Social Choice

CSC200 Lecture 41
Winter 2016

Allan Borodin

Announcements and todays agenda

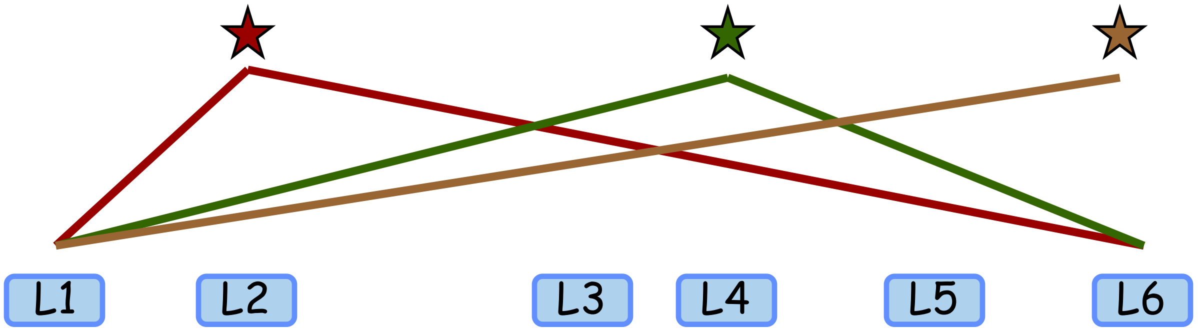
- Today: Continue single peaked preferences and then Information Aggregation
- Announcements
  - **Assignment 4 due next Wed, March 30.**
  - **Very interesting talk today on algorithmic social choice. Two topics mentioned with web sites**
    - **Robovote.org (to be launched in May, 2016) This will provide a non profit site for "ordinary people" to vote on their preferences.**
    - **Spliddit.org Non profit site in operation for about a year with 60,000 users. This enables fair division of resources and continues to evolve by interaction with users**

Are we doomed to possible manipulation?

- Unlike the prevous impossibility theorems, the axioms in the Gibbard Satherthwaite Theorem seem very reasonable.
- But the theorem does imply that all preference profiles are possible which in many applications is not the case.
- Moreover, one of the insights of algorithmic social choice is that while certain voting rules can be manipulated, it may be computationally hard to determine how this manipulation can be done.

Single-peaked Preferences

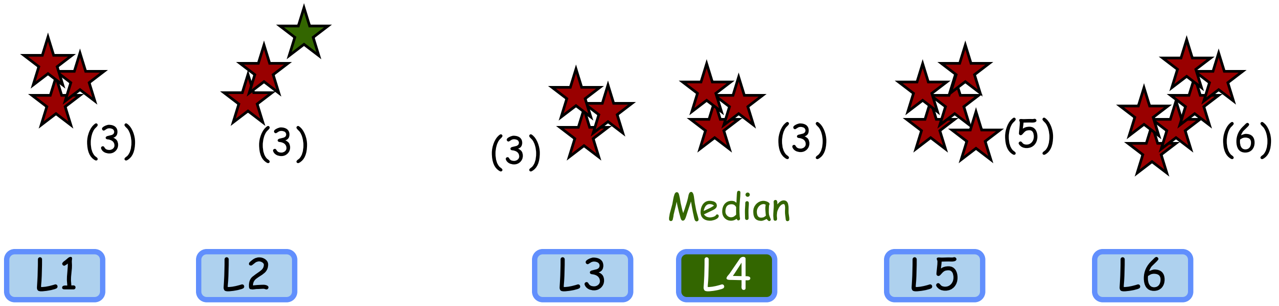
- Special class of preferences for which GS Theorem is circumvented
- Let $\gg$ denote some "natural" ordering over alternatives *A*
  - e.g., order political candidates on left-right spectrum
  - e.g., locations of park, warehouse on a line (e.g., position on a highway)



- Voter *k*'s preferences are *single-peaked* if there an *ideal* alternative, $a^*[k]$, that *k* likes best, and that as you move away from $a^*[k]$ in the ordering $\gg$, alternatives become less and less preferred by *k;* that is:
  - $a^*[k] \succ_k a$ for any $a \neq a^*[k]$
  - $b \succ_k c$ if either: (1) $c \gg b \gg a^*[k]$; or (2) $a^*[k] \gg b \gg c$
- In figure: green voter (★) prefers $L4 \succ L3 \succ L2 \succ L1$ and $L4 \succ L5 \succ L6$

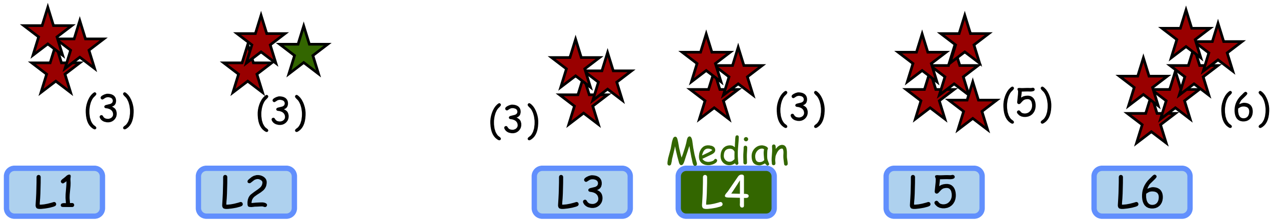
Median Voting

- Suppose all voter's prefs are single-peaked
  - they must be single-peaked w.r.t. the same domain ordering ≫
  - but you can use any ordering you want (as long as it creates SP'ed prefs)
- *Median voting scheme:* voter specifies only her peak; winner is *median* of the reported peaks (Black 1948)



