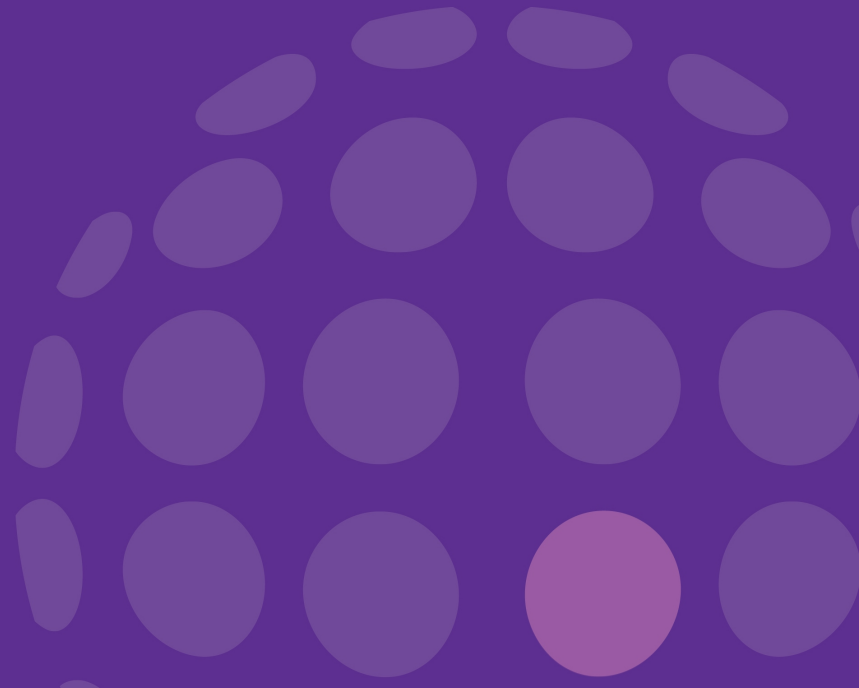




STATISTICAL MISINTERPRETATIONS – AN INTRODUCTION

JOHN SALTER

WEDNESDAY 17 OCTOBER 2012



The views and thoughts contained within this presentation are mine and do not represent the views of Cerafor nor any organisation I have volunteered for nor any client of Cerafor.

DISCLAIMER

WHAT DO I MEAN?

Research has shown that adults have intuitions about probability and statistics that, in many cases, are at odds with accepted theory.

Clifford, K.

Definition of Statistics: The science of producing unreliable facts from reliable figures.

Evan Esar

If your experiment needs statistics, you ought to have done a better experiment.

Ernest Rutherford

Results of the study showed that even students with formal instruction in statistics [at the undergraduate and graduate levels] continue to demonstrate misconceptions .

Hirsch, L. S. and O'Donnell, A. M.

PERCENTAGES: INCREASE VS. DECREASE

Kindle 'swindle': It's 70% cheaper in U.S.



Costly: The Kindle Touch

AMAZON launched its touchscreen Kindles yesterday – but with prices more than 70 per cent higher than in the United States.

The cost gap provided yet more evidence of how international brands appear to see Britain as an easy target for inflated prices.

The Touch eReaders, which are more advanced versions of the device that allows users to read books in an electronic format, come in two versions and have been popular on the other side of the Atlantic where they launched six months ago.

The two models are the Touch, which connects to the internet via wi-fi, and the more sophisticated Touch 3G, which allows users to download on the go from mobile phone networks.

In the U.S., the Kindle Touch can be bought for \$99.99, which equates to £65.72 with sales tax.

However, the UK price is £109, which amounts to a mark-up of £43.28 or 65.8 per cent. The difference on the 3G version is even greater. The British price of £169 equates to a rise of 72.3 per cent on the U.S. price of \$149.99, which converts to £98.05 including sales tax.

Adam Scorer, of the UK's official customer body, Consumer Focus, challenged the huge mark-up.

He said: 'While there may be additional costs for selling the new Kindle in the UK, consumers are likely to question whether such a high mark-up is warranted.'

Amazon claimed the prices were cheap in the U.S. because there are various retail subsidies available. It said without special offers that are available in every city in the U.S., the Kindle Touch would come in at £91.48 including tax and the Touch 3G at £124.39. But even at these prices the UK consumers would still be paying 35 per cent more. It also blamed the cost of doing business in Britain for higher prices here.

However, the company is a virtual retailer and so does not have to pay for stores, retail staff and high street rents.

Similarly, the headquarters of its UK business is in Luxembourg, which allows it to minimise tax.

It recently emerged that the firm, which is the UK's biggest online retailer, paid no corporation tax on the profits made on sales of £3.3 billion here last year.

A spokesman said: 'There are different, and higher, taxes in Europe, as well as higher transport and other operating costs, but we are always working to try and lower them.'

