



Master's study

Field of study: Advanced Analytics – Big Data

Author's first name and surname: Bogdan Bojarin

Student's register No.: 75184

**Development of an intelligent system of
document processing and generation based on
machine learning methods**

Master's thesis:

under the scientific supervision of:

dr hab. Michał Ramsza

written in

Institute of Mathematical Economics

Warsaw 2025

Contents

1	Introduction	5
2	Basic things	6
2.1	Compiling L ^A T _E X files	6
2.2	Basic formatting for a text	6
2.3	Fonts and fonts' sizes	6
3	Mathematics	7
3.1	Basic mathematics	7
3.2	Referencing mathematics and other things	7
3.3	Some more mathematical formulas	8
4	Figures and tables	9
5	Bibliography	12
A	Appendix: Some important stuff	13
	List of tables	15
	List of figures	16
	Streszczenie	17

1 Introduction

This template is for the BSc/MSc papers at the Warsaw School of Economics.

2 Basic things

2.1 Compiling L^AT_EX files

The `.tex` file is just a plain text file. It contains the L^AT_EX formatting codes together with the content of a paper. To get a `.pdf` file you have to compile the `.tex` file using a sequence `pdflatex, biblAtex, pdflatex, pdflatex`. This sequence is a default in most editors designed for use with L^AT_EX.

2.2 Basic formatting for a text

Paragraphs are coded by an empty line. That is if you want to start a new paragraph it is enough to leave an empty line and start typing like that:

This is the first paragraph.

This is the next paragraph.

Everything about the paragraph is formatted for you including all indents and spacings. Again, you don't have to take care of it manually.

Basic text formatting, e.g. bold face and italic, is achieved with the following commands: `\textbf{}`, `\textit{}`, `\underline{}`, producing **text**, *text*, text. I suggest not overusing those commands!

Alignment is done through environments `center`, `flushleft` and `\flushright` giving the following examples.

This is centered.

This is aligned to the left.

This is aligned to the right.

In other environments it is possible to use `\centering` to center content of that environment (like in `figure` or `table` environments).

2.3 Fonts and fonts' sizes

You do not change fonts and fonts' sizes! Technically it can be done but I will reject this.

3 Mathematics

This is testing footnotes¹.

3.1 Basic mathematics

There are two types of mathematics inside a \LaTeX document. The first one is the in-line mathematics and the displayed mathematics. The first one looks like this: $F(x) = \int_{-\infty}^x f(\omega)d\omega$ with the code looking like this: `\(F(x) = \int_{-\infty}^x f(\omega) d\omega \)`. The displayed mathematics looks like that

$$F(x) = \int_{-\infty}^x f(\omega)d\omega$$

with the code

```
\[
F(x) = \int_{-\infty}^x f(\omega) d\omega
\]
```

As you can see the same code is formatted differently depending on the type of mathematics.

3.2 Referencing mathematics and other things

To reference mathematics (only displayed formulas) you use the `equation` environment with a `\label{}` within. The reference is done through the `\ref{}` command. The example is

$$F(x) = \int_{-\infty}^x f(\omega)d\omega. \tag{1}$$

To reference the equation you use the `\ref{}` command giving (1). The `\label{}` / `\ref{}` pair works for anything that can be referenced.

¹This is a footnote. We can put some math here $x^2 - f(x) = g(x^2)$ which is not encouraged but sometimes necessary. The other thing we can do is to put here an URL <https://tex.stackexchange.com/questions/249415/set-font-size-for-footnotes>.

3.3 Some more mathematical formulas

\LaTeX is known for producing beautifully typeset mathematical formulas. The above mathematical formulas are relatively simple. Here are slightly more complex formulas. Let A be a matrix

$$A = \left(\begin{bmatrix} 1 & \alpha^2 \\ 2 & \sqrt{\pi} - \log(x - \sin(y)) \end{bmatrix}^2 - \begin{bmatrix} 1 & f(x) \\ 2 & g(y) \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} \right),$$

where

$$f(x) = \begin{cases} \frac{1}{x} & \text{for } x < -\frac{1}{2}, \\ \frac{1}{1+x^2} & \text{for } x \geq -\frac{1}{2} \end{cases}$$

and

$$g(y) = \sin \left(\frac{\mathbf{E}(X)}{\cos(y) + \log(y)} \right), \quad \text{where } X \sim \mathcal{N}(0, \sigma).$$

Note that the above formulas are parts of a sentence. Thus, you still use proper punctuation. In \LaTeX , we can also typeset diagrams of arbitrary complexity. However, this requires another language for defining graphical scenes: TikZ (<https://tikz.org/>).

