Stat 2911 Lecture Notes

Class 1, 2017

Uri Keich

Topics

STAT 2911 (2017)

Probability and Statistical Models (Advanced)

Uri Keich, Carslaw 821

Office hour: Monday 5-6 pm

If you want to talk to me please try to come to the office hour. If you can't then email me so we can set up a time.

LMS site: course material, turnitin assignments

Consider STAT 2011 (mainstream)

Labs and tutorials start this week

Attempt the tutorials at home: it's critical for your understanding of the material.

Lab reports are due by 5pm on Friday.

Much more demanding than MATH 1905

Probability in general, and this course in particular, has two components:

- The mathematical theory of probability:
  definitions, theorems, <u>proofs</u>

Abstraction of experiments whose outcome is random.

- Modeling: argued rather than proved
  Applications can be confusing as well as ill-defined!
  Useful for describing/summarizing the data, understand features of the data, make accurate predictions.

Terminology. The set of all possible outcomes is called the sample space (denoted usually by Ω).
A point $\omega \in \Omega$ is called a sample point.

Examples 1) Model a roll of a die

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

2) Model a coin flip

$$\Omega = \{H, T\}$$

Should we model the possibility that the coin lands on its side?

More terms:

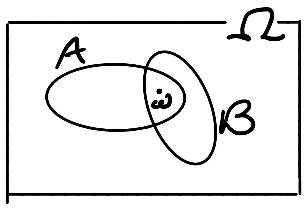
Events are subsets of Ω to which we can assign a probability.

We say an event A occurred if the outcome, or sample point $\omega \in \Omega$ satisfies $\omega \in A$.

Since events are generally not disjoint (zero intersection) more than one event can occur.

Example using a Venn diagram:

If the outcome is ω (as pictured) then both events A and B occurred.



Helps visualize relations between sets.

We denote the probability of an event A by $P(A)$.
Intuitively, $P(A)$ is the rate at which A occurs if we repeat the experiment many times.

Def: A function P defined on the set of all events is a probability measure / function if:

(i) $P(\Omega)=1$

(ii) $P(A) \geqslant 0$ for every event A

(iii) If $A_1, A_2, \ldots$ are mutually disjoint events (for $i \neq j$ $A_i \cap A_j = \emptyset$), then $P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$.

Do you spot something odd about (i) and (iii)?
How do we know Ω and $\bigcup_n A_n$ are events?

Why bother defining events?

Why not define any subset of Ω as an event?!

That often works but sometimes it can fail spectacularly.

Imagine choosing a point at random in a large cube $C \subset \mathbb{R}^3$, where the prob. of the point lying in any subset $A \subset C$ is proportional to its volume $|A|$.

Clearly, if $A, B \subset C$ are related through a rigid motion (translation + rotation) then $|A| = |B|$ so $P(A) = P(B)$.

Similarly if $A = A_1 \cup A_2$ with $A_i \subset C$ disjoint ($A_1 \cap A_2 = \emptyset$) then $P(A) = P(A_1) + P(A_2)$.

Banach-Tarski Paradox

The unit ball $B \subset \mathbb{R}^3$ can be decomposed into 5 (disjoint) pieces $B = B_1 \cup \ldots \cup B_5$, which can be reassembled, using only rigid motions, into two balls of the same size as the original!

$$P(B) = P(B_1) + \ldots + P(B_5) = P(B_1') + \ldots + P(B_5') = 2P(B)$$

Cor. Volume / Prob. cannot be assigned to arbitrary sets, rather to sets that are "measurable".

The collection of measurable sets is captured by the notion of a σ-algebra (σ-field).

<u>Def</u>. A collection of subsets of Ω, $\mathcal{F}$, is a <u>σ-algebra</u> if:

(i) $\Omega \in \mathcal{F}$

(ii) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$ ($\mathcal{F}$ is closed wrt complementing)

(iii) $A_1, A_2, \dots \in \mathcal{F} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$ ($\mathcal{F}$ is closed wrt a <u>countable</u> union)

Countable means you can count it.

