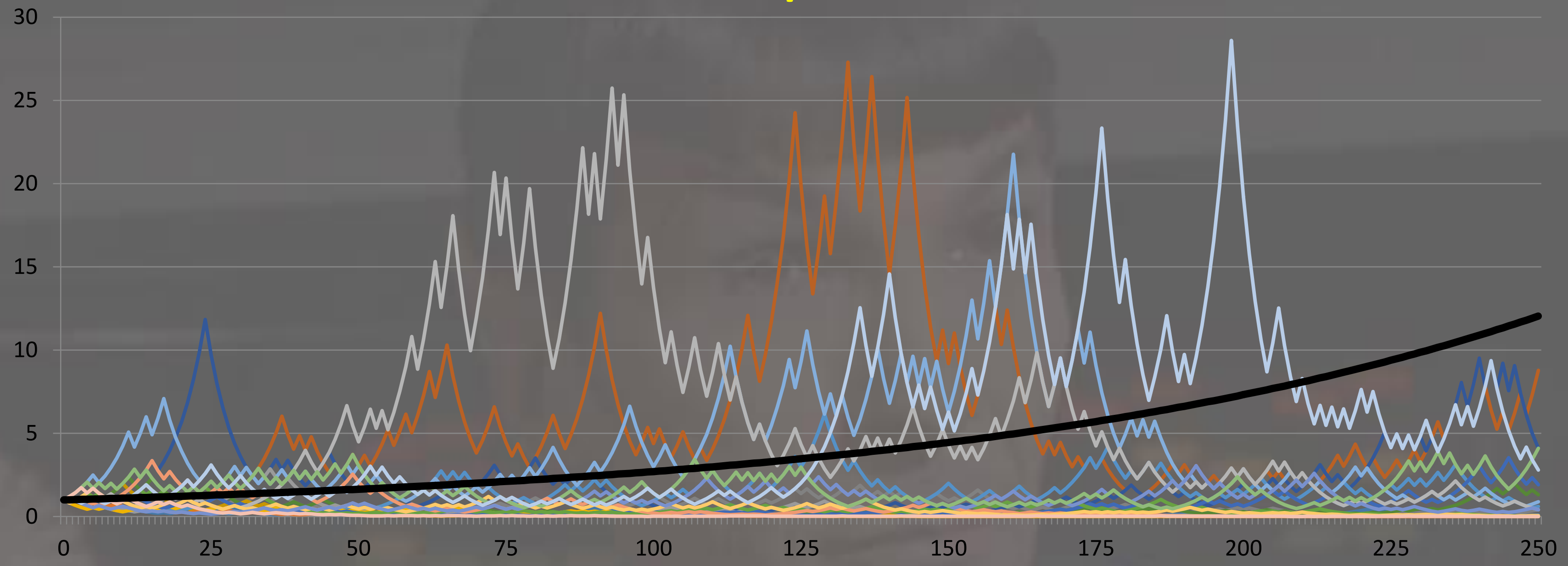
Self-correcting Kelly strategies for skeptical Bayesians

Aaron Brown
AQR Capital Management

16th International Conference on Gambling and Risk Taking

- Consider a gamble that returns +20% or -18% with equal probability
- Standard utility theory says you should accept or reject based on your risk aversion
- Expected return +1%
- Standard deviation of return +19%
- Law of large numbers is your friend, if gamble is repeated many times ratio of expected return to standard deviation increases without bound
- Let’s see what really happens. . .

Simulation of 20 paths of 250 bets



- 50% chance of +20%, 50% chance of -18%
- Vertical axis is wealth starting at 1
- Horizontal axis is number of bets made
- Black line is growth in expected value
- Colored lines are 20 simulated paths of wealth

What happened?

- Most paths go quickly to near zero
- A few paths shoot up far beyond expected value, but eventually crash
- If we ran this longer
    - all paths would go to zero wealth
    - expected value would continue exponential increase

What is the problem?

