

“I’m just an *Average Villager*”

Some Simulation Results of the Werewolf Game

Eric-Jan Wagenmakers

University of Amsterdam

The Werewolf Game (Vanilla Version)

- The WG is a role playing game.
- There are n players in a hypothetical village. A referee has designated an unknown number of them as werewolves.
- The werewolves know who the other werewolves are and conspire to devour all the humans.
- Every night, the werewolves conspire to kill off a single human.

The Werewolf Game (Vanilla Version)

- The villagers must lynch one of their own each day. This is decided by voting (one vote for each man, woman, and werewolf that is still alive in the game).
- The humans have the majority and try to find out who the werewolves are and lynch every last one of them.

The Werewolf Game (Vanilla Version)

- The game can have three results.
 1. If the humans have all been eaten, the werewolves win.
 2. If the werewolves have all been lynched, the humans win.
 3. If one human and one werewolf are the only ones alive when it comes to voting, the game is a tie.
- (did I miss anything?)

Expectations

- The more werewolves, the more likely that the werewolves will win.
- The more humans, the more likely that the humans will win.
- Most often, the “teller”, who referees the game, assigns more werewolves when the group of people is larger.
- In a group of 8 people, there will most likely be 2 werewolves.

Aim

- How successful will the werewolves be in a simplified version of the game?
- Will our expectations be confirmed?

Simplifying Assumptions I

- There are no skill differences between participants.
- When one human and one werewolf are left when it's daytime, it's declared a tie (because they will obviously never agree on who needs to be lynched).

Simplifying Assumptions II: The Voting

- No player can vote for him/herself to be lynched.
- Werewolves never vote for other werewolves (!).
- Other than that, the vote is random.
- Until the very end, voting ties are broken by adding a small random number to the vote counts.

Procedure

- When there is only one werewolf, one can easily calculate how likely it is that this lone werewolf will win the game.
- For instance, suppose we have 4 humans and 1 werewolf.
- With probability $Pr=4/5$, the werewolf will live through the first round of voting, and then proceed to kill a human.

Procedure

- With probability $\text{Pr}=4/5$, the werewolf will live through the first round of voting, and then proceed to kill a human.
- In this case there are two humans and one werewolf left. With probability $\text{Pr}=2/3$, the werewolf will survive the voting process and proceed to kill the remaining human during the night.
- So the overall probability of a single werewolf winning against 4 humans is $(4/5)*(2/3)=8/15$.

Procedure

- For two werewolves, things are a bit more complicated, also because werewolves do not vote for each other.
- So what I did was just simulate the process. For instance, when I want to estimate the probability of 2 werewolves winning against 15 humans, this game is played by the computer 1000 times.

Test Your Intuition

- How much chance do you think a single werewolf has to win against 7 humans?
- How much chance do you think a single werewolf has to win against 14 humans?

NB. If the werewolves win 500 times, tie 100 times, and lose 400 times, the probability is 0.55, as I've counted a tie as gaining 0.5 points (winning gains 1 point, losing 0 points)

Check Your Intuition?

