Life Contingencies: The Mathematics, Statistics, and Economics of Life Insurance

An open text authored by the Actuarial Community

Contents

Preface

Date: 25 April 2022

Book Description

Like the companion book Loss Data Analytics, this book on life contingencies will be an interactive, online, freely available text.

- The online version will contain many interactive objects (quizzes, computer demonstrations, interactive graphs, video, and the like) to promote *deeper learning*. A subset of the book will be available for *offline* reading in pdf and EPUB formats.
- Will focus on data and statistical aspects of life contingent events.
- Will emphasize cash flow fundamentals, an approach that allows users to easily adapt approaches to handle complex products.
- This modular approach emphasizing data and cash flow fundamentals has additional advantages:
  - computational aspects become practically relevant through spreadsheet (e.g., `Microsoft Excel`) and numerical (`R`) examples, and
  - an emphasis on the foundations provides an easy entry point for learners who wish an introduction to the field.

How will the text be used?

This book will be useful in actuarial curricula worldwide. Our primary **target audience** is second or third year undergraduates with little to no experience in insurance. Learners may be international; although the book will be in English, we do not expect knowledge of native idiosyncrasies that might be used in the classroom. It will cover the learning objectives of the major actuarial organizations. Thus, it will be suitable for classroom use at universities as well as for use by independent learners seeking to pass professional actuarial examinations.

A secondary audience is the actuarial practitioner (perhaps international) who wishes to retool and learn about modern approaches in the risk management of life contingent events. Thus, the text will also be useful for the continuing professional development of actuaries and other professionals in insurance and related financial risk management industries.

Why is this good for the profession?

An online text is a type of open educational resource (OER). One important benefit of an OER is that it equalizes access to knowledge, thus permitting a broader community to learn about the actuarial profession. Moreover, it has the capacity to engage viewers through active learning that deepens the learning process, producing analysts more capable of solid actuarial work.

Why is this good for students and teachers and others involved in the learning process? Cost is often cited as an important factor for students and teachers in textbook selection (see a recent post on the $400 textbook). Students will also appreciate the ability to "carry the book around" on their mobile devices.

Life Contingent Calculations

Life contingences is a quantitative discipline, enjoying the rigor and discipline of mathematics. Like any mathematical discipline, one traditionally learns about it through the development of formulaic expressions, that is, their proofs, special cases, analysis of special features, and so on. Users of this text find that we do not shy away from presenting summaries of main conclusions using formulaic expressions. Nonetheless, rather than developing insights from mathematical proofs of the primary findings, we demonstrate their impact through short illustrative examples and links to practical applications.

As with other sources that introduce life contingencies, we utilize spreadsheets extensively. In our teaching, we find that spreadsheets are useful for communication and dynamically visualizing results as they evolve over time. However, unlike other sources, we supplement this with approaches that emphasize programming; in this text, we use `R`. Programming methods such as through `R` (and Python, another good candidate) easily accommodate more complex situations that require more computing and, moreover, are built to graphically portray results in an attractive fashion. Analytics, the process of using data to make decisions, is enjoying tremendous attention from many industries; this is certainly true of in data-driven fields that use life contingent methods. By working with data and using programming methods such as `R` in the study of life contingences, users see the connections within many fields that support the actuarial science discpline. Instruction may emphasize any one of the three approaches, traditional mathematical development, spreadsheets, or a computing approach. However, this text contains all three as we believe that future generations of actuaries need to be familiar with all of these different ways to analyze, and communicate, problems that can be solved using life contingent methods.

Project Goal

The project goal is to have the actuarial community author our textbooks in a collaborative fashion. To get involved, please visit our Open Actuarial Textbooks Project Site.

Acknowledgements

We acknowledge the Society of Actuaries for permission to use problems from their examinations.

We thank Rob Hyndman, Monash University, for allowing us to use his excellent style files to produce the online version of the book.

We thank Yihui Xie and his colleagues at Rstudio for the R bookdown package that allows us to produce this book.

We also wish to acknowledge the support and sponsorship of the International Association of Black Actuaries in our joint efforts to provide actuarial educational content to all.

Contributors

The project goal is to have the actuarial community author our textbooks in a collaborative fashion. The following contributors have taken a leadership role in developing *Life Contingencies*.

- **Vali Asimit**
- **Dani Bauer**
- **Adam Butt**
- **Edward (Jed) Frees**
- **Emiliano Valdez**

For our Readers

Like any book, we have a set of notations and conventions. It will probably save you time if you regularly visit our Appendix Chapter **??** to get used to ours.

Freely available, interactive textbooks represent a new venture in actuarial education and we need your input. Although a lot of effort has gone into the development, we expect hiccoughs. Please let your instructor know about opportunities for improvement, write us through our project site, or contact chapter contributors directly with suggested improvements.

Chapter 1

Exercises

1.1 Chapter 4 Exercises

Section 4.3 Exercises

Exercise 4.3.1. Actuarial Present Values for Select and Ultimate Tables. The select and ultimate tables introduced in Section **??** are complicated – is such complexity is warranted in practice? To get insights into this question, in this exercise you will compare actuarial present values derived from select and ultimate mortality to those from only ultimate mortality. To be specific, we focus consider whole life insurance for using the female Canadian experience introduced in Section **??**; the data are available in Appendix Section 2.4.

For these data,

a. produce A_x for $x = 50, \ldots, 65$ using an interest rate of $i = 0.04$ based on
   i. select and ultimate mortality
   ii. ultimate mortality
b. compare these two sequences of actuarial present values by calculating their ratio.
c. repeat parts (a) and (b) using an interest rate $i = 0.02$.

To give you a feel for the results, Figure 1.1 summarizes the results.

