ANKAM

Statistical Bias and Financial Decision Making

Saurabh Singal

SINGAPORE MANAGEMENT UNIVERSITY   21 FEBRUARY 2017

Behavioral Biases vs. Statistical Paradoxes

The pop-behavioral finance & pop-psychology books have had such a good run that most of us know all about various kinds of biases such as :

Anchoring, Frame Dependence, Mental Accounting, Availability Bias, Representativeness, Overconfidence, Self-Attribution Bias, Illusion of Control, Confirmation Bias, Hindsight Bias, Recency Bias, the Winner's Curse, Endowment effect, Status Quo Bias, Loss Aversion, Regret Avoidance, Lottery Effect,…

However, there are some very important **statistical biases** related to sample selection that are less commonly talked about. Theses biases can induce spurious correlations.

Statistical Paradoxes and Fallacies-Outline

We will first cover examples of these Statistical Paradoxes

- Simpson’s Paradox
- Berkson’s Paradox
- Stage Migration or the Will Rogers’s Effect
- Kelley’s Paradox
- Law of Small Numbers
- Survival Analysis

Warm-up by seeing Bayes' Theorem in action

- Boy Girl Paradox
- Three Prisoners Paradox
- False Positives Paradox

What is Simpson's Paradox?

Simpson's paradox: a reversal of the direction of an association when data from several groups are combined to form a single group.

- For example, When the data are examined as one group, the association between X and Y is positive, but when the data are split into groups based on some other characteristic, say W, the association between X and Y is negative.

- This is a type of **Omitted-Variable Bias (OVB)**: occurs when a model incorrectly leaves out one or more important causal factors. The model compensates for the missing factor by over- or underestimating the effect of one of the other factors.

Simpson's Paradox : *Smoking Increases Longevity*

	Smoker ?	
	Yes	No
Dead	107	132
Alive	174	175
Total	281	307
% Dying	38.10%	43.00%

	Age Group					
	45-54		55-64		65-74	
	Smoker ?		Smoker ?		Smoker ?	
	Yes	No	Yes	No	Yes	No
Dead	27	12	51	40	29	101
Alive	103	66	64	81	7	28
Total	130	78	105	121	36	129
% Dying	20.80%	15.40%	48.60%	33.10%	80.60%	78.30%

Source: http://www.stat.osu.edu/~biostat/newsletters/volume2_2/article_vol2_2.html

Simpson Paradox happens when…

1. $\frac{a+b}{c+d} > \frac{e+f}{g+h}$

2. $\frac{a}{c} \leq \frac{e}{g}$

3. $\frac{b}{d} \leq \frac{f}{h}$

- Finance Examples – use variables like sector, size, …

Simpson's Paradox : *Discrimination In Hiring*

- Hooters has **2** jobs, cooks and service staff.
  - For the job of cooks, 20 men and 5 women apply.
  - For the job of service staff, 5 men and 20 women apply.
- Hooters hires **4** men and **1** woman as cooks.
- Hooter also hires all **5** men and **20** women as service staff.
- Over all, Hooters hired **9** men (out of **25** male applicants) and **21** women (out of the **25** female applicants).
- Total analysis: A candidate has **36%** chance of being hired if male; and **84%** if woman. *Apparent evidence of discrimination.*

Simpson's Paradox : *Discrimination In Hiring -2*

- Granular analysis: A candidate applying for the job of a cook has a 20% chance to be hired**, and this is the same for men and women.**
- A candidate applying for the job of service staff has a 100% chance to be hired, **same for men and women**.
- *There is* ***NO*** *evidence of discrimination.*
- The scary thing is every time we do this sort of thing and get a result it does not mean anything if there is another confounding variable. For example the cooks might really be top chef and sous chef. They only hire men for top chef, only hire women for sous chef. You would miss the discrimination then.
- How far do you go down? Far enough to get the answers you want!!!

Berkson's Paradox : ***Why Are Handsome Men Such Jerks?***

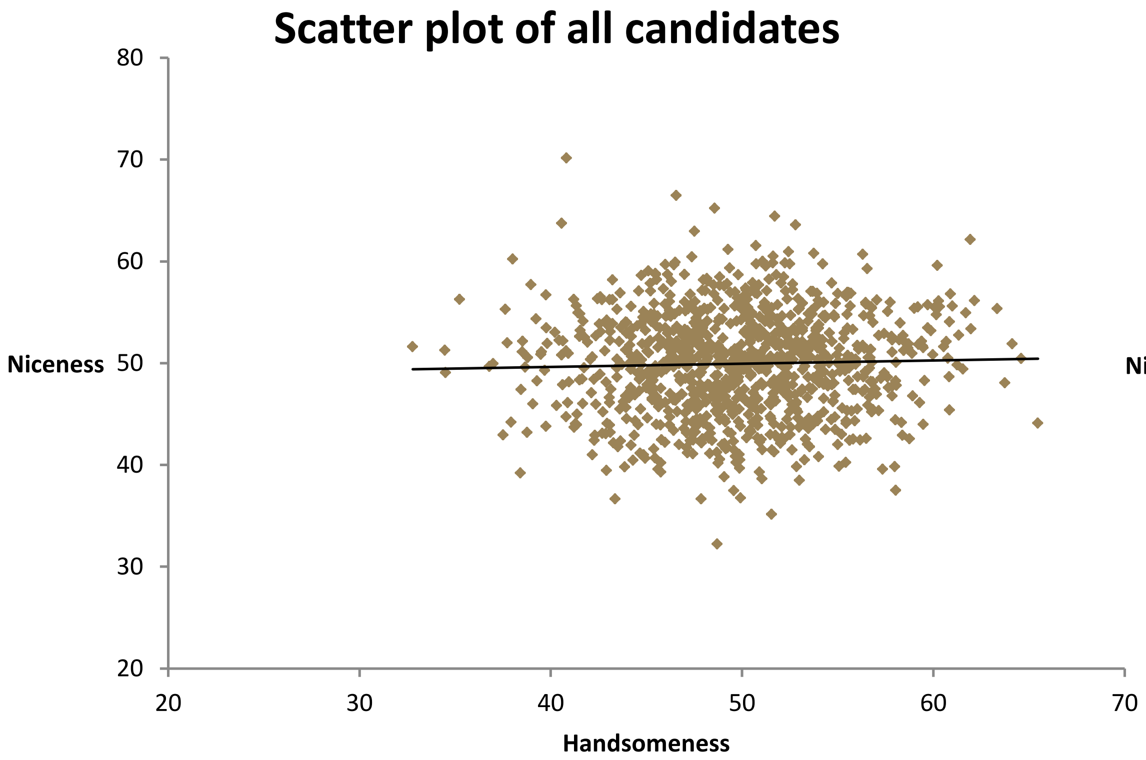
Title of a Slate article by the Wisconsin professor, Jordan Ellenberg

