Judgment aggregation

acknowledgment: Ulle Endriss, University of Amsterdam

Lirong Xia

Rensselaer

March 22, 2016

Your paper presentation(s)

- Two parts
  - presentation: about 1 hour
  - discussion: 30 min
- Meet with me twice before your presentation
  - 1st: discuss content covered in your presentation
  - 2st: go over the slides or notes
- Prepare reading questions for discussion
  - technical questions
  - high-level discussions: importance, pros, cons

Last class: Fair division

- Indivisible goods
  - house allocation: serial dictatorship
  - housing market: Top trading cycles (TTC)
- Divisible goods (cake cutting)
  - $n = 2$: cut-and-choose
  - discrete and continuous procedures that satisfies proportionality
  - hard to design a procedure that satisfies envy-freeness

Judgment aggregation: the doctrinal paradox

	Action p	Action q	Liable? ($p \wedge q$)
Judge 1	Y	Y	Y
Judge 2	Y	N	N
Judge 3	N	Y	N
Majority	**Y**	**Y**	**N**

- p: valid contract
- q: the contract has been breached
- Why paradoxical?
  - issue-by-issue aggregation leads to an illogical conclusion

Formal framework

- An agenda A is a finite nonempty set of propositional logic formulas closed under complementation ($[\varphi \in A] \Rightarrow [\sim\varphi \in A]$)
  - $A = \{ p, q, \sim p, \sim q, p \wedge q \}$
  - $A = \{ p, \sim p, p \wedge q, \sim p \vee \sim q \}$
- A judgment set J on an agenda A is a subset of A (the formulas that an agent thinks is true, in other words, accepts). J is
  - complete, if for all $\varphi \in A$, $\varphi \in J$ or $\sim\varphi \in J$
  - consistent, if J is satisfiable
  - $S(A)$ is the set of all complete and consistent judgment sets
- Each agent (judge) reports a judgment set
  - $D = (J_1,\ldots,J_n)$ is called a profile
- An judgment aggregation (JA) procedure F is a function $(S(A))^n \rightarrow \{0,1\}^A$

Do we want democracy or truth?

- Most previous work took the axiomatic point of view
- Seems truth is better for many applications
  - ongoing work

Some JA procedures

- Majority rule
  - F(φ)=1 if and only if the majority of agents accept φ
- Quota rules
  - F(φ)=1 if and only if at least k% of agents accept φ
- Premise-based rules
  - apply majority rule on "premises", and then use logic reasoning to decide the rest
- Conclusion-based rules
  - ignore the premises and use majority rule on "conclusions"
- Distance-based rules
  - choose a judgment set that minimizes distance to the profile

Axiomatic properties

- A judgment procedure F satisfies
  - unanimity, if [for all j, $\varphi \in J_j$]$\Rightarrow$[$\varphi \in \mathrm{F}(D)$]
  - anonymity, if the names of the agents do not matter
  - independence, if the decision for φ only depends on agents' opinion on φ
  - neutrality, [for all j, $\varphi \in J_j \Leftrightarrow \psi \in J_j$]$\Rightarrow$[$\varphi \in \mathrm{F}(D) \Leftrightarrow \psi \in \mathrm{F}(D)$]
  - systematicity, if for all D, D', φ, ψ [for all j, $\varphi \in J_j \Leftrightarrow \psi \in J_j'$]$\Rightarrow$[$\varphi \in \mathrm{F}(D) \Leftrightarrow \psi \in \mathrm{F}(D')$]
    - =independence + neutrality
  - majority rule satisfies all of these!

Example: Doctrinal paradox

	Action p	Action q	Liable? ($p \wedge q$)
Judge 1	Y	Y	Y
Judge 2	Y	N	N
Judge 3	N	Y	N
Majority	**Y**	**Y**	**N**

- Agenda A = { p, ~p, q, ~q, $p \wedge q$, ~p $\vee$ ~q}
- Profile D
  - J_1={p, q, $p \wedge q$}
  - J_2={p, ~q, ~p $\vee$ ~q}
  - J_3={~p, q, ~p $\vee$ ~q}
- JA Procedure F: majority
- F(D) = {p, q, ~p $\vee$ ~q}

Impossibility theorem

- Theorem. When n>1, no JA procedure satisfies the following conditions
  - is defined on an agenda containing {p, q, p$\wedge$q}
  - satisfies anonymity, neutrality, and independence
  - always selects a judgment set that is complete and consistent

Proof

- Anonymity + systematicity $\Rightarrow$ decision on φ only depends on number of agents who accept φ
- When n is even
  - half approve p half disapprove p
- When n is odd
  - (n-1)/2 approve p and q
  - (n-3)/2 approve ~p and ~q
  - 1 approves p
  - 1 approves q
  - # p = #q = # ~(p$\wedge$q)
    - approve all these violates consistency
    - approve none violates consistency

Avoiding the impossibility

- Anonymity
  - dictatorship
- Neutrality
  - premise-based approaches
- Independence
  - distance-based approach

