PROBABILITY THEORY

MATH 154

Unit 7: Independence

7.1. $A, B \in \mathcal{A}$ in $(\Omega, \mathcal{A}, \mathrm{P})$ are called **independent** if $\mathrm{P}[A \cap B] = \mathrm{P}[A]\mathrm{P}[B]$. A set of events $\{A_i\}_{i \in I} \subset \mathcal{A}$ is called **independent**, if $\mathrm{P}[\bigcap_{j \in J} A_j] = \prod_{j \in J} \mathrm{P}[A_j]$ for any finite subset $J \subset I$,

7.2. Given a pairwise disjoint finite set $A_1, A_2, \dots, A_n \in \mathcal{A}$ satisfying $\bigcup_{j=1}^n A_h - \Omega$ is called a **finite partition** of Ω. Its elements are called **atoms**. It generates a σ-algebra $\mathcal{A} = \{A = \bigcup_{j \in J} A_j \ \}$, where J runs over all subsets of $\{1, \dots, n\}$. It has 2^n elements. Every finite sub-algebra of $\mathcal{A}$ is of this form: Every finite Boolean algebra has 2^n elements for some n.

7.3. Two π-systems $\mathcal{I}, \mathcal{J} \subset \mathcal{A}$ are called **independent**, if for all $A \in \mathcal{I}$ and $B \in \mathcal{J}$, $\mathrm{P}[A \cap B] = \mathrm{P}[A] \cdot \mathrm{P}[B]$. Similarly, two σ-algebras $\mathcal{A}_i \subset \mathcal{A}$ are called **independent** in $(\Omega, \mathcal{A}, \mathrm{P}0$, if any pair $A \in \mathcal{A}, B \in \mathcal{B}$ are independent. An arbitrary family $\prod_{j \in I} \mathcal{A}_j$ of σ-algebras in $\mathcal{A}$ are called independent, if any finite set $A_j \in \mathcal{A}_j$ of events are independent.

7.4. Examples:
1) On $(\Omega = \{1, 2, 3, 4\ \}, 2^\Omega, \mathrm{P}[A] = |A|/|\Omega|)$, the two σ-algebras $\mathcal{A} = \{\emptyset, \{1, 3\ \}, \{2, 4\ \}, \Omega\ \}$ and $\mathcal{B} = \{\emptyset, \{1, 2\ \}, \{3, 4\ \}, \Omega\ \}$ are independent.
2) For independent sets A, B in a probability space, the sub σ-algebras $\mathcal{A} = \{\emptyset, A, A^c, \Omega\}$ and $\mathcal{B} = \{\emptyset, B, B^c, \Omega\}$ are independent.
3) If $(\Omega_j, \mathcal{A}_j, \mathrm{P}[A_j])$ with $j \in I$ are probability spaces then each any subset of them is independent in the product probability space $(\Omega, \mathcal{A}, \mathrm{P})$ where Ω consists of sequences $\Omega = x = \{(x_j)_{j \in I}, x_j \in \Omega_j\}$ and where the σ algebra is generated by the π-system of sets $\prod_{j \in J} A_j$ in which only finitely many are not equal to Ω_j and where the measure P is extended from that π system via Carathéodory once $\mathrm{P}[\prod_{j \in J} A_j] = \prod_{j \in J} \mathrm{P}[A_j]$ is fixed on the π-system.

7.5. Given a probability space $(\Omega, \mathcal{A}, \mathrm{P})$. Let $\mathcal{G}, \mathcal{H}$ be two σ-sub-algebras of $\mathcal{A}$ and $\mathcal{I}$ and $\mathcal{J}$ be two π-systems satisfying $\sigma(\mathcal{I}) = \mathcal{G}$, $\sigma(\mathcal{J}) = \mathcal{H}$. Then $\mathcal{G}$ and $\mathcal{H}$ are independent if $\mathcal{I}$ and $\mathcal{J}$ are independent. Proof: (i) Fix $I \in \mathcal{I}$ and define on $(\Omega, \mathcal{H})$ the measures $\mu(H) = \mathrm{P}[I \cap H], \nu(H) = \mathrm{P}[I]\mathrm{P}[H]$ of total probability $\mathrm{P}[I]$. By independence of $\mathcal{I}$ and $\mathcal{J}$, they coincide on $\mathcal{I}$ and by the extension, they agree on $\mathcal{H}$ and we have $\mathrm{P}[I \cap H] = \mathrm{P}[I]\mathrm{P}[H]$ for all $I \in \mathcal{I}$ and $H \in \mathcal{H}$.
(ii) Define for fixed $H \in \mathcal{H}$ the measures $\mu(G) = \mathrm{P}[G \cap H]$ and $\nu(G) = \mathrm{P}[G]\mathrm{P}[H]$ of total probability $\mathrm{P}[H]$ on $(\Omega, \mathcal{G})$. They agree on $\mathcal{I}$ and so on $\mathcal{G}$ again by extension. We therefore have $\mathrm{P}[G \cap H] = \mathrm{P}[G]\mathrm{P}[H]$ for all $G \in \mathcal{G}$ and all $H \in \mathcal{H}$.

