Adversarial Knapsacks

Morad Elsaify, Benjamin Ewing, Usa Kerdnunvong, Kieran Manzella, Neil Pruthi

CS 590.02
Final Project Presentation
November 28, 2018

Premise

- n players
  - Budget: b_i
- m items
  - Price: p_j
  - Value: v_j
    - Random variable with known distribution
- Utility:
  - Winners are those that choose the knapsack(s) that take the highest value
  - Win: $n - 1$ (split among winners)
  - Lose: -1
  - If all players tie, each receives 0
- Knapsack formation
  - Sequential
  - Simultaneous

Applications

- Talent acquisition
  - Job candidate selection
    - Variance in ability of candidates
  - College admissions
    - Grades, standardized scores, and extracurricular activities are imperfect signals for future student success
  - Sports team recruiting
  - Fantasy sports leagues
    - High performance variance and unknown distribution for new players while older players have previous years' statistics
- Patents and research publications
  - Cannot grant two patents for the same product
  - First company to discovery wins

A bit of theory

Let X_1 and X_2 be independent random variables with symmetric probability distribution functions with $E[X_1] = E[X_2]$

$\rightarrow P[X_1 > X_2] = P[X_2 > X_1] = 0.5$

So...variance does **not** matter -- **only** expected values do!

The sum of symmetric random variables is also symmetric, so this holds for sums of symmetric random variables as long as expected values are the same.

Implications

Intuitively, expected value is the most important consideration

So, how do variance and other factors come in?

- Games with more than two players
- Items with asymmetric distributions

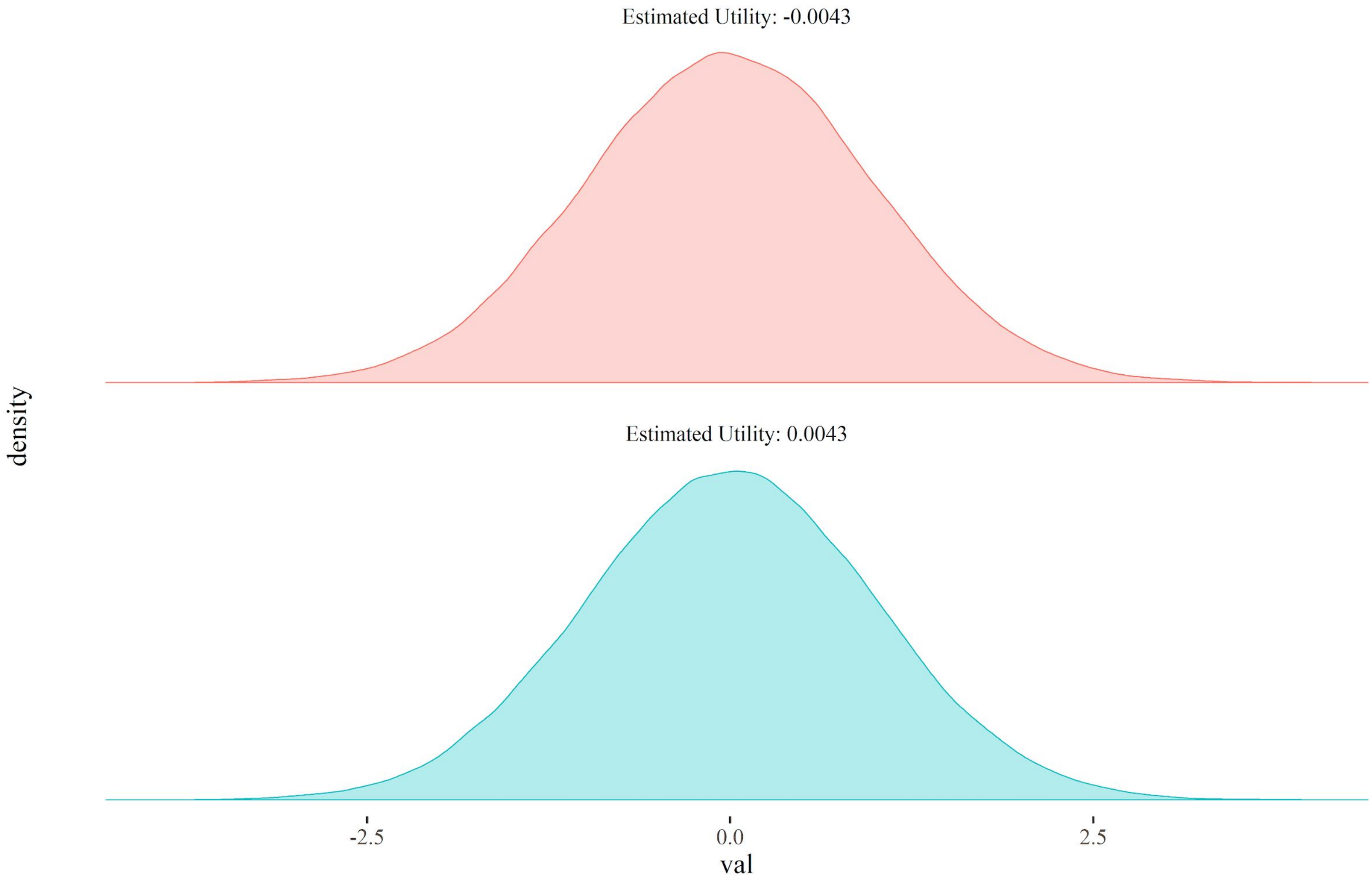
Let’s talk about simultaneous games a little more