Premise-based approaches

- $A = A_p + A_c$
  - A_p=premises
  - A_c=conclusions
- Use the majority rule on the premises, then use logic inference for the conclusions
- Theorem. If
  - the premises are all literals
  - the conclusions only use literals in the premises
  - the number of agents is odd
- then the premise-based approach is anonymous, consistent, and complete

	p	q	$(p \wedge q)$
Judge 1	Y	Y	Y
Judge 2	Y	N	N
Judge 3	N	Y	N
Majority	**Y**	**Y**	**Logic reasoning Y**

Distance-based approaches

- Given a distance function
  - d: $\{0,1\}^A \times \{0,1\}^A \rightarrow \mathrm{R}$
- The distance-based approach chooses

$$\text{argmin}_{J \in \mathrm{S}(A)} \Sigma_{J' \in D}\, \mathrm{d}(J, J')$$

- Satisfies completeness and consistency
- Violates neutrality and independence
  - c.f. Kemeny

Recap

- Doctrinal paradox
- Axiomatic properties of JA procedures
- Impossibility theorem
- Premise-based approaches
- Distance-based approaches

Hypothesis testing (definitions)

An example

- The average GRE quantitative score of
  - RPI graduate students vs.
  - national average: 558(139)
- Method 1: compute the average score of all RPI graduate students and compare to national average
- End of class

Another example

- Two heuristic algorithms: which one runs faster in general?
- Method 1: compare them on all instances
- Method 2: compare them on a few "randomly" generated instances

Simplified problem: one sample location test

- You have a random variable X
  - you know
    - the shape of X: normal
    - the standard deviation of X: 1
  - you don't know
    - the mean of X
- After observing one sample of X (with value x), what can you say when comparing the mean to 0?
  - what if you see 10?
  - what if you see 2?
  - what if you see 1?

Some quick answers

- Method 1
  - if x>1.645 then say the mean is strictly positive
- Method 2
  - if x<-1.645 then say the mean is strictly negative
- Method 3
  - if x<-1.96 or x>1.96 then say the mean is non-zero
- How should we evaluate these methods?

The null and alternative hypothesis (Neyman-Pearson framework)

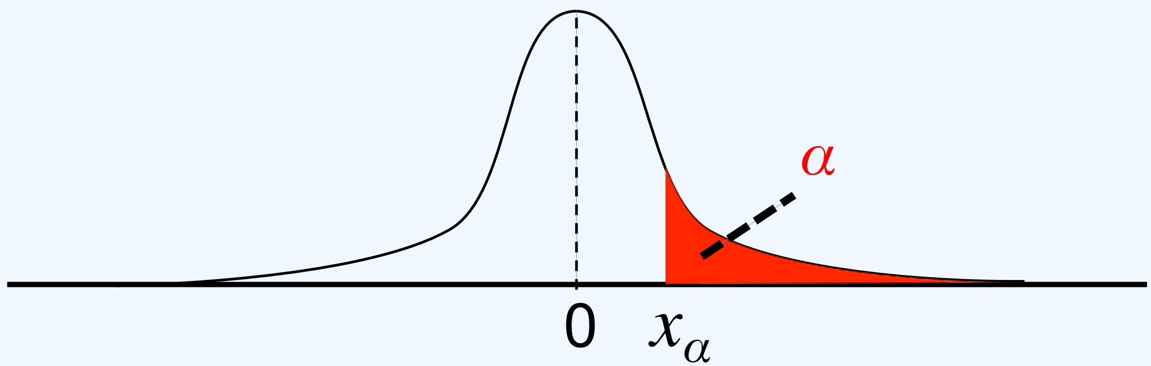
- Given a statistical model
  - parameter space: Θ
  - sample space: S
  - $\Pr(s|\theta)$
- H_1: the alternative hypothesis
  - $H_1 \subseteq \Theta$
  - the set of parameters you think contain the ground truth
- H_0: the null hypothesis
  - $H_0 \subseteq \Theta$
  - $H_0 \cap H_1 = \varnothing$
  - the set of parameters you want to test (and ideally reject)
- Output of the test
  - reject the null: suppose the ground truth is in H_0, it is unlikely that we see what we observe in the data
  - retain the null: we don't have enough evidence to reject the null

One sample location test

- Combination 1 (one-sided, right tail)
  - H_1: mean>0
  - H_0: mean=0 (why not mean<0?)
- Combination 2 (one-sided, left tail)
  - H_1: mean<0
  - H_0: mean=0
- Combination 3 (two-sided)
  - H_1: mean≠0
  - H_0: mean=0
- A hypothesis test is a mapping f : S⟶{reject, retain}