7.6. A random variable X **generates a σ subalgebra** $\sigma(X)$ of $\mathcal{A}$. It is defined as the smallest σ-algebra that contains all events $A = \{X \in [a,b)\}$ with $a < b$. Write $\sigma(X) = X^{-1}(\mathcal{B})$ because $\sigma(X) = \{X^{-1}(B) \mid B \in \mathcal{B}\ \}$, where $\mathcal{B}$ is the Borel σ algebra on $[0,1]$.

7.7. Examples:
1) A constant map $X(x) = c$ defines the **trivial algebra** $\mathcal{A} = \{\emptyset, \Omega\ \}$.
2) The projection map $X(x,y) = x$ from the square $(\Omega = [0,1] \times [0,1], \sigma(\mathcal{B} \times \mathcal{B}), \lambda \times \lambda)$ to the real line $\mathbb{R}$ has $\mathcal{B} = \{A \times [0,1]\ \}$ as a projection.
3) The map X from $\mathbb{Z}_6 = \{0,1,2,3,4,5\}$ to $\{0,1\} \subset \mathbb{R}$ defined by $X(x) = x \bmod 2$ has the value $X(x) = 0$ if x is even and $X(x) = 1$ if x is odd. The σ-algebra generated by X is $\mathcal{A} = \{\emptyset, \{1,3,5\}, \{0,2,4\}, \Omega\ \}$.

7.8. Two random variables X, Y are called **independent**, if they generate independent σ-algebras. They are independent if and only if $A = \{X \in (a,b]\}$ and $B = \{Y \in (c,d]\}$ are independent for all $a < b, c < d$. Independent random variables as two aspects of the laboratory Ω which do not influence each other. An $A = \{a < X(\omega) \leq b\ \}$ defined by X is independent of $B = \{c < Y(\omega) \leq d\ \}$ defined by B.

7.9. Examples:
1) Throwing a dice 3 times is modeled with a laboratory Ω has $6^3 = 216$ elements, where each experiment is a random vector $x = (x_1, x_2, x_3)$. Now, $X_j(x) = x_j \in \{1,2,3,4,5,6\}$ are independent random variables.
2) Given random variables Y_j on $(\Omega_j, \mathcal{A}_j, \mathrm{P}_j)$, w: $X_j(x) = f_j(x_j)$ on a product probability space $(\Omega = \prod_j \Omega_j, \mathcal{A} = \prod_j \mathcal{A}_j, \mathrm{P} = \prod_j \mathrm{P}_j)$ are independent.

7.10. If a σ-algebra $\mathcal{F} \subset \mathcal{A}$ is independent to itself, then $\mathrm{P}[A \cap A] = \mathrm{P}[A] = \mathrm{P}[A]^2$ so that for every $A \in \mathcal{F}$, $\mathrm{P}[A] \in \{0,1\}$. Such a σ-algebra is called P-**trivial**.
The trivial algebra $\mathcal{F} = \{\emptyset, \Omega\}$ is P-trivial in any probability space $(\Omega, \mathcal{A}, \mathrm{P}P)$. (See HW). Independence implies zero **covariance** $\mathrm{Cov}[X,Y] = \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y] = 0$ and zero correlation $\mathrm{Cor}[X,Y] = \mathrm{Cov}[X,Y]/(\sigma[X]\sigma[Y])$.

Theorem 1. *If X, Y are independent $\mathcal{L}^2$ random variables then* $\mathrm{E}[XY] = \mathrm{E}[X]\mathrm{E}[Y]$.

Proof. $\mathcal{L}^2$ assures that $E[XY] \leq \sqrt{E[X^2]|E[Y^2]}$ exists (Cauchy-Schwarz). By approximation it is enough to check for step functions $X = \sum_{i=1}^n a_i 1_{A_i}$, $Y = \sum_{j=1}^m b_j 1_{B_j}$ which generate finite independent σ algebras meaning every pair A_i, B_j is independent for i, j. Because $\mathrm{E}[X] = \sum_{i=1}^n a_i \mathrm{P}[A_i]$ and $\mathrm{E}[Y] = \sum_{j=1}^m b_j \mathrm{P}[B_j]$, we have $\mathrm{E}[XY] = \sum_{i=1}^n \sum_{j=1}^m a_i b_j \mathrm{P}[A_i \cap B_j] = \sum_{i=1}^n \sum_{j=1}^m a_i b_j \mathrm{P}[A_i]\mathrm{P}[B_j] = (\sum_{i=1}^n a_i \mathrm{P}[A_i])(\sum_{j=1}^m b_j \mathrm{P}[B_j])$. $= \mathrm{E}[X]\mathrm{E}[Y]$. ☐

7.11. The moment generating function $M_X(t) = \mathrm{E}[e^{Xt}]$ is defined if $X \in \mathcal{L}^\infty$ and can be expanded as $e^{Xt} = \sum_{k=0}^\infty X^k t^k / k!$. In the HW you show: If X, Y are $\mathcal{L}^\infty$ independent random variables, then X^n, Y^m are independent and e^{tX}, e^{tY} are independent.

Theorem 2. *X, Y are independent in $\mathcal{L}^\infty$ then $M_{X+Y}(t) = M_X(t)M_Y(t)$.*

Proof. $M_{X+Y}(t) = \mathrm{E}[e^{t(X+Y)}] = \mathrm{E}[e^{t(X+Y)}] = \mathrm{E}[e^{tX}e^{tY}] = \mathrm{E}[e^{tX}]\mathrm{E}[e^{tY}] = M_X(t)M_Y(t)$. ☐