

# Life Contingencies Notes Editing

## Contents

<b>Introduction</b>	<b>1</b>
<b>Software Tools</b>	<b>1</b>
R Markdown . . . . .	1
LaTeX . . . . .	2
Pandoc . . . . .	2
<b>Current progress</b>	<b>2</b>
Conversion steps . . . . .	2
Challenges . . . . .	3
Possibly helpful . . . . .	3
<b>Example: Chapter 1</b>	<b>3</b>

## Introduction

The ultimate goal of this project is to produce an interactive, online, freely available text for actuarial mathematics. An example for Loss Data Analytics can be found **here**.

Currently, we are in the process of developing a similar text for life contingencies and are aggregating various sets of notes.

You will be assisting with converting and editing Professor Jay Vadiveloo's (University of Connecticut) notes **from Microsoft Word to R Markdown**. Knowledge of R, LaTeX, and life contingencies will be helpful.

## Software Tools

### R Markdown

**R Markdown** facilitates the integration of Markdown, LaTeX, and R code to produce html and pdf documents. You can install it in R:

```
install.packages("rmarkdown")
```

An R Markdown file has an **.Rmd** file extension. To produce the html or pdf document output, open the **.Rmd** file in **RStudio** and click **Knit** at the top of the tab. You can use this file as an example.

The header at the beginning of this **.Rmd** file, enclosed in **---**s, is the **YAML header** where specifications can be made about the documents produced. For example, you can add automatic numbering of the sections with **number\_sections: true**.

The following are some useful tips/resources:

- Introduction to R Markdown
- R Markdown Reference Guide
- R Markdown Cheatsheet
- Commenting is done by enclosing text in **<!--** and **-->** (or highlight text and press **Ctrl+Shift+C** in **RStudio**)

## LaTeX

**LaTeX** is a typesetting method, especially useful for mathematical expressions. We will be using LaTeX code in the R Markdown documents.

In LaTeX syntax, math expressions are enclosed in  $\dots$  (inline) or  $\displaystyle \dots$  (display). For example,  $\mu_{x+t}$  and  $\mathbb{P}(A) = \int_{t \in A} f(t) dt$  is displayed as

$$\mathbb{P}(A) = \int_{t \in A} f(t) dt$$

The following are some useful tips/resources:

- ShareLaTeX: online latex editor that doesn't require installation
- Mathematical expressions in LaTeX: there are also plenty of other resources online regarding LaTeX syntax
- R markdown can sometimes be touchy about exporting documents with LaTeX code. For example, it doesn't like hanging spaces or unpaired brackets, e.g.  $\text{\textbf{xxx}}$  should be  $\text{\textbf{xxx}}$  and  $\text{\textbf{xxx}}$  should be  $\text{\textbf{xxx}}$ .

## Pandoc

**Pandoc** is a command-line tool that converts documents between many different file formats. In particular, we are interested in the .docx (MS Word) to .md (Markdown) conversion.

The following are some useful tips/resources:

- Pandoc installation
- Pandoc options
- `pandoc -s Chapter1.docx -t markdown --atx-headers -o Chapter1.md` is the command to convert Chapter1.docx to Chapter1.md (run in PowerShell or another command-line shell)
- After the .docx to .md conversion with Pandoc, you can manually change the .md extension to .Rmd

## Current progress

### Conversion steps

1. Open PowerShell or another command-line shell.
2. `cd "<file directory>"`
3. You can check the files in the current directory with `ls`
4. `pandoc -s FILE.docx -t markdown --atx-headers -o FILE.md`
5. Manually change the .md extension of the produced file to .Rmd
6. Open the .Rmd file in RStudio and add the following YAML header to the beginning:

```
---
title: "Chapter 1 or whatever title you want"
output:
  html_document:
    toc: yes
    toc_depth: 3
# pdf_document:
#   toc: yes
#   toc_depth: '3'
---
```

7. Lots of little things to clean up in the `.Rmd` file

## Challenges

- Section/subsection headings in MS Word don't carry through as Markdown headings with `#` after pandoc conversion unless they are actually formatted in MS Word as headings
- Lots of spacebar and tabs used for equation alignment in the `.docx` files. In R Markdown, the LaTeX `array` seems to work well for tables and formula alignment. See this example.
- Other formatting inconsistencies in `.docx` files
- Some chapters (6-10) use MathType (which creates uneditable Word objects) instead of the Word equation editor. Pandoc can't convert these expressions directly. We used MathType to convert math expressions to LaTeX in the original `.docx` files and go from there. This may or may not be the best method.

## Possibly helpful

- MS Word formatting display to show where spaces/tabs were used.
- Pandoc documentation: there may be pandoc options that are helpful for the conversion (most of them are applicable to html or pdf document conversions however, not MS Word or Markdown)
- Maybe focus on the html output for now, and don't worry about the timelines in the `.docx` files. It might be better to include those as images later.

## Example: Chapter 1