STAT 345/445 Lecture 2

Section 1.2: Basics of Probability Theory

Subsections 1.2.1 and 1.2.2

... and some considerations about proofs

What probability is used for

- Experiments have uncertain (unpredictable) outcomes
- But, for *repeated* experiments we may expect a pattern

Experiment	**Sample Space**	
Soccer game	$S = \{\text{win}, \text{loose}, \text{draw}\}$	
Roll a die	$S = \{1, 2, 3, 4, 5, 6\}$	← Equally likely
Sum of two dice	$S = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$	← not equally likely

We use probability distributions to describe the long term behavior of repeated experiments

Probability – mathematically

- Want a rule that can assign a number between 0 and 1 to any event E in a sample space S.
- Easy for finite and countable sample spaces - can enumerate all possible subsets of S.
- Technical problem: If S is an uncountable sample space, “every subspace” is too many to handle.
- Need the concept of σ-algebras.

Sigma-algebra

Definition: σ-algebra

A collection of subsets of S is called a **σ-algebra on S** (or **Borel field**) $\mathcal{B}$ if the following holds:

(i) $\emptyset \in \mathcal{B}$

(ii) If $A \in \mathcal{B}$ then $A^c \in \mathcal{B}$ Closed under complement

(iii) If $A_1, A_2, A_3, \ldots \in \underline{\mathcal{B}}$ then $\displaystyle\bigcup_{i=1}^{\infty} A_i \in \mathcal{B}$ Closed under unions

Mathematicians and Statisticians tend to be minimalist in definitions...

σ-algebra - theorem

Theorem

If $\mathcal{B}$ is a σ-algebra then

(a) $S \in \mathcal{B}$

(b) If $A_1, A_2, A_3, \ldots \in \mathcal{B}$ then $\bigcap_{i=1}^{\infty} A_i \in \mathcal{B}$ Closed also under intersections

Proof:

(a) $\emptyset \in \mathcal{B}$ (i) $\Rightarrow \emptyset^c \in \mathcal{B}$ (ii).

by cond. (i)   by cond. (ii)

σ-algebra - theorem

Theorem

If $\mathcal{B}$ is a σ-algebra then

(a) $S \in \mathcal{B}$

(b) If $A_1, A_2, A_3, \ldots \in \mathcal{B}$ then $\displaystyle\bigcap_{i=1}^{\infty} A_i \in \mathcal{B}$

Proof:

(a) $\emptyset \in \mathcal{B}$ (i) $\Rightarrow \emptyset^c \in \mathcal{B}$ (ii). Since $\emptyset^c = S$ we have that $S \in \mathcal{B}$

σ-algebra - theorem

Theorem

If $\mathcal{B}$ is a σ-algebra then

(a) $S \in \mathcal{B}$

(b) If $A_1, A_2, A_3, \ldots \in \mathcal{B}$ then $\displaystyle\bigcap_{i=1}^{\infty} A_i \in \mathcal{B}$

Proof:

(a) $\emptyset \in \mathcal{B}$ (i) $\Rightarrow \emptyset^c \in \mathcal{B}$ (ii). Since $\emptyset^c = S$ we have that $S \in \mathcal{B}$

(b) Let $A_1, A_2, A_3, \ldots \in \mathcal{B}$. Then $A_1^c, A_2^c, A_3^c, \ldots \in \mathcal{B}$ by (ii).

σ-algebra - theorem

Theorem

If $\mathcal{B}$ is a σ-algebra then

(a) $S \in \mathcal{B}$

(b) If $A_1, A_2, A_3, \ldots \in \mathcal{B}$ then $\displaystyle\bigcap_{i=1}^{\infty} A_i \in \mathcal{B}$

Proof:

(a) $\emptyset \in \mathcal{B}$ (i) $\Rightarrow \emptyset^c \in \mathcal{B}$ (ii). Since $\emptyset^c = S$ we have that $S \in \mathcal{B}$

(b) Let $A_1, A_2, A_3, \ldots \in \mathcal{B}$. Then $A_1^c, A_2^c, A_3^c, \ldots \in \mathcal{B}$ by (ii). Therefore

$$\bigcup_{i=1}^{\infty} A_i^c \in \mathcal{B} \Rightarrow \left(\bigcup_{i=1}^{\infty} A_i^c\right)^c \in \mathcal{B} \qquad \text{by (iii) and (ii)}$$

σ-algebra - theorem

Theorem

If $\mathcal{B}$ is a σ-algebra then

(a) $S \in \mathcal{B}$

(b) If $A_1, A_2, A_3, \ldots \in \mathcal{B}$ then $\displaystyle\bigcap_{i=1}^{\infty} A_i \in \mathcal{B}$

Proof:

(a) $\emptyset \in \mathcal{B}$ (i) $\Rightarrow \emptyset^c \in \mathcal{B}$ (ii). Since $\emptyset^c = S$ we have that $S \in \mathcal{B}$