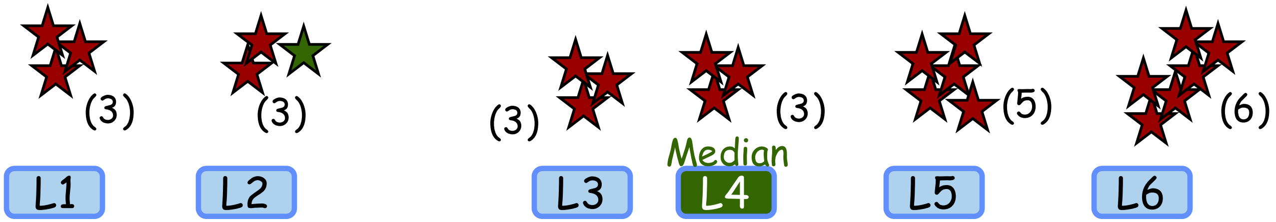
What's Special about Median Voting?

- Assume single-peaked preferences and use median voting
- The winner *W* is Pareto efficient *(in example L4)*
    - *no other choice is better for one person without hurting someone else*
- The winner *W* is a Condorcet winner (if *n* odd): Why?
    - at least *(n+1)/2* voters prefer *W* to anyone *left* of *W* (more if there is more than one voter's peak at the median)
    - at least *(n+1)/2* voters prefer *W* to anyone *right* of *W* (more if there is more than one voter's peak at the median)
    - ***so W wins a majority election against any other candidate***
- Known as the *Median Voter Theorem*



What’s Special about Median Voting?

- Can take Median Voter Theorem a step further, imagine following procedure:
  - place *W* at top of societal ranking, then remove it from candidate set
  - repeat process to find median winner among *remaining* candidates
    - there again must be a Condorcet winner (!)
    - *in example: peaks for all voters stays the same except for those who voted for L4 (those voters each have a new peak, either L3 or L5)*
  - remove and repeat until you’ve ranked all candidates
- Societal ranking must be complete and transitive and respects majoritarian preferences: if A >B in ranking, the majority prefer A to B
  - breaks the Condorcet paradox



The complexity of manipulation

- As mentioned, one of the insights of algorithmic game theory and social choice theory is that one has to take into consideration the complexity of any proposed solution (or a criticism of a solution).
- Depending on the precise voting rule, manipulation may be easy to achieve or (according to our beliefs about NP hard problems) may not be efficiently achievable.
- Determining which) if any voter can manipulate (i.e. change their vote to make a candidate p the winner) is:
  - relatively efficient (i.e. doable in polynomial time) for all positional scoring rules.
  - Seemingly inefficient (i.e. NP hard) for STV
- Determining if there is a coalition of k voters that can manipulate is NP-hard for k > 1 for almost all natural scoring rules with the exception of plurality.

Voting as Information Aggregation

- So far, we've assumed that voting is used to aggregate rankings of people with genuinely different tastes/preferences for alternatives
  - e.g., you prefer Red Bull (or sociology) and I prefer Cheetah Power Surge (or calculus); not a question of which of us is "right" or "wrong"
- However, sometimes voting is used to determine or judge more "objective" facts
  - juries voting on guilt or innocence of defendant.
  - Wisdom of the Crowds, Prediction Markets (Ch.22)
    - polls: who is likely to win an election, sporting contest, …
    - polls to determine other objective facts (e.g., which street in GTA has deepest, most damaging potholes, which highway is fastest, …)
  - Crowdsourcing technologies (e.g., run market to label images, assess quality of text/translations, etc.): Mechanical Turk, CrowdFlower, …

Voting as Information Aggregation

- Many similarities to prediction markets, info cascades, but also some key differences
- Prediction Markets:
  - similarities: aggregating beliefs, will consider simultaneous and sequential models
  - don’t ask people to state (or bet based on) probabilities, but simply cast a vote (Yes or no? Sunny, cloudy, rainy? …)
- Information Cascades:
  - similarities: aggregating beliefs, asking people to vote (or “decide”) based on beliefs rather than state probabilities
  - simultaneous voting doesn’t permit cascades (though sequential could, as we will see)

Our Old Friend: Bayes Rule

- Provides precise means to update our beliefs about some hypothesis in the face of (new) evidence
- **Hypothesis H:**
  - new restaurant is “good”; patient has malaria; 2/3 marbles are blue
- **Evidence (or signal) E:**
  - positive review on Yelp; patient has fever; drew a red marble
- **Given prior probability of hypothesis P(H):**
  - *Pr(RestGood) = 0.3*; *Pr(Malaria) = 0.001*; *Pr(MajBlue) = 0.5*
- **Given probability of evidence given hypothesis P(E|H):**
  - *Pr(+Rev|Good) = 0.9*; *Pr(Fev|Malar) = 0.95*; *Pr(Red|MajBlue) = 1/3*
  - usually written in this direction because of *causal/stable* connection
- We desire: *Pr(H | E) = Pr(E | H) Pr(H) / Pr(E)*
  - *Pr(E)* can be treated as a normalizing constant if we compute all competing hypotheses

