V The psychology of probabilities

2. Other common pitfalls

V.2.1 Selective sampling

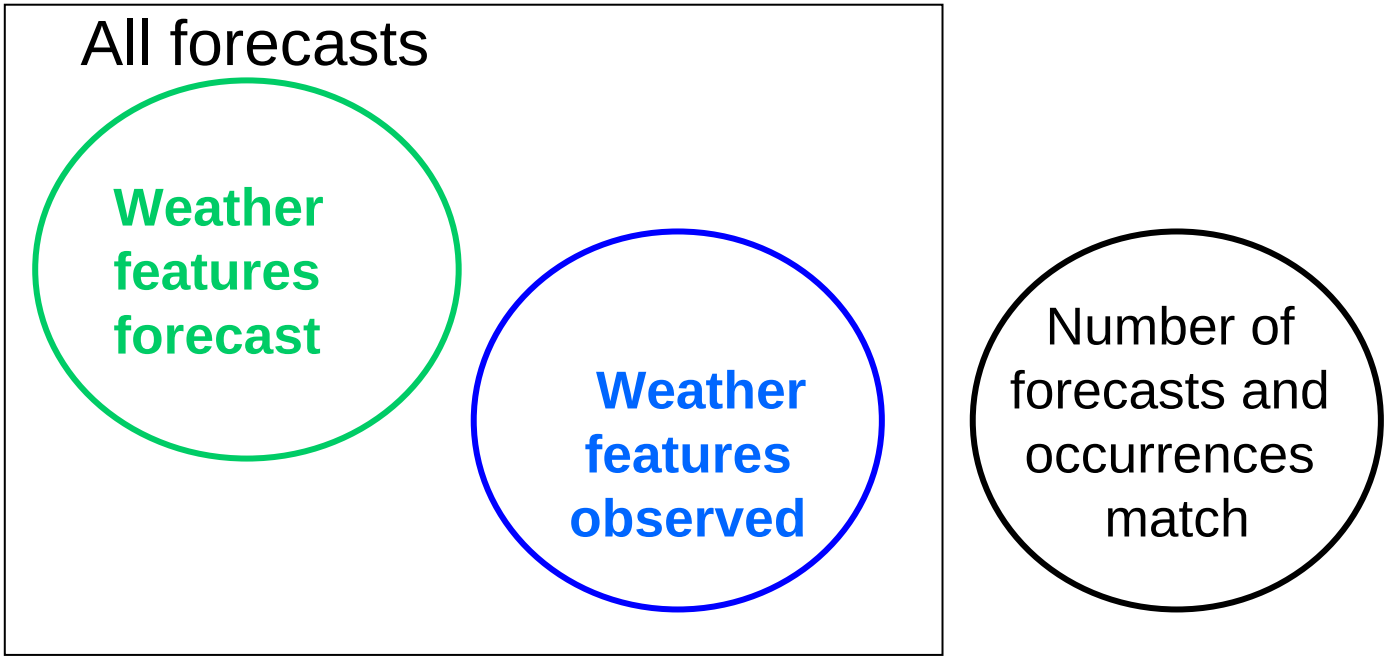
Selective sampling

Common complaints about ECMWF forecasts:

"-There are **too many spurious** tropical cyclones beyond D+6."

"-You are over-forecasting cut-offs at D+6. **Only half of them verify."**

"-When the NWP predicts >25 mm at D+5 it rains **<10 mm/d** on average."



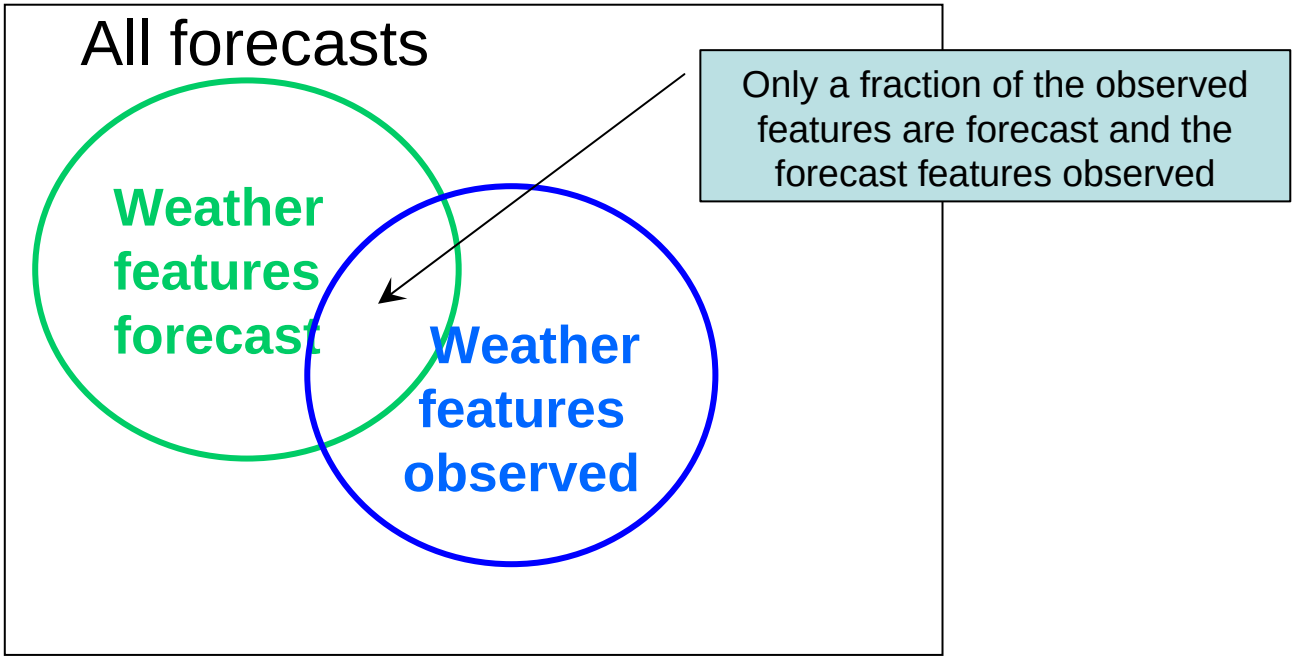
Selective sampling

Common complaints about ECMWF forecasts:

"-There are **too many spurious** tropical cyclones beyond D+6."

"-You are over-forecasting cut-offs at D+6. **Only half of them verify."**

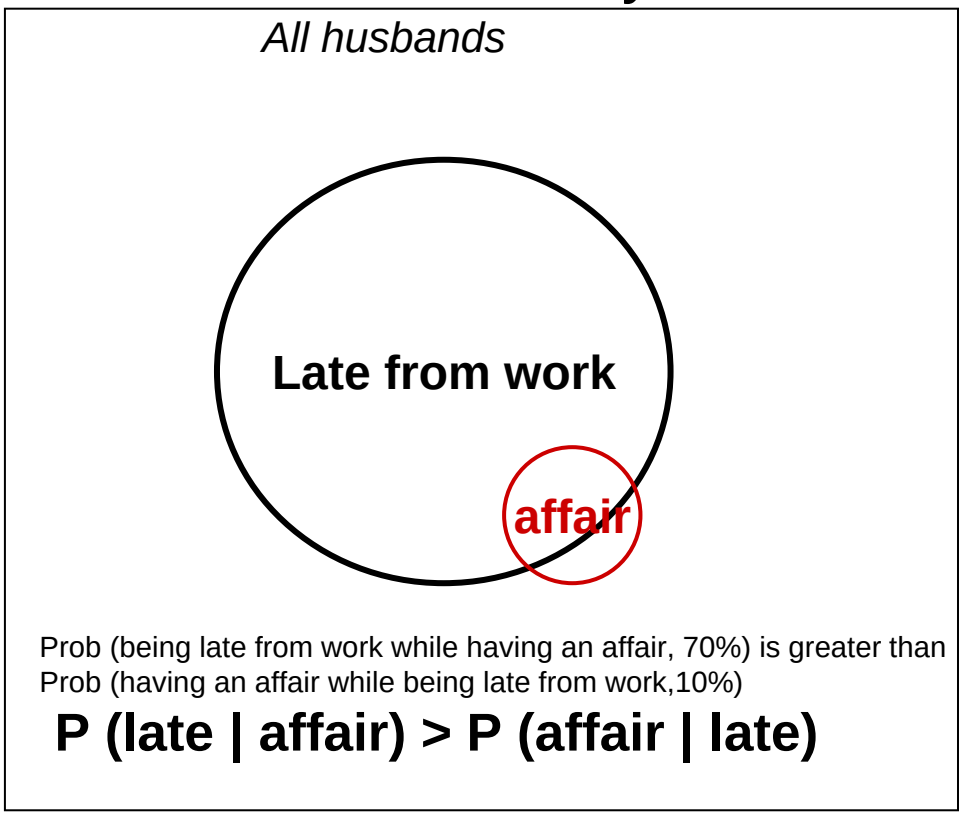
"-When the NWP predicts >25 mm at D+5 it rains **<10 mm/d** on average."