'Orders from customers worldwide for our new line have exceeded our expectations.

'We know that UK customers are eager to get their hands on these new Kindles.'

April 21, 2012; Daily Mail (UK)

£98.05 to £169 is a 72.3% increase ('more expensive')

£169 to £98.05 is a 42.0% decrease ('cheaper')

STATISTICAL SIGNIFICANCE AND CLINICAL SIGNIFICANCE

Definition

Statistical Significance is a statistical assessment of whether observations reflect a pattern, rather than just chance; the fundamental challenge being that any partial picture is subject to observational error.

A result is deemed 'statistically significant' if it is unlikely to have occurred by chance, and hence provides enough evidence to reject the hypothesis of no effect.

Clinical Significance answers the question, how effective is the intervention or treatment, or how much change does the treatment cause? In terms of testing clinical treatments, practical significance optimally yields quantified information about the importance of a finding

STATISTICAL SIGNIFICANCE AND CLINICAL SIGNIFICANCE

Statistical Significance vs. Clinical Significance

The ANOVA Procedure

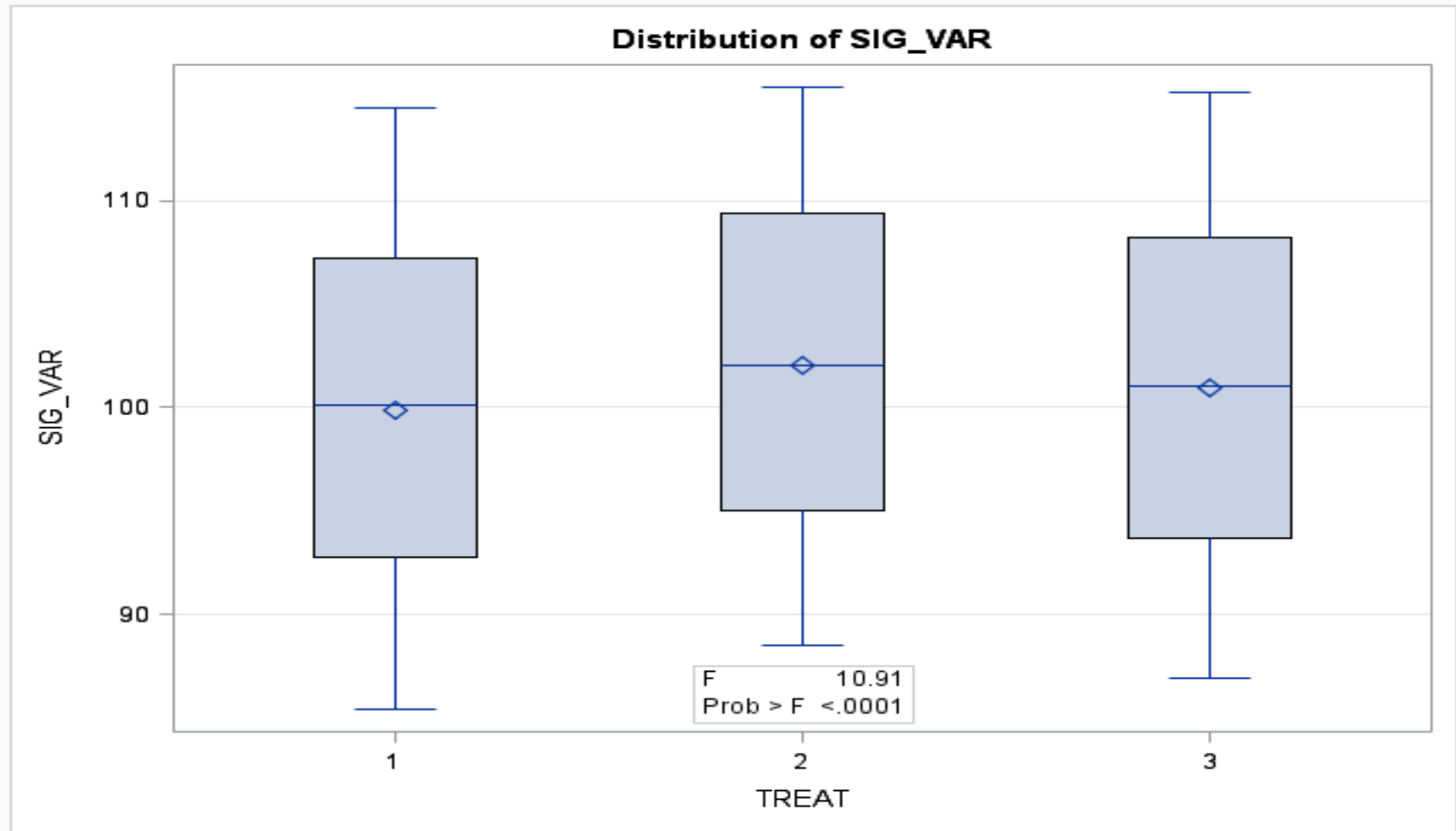
Dependent Variable: SIG_VAR

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	1467.3812	733.6906	10.91	<.0001
Error	1797	120862.8415	67.2581		
Corrected Total	1799	122330.2227			