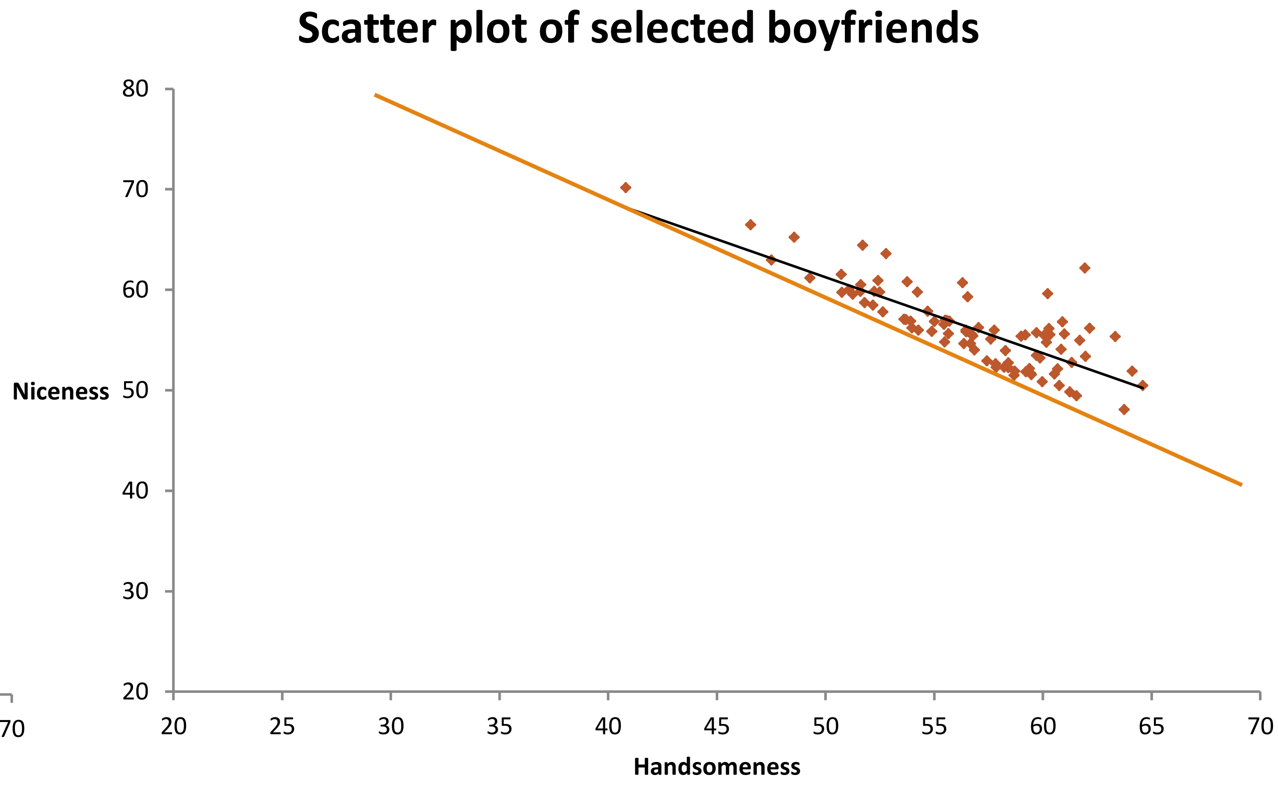
"you will date a man if his niceness plus his handsomeness exceeds some threshold. … So, among the men that you date, the nicer ones are less handsome on average (and vice-versa), even if these traits are uncorrelated in the general population."

This kind of selection bias is called **Berkson's Paradox**.

Beware of this when using sum of two ranks to select a universe.

Candidates Ranked On Handsomeness And Niceness





- The black lines are regression lines. Data on left show no correlation.
- Negatively sloped regression line in right chart shows negative correlation.
- Only points above blue line can be selected, leading to sampling bias.

Stage Migration Or Will Rogers Effect

- Will Rogers remarked: *When the Okies left Oklahoma and moved to California, they raised the average intelligence level in both states.*
- Stage Migration is the name given to this movement of people from the set of healthy people to the set of unhealthy people
  - Improved techniques in detection of illness leads to some people being classified as “unhealthy” at an early stage.
  - This leads to the increase in average life span of both the healthy and unhealthy groups.
  - Both lifespans are statistically lengthened, even if the patients did not live longer.
  - Example : $A=\{1,2,3\}$ and $B=\{4,5,6\}$. If we move 4 from set B to set A, the mean of both sets will increase.

How A Quant Strategy Develops

- Observation
- Common place events
- Unusual Events
- How improbable are these unusual events and what can be learned from them?
- Simplicity - Hemingway spent time forgetting tough words.
- Not too simple - Brain surgery is impossible with a pen knife.
- We have already met the Omitted-Variable Bias.

October 2014

In October 2014,

- the world’s equity markets had a string of big negative days.
- Many equity indices were in “correction” territory or down more than 10%. And talks of bear markets were rampant.
- And then the equity markets also had a string of positive days.