Let’s begin with a very simple case.

Suppose we have two players and two items and knapsacks that can only fit one item each.

And each item has the same expected value and variance.

Which item would you put in your knapsack?

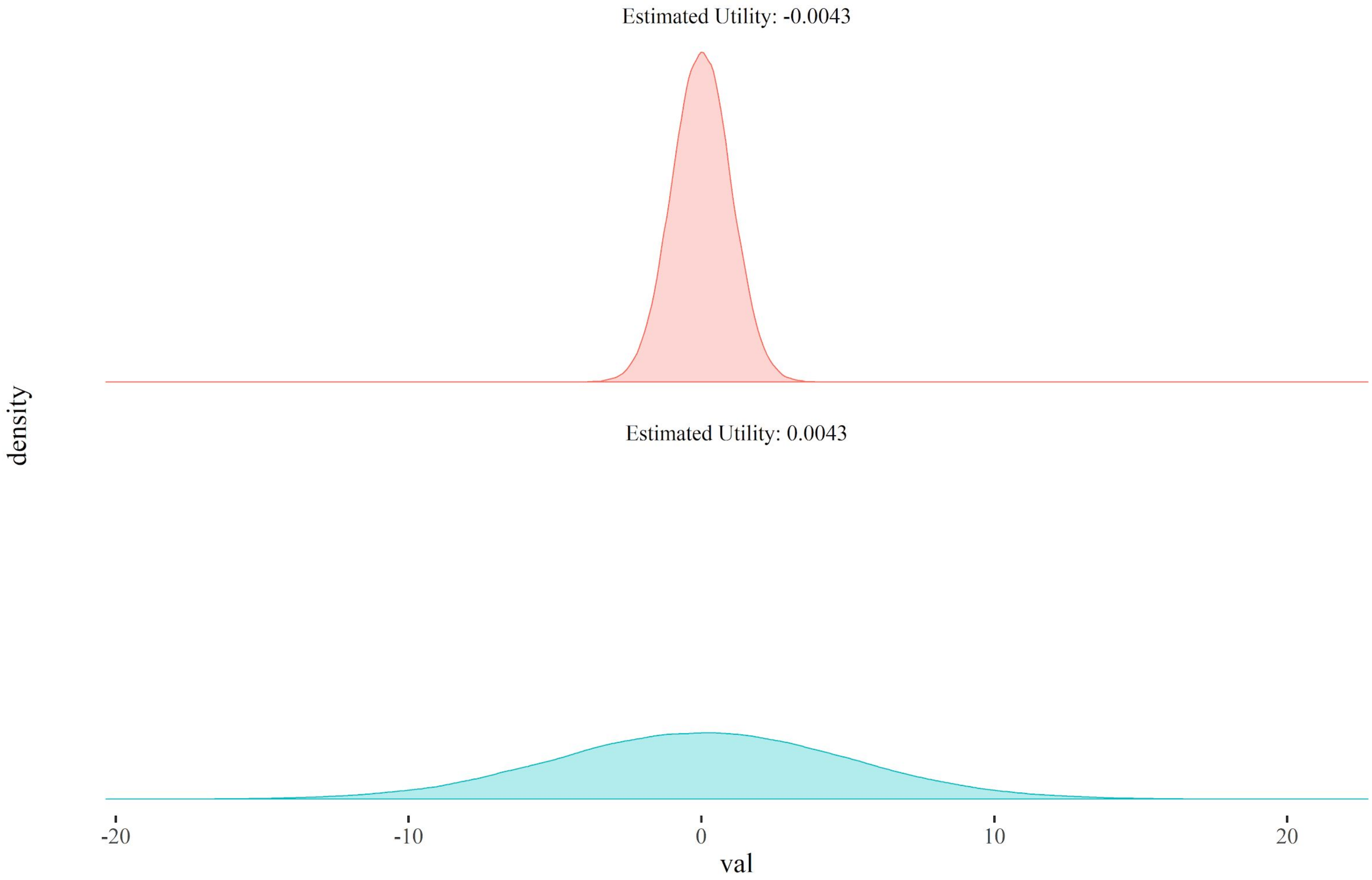

Estimated Utility: -0.0043
Estimated Utility: 0.0043
density
-2.5
0.0
2.5
val

Alright, now what if we have the same game, but items have different variances?