Exercise 4.3.1 Answer

\begin{table}

\caption{Comparison of Life Insurance APVs, Interest is 4%}

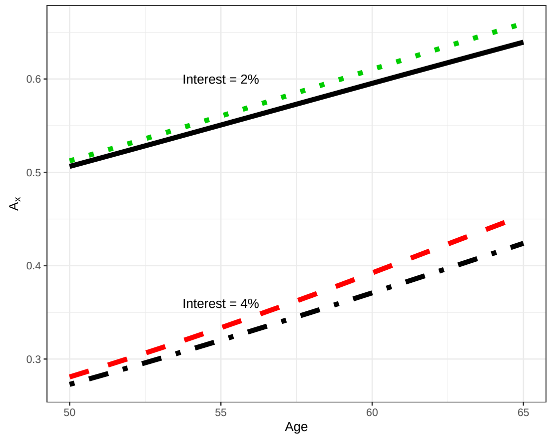

Figure 1.1: **Life Insurance APV by Age, Interest, and Mortality Type**. A plot of life insurance actuarial present value A_x by age x. The top two lines are based on $i = 0.02$, the bottom two based on $i = 0.04$. For each pair, the top line uses ultimate mortality, the bottom line is based on select and ultimate mortality

Age	50	51	52	53	54	55	56	57
Select2 and Ultimate, i=4%	0.2729	0.282	0.2913	0.3007	0.3104	0.3201	0.33	0.3401
Ultimate, i=4%	0.2808	0.2908	0.3011	0.3117	0.3225	0.3336	0.3449	0.3564
Ratio	0.9716	0.9695	0.9673	0.9649	0.9624	0.9598	0.957	0.9542
Age	58	59	60	61	62	63	64	65
Select and Ultimate, i=4%	0.3502	0.3605	0.3709	0.3814	0.3919	0.4025	0.4132	0.424
Ultimate, i=4%	0.3682	0.3801	0.3923	0.4046	0.417	0.4294	0.4417	0.454
Ratio	0.9513	0.9484	0.9455	0.9425	0.9398	0.9375	0.9355	0.934

\end{table}

\begin{table}

\caption{Comparison of Life Insurance APVs, Interest is 2%}

Age	50	51	52	53	54	55	56	57
Select and Ultimate, i=2%	0.5063	0.5151	0.524	0.5329	0.5418	0.5507	0.5596	0.5686
Ultimate, i=2%	0.5122	0.5217	0.5312	0.5408	0.5506	0.5604	0.5702	0.5802
Ratio	0.9885	0.9875	0.9864	0.9853	0.9841	0.9828	0.9814	0.98
Age	58	59	60	61	62	63	64	65
Select and Ultimate, i=2%	0.5775	0.5864	0.5953	0.6042	0.6131	0.6219	0.6307	0.6394
Ultimate, i=2%	0.5901	0.6002	0.6102	0.6203	0.6303	0.6402	0.6499	0.6595
Ratio	0.9786	0.9771	0.9756	0.9741	0.9727	0.9715	0.9704	0.9695

\end{table}

Exercise 4.3.1 Solution using R Code

```r
CanadianSelUlt <- read.csv("Data/CanadianFemaleSelectRead.csv", stringsAsFactors = FALSE)

temp <- CanadianSelUlt[131:201,1:2]
temp1 <- as.matrix(temp)
CanadianUlt50 <- matrix(as.numeric(temp1), ncol=2)
colnames(CanadianUlt50) <- c("Age", "qx_Ult")
rownames(CanadianUlt50) <- NULL

tempSel <- CanadianSelUlt[51:81,]
tempSel1 <- as.matrix(tempSel)

CanadianSel50 <- matrix(as.numeric(tempSel1), ncol=ncol(tempSel))
rownames(CanadianSel50) <- NULL

CanadianSel50.toUlt <- matrix(1, nrow = nrow(CanadianSel50), ncol = 120-50+1)
CanadianSel50.toUlt[1:nrow(CanadianSel50),1:ncol(CanadianSel50)] <- CanadianSel50

for (irow in (1:nrow(CanadianSel50)) ){
for (icol in (1: (120-50-15-irow)) ){
  CanadianSel50.toUlt[irow,1+15+icol ] <- CanadianUlt50[irow+14+icol,2]
   }
}

#  Actuarial Present Value Calculations
APV.Ult <- function(xage, q.x, int=0.00){
  v = 1/(1+int)
  num = length(xage)
  A.x <- 0*xage -> add.x -> e.x
  A.x[num] = v*q.x[num]
  add.x[num] = 1 + (1-q.x[num])/(int + 1e-12)
  e.x[num] = 1-q.x[num]
  for (i in 1:(num-1)){
    A.x[num-i] = v*q.x[num-i] + v*(1-q.x[num-i])*A.x[num-i+1]
    add.x[num-i] = 1 + v*(1-q.x[num-i])*add.x[num-i+1]
    e.x[num-i] =  (1-q.x[num-i])*(1+e.x[num-i+1])
    }
  tablelife <- cbind(xage,q.x,A.x,add.x,e.x)
  return(tablelife)
}
AxUlt04 <- APV.Ult(xage= CanadianUlt50[,1], q.x=CanadianUlt50[,2], int=0.04)[,3]
AxUlt04.5065 <- AxUlt04[1:16]
AxUlt02 <- APV.Ult(xage= CanadianUlt50[,1], q.x=CanadianUlt50[,2], int=0.02)[,3]
AxUlt02.5065 <- AxUlt02[1:16]
```

```
#  Actuarial Present Value Calculations

APV.Sel <- function(int=0.00){
AxSel04 <- rep(0, 16)
  v = 1/(1+int)
for (iage in 1:16){
#iage =2
  AxSel04[iage] =  v*CanadianSel50.toUlt[iage, 2]
kpx = 1
  for (k in 1:(ncol(CanadianSel50.toUlt)-2)) {
    kpx = kpx*(1-CanadianSel50.toUlt[iage, k+1])
   AxSel04[iage] =  AxSel04[iage] + v**(k+1)*kpx*CanadianSel50.toUlt[iage, k+2]
  }
}
return(AxSel04)
}
AxSel04 <- APV.Sel(int=0.04)
AxSel02 <- APV.Sel(int=0.02)

Age1 <- c("50", "51", "52", "53", "54", "55", "56", "57")
Age2 <- c("58", "59", "60", "61", "62", "63", "64", "65")

Ratio04Sel.Ult <- AxSel04/AxUlt04.5065
OutPartA <- rbind(50:57, AxSel04[1:8],  AxUlt04.5065[1:8],  Ratio04Sel.Ult[1:8],
                  58:65, AxSel04[9:16], AxUlt04.5065[9:16], Ratio04Sel.Ult[9:16] )
OutPartA1 <- round(OutPartA, digits = 4)
OutPartA1[1,] <- Age1
OutPartA1[5,] <- Age2
rownames(OutPartA1) <- c(
  "Age",  "Select2 and Ultimate, i=4%", "Ultimate, i=4%","Ratio",
  "Age",  "Select and Ultimate, i=4%", "Ultimate, i=4%", "Ratio")
#knitr::kable(OutPartA1, align = "r", caption="Select and Ultimate Female Canadian APV

Ratio02Sel.Ult <- AxSel02/AxUlt02.5065
OutPartB <- rbind(50:57, AxSel02[1:8],  AxUlt02.5065[1:8],  Ratio02Sel.Ult[1:8],
                  58:65, AxSel02[9:16], AxUlt02.5065[9:16], Ratio02Sel.Ult[9:16] )
OutPartB1 <- round(OutPartB, digits = 4)
OutPartB1[1,] <- Age1
OutPartB1[5,] <- Age2
rownames(OutPartB1) <- c(
  "Age",  "Select and Ultimate, i=2%", "Ultimate, i=2%","Ratio",
  "Age",  "Select and Ultimate, i=2%", "Ultimate, i=2%", "Ratio")
#knitr::kable(OutPartB1, align = "r", caption="Select and Ultimate Female Canadian APV
```

