

PROBABILITY THEORY

MATH 154

Unit 9: Tail algebras

9.1. Given a family $\{\mathcal{A}_i\}_{i \in I}$ of σ -subalgebras of $\mathcal{A}$. For any nonempty set $J \subset I$, let $\mathcal{A}_J := \bigvee_{j \in J} \mathcal{A}_j$ be the σ -algebra generated by $\bigcup_{j \in J} \mathcal{A}_j$. Define also $\mathcal{A}_\emptyset = \{\emptyset, \Omega\}$. The **tail σ -algebra** $\mathcal{T}$ of $\{\mathcal{A}_i\}_{i \in I}$ is defined as $\mathcal{T} = \bigcap_{J \subset I, J \text{ finite}} \mathcal{A}_{J^c}$, where $J^c = I \setminus J$. Recall that an algebra $\mathcal{A}$ is **trivial** if $P[A] = 0$ or $P[A] = 1$ for all $A \in \mathcal{A}$. As you have seen in a homework, there are σ algebras with infinitely many elements that are trivial. It does not have to be the smallest possible σ -algebra $\{\Omega, \emptyset\}$ which is trivial as a sub-algebra of any choice of probability space $(\Omega, \mathcal{A}, P)$.

Theorem 1 (Kolmogorov's 0 – 1 law). *If $\{\mathcal{A}_i\}_{i \in I}$ are independent σ -algebras, then the tail σ -algebra $\mathcal{T}$ is trivial.*

Proof. (i) Assume $F, G \subset I$ are disjoint sets. Then $\mathcal{A}_F$ and $\mathcal{A}_G$ are independent.

Proof. Define for $H \subset I$ the π -system $\mathcal{I}_H = \{A \in \mathcal{A} \mid A = \bigcap_{i \in K} A_i, K \subset_f H, A_i \in \mathcal{A}_i\}$. The π -systems $\mathcal{I}_F$ and $\mathcal{I}_G$ are independent and generate the σ -algebras $\mathcal{A}_F$ and $\mathcal{A}_G$.

(ii) Especially: $\mathcal{A}_J$ is independent of $\mathcal{A}_{J^c}$ for every $J \subset I$.

(iii) $\mathcal{T}$ is independent of $\mathcal{A}_I$.

Proof. $\mathcal{T} = \bigcap_{J \subset_f I} \mathcal{A}_{J^c}$ is independent of any $\mathcal{A}_K$ for $K \subset_f I$. It is therefore independent to the π -system $\mathcal{I}_I$ which generates $\mathcal{A}_I$.

(iv) $\mathcal{T}$ is a sub- σ -algebra of $\mathcal{A}_I$. Therefore $\mathcal{T}$ is independent of itself which implies that it is P -trivial. $\square$

9.2. Example: Let X_n be a sequence of independent random variables and let $A = \{\omega \in \Omega \mid \sum_{n=1}^{\infty} X_n \text{ converges}\}$. Then $P[A] = 0$ or $P[A] = 1$. Proof. Because $\sum_{n=1}^{\infty} X_n$ converges if and only if $Y_n = \sum_{k=n}^{\infty} X_k$ converges, we have $A \in \sigma(A_n, A_{n+1}, \dots)$. Therefore, A is in $\mathcal{T}$, the tail σ -algebra defined by the independent σ -algebras $\mathcal{A}_n = \sigma(X_n)$. If for example, if X_n takes values $\pm 1/n$, each with probability $1/2$, then $P[A] = 0$. If X_n takes values $\pm 1/n^2$ each with probability $1/2$, then $P[A] = 1$. The decision whether $P[A] = 0$ or $P[A] = 1$ is related to the convergence or divergence of a series. This will be discussed later again in the context of limit theorems.

9.3. Example: Let $\{A_n\}_{n \in \mathbb{N}}$ be a sequence of subsets of Ω . The set $A_\infty := \limsup_{n \rightarrow \infty} A_n = \bigcap_{m=1}^{\infty} \bigcup_{n \geq m} A_n$ consists of the set $\{\omega \in \Omega\}$ such that $\omega \in A_n$ for infinitely many $n \in \mathbb{N}$. The set A_∞ is contained in the tail σ -algebra of $\mathcal{A}_n = \{\emptyset, A_n, A_n^c, \Omega\}$. It follows from Kolmogorov's 0 – 1 law that $P[A_\infty] \in \{0, 1\}$ if $A_n \in \mathcal{A}$ and $\{A_n\}$ are P -independent.

