

The background of the slide is a solid blue color with a large, faint watermark of the University of Bonn seal. The seal features a central figure, likely a scholar or saint, seated and holding a book. Above the figure are three smaller figures in niches. The seal is surrounded by Latin text in a circular border.

# Probability Theory

## 19. Conditional independence

Peter Pfaffelhuber

June 25, 2024

# Conditional probability and independence

- ▶ Lemma 11.12:

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

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

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

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

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

## Example for conditional independence

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

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

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

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

$$\begin{aligned} \mathbb{E}[1_{Y_1=y_1, \dots, Y_n=y_n}, U \in I] &= \mathbb{P}(Y_1 = y_1, \dots, Y_n = y_n, U \in I) \\ &= \int_I u^k (1-u)^{n-k} du = \mathbb{E}[U^k (1-U)^{n-k}, U \in I], \end{aligned}$$

so

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

and

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

## Conditional probability and conditional independence

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

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

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

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

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

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

$$\begin{aligned} \Leftarrow: P(G \cap H|\mathcal{K}) &= E(1_G 1_H|\mathcal{K}) = E(E[1_G|\sigma(\mathcal{H}, \mathcal{K})]1_H|\mathcal{K}) \\ &= E(P(G|\sigma(\mathcal{H}, \mathcal{K}))1_H|\mathcal{K}) = E[P(G|\mathcal{K}), H|\mathcal{K}] = P(G|\mathcal{K}) \cdot P(H|\mathcal{K}) \end{aligned}$$

## Example: Markov chains

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

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

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