We could ask some questions:

- What is the probability of a 1-day move>1%?
- Or, drill down : in a bull market, what is the probability of a string of large 1-day moves
- Or invert the question:

Regression to the Mean

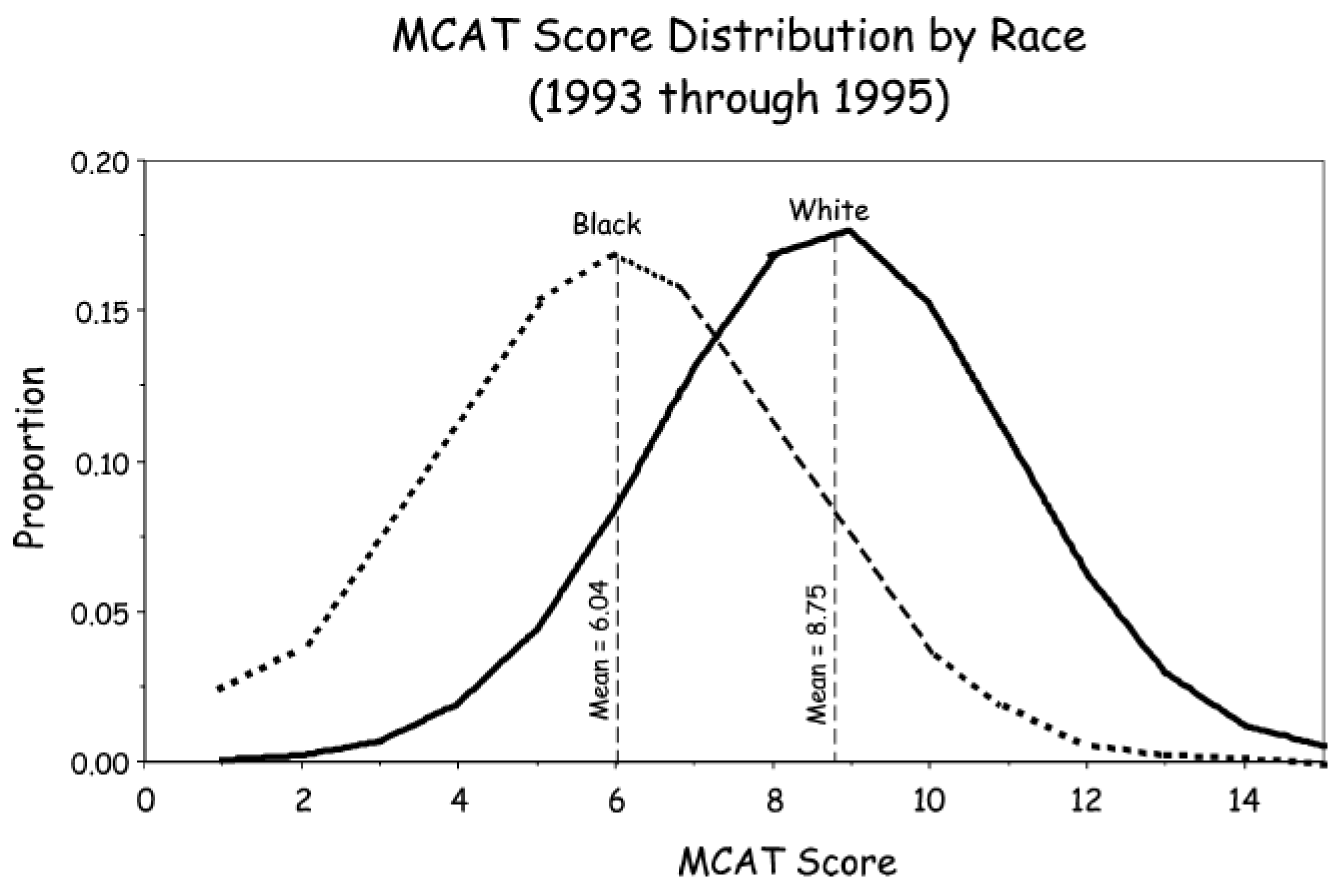
- **Regression to the mean** - if a variable is extreme on its first measurement, the second measurements of the same variable tends to less extreme , *i.e.*, tends to be closer to the average on its second measurement—and vice versa. An initial sampling bias may disappear as new, repeated, larger samples display sample means that are closer to the true underlying population mean.
- Francis Galton observed the phenomenon, when analyzing the heights of fathers and sons. Very tall fathers had sons who were tall but not as tall. He wrote *Regression towards mediocrity in hereditary stature).* In fact this is how linear regression was invented.
- One example is exams. A students performance on an exam depends on skill but also on luck. The students with above average scores could be either the skilled ones with not too much bad luck on that day OR somewhat less skilled ones with lots of good luck. The skill will not change much from day to day but the randomness or luck will change.
- If a class of students takes two tests similar in difficulty and testing the same type of ability on two successive days, the highest performers one first day will do less well on the second day. Similarly, the worst performers on the first day will tend to do better the second day. And vice versa.

Kelley's Paradox

- STRIVER program - black candidates with lesser score on the MCAT were considered to have similar ability than white candidates with the same score. The argument is that some one from a disadvantaged background or hailing from a group with smaller mean has better chance of doing well in future when he is given the same opportunity as candidates from the other, higher mean group.
- But the thinking is inverted! And regression to the mean is at play…

Source :Strivers example from Wainer, Howard and Lisa M. Brown. "Three Statistical Paradoxes in the Interpretation of Group Differences: Illustrated with Medical School Admission and Licensing Data." Handbook of Statistics, Vol. 26, 2007. Elsevier B.V.

Kelley's Paradox-2 : *example from Wainer & Brown*



Kelley's Paradox and Regression to the Mean

Kelley's equation relates the estimated true score ($\hat{\tau}$) to the observed score (x).

