1.1. Measures

Let \((X, \mathcal{X})\) be a measurable space.

a) Give the definitions of the set \(\mathcal{M}(X)\) of \(\sigma\)-finite measures and the subset \(\mathcal{P}(X)\) of probability measures.

b) The spaces \(\mathcal{M}(X)\) and \(\mathcal{P}(X)\) are endowed with the \(\sigma\)-algebra generated by the mappings \(\pi_A : \mu \mapsto \mu(A), A \in \mathcal{X}\). Give the definition of this \(\sigma\)-algebra and show that \(\mathcal{P}(X)\) is a measurable subset of \(\mathcal{M}(X)\).

c) Show that the mappings \(\pi_f : \mu \mapsto \int f d\mu\), where \(f\) runs through the set of bounded measurable functions on \(X\), generate the same \(\sigma\)-algebra.

1.2. Kernels

Fix two measurable spaces \((X, \mathcal{X})\) and \((Y, \mathcal{Y})\).

a) Give the definitions of kernels and probability kernels from \(X\) to \(Y\).

b) Show that the mapping \(\delta : X \times \mathcal{Y} \to [0, 1], (x, A) \mapsto 1_A(f(x))\) is a probability kernel from \(X\) to \(Y\) iff \(f : X \to \mathcal{Y}\) is measurable.

c) Show that \(P : X \times \mathcal{Y} \to [0, 1]\) is a probability kernel from \(X\) to \(Y\) iff \(P : X \to \mathcal{P}(Y)\) is measurable.
1.3. Conditioning

Let $X$ and $Y$ be random variables on a probability space $(\Omega, \mathcal{F}, P)$ with values in measurable spaces $(X, \mathcal{X})$ and $(Y, \mathcal{Y})$. Moreover, let $\mathcal{G}$ be a $\sigma$-algebra contained in $\mathcal{F}$.

a) What are the definitions of the regular conditional probabilities $P_{X|\mathcal{G}}$ and $P_{X|Y}$?

b) Are $P[X \in \cdot | Y]$ and $P_{X|Y}(Y, \cdot)$ random probability measures, i.e., probability kernels on $(\Omega, \mathcal{F})$? Which properties need to be verified?

c) Show that $P_{X|\sigma(Y)}(\omega, A) = P_{X|Y}(Y(\omega), A)$ holds identically.

1.4. Bayes' formula

Let $X$ and $Y$ be random variables with values in measurable spaces $(X, \mathcal{X})$ and $(Y, \mathcal{Y})$, respectively. Assume that the law of $(X, Y)$ can be written in the form $\Lambda(x, y) \mu_X(dx) \mu_Y(dy)$ for some non-negative measurable function $\Lambda : X \times Y \to \mathbb{R}$ and some probability measures $\mu_X \in \mathcal{P}(X)$ and $\mu_Y \in \mathcal{P}(Y)$.

a) Show that the following is a regular conditional probability of $X$ given $Y$:

$$P_{X|Y}(y, A) = \frac{\int_A \Lambda(x, y) \mu_X(dx)}{\int \Lambda(x, y) \mu_X(dx)}$$

(In the expression above, the fraction is set to zero if the denominator vanishes.)

b) Use a) to calculate the regular conditional probability of $X$ given $Y$ if $(X, Y)$ is multivariate Gaussian.

Hint: the solution is $P_{X|Y}(y, dx) \sim \mathcal{N}(\hat{X}, \hat{\Sigma})$ with

$$\hat{X} = \mathbb{E}[X] + \text{Cov}(X, Y) \text{Var}(Y)^{-1}(y - \mathbb{E}[Y]),$$
$$\hat{\Sigma} = \text{Var}(X) - \text{Cov}(X, Y) \text{Var}(Y)^{-1} \text{Cov}(Y, X),$$

where $\text{Var}(Y)^{-1}$ is the generalized inverse of the (co-)variance matrix of $Y$. 
1.5. Hidden Markov Models

Let \((X, Y)\) be a Markov process. Show that the following statements are equivalent:

a) The transition kernel of \((X, Y)\) is of the form \(P(x, dx')K(x', dy')\).

b) \(X\) is Markov in the filtration generated by \((X, Y)\), and \(Y_{k+1}\) is independent of \((X_k, Y_k)\) conditional on \(X_{k+1}\), for each \(k \in \mathbb{N}\).

c) \(X_{k+1}\) is independent of \(Y_k\) conditional on \(X_k\), and \(Y_{k+1}\) is independent of \((X_k, Y_k)\) conditional on \(X_{k+1}\), for each \(k \in \mathbb{N}\).

1.6. Hidden Markov Models

a) Show that the process \((X, Y)\) defined by
\[
X_{k+1} = f(X_k, \alpha_k), \quad Y_{k+1} = g(X_{k+1}, \beta_k), \quad k = 0, 1, 2, \ldots
\]
is a Hidden Markov Model, where \(f : \mathbb{X} \times [0, 1] \to \mathbb{X}\) and \(g : \mathbb{X} \times [0, 1] \to \mathbb{Y}\) are measurable functions, and where \((\alpha_k)_{k \in \mathbb{N}}\) and \((\beta_k)_{k \in \mathbb{N}}\) are independent sequences of i.i.d. \([0, 1]\)-valued random variables.

b) Conversely, show that any Hidden Markov Model is of this form if \(X\) and \(Y\) are Borel spaces.

Hint. You may use [1, Theorem 6.10] and the discussions preceding this theorem.

References