

P5.2 Statistics for Medicine

Massimo Borelli

Master of Advanced Studies in Medical Physics



UNIVERSITÀ
DEGLI STUDI DI TRIESTE



The Abdus Salam
International Centre
for Theoretical Physics



1 Brief recalls on random variables

Finite random variables

The otitis dataset

(number of episodes of otitis media in the first two years of life)

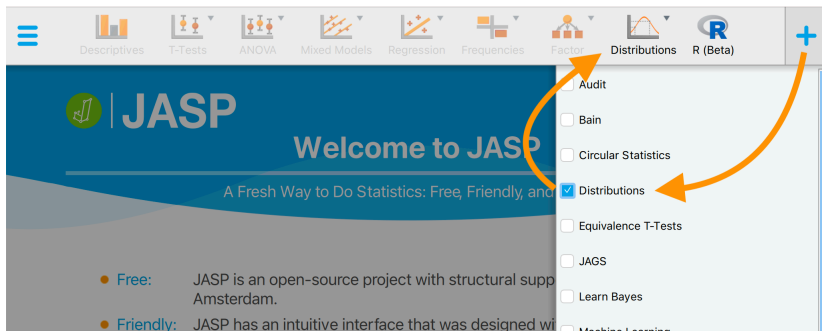
$$\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 \\ 0.129 & 0.264 & 0.271 & 0.185 & 0.095 & 0.039 & 0.017 \end{pmatrix}$$

- **probability mass / discrete density** function

Example (graph)

Draw the barplot with JASP

Commonly used random variables



The screenshot displays the JASP software interface. The top navigation bar includes icons for Descriptives, T-Tests, ANOVA, Mixed Models, Regression, Frequencies, Factor, Distributions, and R (Beta). The 'Distributions' menu is open, showing a list of statistical tests with checkboxes. The 'Distributions' option is selected, indicated by a blue checkmark. Two orange arrows highlight the 'Distributions' menu item and the '+' icon in the top right corner of the navigation bar.

JASP
Welcome to JASP
A Fresh Way to Do Statistics: Free, Friendly, and

- **Free:** JASP is an open-source project with structural support from the University of Amsterdam.
- **Friendly:** JASP has an intuitive interface that was designed with

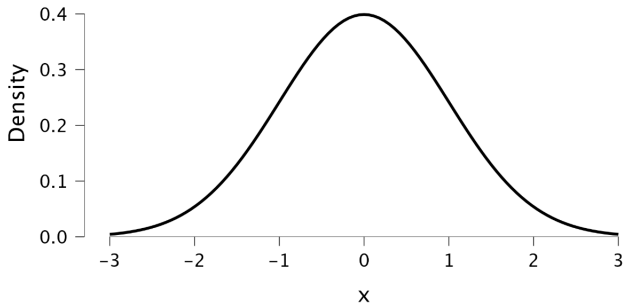
Menu items (checked items are highlighted):

- Audit
- Bain
- Circular Statistics
- Distributions
- Equivalence T-Tests
- JAGS
- Learn Bayes
- Machine Learning

the Normal distribution

Probability Density Function

Density Plot



the Normal distribution

▼ Show Distribution

Parameters μ, σ^2 ▼

Mean: μ 0

Variance: σ^2 1

Display

- Explanatory text
- Parameters, support, and moments
- Probability density function
- Cumulative distribution function
- Quantile function

Options

Range of x from -3 to 3

Highlight

- Density
- Probability

Interval

- from 0 to 1
- from $-\infty$ to 0
- from 0 to ∞

the Normal distribution

Example (B. Rosner, example 5.22)

The cerebral blood flow (CBF) in the general population is, approximately, normally distributed with mean $\mu = 75$ and standard deviation $\sigma = 17$. Which could be the percentage of persons having a $\text{CBF} < 40$?

the Normal distribution

Example (B. Rosner, example 5.23)