Section 4.5 Exercises

Exercise 4.5.1. Dog Survival Distributions and Actuarial Present Values. The analysis in Section **??** suggests that breed is an important factor for dog survival. To underscore this point for a broader audience, let us look at survival for a type of small dog, a "Jack Russell Terrier", and a large dog, a "German Shepherd Dog". We can make differences in survival distributions between small and large dogs even more meaningful by also computing selected actuarial present values that summarize future expected costs.

You should begin your analysis by downloading the data, available in Appendix Section 2.3.

The data have been re-worked so that they can be used for your analysis. Here is a bit more code needed to convert data fields appropriately and to separate the file into two subsets, one for Terriers and one for German Shepherds.

Show R Code to Pre-Process the Data

```
# Read in Revised Data
DogMortA <- read.csv("DogSurvivalData1.csv")
DogMort <- DogMortA[,-1]
DogMort$dateBirth <- as.Date(DogMort$dateBirth, "%Y-%m-%d")
DogMort$dateDeath <- as.Date(DogMort$dateDeath, "%Y-%m-%d")
DogMort$AgeAtDeath <- (as.numeric(DogMort$dateDeath) - as.numeric(DogMort$dateBirth))/365.25
DogTerr <- subset(DogMort, Breed == "Jack Russell Terrier")
DogGShep <- subset(DogMort, Breed == "German Shepherd Dog")
```

For these data:

a. Compute one-year mortality rates for both Terriers and German Shepherds. Do this over $x = 0, \ldots, 20$. Graph the results.
b. Because you may be working with products with events occurring more frequently than once a year, you decide to follow the strategy introduced in Section **??** and look at mortality rates occurring on a quarterly basis. Repeat the analysis from part (a) using quarterly rates.
c. The results of earlier analyses indicate substantial volatility in the both quarterly rates and annual rates at later ages. One approach to mitigate this volatility is to smooth the rates. A simpler approach is to use longer periods (hence increasing exposure and reducing volatility). Repeat the analysis in part (a) using rates on a once every two year basis.
d. To see how these rates might be used, you decide to repeat the analysis in Section **??** and compute an actuarial present value of the lifetime cost of surgical vet care. Use the same assumptions as in that section with the initial annual cost as 458, $i = 0.03$ interest rate and a 11 percent annual growth in surgical care costs. However, use the dog mortality from part (a). Do this for $x = 2, 5, 12$, for both breeds. You will see that that your results on German Shepherds differ slightly from those presented in Section **??**; comment on why this is so.

Exercise 4.5.1 Answers

```
library(data.table)
library(survival)
# Read in Revised Data
DogMortA <- read.csv("Data/DogSurvivalData1.csv")
DogMort <- DogMortA[,-1]
DogMort$dateBirth <- as.Date(DogMort$dateBirth, "%Y-%m-%d")
DogMort$dateDeath <- as.Date(DogMort$dateDeath, "%Y-%m-%d")
DogMort$AgeAtDeath <- (as.numeric(DogMort$dateDeath) - as.numeric(DogMort$dateBirth))/3
DogTerr <- subset(DogMort, Breed == "Jack Russell Terrier")
DogGShep <- subset(DogMort, Breed == "German Shepherd Dog")

fitTerr <- survfit(Surv(AgeAtDeath) ~ 1, data=DogTerr)
fitGShep <- survfit(Surv(AgeAtDeath) ~ 1, data=DogGShep)
```

Part A

```
seqA1 <- 0:20
seqA2 <- seqA1 + 1
qxTerrA  <- (summary(fitTerr,  time = seqA1)$surv -   summary(fitTerr,  time = seqA2)$s
qxGShepA <- (summary(fitGShep, time = seqA1)$surv - c(summary(fitGShep, time = seqA2)$s
plot(seqA1,qxTerrA, xlab = "Age", ylab = expression(q[x]))
lines(0:16, qxGShepA, col = "blue")
```

Part B

```
seqB1 <- seq(from = 0, to = 20, by = 0.25)
seqB2 <- seqB1 + 0.25
qxTerrB  <- (summary(fitTerr,  time = seqB1)$surv -   summary(fitTerr,  time = seqB2)$s
qxGShepB <- (summary(fitGShep, time = seqB1)$surv - c(summary(fitGShep, time = seqB2)$s
plot(seqB1,qxTerrB, xlab = "Age", ylab = expression(q[x]))
lines(seqB1[1:67], qxGShepB, col = "blue")
```

Part C

```
seqC1 <- seq(from = 0, to = 20, by = 2)
seqC2 <- seqC1 + 2
qxTerrC  <- (summary(fitTerr,  time = seqC1)$surv - c(summary(fitTerr,  time = seqC2)$s
qxGShepC <- (summary(fitGShep, time = seqC1)$surv - c(summary(fitGShep, time = seqC2)$s
plot(seqC1,qxTerrC, xlab = "Age", ylab = expression(q[x]))
lines(seqC1[1:9], qxGShepC, col = "blue")
```

Part D

```
DogAnn <- function(x=x, q.x=qx){
  Term <- length(q.x) - (x+1)
  kpx <- 1
  APV  <- 0
```

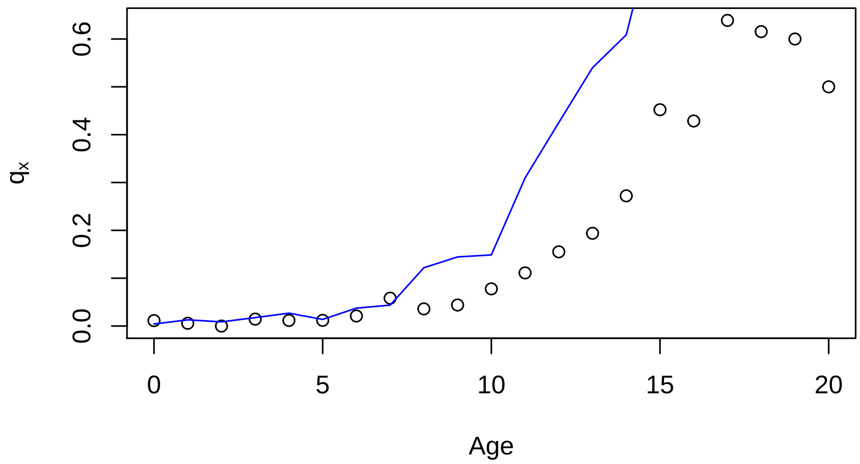

Figure 1.2: One-year Dog Mortality Rates, Blue Solid Line is for German Shepherds

