

# Statistical Errors

**Krikamol Muandet**

Empirical Inference Department  
Max Planck Institute for Intelligent Systems

February 20, 2014

# Tea Talk

## Agenda

- ▶ The goal is to stimulate a fruitful discussion/environment.
- ▶ You are free to choose the topic (technical/howto/vacation/etc).
- ▶ Keep it short (15 mins talk  $\pm$  5 mins discussion).
- ▶ Tea talk is for everyone (undergrad/grad/intern/postdoc/staff/director).

# Tea Talk

## Agenda

- ▶ The goal is to stimulate a fruitful discussion/environment.
- ▶ You are free to choose the topic (technical/howto/vacation/etc).
- ▶ Keep it short (15 mins talk  $\pm$  5 mins discussion).
- ▶ Tea talk is for everyone (undergrad/grad/intern/postdoc/staff/director).

## Issue

- ▶ Each of us tends to give tea talk more frequently.
- ▶ Some of us tends to spend too much time on preparing tea talk.
- ▶ Unable to give a talk? No problem!

# Tea Talk

## Agenda

- ▶ The goal is to stimulate a fruitful discussion/environment.
- ▶ You are free to choose the topic (technical/howto/vacation/etc).
- ▶ Keep it short (15 mins talk  $\pm$  5 mins discussion).
- ▶ Tea talk is for everyone (undergrad/grad/intern/postdoc/staff/director).

## Issue

- ▶ Each of us tends to give tea talk more frequently.
- ▶ Some of us tends to spend too much time on preparing tea talk.
- ▶ Unable to give a talk? No problem!

**Empirical Inference Journal Club (EIJC):** Friday 11am (Philosophy of causation).



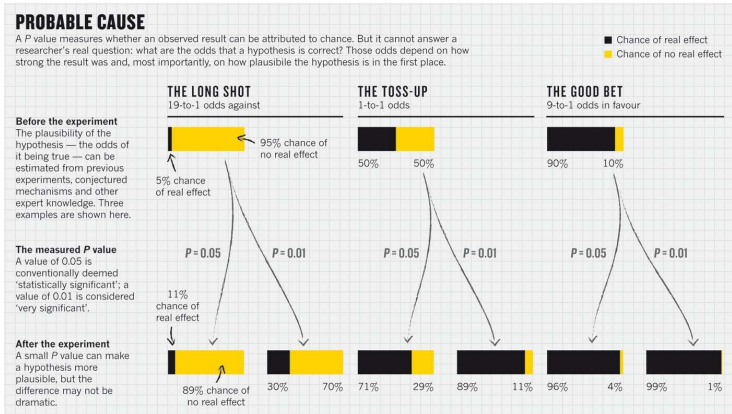
# STATISTICAL ERRORS

*P* values, the ‘gold standard’ of statistical validity, are not as reliable as many scientists assume.

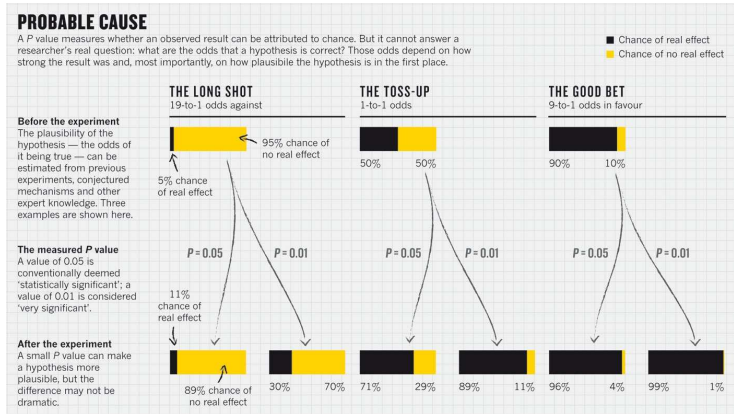
BY REGINA NUZZO

- ▶ Extremists quite literally see the world in black and white ( $p$ -value = 0.01).
- ▶ After replication with extra data,  $p$ -value becomes 0.59!

- ▶ Extremists quite literally see the world in black and white ( $p$ -value = 0.01).
- ▶ After replication with extra data,  $p$ -value becomes 0.59!



- ▶ Extremists quite literally see the world in black and white ( $p$ -value = 0.01).
- ▶ After replication with extra data,  $p$ -value becomes 0.59!



- ▶ The finding has a greater than one in ten chance of being a false alarm.
- ▶ The probability of replicating original result was not 99%, but something closer to 73%—or only 50% if another 'very significant' results are required.





*“The  $P$  value was never meant to be used the way it’s used today.”*

*"Laws handed down from God are no longer handed down from God. They're actually handed down to us by ourselves, through the methodology we adopt"*  
– Steven Goodman (Stanford)

*"Laws handed down from God are no longer handed down from God. They're actually handed down to us by ourselves, through the methodology we adopt"*  
– Steven Goodman (Stanford)

