

## P5.2 Statistics for Medicine

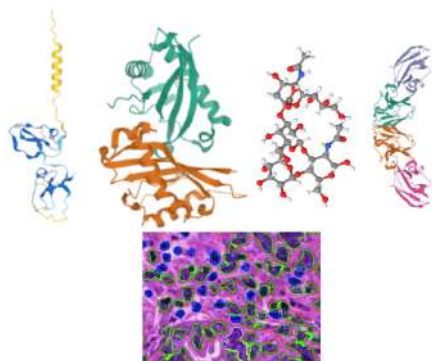
Massimo Borelli

Master of Advanced Studies in Medical Physics



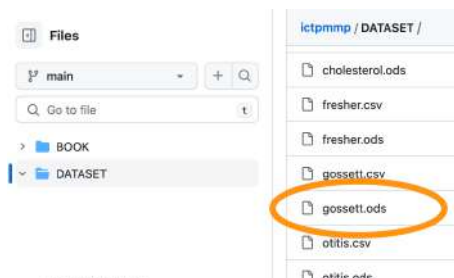
Massimo Borelli P5.2 Statistics for Medicine

## Our goal /1



Massimo Borelli P5.2 Statistics for Medicine

## Today



Massimo Borelli P5.2 Statistics for Medicine

## Recap /1

1	To start	5
1.1	What are we talking about	5
1.1.1	Shifting Statistics from Physics to Medicine	6
1.2	Which is the 'best' software for medical statistics?	6
1.2.1	The R language	6
1.2.2	The R language user graphical interfaces	8
1.2.3	JASP	8
1.3	Exercises	9
2	Data Presentation	11
2.1	Background	11
2.2	Descriptive Statistics in JASP	11
2.2.1	Numerical summaries	13
2.2.2	A picture is worth a thousand words	15
2.3	Which are 'the best' Descriptives?	17
2.4	Exercises	18

Massimo Borelli P5.2 Statistics for Medicine

## Our goal /2

### In vivo dosimetry and shielding disk alignment verification by EBT3 GAFCHROMIC film in breast IOERT treatment

Mara Severgnini,<sup>1a</sup> Mario de Denaro,<sup>1</sup> Marina Bortul,<sup>2</sup> Cristiana Vidali,<sup>3</sup>

TABLE 1. Results for the 37 patients treated in this study. In the first six patients, the dimensions of GAFCHROMIC film were smaller than the disk's and it is not possible to estimate the area of the radiation field that escapes outside the shield; the data is not available (n.a.).

N.Rt	Energy MeV	Collimator diam cm	Collimator FLD cm	Distance Setup vs. target cm	Area Outside Shielding cm <sup>2</sup>
1	9	n.a.	n.a.	n.a.	n.a.
2	9	n.a.	n.a.	n.a.	n.a.
3	9	n.a.	n.a.	n.a.	n.a.
4	9	n.a.	n.a.	n.a.	n.a.
5	6	n.a.	n.a.	n.a.	n.a.
6	9	n.a.	n.a.	n.a.	n.a.
7	9	9	9	1.2	8.2
8	9	9	9	1	4.9
9	6	9	9	21.2	8.7
10	9	9	9	1	4.9
11	9	9	9	21.2	8.7
12	6	9	9	1	4.9

Massimo Borelli P5.2 Statistics for Medicine

## the T test: the basics

Massimo Borelli P5.2 Statistics for Medicine

## Recap /2

3	Probability in medicine	19
3.1	Brief recalls on random variables	19
3.2	Commonly used random variables	20
3.2.1	The Normal Distribution	20
3.2.2	The Lognormal Distribution	23
3.2.3	The Binomial Distribution	24
3.2.4	The Poisson Distribution	25
3.3	Evaluating odds and risks: Bayes theorem	26
3.3.1	Bayes theorem	28
3.3.2	The Bayes factor	29
3.4	Sample and population: approaching inference	30
3.5	Mismatching variability with reliability	32
3.6	Exercises	35

Massimo Borelli P5.2 Statistics for Medicine

## Today

the univariate inferential analysis



Massimo Borelli P5.2 Statistics for Medicine



Massimo Borelli P5.2 Statistics for Medicine

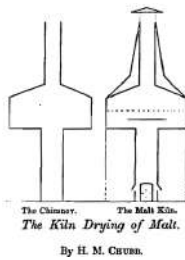
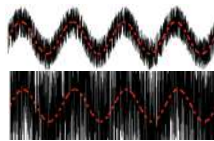