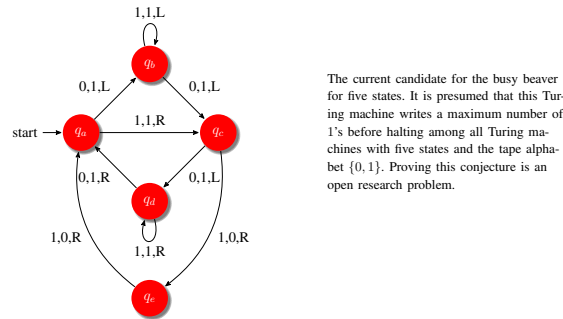


Figure 1: This is a simple diagram of a Turing machine. With TikZ, we can prepare diagrams of any complexity. *Source:* <https://tikz.org/>.

For those of you doing game theory, the TikZ is a great solution for visualizing extensive-form games. For normal-form games, we have a simpler solution. It is very easy to typeset a normal-form game. Below is an example of such a game.

	L	M	H
L	16, 9	3, 13	0, 3
M	21, 1	10, 4	-1, 0
H	9, 0	5, -4	-5, -15

4 Figures and tables

Both figures and tables use the same ideas. To insert a table, you use the `table` environment. The following tables are just examples of what can be automatically generated with the R and Python programming languages.

Table 1: This is an example of a table generated in the R programming language. The script generating the table is `example.R`.

	Values x	Values y	Class
1	-0.12	0.73	Down
2	-1.54	-2	Up
3	-0.64	-0.36	Down
4	-0.96	-0.43	Up
5	0.92	1.72	Down

Table 2: This is another example of a table generated in the R programming language. This table is automatically generated from the linear regression model.

<i>Dependent variable:</i>	
	y
x	1.989*** (0.032)
Constant	1.020*** (0.033)
Observations	1,000
R ²	0.796
Adjusted R ²	0.796
Residual Std. Error	1.029 (df = 998)
F Statistic	3,903.749*** (df = 1; 998)
<i>Note:</i> *p<0.1; **p<0.05; ***p<0.01	

To insert a figure, you need to have a figure. In the `./figs` directory, there are figures generated with the R and Python scripts, and the following is an example of the `figure` environment. Figure 3 is slightly more complex than just a simple figure, but it is useful to have such a template. It is possible to reference subfigures as 3a and 3b.

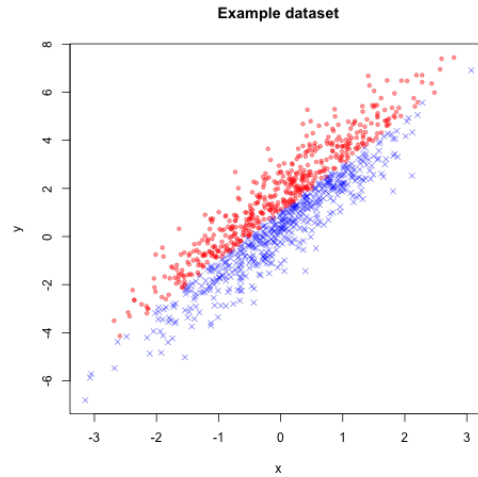
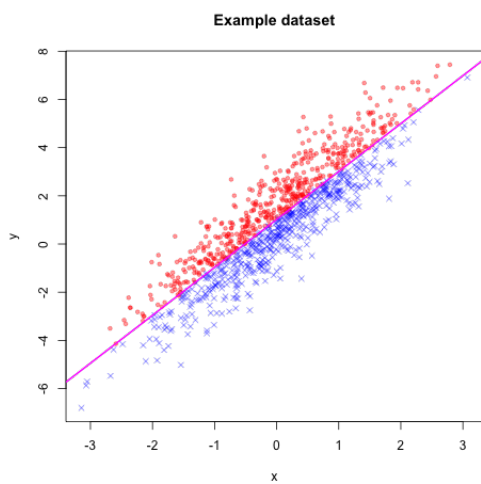


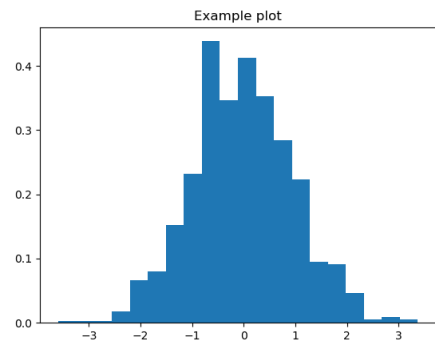
Figure 2: This is an example figure generated in the R programming language.
Source: own calculations.