Bayes Rule in Action (A Review)

- *Pr(Malaria) = 0.001;* *Pr(no Mal) = 1 – 0.001 = 0.999*
- *Pr(Fever | Mal) = 0.9;* *Pr(Fever | no Mal) = 0.05*
  - *Now we observe someone with Fever: what are odds of Malaria?*

$$\Pr(Mal \mid Fev) = \frac{1}{\Pr(Fev)} \Pr(Fev|Mal) \Pr(Mal) = \frac{0.9 \cdot 0.001}{\Pr(Fev)} = \frac{0.0009}{\Pr(Fev)}$$

$$\Pr(NoMal \mid Fev) = \frac{1}{\Pr(Fev)} \Pr(Fev|NoMal) \Pr(NoMal) = \frac{0.05 \cdot 0.999}{\Pr(Fev)} = \frac{0.04995}{\Pr(Fev)}$$

- *Pr(Fev)* is just a normalizing constant (same in both)
  - notice: *Pr(M|F) + Pr(NM|F) = 1*; and …
  - … *Pr(F) = Pr(F|M)Pr(M) + Pr(F|NM)Pr(NM) = 0.0009+0.04995 = 0.05085*
  - … key is *likelihood ratio Pr(F|M)Pr(M) : Pr(F|NM)Pr(NM)*

$$\Pr(Mal \mid Fev) = \frac{0.0009}{0.05085} = 0.0177; \qquad \Pr(NoMal \mid Fev) = 1 - 0.0177 = 0.9823$$

Bayes Rule in Action (A Review)

- *Pr(MajBlue) = 0.5; Pr(MajRed) = 0.5*
- *Pr(Blue | MajBlue) = 2/3; Pr(Blue | MajRed) = 1/3*

$$\Pr(MBlue|\ Blue) = \frac{1}{\Pr(Blue)} \Pr(Blue|MBlue) \Pr(MBlue) = \frac{2/3 \cdot 1/2}{\Pr(Blue)}$$

$$\Pr(MRed|\ Blue) = \frac{1}{\Pr(Blue)} \Pr(Blue|MRed) \Pr(MRed) = \frac{1/3 \cdot 1/2}{\Pr(Blue)}$$

- With *equal* prior for two hypotheses, likelihood ratio is just ratio of odds of evidence under one hypothesis vs. the other

$$\frac{\Pr(MBlue|Blue)}{\Pr(MRed|Blue)} = \frac{2/3 \cdot 1/2}{1/3 \cdot 1/2} = \frac{2/3}{1/3} = \frac{2}{1}$$

- So *Pr(MBlue|Blue) = 2/3*

Voting as Information Aggregation

- Suppose you need to assess two competing hypotheses, X and Y
  - exactly one of X or Y must be true
- For simplicity, X and Y are equally likely: your prior $Pr(X) = Pr(Y) = ½$
- You gather a collection of "experts" to *vote* on X and Y
  - e.g., doctors diagnosing a rare disease, jury voting on guilt/innocence, crowd predicting winner of election, stock traders predicting success or failure of a company

Voting as Information Aggregation

- Experts each independently assess *X*, *Y*
  - conduct their own evidence gathering
  - … or use their own thought processes, interpretation of evidence
- Experts are *fallible*, but more likely to vote correctly than not:
  - $Pr(voteX \mid X) = q > ½$ and $Pr(voteY \mid X) = 1\text{-}q < ½$
  - $Pr(voteY \mid Y) = q > ½$ and $Pr(voteX \mid Y) = 1\text{-}q < ½$
- You must select and act on one of the two hypotheses. Is using the majority vote a good idea?

Majority Voting

- With an equal prior, *for any number of experts*, the hypothesis the majority vote being the winner is more likely to be true than the other
  - e.g., suppose 9 voters: 5 votes for *X*, 4 votes for *Y*

$$\Pr(X|\,5\,\text{"X"},4\,\text{"}Y\text{"}) = \frac{1}{\mathrm{Z}}\Pr(5X,4Y|X)\Pr(X) = \frac{q^5(1-q)^4}{\mathrm{Z}}\cdot\frac{1}{2}$$
$$\Pr(Y|\,5\,\text{"X"},4\,\text{"}Y\text{"}) = \frac{1}{\mathrm{Z}}\Pr(5X,4Y|Y)\Pr(Y) = \frac{q^4(1-q)^5}{\mathrm{Z}}\cdot\frac{1}{2}$$

  - likelihood ratio *X:Y* greater than 1 (since *q* > *1-q*), so *X* more probable

$$\frac{\Pr(X|\,5\,\text{"X"},4\,\text{"}Y\text{)}}{\Pr(Y|\,5\,\text{"X"},4\,\text{"}Y\text{)}} = \frac{q^5(1-q)^4}{q^4(1-q)^5} = \frac{q}{1-q} > 1$$

