Module 4 : Fisherian inference

Marie-Abèle Bind

STAT 140

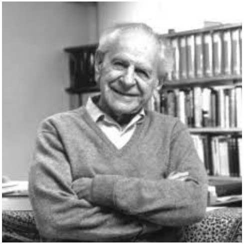
Spring 2026

Quote #1

"Science must begin with myths, and with the criticism of myths."

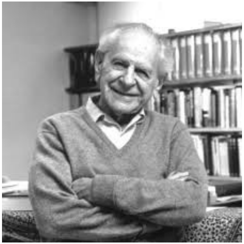
Quote #1

"Science must begin with myths, and with the criticism of myths."



Quote #1

"Science must begin with myths, and with the criticism of myths."
Popper, 1962

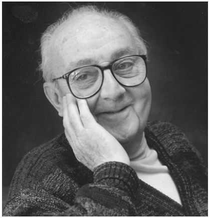


Quote #2

"All models are wrong but some are useful."

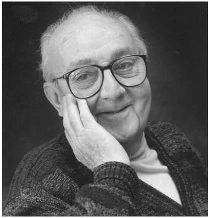
Quote #2

"All models are wrong but some are useful."



Quote #2

"All models are wrong but some are useful."
Box, 1979



Quote #3

"Truth is much too complicated to allow anything but approximations."

Quote #3

"Truth is much too complicated to allow anything but approximations."

Quote #3

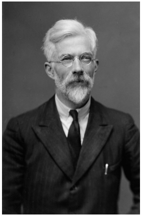
"Truth is much too complicated to allow anything but approximations."
von Neumann, 1947

A Popperian perspective in defense of ...

- "Science must begin with myths, and with the criticism of myths."

A Popperian perspective in defense of ...

- "Science must begin with myths, and with the criticism of myths."
- A p-value, as described by Fisher (1925) in the context of a randomized experiment is a mathematical formalization of the Popperian perspective : a stochastic proof by contradiction of the plausibility of an empirical hypothesis when confronted with observed data.



Extreme opposition

- Desires to prohibit the reporting of p-values to help assess the compatibility of proposed models with observed data (Traffimow and Marks, 2015).


nature
International journal of science

Social Selection | Published: 26 February 2015

Psychology journal bans P values

Chris Woolston

Potential outcomes (simplest situation)

- Consider N units, indexed by i, in an experiment randomized to a binary treatment W, either active (i.e., $W_i=1$) or control (i.e., $W_i=0$).

Potential outcomes (simplest situation)

- Consider N units, indexed by i, in an experiment randomized to a binary treatment W, either active (i.e., W_i=1) or control (i.e., W_i=0).
- An outcome variable Y is measured after exposure to the treatment.

Potential outcomes (simplest situation)

- Consider N units, indexed by i, in an experiment randomized to a binary treatment W, either active (i.e., $W_i=1$) or control (i.e., $W_i=0$).
- An outcome variable Y is measured after exposure to the treatment.
- No hidden treatment and no interference between units (SUTVA). $Y_i(W_i=w)$ is a function of unit i and treatment w.

Potential outcomes (simplest situation)

- Consider N units, indexed by i, in an experiment randomized to a binary treatment W, either active (i.e., W_i=1) or control (i.e., W_i=0).
- An outcome variable Y is measured after exposure to the treatment.
- No hidden treatment and no interference between units (SUTVA). $Y_i(W_i=w)$ is a function of unit i and treatment w.
- For each unit i=1,..., N :
  - $Y_i(W_i=1)$=value of Y when i is exposed to the active treatment
  - $Y_i(W_i=0)$=value of Y when i is exposed to the control treatment.

Potential outcomes (simplest situation)

- Consider N units, indexed by i, in an experiment randomized to a binary treatment W, either active (i.e., W_i=1) or control (i.e., W_i=0).
- An outcome variable Y is measured after exposure to the treatment.
- No hidden treatment and no interference between units (SUTVA). $Y_i(W_i$=w) is a function of unit i and treatment w.
- For each unit i=1,..., N :
  - $Y_i(W_i$=1)=value of Y when i is exposed to the active treatment
  - $Y_i(W_i$=0)=value of Y when i is exposed to the control treatment.
- Because of SUTVA, the science table simplifies to 2 columns and N rows.

