Actuarial 3 Tarea 1

Actuarial 3

Marcelino Sánchez

14/9/23

Ejercicio 1

Se sabe que el monto de las reclamaciones de un seguro de crédito es Exponencial con media de 80 mil pesos, y que el número de reclamaciones de cada uno de los 48 asegurado (Nj) es () Po e independiente de los demás. En este seguro cada asegurado puede hacer más de una reclamación, pero la probabilidad de ese evento es de sólo 0.0265. La prima recargada agregada (bajo el principio de la desviación estándar) se define por [] [] ESS $\Pi = +$, donde 0 es el margen de seguridad.

ii) ¿Cuál es la prima recargada individual que se debe cobrar a cada asegurado si se considera un margen de seguridad de 0.5? ¿Cómo se compara frente a la respectiva prima neta individual? [10]

```
mediaDistPerd <- 80

varDistPerd <- 80^2

# Averiguando valor de lambda

funcionLambda <- function(lambda){
    return((1+lambda)*exp(-lambda)-1+.0265)
}

# Encontrar la raíz usando uniroot
    resultado <- uniroot(funcionLambda, interval = c(0, 1))

# Imprimir el resultado
    cat("El valor de lambda que hace que la función sea aproximadamente cero es:", resultado$r</pre>
```

El valor de lambda que hace que la función sea aproximadamente cero es: 0.2500015

#Prima recargada

```
mediaDistFrec <- 48*resultado$root</pre>
  vardDistFrec <- 48*resultado$root</pre>
  esperanzaS <- mediaDistPerd*mediaDistFrec
  varianzaS <- varDistPerd*mediaDistFrec + vardDistFrec*mediaDistPerd^2
  paramRecargada <- 0.5
  primaRecargada <- 1000*(esperanzaS+paramRecargada*sqrt(varianzaS))/48</pre>
  cat("La prima recargada agregada es:",primaRecargada , "\n")
La prima recargada agregada es: 24082.62
  cat("La prima recargada individual es:", (primaRecargada/48)*1000, "\n")
La prima recargada individual es: 501721.2
  primaNetaIndividual <- (esperanzaS/48)*1000</pre>
  cat("La prima neta individual es:", primaNetaIndividual, "\n")
La prima neta individual es: 20000.12
  cat("Con lo cual la recargada es ",primaRecargada/primaNetaIndividual, " veces la neta", "
Con lo cual la recargada es 1.204124 veces la neta
```

iii) Aproxime la distribución de S mediante el Teorema Central del Límite para encontrar el margen de seguridad que garantice con 99.5% de probabilidad que la prima recargada agregada será suficiente para hacer frente a las obligaciones de este seguro. ¿Cuál sería

la prima recargada individual en este caso? [10]

```
# Aproximación de la distribución de S mediante el TCL
  # Ya tenemos la media y varianza de S
  # Calculamos la alfa que nos da el 99.5% de seguridad
  alfa <- qnorm(0.995)
  # Calculamos la prima recargada
  (primaRecargada <- (esperanzaS+alfa*sqrt(varianzaS))*1000)</pre>
[1] 1969524
  (primaRecargadaInd <- (primaRecargada/48))</pre>
[1] 41031.74
Ejercicio 2
  # Cargamos librerías
  library(readxl)
  library(openxlsx)
  library(lubridate)
Attaching package: 'lubridate'
The following objects are masked from 'package:base':
    date, intersect, setdiff, union
  library(dplyr) # Load the dplyr package for data manipulation
```

```
Attaching package: 'dplyr'
The following objects are masked from 'package:stats':
    filter, lag
The following objects are masked from 'package:base':
    intersect, setdiff, setequal, union
  library(ggplot2)
  library(tidyr)
  # Set the path to the Excel file and the sheet name
  file_path <- paste0(getwd(), "/bases/Tarea1.xlsx")</pre>
  sheet name <- "2"
  # Read the entire sheet into a data frame
  sheet <- read_excel(file_path, sheet = sheet_name, range = "A7:D427")</pre>
  # Extract year and month from the Fecha column
  sheet$Month <- lubridate::month(sheet$Fecha)</pre>
  sheet$Year <- lubridate::year(sheet$Fecha)</pre>
  sheet$Day <- lubridate::day(sheet$Fecha)</pre>
  sheet$Year <- as.factor(sheet$Year)</pre>
  sheet$Month <- as.factor(sheet$Month)</pre>
  sheet$Ramo <- as.factor(sheet$Ramo)</pre>
  # Print the table
  head(sheet)
# A tibble: 6 x 7
  Fecha
                         Año Ramo
                                                   Monto Month Year
                       <dbl> <fct>
                                                   <dbl> <fct> <fct> <int>
  <dttm>
1 2020-01-02 00:00:00 2020 Incendio y propiedad 2.91 1
                                                                2020
2 2020-01-02 00:00:00 2020 Marítimo
                                                    1.23 1
                                                                2020
                                                                           2
3 2020-01-05 00:00:00 2020 Marítimo
                                                    7.26 1
                                                                2020
                                                                           5
```

```
4 2020-01-06 00:00:00 2020 Incendio y propiedad 7.70 1 2020 6 5 2020-01-09 00:00:00 2020 Incendio y propiedad 1.93 1 2020 9 6 2020-01-12 00:00:00 2020 Incendio y propiedad 1.43 1 2020 12
```

