Figures and other stuff for statistics using package `tikz` and `pgfplot`

Erika Griechisch

Last update: Friday 15th September, 2017

Abstract

I teach statistics since 2014, therefore I created several exercise sheets, lecture slides for statistics teaching and besides that, sometimes I need figures when I prepare my posters. During the last few years it was a bit painful to find the best way to create figures, and to spare some time for others, hereby I include all figures in a folder I found and used from different forums (mainly on http://www.texample.net/tikz/examples and https://tex.stackexchange.com/questions/tagged/tikz-pgf).

All figures are stored in separate files using mainly `standalone` documentclass. I tried to

- modify the original sources to make them as general as possible.;
- keep the original source (website) at the beginning of the file. If I did not succeed and you know the original source, please let me now
- keep every example clear and flexible.

  `TODO` comments mark the lines you have to change to adjust the figure to your data.

I have to tell here I am not a pro in LaTeX, though I use for a long time. I take examples from websites and modify them if it is necessary. If you have any idea to make the figures nicer, the sources cleaner and/or easier to understand, feel free to submit a commit or leave me a message!

Contents

1 Form

First of all here is a form using the `paperandpencil.sty` file. Finding an appropriate package to create forms or questionnaires, was really difficult. I think this solution is not the best, but the best I found. Please let me know if you know better way of creating forms in LaTeX!

1. Name

_____

2. Gender

1. male
2. male

3. Age (in years)

☐☐

4. Education level

1. no education
2. primary school
3. secondary school
4. university
5. PhD

5. Body mass (kg)

☐☐☐

6. Height (cm)

☐☐☐

7. Eye color

☐ blue
☐ green
☐ gray
☐ brown
☐ black
☐ other

8. Hobbies

☐ sport
☐ music
☐ stamp
☐ dance
☐ art
☐ other: _____

9. Please evaluate composer W. A. Mozart on a scale. Zero means horrible, ten means fantastic.

0	1	2	3	4	5	6	7	8	9	10
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

10. Date

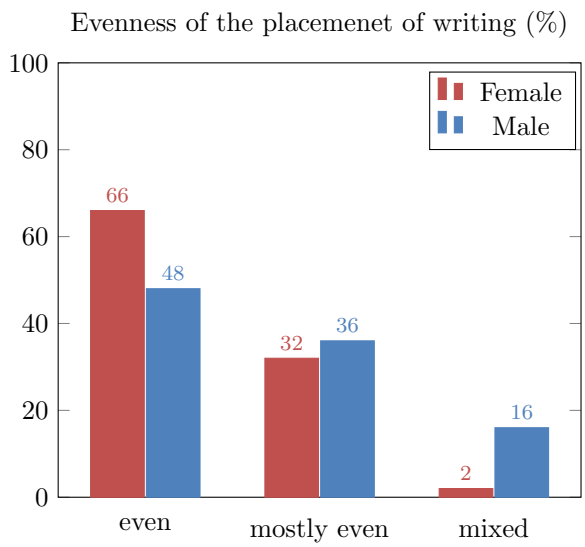
Month ☐☐ Year ☐☐☐☐

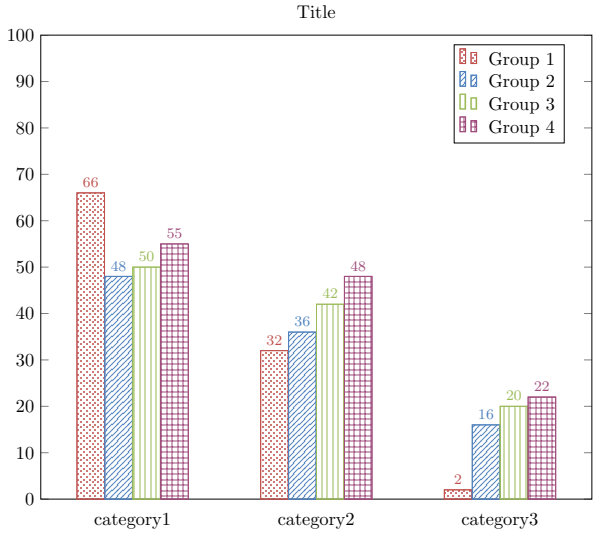
2 Discrete variables

2.1 Barplot

I found 2 ways of plotting barplots for two groups. In my case I compared female and male handwriting based on the evenness of the handwriting, which has 3 possible values: even, mostly even and mixed.