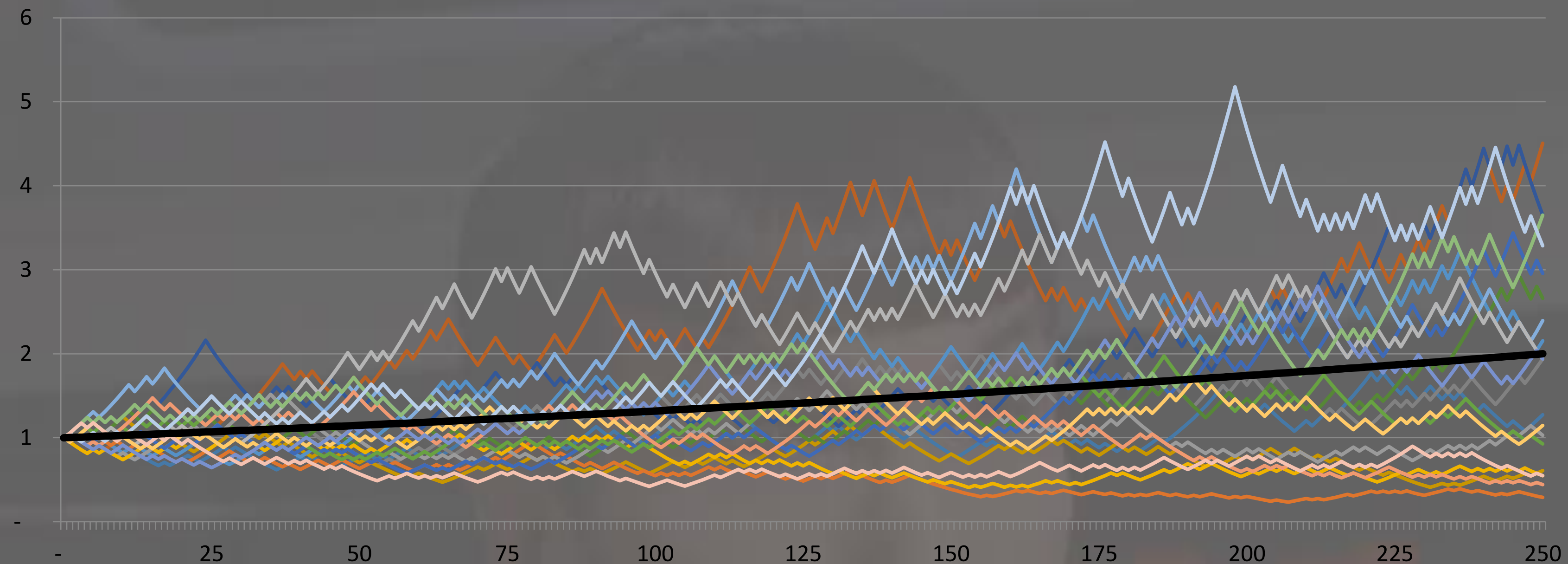
- Expected return of gamble is +1%
- Median return is -0.8%
  - A win plus a loss leaves you with 1.2 x 0.82 = 0.984
  - A loss plus a win is the same
  - Median result is 0.984^0.5 = 0.992 or a 0.8% loss
- Law of large numbers is your enemy
  - Expected wealth increases 1% every gamble
  - Median wealth declines 0.8% every gamble
  - Long-term distribution is microscopic chance of astronomical gain, virtual certainty of ruin
  - You need over 52% win rate to break even, as repetitions increase that win rate becomes nearly impossible

It gets worse. . .

- Imagine that the 20 colored paths represent:
  - 20 new traders starting at a prop desk
  - 20 new hedge funds
  - 20 new business initiatives/politicians/military adventures
- Some of them achieve extraordinary success
- Big drawdowns are followed by even bigger recoveries, teaching people to keep the faith and double up after losses
- Successful realizations will attract investment, imitation, larger limits, more freedom
- All will inevitably crash, and crash with far higher exposures than they carried on their upswings
- How does the financial system (or anything) survive?