Theorem 2 (Borel-Cantelli Lemma). *Take any sequence $A_n \in \mathcal{A}$.*

a) $\sum_{n \in \mathbb{N}} P[A_n] < \infty \Rightarrow P[A_\infty] = 0$ always holds.

b) $\sum_{n \in \mathbb{N}} P[A_n] = \infty \Rightarrow P[A_\infty] = 1$, if A_n are independent.

Proof. a) $P[A_\infty] = \lim_{n \rightarrow \infty} P[\bigcup_{k \geq n} A_k] \leq \lim_{n \rightarrow \infty} \sum_{k \geq n} P[A_k] = 0$.

b) For every integer $n \in \mathbb{N}$,

$$P\left[\bigcap_{k \geq n} A_k^c\right] = \prod_{k \geq n} P[A_k^c] = \prod_{k \geq n} (1 - P[A_k]) \leq \prod_{k \geq n} e^{-P[A_k]} = e^{-\sum_{k \geq n} P[A_k]}.$$

From

$$P[A_\infty^c] = P\left[\bigcup_{n \in \mathbb{N}} \bigcap_{k \geq n} A_k^c\right] \leq \sum_{n \in \mathbb{N}} P\left[\bigcap_{k \geq n} A_k^c\right] = 0$$

follows $P[A_\infty^c] = 0$. □

9.4. The following example illustrates that independence is necessary in the part b) of the Borel-Cantelli lemma: take the probability space $([0, 1], \mathcal{B}, P)$, where $P = \lambda$ is the Lebesgue measure on the Borel σ -algebra $\mathcal{B}$ of $[0, 1]$. For $A_n = [0, 1/n]$ we get $A_\infty = \emptyset$ and so $P[A_\infty] = 0$. But because $P[A_n] = 1/n$ we have $\sum_{n=1}^{\infty} P[A_n] = \sum_{n=1}^{\infty} \frac{1}{n} = \infty$ because the **harmonic series** $\sum_{n=1}^{\infty} 1/n$ diverges: $\sum_{n=1}^R \frac{1}{n} \geq \int_1^R \frac{1}{x} dx = \log(R)$.

9.5. Writing a novel amounts to enter a sequence of N symbols into a computer. For "Hamlet", Shakespeare had to enter $N = 180'000$ characters. Pop-culture ¹ imagines a monkey typing randomly for an indefinite time, producing a random text. Call A_n the event that Monkey types Hamlet on the interval $[nN, \dots, nN + N]$. These sets A_n are all independent and have probability 26^{-N} . Since $\sum_n A_n = \infty$ Borel-Cantelli assures that the even appears infinitely often Reality produces constraints like that monkeys like humans live less than $4 \cdot 10^9$ seconds but mathematicians do not care about such things. Their ideas live for ever!

9.6. A nice application of Borel-Cantelli are **percolation problems**. If we take an infinite connected network = graph and delete each bond=edge randomly with probability p , then there will be a threshold p_c such that for $p > p_c$ the network has no infinite cluster and for $p < p_c$ there is an infinite cluster. The event "there is an infinite cluster" is in the tail σ algebra of a set of σ algebras $\mathcal{A}_e = \{0, 1, N_e, N_e^c\}$, where N_e are the set of networks for which bond e is active and N_e^c the set of networks for which edge e is broken. The index set $J = E$ enumerates all the edges of the network and the tail σ -algebra $\mathcal{T}$ consists of all events that do not change if a finite part of the network is altered.

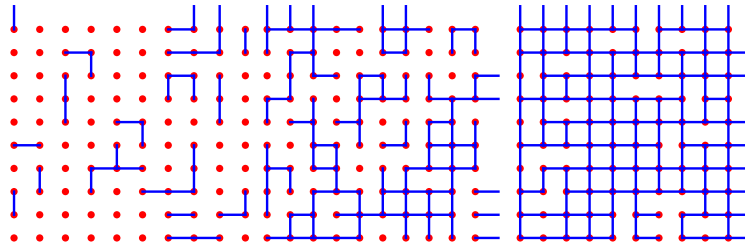


FIGURE 1. The bond percolation threshold in dimension 2 is known to be $p = 1/2$. We see random lattice networks with $p = 0.2, p = 0.5, p = 0.8$.

¹First appearing in Feller's book from 1950