(b) Let $A_1, A_2, A_3, \ldots \in \mathcal{B}$. Then $A_1^c, A_2^c, A_3^c, \ldots \in \mathcal{B}$ by (ii). Therefore

$$\bigcup_{i=1}^{\infty} A_i^c \in \mathcal{B} \Rightarrow \left(\bigcup_{i=1}^{\infty} A_i^c\right)^c \in \mathcal{B} \qquad \text{by (iii) and (ii)}$$

$$\Rightarrow \bigcap_{i=1}^{\infty} (A_i^c)^c = \bigcap_{i=1}^{\infty} A_i \in \mathcal{B} \qquad \text{by DeMorgan and compl.}$$

Exmples of σ-algebras

- The trivial σ-algebra: $\mathcal{B} = \{\emptyset, S\}$
- If S is finite:

$$\mathcal{B} = \{\text{all subsets of } S\text{, including } S \text{ and } \emptyset\ \}$$

  - If $S = \{1, 2, 3, \ldots, n\}$, how many subsets are there? 2^n

- If $S = \mathbb{R}$ the most common σ-algebra is the collection of all open and closed intervals, and their unions

= all sensible subsets

$$\mathcal{B} = \text{all sets that can be written as a union of intervals}$$

$$[a, b], [a, b), (a, b], \text{or } (a, b) \text{ where } a, b \in \mathbb{R}, a < b$$

e.g. $\{a\} = (a-1, a] \cup [a, a+1)$

Probability function

$P : \mathcal{B} \longmapsto \mathbb{R}$ (actually $\longmapsto [0,1]$)

Definition: Probability fuction

Let S be a sample space and let $\mathcal{B}$ be an associated σ-algebra.
A **probability function** is a function $P(\cdot)$ with domain $\mathcal{B}$ that satisfies the **Kolmogorov axioms**:

prob. is not negative

(i) $P(A) \geq 0 \quad \forall A \in \mathcal{B}$

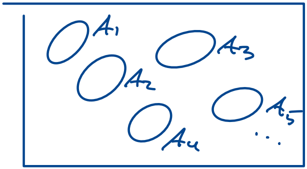
"Prob. 1 that something happens"

(ii) $P(S) = 1$

(iii) If $A_1, A_2, A_3, \ldots \in \mathcal{B}$ are mutually exlusive then $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$

All probability theory is based on these Kolmogorov axioms

Kolm. Ax. (iii):



mutually exclusive

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

Probability model

- The triple $(S, \mathcal{B}, P)$ is called a **probability model**

Creating a probability model

- Say $S = \{s_1, s_2, \ldots, s_n\}$
- Let $\mathcal{B}$ be the collection of every subset of S
- Probability function?

Let $p_1, p_2, \ldots, p_n \in [0,1]$ where $\sum_{i=1}^{n} p_i = 1$ (**)

Define: For any $A \in \mathcal{B}$ set

$$P(A) = \sum_{\{i : s_i \in A\}} p_i$$

Is P a prob. func.?
$\rightarrow$ Proof on whiteboard

e.g if $A = \{s_2, s_5, s_{10}\}$: $\sum_{i \in \{2,5,10\}} p_i = p_2 + p_5 + p_{10}$

Calculus of Probabilities

In the following: Let $(S, \mathcal{B}, P)$ be a probability model

Theorem 1

Let $A \in \mathcal{B}$. Then

(a) $P(A^c) = 1 - P(A)$

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

(c) $P(A) \leq 1$

Proof:

(a) A and A^c are disjoint. $\Rightarrow$ by axiom (iii): $P(A \cup A^c) = P(A) + P(A^c)$.

Calculus of Probabilities

In the following: Let $(S, \mathcal{B}, P)$ be a probability model

Theorem 1

Let $A \in \mathcal{B}$. Then

(a) $P(A^c) = 1 - P(A)$

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

(c) $P(A) \leq 1$

Proof:

(a) A and A^c are disjoint. $\Rightarrow$ by axiom (iii): $P(A \cup A^c) = P(A) + P(A^c)$.
$A \cup A^c = S$ so by axiom (ii): $P(A \cup A^c) = 1$.

Calculus of Probabilities

In the following: Let $(S, \mathcal{B}, P)$ be a probability model

Theorem 1

Let $A \in \mathcal{B}$. Then

(a) $P(A^c) = 1 - P(A)$

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

(c) $P(A) \leq 1$

Proof:

(a) A and A^c are disjoint. $\Rightarrow$ by axiom (iii): $P(A \cup A^c) = P(A) + P(A^c)$.
$A \cup A^c = S$ so by axiom (ii): $P(A \cup A^c) = 1$. $\Rightarrow P(A^c) = 1 - P(A)$

(b) $P(\emptyset) = P(S^c) = 1 - P(S)$ by part (a).