The first figure one on the left is a colorful barchart. The second one is colorful as well, but includes patterns on the rectangles, so it can be used if it is uncertain whether the plot will be printed e.g. in colors or in black-and-white.





Change values and category

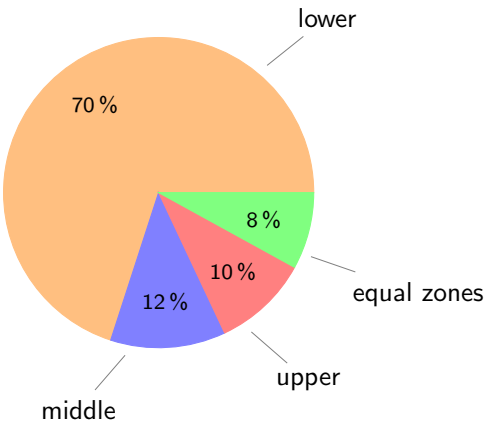
To show the category names below the x axis, use the `symbolic x coords={even,mostly even,mixed}` and change the `even, mostly even, mixed` to your category names. Besides this you have to change the `addplot` line too and instead of using there `{(even,66) (mostly even,32) (mixed,2)}` replace the category names according to the previously mentioned `symbolic x coords`. E.g. `symbolic x coords={category1, category2, category3}` and `{(category1,66) (category2,32) (category3,2)}`.

Change group names

`\legend{Female,Male}` ⇒ `\legend{Group1, Group2}`

2.2 Piechart

I found a nice way to visualize frequencies. Check the source of the `pie1.tex` file, the original source with detailed description is given there. It was not obviuous to me how to interpret the whole file, but it is not necessary to understand each line.



Change percentages and category names

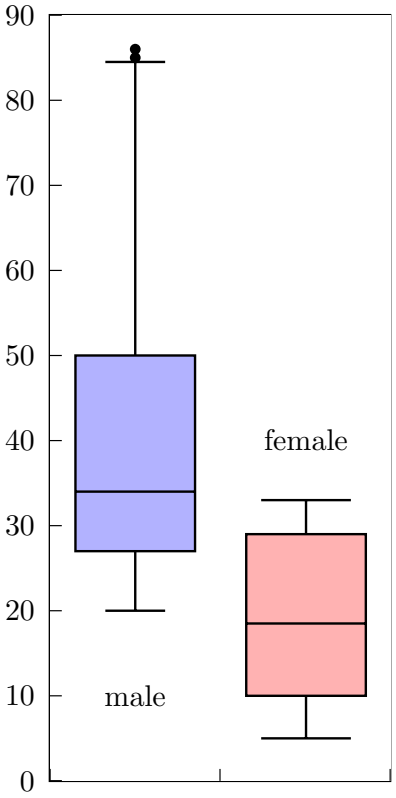
Find the comment starting with `TODO`, the next few lines have to be changed in order to produce piechart with different percentages.

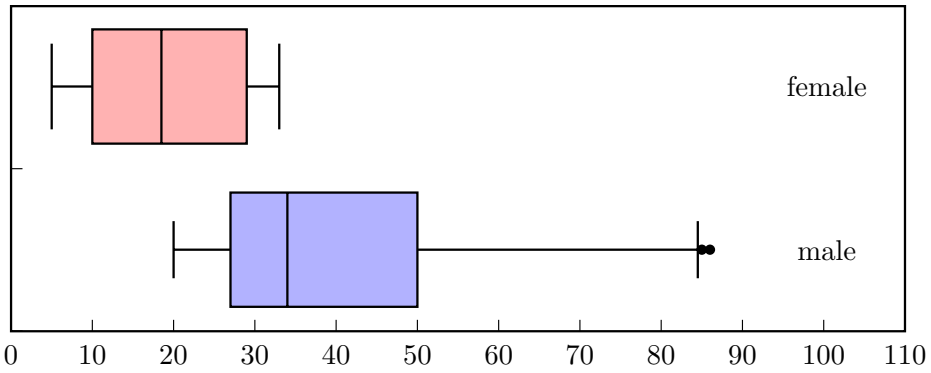
Change colors

The `\def\cyclelist{{"orange","blue","red","green"}}` line defines the colors of the piechart.

