

MATH 4753 Laboratory 6

Continuous distributions

This lab will investigate some of the continuous distributions. You are expected to learn the skills needed to make plots and areas under density curves.

The Lab progresses from the Normal to the Beta. Once you understand how the Normal works it is an easy matter to extend the techniques you have learnt to other distributions.

It is expected in this lab that you be able to use available information in R to work out how the Beta distribution works by using `?dbeta` in R.

Make sure that all plots are legible – this will mean adjusting `ylim` and `xlim` options in the `curve()` function.

Objectives

In this lab you will learn how to:

1. Plot the densities.
2. Make areas under the curve.
3. Calculate probabilities.
4. Create a function that calculates and displays areas
5. Add the above function to a package

Tasks

Use RMD to knit your final HTML document. Upload both files to canvas at the end.

Note: All plots you are asked to make should be recorded through the correct use of RMD.

- Task 1
 - Make a folder LAB6
 - Download the file “lab6.r”
 - Place this file with the others in LAB6.
 - Start Rstudio
 - Open “lab6.r” from within Rstudio.
 - Go to the “session” menu within Rstudio and “set working directory” to where the source files are located.
 - Issue the function `getwd()`.

- Task 2
 - Open Lab 6.R and make space at the top for the task code needed for the current lab.
 - Use the hash # symbol and write your own comments in the code file explaining what the code does.
 - Use layout() to construct a plotting region that will take 4 plots.
 - Use it to plot 4 normal curves with the following parameters. Hint: You will need to adjust the xlim option to show all of what is needed.
 - $\mu = 10, \sigma = 4$
 - $\mu = 10, \sigma = 2$
 - $\mu = 5, \sigma = 10$
 - $\mu = 5, \sigma = \frac{1}{2}$
 - Plot the following regions and probabilities using one graphical interface per plot (use 4 dec. places)
 - $Y \sim N(0, 1), P(Y \geq 2)$
 - $Y \sim N(\mu = 4, \sigma = 2), P(1 \leq Y < 5)$
 - $Y \sim N(\mu = 10, \sigma = 4), P(Y < 10)$
 - $Y \sim N\left(\mu = -2, \sigma = \frac{1}{2}\right), P(-3 < Y \leq -2)$
- Task 3
 - We will now study the gamma distribution. This has two parameters, shape (>0) and scale (>0). You will use dgamma() and pgamma()
 - Notice that the random variable Y takes only positive values.
 - On the one plotting surface overlay the following three plots (NB – Don't use layout)
 - $Y \sim \text{Gamma}(\text{shape} = 1, \text{scale} = 1)$
 - $Y \sim \text{Gamma}(\text{shape} = 3, \text{scale} = 1)$
 - $Y \sim \text{Gamma}(\text{shape} = 5, \text{scale} = 1)$
 - Plot the following regions and probabilities in R
 - $Y \sim \text{Gamma}(\text{shape} = 3, \text{scale} = 2), P(2 < Y < 5)$
 - $Y \sim \text{Gamma}(\text{shape} = 6, \text{scale} = 3), P(1 \leq Y \leq 4)$
 - $Y \sim \text{Gamma}(\text{shape} = 2, \text{scale} = 4), P(3 \leq Y < 6)$
- Task 4
 - The chi-square distribution is a special case of the gamma with shape = $\frac{\nu}{2}$ and scale = 2, where ν is called the number of degrees of freedom.
 - Use layout() and break up the graphical interface so that there are four squares to plot the following on:
 - $Y \sim \text{chisq}(\text{df} = 1)$
 - $Y \sim \text{chisq}(\text{df} = 2)$
 - $Y \sim \text{chisq}(\text{df} = 4)$
 - $Y \sim \text{chisq}(\text{df} = 20)$
 - Plot the following regions and probabilities in R on separate plotting surfaces.
 - $Y \sim \text{chisq}(\text{df} = 2), P(2 \leq Y \leq 4)$
 - $Y \sim \text{chisq}(\text{df} = 3), P(3 \leq Y \leq 5)$
 - $Y \sim \text{chisq}(\text{df} = 20), P(10 < Y \leq 21)$
- Task 5
 - The Weibull measures length of life like a gamma. $0 \leq Y < \infty$ and it has two parameters, shape and scale both > 0 .
 - Use dweibull() and pweibull().
 - Make four plots as above with any shape and scale combinations you wish (Using layout()).

- Calculate any three areas and probabilities you choose and show them on the graphs as above (separate plots).
- Task 6
 - Investigate the beta distribution $0 \leq Y \leq 1$, it has two parameters.
 - Make sure you understand the parameters by using `?dbeta` in R
 - Repeat task 5 for the beta distribution.
- Task 7
 - Look at the function `myncurve()` below and edit it so that it will display the curve, shaded area between the curve and x axis from $-\infty$ to $x=a$, and calculate the area (probability, $P(X \leq a)$) which is released to the command-line in a list.

```
myncurve = function(mu, sigma){
  curve(dnorm(x,mean=mu,sd=sigma), xlim = c(mu-3*sigma, mu +
  3*sigma))
}
```

- Add the completed function to your package – you can consult <http://r-pkgs.had.co.nz/> for more information about building R packages.
- Build and install the package
- And run the function in an R chunk (part of Lab6.Rmd) to prove the package is working by calculating
`ILAS2019::myncurve(mu=10,sigma=5, a=6)`

LAB FINISHES HERE

- Task 8 – Extra for experts
 - Write some code that will use the gamma density to produce exponential densities. Check the results by using the built in `dexp()` function.