Calculus of Probabilities

In the following: Let $(S, \mathcal{B}, P)$ be a probability model

Theorem 1

Let $A \in \mathcal{B}$. Then

(a) $P(A^c) = 1 - P(A)$

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

(c) $P(A) \leq 1$

Proof:

(a) A and A^c are disjoint. $\Rightarrow$ by axiom (iii): $P(A \cup A^c) = P(A) + P(A^c)$. $A \cup A^c = S$ so by axiom (ii): $P(A \cup A^c) = 1$. $\Rightarrow P(A^c) = 1 - P(A)$

(b) $P(\emptyset) = P(S^c) = 1 - P(S)$ by part (a). $\Rightarrow P(\emptyset) = 1 - 1$ by axiom (ii)

(c) $1 = P(S) = P(A \cup A^c) = P(A) + P(A^c)$ by axioms (ii) and (iii). Since $P(A^c) \geq 0$ (axiom (i)) we have $1 \geq P(A)$.

Theorem 2

Let $A, B \in \mathcal{B}$. Then

(a) $P(A) = P(A \cap B) + P(A \cap B^c)$

(b) $P(A \setminus B) = P(A) - P(A \cap B)$

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

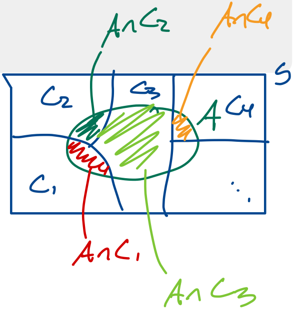
(d) If $A \subseteq B$ then $P(A) \leq P(B)$

Proof: Homework 1

Theorem 3: Law of total probability

Let $A \in \mathcal{B}$ and let $C_1, C_2, \ldots \in \mathcal{B}$ be a partition of S. Then

$$P(A) = \sum_{i=1}^{\infty} P(A \cap C_i)$$



proof...

$$P(A) = P(A \cap S) = P\left(A \cap \left(\bigcup_{i=1}^{\infty} C_i\right)\right)$$

$$= P\left(\bigcup_{i=1}^{\infty} A \cap C_i\right)$$

$$= \sum_{i=1}^{\infty} P(A \cap C_i) \quad \text{since}$$

$A \cap C_1, A \cap C_2, \ldots$ are disjoint:

$\forall i, j$ with $i \neq j$

$(A \cap C_i) \cap (A \cap C_j) = A \cap C_i \cap C_j = \emptyset$

Distributive law:

$A \cap (C_1 \cup C_2 \cup \cdots)$

$= (A \cap C_1) \cup (A \cap C_2) \cup \cdots$

Theorem 4: Boole's Inequality

Let $A_1, A_2, A_3, \ldots \in \mathcal{B}$. Then

(a) $P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} P(A_i)$, for $n = 1, 2, 3, \ldots$

(b) $P\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} P(A_i)$

upper bound for $\cup$ of sets

Proof:

(a) ~~1.~~ Follows from (b) but we can also use induction ...

(b) ~~2.~~ STAT 445: Read proof in book

proof on whiteboard

Proof by induction

- Want to show that some property $Q(n)$ holds for all $n = 1, 2, 3, \ldots$
  - Checking $Q(n)$ for all n will literally take forever!

Strategy:

1. Prove that $Q(1)$ holds
   - Sometimes it is easier to also prove $Q(2)$
2. For an arbitrary integer k *assume* that $Q(k)$ holds. Then prove that $Q(k+1)$ holds.

Theorem 5: Bonferroni Inequality

Let $A_1, A_2, A_3, \dots A_n \in \mathcal{B}$. Then

(a) $P\left(\bigcap_{i=1}^{n} A_i\right) \geq \sum_{i=1}^{n} P(A_i) - (n-1)$

(b) $P\left(\bigcap_{i=1}^{n} A_i\right) \geq 1 - \sum_{i=1}^{n} P(A_i^c)$

lower bound for $\cap$ of sets

Proof: Going from Bool to Bonferroni i.e. from $\cup$ to $\cap$ => Think DeMorgan

proof on whiteboard

Use of Bonferroni

- Want $P(A \cap B)$.
  - Know that $P(A) = p_1$ and $P(B) = p_2$

  A lower bound on $P(A \cap B)$:

$$P(A \cap B) \geq 1 - ((1 - p_1) + (1 - p_2))$$

- Example: Confidence intervals - familywize confidence level
  - Say we have 95% confidence intervals based on the following probability statements

$$P\left(\overline{X}_1 - z_{\alpha/2}\sigma_1 \leq \mu_1 \leq \overline{X}_1 + z_{\alpha/2}\sigma_1\right) = 1 - \alpha$$
$$P\left(\overline{X}_2 - z_{\alpha/2}\sigma_2 \leq \mu_2 \leq \overline{X}_2 + z_{\alpha/2}\sigma_2\right) = 1 - \alpha$$

    Then both (i.e. intersection) probability statements hold with probability *at least* $1 - 2\alpha$