Go back to our simulation

- Bet wins +20% half the time and loses 18% the other half
- Reduce the size to 5/18 so the bet wins 5.56% half the time and loses 5% the other half
- Now expected return is only 0.28% instead of 1%
- But median return is +0.14% instead of -0.8%
- We make a profit as long as we win 49% of our bets, instead of 52%
- Law of large numbers is our friend again



- Reduced bet simulation using same win/loss results as previous simulation
- Black line is exponential growth at the expected value rate, 0.28% per bet
- Long run results will cluster around growth at half that rate, 0.14%
- Milder ups and downs, no tendency to soar or blow up

Kelly myth #1

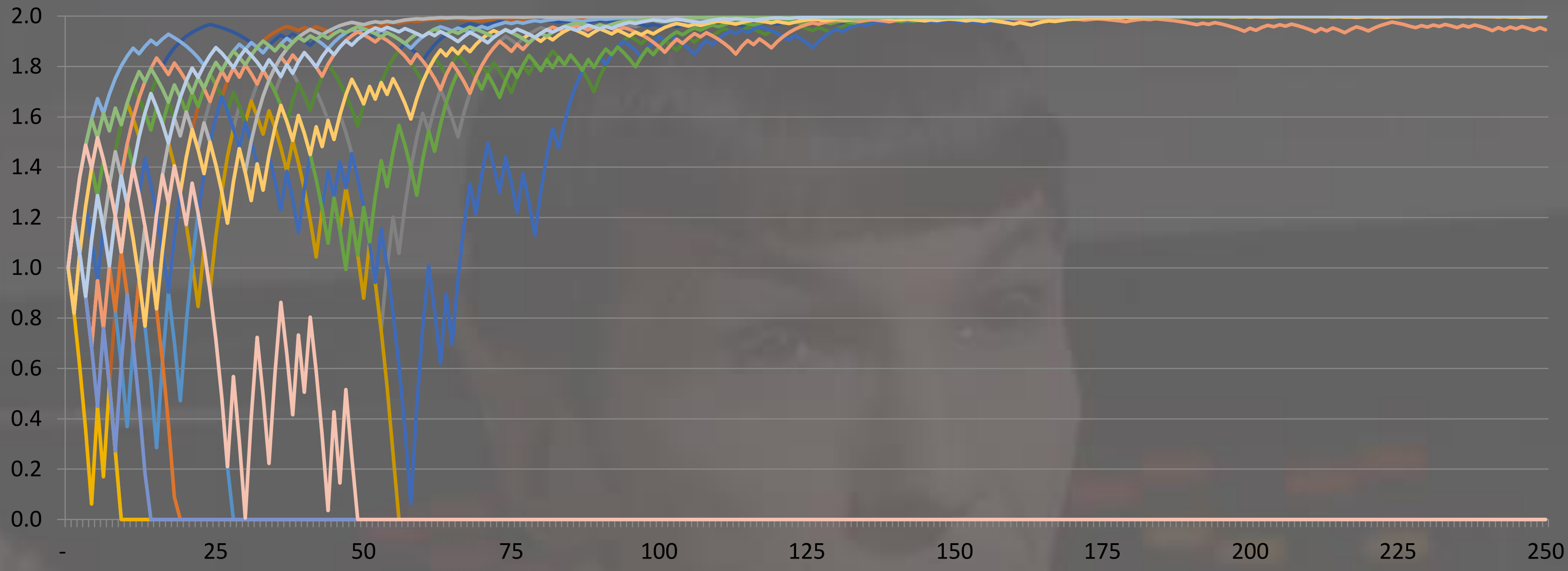
Overbetting is the error that leads to predictable crashes

What really changes

- Suppose we gave the original risk takers who were betting +20%/-18% initial capital of 3.6 instead of 1 and let them make the same dollar size bets as before
- Now their gains and losses are +5.56%/-5%
- What if we lied and just told them they had extra capital?
- Obviously it's not the extra capital that makes a difference, nor the dollar size of the bets
- What matters is how much the risk takers increase their dollar bets after wins and reduce them after losses