Now what might we expect to happen?

Well, recalling what Usa talked about, variance shouldn't matter here.

After all, we have two players, and items have equal expected values and symmetric distributions.


Estimated Utility: -0.0043
Estimated Utility: 0.0043
density
-20
-10
0
10
20
val

Let’s talk about more than two players, and more than two items

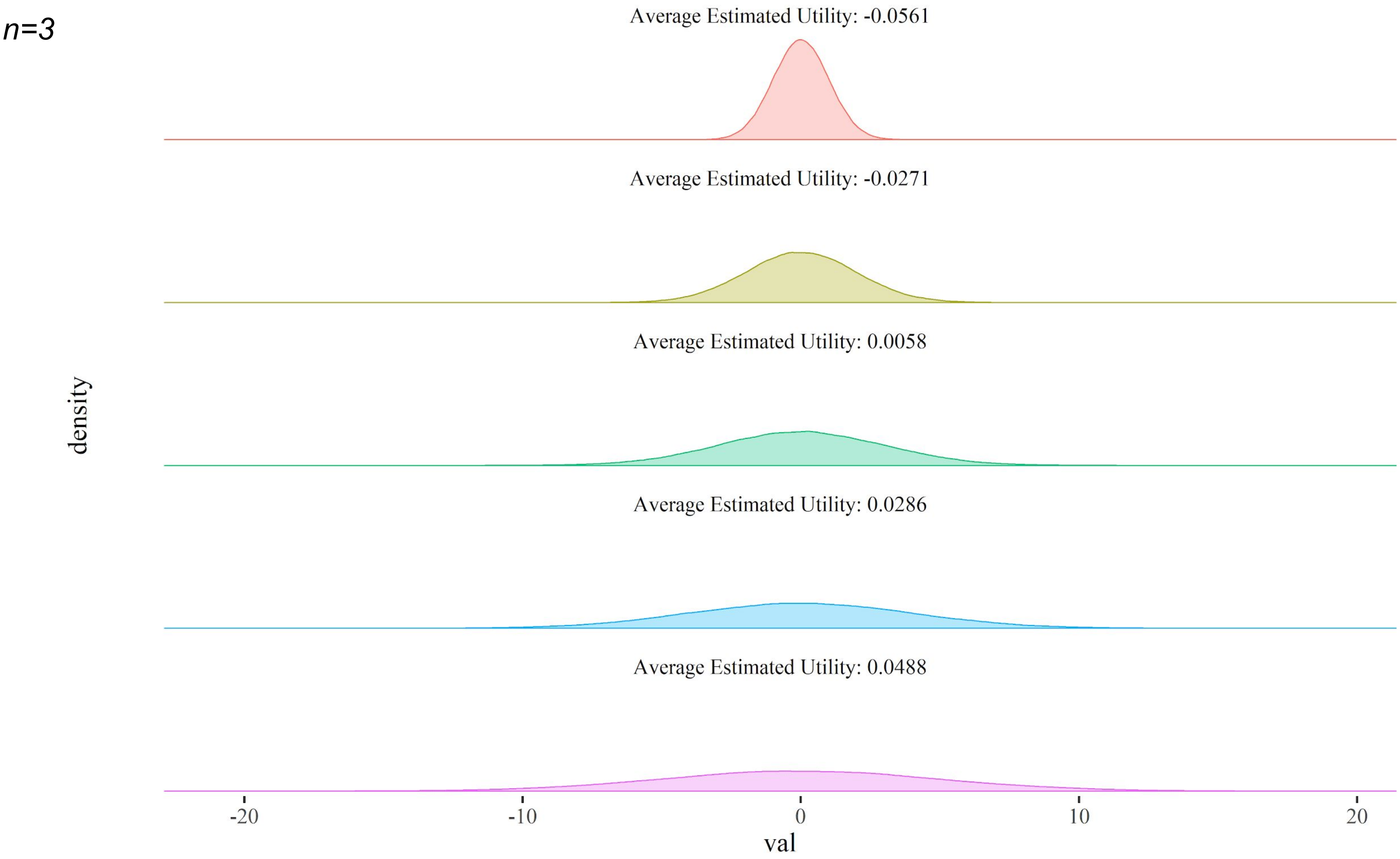
Suppose we have *n* players and *m* items, where $m > n$.