- "True Score": average of the person's observed scores if he takes identical tests several times.
- The best estimate of true score: regress the observed score in the direction of the mean score(μ) of the group that the examinee came from. The amount of the regression is determined by the reliability (ρ) of the test.
- Kelley's equation : $\boldsymbol{\hat{\tau} = \rho(x) + (1 - \rho)\mu}$
- A test can be completely unreliable (ρ = 0), perfectly reliable (ρ = 1) but usually it is in between these two extremes

Kelley's Paradox and Nobel Prize Winners

- Horace Secrist wrote a book , *The Triumph of Mediocrity in Business* . He claimed that the great depression was caused by firms becoming mediocre over time. The companies with the highest earnings one decade ago were performing only a little better than average. And vice versa for the worst performers. The great statistician, Harold Hotelling, showed that this was a statistical fallacy, caused by failure to note that regression to the mean was at work.

- William Sharpe explained the convergence, of the most profitable and least profitable companies between 1966 and 1980. He said that *ultimately economic forces will force the convergence of profitability and growth rates of different firms*. Another Nobel prize winner, Milton Friedman agreed with Sharpe's results but disagreed with his explanation. Friedman sowed that the true reason for convergence was statistical, not economic. Sharpe later dropped this chapter from the next edition of his book.

Nobel laureates are not immune to Kelley's Paradox

- Galton - Father and son's heights
- Secrist and Hotelling
- William Sharpe and Milton Friedman

Probability Inversions

- Is it the case that bear markets have more large up days than bull markets?

- Given that we saw a large number of big positive days, what is the probability we are in a bear market?

Examine Not $P(\text{many big positive days} \mid \text{bull market})$

but $P(\text{we are in a bear market} \mid \text{many big positive days})$?

Need to use Bayes' Theorem

Law Of Small Numbers

- The Bill and Melinda Gates Foundation has funded the movement that supports smaller schools.
- *It is claimed that outstanding schools are small schools.*
- *But weak schools tended to be smaller than average as well!*
- *Evidence that Smaller Schools do not improve Student Achievement.* This is the title of an article by *Howard Wainer & Harris L. Zwerling*.
- ***Which sample is likely to have the highest average? The answer is the sample with the smallest number of observations (Wainer & Zwerling 2006). The reason is that the variability of an average declines with the number of observations.***

Law Of Small Numbers - 2

- The counties in which the incidence of kidney cancer is ***lowest*** are mostly rural, sparsely populated, and located in traditionally Republican states in the Midwest, the South, and the West.

- The incidence of kidney cancer is ***highest*** in counties which tend to be mostly rural, sparsely populated, and located in traditionally Republican states in the Midwest, the South, and the West.

Law Of Small Numbers - 3

- The very small population in the counties under question makes it drawing inferences prone to error.

- A county with small population has a much greater chance to be healthier, cleaner, dirtier, richer, …, than average. But it also has a much greater chance to be less healthy, dirtier, poorer,…, than average.

- If a Nobel laureate decided to move to a village with a small population, of say 49 people, suddenly this village would rightfully boast that 2% of its residents are Nobel prize winners. (But if you or I decide to subsequently decide to reside in the same village, our chances of getting a Nobel are not going to be 2%!)

Law Of Small Numbers - 4

- Daniel Kahneman in ***Thinking Fast and Slow*** illustrates this further. One person tosses 4 fair coins repeatedly. Second person tosses 7 fair coins repeatedly.
- Now note the proportion of times each person gets a homogeneous result (either all heads or all tails). The respective probabilities are 0.5^4 X 2 and 0.5^7 X 7. Or 12.5% and 1.56% respectively, the larger number being 8 times the smaller.
- If a coin toss is replaced by probability of cancer and number of coins (4 or 7) is replaced by population of a county, then the smaller counties will dominate the top and bottom rankings.
- Imagine adding two more people that toss a single coin each. And a third who tosses a million coins
- ***This is just a statistical fact, and we must be cognizant of this.***

Survivorship Bias : *Armour On Airplanes*

- In World War II, bullet holes on fighter planes returning from combat were analyzed to determine where more armour should be put.
- Too much armour makes the plane heavier and difficult to maneuver and also unable to fly longer distances. Too little armour makes it more vulnerable to enemy fire.
- Initial suggestion by researchers: maximum damage was on fuselage (wings, nose and tail), so put more armour there. Less damage around engines, so reduce armour there.

Abraham Wald's suggestion: put more armour on engines! A hit there causes no survivors. A hit to the fuselage might still lead to survival. The initial reaction was to reinforce where damage had been found but in fact it is best to reinforce other areas because those were the planes that didn't return!

- A person shot in the hand or feet might survive; a shot to the head is likely to be fatal.
- Finance Examples –studies dealing with time series of stocks, index membership etc.

Biases Related to Bayes' Rule

$$P(X|D) = \frac{P(X \cap D)}{P(D)} = \frac{P(D|X) \times P(X)}{P(D)}$$

Which yields

$$P(X|D) \propto P(D|X) \times P(X)$$

Where

$P(X|D)$ is the posterior probability

$P(D|X)$ is the likelihood

$P(X)$ is the prior probability

The three examples (Prosecutor's Fallacy, Defence Attorney's Fallacy and errors in interpreting DNA match probability) all are connected to conditional probability and Bayes' Theorem.

The Boy-Girl Paradox

Let us start with a fun example which can be solved with or without the use of Bayes Theorem. This will give us a taste for what Bayes' Theorem can do for us!

A new neighbor moves in next door. You learn that he has **2** children.

You see that one of the children is a boy.

What is the probability that the other child is a girl?

Boy-Girl Paradox: Solution Using Bayes' Theorem

Cases:

1. Both are girls (GG) : P(GG) = **1/4**,
2. Both are boys (BB) : P(BB) = **1/4**,
3. One boy, one girl- two case, (GB) and (BG)- with probability= **(1/4)*2**= **1/2**

- Now we add the additional assumption that "at least one is a boy" : call this B.

Boy-Girl Paradox: Solution Using Bayes' Theorem (contd.)