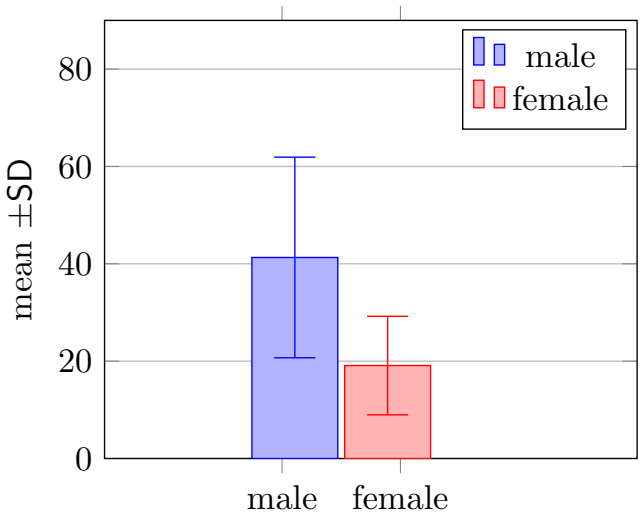
2.3 Boxplot

I could only find one nice way to visualise boxplots in LaTeX using the `pst-plot` package. It can be compiled with `xelatex` instead of `pdflatex`. I am not really familiar with this package, but I played with the parameters I created a horizontal version of a boxplot too. If anyone would extend the sources with descriptions of the parameter in `boxplot_vertical.tex` and `boxplot_horizontal.tex` I would really appreciate :) Still I hope one day I will have enough time to provide a more general source.





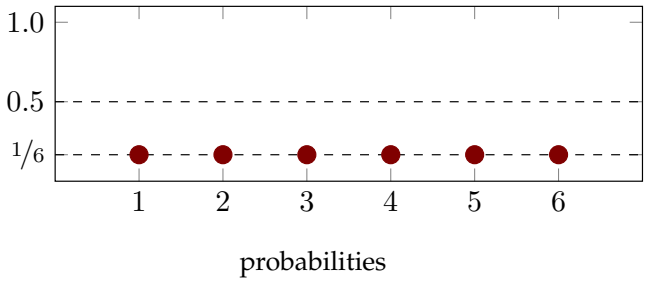
2.4 Mean-SD diagram



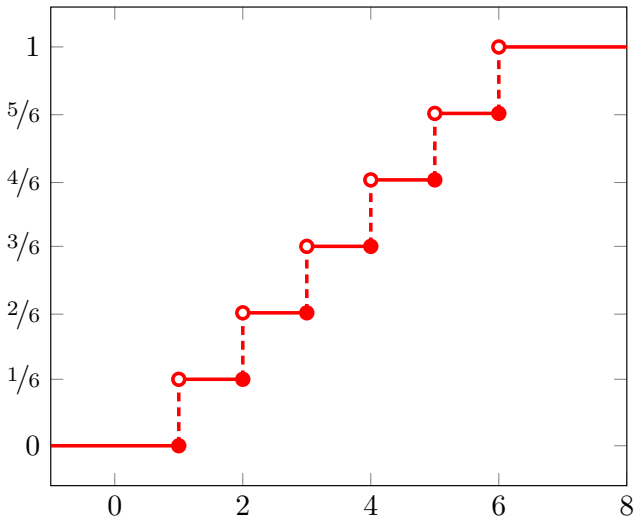
TODO: how to enlarge the distance between the two mean-SD diagram?