summary(sheet)

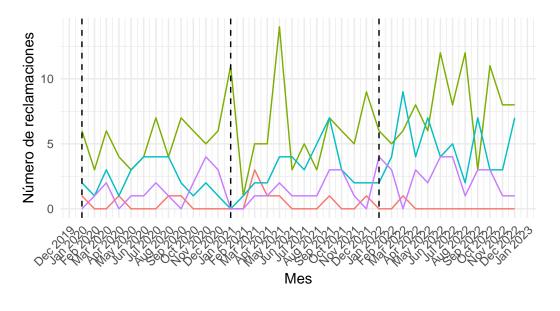
```
Fecha
                                       Año
                                                                     Ramo
Min.
       :2020-01-02 00:00:00.00
                                         :2020
                                  Min.
                                                 Catastrofico
                                                                       : 12
                                  1st Qu.:2020
1st Qu.:2020-12-12 06:00:00.00
                                                 Incendio y propiedad: 228
Median :2021-10-05 00:00:00.00
                                  Median:2021
                                                 Marítimo
                                                                       :120
       :2021-09-06 14:41:08.56
                                  Mean
                                         :2021
                                                 Responsabilidad civil: 60
3rd Qu.:2022-06-08 06:00:00.00
                                  3rd Qu.:2022
       :2022-12-31 00:00:00.00
Max.
                                  Max.
                                         :2022
    Monto
                     Month
                                  Year
                                                Day
Min.
      : 0.113
                        : 43
                                2020:110
                 5
                                                  : 1.00
                                           Min.
1st Qu.: 2.571
                        : 41
                                2021:130
                                           1st Qu.: 7.00
Median : 4.463
                 7
                        : 39
                                2022:180
                                           Median :15.00
Mean : 6.733
                 3
                        : 38
                                           Mean :15.01
3rd Qu.: 7.201
                                           3rd Qu.:22.00
                 10
                        : 38
Max.
       :77.042
                 12
                        : 38
                                           Max. :31.00
                 (Other):183
```

b) Construya una gráfica que permita analizar la evolución mensual del número de reclamaciones por ramo de 2020 a 2022

[`]summarise()` has grouped output by 'Year', 'Month'. You can override using the `.groups` argument.

```
# Left join the grid with your data
grouped_data <- all_combinations %>%
  left_join(agg_data, by = c("Year", "Month", "Ramo")) %>%
  mutate(Num_Reclamaciones = ifelse(is.na(Num_Reclamaciones), 0, Num_Reclamaciones))

# Create the plot with faceting by year and connected lines across months
ggplot(grouped_data, aes(x = as.Date(paste(Year, Month, "01", sep = "-")), y = Num_Reclamacione() +
  geom_vline(aes(xintercept = as.Date(paste(Year, "01-01", sep = "-"))), color = "black",
  labs(x = "Mes", y = "Número de reclamaciones", color = "Ramo") +
  scale_x_date(date_labels = "%b %Y", date_breaks = "1 month") +
  theme_minimal() +
  theme(legend.position = "bottom") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) # Rotate x-axis labels for real
```

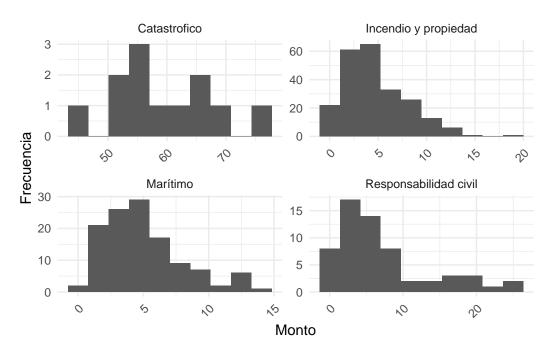


c) Con la información de los 3 años construya 4 histogramas de los montos individuales de los siniestros de cada uno de los ramos. [10]

