

# Section 1: Introduction to probability

ARE 210

August 29, 2017

- Introduction (10 min)
- A few tips (10 min)
- Practice questions (30 min)

The section notes are available on the section Github at [github.com/johnloeser/are210](https://github.com/johnloeser/are210) in the “section1” folder.

## 1 Definitions

- A **probability space** is a triple  $(\Omega, \mathbf{F}, P)$ 
  - The **sample space**  $\Omega$  is a set
    - \*  $A \subseteq \Omega$  is an **event**
  - $\mathbf{F} \subseteq 2^\Omega$  is a  **$\sigma$ -algebra**
    - \* A set  $\mathbf{F}$  is a  **$\sigma$ -algebra** if
      1.  $\emptyset \in \mathbf{F}$
      2.  $A \in \mathbf{F} \Rightarrow A^c \in \mathbf{F}$
      3.  $A_1, A_2, \dots \in \mathbf{F} \Rightarrow \cup_{i=1}^\infty A_i \in \mathbf{F}$
  - The **probability measure**  $P : \mathbf{F} \rightarrow [0, 1]$  such that
    1.  $P(\Omega) = 1$
    2.  $A_1, A_2, \dots$  pairwise disjoint  $\Rightarrow P(\cup_{i=1}^\infty A_i) = \sum_{i=1}^\infty P(A_i)$
- A **measurable space** is a 2-tuple  $(\Omega, \mathbf{F})$ , where  $\Omega$  is a set and  $\mathbf{F}$  is a  $\sigma$ -algebra over  $\Omega$ 
  - Common measurable spaces are  $(\Omega, 2^\Omega)$  and  $(\mathbf{R}^k, \mathbf{B}^k)$  (where  $\mathbf{B}$  is the **Borel  $\sigma$ -algebra**, the smallest  $\sigma$ -algebra containing all open sets in  $\mathbf{R}$ )
- A **partition** of  $\Omega$  is a set of disjoint sets whose union is  $\Omega$
- The probability of  $A$  conditional on  $B$  is  $P(A|B) = \frac{P(A \cap B)}{P(B)}$
- **Bayes rule** is  $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$
- $\{A_i\}_{i=1}^N$  are **mutually independent** if  $\forall I \subseteq \{1, \dots, N\}, P(\cap_{i \in I} A_i) = \prod_{i \in I} P(A_i)$
- A **random variable**  $X : (\Omega, \mathbf{F}) \rightarrow (E, \mathbf{E})$ , where  $X$  is a measurable function, and  $(\Omega, \mathbf{F})$  and  $(E, \mathbf{E})$  are measurable spaces
  - Typically,  $(E, \mathbf{E}) = (\mathbf{R}^k, \mathbf{B}^k)$

## 2 Some useful tips for the homework

- Set logic
  - $a \in A \cup B \Leftrightarrow a \in A \vee a \in B$
  - $a \in A \cap B \Leftrightarrow a \in A \wedge a \in B$
  - $a \in A^c \Leftrightarrow a \notin A$
  - $A = B \Leftrightarrow (a \in A \Rightarrow a \in B) \wedge (a \in B \Rightarrow a \in A)$
  - $A \subseteq \Omega \Leftrightarrow A \in 2^\Omega$
  - $a \in X^{-1}(A) \Leftrightarrow X(a) \in A$
- Let  $C$  be a set of sets, and  $\sigma(C)$  be the smallest  $\sigma$ -algebra over  $\Omega$  containing  $C$ . It's useful to conceptualize  $\sigma(C)$  as countable unions of the smallest partition of  $\Omega$  which can be used to construct any element of  $C$  by countable union.

Also worth noting - Lecture 1 covered sufficient material to answer questions 1 through 5, while Lecture 2 covered sufficient material to answer questions 6 through 8. If any questions seem difficult to you, I recommend coming to office hours to discuss why it seemed tough. If the questions seemed easy, it might be useful to wait to complete the problem set after we've covered all the relevant material as a refresher.

## 3 Some example questions

1) **Conditional independence:** Define  $A_1$  and  $A_2$  to be independent conditional on the event  $B$  occurring if  $P(A_1 \cap A_2|B) = P(A_1|B)P(A_2|B)$ . Does independence conditional on each event in a partition of the sample space imply independence? If so, why? If not, construct a counterexample.

2) **Practice constructing counterexamples:** Recall our formula for Bayesian updating,  $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$ . Suppose we observe the events  $B_1, B_2, \dots, B_N$ . Consider calculating  $P(A|\cap_{i=1}^N B_i)$  via **batch Bayesian updating** and **sequential Bayesian updating**. Define the batch Bayesian updating approach by

$$P(A|\cap_{i=1}^N B_i) = \frac{P(\cap_{i=1}^N B_i|A)P(A)}{P(\cap_{i=1}^N B_i)}$$

Define the sequential Bayesian updating approach recursively by

$$P^{SEQ}(A|B_1) = \frac{P(B_1|A)P(A)}{P(B_1)}$$

$$P^{SEQ}(A|\cap_{i=1}^K B_i) = \frac{P(B_K|A)P^{SEQ}(A|\cap_{i=1}^{K-1} B_i)}{P(B_K)}$$

a) Propose two sufficient conditions for the two approaches to yield the same answer. Which approach do you think is “correct”? Hint:  $\frac{P(B_K|A)P^{SEQ}(A|\cap_{i=1}^{K-1} B_i)}{P(B_K)} = \left(\frac{P(B_K|A)}{P(B_K)}\right) P^{SEQ}(A|\cap_{i=1}^{K-1} B_i)$ .

b) How can you correct the sequential (batch) approach so it returns the same answer as the batch (sequential) approach?

c) For  $N = 2$ , construct an example where the two approaches yield different answers, but the second condition is not violated.

d) For  $N = 2$ , construct an example where the two approaches yield different answers, but the first condition is not violated.