Table 3: This is another table generated by the Python script `example.py`.
This table looks a little bit different, but it's acceptable.

Class	Values
j	0.953570
M	0.183956
X	1.243109
I	-1.032789
N	0.443236
K	-1.602915
l	-1.273745
l	2.209001
r	0.190158
R	0.873841



(a) This is another visualization done in the R programming language in the script `example.R`. This caption is wrapped at the right width, and the height is being compensated.



(b) This figure was generated in the Python programming language. The script `example.py` creates this figure and the additional table.

Figure 3: This is the main caption and it is below the figures. Both figures were automatically created in scripts. If we want to change the figure, we change the script only. *Source:* own calculations

5 Bibliography

The content for the bibliography is in a different file named `refs.bib`. You can change the name but then you have to change the information in this file from `\bibliography{refs}` to `\bibliography{new-name}` where `new-name` is the name of your file. The file `refs.bib` contains some examples for books and papers.

The process of citation is simple. The command `\textcite{garland2010}` gives this Tucker (2010) and puts all information into the bibliography section at the end. Everything is sorted and formatted, so you don't have to worry about this. An example of a paper with many authors is Benaim and Weibull 2003 or Osborne and Rubinstein 1998. We can cite online resources Overleaf Team 2023 or Cole 2023. We can use the following citation (Benaim and Weibull 2003) or like this² or like this³.

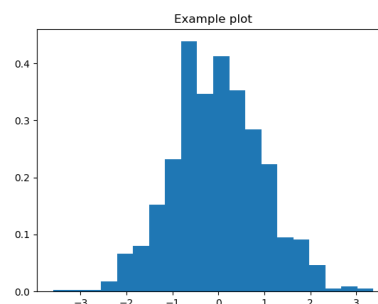


Figure 4: This is how one can wrap a text around a figure. *Source:* own calculations

²Benaim and Weibull 2003.

³Devan Cole (2023). *Supreme Court rejects Jack Smith's request for justices to quickly hear Trump immunity dispute*. CNN. URL: <https://edition.cnn.com/2023/12/22/politics/supreme-court-trump-immunity-jack-smith/index.html>.

A Appendix: Some important stuff

This is an appendix. This is the place to put it if you have some additional figures, tables, or a code. The really long tables or really wide tables should be placed in additional files e.g., XLSX.

```
1 ### Creating another visualization
2 with(d, {
3     plot(x, y,
4         main = "Example dataset",
5         pch = ifelse(class == "Up", 20, 4),
6         col = ifelse(class == "Up", rgb(1, 0, 0, .4), rgb(0, 0, 1, .5)))
7     abline(model, lwd = 2, col = "magenta")
8 })
9 dev.copy(device = png, "./paper/figs/fig_01.png")
10 dev.off()
```

Below, there is a fragment of the `example.py` script that creates a Pandas tabel and exports to \LaTeX .

```
1 # Tworzenie przykładowej tabeli
2 import pandas as pd
3
4 df = pd.DataFrame({'Class':x_letters, 'Values':x_numbers})
5 print(df)
6
7 # Zmiana formatu na LaTeX i eskport do pliku .tex
8 df_tex = df.style.hide(axis=0).to_latex(hrules=True)
9
10 tab_file = os.path.join(os.getcwd(), './paper/tabs/tab_02.tex')
11
12 with open(tab_file, 'w') as file:
13     file.write(df_tex)
14 file.close()
```

References

- Benaim, M. and J. W. Weibull (2003). “Deterministic approximation of stochastic evolution in games”. In: *Econometrica* 71, pp. 873–903.
- Osborne, M. and A. Rubinstein (1998). “Games with procedurally rational players”. In: *American Economic Review* 88, pp. 834–847.
- Tucker, G. S. (2010). *The High Tide of American Conservatism: Davis, Coolidge, and the 1924 Election*. Emerald Book.

Online references

- Cole, Devan (2023). *Supreme Court rejects Jack Smith’s request for justices to quickly hear Trump immunity dispute*. CNN. URL: <https://edition.cnn.com/2023/12/22/politics/supreme-court-trump-immunity-jack-smith/index.html>.
- Overleaf Team (2023). *Bibliography management with biblatex*. Overleaf. URL: https://www.overleaf.com/learn/latex/Bibliography_management_with_biblatex.

List of Tables

1	Short name for a table	9
2	Short name for a table	9
3	Pandas table	10

List of Figures

1	Example of a diagram	8
2	Short name	10
3	Short caption 2	11
4	Short caption 2	12

Streszczenie

Tutaj zamieszczają Państwo streszczenie pracy. Streszczenie powinno być długości około pół strony.