Probability: Definitions

Def: Measurable mapping

Let $\mathcal{A}, \mathcal{B}$ be σ-algebras in Λ, Ω. A mapping $X : \Lambda \to \Omega$ is called *measurable* if $X^{-1}(B) \in \mathcal{A}$ for all $B \in \mathcal{B}$.

Interpretation: "F measurable" means that expression "$\mathbb{P}(X^{-1}(B))$" makes sense.

Def: Random variables

A *random variable* X is a measurable mapping from an abstract probability space $(\Lambda, \mathcal{A}, \mathbb{P})$ into an observation space $(\Omega, \mathcal{B}(\Omega))$.

Image Measure

The measure $\mathbb{P}$ is not known explicitly. We work with the distribution μ_X of random variable X defined as the *image measure*:

$$\mu_X := X(\mathbb{P}) \qquad \text{i.e.} \qquad \mu_X(A) := \mathbb{P}(X^{-1}(A))$$

Conditioning

Note

Defining conditional measures requires some effort.

Direct approach

Conditional probability of $X(\omega) \in A$ given that $X(\omega) \in B$:

$$\mu(A|B) := \frac{\mu(A \cap B)}{\mu(B)}$$

$\to$ no use if $\mu(B) = 0$ (think of Bayesian model on $\mathbb{R}^d$)

For now:

- We will just write $\mu(X|Y)$ for the conditional probability of X given Y and forget about details.
- If X, Y have a joint density, $\mu(X|Y)$ has a conditional density $p(x|y)$.

Parametric Model

Parametric model

Let $X : (\Lambda, \mathcal{A}) \to (\Omega_x, \mathcal{B}_x)$ and $\Theta : (\Lambda, \mathcal{A}) \to (\Omega_\theta, \mathcal{B}_\theta)$ be two random variables, and $\mu_X = X(\mathbb{P})$. Then the conditional distribution $\mu_X(X|\Theta)$ is called a *parametric family* of models (parameterized by $\theta \in \Omega_\theta$).

Bayesian model

If X observed and Θ unobserved, we call:

- $\mu_\Theta := \Theta(\mathbb{P})$ the *prior measure*
- $\mu_\Theta(\Theta|X)$ the *posterior measure*
- The overall model is called a Bayesian model.

Note: Not defined by a Bayes equation!

Bayes' Theorem

Problem:

Given the prior and the data, how can we determine the posterior? (Without exhaustive knowledge of $\mathbb{P}$, $\mathcal{A}$ etc)

Bayes Theorem

If the sampling model $\mu_X(X|\Theta)$ has density $p_{X|\theta}$, then:

$$\frac{d\mu_{\Theta|X}}{d\mu_\Theta}(\theta|x) = \frac{p_{X|\theta}}{\int p_{X|\theta} d\mu_\theta(\theta)}$$

for all x with $\int p_{X|\theta} d\mu_\theta(\theta) \notin \{0, \infty\}$.