3 Discrete distributions

3.1 Rolling a dice

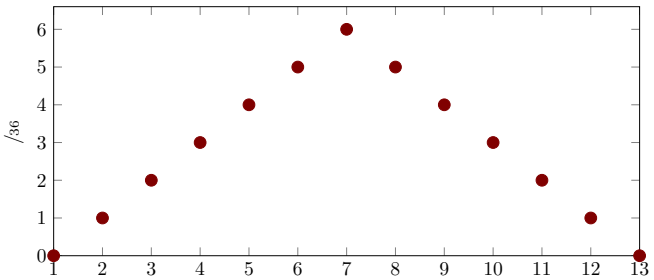


probabilities

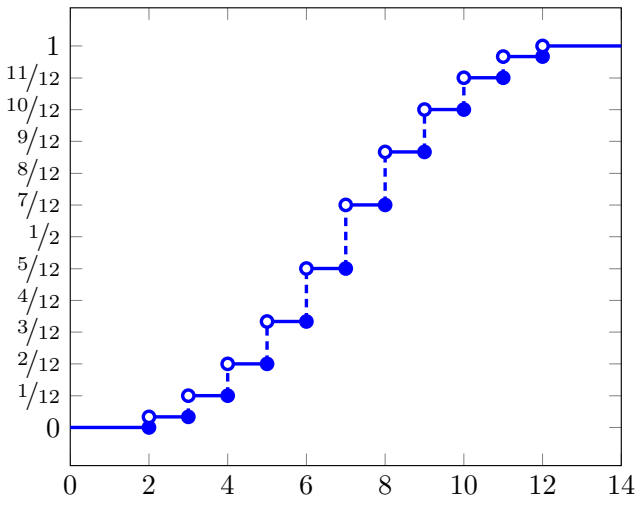


cumulative distribution function (CDF)

3.2 Rolling two dices – sum of the values

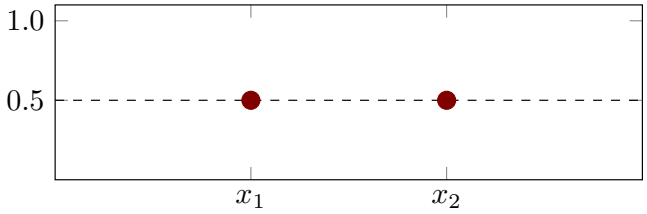


probabilities

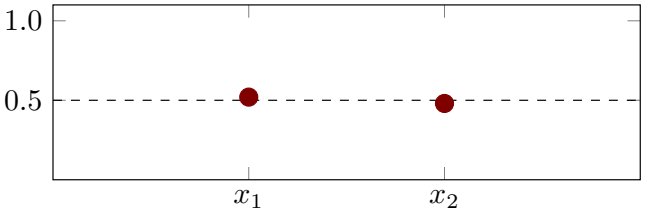


cumulative distribution function (CDF)

3.3 Tossing a coin (head or tail)



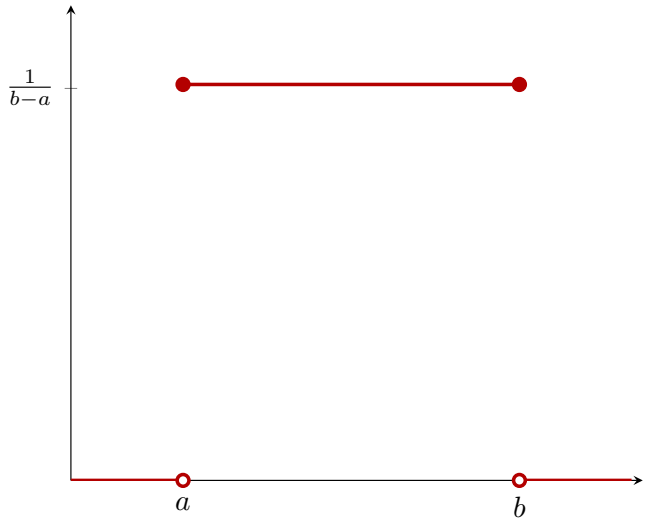
theoretical distribution



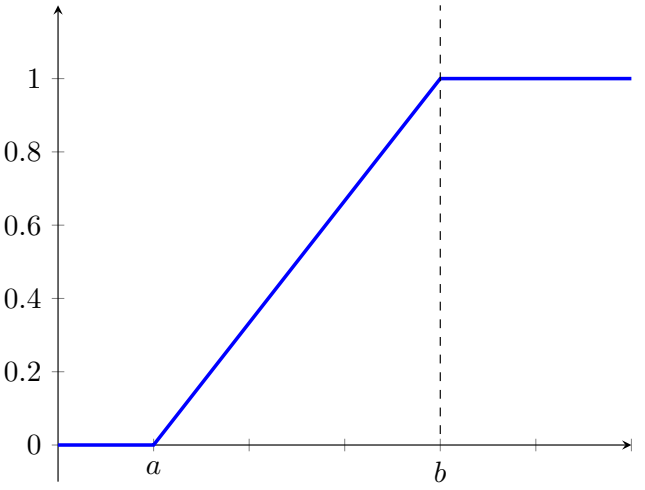
an empirical distribution
52% head (x_1), 48% tail (x_2)