V.2.2 Examples of misinterpretations of conditional probabilities

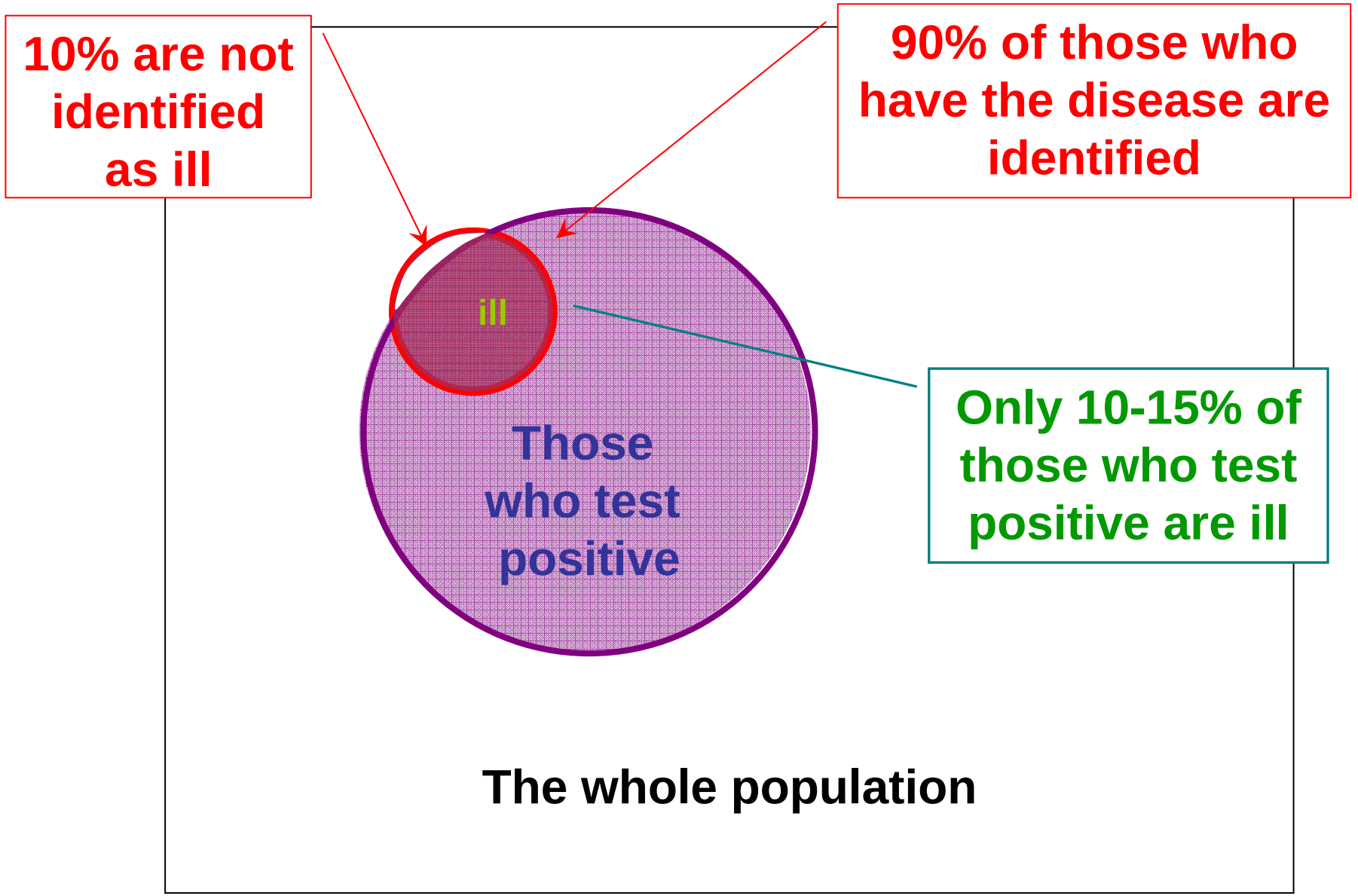
a) From the 2004 movie “Shall we dance?”

This shows that she is not an educated Bayesian!



b) The medical test

- Paul is afraid he has a fatal disease
- **It affects 1% of the population**
- His doctor subjects him to a test
- **The test is “90% accurate”**
- Paul tests positive
- Paul commits suicide


10% are not identified as ill
90% of those who have the disease are identified
ill
Those who test positive
Only 10-15% of those who test positive are ill
The whole population

c) O.J.Simpon battered his wife

His defence lawyer:

> "Only 1 partner per 2500 battered wives go on murdering them…O.J. is not likely to have murdered her"

His defence lawyer:

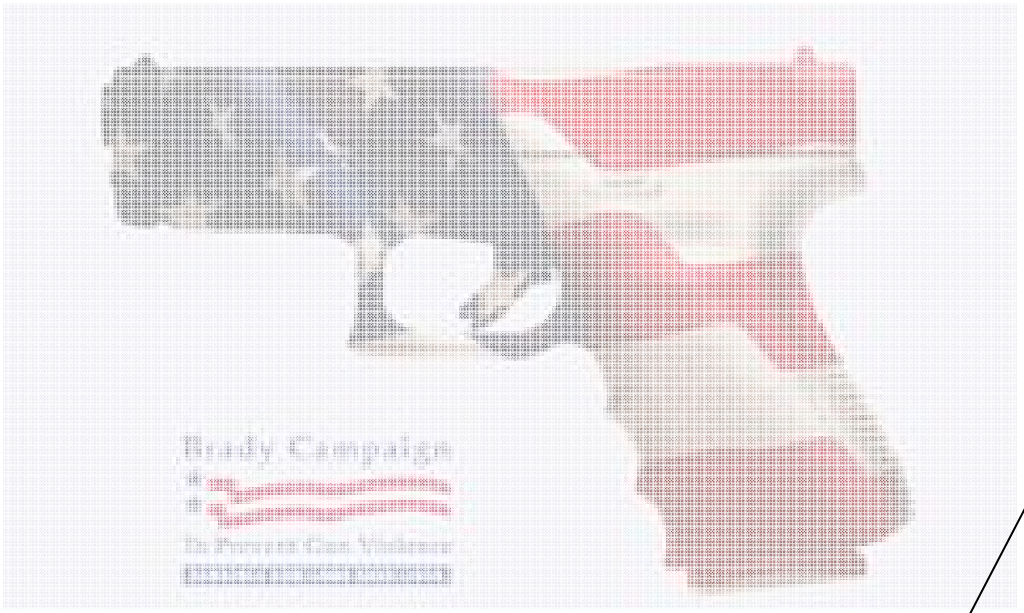
"Only 1 partner per 2500 battered wives go on murdering them…O.J. is not likely to have murdered her"

Oversight by prosecution:

In *cases of murder*, 8 out of 9 were committed by the partner

This makes O.J. by far the most likely suspect

d) “90 % of gun murders are committed by gun owners”



Gun owners

Murders by gun owners

All gun murders

Gun lobby:
1 % of gun owners commit gun murder

Anti-gun lobby:
Gun owners commit 90 % of gun murders

e) I have got 2 children of which one is a boy:

-What is the probability that the other is also a boy? Both sexes equally likely 50%

a)Totally four possibilities with two children:

BB **BG** **GB** ~~**GG**~~ since there already is a boy

. . . which leaves three possibilities of which only one **BB**

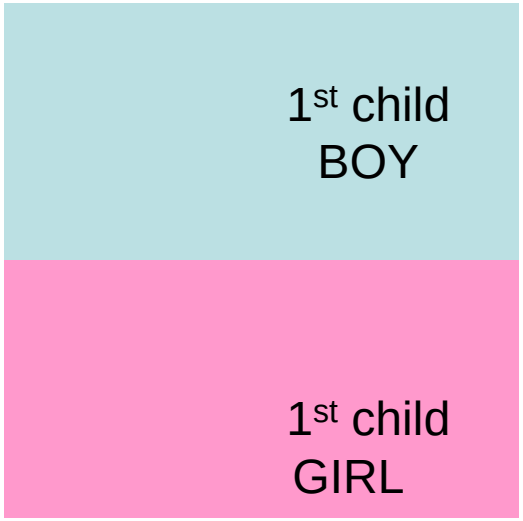
I have got 2 children of which one is a boy:

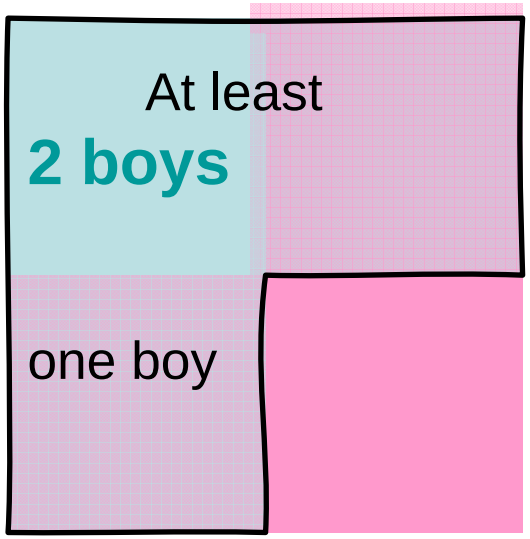
-What is the probability that the other is also a boy?