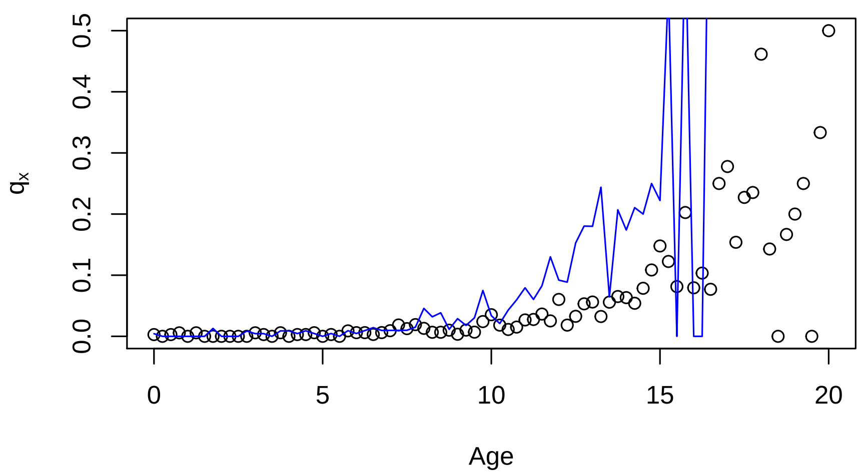

Figure 1.3: Quarterly Dog Mortality Rates, Blue Solid Line is for German Shepherds

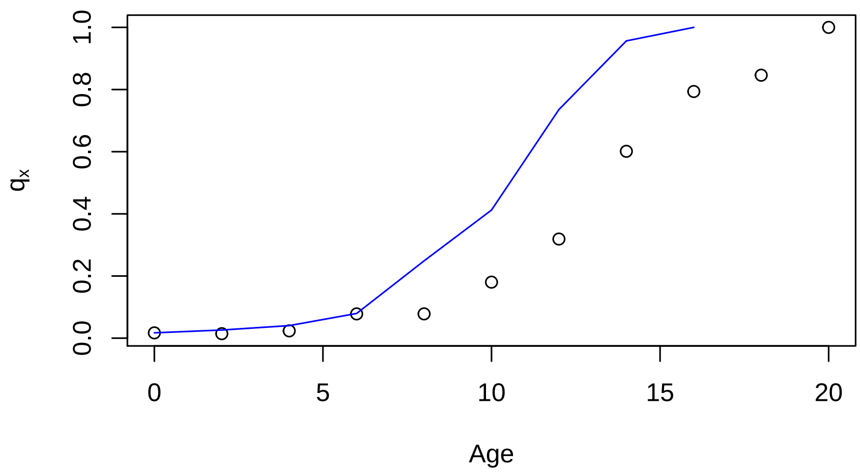

Figure 1.4: Two-year Dog Mortality Rates, Blue Solid Line is for German Shepherds

```
 if (Term>0){ for(k in 0:(Term-1)){
    if (k>1) {kpx <- kpx*(1-q.x[x+1+k])}
    APV  <- APV  + kpx *  (1/1.03)^k   *1.11^k *458
    }}
 return(APV)
  }
DogAnn(x=2, q.x=qxTerrA)
DogAnn(x=5, q.x=qxTerrA)
DogAnn(x=12, q.x=qxTerrA)
DogAnn(x=2, q.x=qxGShepA)
DogAnn(x=5, q.x=qxGShepA)
DogAnn(x=12, q.x=qxGShepA)
```

```
[1] 7605.312
[1] 5111.737
[1] 1795.301
[1] 5444.105
[1] 3420.25
[1] 1184.634
```

Note that the mortality estimates in this exercise are based only on data for a specific breed. In contrast, the analysis in Section **??** used a Cox proportional hazards to regression model to use data from all breeds to form mortality estimates.

Chapter 2

Appendix: Data Dictionary

This appendix describes the datasets used in this book.

For each set of data, we provide download buttons so that you can easily access the data in standard .csv (comma separated value) format. This allows you replicate and experiment with the methods developed in the book as well as sharpen your understanding through exercises.

We provide the source of each dataset. We also recommend, for deeper understanding, that you occasionally refer to these original sources to further develop your appreciation of the data underpinning the analytics developed in this book.

2.1 United States Population Mortality Counts

Source: The Human Mortality Database (HMD).

Description: We now bring into consideration *mortality experience* for populations that had been observed over time, which is available at the Human Mortality Database (HMD) for a wide range of countries. The available data include *Exposures* by age, sex, and calender year period, i.e. how many people of a given age and sex lived in the country's population during a given period of time, and corresponding *Deaths*, i.e. how many of these individuals had died.

The data are available using this download button: Download the United States Population Mortality Counts Data

2.2 Synthetic Insurer Data

Description: We provide survival information for a hypothetical insurance company in `SyntheticInsurerData.csv`. The company has sold *whole life*

Table 2.1: **US Population Mortality Counts First Five Rows**

Year_start	Year_end	Age	Female	Male	Total
1935	1939	0	4869267	5057569	9926836
1935	1939	1	4802597	4936238	9738835
1935	1939	2	5119574	5244634	10364208
1935	1939	3	5159494	5287402	10446896
1935	1939	4	5189350	5307754	10497104

Table 2.2: **US Population Mortality Counts Last Five Rows**

Year_start	Year_end	Age	Female	Male	Total
2010	2014	106	3853.91	415.44	4269.35
2010	2014	107	2031.89	210.41	2242.30
2010	2014	108	1083.62	106.29	1189.91
2010	2014	109	535.32	58.42	593.73
2010	2014	110	576.06	80.03	656.09