4 Continuous distributions

4.1 Uniform distribution

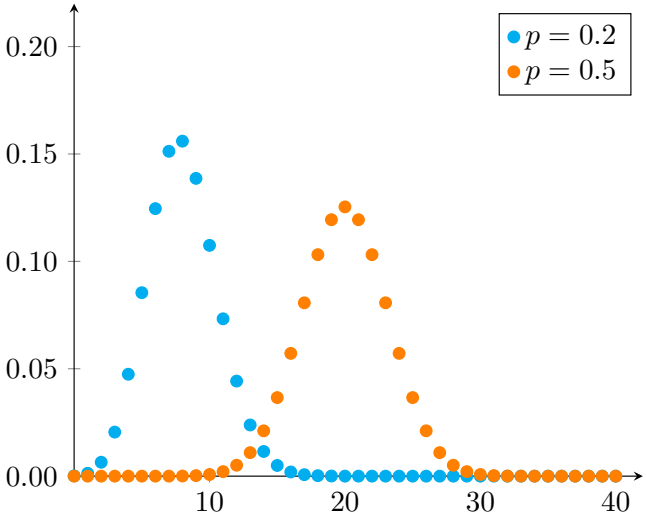


Uniform density function
TODO how to change the value at a to 0 instead of $1/b-a$?