Both sexes equally likely 50%

a) Totally three possibilities: BB BG GB ~~GG~~ of which only one BB

b) Venn diagram





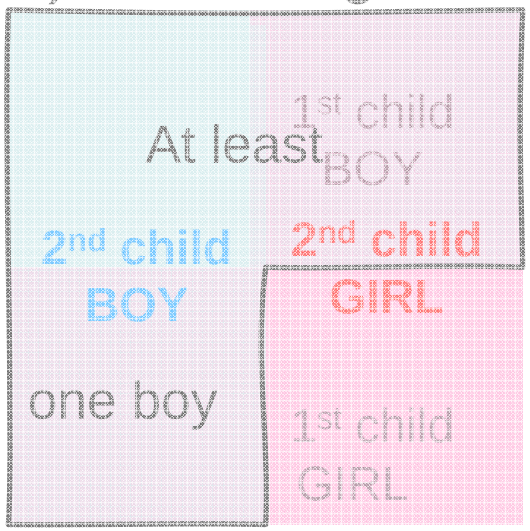
I have got 2 children of which one is a boy:

-What is the probability that the other is also a boy?

Both sexes equally likely 50%

a) Totally three possibilities: BB BG GB ~~GG~~ of which only one BB

b) Venn diagram



c) Bayes's Rule

Prob(**BB** | **B** of CC)=

$$\frac{\text{prob}(\mathbf{BB})\cdot\text{prob}(\mathbf{B}\text{ of CC}|\mathbf{BB})}{\text{prob}(\mathbf{B}\text{ of CC})}$$

33%

25%

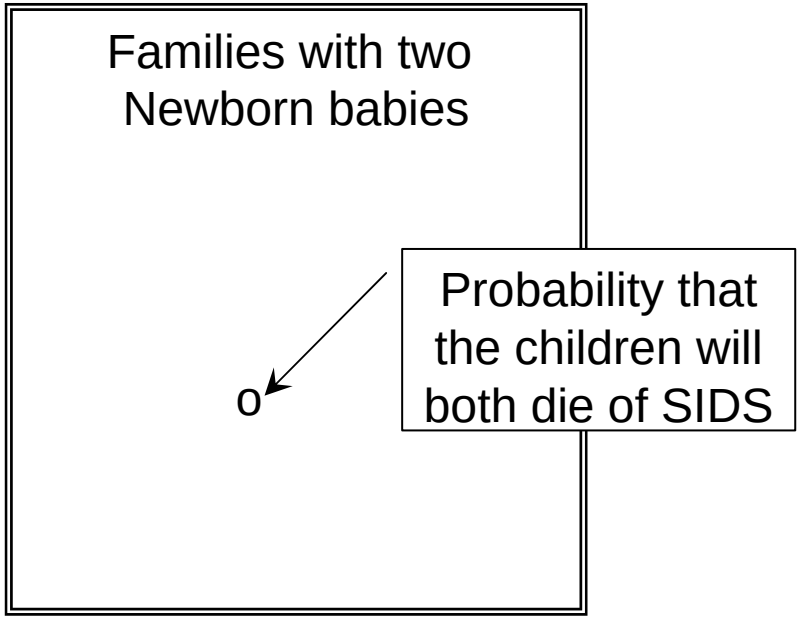
75%

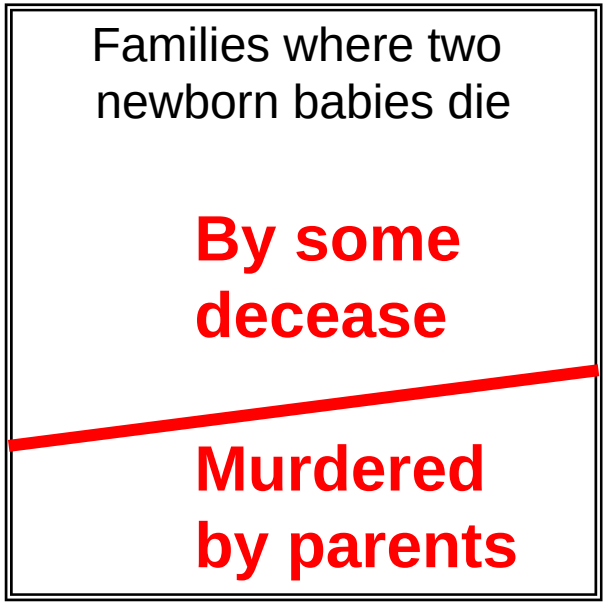
100%

f) The Sally Clarke tragedy (jailed for having killed her two babies in the late 1990's)

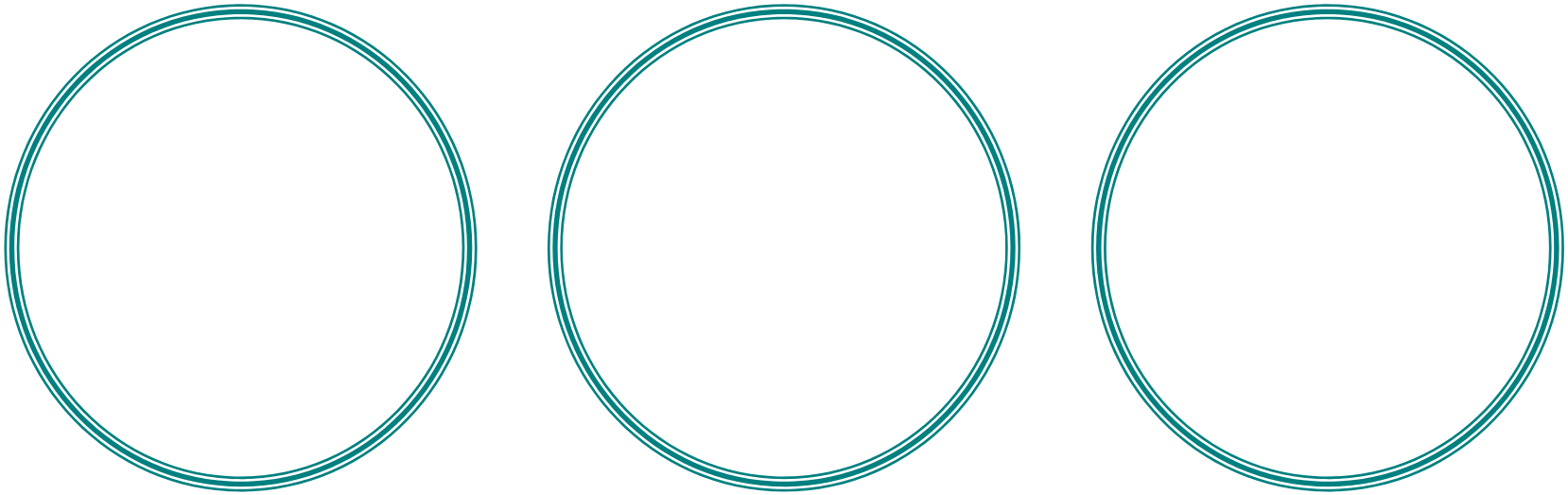
<u>The question should ***not*** be:</u> What is the probability that two children in the same family will die of SIDS?