Table 2.3: **US Population Mortality Counts Summary Statistics**

	Year_start	Year_end	Age	Female	Male	Total
	Min. :1935	Min. :1939	Min. : 0	Min. : 56	Min. : 21	Min. : 80
	1st Qu.:1954	1st Qu.:1958	1st Qu.: 27	1st Qu.: 1293833	1st Qu.: 823082	1st Qu.: 21
	Median :1972	Median :1976	Median : 55	Median : 5263300	Median : 4850955	Median :10
	Mean :1972	Mean :1976	Mean : 55	Mean : 4946072	Mean : 4758380	Mean : 970
	3rd Qu.:1991	3rd Qu.:1995	3rd Qu.: 83	3rd Qu.: 8238099	3rd Qu.: 8407376	3rd Qu.:16
	Max. :2010	Max. :2014	Max. :110	Max. :11606030	Max. :11674875	Max. :2328

Table 2.4: **Synthetic Insurer First Five Rows**

Month_of_Sale	Age	Sex	Smoking	BMI	BloodPressure	Claim	Time_of_death
1	27	0	0	25.8	117	YES	55.63
1	51	1	0	17.6	109	YES	18.53
1	59	1	0	22.5	132	YES	15.88
1	37	1	0	22.9	109	YES	57.40
1	62	0	0	30.9	147	YES	27.83

Table 2.5: **Synthetic Insurer Last Five Rows**

Month_of_Sale	Age	Sex	Smoking	BMI	BloodPressure	Claim	Time_of_death
780	57	1	1	19.30	118	NO	NA
780	40	0	0	20.10	117	NO	NA
780	27	1	0	20.60	90	NO	NA
780	55	1	1	20.10	118	NO	NA
780	23	1	0	19.03	82	NO	NA

insurance policies, i.e. policies that pay a *death benefit* upon the policyholder's death, since 1955. Policyholders have to go through an underwriting examination, and in addition to policyholders' age, sex (0 for female, 1 for male), smoking status (0 for non-smoker, 1 for smoker), and the month of sale, the company records the applicants body-mass index (BMI) and the systolic blood pressure at the time of underwriting. Finally, for those policyholders with a claim, i.e., for the policyholders that have died, the company records the time of death (relative to the month of underwriting). The data are organized in the order of sales, so that the oldest entries are at the top of the data and the newest entries are at the bottom:

The data are available using this download button: Download the Synthetic Insurer Data

Table 2.6: **Synthetic Insurer Summary Statistics**

	Month_of_Sale	Age	Sex	Smoking
	Min. : 1.0	Min. :19.00	Min. :0.0000	Min. :0.0000
	1st Qu.:209.0	1st Qu.:31.00	1st Qu.:0.0000	1st Qu.:0.0000
	Median :382.0	Median :39.00	Median :1.0000	Median :0.0000
	Mean :394.6	Mean :39.99	Mean :0.6988	Mean :0.2985
	3rd Qu.:591.0	3rd Qu.:48.00	3rd Qu.:1.0000	3rd Qu.:1.0000
	Max. :780.0	Max. :65.00	Max. :1.0000	Max. :1.0000

	BMI	BloodPressure	Claim	Time_of_death
	Min. :16.10	Min. : 57.0	NO :101399	Min. : 0.00
	1st Qu.:19.50	1st Qu.:104.0	YES: 59382	1st Qu.:17.99
	Median :21.70	Median :114.0		Median :28.36
	Mean :22.79	Mean :114.7		Mean :28.27
	3rd Qu.:25.00	3rd Qu.:125.0		3rd Qu.:38.51
	Max. :69.60	Max. :208.0		Max. :64.20
				NA's :101399

Table 2.7: **Canine Mortality First Five Rows**

animalid	Sex	Neutered	Breed	CauseDeath	Method.of.death
3534394	Female	Entire	Boston Terrier	Disorder not diagnsosed	Euthanasia
11836307	Female	Entire	Bulldog	Behaviour disorder	Euthanasia
12090746	Male	Entire	Chinese Shar-Pei	Disorder not diagnsosed	Euthanasia
12092145	Female	Neutered	Crossbreed	Collapsed	Euthanasia
11739930	Female	Entire	German Shepherd Dog	Disorder not diagnsosed	Euthanasia

2.3 Canine Mortality

Source: Canine mortality data are publicly available at Pegram et al. (2021b). See Pegram et al. (2021a) for additional background.

Description: In this dataset, there are 29865 observations used to study euthanasia. Our interest is in survival patterns by age and so we remove records where either the date of birth or death is missing or incorrect (we also removed 17 records with missing sex information); this results in 5933 records for analysis.

The data are available using this download button: Download the Canine Mortality Data

Show R Code to Pre-Process the Canine Mortality Data

```
# Read in Revised Data
DogMortA <- read.csv("Data/DogSurvivalData1.csv")
```

Table 2.8: **Canine Mortality Last Five Rows**

animalid	Sex	Neutered	Breed	CauseDeath	Method.of.death
11718193	Female	Entire	Border Terrier	Heart disease	Euthanasia
11514456	Male	Entire	Jack Russell Terrier	Musculoskeletal disorder	Euthanasia
6916114	Female	Entire	Jack Russell Terrier	Disorder not diagnsosed	Euthanasia
11542454	Male	Neutered	Crossbreed	Disorder not diagnsosed	Unassisted
3963609	Male	Neutered	Jack Russell Terrier	Heart disease	Euthanasia

Table 2.9: **Canine Mortality Summary Statistics**

	animalid	Sex	Neutered	Breed
	Min. : 881900	Female:2841	Entire :2529	Crossbreed :1259
	1st Qu.: 2401823	Male :3092	Neutered:3404	Staffordshire Bull Terrier: 507
	Median : 6846177			Labrador Retriever : 481
	Mean : 6824982			Jack Russell Terrier : 352
	3rd Qu.:11559402			German Shepherd Dog : 234
	Max. :12493098			Yorkshire Terrier : 192
				(Other) :2908

	CauseDeath	Method.of.death	dateBirth	dateDeath
	Disorder not diagnsosed:1126	Euthanasia:5325	2005-01-03: 28	2016-12-12: 41
	Collapsed : 523	Unassisted: 462	2005-01-05: 26	2016-10-07: 38
	Neoplasia : 522	Unrecorded: 146	2004-01-01: 25	2016-08-09: 37
	Mass : 375		2005-01-06: 25	2016-09-12: 37
	Brain disorder : 350		2003-01-05: 24	2016-11-09: 37
	Behaviour disorder : 341		2003-01-10: 24	2016-04-04: 35
	(Other) :2696		(Other) :5781	(Other) :5708

```r
DogMort <- DogMortA[,-1]
DogMort$dateBirth <- as.Date(DogMort$dateBirth, "%Y-%m-%d")
DogMort$dateDeath <- as.Date(DogMort$dateDeath, "%Y-%m-%d")
DogMort$AgeAtDeath <- (as.numeric(DogMort$dateDeath) - as.numeric(DogMort$dateBirth))/365.25
DogTerr <- subset(DogMort, Breed == "Jack Russell Terrier")
DogGShep <- subset(DogMort, Breed == "German Shepherd Dog")
```

2.4 Canadian Female Select and Ultimate Mortality Table

Source

- Individual Life Experience Subcommittee, “Construction of CIA9704 Mortality Tables for Canadian Individual Insurance based on data from 1997 to 2004”, Canadian Institute of Actuaries (2010).
- Accessed: October, 2021 from http://www.cia-ica.ca/docs/default-source/2010/210028e.pdf
- Data retrieved from https://mort.soa.org/ (Table 1458)

Select data are available using this download button: Download the Canadian Female Select Mortality Table.