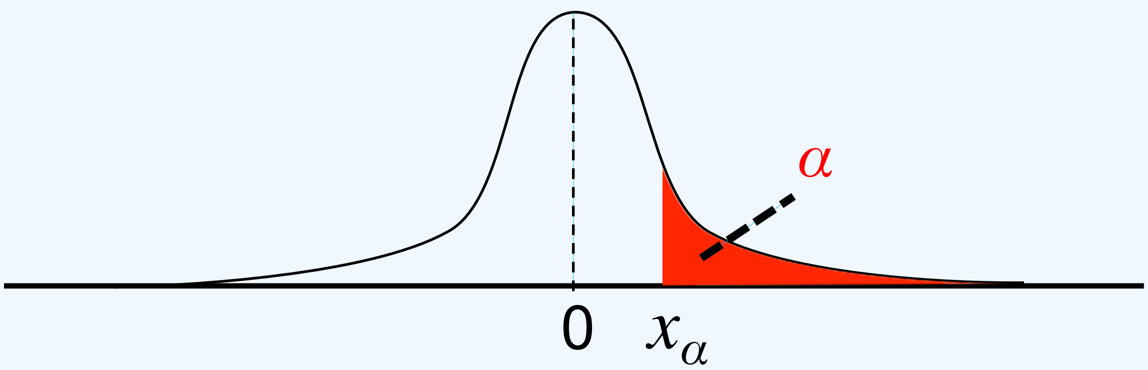
One-sided Z-test

- H_1: mean>0
- H_0: mean=0
- Parameterized by a number $0<\alpha<1$
  - is called the level of significance
- Let x_α be such that $\Pr(X>x_\alpha|H_0)=\alpha$
  - x_α is called the critical value



- Output reject, if
  - $x>x_\alpha$, or $\Pr(X>x|H_0)<\alpha$
    - $\Pr(X>x|H_0)$ is called the p-value
- Output retain, if
  - $x\leq x_\alpha$, or p-value$\geq\alpha$

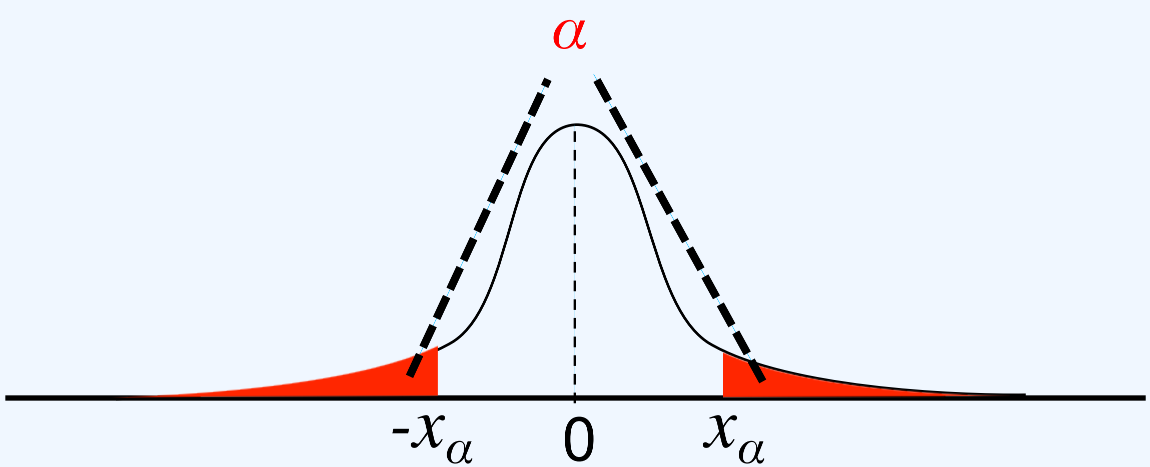
Interpreting level of significance



- Popular values of α:
  - 5%: x_α= 1.645 std (somewhat confident)
  - 1%: x_α= 2.33 std (very confident)
- α is the probability that given mean=0, a randomly generated data will leads to "reject"
  - Type I error

Two-sided Z-test

- H_1: mean≠0
- H_0: mean=0
- Parameterized by a number $0<\alpha<1$
- Let x_α be such that $2\Pr(X>x_\alpha|H_0)=\alpha$



- Output reject, if
  - $x>x_\alpha$, or $x<x_\alpha$
- Output retain, if
  - $-x_\alpha \le x \le x_\alpha$

What we have learned so far…

- One/two-sided Z test: hypothesis tests for one sample location test (for different H_1's)
- Outputs either to “reject” or “retain” the null hypothesis
- And defined a lot of seemingly fancy terms on the way
  - null/alternative hypothesis
  - level of significance
  - critical value
  - p-value
  - Type I error

Questions that haunted me when I first learned these

- Isn't point estimation H_0 never true?
  - the "chance" for the mean to be exactly 0 is negligible
  - fine, but what made you believe so?
- What the heck are you doing by using different H_1?
  - the description of the tests does not depend on the selection of H_1
  - if we reject H_0 using one-sided test (mean>0), shouldn't we already be able to say mean≠0? Why need two-sided test?
- What the heck are you doing by saying "reject" and "retain"
  - Can't you just predict whether the ground truth is in H_0 or H_1?

Next class

- Evaluation of hypothesis testing methods
- Statistical decision theory