$A = \{1, 3, 5\}$ is a finite countable set

$\mathbb{N}_1 = \{1, 2, 3, \dots\}$ is an infinite " "

$\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$ is also countable, moreover although $\mathbb{Z} \supset \mathbb{N}_1$ it has the same cardinality (size) as $\mathbb{N}_1$.

The same goes for $\mathbb{Q} = \{\frac{m}{n} : m \in \mathbb{Z}, n \in \mathbb{N}_1\}$

$\mathbb{R}$ is <u>not</u> countable (Cantor).

A countable infinity is the smallest infinity.

<u>Example</u> (of a σ-algebra)

$\Omega = \{1, 2, \dots, 6\}$

$\mathcal{F}$ = power set of Ω = set of all subsets of Ω

$= \{ \emptyset,$ (1)

$\{1\}, \{2\}, \dots, \{6\},$ (6)

$\{1,2\}, \{1,3\}, \dots, \{5,6\},$ (15)

$\vdots$

$\Omega \}$ (1)

$= 2^{\Omega}$ (notation)

Questions:

- What is the cardinality (size) of $\mathcal{F}$, or how many subsets of Ω does it contain?
- Is 2^{Ω} always a σ-algebra?
- For any sample space $\Omega \neq \emptyset$, what is the smallest σ-algebra you can think of?

Def. A triplet $(\Omega, \mathcal{F}, P)$ is a <u>probability space</u> if

(i) Ω is a sample space

(ii) $\mathcal{F}$ is a σ-algebra of events in Ω

(iii) $P: \mathcal{F} \to \mathbb{R}$ is a prob. measure.

<u>Example</u>. Model a roll of a fair die

$$\Omega = \{1, 2, \dots, 6\}$$

$$\mathcal{F} = 2^{\Omega}$$

$$P(\{i\}) = 1/6 \qquad i = 1, 2, \dots, 6$$

$$P(A) = |A|/6 \qquad \text{for } A \in \mathcal{F}.$$

By varying $P(\{i\}) = p_i \in (0,1)$ we can model a biased die as well.

What can we say about P?

<u>Claim</u>. If A and B are disjoint events then

$$P(A \cup B) =$$

How do you know $A \cup B \in \mathcal{F}$?!

<u>Ex</u>. (i) $\emptyset \in \mathcal{F}$

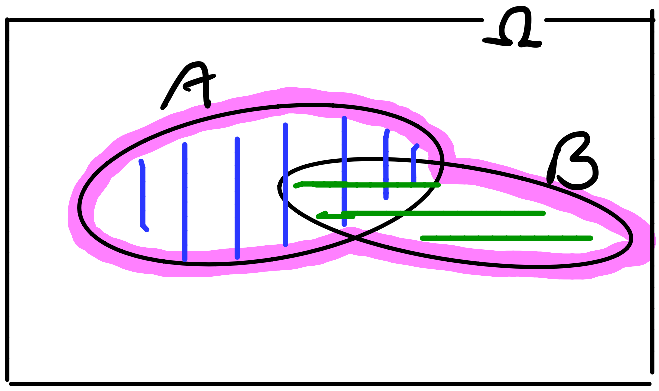
(ii) If $A_1, \ldots, A_n \in \mathcal{F}$ then $\bigcup_{i=1}^{n} A_i \in \mathcal{F}$ (hint: use (i))

(iii) $P(\emptyset) = 0$

Using the exercise prove that $P(A \cup B) = P(A) + P(B)$ above.

What if A and B are not disjoint?

You can use a Venn diagram to get an idea:



prob. ~ area

$$P(A\cup B) = P(A) + P(B) - P(A\cap B)$$

How do we know that $A\cap B \in \mathcal{F}$?

Ex. Using de Morgan laws: $(A\cup B)^c = A^c\cap B^c$

$(A\cap B)^c = A^c\cup B^c$

show that if $A,B\in\mathcal{F}$ then $A\cap B\in\mathcal{F}$.