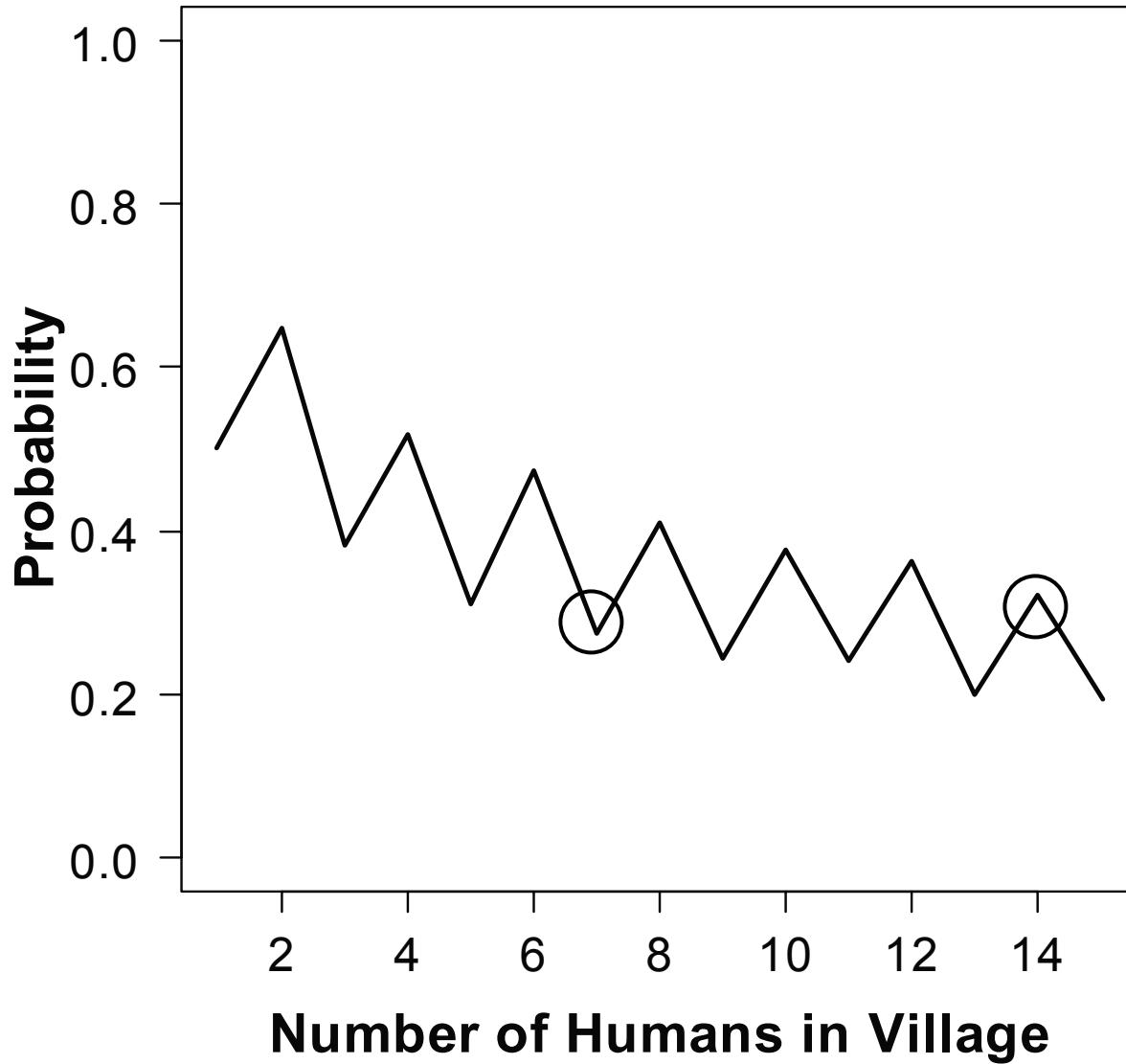
- How much chance do you think a single werewolf has to win against 7 humans?

$\text{Pr}(\text{werewolf wins}) = .28$

- How much chance do you think a single werewolf has to win against 14 humans?