And items each have the same expected value, but different variances.

Which item would you choose?

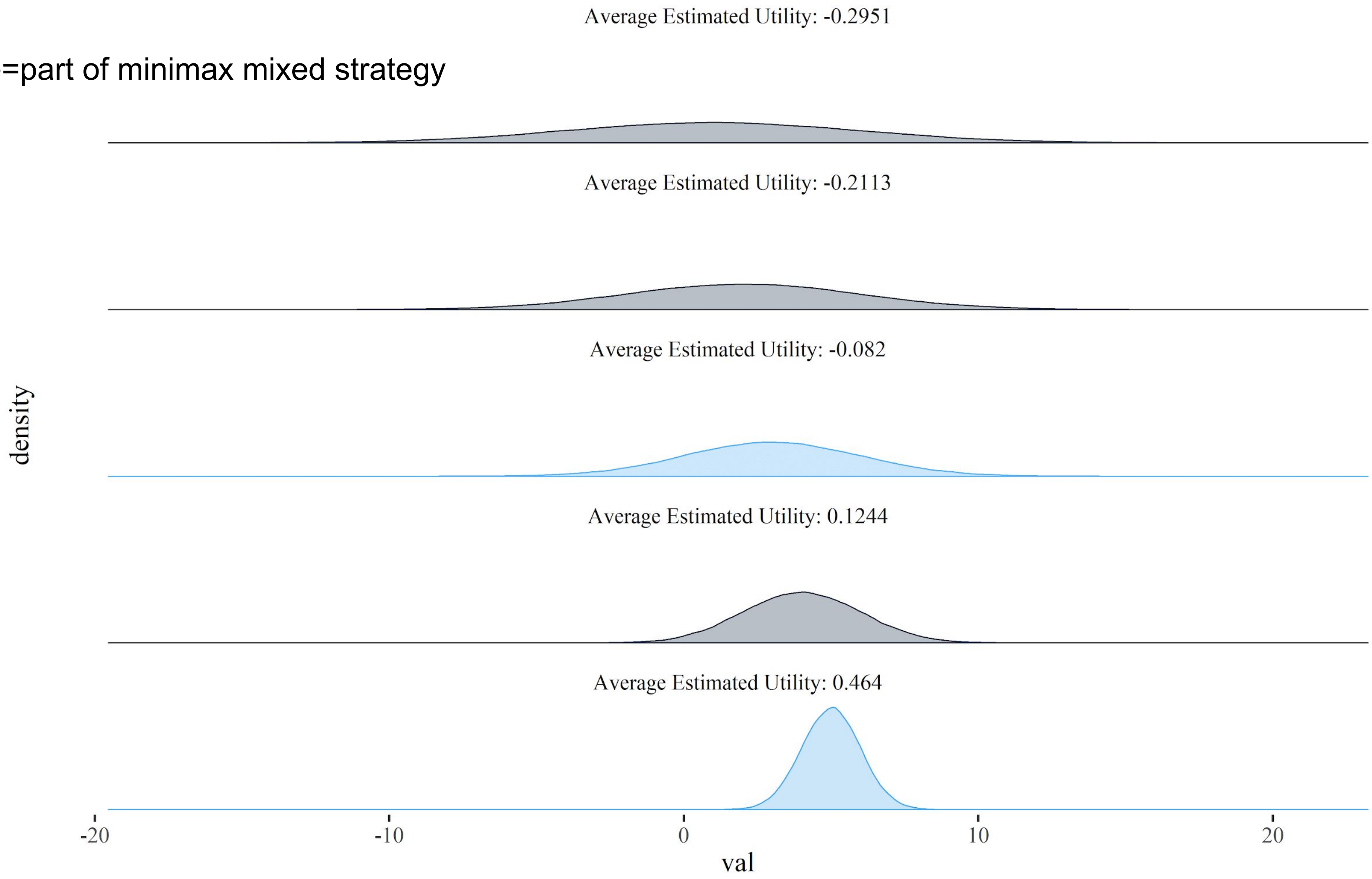
Now we’d expect variance to matter, because players want to **win**.