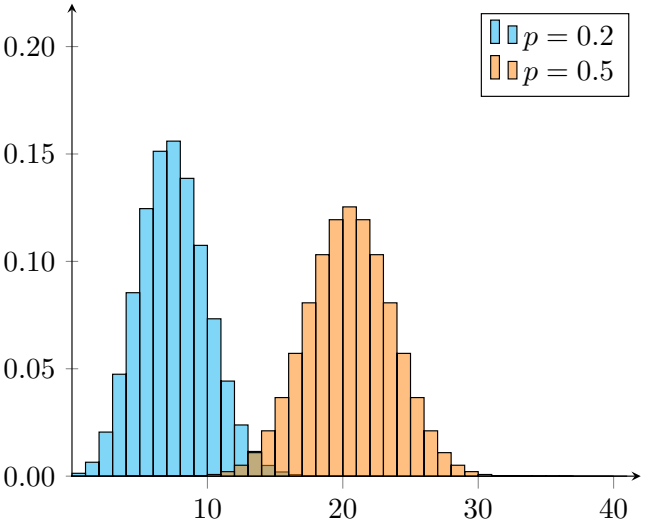


Uniform cumulative distribution function

4.2 Binomial distribution

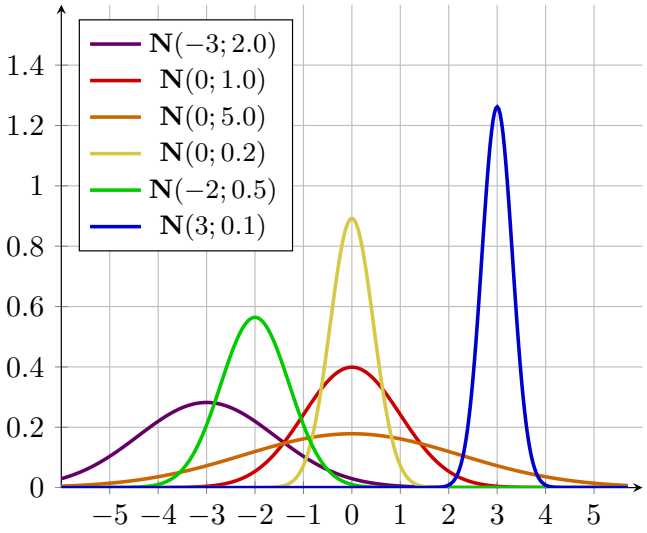


Binomial distribution with 10 repetition ($n = 10$)
and different p values

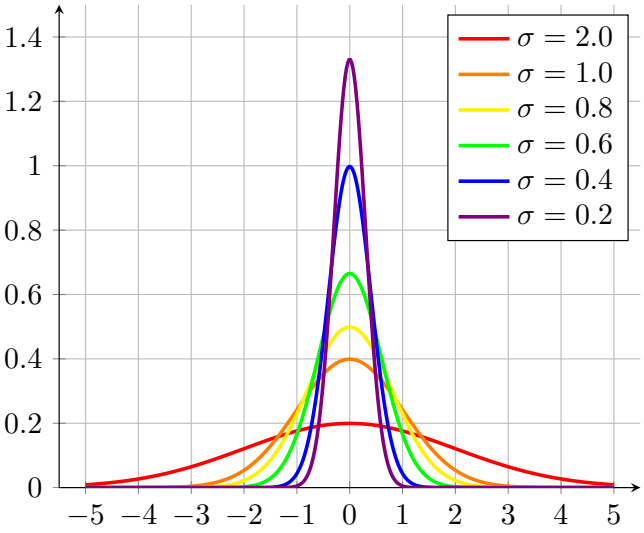


The previous using histogram

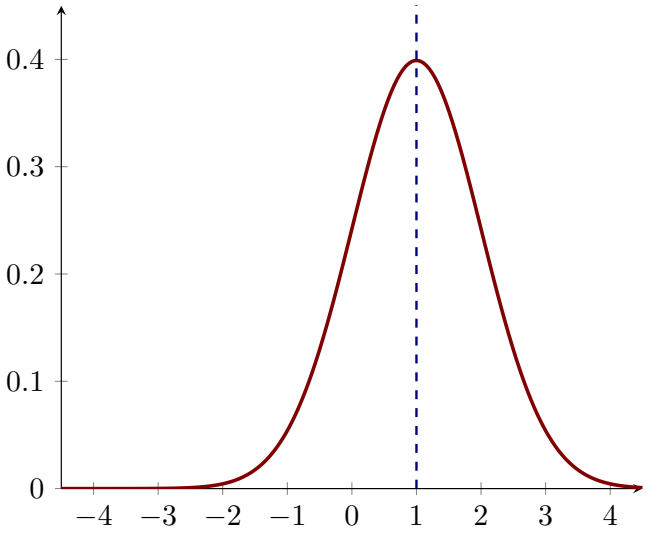
4.3 Normal distribution



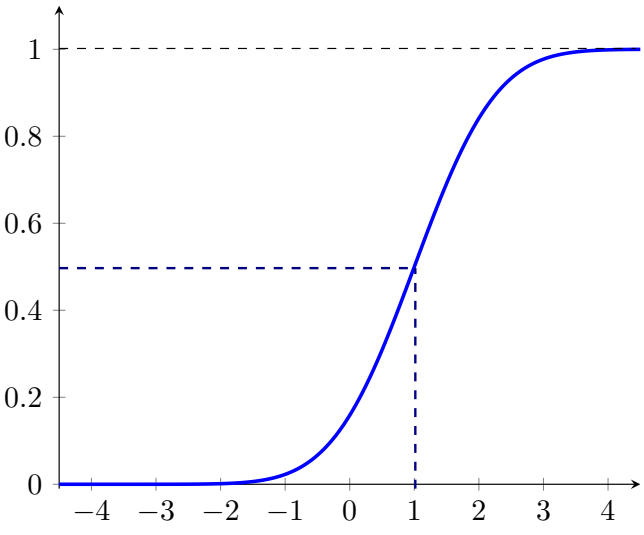
Normal distribution with different parameters



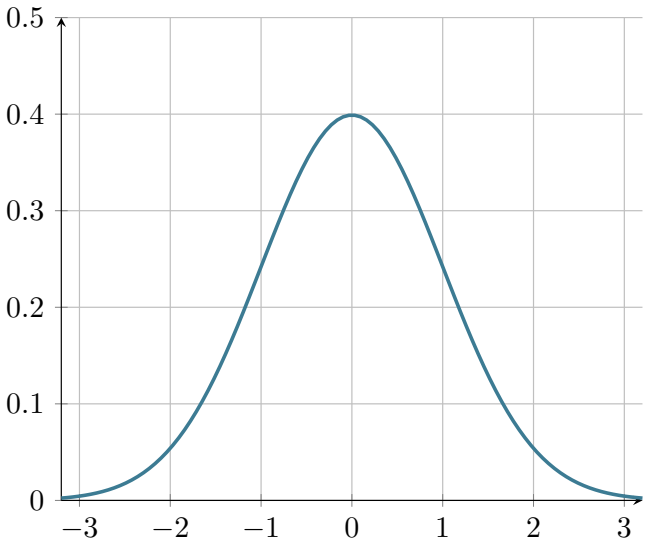
Normal distribution with $\mu = 0$ and different σs