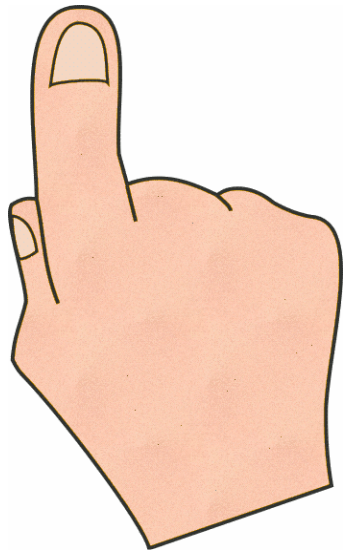
But: What is the probability that two children who die in the same family suffered from SIDA



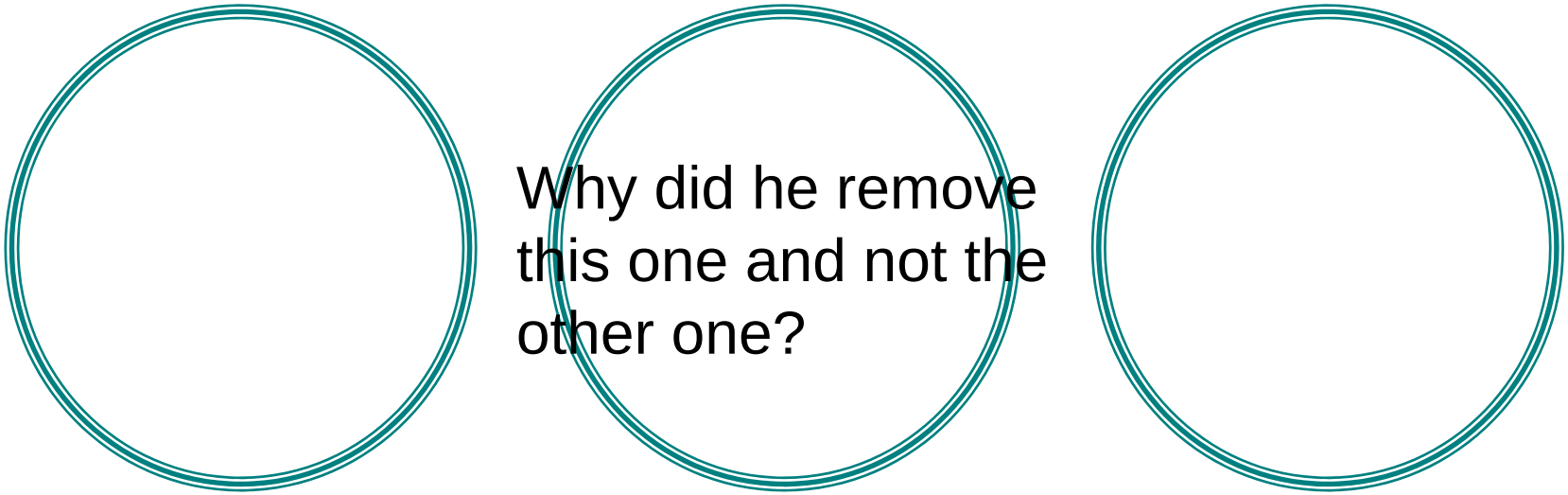


g) Where is the Euro coin hidden?





Where is the Euro coin hidden?



Why did he remove this one and not the other one?

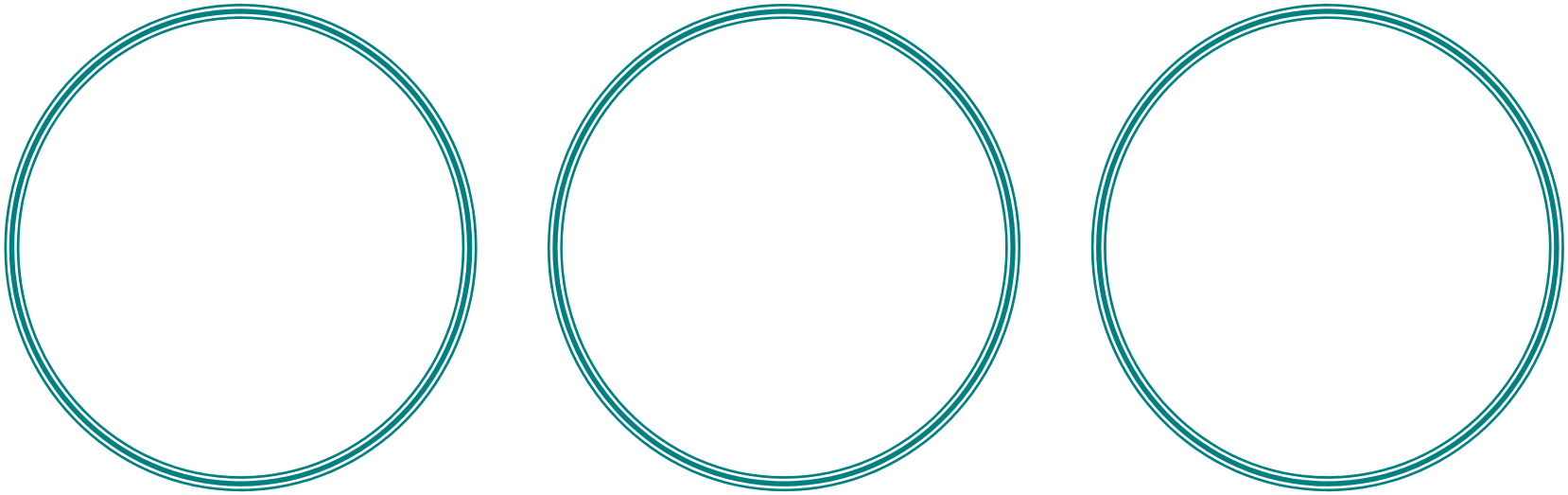
2/3 chance

1/3 chance

1/3 chance

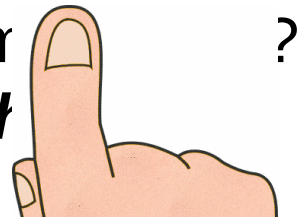
This would never have been moved

Where is the Euro coin hidden?



Changing or not????? Which choice would make m ???

If I change I will lose – ONLY if I am right from th



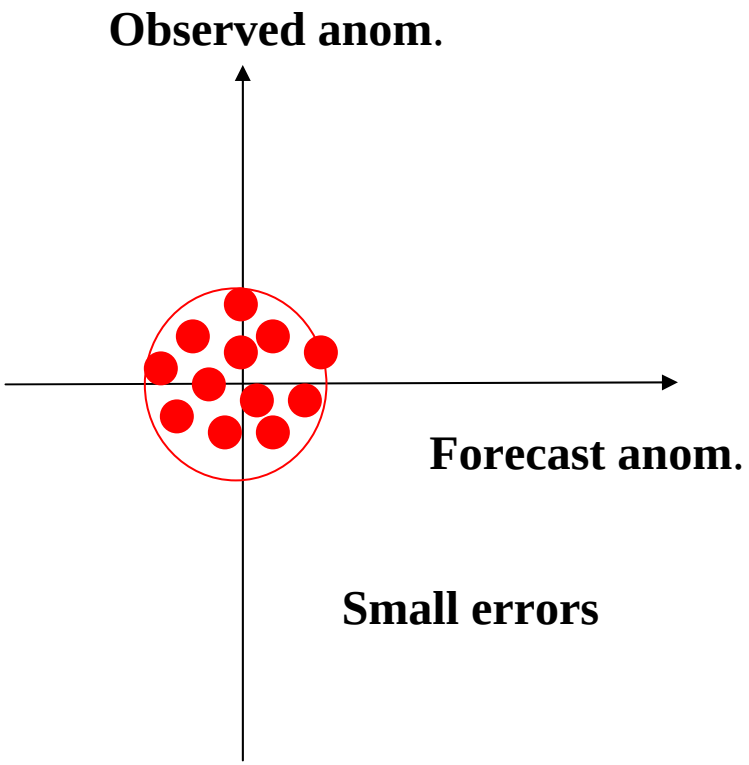
Probability being right from the start = 1/3