Another example

- Suppose someone raised bets 20% after losses and cut them 18% after wins?
- Same average bet size as original bettors since wins and losses are equally likely



- 40% of paths go to 0, 60% go to 2

Kelly insight #1

- Don’t focus on absolute bet size
- Focus on how bets are increased or decreased in response to events
- Dramatic bet increases after success, and decreases after failure, exploit opportunities while limiting downside, but cause volatility drag
- The reverse strategy limits upside and can go to zero, but profits from volatility drag
- Intermediate strategies trade off volatility drag for convexity

Kelly myth #2

There is a single optimal risk strategy

Kelly insight #2

- Risk strategy depends on situation
- Project with fixed upside and resources already spent
  - Decrease risk after success, increase after failure
  - Fail fast or succeed
- Venture with small costs and improbable gains
  - Increase risk faster than Kelly after success, and decrease faster after failure
  - The right tail matters, the median is little better than failure
- Typical intermediate case
  - Right tail is illusory
  - Extended failure to grow or opportunity changes, not losing everything, ends betting
  - Exponential opportunity is key, rate of growth is secondary

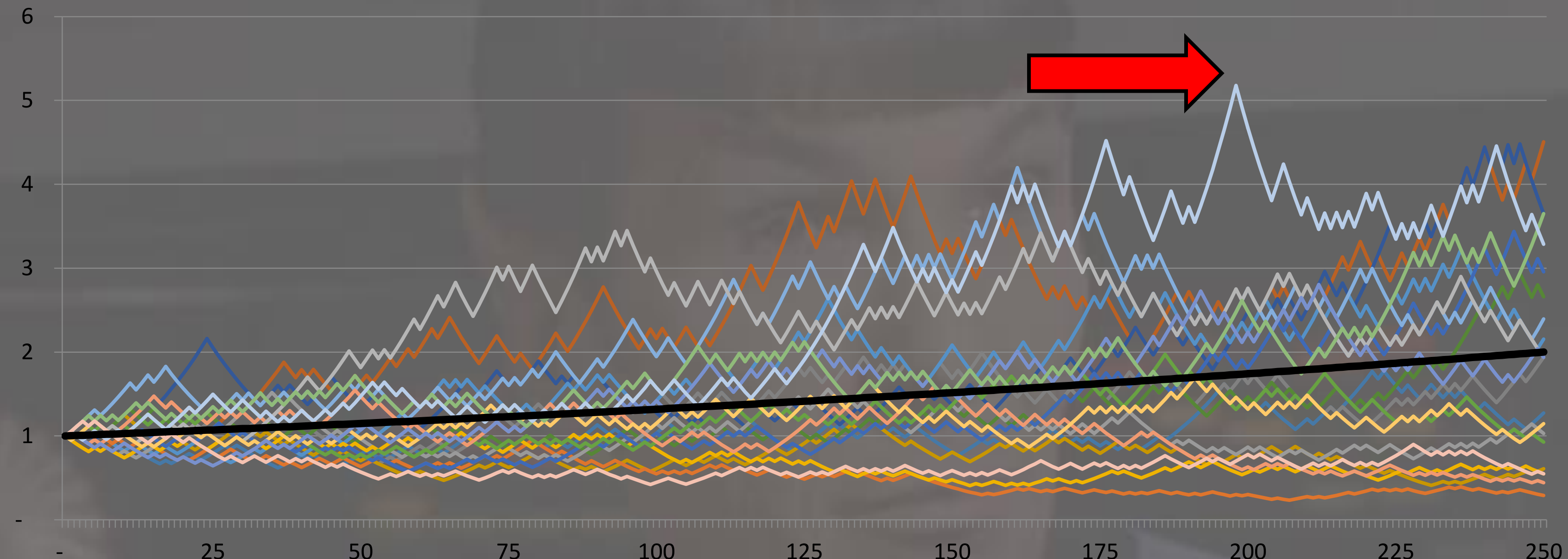
Kelly myth #3

Bet size is what matters

Kelly insight #3

- Smaller than Kelly (take chips off the table)
  - You have more wealth left if the venture fails
  - You get the Kelly exponential growth rate, but applied to a lower base
    - For permanent opportunities, not a big loss
    - For transitory opportunities, it can matter a lot
- Larger than Kelly (pretend capital)
  - Wealth can go to zero (or below)
  - You get more out of transitory opportunities
- Common sense rule: Set absolute bet size so you run out of capital at the point you'd give up or get stopped out anyway