And, it does.


n=3
Average Estimated Utility: -0.0561
Average Estimated Utility: -0.0271
Average Estimated Utility: 0.0058
Average Estimated Utility: 0.0286
Average Estimated Utility: 0.0488
density
-20
-10
0
10
20
val

Let’s get a little crazy now and have items with different variances *and* different expected values

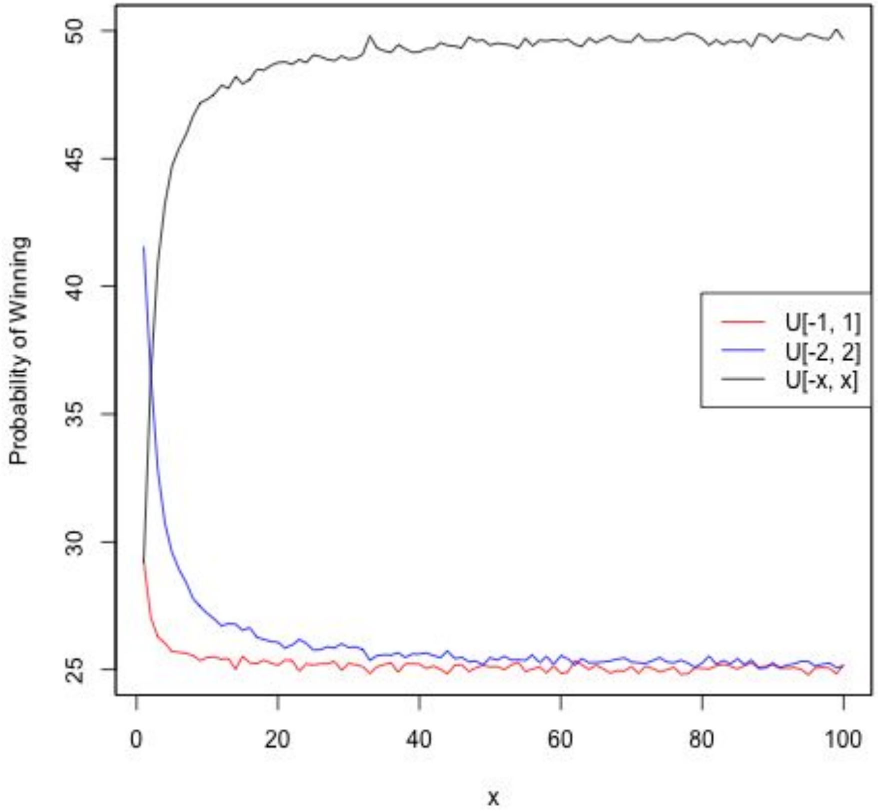
n=3
blue=part of minimax mixed strategy



>3 Players: Variance Matters

Three items: U[-1, 1], U[-2, 2], U[-*x*, *x*]