Ramo — Catastrofico — Incendio y propiedad — Marítimo — Responsabilidad c

```
# Cargar el paquete necesario
library(ggplot2)
# Crear los histogramas
```

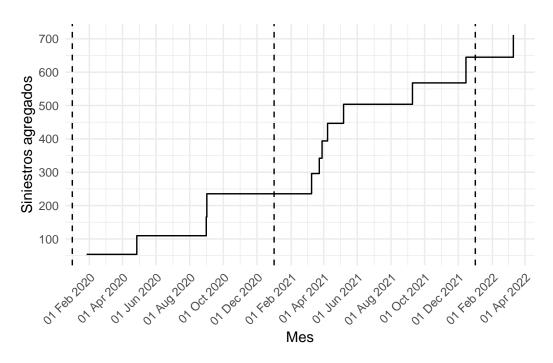
```
ggplot(sheet, aes(x = Monto)) +
  geom_histogram(bins = 10) +
  facet_wrap(~ Ramo, scales = "free" ) +
  labs(x = "Monto", y = "Frecuencia") +
  theme_minimal() +
  theme(legend.position = "bottom") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) # Rotate x-axis labels for real
```



- d) Se quiere analizar exclusivamente el ramo Catastrófico de 2020 a 2022:
- e) Grafique el proceso de siniestros agregados. [10]

```
catSheet <- sheet %>%
  filter(Ramo == "Catastrofico") %>%
  mutate(acumMonto = cumsum(Monto))

ggplot(catSheet, aes(x = as.Date(paste(Year, Month, Day,
  sep = "-")), y = acumMonto)) + geom_step() +
  geom_vline(aes(xintercept = as.Date(paste(Year, "01-01", sep = "-"))), color = "black",
  labs(x = "Mes", y = "Siniestros agregados", color = "Ramo") +
  scale_x_date(date_labels = "%d %b %Y", date_breaks = "2 month") +
  scale_y_continuous(breaks = seq(0,800, 100)) +
  theme_minimal() +
  theme(legend.position = "bottom") +
```



ii) Suponga que la aseguradora cobró como prima neta agregada el monto total de los 12 siniestros y que cobró una tercera parte al inicio de cada año a sus asegurados. Grafique el proceso de primas recolectadas. [5]

```
prima_agregada_final <- tail(catSheet$acumMonto, 1)

prima_agregada_func <- function(x){
    return(x*prima_agregada_final/3)
}

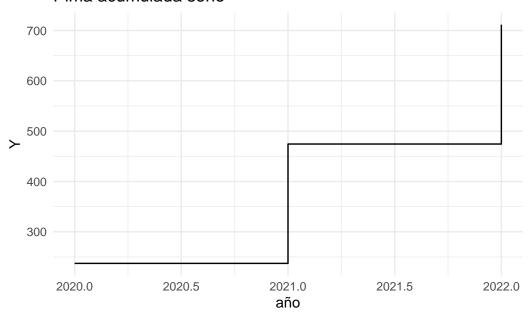
x_values <- seq(2020, 2022,1)
y_values <- sapply(x_values+1-2020, prima_agregada_func)

# Create an igraph graph
prima_agregada_serie <-data.frame(x_values = x_values, y_values = y_values)

# Plot the graph using ggplot2
ggplot(prima_agregada_serie, aes(x = x_values, y = y_values)) +
    geom_step() +
    scale_y_continuous(breaks = seq(0,800, 100)) +</pre>
```

```
labs(x = "año", y = "Y", title = "Pima acumulada serie") +
theme_minimal()
```