R-Square	Coeff Var	Root MSE	SIG_VAR Mean
0.011995	8.124282	8.201105	100.9456

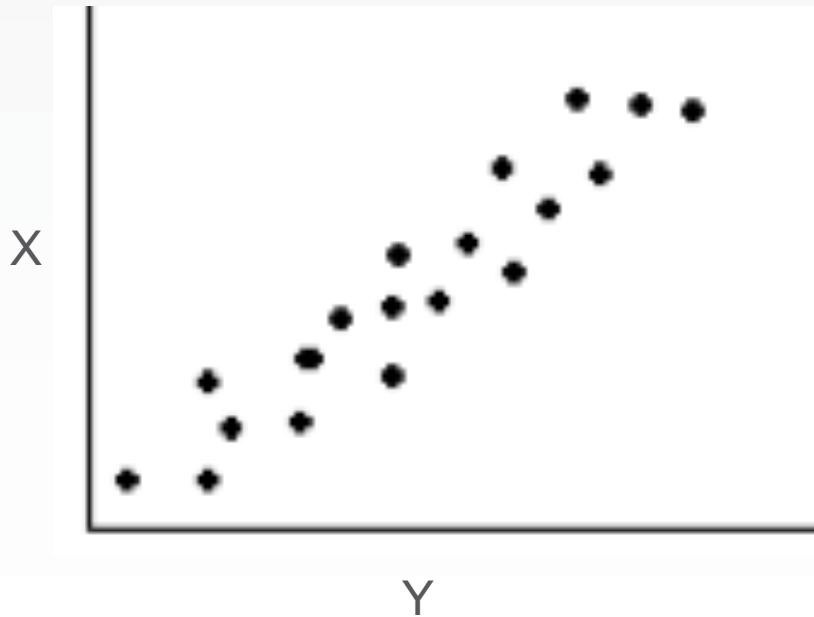
Source	DF	Anova SS	Mean Square	F Value	Pr > F
TREAT	2	1467.381156	733.690578	10.91	<.0001