Figure: William Sealy Gosset

Not Kiln-Dried	Kiln-Dried	Difference
1903	2009	+106
1935	1915	-20
1910	2011	+101
2496	2463	-33
2108	2180	+72
1961	1925	-36
2060	2122	+62
1444	1482	+38
1612	1542	-70
1316	1443	+127
1511	1535	+24

Not Kiln-Dried	Kiln-Dried	Difference
1903	2009	+106
1935	1915	-20
1910	2011	+101
2496	2463	-33
2108	2180	+72
1961	1925	-36
2060	2122	+62
1444	1482	+38
1612	1542	-70
1316	1443	+127
1511	1535	+24



• Detecting a signal from noise

$$t = \frac{m - \mu}{s/\sqrt{n}}$$

difference

Valid	11
Mean	33.727
Std. Deviation	66.171
Std. Error of Mean	19.951

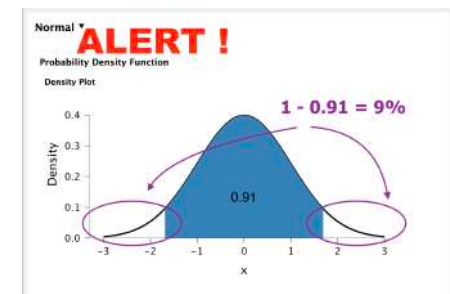
## Gosset discoveries /2

$$t = \frac{m - \mu}{s/\sqrt{n}}$$

- (independency) in a random sample from a gaussian distribution  $N(\mu, \sigma)$ , estimating the sample mean  $m$  do not convey any information in estimating the sample standard deviation  $s$ , and vice versa.
- (a novel random variable) the random variable  $t = \frac{m - \mu}{s/\sqrt{n}}$  possesses an explicit density function, which is not a gaussian, but can be numerically computed.

$$t = \frac{33.727 - 0}{66.171/\sqrt{11}} = \frac{33.727}{19.951} \approx 1.690$$

## Gosset discoveries /1



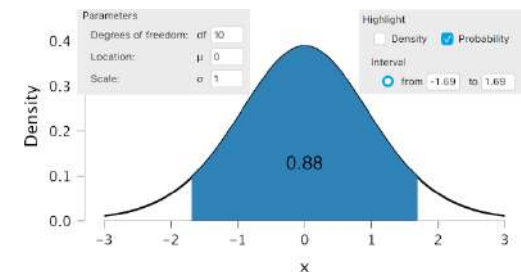
normal distribution does not work!

## BIOMETRIKA.

## THE PROBABLE ERROR OF A MEAN.

BY STUDENT.

## JASP: Scaled Shifted Student's t

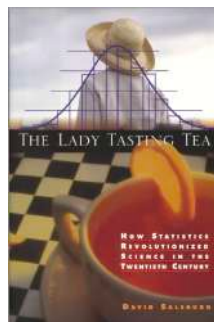




## JASP: Classical One Sample T-Test

Table: One Sample T-Test

	t	df	p
difference	1.690	10	0.122



## JASP: Classical One Sample T-Test

Table: One Sample T-Test

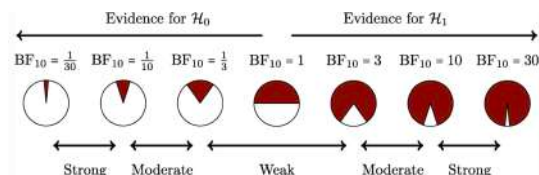
	t	df	p
difference	1.690	10	0.122

$$BF_{10} = \frac{P(D|M_1)}{P(D|M_0)} = 0.885$$

BF <sub>10</sub>	log <sub>10</sub> BF <sub>10</sub>	Evidence	In favour of
>100	>4.6	Decisive	Alternative hypothesis
30 to 100	3.4 to 4.6	Very strong	Alternative hypothesis
10 to 30	2.3 to 3.4	Strong	Alternative hypothesis
3 to 10	1.1 to 2.3	Moderate	Alternative hypothesis
1 to 3	0 to 1.1	Anecdotal	Alternative hypothesis
1	0	No evidence	Neither
1 to 0.33	0 to -1.1	Anecdotal	Null hypothesis
0.33 to 0.1	-1.1 to -2.3	Moderate	Null hypothesis
0.1 to 0.033	-2.3 to -3.4	Strong	Null hypothesis
0.033 to 0.01	-3.4 to -4.6	Very strong	Null hypothesis
<0.01	<-4.6	Decisive	Null hypothesis

However, these are merely a simplified heuristic for interpreting Bayes factors, but that the Bayes factor really is a continuous metric of evidence.

