18.440: Lecture 4

Axioms of probability and inclusion-exclusion

Scott Sheffield

MIT

Outline

Axioms of probability

Consequences of axioms

Inclusion exclusion

Outline

Axioms of probability

Consequences of axioms

Inclusion exclusion

Axioms of probability

- $P(A) \in [0,1]$ for all $A \subset S$.
- $P(S) = 1$.
- Finite additivity: $P(A \cup B) = P(A) + P(B)$ if $A \cap B = \emptyset$.
- Countable additivity: $P(\cup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} P(E_i)$ if $E_i \cap E_j = \emptyset$ for each pair i and j.

- **Neurological:** When I think "it will rain tomorrow" the "truth-sensing" part of my brain exhibits 30 percent of its maximum electrical activity.
- **Frequentist:** $P(A)$ is the fraction of times A occurred during the previous (large number of) times we ran the experiment.
- **Market preference ("risk neutral probability"):** $P(A)$ is price of contract paying dollar if A occurs divided by price of contract paying dollar regardless.
- **Personal belief:** $P(A)$ is amount such that I'd be indifferent between contract paying 1 if A occurs and contract paying $P(A)$ no matter what.

Axiom breakdown

- What if personal belief function doesn't satisfy axioms?
- Consider an A-contract (pays 10 if candidate A wins election) a B-contract (pays 10 dollars if candidate B wins) and an A-or-B contract (pays 10 if either A or B wins).
- Friend: "I'd say A-contract is worth 1 dollar, B-contract is worth 1 dollar, A-or-B contract is worth 7 dollars."
- **Amateur response:** "Dude, that is, like, so messed up. Haven't you heard of the axioms of probability?"
- **Professional response:** "I fully understand and respect your opinions. In fact, let's do some business. You sell me an A contract and a B contract for 1.50 each, and I sell you an A-or-B contract for 6.50."
- Friend: "Wow... you've beat by suggested price by 50 cents on each deal. Yes, sure! You're a great friend!"
- Axioms breakdowns are money-making opportunities.

- **Neurological:** When I think "it will rain tomorrow" the "truth-sensing" part of my brain exhibits 30 percent of its maximum electrical activity. Should have $P(A) \in [0,1]$, maybe $P(S) = 1$, not necessarily $P(A \cup B) = P(A) + P(B)$ when $A \cap B = \emptyset$.
- **Frequentist:** $P(A)$ is the fraction of times A occurred during the previous (large number of) times we ran the experiment. Seems to satisfy axioms...
- **Market preference ("risk neutral probability"):** $P(A)$ is price of contract paying dollar if A occurs divided by price of contract paying dollar regardless. Seems to satisfy axioms, assuming no arbitrage, no bid-ask spread, complete market...
- **Personal belief:** $P(A)$ is amount such that I'd be indifferent between contract paying 1 if A occurs and contract paying $P(A)$ no matter what. Seems to satisfy axioms with some notion of utility units, strong assumption of "rationality"...

Outline

Outline

Axioms of probability

Consequences of axioms

Inclusion exclusion

Intersection notation

- We will sometimes write AB to denote the event $A \cap B$.

Consequences of axioms

- Can we show from the axioms that $P(A^c) = 1 - P(A)$?
- Can we show from the axioms that if $A \subset B$ then $P(A) \leq P(B)$?
- Can we show from the axioms that $P(A \cup B) = P(A) + P(B) - P(AB)$?
- Can we show from the axioms that $P(AB) \leq P(A)$?
- Can we show from the axioms that if S contains finitely many elements $x_1, \ldots, x_k$, then the values $\big(P(\{x_1\}), P(\{x_2\}), \ldots, P(\{x_k\})\big)$ determine the value of $P(A)$ for any $A \subset S$?
- What k-tuples of values are consistent with the axioms?