As *x* goes to infinity, the probability that the third item has a value between -2 and 2 approaches zero. The probability that the third has a value greater than 2 approaches 0.5.



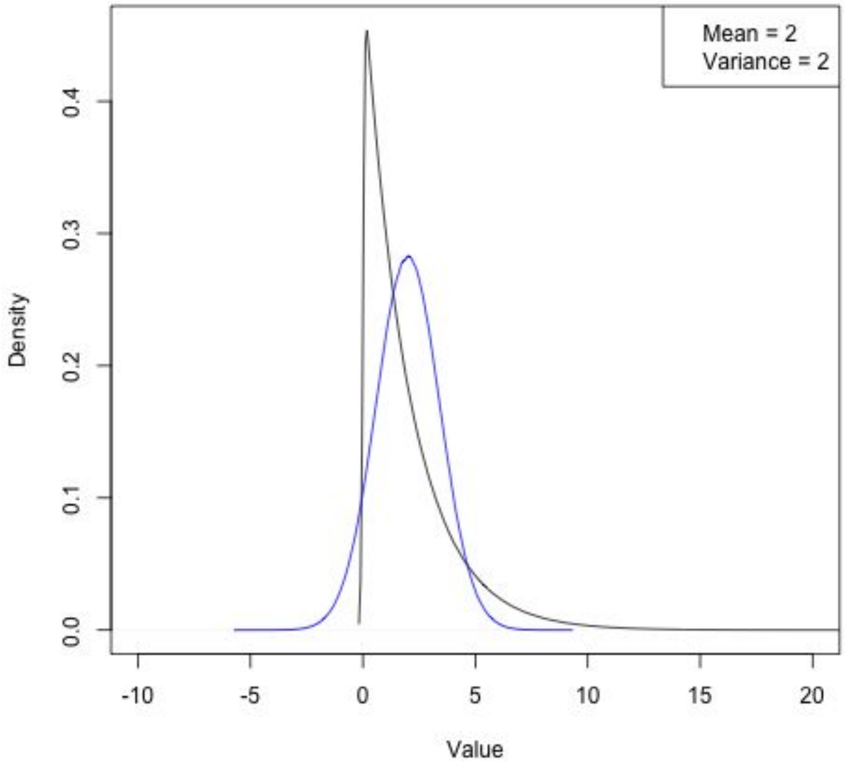
Skewness Matters

Suppose you're trying to decide between the two items on the right.

Which would you choose?

The normal distribution beats the Gamma 58% of the time, **despite** having the same mean **and** variance!

Gamma has positive skewness, normal has zero.



Transitivity in Preferences?

Suppose we have three items (a, b, and c) and each has a payoff from a different distribution.

Is it possible to have a Condorcet cycle ($a > b > c > a$)?

We looked at 13 distributions and 22 parameterized distributions, all with the same means.

There were no cycles. Discrete distributions performed best, while log-normal performed worst.

Summary

- Small-scale (2 player, 2 item) situations match intuition
  - Picking from symmetric and fair (identical expected value) is difficult to optimize.
  - Skew can change results.
    - Negative vs positive- how often is negative an option?
- Larger-scale situations also match intuition
  - Picking winners benefits from high variance.
  - A knapsack with one high-performing item has a significant advantage.
- Optimal strategies are universal
  - In all simulations, optimal choices exist conditionally on other players, but only regarding availability of options.
- Further questions
  - Correlated distributions- how does correlated performance affect optimal strategy?
  - Where else could we apply this?