Table 2.10: **Canadian Female Select Mortality First Five Rows**

Age	Select0	Select1	Select2	Select3	Select4	Select5	Select6	Select7	Select8	Select9	Select10	Select11	Select12	Select13	Select14
0	0.00035	0.00011	0.00011	1e-04	9e-05	9e-05	9e-05	9e-05	0.00009	0.00009	0.00010	0.00010	0.00011	0.00012	0.00014
1	0.00023	0.00011	0.00010	1e-04	9e-05	9e-05	9e-05	9e-05	0.00009	0.00009	0.00010	0.00011	0.00012	0.00013	0.00014
2	0.00011	0.00010	0.00010	9e-05	9e-05	9e-05	9e-05	9e-05	0.00009	0.00010	0.00011	0.00012	0.00013	0.00014	0.00016
3	0.00010	0.00010	0.00009	9e-05	9e-05	9e-05	9e-05	9e-05	0.00010	0.00011	0.00012	0.00013	0.00014	0.00016	0.00017
4	0.00010	0.00009	0.00009	9e-05	9e-05	9e-05	9e-05	1e-04	0.00011	0.00012	0.00013	0.00014	0.00016	0.00017	0.00019

Table 2.11: **Canadian Female Select Mortality Last Five Rows**

Age	Select0	Select1	Select2	Select3	Select4	Select5	Select6	Select7	Select8	Select9	Select10	Select11	Select12	Select13	Select14
76	0.00544	0.00819	0.01031	0.01236	0.01440	0.01799	0.02341	0.02963	0.03723	0.04634	0.05706	0.06945	0.08367	0.09994	0.11842
77	0.00566	0.00853	0.01073	0.01285	0.01633	0.02153	0.02751	0.03484	0.04363	0.05398	0.06599	0.07978	0.09560	0.11358	0.13176
78	0.00588	0.00887	0.01115	0.01457	0.01955	0.02530	0.03235	0.04082	0.05082	0.06243	0.07580	0.09115	0.10864	0.12638	0.14378
79	0.00609	0.00920	0.01264	0.01743	0.02298	0.02975	0.03791	0.04755	0.05877	0.07172	0.08661	0.10359	0.12088	0.13790	0.15707
80	0.00630	0.01043	0.01513	0.02049	0.02702	0.03486	0.04415	0.05499	0.06751	0.08194	0.09843	0.11526	0.13190	0.15065	0.17193

Ultimate data are available using this download button: Download the Canadian Female Ultimate Mortality Table.

Table 2.12: **Canadian Ultimate Mortality First Five Rows**

Age	qx
15	0.00015
16	0.00016
17	0.00018
18	0.00019
19	0.00021

Table 2.13: **Canadian Ultimate Mortality Last Five Rows**

Age	qx
116	0.45
117	0.45
118	0.45
119	0.45
120	1.00

Table 2.14: **Korean Female Insured Lives First Five Rows**

Age	qx
0	0.00505
1	0.00044
2	0.00034
3	0.00025
4	0.00019

Show R Code to Pre-Process the Canadian Female Select and Ultimate Mortality Data

```
CanadianSelect <- read.csv("Data/CanadianFemaleSelectRead.csv", stringsAsFactors = FALSE)
CanadianSelect1 <- subset(CanadianSelect, Age >= 50 & Age <= 69)
CanadianSelect2 <- CanadianSelect1[2:6,c(1:6,12,16)]
CanadianSelect3 <- CanadianSelect1[37:41,1:2]
CanadianSelect4 <- cbind(CanadianSelect2,CanadianSelect3[,2],CanadianSelect3[,1])
colnames(CanadianSelect4) <- c("Age", "Select=0","Select=1","Select=2","Select=3","Select=4","Sel
rownames(CanadianSelect4) <- NULL
knitr::kable(CanadianSelect4, align = "r", caption="Select and Ultimate Female Canadian Mortality
```

2.5 Korean Mortality by Insured Status

Sources:

- Insured lives and annuitant rates were drawn from a database organized by the Society of Actuaries, https://mort.soa.org/. These represent (beginning of the year) mortality rates (q_x).
- The general population data are drawn from the Human Mortality Database. General population data are central death rates (m_x).

Korean Mortality - Insured Lives

The data are available using this download button: Download the 2007 Korean Female Insured Lives Data.

Table 2.15: **Korean Female Insured Lives Last Five Rows**

Age	qx
108	0.76028
109	0.80509
110	0.84620
111	0.88274
112	1.00000

Table 2.16: **Korean Female Annuitant First Five Rows**

Age	qx
45	0.00037
46	0.00041
47	0.00045
48	0.00048
49	0.00051

Korean Mortality - Annuitants

The data are available using this download button: Download the 2007 Korean Female Annuitants Data.

Korean Mortality - Population

The data are available using this download button: Download the Korean Population Mortality Data.

2.6 Mortality by Country

Source: The Human Mortality Database. In addition to classification by country, we also look at experience by sex and age as these distinctions are well known. General population data are central death rates (m_x).

Table 2.17: **Korean Female Annuitant Last Five Rows**

Age	qx
108	0.66236
109	0.71596
110	0.76430
111	0.80563
112	1.00000

Table 2.18: **Korean Female Population First Five Rows**

Year	Age	Female	Male	Total
2003	0	0.004846	0.005998	0.005447
2003	1	0.000477	0.000418	0.000446
2003	2	0.000291	0.000409	0.000353
2003	3	0.000286	0.00031	0.000298
2003	4	0.000261	0.000347	0.000306

Table 2.19: **Korean Female Population Last Five Rows**

Year	Age	Female	Male	Total
2018	106	0.556191	0.667144	0.570571
2018	107	0.615315	0.848961	0.642871
2018	108	0.698276	1.426573	0.761357
2018	109	0.874699	4.285714	1.003864
2018	110	1.804822	.	1.804822

Population Mortality - Japan

The data are available using this download button: Download the Japan Population Mortality Data.