Pima acumulada serie



iii) Si la Aseguradora tenía un capital inicial para estos riesgos de 50 millones de dólares, combine los procesos anteriores para graficar el proceso de excedentes. ¿Qué puede concluir? (Sugerencia: Considere el tiempo en el que se presenta el primer excedente por debajo del capital inicial y el tiempo de ruina). [10]

```
# Crear un data frame con los datos de los excedentes

serieMontos <- catSheet %>%
    select(Year, Month, Day, Monto)

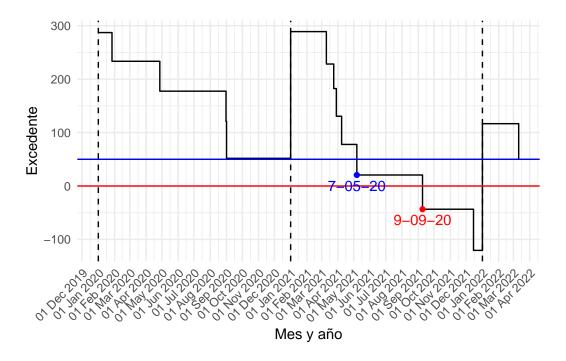
for (i in 1:3) {
    new_row <- c(as.character(2020 + i-1), "1", "1", -prima_agregada_func(1))
    serieMontos <- rbind(serieMontos, new_row)
}

serieMontos <- serieMontos %>%
    arrange(Year, Month, Day) %>%
    mutate(excedente = 50-cumsum(Monto))
```

```
# Filtrar rows donde el excedente es negativo y
  #donde el excedente es menor a 50
  serieMontos <- serieMontos %>%
    mutate(excedenteMenor50 = excedente < 50,</pre>
           excedenteNegativo = excedente < 0)</pre>
  print(serieMontos[serieMontos$excedenteNegativo,])
# A tibble: 2 x 7
                   Monto excedente excedenteMenor50 excedenteNegativo
 Year Month Day
 <fct> <fct> <chr> <chr>
                               <dbl> <lgl>
                                                      <1g1>
1 2021 9
             9 64.071
                              -43.5 TRUE
                                                      TRUE
2 2021 12
             15
                   77.042
                             -121. TRUE
                                                      TRUE
  print(serieMontos[serieMontos$excedenteMenor50,])
# A tibble: 4 x 7
 Year Month Day
                   Monto excedente excedenteMenor50 excedenteNegativo
 <fct> <fct> <chr> <chr>
                               <dbl> <lgl>
                                                      <lgl>
1 2021 5
             7
                   57.22
                               20.6 TRUE
                                                      FALSE
2 2021 9
             9
                   64.071
                              -43.5 TRUE
                                                      TRUE
3 2021 12 15 77.042
                             -121. TRUE
                                                      TRUE
4 2022 3
                   66.663
             11
                               50.0 TRUE
                                                     FALSE
  # Extract points where excedenteNegativo is TRUE
  points_excedente_negativo <- head(serieMontos[serieMontos$excedenteNegativo,],1)</pre>
  # Extract points where excedenteMenor50 is TRUE
  points_excedente_menor_50 <- head(serieMontos[serieMontos$excedenteMenor50,],1)</pre>
  # Graficar el proceso de excedentes
  ggplot(serieMontos, aes(x = as.Date(paste(Year, Month, Day,
  sep = "-")), y = excedente)) + geom_step() +
    geom_vline(aes(xintercept = as.Date(paste(Year, "01-01", sep = "-"))), color = "black",
    labs(x = "Mes y año", y = "Excedente") +
    scale_x_date(date_labels = "%d %b %Y", date_breaks = "1 month") +
    geom_hline(yintercept = 0, color = "red") +
    geom_hline(yintercept = 50, color = "blue") +
    theme_minimal() +
    theme(legend.position = "bottom") +
```

```
theme(axis.text.x = element_text(angle = 45, hjust = 1)) + # Rotate x-axis labels for reageom_point(data = points_excedente_negativo, aes(x = as.Date(paste(Year, Month, Day, sepgeom_point(data = points_excedente_menor_50, aes(x = as.Date(paste(Year, Month, Day, sepgeom_text(data = points_excedente_negativo, aes(label = as.Date(paste(Day, Month, Year, geom_text(data = points_excedente_menor_50, aes(label = as.Date(paste(Day, Month, Year,
```

Warning in strptime(x, ff): datetimes before 1902 may not be accurate: warns once per session



Ejercicio 4

Situación Financiera de las Aseguradoras.

