathcal{F}$

$:= \int f \mathbb{1}_F \, d\lambda$

- WLOG $X \geq 0$ (otherwise decompose $X = X^+ - X^-$)
- Define $\nu(F) := \mathbb{E}[X \mathbb{1}_F]$, $F \in \mathcal{F}$.
- Then ν is a finite measure on $(\Omega, \mathcal{F})$

since $\mathbb{E}X < \infty$

$\forall$ disjoint $F_1, F_2, \ldots \in \mathcal{F}$: by monotone convergence thm

$$\nu\left(\bigcup_{i=1}^{\infty} F_i\right) = \mathbb{E}\left[\sum_{i=1}^{\infty} X \mathbb{1}_{F_i}\right] = \sum_{1}^{\infty} \mathbb{E}[X \mathbb{1}_{F_i}] = \sum_{1}^{\infty} \nu(F_i)$$

- $\nu \ll \mathbb{P}$ (if $\mathbb{P}(F)=0$ then $\mathbb{E}[X \mathbb{1}_F] = 0$)
- Apply RNT $\Rightarrow$ $\exists$ $f =: Y$ which is $\mathcal{F}$-integrable and

$$\mathbb{E}[X \mathbb{1}_F] = \int_F Y \, d\mathbb{P} = \mathbb{E}[Y \mathbb{1}_F] \quad \forall \; F \in \mathcal{F}. \qquad \square$$

Proof of Uniqueness Assume $\mathbb{E}[X \mathbb{1}_F] = \mathbb{E}[Y \mathbb{1}_F] = \mathbb{E}[Y' \mathbb{1}_F]$ $\forall F \in \mathcal{F}$

Subtract $\Rightarrow$ $\mathbb{E}[(Y-Y') \mathbb{1}_F] = 0$. Apply for $F := \{Y - Y' \geq 0\}$ $\Rightarrow$

$\mathbb{E}(Y-Y')^+ = 0$. Similarly, $\mathbb{E}(Y-Y')^- = 0$.

Add $\Rightarrow$ $\mathbb{E}|Y-Y'| = 0$ $\Rightarrow$ $Y = Y'$ a.s. $\square$