Using Bayes' Theorem, we find:

- P(BB|B) = P(B|BB) · P(BB) / P(B)

- P(B|BB) = probability of at least one boy given both are boys = 1

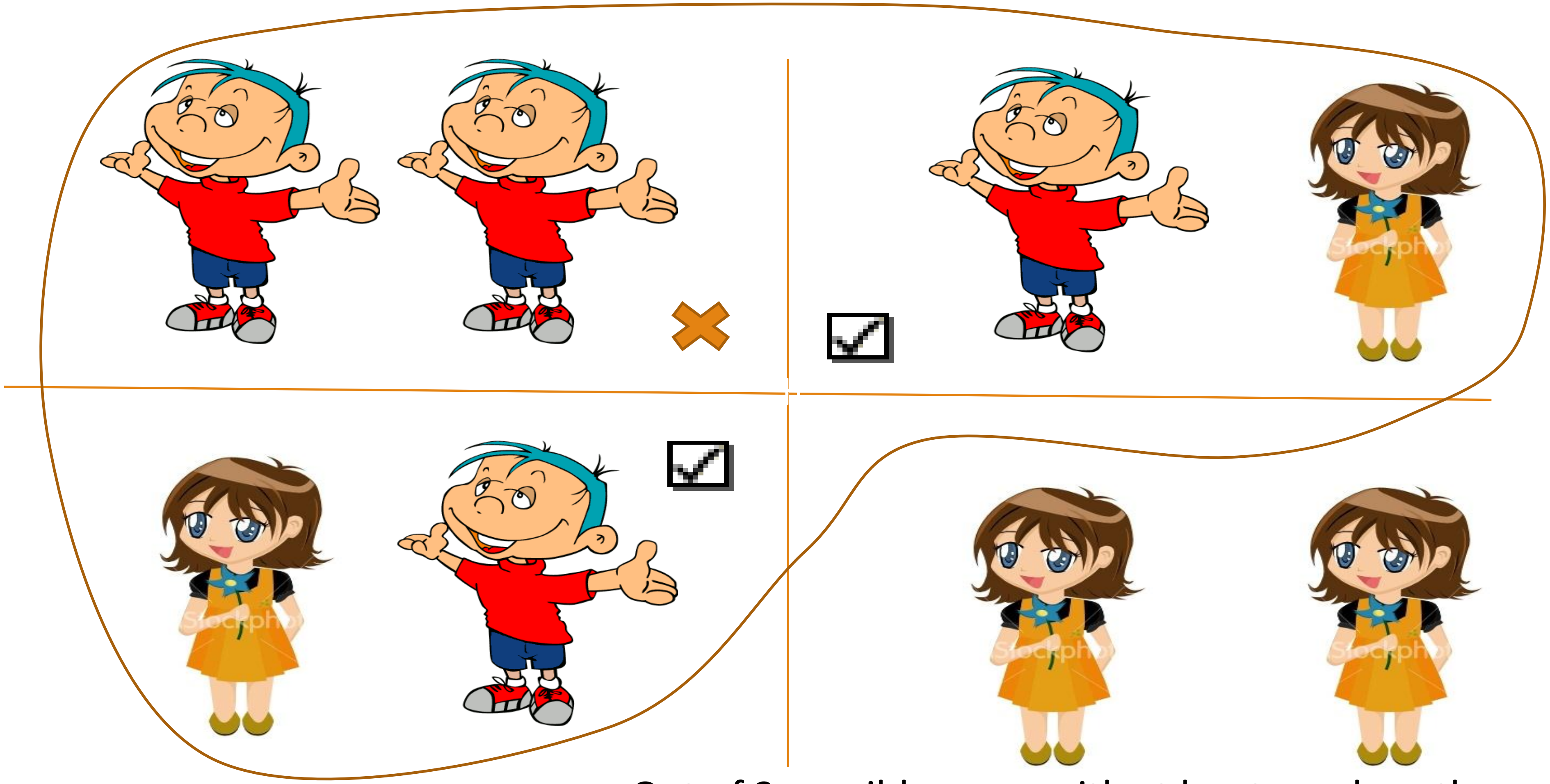
- P(B)= probability of at least one being a boy, which includes cases (2) and (3)

  = **1/4** + **1/2** =**3/4**

Finally, **P(BB|B) = (1 * 1/4)/ (3/4) = 1/3**

So the probability other child is a girl is 2/3

Boy Girl Paradox: Intuitive Solution



Out of 3 possible cases with at least one boy, there are two cases with a girl-

$Prob\,(other\ child\ is\ a\ girl) = \frac{2}{3}$

Bertrand’s Box Paradox

There are **3** boxes with gold and silver coins.

- Box **A** has **2** gold coins,
- Box **B** has **2** silver coins and
- Box **C** has **1** gold coin and **1** silver coin.

You randomly pick a box and then draw one coin. Suppose it is a silver coin.

What is the probability the other coin in this box is also silver?

Bertrand's Box Problem: The Solution

- Let us label the two gold coins in Box A as G_1 and G_2
- The two silver coins in Box B as S_1 *and* S_2.
- And the coins in Box C as S_3 and G_3.
- So there are three possibilities that could lead to picking a silver coin:
- I chose box B and picked S_1, the other coin is S_2
- I chose box B and picked S_2, and the other coin is S_1
- I chose box C and picked S_3, the other coin is G_3

Therefore 2 favourable cases out of 3 => 2/3 probability that the other coin is of the same type, in this case silver.

Calculation Process

$$\frac{P(S|SS\ box)}{P(S|GG\ box) + P(S|GS\ box) + P(S|SS\ box)} = \frac{1}{1+0+\frac{1}{2}} = \frac{2}{3}$$

Here, SS Box is the box with 2 silver coins, or Box B
GG Box is the box with 2 gold coins, Box A
GS Box is the box with a gold and a silver coin, or Box C.
Applying Bayes' Theorem gives us the answer, which is 2/3

Three Prisoner's Paradox

Three prisoners, A, B and C are sentenced to death.

One of the prisoners will be pardoned.