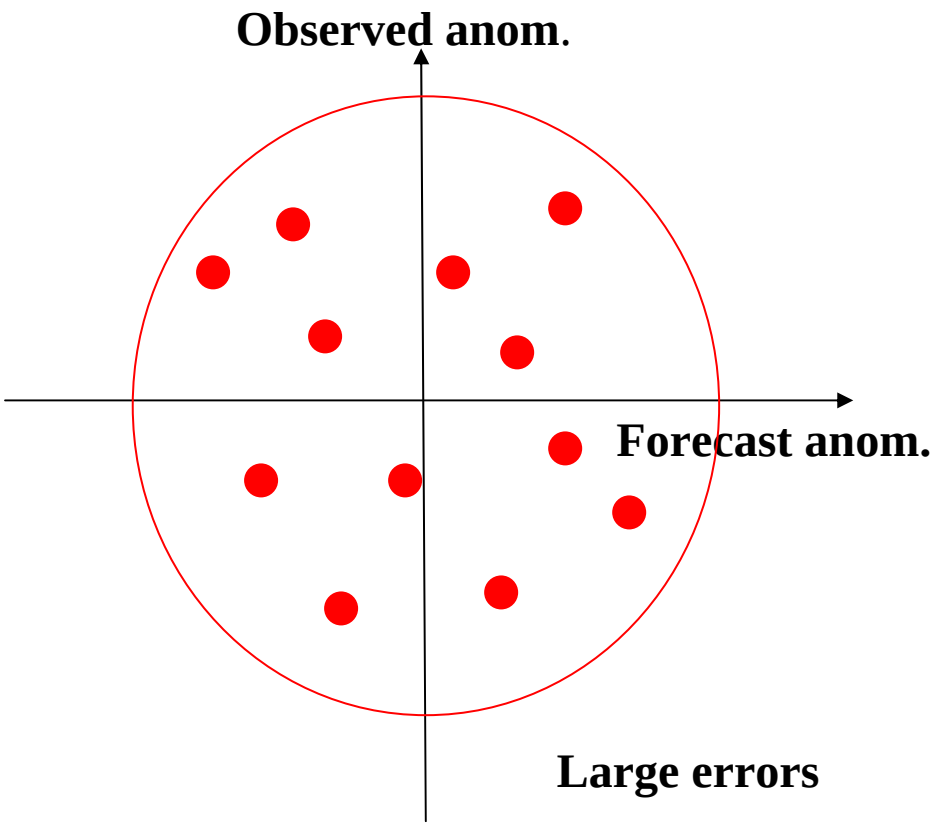
Probability losing after change = 1/3

And therefore the probability winning = 2/3 !!!

V.2.3 Interpreting the ACC

The ACC does not measure accuracy (error) but the correspondence between forecast and analysis (reality)

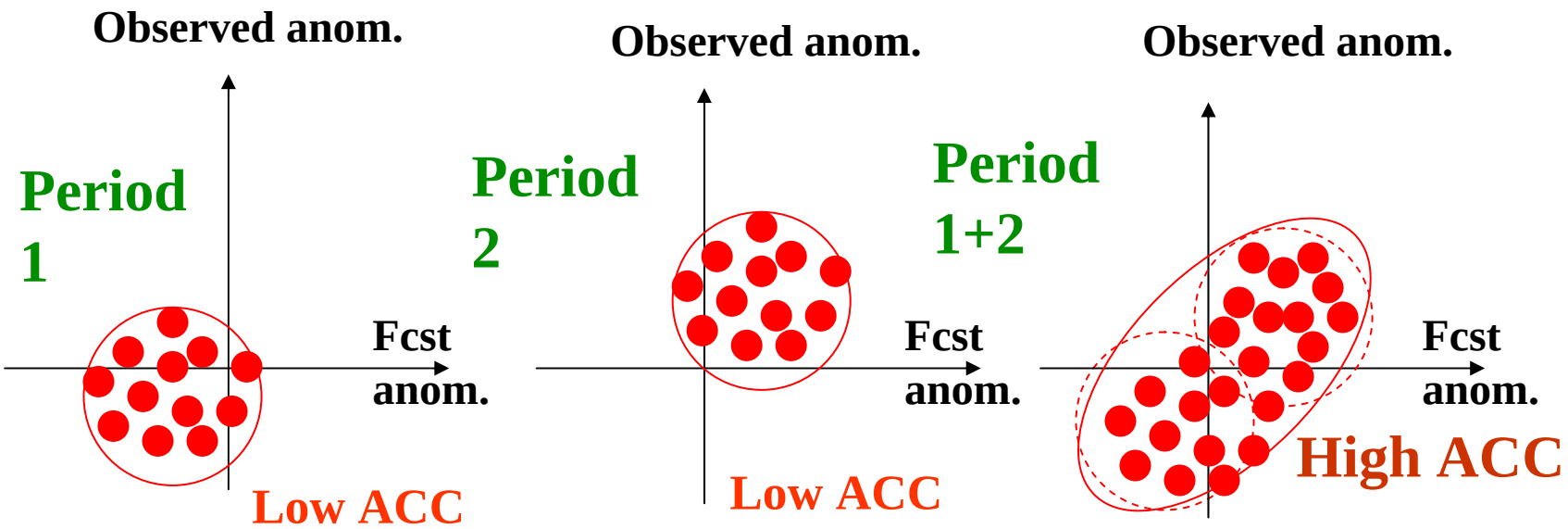




Both sets of forecasts correlate ”badly”, but only the forecast to the right might be considered ”bad”

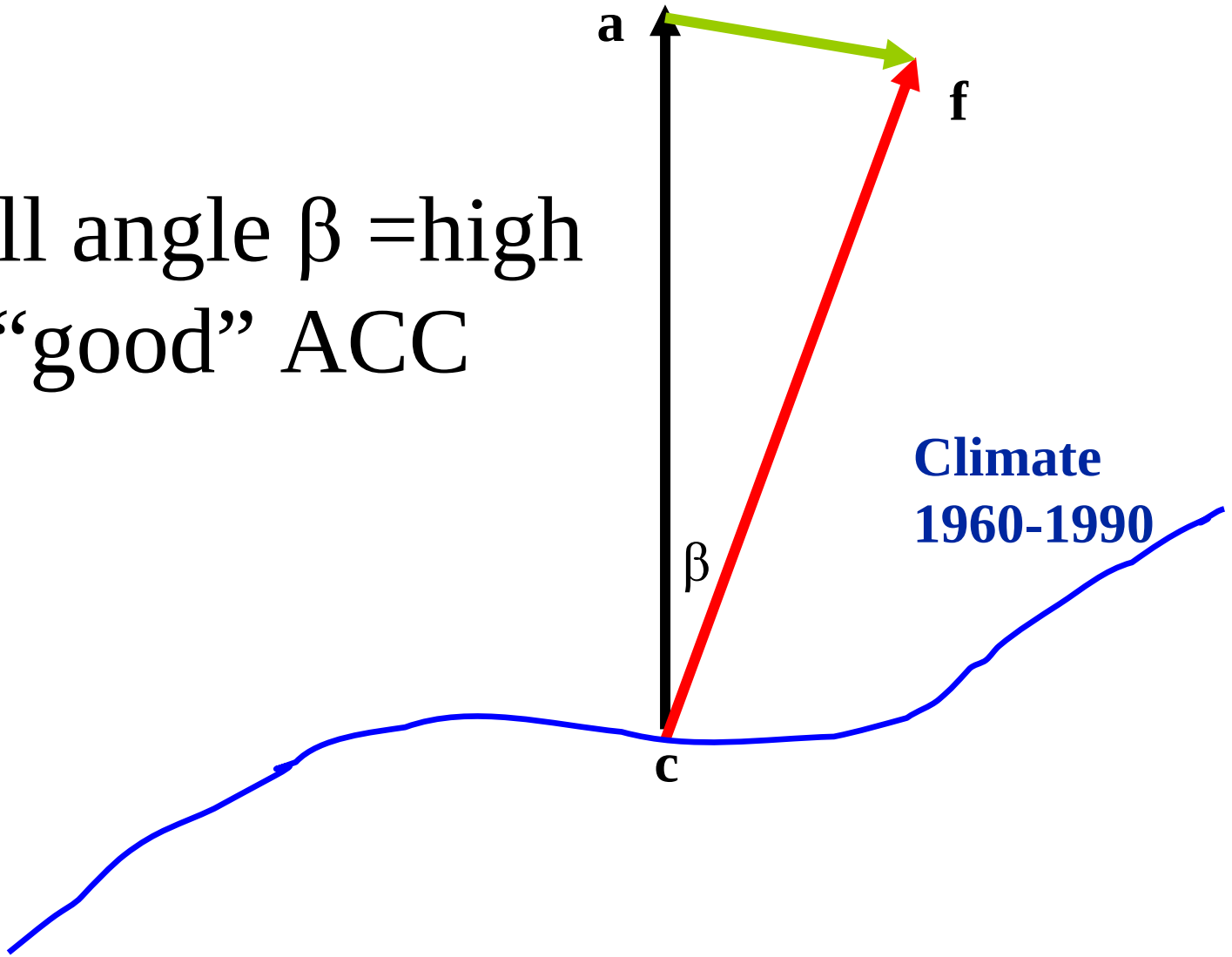
Two short periods can both have low correlation