Masters of the universe



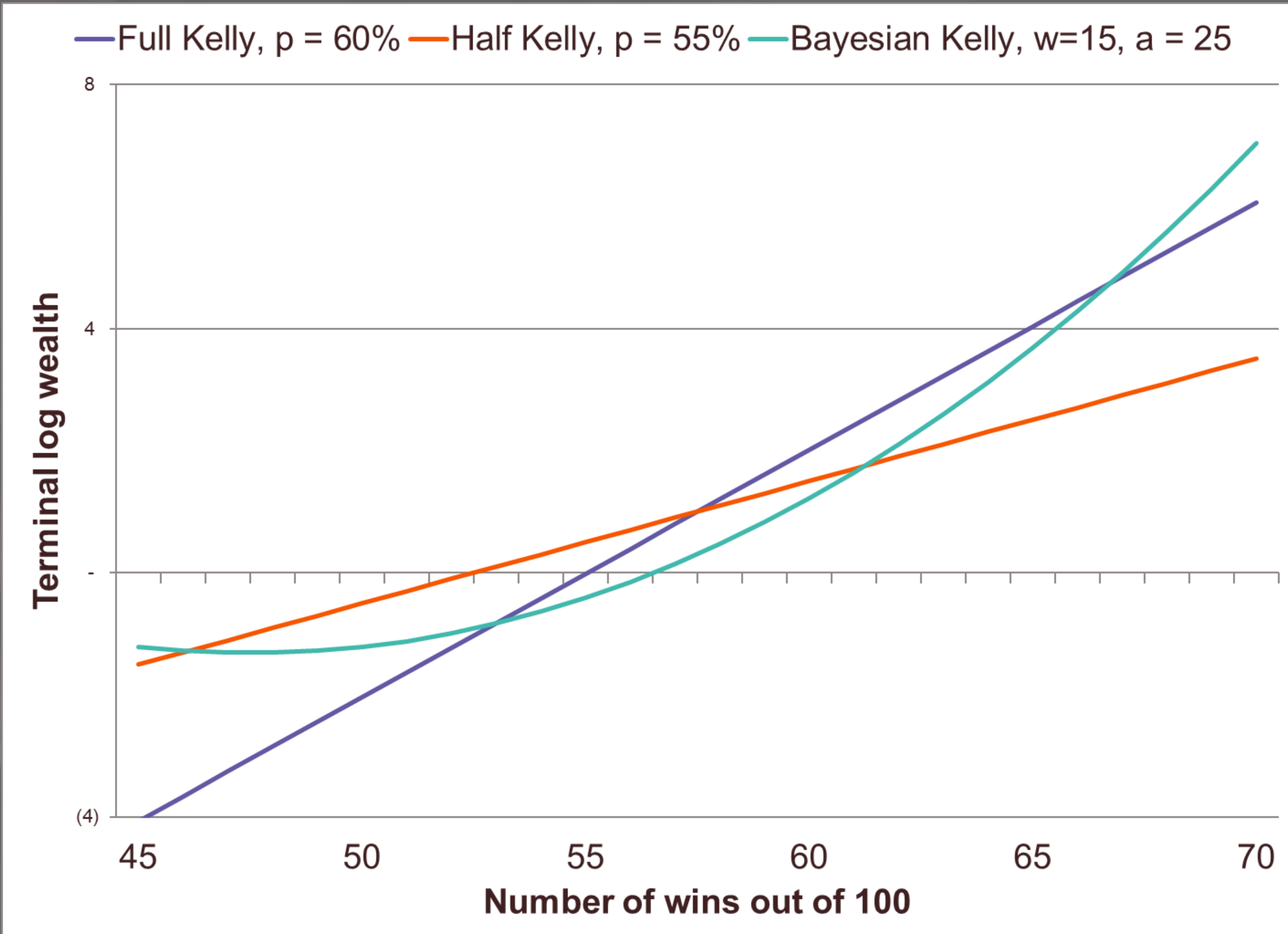
- Consider the highest peak in the first simulation, the light blue line with a wealth of 29 after the 198th bet
- This risk taker had won 112 of 198 bets, 57%