Prisoner A requests the warden to tell him the identity of one of the others who is going to be executed. If B is to be pardoned, give me C's name. If C is to be pardoned, give me B's name. And if I'm to be pardoned, randomly flip a coin and based on it, tell me either B or C.

The warden tells A that B is to be executed.

Prisoner A is pleased because he believes that his probability of surviving has gone up from **1/3** to ½ .

Is he justified? And what is the probability of surviving for C now?

Three Prisoner's Paradox : solution using Bayes' Theorem

Call ***A,B, C*** the events that the corresponding prisoner will be pardoned, and ***b*** the event that the warden mentions prisoner B as the one not being pardoned, then, using Bayes' formula, the posterior probability of A being pardoned, is:

$$P(A|b) = \frac{P(b|A)P(A)}{P(b|A)P(A) + P(b|B)P(B) + P(b|C)P(C)}$$

$$= \frac{\frac{1}{2} \times \frac{1}{3}}{\frac{1}{2} \times \frac{1}{3} + 0 \times \frac{1}{3} + 1 \times \frac{1}{3}} = \frac{1}{3}$$

Three Prisoner's Paradox: intuitive explanation

If B will be executed, it is because

- either C will be pardoned (**1/3** chance), or
- A will be pardoned (**1/3** chance) and the B/C coin the warden flipped came up B (**1/2** chance; for a total of a **1/6** chance B was named because A will be pardoned).

Prisoner A does not gain any new information, so his probability of survival should be the same as earlier, i.e., 1/3.

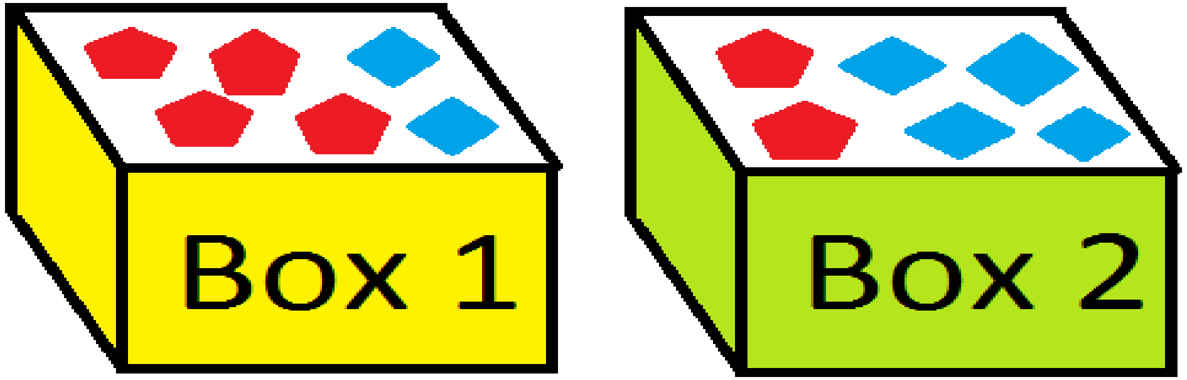
That means there is a **2/3** chance for C to get a pardon.

Let's Play a Game!

There are two boxes containing red and blue checkers.

Box 1 has **100** red and 50 blue checkers.

Box 2 has **50** red and **100** blue checkers.



Let's Play a Game -2

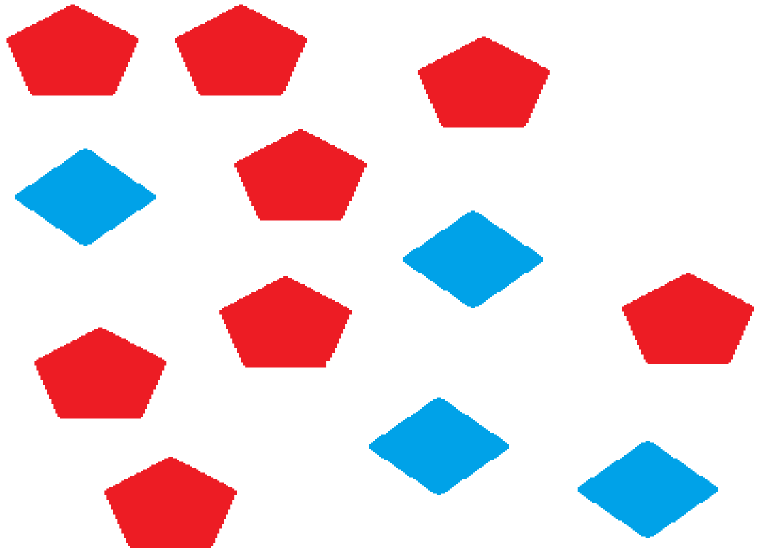
We select a box randomly

Assume there is **50%** chance of either box being selected.

Sampling with replacement, we pull **12** checkers out of this same box

Of the **12** checkers, **8** are red and **4** are blue.

What is the chance we picked them from the second box? (the one with **50** red and **100** blue checkers)



Let’s Play a Game: Answer

It is difficult at first for most of us to guess the correct answer: as ***1/17***.

- The prior probability of either box being selected is **50%**.
- When we see more red than blue checkers, we can assume there is a higher than **50%** chance that they are drawn from box **1**, which has more red checkers.
- so the chance of box **2** having been selected is < **50%**.
- It is difficult to guess it is so much lower.

The human brain is very good at pattern recognition but poorly designed for computing probabilities, esp. conditional probabilities.

Let’s Play a Game: Motivation

What if the red checkers were down days and Blue checkers are up days; Box 2 was a Bull market and Box 1, a Bear market

AND

If we had to guess which market regime we are in?

Application Of Bayes' Theorem In Finance

Black-Litterman Framework

- In the Markowitz mean-variance optimization framework – the stocks or assets have a multivariate-normal distribution.

- If we want to revise our estimates of return on individual stocks the proper way is to go through the variance-covariance matrix due to the correlations – i.e., to compute the distribution conditional upon our views. This is the Black-Litterman Framework.