combined the correlation might increase considerably



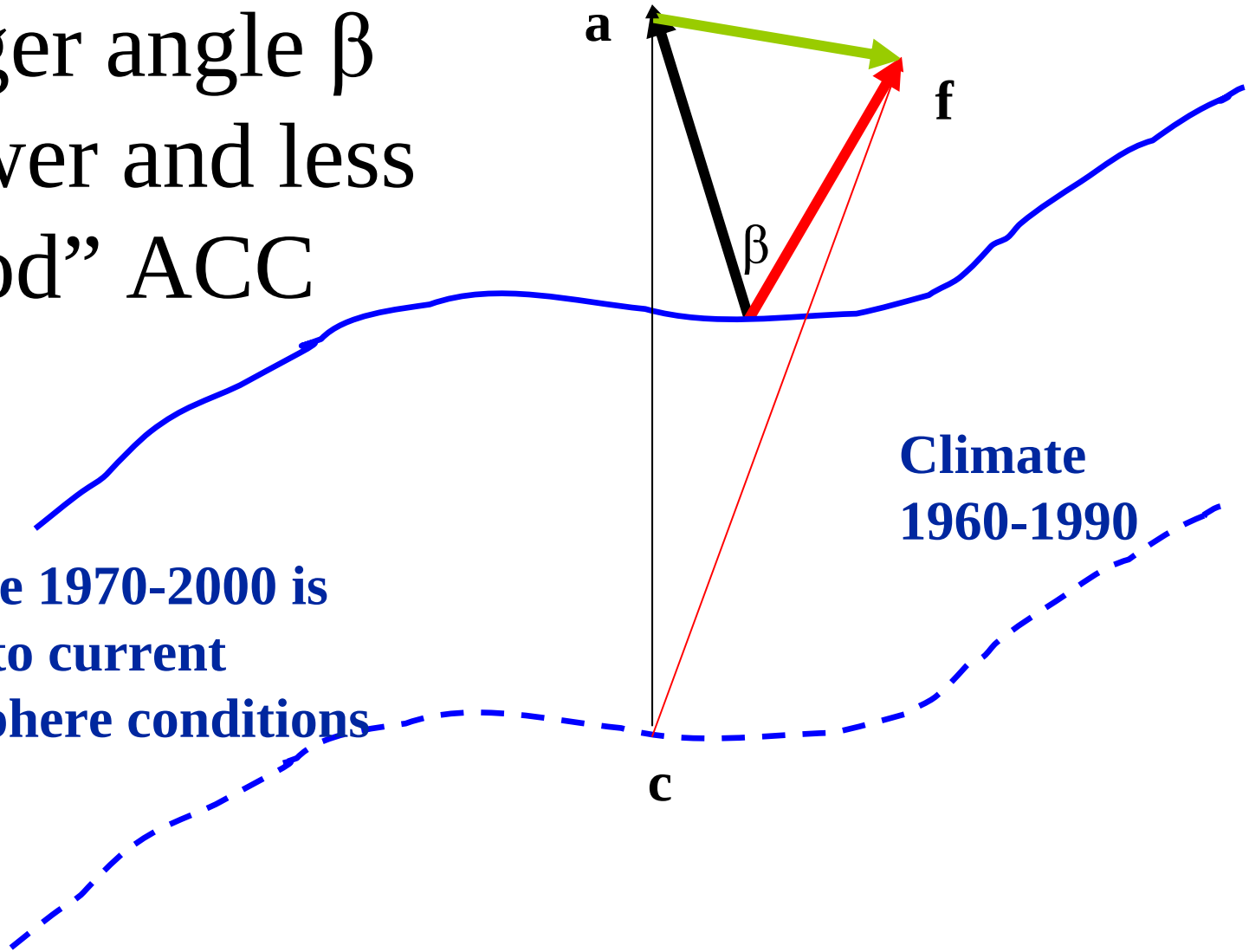
Anomaly correlations and climate reference

Small angle β =high and “good” ACC



Anomaly correlations and climate reference

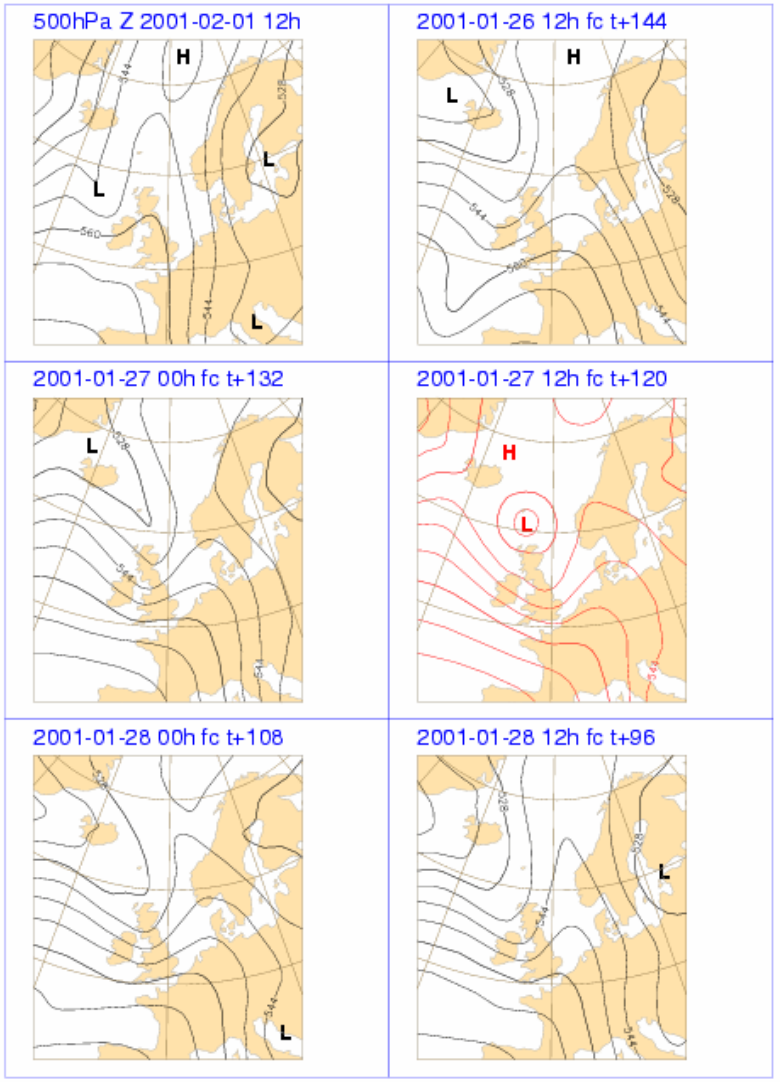
Larger angle β
=lower and less
“good” ACC