Skeptical Bayesian

- You can show her that she paid more in volatility drag than she gained from betting aggressively
- She would have done better with smaller bets
- At an assumed 57% win rate, even her optimal bet size would lead to blow up given the actual 50% win rate
- Key is to persuade her to change bet size based on Bayesian principles

Some quick math

- Conjugate prior for Bernouli probability is Beta distribution
- Has convenient properties
  - Parameters w and a are equivalent to observing w successes in a attempts before seeing evidence
  - Observing W wins in A attempts leads to a posterior expectation of $p = (W + w) / (A + a)$
  - Optimal Kelly bet is just the bet at the expected p
  - Wealth after a series of Kelly bets based on this prior does not depend on the order of the wins and losses
- Example, bet +20% or -20%, estimate $p = 0.6$



Full Kelly, p = 60%
Half Kelly, p = 55%
Bayesian Kelly, w=15, a = 25
Terminal log wealth
8
4
-
(4)
45
50
55
60
65
70
Number of wins out of 100

Back to master of the universe

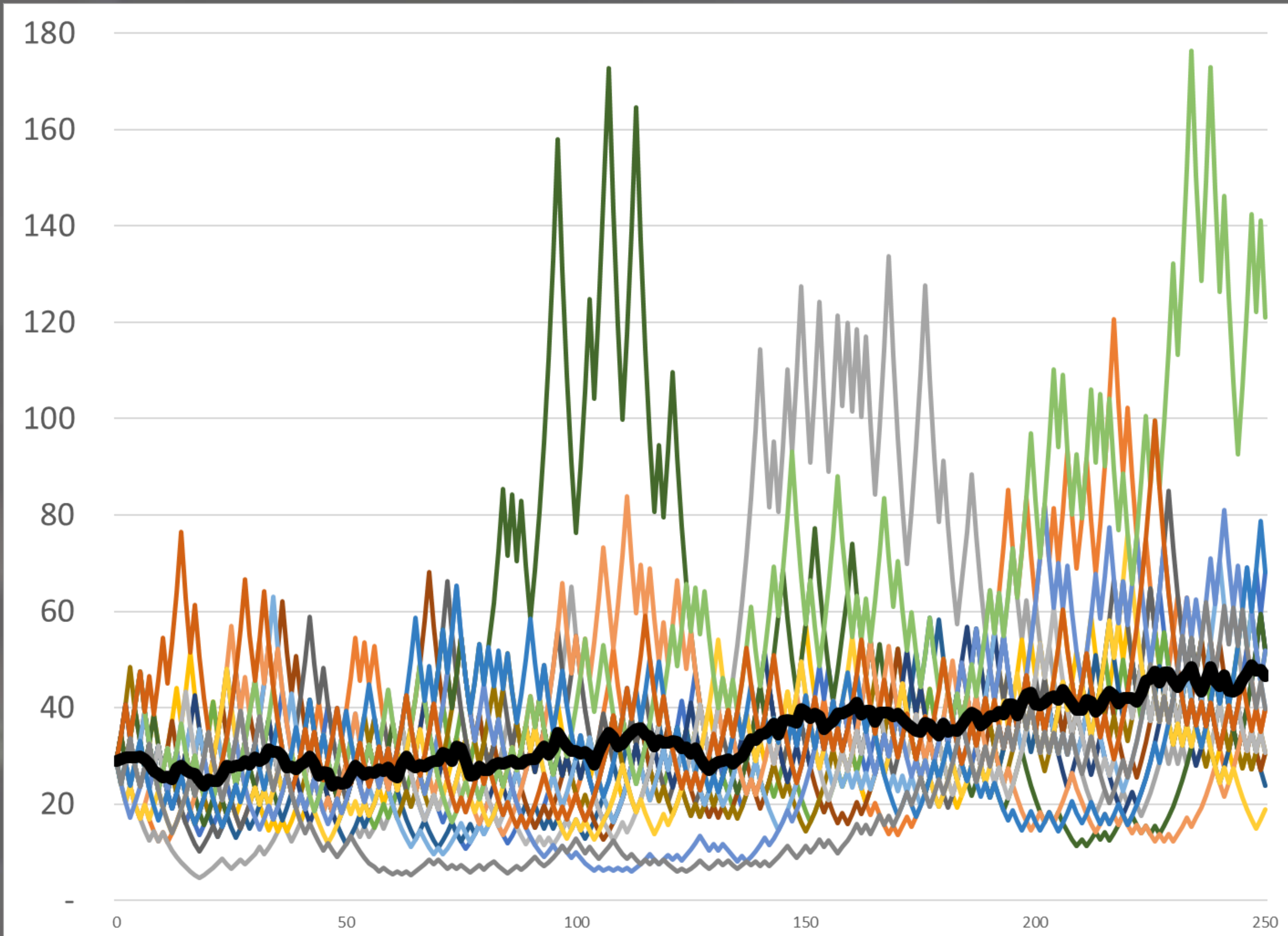
- Won 112 of 198 bets, 57%
- Kelly bet at p = 0.57 is 97% of wealth
- Too high to avoid blow up if actual p = 0.50
- If she accepted p = 5/9 (0.55), median profit is zero
- Need smaller p to be profitable
- If she accepted p = 0.50, profit is maximized