Population Mortality - Poland

The data are available using this download button: Download the Poland Population Mortality Data.

Population Mortality - United States

The data are available using this download button: Download the USA Population Mortality Data.

Show R Code to Pre-Process the Mortality Data by Country

Table 2.20: **Japan Population First Five Rows**

Year	Age	Female	Male	Total
1947	0	0.083595	0.095448	0.089645
1947	1	0.035643	0.037017	0.036341
1947	2	0.017271	0.017203	0.017237
1947	3	0.011227	0.01133	0.011279
1947	4	0.007023	0.007468	0.007248

Table 2.21: **Japan Population Last Five Rows**

Year	Age	Female	Male	Total
2019	106	0.522855	0.69263	0.537623
2019	107	0.534649	0.66628	0.544456
2019	108	0.569069	0.607969	0.571733
2019	109	0.608248	0.806337	0.619736
2019	110	0.630664	0.927907	0.643787

Table 2.22: **Poland Population First Five Rows**

Year	Age	Female	Male	Total
1958	0	0.065617	0.08371	0.074886
1958	1	0.004736	0.005034	0.004888
1958	2	0.001734	0.002034	0.001888
1958	3	0.001191	0.001463	0.00133
1958	4	0.000853	0.001073	0.000965

Table 2.23: **Poland Population Last Five Rows**

Year	Age	Female	Male	Total
2019	106	0.423814	0.307903	0.404767
2019	107	0.394572	0.350672	0.387116
2019	108	0.657174	0.760938	0.673882
2019	109	0.529723	0	0.449214
2019	110	1.753104	0	1.517067

Table 2.24: **USA Population First Five Rows**

Year	Age	Female	Male	Total
1933	0	0.054177	0.068175	0.061292
1933	1	0.008866	0.010039	0.009459
1933	2	0.004025	0.004671	0.004351
1933	3	0.002869	0.003333	0.003104
1933	4	0.002230	0.002537	0.002386

Table 2.25: **USA Population Last Five Rows**

Year	Age	Female	Male	Total
2019	106	0.517094	0.521663	0.517643
2019	107	0.509637	0.586390	0.518399
2019	108	0.637707	0.724113	0.647031
2019	109	0.550384	0.373599	0.531748
2019	110	0.598467	0.509626	0.588324

```
JapanPopFemaleqx  <- as.numeric(JapanPop2019$Female) /(1 + as.numeric(JapanPop2019$Female) /2)
JapanPopMaleqx    <- as.numeric(JapanPop2019$Male)   /(1 + as.numeric(JapanPop2019$Male)   /2)
PolandPopFemaleqx <- as.numeric(PolandPop2019$Female)/(1 + as.numeric(PolandPop2019$Female)/2)
PolandPopMaleqx   <- as.numeric(PolandPop2019$Male)  /(1 + as.numeric(PolandPop2019$Male)  /2)
USAPopFemaleqx    <- as.numeric(USAPop2019$Female)   /(1 + as.numeric(USAPop2019$Female)   /2)
USAPopMaleqx      <- as.numeric(USAPop2019$Male)     /(1 + as.numeric(USAPop2019$Male)     /2)

PolandPopMaleqx[110] <- 1 -> PolandPopMaleqx[111]

IntlMort <- data.frame(cbind(JapanPop2019$Age, JapanPopFemaleqx, JapanPopMaleqx, PolandPopFemalec
                             PolandPopMaleqx, USAPopFemaleqx, USAPopMaleqx))
names(IntlMort) <- c("Age", "qx_Female_Japan", "qx_Male_Japan", "qx_Female_Poland",
                            "qx_Male_Poland",  "qx_Female_USA", "qx_Male_USA")
#str(IntlMort)
```

2.7 Disability Income

Source: Zayatz (2015), Tables 14A, 14B, 21A, 21B

Description:

- Entitlement age `Entryx` denotes age last birthday at entitlement to disability benefits.
- The duration `Select.y` is measured in years since entitlement.
- Attained age `Attainedx` is calculated as sum of `Entryx` and duration.
- The select and ultimate table is read across the row for 0-10 years of entitlement, and down the last (ultimate) column for 10 or more years of entitlement.

Disability Recovery Rates - Male

The data are available using this download button: Download the Male Disability Recovery Rates Data

Table 2.26: **Male Disability Recovery Rates First Five Rows**

Entryx	Select.0	Select.1	Select.2	Select.3	Select.4	Select.5
16	0.009456	0.021605	0.022333	0.012668	0.076435	0.102141
17	0.008263	0.016584	0.014498	0.012572	0.069437	0.087636
18	0.006705	0.012183	0.008357	0.013495	0.058735	0.069197
19	0.004731	0.008923	0.006305	0.013196	0.046345	0.052405
20	0.003801	0.007313	0.005540	0.012486	0.037918	0.043175

Table 2.27: **Male Disability Recovery Rates Last Five Rows**

Entryx	Select.0	Select.1	Select.2	Select.3	Select.4	Select.5
60	0.000623	0.000626	0.000110	0.000793	0.001620	0.001171
61	0.000475	0.000229	0.000088	0.000624	0.001386	NA
62	0.000260	0.000164	0.000070	0.000265	NA	NA
63	0.000250	0.000241	0.000109	NA	NA	NA
64	0.000259	0.000297	NA	NA	NA	NA

Select.6	Select.7	Select.8	Select.9	Ultimate	Attainedx
0.112060	0.075165	0.056866	0.056944	0.040505	26
0.087495	0.063740	0.050628	0.048795	0.035679	27
0.062385	0.051173	0.043173	0.038157	0.030488	28
0.047538	0.040872	0.035954	0.030026	0.025868	29
0.038960	0.032915	0.029194	0.022511	0.020527	30

Select.6	Select.7	Select.8	Select.9	Ultimate	Attainedx
NA	NA	NA	NA	NA	70
NA	NA	NA	NA	NA	71
NA	NA	NA	NA	NA	72
NA	NA	NA	NA	NA	73
NA	NA	NA	NA	NA	NA

Disability Recovery Rates - FeMale

The data are available using this download button: Download the Female Disability Recovery Rates Data

Select.6	Select.7	Select.8	Select.9	Ultimate	Attainedx
0.076092	0.068259	0.070327	0.060100	0.047138	26
0.064836	0.058914	0.060321	0.049663	0.038233	27
0.053373	0.049032	0.048837	0.037076	0.030978	28
0.040477	0.039143	0.039245	0.028696	0.025747	29
0.034821	0.030852	0.030893	0.023135	0.020749	30