Condorcet Jury Theorem

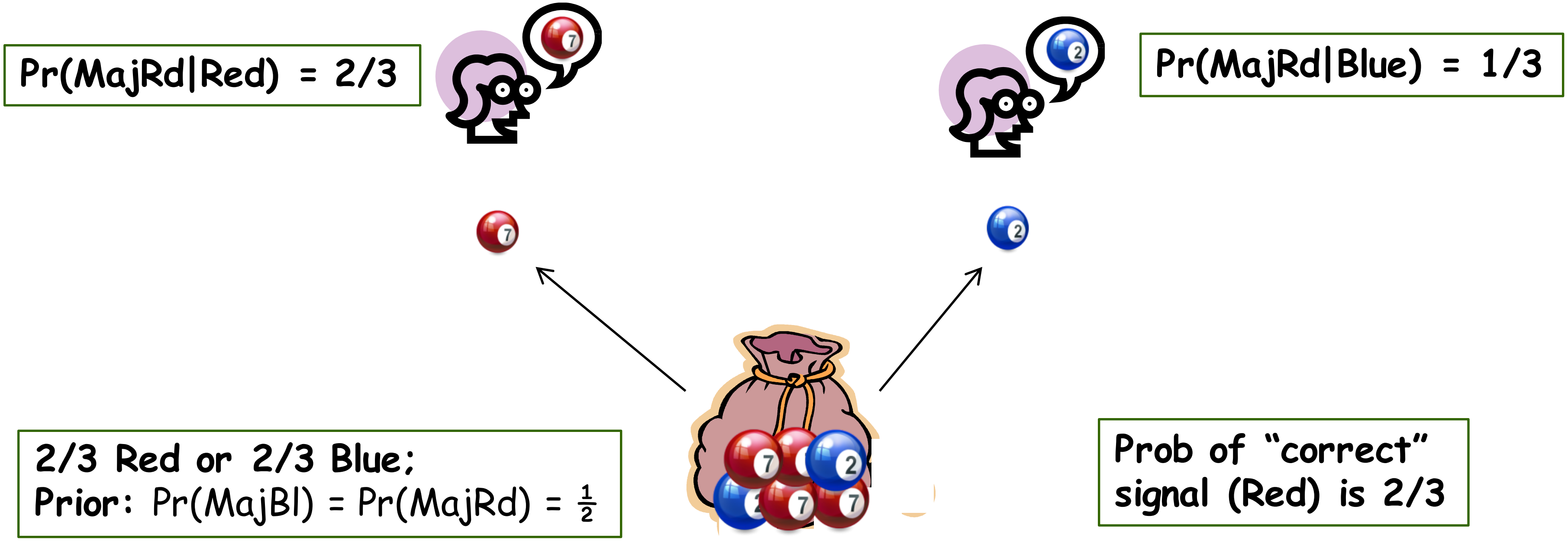
- Not only is majority winner more likely to be true, but if *q* > ½, then as number of experts grows, the probability of making the correct decision converges to 1. (See Chernoff bounds.)
  - this observation dates back to Condorcet in 1785
  - discussed in the context of jury trials, hence the name
- Some examples (compute using a binomial calculator)
  - *e.g., probability of getting right decision with nine experts and q= 0.6 is like asking "What is probability, if your flip nine biased coins with Pr(heads)=0.6,* ***that five or more*** *come up heads?"*

Odds of Correct Majority Vote

	9 experts	15 experts	101 experts	1001 experts
q = 0.51	0.5246	0.5314	0.5799	0.7366
q = 0.6	0.7334	0.7869	0.9791	~1.0
q = 0.9	0.9991	~1.0	~1.0	~1.0

Modeling Expert Evidence

- It could be that the expert's votes are determined by their own evidence, which provides a noisy signal about the truth
- If so, their beliefs (hence votes) are determined by Bayes Rule too
  - Remember information cascades (but without sequencing)