## In conclusion



## Ronald Fisher's idea on significance level

- 1 The conventional significance level of 5%
- 2 The freedom to choose the significance level
- 3 significance level and sample size impact on the test power
- 4 statistical or clinical significance?
- 5 Absence of evidence, or evidence of absence?

## JASP: Classical One Sample T-Test

Table: One Sample T-Test

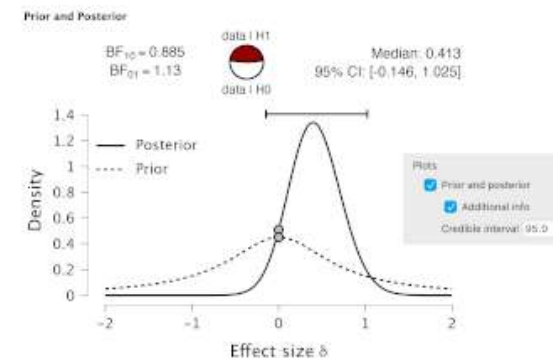
	t	df	p
difference	1.690	10	0.122

Table: Bayesian One Sample T-Test

	BF <sub>10</sub>	error %
difference	0.885	0.004

Table: One Sample T-Test

	t	df	p
difference	1.690	10	0.122



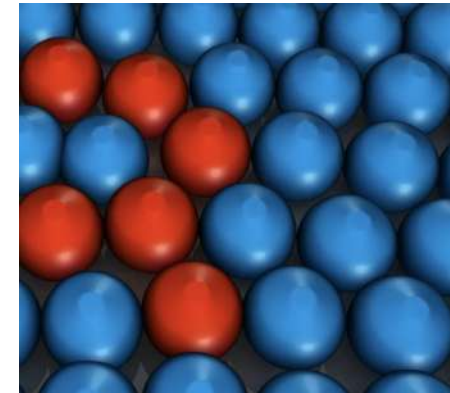
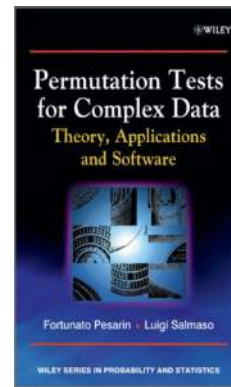
## Statistics for Medicine

Massimo Borelli

Master of Advanced Studies in Medical Physics

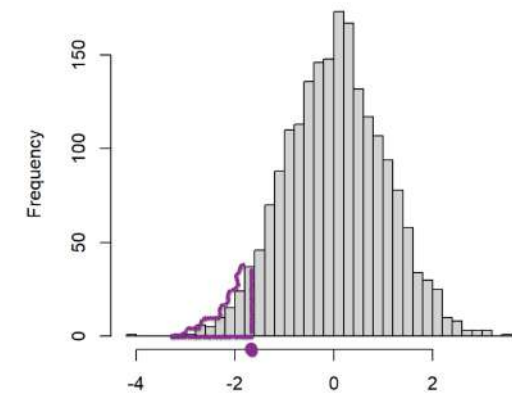
UNIVERSITÀ  
DEGLI STUDI DI TRIESTEThe Abdus Salam  
International Centre for  
Theoretical Physics

L12	Statistics for medicine	1	12	lesson	W
L13	introduction to Monte Carlo simulation	1	12	lab	W



Area	Area	Area	Area	Area	Area
NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA
1.2	1.2	1.2	1.2	1.2	1.2
8.2	8.2	8.2	8.2	8.2	8.2
1	1	1	1	1	1
4.9	4.9	4.9	4.9	4.9	4.9
21.2	21.2	21.2	21.2	21.2	21.2
8.7	8.7	8.7	8.7	8.7	8.7
5.5	5.5	5.5	5.5	5.5	5.5
2.1	2.1	2.1	2.1	2.1	2.1
13.3	13.3	13.3	13.3	13.3	13.3
5.8	5.8	5.8	5.8	5.8	5.8
0	0	0	0	0	0
0	0	0	0	0	0
14.7	14.7	14.7	14.7	14.7	14.7

repetition	red mean	blue mean	test statistic
1	4.3	4.8	...
2	5.1	4.4	...
3	3.6	5.1	...
4	4.7	4.9	...
5	5.1	4.3	...
...	...	...	...
1000000	4.4	4.8	...



## Differences between two groups

the roma dataset

Variable	Level	Counts	Total	Proportion
Histology	benign	171	210	0.814
	malignant	39	210	0.186

- can we exploit logHE4 to predict Histology?
- can we exploit logCA125 to predict Histology?
- can we exploit logCA19-9 to predict Histology?
- can we exploit logCEA to predict Histology?

## Differences between two groups