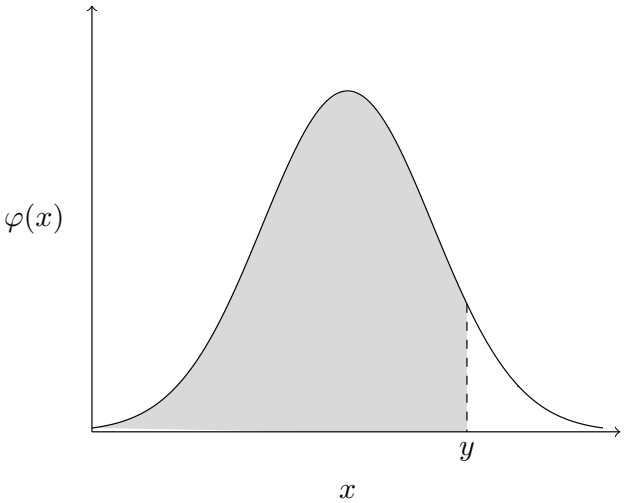
Normal density function
with $\mu = 1$ and $\sigma = 1$



Normal cumulative distribution function
with $\mu = 1$ and $\sigma = 1$

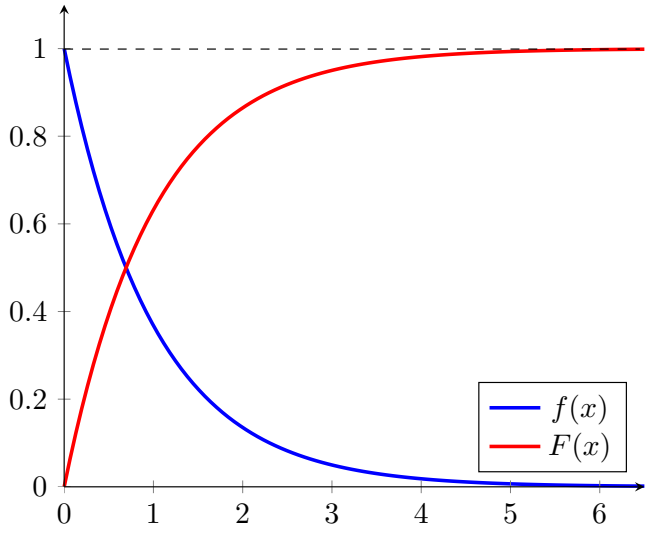


Standard normal density function

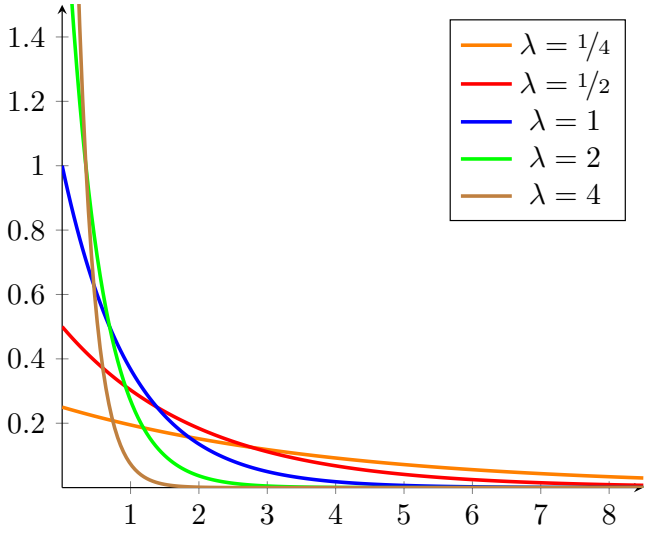


Standard normal density function
with shaded area at y

4.4 Exponential distribution



Exponential distribution, $\lambda = 1$
density function and cumulative distribution function



Exponential distributions
with only density functions and different λ