Skeptical

- Observed her cohort of 20 risk-takers averaged 50%
- Admits that best of 20 risk-takers at peak wealth is more likely to have been lucky than unlucky
- Accepts that long-term win rate is probably less than 57% historical rate
- Suppose she is skeptical of your arguments, and will only accept that her true long-term win rate is probably 56% (still too high to survive)

Bayesian

- 56% is consistent with a prior belief of 0, 2
- p = (112 + 0) / (198 + 2)
- There is only a small decrease in current bets compared to p = 112 / 108
- There is a dramatic difference in long-term returns
- 14, 27 implies a prior expectation of 52% success rate
- (112 + 14) / (198 + 27) = 56%


180
160
140
120
100
80
60
40
20
-
0
50
100
150
200
250

Why solve for optimal solutions when you know your model is oversimplified and your parameters are just guesses?

- I'm a quant. That's what we do.
- Most solutions cannot work. At least a solution that is optimal under some conditions might work.
- Optimal solutions decided under calm conditions and applied consistently avoid the many behavioral biases that sabotage ad hoc solutions.
- Some individuals probably can use intuition to do better than simple optimal quant solutions, but. . .
    - You don't know for sure who they are
    - Their abilities can wax and wane, things change
    - You cannot systematically test them against a wide range of situations over a long period of time
    - You cannot continuously improve them

Bayesian Kelly approach

- Parametrize results (e.g. lognormal random walk)
- Set prior distribution over parameters
- Set optimal Kelly bet based on posterior distribution
- Update posterior as results come in
- Assumes
  - Wealth is the constraint
  - Wealth is the goal
  - Probability of zero wealth must be zero
  - Time horizon is infinite

Realistic example

- Hiring a trader, investing in a hedge fund, trying out a computer trading algorithm
- How big to start out?
- How to vary size with success?
- When to give up?

Realistic Bayesian Kelly

- Define terminal events
  - Success: you are confident there is a positive edge
  - Failure: you give up hope of a positive edge
- Given parametrization and prior, simulate results for a range of nominal wealth levels
- Estimate probability of success versus failure
- Estimate gain before success, loss before failure
- Add in future value of success
- Select optimal nominal wealth level