Probability Theory
19. Conditional independence

Peter Pfaffelhuber

June 25, 2024

Conditional probability and independence

- Lemma 11.12:

$$\mathcal{G}, \mathcal{H} \subseteq \mathcal{F} \text{ independent} \iff \mathrm{P}(G|\mathcal{H}) = \mathrm{P}(G), \quad G \in \mathcal{G}.$$

Proof: $\Rightarrow$: For $G \in \mathcal{G}, H \in \mathcal{H}$

$$\mathrm{E}[\mathrm{P}(G), H] = \mathrm{P}(G)\mathrm{P}(H) = \mathrm{P}(G \cap H) = \mathrm{E}[1_G, H] = \mathrm{E}[\mathrm{P}(G|\mathcal{H}), H].$$

'$\Leftarrow$': With $\mathrm{P}(G|\mathcal{H}) = \mathrm{P}(G)$, it follows for $H \in \mathcal{H}$

$$\mathrm{P}(G \cap H) = \mathrm{E}[1_G, H] = \mathrm{E}[\mathrm{P}(G|\mathcal{H}), H] = \mathrm{E}[\mathrm{P}(G), H] = \mathrm{P}(G) \cdot \mathrm{P}(H).$$

Example for conditional independence

- Let $(X_t)_{t=0,1,2,...}$ be a Markov chain. Simple Example:

  $Y_1, Y_2, \dots$ uiv with $P(Y_1 = 1) = 1 - P(Y_1 = -1) = p$ for a $p \in [0, 1]$. Further $X_t = Y_1 + \cdots + Y_t$. Then

$$P(X_{t+1} = k | X_0, \dots, X_t) = P(X_{t+1} = k | X_t) = \begin{cases} p, & k = X_t + 1, \\ q, & k = X_t - 1. \end{cases}$$

- For Markov chains:

  Given X_t, X_{t+1} is independent of $X_0, \dots, X_{t-1}$.

  Or in terms of σ-algebras:

  Given $\sigma(X_t)$, $\sigma(X_{t+1})$ is independent of $\sigma(X_0, \dots, X_{t-1})$.

Conditional independence

- Definition 11.14: Let $\mathcal{G} \subseteq \mathcal{F}$. A family $(\mathcal{C}_i)_{i \in I}$ is called *independently given* $\mathcal{G}$, if

$$\mathbb{E}\Big[1\Big(\bigcap_{j \in J} A_j\Big)\Big|\mathcal{G}\Big] = \mathrm{P}\Big(\bigcap_{j \in J} A_j|\mathcal{G}\Big) = \prod_{j \in J} \mathrm{P}(A_j|\mathcal{G}) = \prod_{j \in J} \mathrm{E}[1_{A_j}|\mathcal{G}]$$

  applies to all $J \subseteq_f I$ and $A_j \in \mathcal{C}_j, j \in J$.
- Examples:
  - If $\mathcal{G} = \mathcal{F}$, then $(\mathcal{C}_i)_{i \in I}$ is always independent given $\mathcal{G}$.
  - For $\mathcal{G} = \{\emptyset, \Omega\}$,

$$(\mathcal{C}_i)_{i \in I} \text{ independently given } \mathcal{G} \iff (\mathcal{C}_i)_{i \in I} \text{ independent.}$$

Example: Random probability of success

- $U \sim U([0,1])$; given U let $Y_1, ..., Y_n \sim B(n, U)$ be are independent and $X = Y_1 + \cdots + Y_n \sim B(n, U)$. Then for $I \subseteq [0,1]$ and $y_1 + \cdots + y_n = k$

$$\mathrm{E}[1_{Y_1=y_1,\ldots,Y_n=y_n}, U \in I] = \mathrm{P}(Y_1 = y_1, \ldots, Y_n = y_n, U \in I)$$
$$= \int_I u^k (1-u)^{n-k} du = \mathrm{E}\big[U^k (1-U)^{n-k}, U \in I\big],$$

so

$$\mathrm{P}(Y_1 = y_1, \ldots, Y_n = y_n | U) = U^k (1-U)^{n-k}$$

and

$$\mathrm{P}(Y_1 = y_1, \ldots, Y_n = y_n | U) = \prod_{i=1}^n \mathrm{P}(Y_i = y_i | U).$$

Conditional probability and conditional independence

- Proposition 11.17: $\mathcal{K}, \mathcal{G}, \mathcal{H} \subseteq \mathcal{F}$. Then,

$$\mathcal{G}, \mathcal{H} \text{ independently given } \mathcal{K} \iff \mathrm{P}(G|\sigma(\mathcal{H}, \mathcal{K})) = \mathrm{P}(G|\mathcal{K}), \ G \in \mathcal{G}.$$

Proof with $G \in \mathcal{G}, H \in \mathcal{H}, K \in \mathcal{K}$: $\Rightarrow$

$$\mathrm{E}[\mathrm{P}(G|\mathcal{K}), H \cap K] = \mathrm{E}[\mathrm{P}(G|\mathcal{K})\mathrm{P}(H|\mathcal{K}), K] = \mathrm{E}[\mathrm{P}(G \cap H|\mathcal{K}), K]$$

$$= \mathrm{P}(G \cap H \cap K) = \mathrm{E}[1_G, H \cap K] = \mathrm{E}[\mathrm{P}(G|\sigma(\mathcal{H}, \mathcal{K})), H \cap K].$$

The following is a $\cap$-stable Dynkin system:

$$\mathcal{D} := \{A \in \sigma(\mathcal{H}, \mathcal{K}) : \mathrm{E}[\mathrm{P}(G|\mathcal{K}), A] = \mathrm{P}(G \cap A)\}$$

$\Leftarrow$: $\quad \mathrm{P}(G \cap H|\mathcal{K}) = \mathrm{E}(1_G 1_H|\mathcal{K}) = \mathrm{E}(\mathrm{E}[1_G|\sigma(\mathcal{H}, \mathcal{K})]1_H|\mathcal{K})$

$$= \mathrm{E}(\mathrm{P}(G|\sigma(\mathcal{H}, \mathcal{K}))1_H|\mathcal{K}) = \mathrm{E}[\mathrm{P}(G|\mathcal{K}), H|\mathcal{K}] = \mathrm{P}(G|\mathcal{K}) \cdot \mathrm{P}(H|\mathcal{K})$$

Example: Markov chains

- Markov chain $(X_t)_{t=0,1,2,\ldots}$, i.e.

$$P(\underbrace{X_{t+1} \in A}_{\in \mathcal{G}} \,|\, \sigma(\underbrace{\sigma(X_t)}_{=:\mathcal{K}}, \underbrace{\sigma(X_1, \ldots, X_{t-1})}_{=:\mathcal{H}})) = P(\underbrace{X_{t+1} \in A}_{\in \mathcal{G}} \,|\, \underbrace{\sigma(X_t)}_{=:\mathcal{K}}).$$

Thus X_{t+1} is independent of $X_1, \ldots, X_{t-1}$ given X_t.