$\text{Pr}(\text{werewolf wins}) = .32$

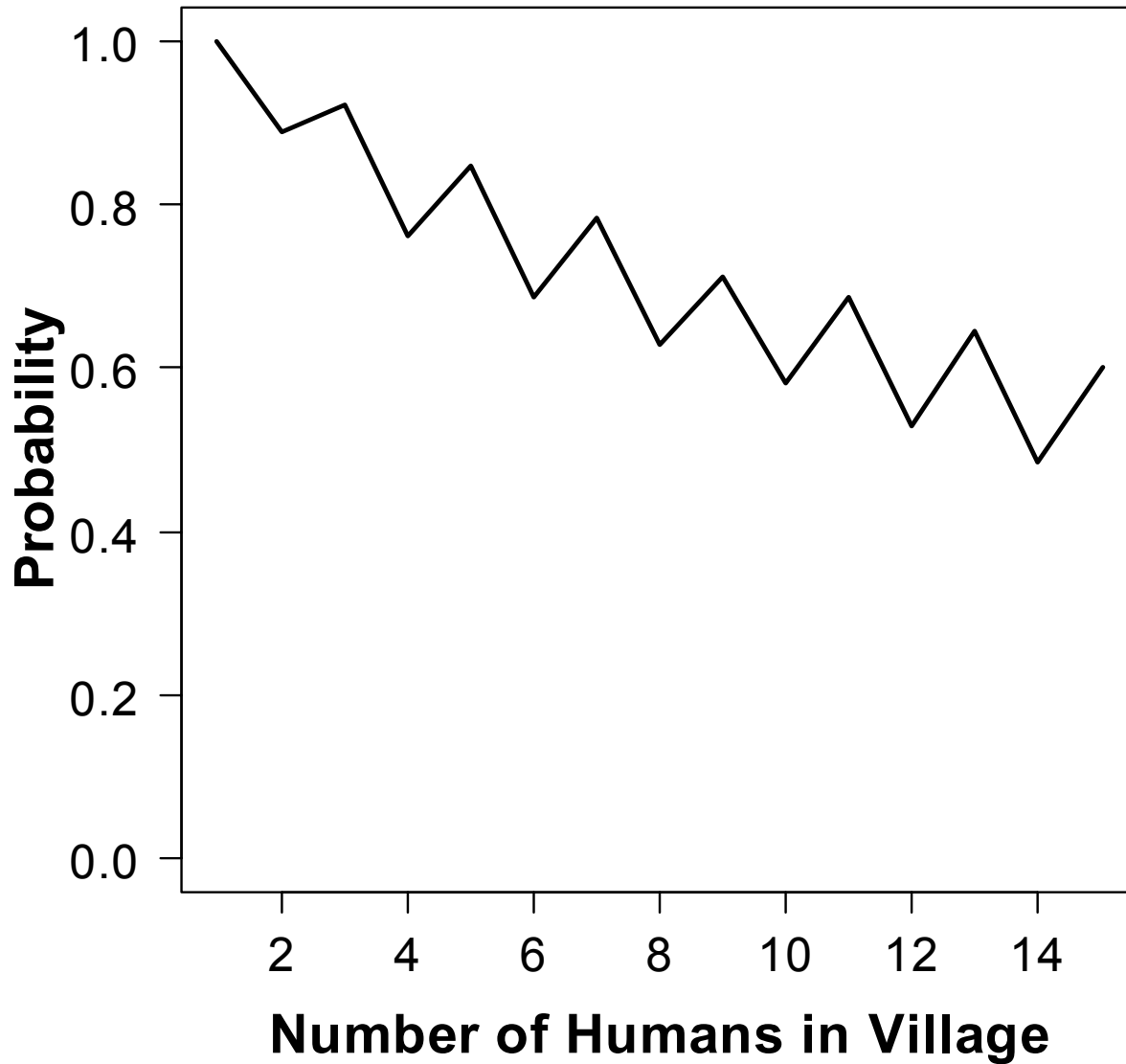
Probability of 1 Werewolf Winning



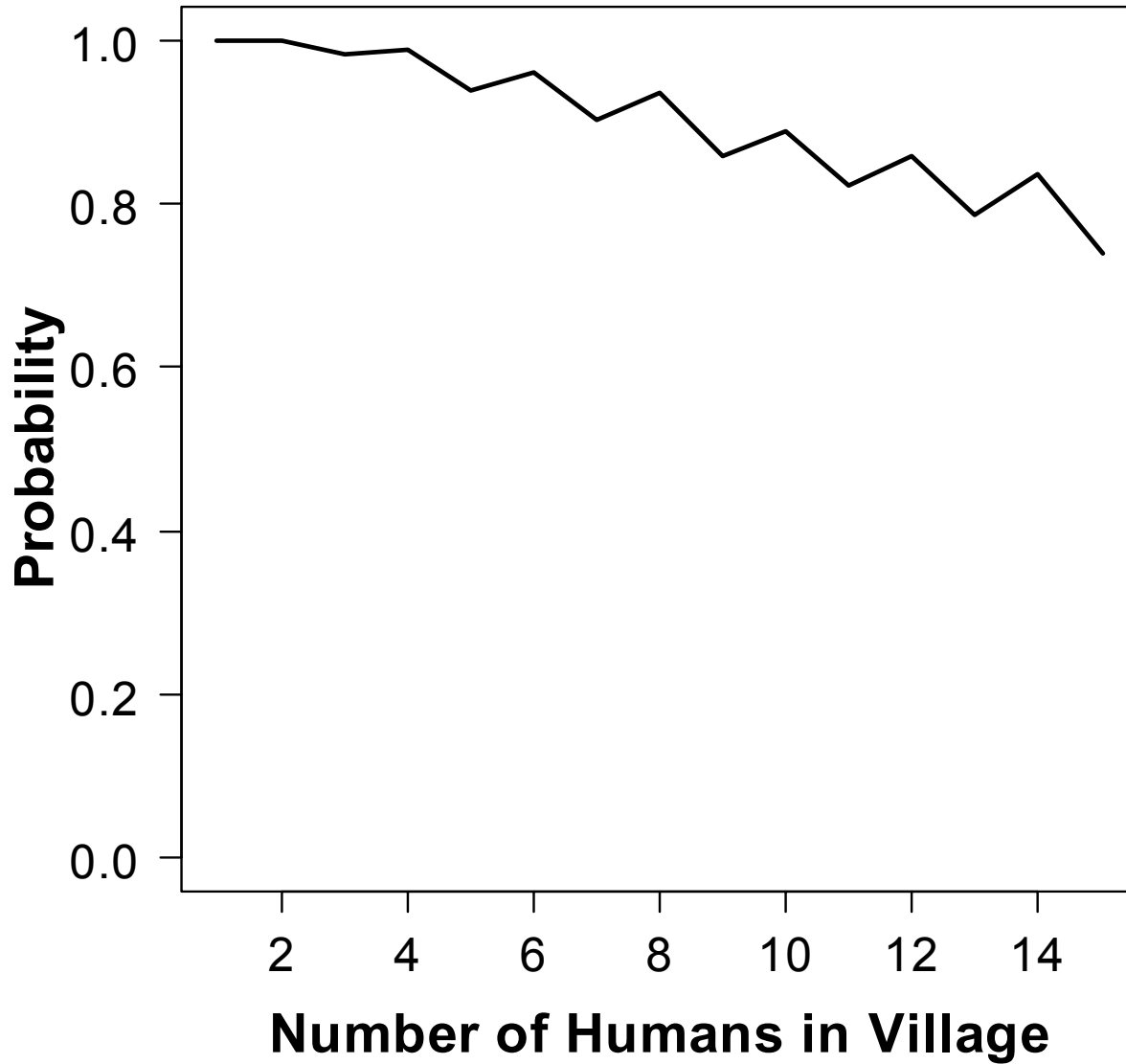
What is Going on Here?

- Consider 3 humans (any odd number, really) and 1 werewolf.
- After surviving the voting process, and subsequently killing a human villager, we are left with 1 human and 1 werewolf – it's a tie! With an odd number of players, the best a single werewolf can possibly do is to tie the game!
- With an *even* numbers of players, the werewolf can use the night time to eliminate the remaining human villager.