Science and observed tables

- Science table :

i	$Y_i(W_i=0)$	$Y_i(W_i=1)$
1	$Y_1(W_1=0)$	$Y_1(W_1=1)$
...	...	...
N	$Y_N(W_N=0)$	$Y_N(W_N=1)$

Science and observed tables

- Science table :

i	$Y_i(W_i{=}0)$	$Y_i(W_i{=}1)$
1	$Y_1(W_1{=}0)$	$Y_1(W_1{=}1)$
...	...	...
N	$Y_N(W_N{=}0)$	$Y_N(W_N{=}1)$

- Observed data table :

i	W_i^{obs}	$Y_i(W_i{=}0)$	$Y_i(W_i{=}1)$
1	0	Y_1^{obs}	?
...	...	...	...
N	1	?	Y_N^{obs}

Science and observed tables

- Science table :

i	$Y_i(W_i{=}0)$	$Y_i(W_i{=}1)$
1	$Y_1(W_1{=}0)$	$Y_1(W_1{=}1)$
...	...	...
N	$Y_N(W_N{=}0)$	$Y_N(W_N{=}1)$

- Observed data table :

i	W_i^{obs}	$Y_i(W_i{=}0)$	$Y_i(W_i{=}1)$
1	0	Y_1^{obs}	?
...	...	...	...
N	1	?	Y_N^{obs}

$$Y_i^{obs} = W_i^{obs} Y_i(W_i = 1) + (1 - W_i^{obs}) Y_i(W_i = 0)$$

Stochastic proof by contradiction

- We follow three steps to assess the plausibility of a Fisher sharp null hypothesis (H_0), *the myth*.

Stochastic proof by contradiction

- We follow three steps to assess the plausibility of a Fisher sharp null hypothesis (H_0), *the myth*.

$$\underline{H_0} : \quad \forall i \in \{1, ..., N\} \quad \mathrm{Y}_i(\mathrm{W}_i{=}1) = \mathrm{Y}_i(\mathrm{W}_i{=}0)$$

Stochastic proof by contradiction

- We follow three steps to assess the plausibility of a Fisher sharp null hypothesis (H_0), *the myth*.
  1. Guided by scientific considerations, choose an appropriate test statistic, T, and define more extreme (e.g., which group is expected to have a better outcome).

Stochastic proof by contradiction

- We follow three steps to assess the plausibility of a Fisher sharp null hypothesis (H_0), *the myth*.
  1. Guided by scientific considerations, choose an appropriate test statistic, T, and define more extreme (e.g., which group is expected to have a better outcome).

$$\underline{\text{Example of T}}: \quad T(W, Y^{obs}) = \frac{1}{N_1} \sum_{i:W_i=1} Y_i^{obs} - \frac{1}{N_0} \sum_{i:W_i=0} Y_i^{obs}$$

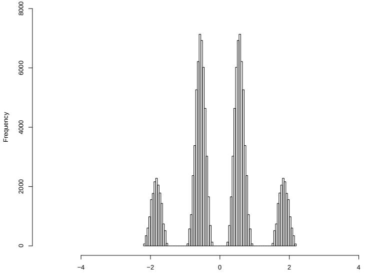
- Statistic : function of the observed data.
- Sensitive to expected departures from the null hypothesis.
- Example of more extreme : $T(W, Y^{obs}) > T^{obs}$

Stochastic proof by contradiction

- We follow three steps to assess the plausibility of a Fisher sharp null hypothesis (H_0), *the myth*.
  1. Guided by scientific considerations, choose an appropriate test statistic, T, and define more extreme (e.g., which group is expected to have a better outcome).
  2. Assuming H_0, calculate the value of T for all possible randomized allocations to obtain the null randomization distribution of T.

Stochastic proof by contradiction

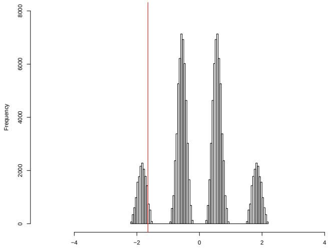
- We follow three steps to assess the plausibility of a Fisher sharp null hypothesis (H_0), *the myth*.
  1. Guided by scientific considerations, choose an appropriate test statistic, T, and define more extreme (e.g., which group is expected to have a better outcome).
  2. Assuming H_0, calculate the value of T for all possible randomized allocations to obtain the null randomization distribution of T.



Example of a null randomization distribution

Stochastic proof by contradiction

- We follow three steps to assess the plausibility of a Fisher sharp null hypothesis (H_0), *the myth*.
  1. Guided by scientific considerations, choose an appropriate test statistic, T, and define more extreme (e.g., which group is expected to have a better outcome).
  2. Assuming H_0, calculate the value of T for all possible randomized allocations to obtain the null randomization distribution of T.



  3. Locate the observed value of the test statistic, T^{obs}, in the null randomization distribution constructed in Step 2.

