

M378K Introduction to Mathematical Statistics

Problem Set #1

Probability spaces. Random variables.

1.1. Probability distributions. Consider an **outcome space** (also known as a **sample space**) Ω . In statistics, it's convenient to imagine it as containing your entire target population. The individual elements $\omega \in \Omega$ are known in probability as **elementary outcomes**; in statistics, they can frequently (but not always!) be understood as individuals in your target population.

We are usually not interested that much in individual ω , but want to consider **events** E in Ω . In full mathematical generality, the set Ω can be complicated so that the family of all of its subsets is also a complicated object. Due to mathematical difficulties, it is sometimes impossible to measure in any meaningful way absolutely all subsets of Ω ¹. For the purposes of this course, we will not have to dwell on these matters. So, we will refer to any (nice) subset of Ω as an **event**.

¹See https://en.wikipedia.org/wiki/Banach-Tarski_paradox

We will inherit all the notation and rules from set theory while adding some more vocabulary and context. Most notably, we will consider *intersections*, *unions*, and *complements* of events. These are best understood via Venn diagrams.

Moreover, in a probabilistic setting, we have the following definition:

Definition 1.1. Let E and F be two events on the same Ω such that

$$E \cap F = \emptyset.$$

Then, we say that E and F are mutually exclusive (or disjoint).

Now, we are ready for the following (crucial!) definition:

Definition 1.2. Consider a mapping \mathbb{P} from the set of all events on Ω to \mathbb{R} . We say that \mathbb{P} is a probability (distribution, measure) if it satisfies the following three conditions:

- $\mathbb{P}[E] \geq 0$ for all events E ;
- $\mathbb{P}[\Omega] = 1$;
- for all **pairwise disjoint** sequences of events $\{E_j : j = 1, 2, \dots\}$, we have that

$$\mathbb{P}\left[\bigcup_{j=1}^{\infty} E_j\right] = \sum_{j=1}^{\infty} \mathbb{P}[E_j].$$

One immediately useful consequence is the **inclusion-exclusion formula**. This is its simplest version.

Proposition 1.3. *Let E and F be two events on Ω . Then,*

$$\mathbb{P}[E \cup F] = \mathbb{P}[E] + \mathbb{P}[F] - \mathbb{P}[E \cap F].$$

Of course, the above formula can be generalized to arbitrary unions of finitely many events. *Try to figure it out!*

Problem 1.1. Source: An old P exam problem.
For two events A and B , you are given that

$$\mathbb{P}[A \cup B] = 0.7 \quad \text{and} \quad \mathbb{P}[A \cup B^c] = 0.9.$$

Calculate $\mathbb{P}[A]$.

1.2. Random variables. Informally speaking, any "nice" mapping/function from Ω to a target set \mathcal{S} is a *random element*². When \mathcal{S} is \mathbb{R} , we like to use the term *random variable*. When \mathcal{S} is \mathbb{R}^n for some n , we like to use the term *random vector*.

Let's consider a classroom of students as our Ω and give examples of a

- random element
- random variable
- random vector

²In practice, people like to use the term *random variable* even in more general context when there is no source of confusion. We will habitually do this.

To keep track of what values a random variable is "allowed" to take, we use the following terminology³:

Definition 1.4. *Given a set B , we say that a random variable Y is B -valued if*

$$\mathbb{P}[Y \in B] = 1.$$

³Read your lecture notes: https://web.ma.utexas.edu/users/gordanz/notes/discrete_probability_color.pdf