STATISTICAL SIGNIFICANCE AND CLINICAL SIGNIFICANCE



CORRELATION AND CAUSALITY

- A statistical term that indicates whether two variables 'move' together



X causes Y

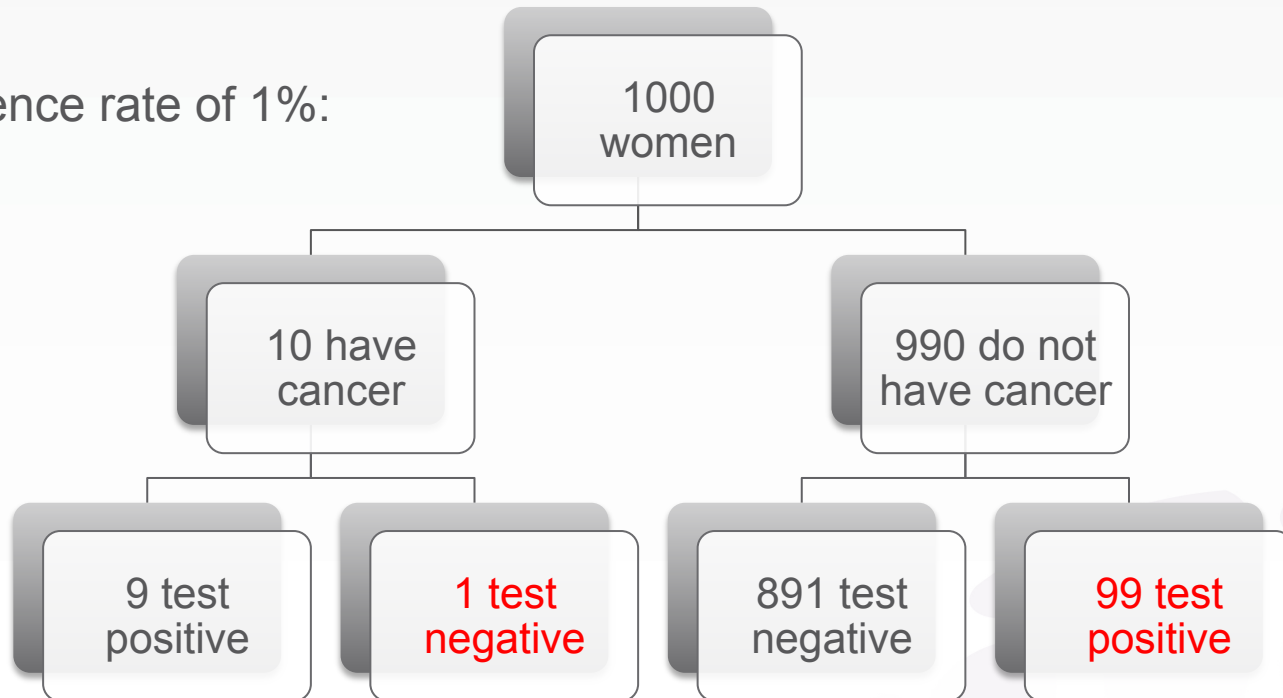
Y causes X

Z causes X and Y

??? Coincidence ???

PREVALENCE OF A DISEASE

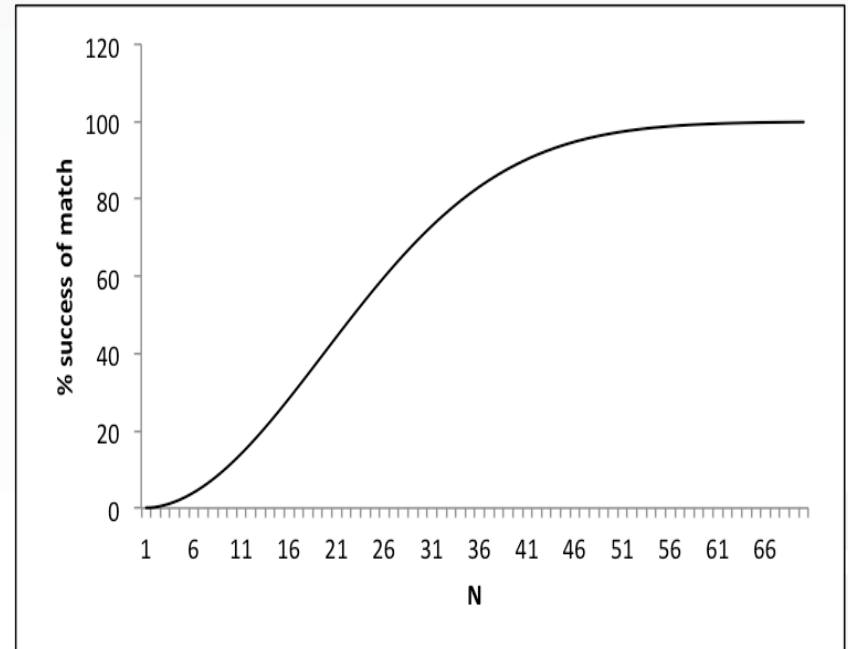
- Consider a screening test for cancer that is 90% accurate
- If you are screened and test POSITIVE, what are the chances you have cancer?
- Incidence rate of 1%:



HAPPY BIRTHDAY



- How many people are needed in a sample in order to have a >50/50 chance of at least two of them sharing a birthday?



DEAL OR NO DEAL (EXPEDITED)



- You are on a game show, and you are given the choice of three doors:
- Behind one door is a car; behind the others, goats.
- You pick a door, say No. 1 [but the door is not opened], and the host, who knows what is behind every door, opens another door, say No. 3, which has a goat. He then says to you, "Do you want to pick door No. 2?"
- Is it to your advantage to switch your choice?

DEAL OR NO DEAL (EXPEDITED)

Consider the discrete random variables, all taking values in the set of door numbers $\{1,2,3\}$:

C : the number of the door hiding the car,

P : the number of the door selected by the player, and

H : the number of the door opened by the host.

Then, if the player initially selects Door 1, and the host opens Door 3, the probability of winning by switching is:

$$\begin{aligned} P(C = 2|H = 3, P = 1) &= \frac{P(H = 3|C = 2, P = 1)P(C = 2|P = 1)}{\sum_{i=1}^3 P(H = 3|C = i, P = 1)P(C = i|P = 1)} \\ &= \frac{1 \times \frac{1}{3}}{(\frac{1}{2} \times \frac{1}{3}) + (1 \times \frac{1}{3}) + (0 \times \frac{1}{3})} = \frac{2}{3} \end{aligned}$$

as due to the random placement of the car and the independence of the initial choice of the player and the placement of the car:

$$P(C = 1|P = 1) = P(C = 2|P = 1) = P(C = 3|P = 1) = \frac{1}{3}$$

ANY QUESTIONS?

It is the mark of a truly intelligent person to
be moved by statistics.

George Bernard Shaw



cerafor



Cerafor Kent Innovation Centre, Millennium Way, Broadstairs, Kent CT10 2QQ

T. +44 1843 609279 F. +44 1227 200125 E. info@cerafor.com

www.cerafor.com