Pr(MajRd|Red) = 2/3

Pr(MajRd|Blue) = 1/3

2/3 Red or 2/3 Blue;
Prior: Pr(MajBl) = Pr(MajRd) = ½

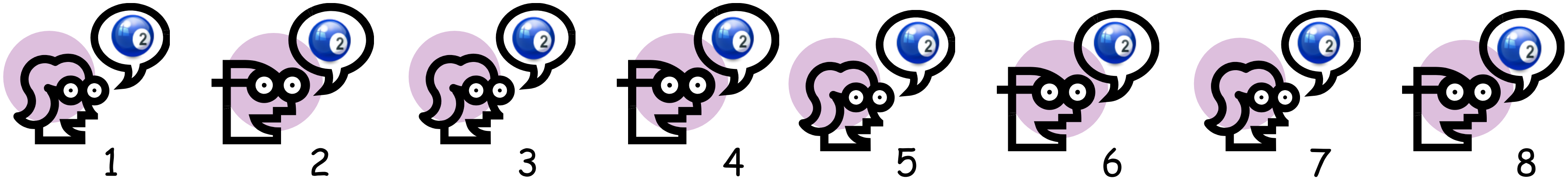
Prob of "correct" signal (Red) is 2/3

Connection to Information Cascades

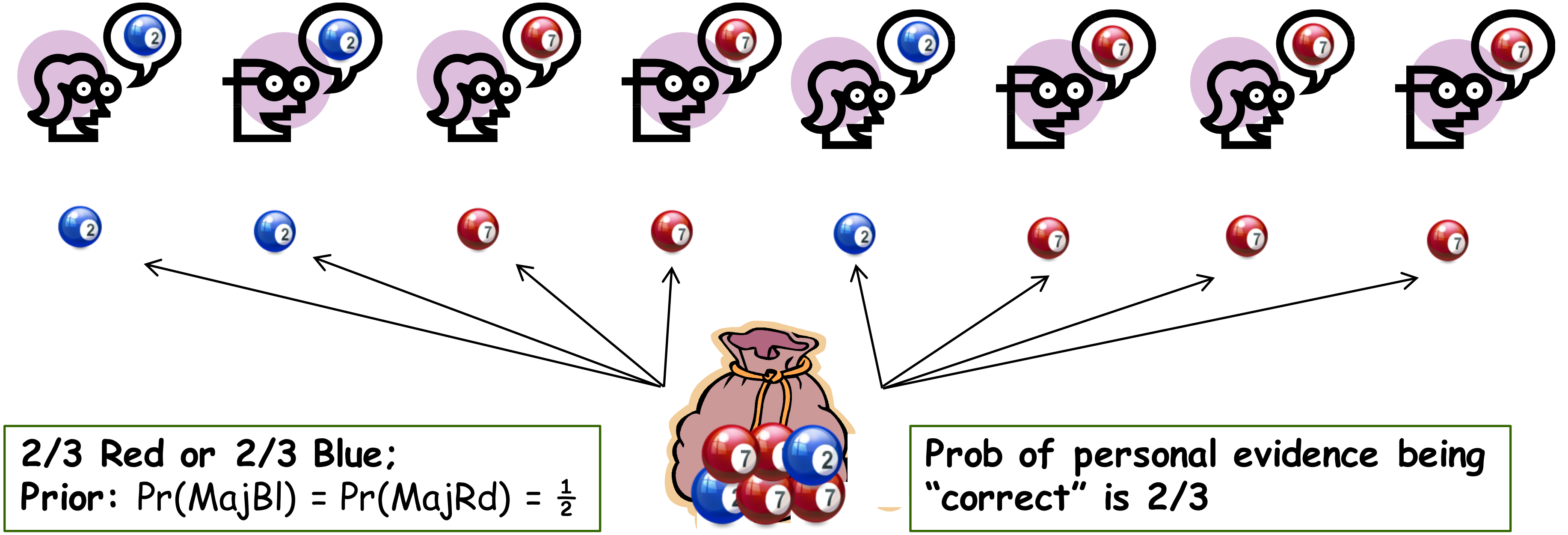
- Key difference with our model of information cascades (Ch.16)
  - in herding experiments, people voted (or made decisions) in *sequence* rather than *simultaneously*, so votes are not independent
  - jury theorem analysis depends critically on ***independent voting***
- What effect does sequencing have on voting?
  - in sequential voting, voter *k* weighs her “personal” observations (evidence) against what she infers about the about the observations made by those (*1* thru *k-1*) who voted before her
  - this influences her vote in such a way that her observation may not play a role in her decision at all
  - in the extreme—if a cascade forms starting with voter *k*—then the evidence of all subsequent voters is completely lost

Sequential vs. Simultaneous Voting

Sequential → Majority Vote is Blue



Simultaneous → Majority Vote is Red



2/3 Red or 2/3 Blue;
Prior: Pr(MajBl) = Pr(MajRd) = ½

Prob of personal evidence being "correct" is 2/3

Connection to Information Cascades

- Generally, probability of wrong decision is much higher with sequential voting
  - once a cascade forms, none of the remaining voter's evidence is ever taken into account: the previous votes completely determine what the remaining voters say (independent of their evidence)
  - e.g., in **2/3 Red, 1/3 Blue example**, probability of first two experts announcing Blue is 1/9, at which point *an incorrect cascade* has formed
  - so probability of incorrect decision is at least 1/9
    - in fact, it is much greater, since incorrect cascade can form *later* if it doesn't happen with expert 1 and 2
  - adding more voters does nothing to decrease this probability of error
- Simultaneous voting is more reliable
  - adding more voters only improves the situation
- **Notice we assume experts vote for mostly likely hypothesis!!**
  - if experts are trying to ensure a correct "decision", *picture can change…*

Strategic Experts

- We assumed expert votes for most likely outcome given her personal beliefs based on own evidence (or interpretation), i.e., “truthfully”
- But suppose expert E_1 knows how judge will make decision, and wants to vote *strategically* to ensure the decision is the best one or the “right” one (according to her own beliefs)
- Quick warm up (either *X* or *Y* is true):
  - assume judge uses majority vote for decision, odd number *n* of experts
  - does it still make sense for E_1 to vote “truthfully”?
  - if more than half (i.e., $(n+1)/2$ or more) vote *X* (or *Y*), E_1’s vote won’t affect decision, so she should (or “may as well”) vote truthfully
  - if exactly half vote *X* and half *Y*, then her vote breaks the tie; so she *should* vote truthfully

Why not majority voting?

- Should decisions be made on basis of “most likely” events?
  - would you go to a restaurant with a 51% of being good, 49% awful
  - would you send somebody to a life in prison if you believed they had a 51% chance of being guilty
- Decisions are based on probabilities of outcomes *and consequences* (the latter modeled using utilities or valuations)
- Hence we use *unanimity* (12/12 must vote guilty) in most jury trials
  - Why? Convicting an innocent person considered more negative than letting a guilty offender go free (at least by most people, societies)
  - “Beyond a reasonable doubt”, “burden of proof”, etc.
  - Look back at table: if each juror votes independently and has a 0.6 chance of getting it right, going with simple majority still has
    - 27% chance of wrong decision with 9 jurors
    - 21% of getting it wrong with 15 jurors
    - Is that “beyond a reasonable doubt”?