Bayes-Stein shrinkage: We "shrink" stock beta towards unity. Very high values of beta are lowered somewhat, and very low values get increased.

Regime Detection, Bayesian Networks, Hidden Markov Models

Application Of Bayes' Theorem In Finance Less-Technical Example

- Finance people and ordinary people are somewhat Bayesians without being conscious of it. For example, the research analysts revise their price targets based on the current or historical prices.

- And any one who takes into account the long term drift in equities is accounting for the base rate. For the S&P 500, about **54%** days are positive, as are about **60%** months, **66%** quarters and **75%** years.

Dangers of Chart Reading

- Human eye-brain system is good at the cognitive task of comparing position along a common scale.
- However, it is not so good at elementary perceptual tasks that involve making judgment about length, area, shading, direction, angle, volume, and curvature.
- Source: William S. Cleveland and Robert McGill, “*Graphical Perception: Theory, Experimentation and Application to the Development of Graphical Methods*.”
- Use of charts can be dangerous as we can fall prey to illusions.

Three Optical Illusions

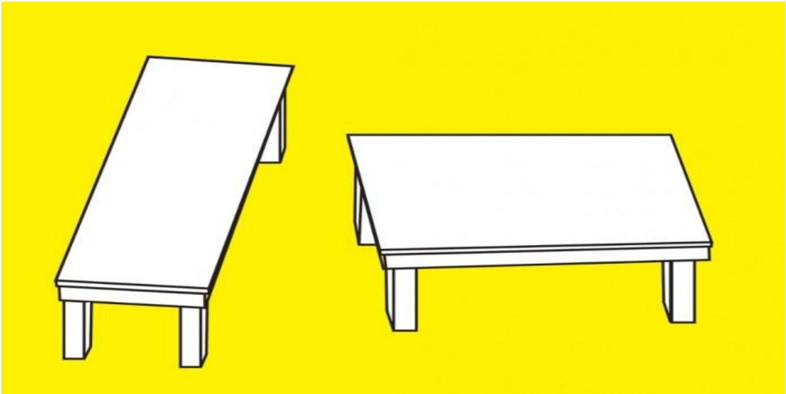


Figure 1.

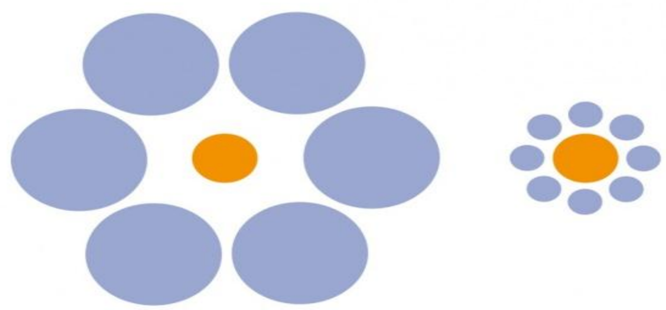


Figure 2.

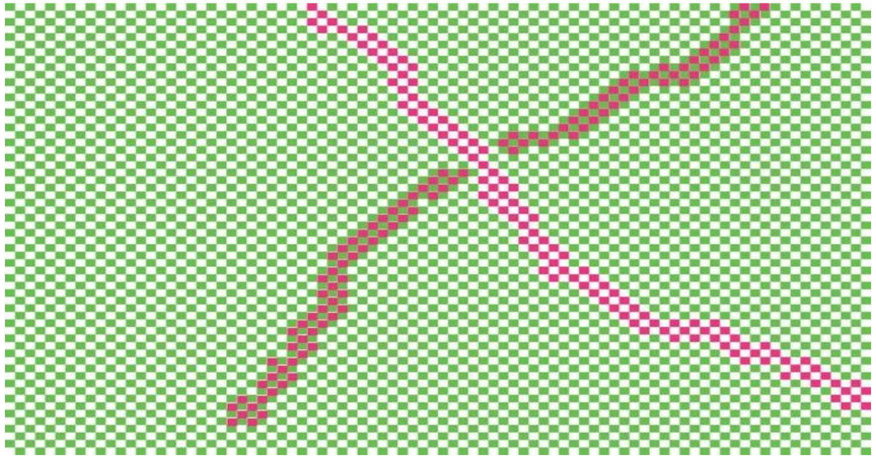


Figure 3.

Figure 1: The two tables have same length.
Figure 2: The two orange circles have the same radius.
Figure 3. The two lines of pixels which appear pink and red are of the exact same shade.
Source: Eye Benders by Clive Gifford.

False Positive Paradox: *Testing For HIV*

Here **positive** test MEANS test says “**there is an infection**”.

Prevalence rate : About **0.01 % men are infected with HIV.**

Sensitivity: If an ***infected*** man is tested, there is a **99.9%** chance the test result is ***positive***.

Specificity: If a man is ***not infected***, there is a **99.99%** chance he will test ***negative***.

What is the chance that a man who tests positive is infected with the HIV virus?

False Positive Paradox - Answer

$$Prob(Infected|Test\ Positive = \frac{Prob(Infected\ AND\ Test\ Positive)}{Prob(Test\ Positive)} =$$

$$\frac{Prob(Test+ve|Infected)\ \times Prob(Infected)}{prob(Test+ve|infected)\times Prob(Infected)+Prob(Test+ve\ |NOT\ Infected)\ \times Prob(NOT\ Infected)} =$$

$$\frac{00.0\%\ \times 0.01\%}{99.9\%\ \times 0.01\%+0.01\%\times 99.99\%} \approx \frac{0.01\%}{0.01\%+0.01\%} = \mathbf{\frac{1}{2}}$$

Probability(Infected | test is positive)

= 49.98% ≈50%.

This is far lower than what most people guess.

It is easy to forget that the incidence rate of this disease is very low.