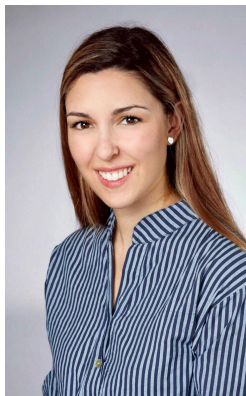
Glaucoma is characterized by intraocular pressure greater than 20 mmHg, while in normal population intraocular pressure X has mean $\mu = 16$ and standard deviation $\sigma = 9$. How much it could be $P(12 \leq X \leq 20)$?

the Normal distribution

Example (B. Rosner, example 5.24)

In adult male, the diastolic pressure is normally distributed with mean $\mu = 80$ and variance $\sigma^2 = 144$. Find the upper and the lower fifth percentile.

the Normal distribution: the QQ plot



S. Najaf
the roma dataset

- Histology
- AgePatient
- Menopause status
- four biomarkers (log transformed):
 - logHE4,
 - logCA125
 - logCA19.9
 - logCEA

the Normal distribution

logHE4
logCA125
logCA19-9
logCEA
AgePatient
Menopause
Histology

Get variable from data set

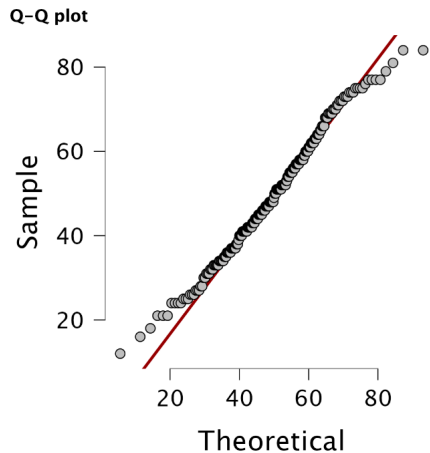
Statistics

- Descriptives
- First observed moments

Plots

- Histogram with bins
- Empirical cumulative distribution

the Normal distribution



caveat: the Normal distribution

Do two dromedaries make a camel? Bernard Rosner

.. linear combination of normal random variables are often of specific concern. It can be shown that any linear combination of normal random variables is itself normally distributed.

Martin Bland:

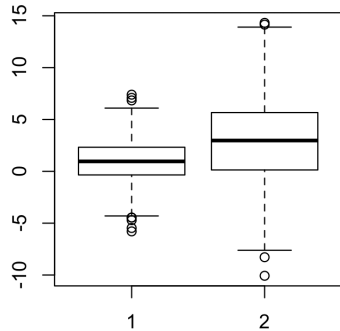
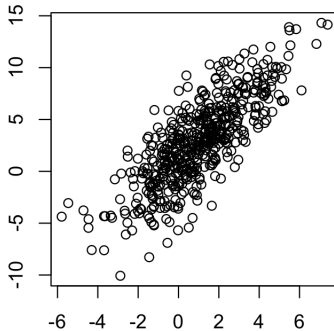
... If we add two variables from Normal distributions together, even with different means and variances, the sum follows a Normal distribution.

caveat: the Normal distribution

Do two dromedaries make a camel?



caveat: the Normal distribution



C. Kowalski. 1973

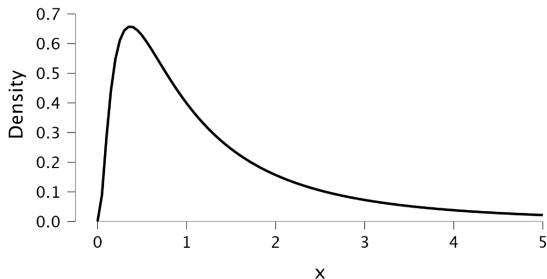
Non-Normal Bivariate Distributions with Normal Marginals

<https://www.tandfonline.com/doi/abs/10.1080/00031305.1973.10479002>

the log-Normal distribution

Probability Density Function

Density Plot



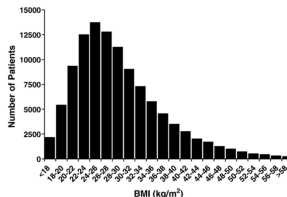
Eckhard Limpert, et al.