Strategic Voting under Unanimity

- Suppose juror E_1 believes INN (innocent): how to vote?
  - If she votes INN, then person will be acquitted surely
  - But if anyone else votes INN, person acquitted regardless of her vote
  - The only time her INN vote changes the decision is if all 11 others vote GLTY (guilty); we say her vote is *pivotal* in such a case
  - But if all others vote GLT, assuming independence of their votes, then person is most likely guilty (outweighs her personal evidence, just like in sequential voting, information cascades)
  - So the only rational thing to do is to vote GLTY
  - If everyone reasons like E_1, even the innocent are guaranteed to go to jail!

Swing Voter's Curse

- This phenomenon sometimes called the *"Swing voter's curse":*
  - knowing your vote is pivotal means you should probably vote differently than if you knew nothing about the votes of others
  - sometimes used to explain voting abstention ("uninformed" voters)

- Doesn't depend on unanimity: even in majority voting it can arise
  - e.g., suppose city counsellors always re-elected unless they are idiots
    - people pay little att'n, only turn out (or vote against) if idiotic behavior
  - suppose you have little knowledge of incumbent Joe
  - If you are a *swing voter*, then 50% of voters think Joe is an idiot
  - If 50% think Joe is an idiot, then he's probably an idiot
  - So you should probably vote against Joe (even if you know nothing, or even if you believe he's competent)

- It all sounds convincing, but…

Equilibrium Voting

- What's wrong with this reasoning?

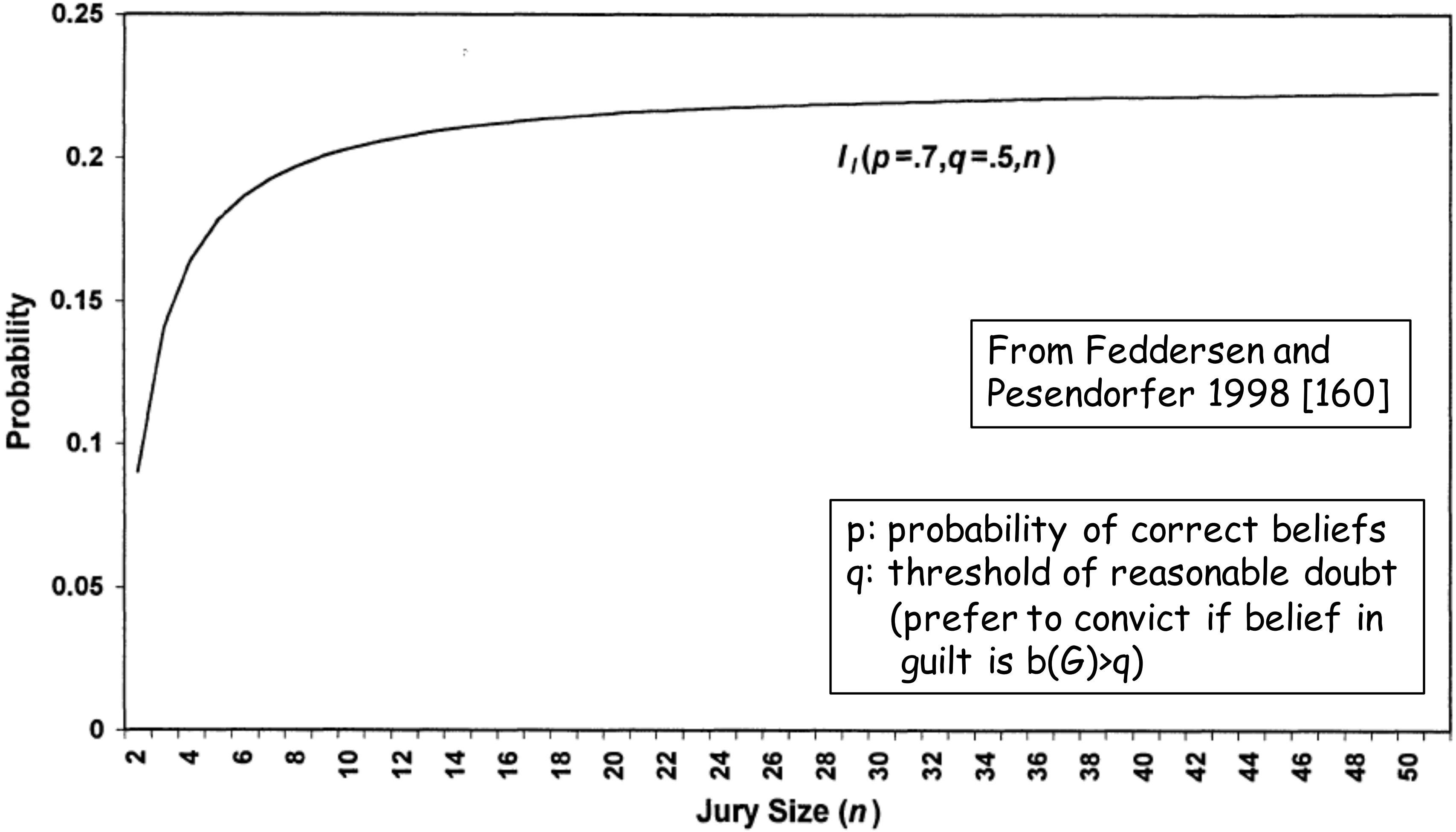
- E_1 votes strategically (GLTY) *assuming everyone else votes truthfully*
- But this undesirable result (everyone voting GLTY) *only occurs when everyone votes strategically*
- So if E_1 thinks this way, she should consider that others will think this way
- In other words, we need to consider a Nash equilibrium in a voting game
  - much like the equilibrium analysis in (first-price) auctions
  - BTW, remember winner's curse?