Probability of 2 Werewolves Winning



Probability of 3 Werewolves Winning



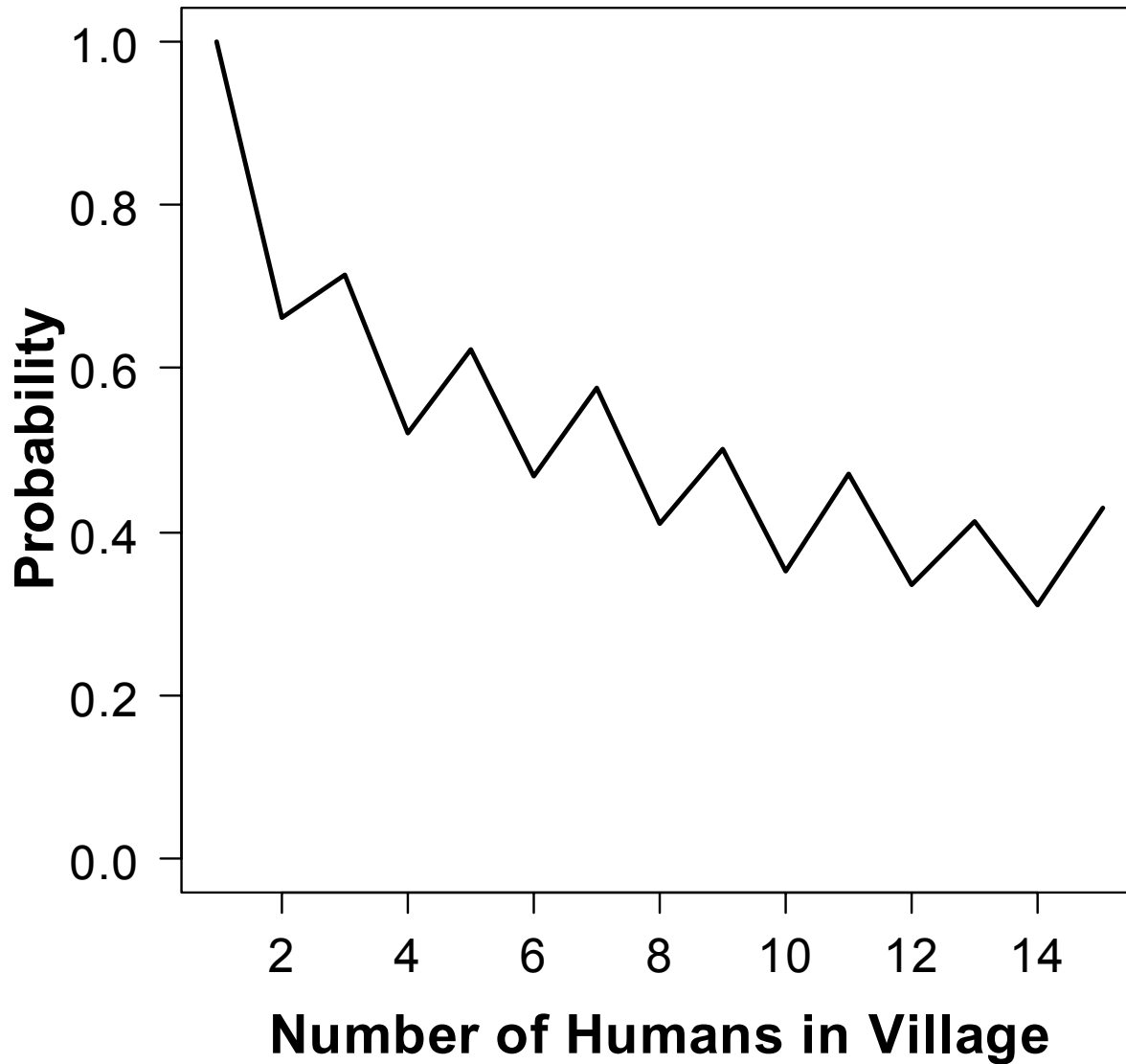
Take Home Messages

- It is important to know at what moment the next-to-last werewolf gets killed.
- Werewolves do better in the simulations than in real life. Apparently, people can tell when somebody is lying – or is it something else?
- Generally, intuition is inferior to just simulating the process under investigation.

Ties...

- So what if there are no ties: when one human and one human are left, they will not agree on who gets lynched. Night will fall, and the human gets eaten.
- Will this eliminate the saw-tooth pattern?

Probability of 1 Werewolf Winning



The End