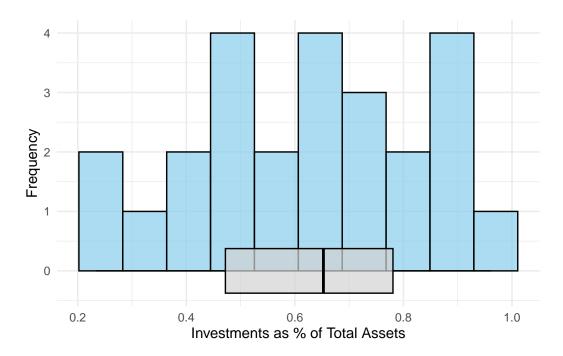
En el archivo "Tarea 1.xls" aparecen Estados Financieros de las 83 Compañías Aseguradoras en México al 31 de diciembre de 2022. Considere la muestra integrada por las 25 principales Compañías Aseguradoras en términos de prima emitida para responder lo siguiente.

```
# Set the path to the Excel file and the sheet name
file_path <- pasteO(getwd(), "/bases/Tarea1.xlsx")
sheet_name <- "4"</pre>
```

```
# Read the entire sheet into a data frame
sheet2 <- read_excel(file_path, sheet = sheet_name, range = "A8:AP91")
muestrasEmpresas <- head(sheet2[order(sheet2$"Primas Emitidas", decreasing = T),],25)</pre>
```

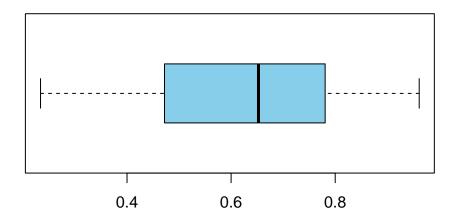
CC a) Construya el histograma y el diagrama de caja y brazos para las inversiones como porcentaje del activo total. Interprete sus resultados en términos de la estructura financiera de las aseguradoras. [5]

```
muestrasEmpresas$percentage_column <- muestrasEmpresas$Inversiones/muestrasEmpresas$"Active
ggplot(muestrasEmpresas, aes(x = percentage_column)) +
    geom_histogram(bins= 10, fill = "skyblue", color = "black", alpha = 0.7) +
    geom_boxplot( fill = "lightgray", color = "black", alpha = 0.7,outlier.colour = "red") +
    labs(x = "Investments as % of Total Assets", y = "Frequency") +
    theme_minimal()</pre>
```



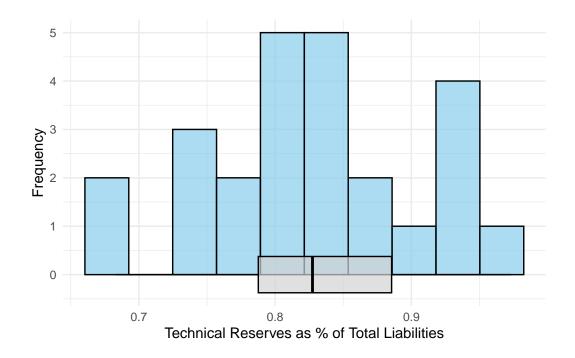
boxplot(muestrasEmpresas\$percentage_column, main = "Investments as % of Total Assets", col

Investments as % of Total Assets



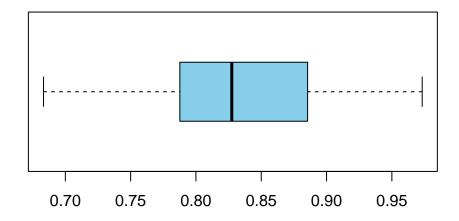
CC b) Construya el histograma y el diagrama de caja y brazos para las reservas técnicas como porcentaje del pasivo total. Interprete sus resultados en términos de la estructura financiera de las aseguradoras. [5]

```
muestrasEmpresas$percentage_column2 <- muestrasEmpresas$"Total Reservas Técnicas"/muestras
ggplot(muestrasEmpresas, aes(x = percentage_column2)) +
   geom_histogram(bins= 10, fill = "skyblue", color = "black", alpha = 0.7) +
   geom_boxplot( fill = "lightgray", color = "black", alpha = 0.7,outlier.colour = "red") +
   labs(x = "Technical Reserves as % of Total Liabilities", y = "Frequency") +
   theme_minimal()</pre>
```



boxplot(muestrasEmpresas\$percentage_column2, main = "Technical Reserves as % of Total Liab