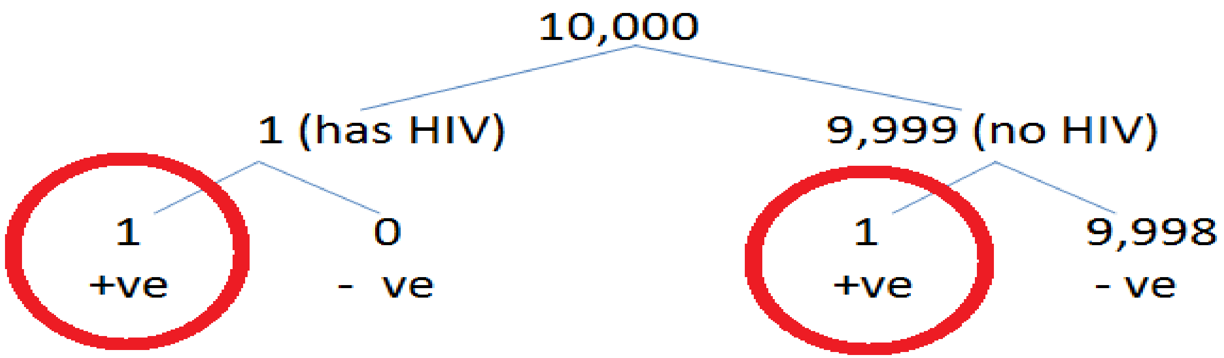
*Both the **test accuracy**, and the **incidence rate** will influence the probability of a positive test result.*

False Positive Paradox – Answer (without tears)

Assume **10,000** men.

Of these **10,000** men, **1** is infected. He will test positive.

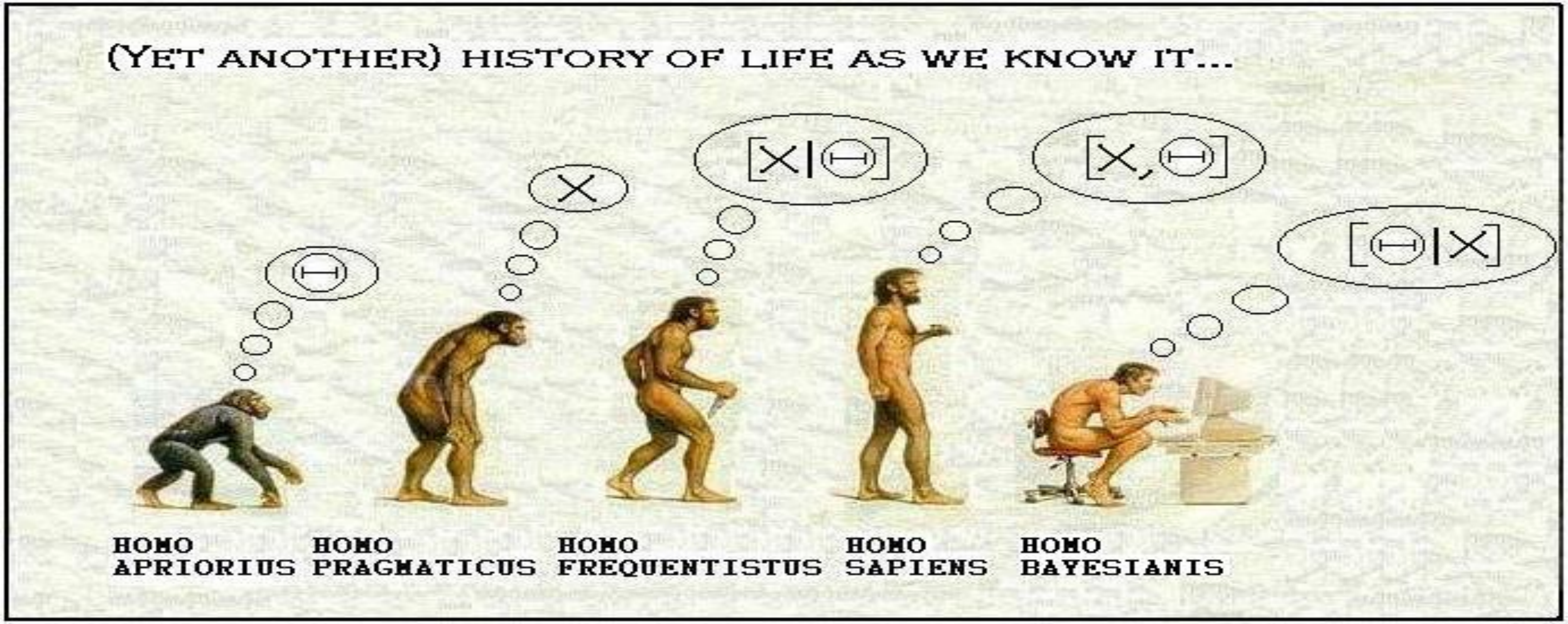
Of the **9,999** men, **1** tests positive (***9999*** × ***0.0001 = 0.9999***)



Two men test positive, of which only one has HIV.

So if he tests positive, there is only a **50%** chance a man has HIV.

We Are All Bayesians Now!



Explanation Of Cartoon And Conclusion

Θ - Will I be attacked?
X - I heard a noise nearby

- **A priori** - Oh no! someone might attack me!!
- **Pragmatic** - What's that noise? Prepare for an attack!
- **Frequentist** - How often did I hear a noise before I was attacked?
- **Sapiens** - Thinks about how often he hears noises and how often he was attacked, but does not have the intuitive or instinctive ability to combine the ideas, as we saw in this talk….
- **Bayesian** - Answers the question, ***if I hear a noise, what is the chance of me being attacked?***

So be aware that our brains do not intuitively give Bayesian results.

You may not need a computer but if you find yourself wrestling with a problem like the one above, at least reach for a pencil and paper.

Conclusion

In particular, watch out for

1) Omitted variables that are confounded - If they were obvious they wouldn't have been omitted.

2) Datasets that are too small - check the significance numbers

3) Interesting results get reported, non interesting ones buried - ask to the analyses that your staff didn't show you.

4) Watch out for sloppy methods - think of your grumpiest professor. Would he be convinced?

5) Finally "Think!"