Famous 1982 Tversky-Kahneman study (see wikipedia)

- People are told "Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations."
- They are asked: Which is more probable?
  - Linda is a bank teller.
  - Linda is a bank teller and is active in the feminist movement.
- 85 percent chose the second option.
- Could be correct using neurological/emotional definition. Or a "which story would you believe" interpretation (if witnesses offering more details are considered more credible).
- But axioms of probability imply that second option cannot be more likely than first.

Outline

Outline

Axioms of probability

Consequences of axioms

Inclusion exclusion

Inclusion-exclusion identity

- Imagine we have n events, $E_1, E_2, \ldots, E_n$.
- How do we go about computing something like $P(E_1 \cup E_2 \cup \ldots \cup E_n)$?
- It may be quite difficult, depending on the application.
- There are some situations in which computing $P(E_1 \cup E_2 \cup \ldots \cup E_n)$ is a priori difficult, but it is relatively easy to compute probabilities of *intersections* of any collection of E_i. That is, we can easily compute quantities like $P(E_1 E_3 E_7)$ or $P(E_2 E_3 E_6 E_7 E_8)$.
- In these situations, the inclusion-exclusion rule helps us compute unions. It gives us a way to express $P(E_1 \cup E_2 \cup \ldots \cup E_n)$ in terms of these intersection probabilities.

Inclusion-exclusion identity

- Can we show from the axioms that $P(A \cup B) = P(A) + P(B) - P(AB)$?
- How about $P(E \cup F \cup G) = P(E) + P(F) + P(G) - P(EF) - P(EG) - P(FG) + P(EFG)$?
- More generally,

$$P(\cup_{i=1}^{n} E_i) = \sum_{i=1}^{n} P(E_i) - \sum_{i_1 < i_2} P(E_{i_1} E_{i_2}) + \ldots$$
$$+ (-1)^{(r+1)} \sum_{i_1 < i_2 < \ldots < i_r} P(E_{i_1} E_{i_2} \ldots E_{i_r})$$
$$+ \ldots + (-1)^{n+1} P(E_1 E_2 \ldots E_n).$$

- The notation $\sum_{i_1 < i_2 < \ldots < i_r}$ means a sum over all of the $\binom{n}{r}$ subsets of size r of the set $\{1, 2, \ldots, n\}$.

Inclusion-exclusion proof idea

- Consider a region of the Venn diagram contained in exactly $m > 0$ subsets. For example, if $m = 3$ and $n = 8$ we could consider the region $E_1 E_2 E_3^c E_4^c E_5 E_6^c E_7^c E_8^c$.
- This region is contained in three single intersections (E_1, E_2, and E_5). It's contained in 3 double-intersections ($E_1 E_2$, $E_1 E_5$, and $E_2 E_5$). It's contained in only 1 triple-intersection ($E_1 E_2 E_5$).
- It is counted $\binom{m}{1} - \binom{m}{2} + \binom{m}{3} + \ldots \pm \binom{m}{m}$ times in the inclusion exclusion sum.
- How many is that?
- Answer: 1. (Follows from binomial expansion of $(1-1)^m$.)
- Thus each region in $E_1 \cup \ldots \cup E_n$ is counted exactly once in the inclusion exclusion sum, which implies the identity.

Famous hat problem

- n people toss hats into a bin, randomly shuffle, return one hat to each person. Find probability nobody gets own hat.
- Inclusion-exclusion. Let E_i be the event that ith person gets own hat.
- What is $P(E_{i_1}E_{i_2}\ldots E_{i_r})$?
- Answer: $\frac{(n-r)!}{n!}$.
- There are $\binom{n}{r}$ terms like that in the inclusion exclusion sum. What is $\binom{n}{r}\frac{(n-r)!}{n!}$?
- Answer: $\frac{1}{r!}$.
- $P(\cup_{i=1}^n E_i) = 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \ldots \pm \frac{1}{n!}$
- $1 - P(\cup_{i=1}^n E_i) = 1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \ldots \pm \frac{1}{n!} \approx 1/e \approx .36788$

18.440 Probability and Random Variables
Spring 2014