Technical Reserves as % of Total Liabilities



CC c) Obtenga una gráfica de barras apiladas que permita comparar el tamaño de las 4 reservas técnicas como porcentaje del pasivo total para las 25 principales aseguradoras. ¿Qué puede concluir? [10]

```
muestrasEmpresasReservas <- muestrasEmpresas %>%
  select("Aseguradora", "Reserva de Riesgos en Curso", "Reserva para Obligaciones Pendient
muestrasEmpresasReservas$percentage_column11 <- muestrasEmpresasReservas$"Reserva de Riesg
muestrasEmpresasReservas$percentage_column12 <- muestrasEmpresasReservas$"Reserva para Obl
muestrasEmpresasReservas$percentage_column13 <- muestrasEmpresasReservas$"Reserva de Conti
muestrasEmpresasReservas$percentage_column14 <- muestrasEmpresasReservas$"Reservas de Ries
muestrasEmpresasReservas$percentage_column1 <- muestrasEmpresasReservas$"Total Reservas Té
agrupamosReservas <- muestrasEmpresasReservas %>%
  select(
    Aseguradora,
    percentage_column11,
   percentage_column12,
    percentage_column13,
    percentage_column14,
    "Primas Emitidas"
# Assuming you have aggregated data in the "agrupamosReservas" dataframe
# Reshape the data to a longer format
agrupamosReservas$Aseguradora <- factor(agrupamosReservas$Aseguradora, levels = agrupamosRe
agrupamosReservas_long <- agrupamosReservas %>%
  pivot_longer(cols = starts_with("percentage_column"), names_to = "Reserva_Type", values_
# Create the stacked bar plot
ggplot(agrupamosReservas_long, aes(x = Aseguradora, y = Percentage, fill = Reserva_Type))
  geom_bar(stat = "identity") +
  labs(x = "Aseguradora", y = "Porcentaje") +
  ggtitle("Reservas Técnicas como % del Pasivo Total") +
  scale_fill_discrete(labels = c(
    "Reserva de Riesgos en Curso",
    "Reserva para Obligaciones Pendientes de Cumplir",
    "Reserva de Contingencia",
    "Reservas de Riesgos Catastróficos",
```

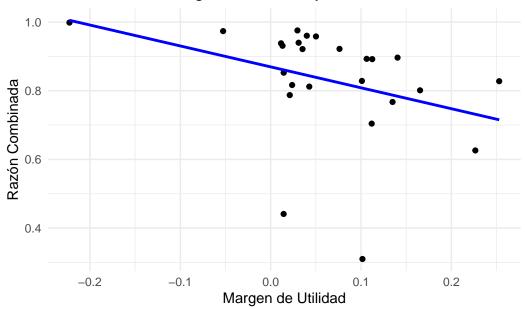
```
"Total Reservas Técnicas"
                          )) +
                          theme_minimal() +
                          theme(legend.position = "bottom") +
Euro Mexica control for the state of the sta
                             theme(axis.text.x = element_text(angle = 45, hjust = 1)) # Rotate x-axis labels for rea
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BENZ Segure Mexico, S.A. de C.V. Grupo Medille Mexico.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         or sedifordings, company of sedifordings of se
```

Aseguradora

CC d) Analice mediante el diagrama de dispersión y el coeficiente de correlación lineal la relación entre: (i) el margen de utilidad (resultado del ejercicio / prima emitida) vs. la razón combinada;

```
muestrasEmpresas$margenUtilidad <- muestrasEmpresas$"Utilidad (Pérdida) del Ejercicio"/mu
muestrasEmpresas$razonCombinada <- (muestrasEmpresas$"Costo Neto de Siniestralidad, Reclam
# Create a scatter plot
ggplot(muestrasEmpresas, aes(x = margenUtilidad, y = razonCombinada)) +
  geom_point() +
  geom_smooth(method = "lm", se = FALSE, color = "blue") +
  labs(x = "Margen de Utilidad", y = "Razón Combinada") +
  ggtitle("Relación entre Margen de Utilidad y Razón Combinada") +
  theme_minimal()
```

Relación entre Margen de Utilidad y Razón Combinada



```
# Calculate the linear correlation coefficient
correlation <- cor(muestrasEmpresas$margenUtilidad, muestrasEmpresas$razonCombinada)
print(paste("Coeficiente de correlación lineal:", correlation))</pre>
```

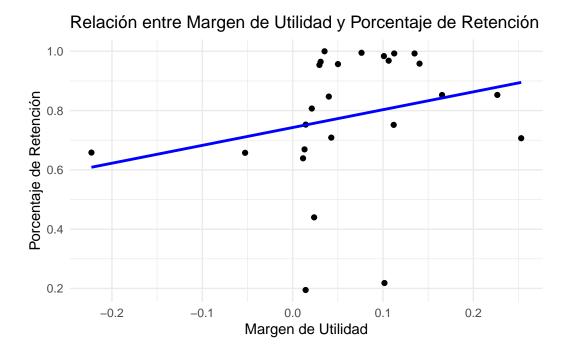
[1] "Coeficiente de correlación lineal: -0.342509417544556"

y (ii) el margen de utilidad vs. el porcentaje de retención (prima de retención / prima emitida). ¿Qué pude concluir? [10]

```
muestrasEmpresas$porcentRetencion <- muestrasEmpresas$"Primas De Retención"/muestrasEmpres

# Create a scatter plot
ggplot(muestrasEmpresas, aes(x = margenUtilidad, y = porcentRetencion)) +
    geom_point() +
    geom_smooth(method = "lm", se = FALSE, color = "blue") +
    labs(x = "Margen de Utilidad", y = "Porcentaje de Retención") +
    ggtitle("Relación entre Margen de Utilidad y Porcentaje de Retención") +
    theme_minimal()</pre>
```