Symmetric Equilibria

- There may be many Nash equilibria (like in 1st price auctions)
- What are possible (symmetric) equilibria?
- Everyone voting GLTY is not an equilibrium; E_1's incentive to vote GLTY (and *ignore* her beliefs) depended on others voting truthfully
- Everyone voting INN ***is*** an equilibrium (jurors ignore their beliefs, nobody ever goes to jail)
- There is a unique mixed Nash eq. where jurors randomly vote (the vote probabilities depend on probabilities of jurors having "correct" beliefs)
  - If *E* believes INN, will vote GLTY with some probability anyway (to correct for chance of being the only INN vote
  - But if *E* believes GLTY, will vote GLTY
- *One issue:* unlike nonstrategic voting, the probability of getting the right decision does not converge to 1 as you increase the number of voters

Equilibrium Voting

FIGURE 1. The Probability an Innocent Defendant Is Convicted as a Function of Jury Size

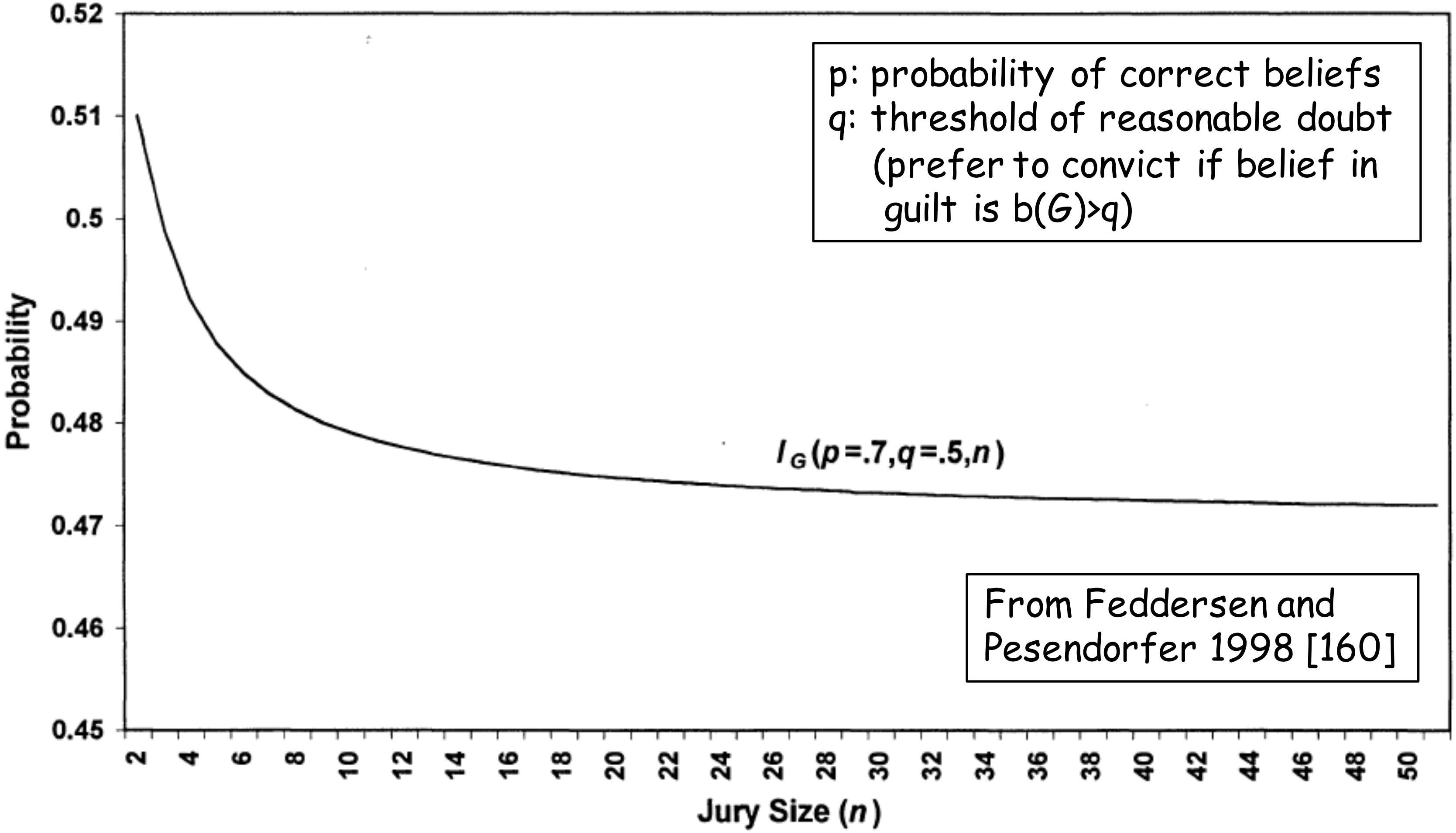


From Feddersen and Pesendorfer 1998 [160]

p: probability of correct beliefs
q: threshold of reasonable doubt
(prefer to convict if belief in guilt is b(G)>q)