V.2.4 Jumpiness

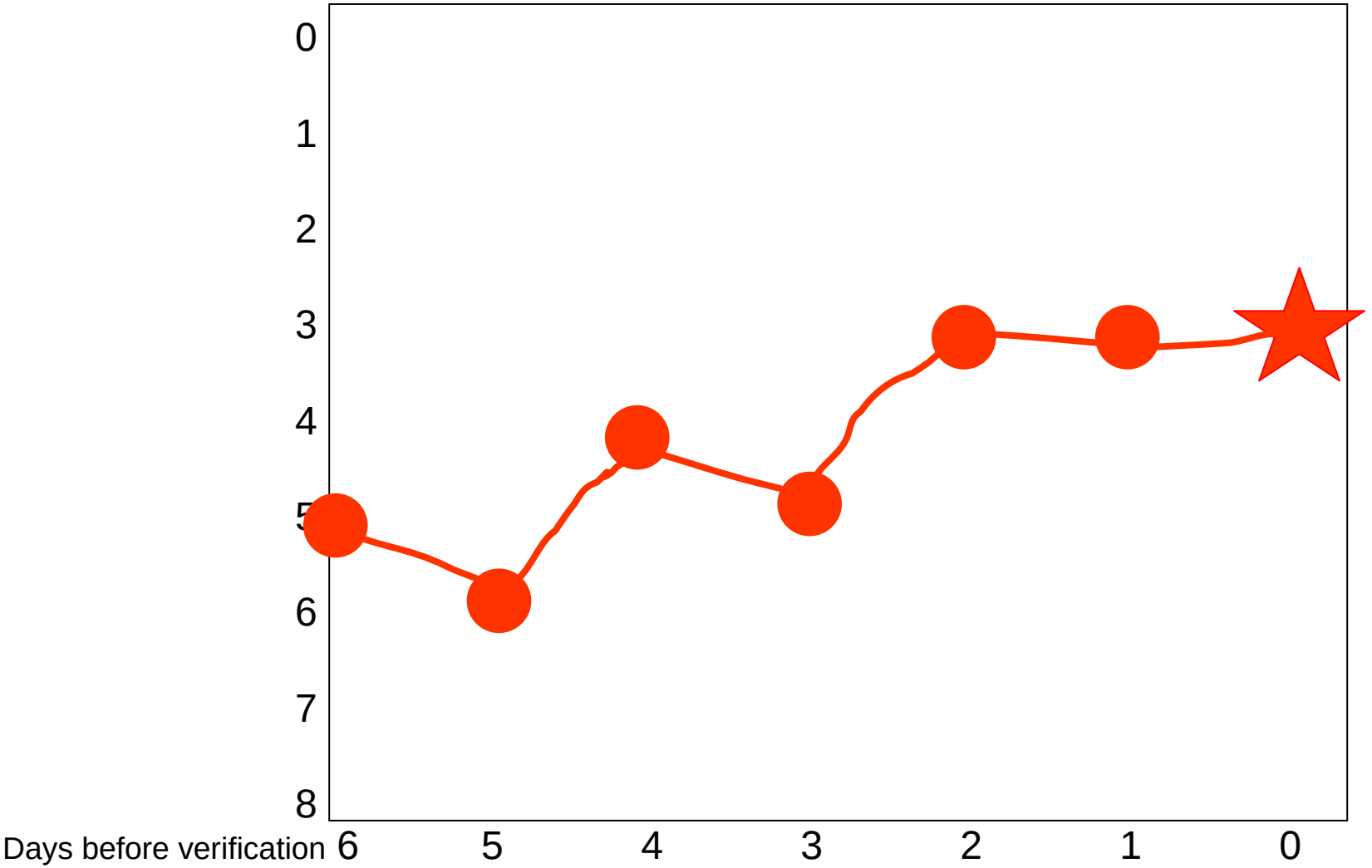
The “jumpy” D+5 forecast

- The D+5 forecast from 27 January 12 UTC is clearly “the odd man out”
- *But at the time there were few if any possibilities to know that this was a bad forecast*



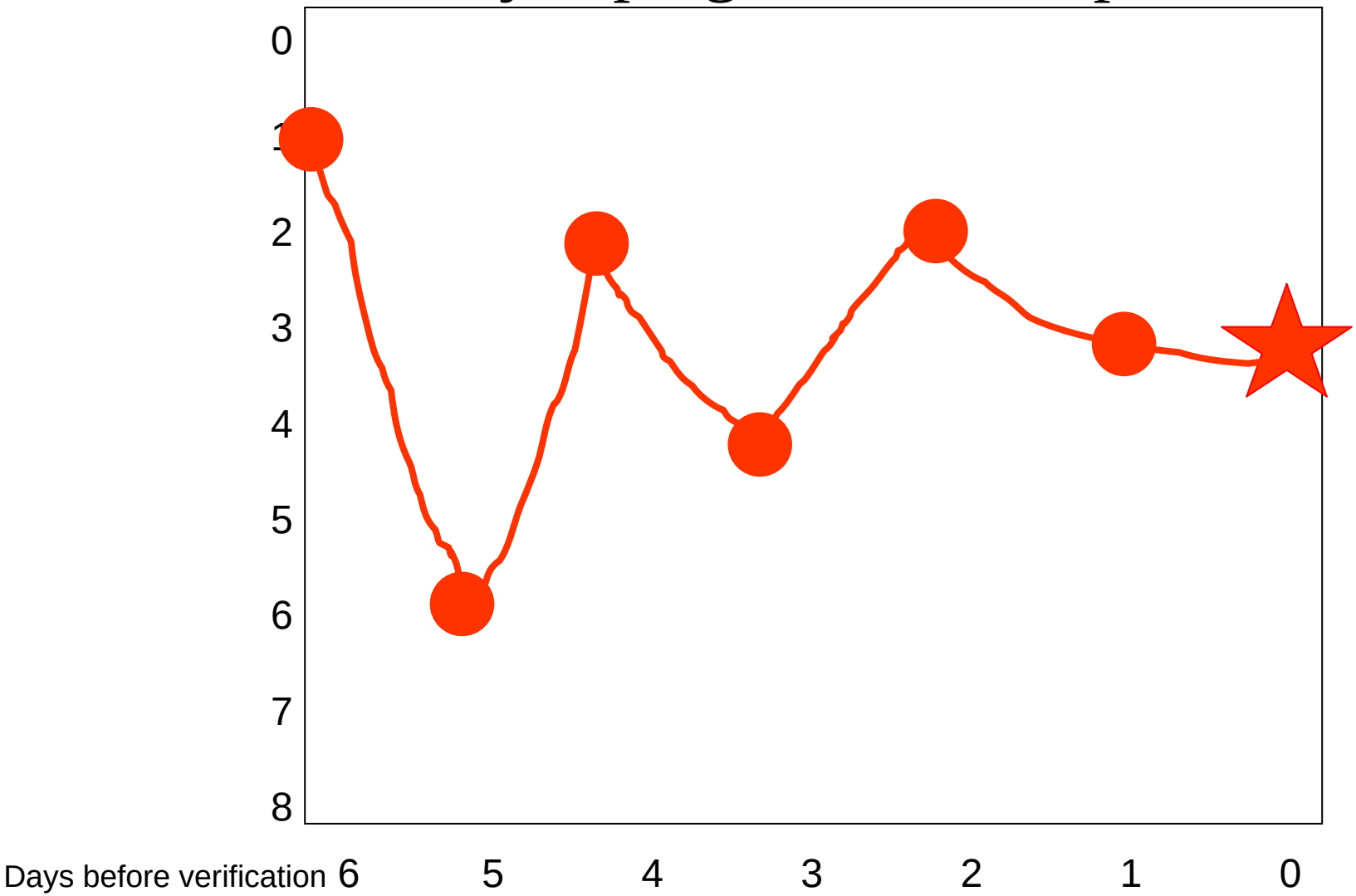
How it should be:

Forecasts continually approaching the truth

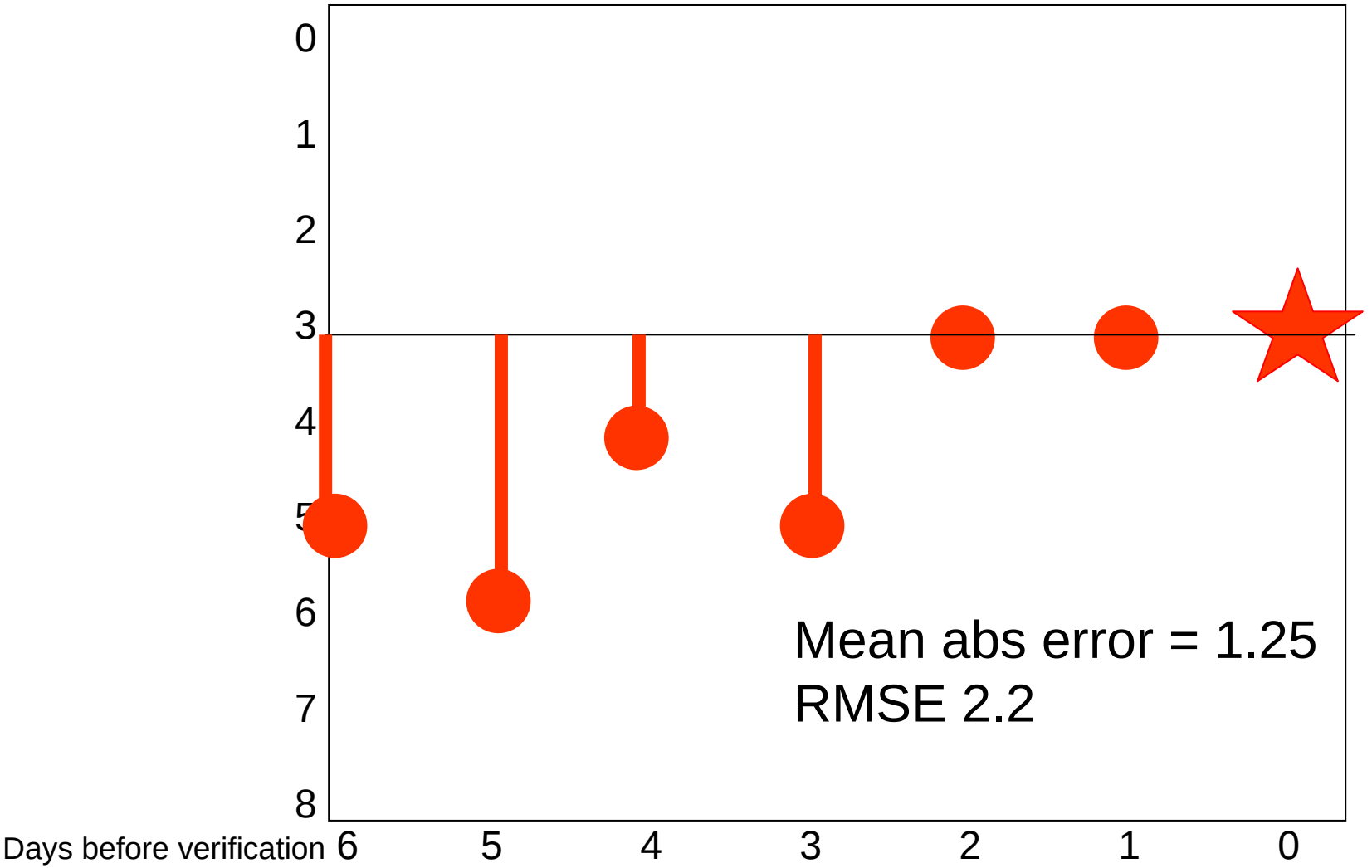


How it often is:

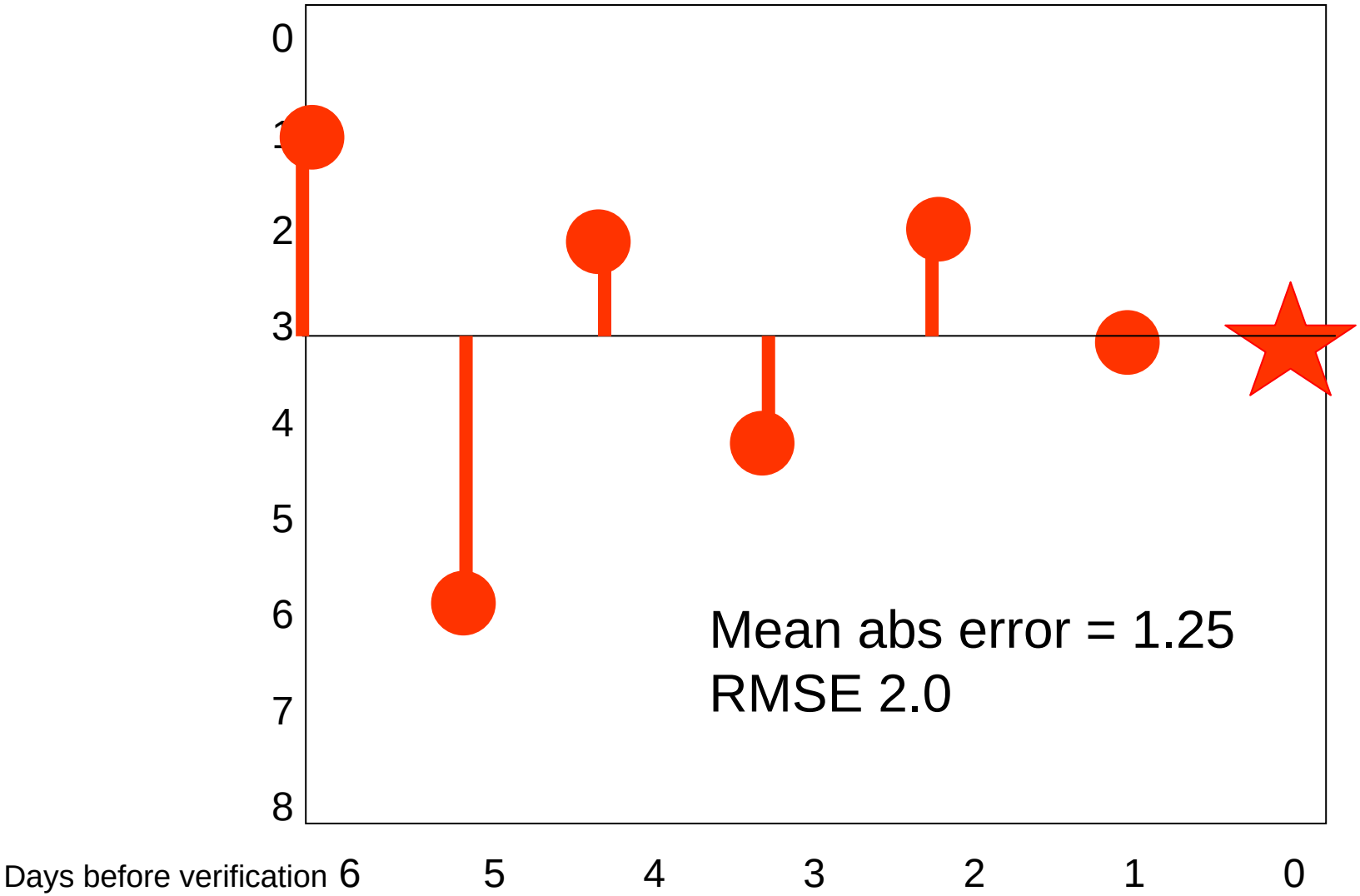
Forecasts jumping all over the place



Forecasts continually approaching the truth may not be better……



…than forecasts jumping all over the place

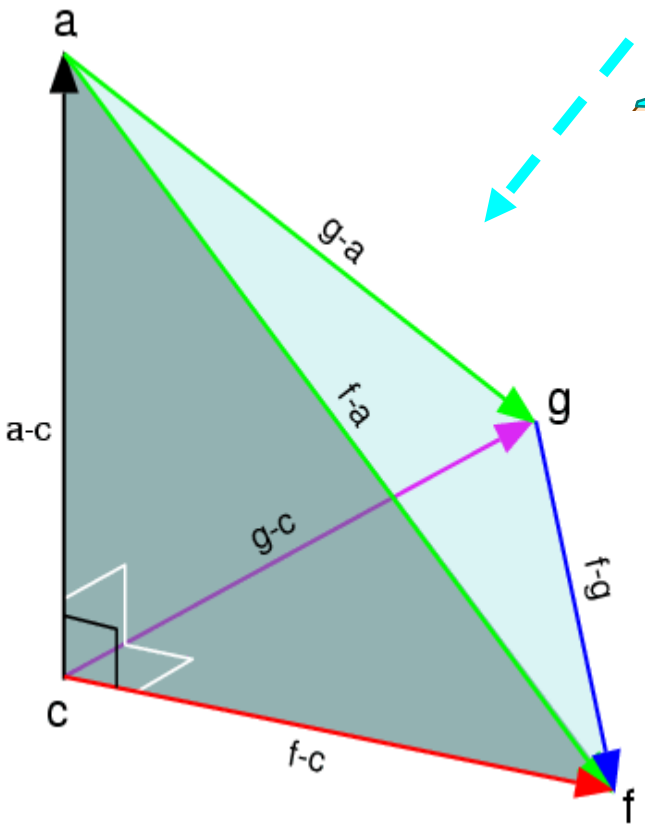


Are “jumpy” forecasts worse?

There is a 30% correlation between D+6/D+5 “jumpiness” and D+5 error

The spurious consistency-skill correlation

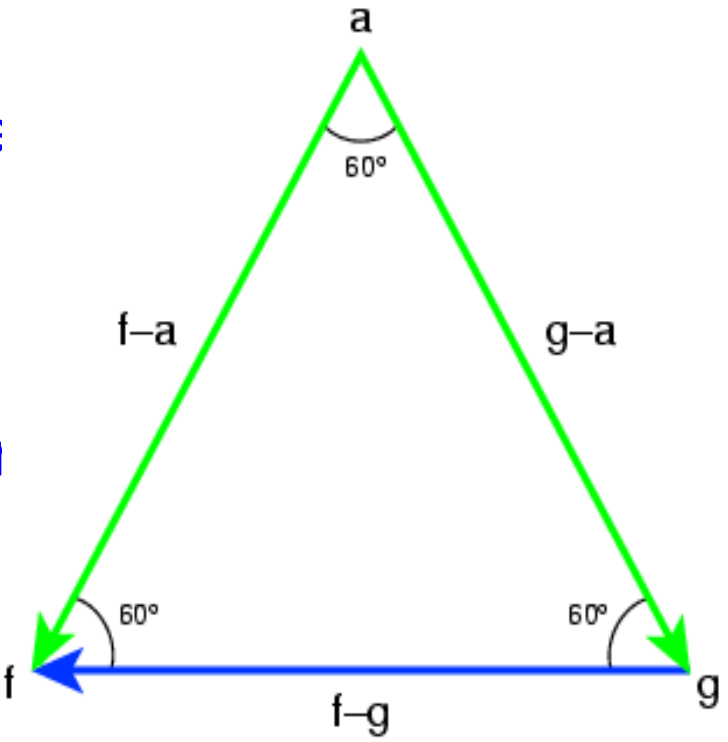
- Two forecasts systems (f) and (g) lack predictive skill and are mutually uncorrelated.
- **This implies that all three vectors are perpendicular (90°)**



Let us now watch this 3D figure from upper right...

The spurious consistency-skill correlation

- Whereas the analysis vector (a) and the forecast vectors (f and g) are perpendicular, *their difference are not!* Their mutual angles are 60° which implies correlations of 50%.
- It is when the forecasts start to display skill and mutual correlation that the 50% correlation starts to decrease the 30% level sometimes reported at a D+5 or D+6 range



END