`geom_smooth()` using formula = 'y ~ x'



Calculate the linear correlation coefficient
correlation <- cor(muestrasEmpresas\$margenUtilidad, muestrasEmpresas\$porcentRetencion)
print(paste("Coeficiente de correlación lineal:", correlation))</pre>

[1] "Coeficiente de correlación lineal: 0.244267414864451"

No hay nada que se pueda decir, realmente no hay una correlación fuerte entre las variables

Ejercicio 13

Método de momentos vs. método de percentiles.

Para una muestra de 12 siniestros se sabe que la media es 35,000, la desviación estándar 75,000, la mediana 10,000 y el percentil 90% es 100,000. Si la muestra proviene de una distribución Weibull (γ , θ), con

$$F(x) = 1 - e^{-\left(\frac{x}{\theta}\right)^{\gamma}}$$
 para $x \ge 0$

1. Utilice el método de momentos para estimar () y (). [10 puntos]

Primero notemos que los momentos importantes (primer momento y segundo momento central) de una Weibull son:

$$\mu = \theta \Gamma (1 + \frac{1}{\alpha})$$

$$\sigma^2 = \theta^2 \Gamma(1 + \frac{2}{\alpha}) - \mu^2$$

Por lo que desarrollando el sistema de ecuaciones tenemos que:

$$\sigma^2 + \mu^2 = \mu^2 \frac{\Gamma(1 + \frac{2}{\alpha})}{(\Gamma(1 + \frac{1}{\alpha}))^2}$$

Y siguiendo el método de momentos haremos que:

$$\mu = \bar{X}$$

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$$

Con lo cual procederemos a resolver el sistema de ecuaciones numéricamente:

```
library(rootSolve)

# Define the system of equations

eq_system <- function(vars) {
    # Define the variables
    alfa <- vars

f1 <- (gamma(1+2/alfa)*35^2)/(gamma(1+1/alfa)^2) - (11*75^2)/12 - 35^2
    return(f1)
}

# Initial estimates
lower_bound <- 0.1
upper_bound <- 100

# Find the root</pre>
```

```
results <- uniroot(f = eq_system, interval = c(lower_bound, upper_bound))
# Display results
cat("El valor de alfa es:", results$root, "\n")</pre>
```

El valor de alfa es: 0.532441

```
alfa <- results$root
theta <- 35/gamma(1+1/alfa)
cat("El valor de theta es:", theta, "\n")</pre>
```

El valor de theta es: 19.52449

2. Utilice el método de percentiles para estimar () y (). [10 puntos]

Notemos que al encontrar los cuantiles de una Weibull tenemos que:

$$Q(p) = \theta(-\ln(1-p))^{\frac{1}{\gamma}}$$

la cual fue encontrada al desarrollar la función de distribución acumulada de la Weibull a partir de esta ecuación:

$$F(x) = 1 - e^{-\left(\frac{x}{\theta}\right)^{\gamma}}$$
 para $x \ge 0$

Con lo cual, dados los datos que tenemos, procederemos a crear el siguiente sistema de ecuaciones para con el método de momentos:

$$Q(0.9) = 100$$

$$Q(0.5) = 10$$

Desarrollando el sistema obtenemos lo siguiente:

$$\gamma = \frac{ln(\frac{ln(2)}{ln(10)})}{-ln(10)}$$

$$\theta = \frac{10}{(\ln(2))^{\frac{1}{\gamma}}}$$

```
gamma <- -log(log(2)/log(10))/log(10)

theta <- 10/(log(2)^(1/gamma))

cat("El valor de gamma es:", gamma, "\n")</pre>
```

El valor de gamma es: 0.5213902

```
cat("El valor de theta es:", theta, "\n")
```

El valor de theta es: 20.19708

Ejercicio 14

Estimación por Máxima Verosimilitud

Suponga que (X_1,X_2,\dots,X_n) es una muestra aleatoria de $N(\mu,\sigma^2)$, donde $\mu\in\mathbb{R}$ y $\sigma>0$.