Equilibrium Voting

FIGURE 2. **The Probability a Guilty Defendant Is Acquitted as a Function of Jury Size**



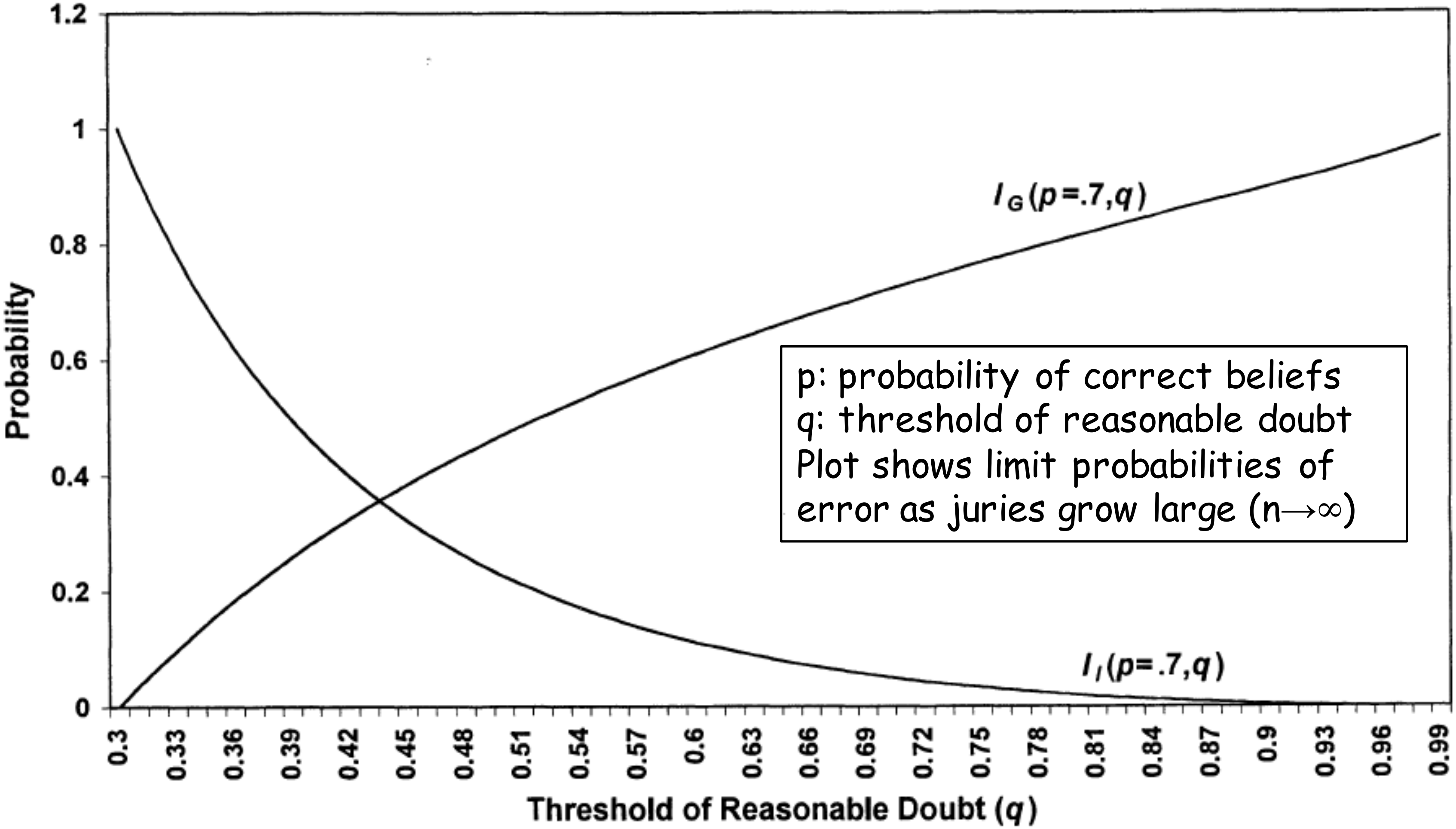
p: probability of correct beliefs
q: threshold of reasonable doubt (prefer to convict if belief in guilt is b(G)>q)

From Feddersen and Pesendorfer 1998 [160]

Equilibrium Voting

From Feddersen and Pesendorfer 1998 [160]

FIGURE 3. Limit Error Probabilities as a Function of the Threshold of Reasonable Doubt



p: probability of correct beliefs
q: threshold of reasonable doubt
Plot shows limit probabilities of error as juries grow large ($n \to \infty$)

Threshold Voting

- Instead of unanimity, we can instead require that only a (reasonable) fraction of experts to vote GLTY, call this alpha-majority
- Then you get more reasonable behavior; **Feddersen and Pesendorfer [160]**
- Specifically, for alpha between 0,1, the probability of a correct decision does converge to 1 as the jury grows
  - This is true despite the fact that experts vote strategically, assuming a symmetric equilibrium (not counting the trivial equilibrium where everyone votes INN)