Table 2.28: **Female Disability Recovery Rates First Five Rows**

Entryx	Select.0	Select.1	Select.2	Select.3	Select.4	Select.5
16	0.012151	0.018994	0.014082	0.012065	0.047060	0.072175
17	0.009336	0.014397	0.009586	0.010991	0.046968	0.061956
18	0.006091	0.010576	0.007151	0.011074	0.044637	0.049647
19	0.003622	0.008175	0.005680	0.010073	0.036436	0.038087
20	0.003286	0.006613	0.005661	0.009786	0.029334	0.035754

Table 2.29: **Female Disability Recovery Rates Last Five Rows**

Entryx	Select.0	Select.1	Select.2	Select.3	Select.4	Select.5
61	0.000425	0.000412	0.000138	0.000715	0.001145	NA
62	0.000141	0.000276	0.000143	0.000267	NA	NA
63	0.000203	0.000255	0.000084	NA	NA	NA
64	0.000222	0.000223	NA	NA	NA	NA
65	0.000212	NA	NA	NA	NA	NA

Select.6	Select.7	Select.8	Select.9	Ultimate	Attainedx
NA	NA	NA	NA	NA	71
NA	NA	NA	NA	NA	72
NA	NA	NA	NA	NA	73
NA	NA	NA	NA	NA	74
NA	NA	NA	NA	NA	75

Death and Disability Recovery Rates - Male

The data are available using this download button: Download the Male Death and Disability Recovery Rates Data

Table 2.30: **Male Death and Disability Recovery Rates First Five Rows**

Entryx	Select.0	Select.1	Select.2	Select.3	Select.4	Select.5
16	0.013461	0.028723	0.027576	0.018476	0.080463	0.104575
17	0.013996	0.024853	0.020698	0.019304	0.074251	0.092377
18	0.014013	0.021066	0.015289	0.020650	0.064944	0.076273
19	0.014643	0.017300	0.013588	0.020882	0.052550	0.058905
20	0.013950	0.015959	0.012445	0.020439	0.044388	0.049337

Table 2.31: **Male Death and Disability Recovery Rates Last Five Rows**

Entryx	Select.0	Select.1	Select.2	Select.3	Select.4	Select.5
61	0.074628	0.049784	0.041770	0.040088	0.040294	0.043873
62	0.081799	0.054262	0.044259	0.043029	0.045562	0.045980
63	0.095675	0.064214	0.051854	0.050283	0.050324	0.054444
64	0.115280	0.070667	0.055579	0.055825	0.056144	0.062203
65	0.124020	0.074767	0.055770	0.060810	0.060725	0.062011

Table 2.32: **Female Death and Disability Recovery Rates First Five Rows**

Entryx	Select.0	Select.1	Select.2	Select.3	Select.4	Select.5
16	0.014809	0.023141	0.017968	0.017366	0.049789	0.075158
17	0.014491	0.020749	0.014300	0.016203	0.050781	0.066032
18	0.013639	0.018872	0.012580	0.016785	0.049449	0.054591
19	0.013372	0.016900	0.011715	0.015952	0.042374	0.042293
20	0.012495	0.015506	0.011147	0.016198	0.034865	0.040256

Select.6	Select.7	Select.8	Select.9	Ultimate	Attainedx
0.121258	0.082287	0.062582	0.063157	0.048641	26
0.094508	0.070856	0.057023	0.054308	0.042584	27
0.067078	0.057816	0.049906	0.043654	0.036840	28
0.052107	0.046330	0.043202	0.035403	0.031748	29
0.043401	0.039036	0.036866	0.027852	0.026635	30

Select.6	Select.7	Select.8	Select.9	Ultimate	Attainedx
0.047364	0.048433	0.052576	0.054822	0.061678	71
0.049717	0.052812	0.055874	0.057919	0.065458	72
0.056046	0.060056	0.065220	0.065203	0.069131	73
0.061766	0.064501	0.073359	0.073558	0.073835	74
0.064728	0.066258	0.081121	0.081981	0.079210	75

Death and Disability Recovery Rates - Female

The data are available using this download button: Download the Female Death and Disability Recovery Rates Data

Select.6	Select.7	Select.8	Select.9	Ultimate	Attainedx
0.080785	0.077056	0.076726	0.063991	0.053156	26
0.070660	0.066819	0.065950	0.054640	0.044335	27
0.060112	0.056380	0.053425	0.043089	0.036942	28
0.047345	0.046225	0.044149	0.035500	0.032151	29
0.041520	0.036651	0.036679	0.030169	0.028106	30

Table 2.33: **Female Death and Disability Recovery Rates Last Five Rows**

Entryx	Select.0	Select.1	Select.2	Select.3	Select.4	Select.5
61	0.062420	0.042505	0.033209	0.031489	0.028894	0.032939
62	0.067121	0.048084	0.038302	0.033147	0.034093	0.034971
63	0.079963	0.055521	0.041415	0.040768	0.037425	0.040520
64	0.100669	0.060609	0.049733	0.043582	0.046125	0.044295
65	0.108699	0.068326	0.056379	0.046410	0.051687	0.045036

Select.6	Select.7	Select.8	Select.9	Ultimate	Attainedx
0.032992	0.036253	0.039311	0.041570	0.043767	71
0.036726	0.038876	0.042402	0.045475	0.046781	72
0.043163	0.043178	0.048631	0.048940	0.050292	73
0.049059	0.049045	0.054979	0.055358	0.053434	74
0.053099	0.054934	0.061211	0.061746	0.057982	75

Test a few things

Access from Github instead of OneDrive???

Access button Access the Canadian Female Select and Ultimate Mortality Table Github .

Then, do a “File==> Save As” to download the data….

Time taken for this draft of the book: 0.12 minutes.

Bibliography

Pegram, Camilla, Carol Gray, Rowena MA Packer, Ysabelle Richards, David B Church, Dave C Brodbelt, and Dan G O'Neill (2021a). "Proportion and risk factors for death by euthanasia in dogs in the UK," *Scientific reports*, Vol. 11, pp. 1–12.

——— (2021b). "Proportion and risk factors for death by euthanasia in dogs in the UK– Supporting Data," URL: https://researchonline.rvc.ac.uk/id/eprint/13486/, [Retrieved on 15 Dec, 2021].

Zayatz, Tim (2015). "Social Security Disability Insurance Program Worker Experience," *Actuarial Study*, Vol. 123, URL: https://ai.ssa.gov/OACT/NOTES/pdf_studies/study123.pdf.