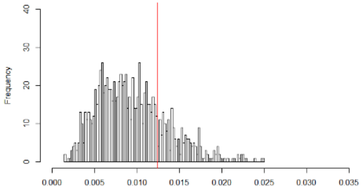
Stochastic proof by contradiction

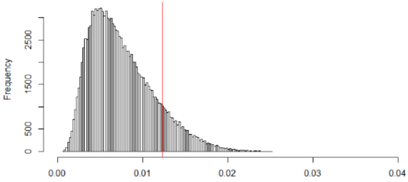
- We follow three steps to assess the plausibility of a Fisher sharp null hypothesis (H_0), *the myth*.
  1. Guided by scientific considerations, choose an appropriate test statistic, T, and define more extreme (e.g., which group is expected to have a better outcome).
  2. Assuming H_0, calculate the value of T for all possible randomized allocations to obtain the null randomization distribution of T.
  3. Locate the observed value of the test statistic, T^{obs}, in the null randomization distribution constructed in Step 2.

The Fisher-exact p-value corresponds to the proportion of values of the test statistic that are as extreme or more extreme than the observed value of that test statistic.

If computationally impossible, approximate ?

For computational reasons, when N is large, the null randomization distribution of the test statistic may be approximated by randomly drawing fewer permutations than $N_{randomizations}$.

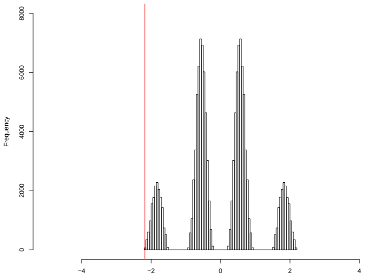


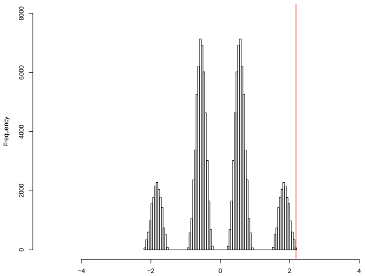


Approximated p-value (left)$=\frac{1}{1{,}000}\sum_{r=1}^{1{,}000} \mathbb{1}(T^{(r)} \geq T^{obs})$=0.20

Fisher-exact p-value (right)$=\frac{1}{N_{randomizations}}\sum_{r=1}^{N_{randomizations}} \mathbb{1}(T^{(r)} \geq T^{obs})$=0.07

Smallest and largest attainable p-values?



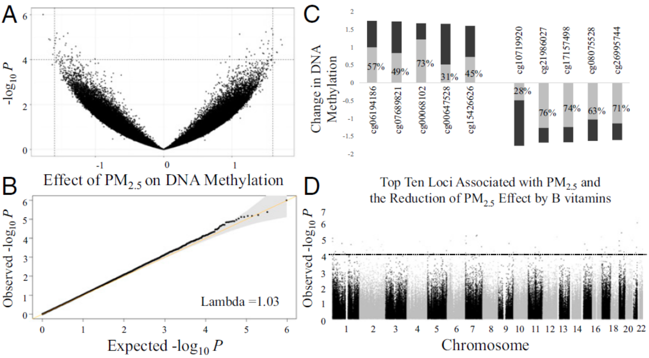


Fisher-exact p-value$=\frac{1}{N_{randomizations}} \sum_{r=1}^{N_{randomizations}} \mathbb{1}(T^{(r)} \geq T^{obs})$

Smallest p-value is important

B vitamins attenuate the epigenetic effects of ambient fine particles in a pilot human intervention trial

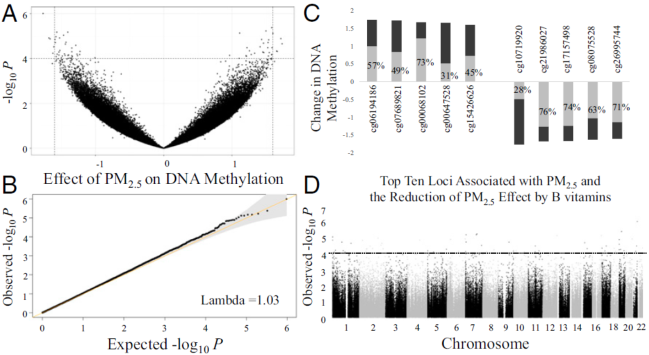
PNAS March 28, 2017 114 (13) 3503-3508; first published March 13, 2017



Smallest p-value is important

B vitamins attenuate the epigenetic effects of ambient fine particles in a pilot human intervention trial

PNAS March 28, 2017 114 (13) 3503-3508; first published March 13, 2017



. $\frac{1}{2^{10}} \approx 0.00098$