Log-normal Distributions across the Sciences: Keys and Clues

<https://academic.oup.com/bioscience/article/51/5/341/243981>

the logNormal distribution

Figure 1



Distribution of BMI values. A histogram of the BMI values (expressed in kilograms per square meter) among the 108 927 hospitalization

Example (Gregg Fonarow et al. - summarizing body mass index)

Suppose that you are required to lead a pilot study concerning radiation dosimetry in 25 obese patients. How do you think you are going to describe the data? Using the mean and the standard deviation, or the median and the quartiles? What are here the difficulties?

the binomial distribution

Free parameter	Fixed parameter
Probability of success: p <input type="text" value="0.186"/>	Number of trials: n <input type="text" value="210"/>
Display	Options
<input type="checkbox"/> Explanatory text	Range of x from <input type="text" value="20"/> to <input type="text" value="60"/>
<input type="checkbox"/> Parameters, support, and moments	Highlight
<input checked="" type="checkbox"/> Probability mass function	<input type="checkbox"/> Mass <input checked="" type="checkbox"/> Cumulative Probability
<input type="checkbox"/> Cumulative distribution function	Interval <input type="text" value="30"/> $\leq X \leq$ <input type="text" value="50"/>

Example (probability)

Suppose that you collect a new sample of 210 women with the same symptoms of those enrolled in roma. Obviously, only by chance you will observe exactly '39' malignancies. Can you compute the probability to observe a number of malignancy between 30 and 50?

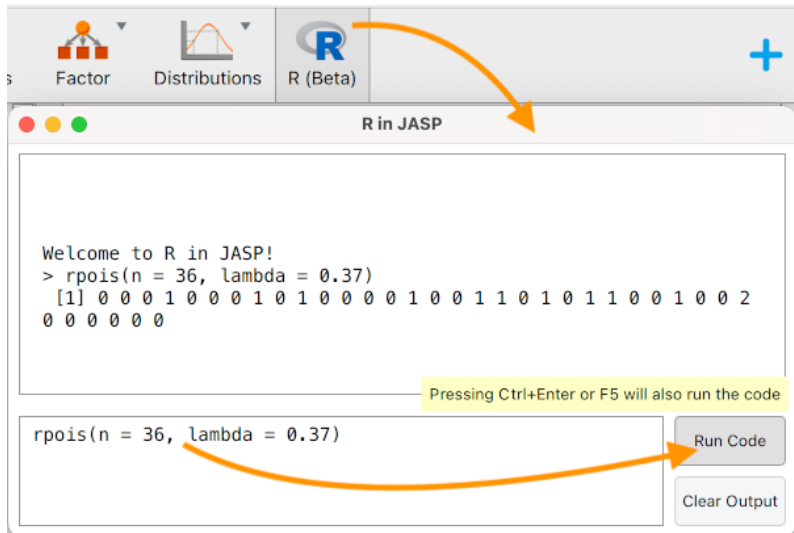
the Poisson distribution

1				1	2
2	3	1	2	1	
	1		2		1
1	1	2		4	1
1		1		3	1
	2	1	1		

Example (probability)

Use JASP to discover in a $\lambda = 0.37$ Poisson distribution how many, in probability, cells could have a value greater or equal than 2.

the Poisson distribution



The screenshot shows the R in JASP interface. At the top, there are three tabs: "Factor", "Distributions", and "R (Beta)". An orange arrow points from the "R (Beta)" tab to the main window. The main window has a title bar "R in JASP" and a large text area containing the following text:

```
Welcome to R in JASP!  
> rpois(n = 36, lambda = 0.37)  
[1] 0 0 0 1 0 0 0 1 0 1 0 0 0 0 1 0 0 1 1 0 1 0 1 1 0 0 1 0 0 2  
0 0 0 0 0 0
```

Below the text area, there is a yellow highlight with the text "Pressing Ctrl+Enter or F5 will also run the code". Below this, there is a code input field containing the code:

```
rpois(n = 36, lambda = 0.37)
```

An orange arrow points from the code input field to the "Run Code" button. Below the "Run Code" button is a "Clear Output" button.