a) Encuentre los EMV de μ y σ^2 (donde $\sigma^2>0$). No olvide verificar que maximizan a la función de verosimilitud. [10 puntos]

La función de verosimilitud dada una muestra de tamaño n de X es:

$$\begin{split} L(\mu,\sigma^2) &= \prod_{i=1}^n f(x_i;\mu,\sigma^2) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} \\ &= (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{\sum (x_i-\mu)^2}{2\sigma^2}} \end{split}$$

De la cual obtenemos la función logarítmica de verosimilitud:

$$\ln L(\mu,\sigma^2) = -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{\sum(x_i-\mu)^2}{2\sigma^2}$$

Derivando respecto a μ e igualando a cero, obtenemos:

$$\frac{\partial \ln L(\mu, \sigma^2)}{\partial \mu} = \frac{\sum (x_i - \mu)}{\sigma^2} = 0$$

Con lo cual obtenemos que:

$$\frac{\sum x_i}{n} = \hat{\mu}$$

Derivando respecto a σ^2 e igualando a cero, obtenemos:

$$\frac{\partial \ln L(\mu, \sigma^2)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{\sum (x_i - \mu)^2}{2\sigma^4} = 0$$

Con lo cual obtenemos que:

$$\frac{\sum (x_i - \mu)^2}{n} = \hat{\sigma^2}$$

Los cuales serían los estimadores de máxima verosimilitud si verificamos que maximizan a la función de verosimilitud. Para esto notemos que la primera entrada de la diagonal de la matriz hessiana es:

$$\frac{\partial^2 \ln L(\mu, \sigma^2)}{\partial \mu^2} = -\frac{n}{\sigma^2} < 0$$

y que el determinante de la matriz hessiana es:

$$\det H(\mu, \sigma^2) = \frac{n}{\sigma^2} \frac{\sum (x_i - \mu)^2}{2\sigma^6} > = 0$$

Con lo cual tenemos que la matriz hessiana es negativa semidefinida en todo punto, por lo que el punto crítico es un máximo.

- b) Suponga X_i representa la pérdida del i-ésimo siniestro de cierto seguro.
- i) Una muestra de 100 pérdidas revela que 62 se ubicaron por debajo de 1,000 dólares y 38 por arriba de 1,000 dólares. Si se sabe que $\sigma = 200$, ¿cuál es el estimador de máxima verosimilitud de μ usando sólo esta información? [10 puntos]

Por el principio de máximo verosímil tenemos que la función de máxima verosimilitud es:

$$L(\mu) = P(x_i <= 1000)^{62} P(x_i > 1000)^{38}$$

Procederemos a maximar la función numéricamente:

```
library(stats)
# Define the system of equations

eq_system <- function(vars) {
    # Define the variables
    mu <- vars

    f1 <- 62*log(pnorm(1000, mu, 200)) + 38*log(1-pnorm(1000, mu, 200))
    return(f1)
}

# Use optimize to find the maximum over a reasonable range
mu <- optimize(eq_system, interval = c(1, 4000), maximum = TRUE)

# Display the results
cat("El valor de mu es:", mu$maximum, "\n")</pre>
```

El valor de mu es: 938.9038

ii) Si además se sabe que las pérdidas por debajo de 1,000 sumaron 42,561 dólares y la suma de sus cuadrados fue de 35,097,400, utilice esta información adicional a la del inciso anterior para estimar a (\$\$). [10 puntos]

Dado lo que sabemos y por el principio de máxima verosimilitud, tenemos que la función de máxima verosimilitud es:

$$L(\mu) = \prod_{i=1}^{62} (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{x_i - \mu}{2\sigma^2}} P(x_i > 1000)^{38}$$

Por lo que procederemos a maximizar la función numéricamente:

```
# Define the system of equations
x_barra <- 42561/62
x_cuadrada_suma <- 35097400
sigma_cuadrada <- 200^2
eq_system <- function(vars) {
    # Define the variables</pre>
```

```
mu <- vars

f1 <- - (62*(x_barra-mu)^2)/(2*sigma_cuadrada) + 38*log(1-pnorm(1000, mu, 200))

return(f1)
}

# Use optimize to find the maximum over a reasonable range

mu <- optimize(eq_system, interval = c(1, 10000), maximum = TRUE)

# Display the results
cat("El valor de mu es:", mu$maximum, "\n")</pre>
```

El valor de mu es: 849.5584