Universidad Nacional Autónoma de México.

Facultad de Ciencias.

Teoría del Riesgo
2021-2

TAREA-EXAMEN I

Profesor:
Salvador Zamora Muñoz

Ayudante:
Mónica Olivares Alvarado
Bruno Gómez Labra
Daniela Betancourt Cruz

Coautores:
Córdoba Ríos Ingrid
Flores Alonso Oscar David
González Eslava Rodrigo Ernesto
Varela López Ana Karen

23 de julio de 2021

1) Las pérdidas de cierta compañía siguen una distribución $Pareto(\alpha, \theta)$ con θ conocida. El décimo percentil es $\theta - k$, el percentil 90 es $5\theta - 3k$, determina el valor de α

   Sabemos que θ es conocido, por lo que al plantear el siguiente sistema de ecuaciones tenemos:

   $$F(x) = 1 - \left[\frac{\theta}{x+\theta}\right]^{\alpha}$$

   $$= \begin{cases} 1 - \left[\frac{\theta}{2\theta-k}\right]^{\alpha} = 0.1 \\ 1 - \frac{\theta}{6\theta-3k}\left[\right]^{\alpha} = 0.9 \end{cases}$$

   Que también podemos ver como:

   $$\left(\frac{\theta}{2\theta-k}\right)^{\alpha} = 0.9$$

   $$\left(\frac{\theta}{6\theta-k}\right)^{\alpha} = 0.1$$

   Supongamos que $k > 0$, pero $k < \theta$

   $$\left(\frac{\theta}{2\theta-k}\right)^{\alpha} = 0.9 \qquad (1)$$

   $$3^{-\alpha}\left(\frac{\theta}{2\theta-k}\right)^{\alpha} = 0.1 \qquad (2)$$

   Si dividimos la (1) entre (2)

   $$\frac{\left(\frac{\theta}{2\theta-k}\right)^{\alpha}}{3^{-\alpha}\left(\frac{\theta}{2\theta-k}\right)^{\alpha}} = 9$$

   $=> 3^{\alpha} = 9$

   $\alpha = 2$

2) Una muestra aleatoria de 1000 contratos de seguros de salud reportan una media muestral de 1300 por pago de beneficios anuales con una desviación estándar de 400. Se espera que 2500 contratos sean emitidos para el siguiente año. Usa el TCL para estimar la probabilidad de que el pago de los beneficios sea mayor que el 101 % del monto esperado.

Solución:

Teorema del límite central.- Se define la v.a. S como la suma de 'n' v.a.i.i.d. cada una con media μ y σ^2, es decir, $S_n = X_1 + ... + X_n = \sum_{i=1}^{n} X_i$. Entonces, $E[S_n] = n\mu$, $Var(S_n) = n\sigma^2$ y $Z_n = \frac{S_n - n\mu}{\sqrt{n}\sigma} \sim N(0,1)$

Ley de los grandes números.- Sea $X_1, ..., X_n$ una sución de v.a.i.i.d. con media finita μ, entonces $\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{n\to\infty} \mu$

Sea X_i la v.a. de pago de los beneficios del contrato i, por tanto, $S_{2500} = \sum_{i=1}^{2500} X_i$
Dado que tenemos una muestra lo suficientemente grande, por la LDGN, podemos suponer $\mu = 1300$ y $\sigma = 400$. Así, para las 2500 muestras obtenemos $E[S_{2500}] = 2500(1300) = 3,250,000$ y $Var[S_{2500}] = \sqrt{2500}(400) = 20,000$.

De acuerdo al TLC:

$$P[S_{2500} > 101\,\%\mu_{2500}] = 1 - P[Z_{2500} < \frac{1.01\mu_{2500} - \mu_{2500}}{\sigma_{2500}}] = 1 - P[Z_{2500} < \frac{0.01\mu_{2500}}{\sigma_{2500}}]$$
$$= 1 - \Phi(1.625) = 1 - 0.9479187$$
$$\therefore P[S_{2500} > 101\,\%\mu_{2500}] = 0.05208128$$

■

3) La v. a. X tiene una fdp $f(x) = (1 + 2x^2)e^{-2x}$, $x \geq 0$. Determina la $F(x)$, $S(x)$, $h(x)$ y $e(x)$.

a) F(t) = $\int_0^t (1 + 2x^2)e^{-2x}dx$
Por partes:

$$f = 2x^2 + 1 \qquad f' = 4x$$
$$g = -\frac{e^{-2x}}{2} \qquad g' = e^{-2x}$$

$$\Rightarrow F(t) = -\frac{(2x^2+1)e^{-2x}}{2} + 2\int_0^t xe^{-2x}dx$$

Nuevamente por partes para la antiderivada

$$f = x \qquad f' = 1$$
$$g = -\frac{e^{-2x}}{2} \qquad g' = e^{-2x}$$

$$\begin{aligned}\Rightarrow F(t) =& -\frac{(2x^2+1)e^{-2x}}{2} - \frac{2}{2}xe^{-2x} + \frac{2}{2}\int e^{-2x}dx \\ =& -\frac{(2x^2+1)e^{-2x}}{2} - xe^{-2x} - \frac{e^{-2x}}{2} \\ =& -(x^2+x+1)e^{-2x}|_0^t \\ =& -(t^2+t+1)e^{-2t} + e^0 \\ \therefore F(t) =& 1 - (t^2+t+1)e^{-2t}\end{aligned}$$

b) S(x) = 1 - F(x)

$$S(x) = 1 - F(x) = 1 - (1 - (x^2+x+1)e^{-2x}) = (x^2+x+1)e^{-2x}$$
$$\therefore S(x) = (x^2+x+1)e^{-2x}$$

■

c) h(x)= $\frac{f(t)}{s(t)}$

$$\begin{aligned}h(x) =& \frac{f(t)}{s(t)} \\ =& \frac{(1+2x^2)e^{-2x}}{(x^2+x+1)e^{-2x}} \\ \therefore h(x) =& \frac{1+2x^2}{x^2+x+1}\end{aligned}$$

■

d) $\frac{\int_d^{\infty} S(x)}{S(d)} dx$

$$\Rightarrow \int S(x) = \int_d^{\infty} (x^2 + x + 1)e^{-2x} dx$$

$$f = x^2 + x + 1 \qquad f' = 2x + 1$$
$$g = -\frac{e^{-2x}}{2} \qquad g' = e^{-2x}$$

$$\Rightarrow \int S(x) = -\frac{(x^2 + x + 1)}{2} e^{-2x} + \frac{1}{2} \int (2x + 1)e^{-2x} dx$$

Por partes para la siguiente integral:

$$f = 2x + 1 \qquad f' = 2$$
$$g = -\frac{e^{-2x}}{2} \qquad g' = e^{-2x}$$

$$\begin{aligned}
&= -\frac{x^2 + x + 1}{4} e^{-2x} - \frac{1}{2}(\frac{(2x + 1)}{2} e^{-2x} + \int e^{-2x} dx) \\
&= -\frac{2(x^2 + x + 1)}{4} e^{-2x} - \frac{(2x + 1)e^{-2x}}{4} - \frac{e^{-2x}}{4} \\
&= \frac{(x^2 + 2x + 2)e^{-2x}}{2}
\end{aligned}$$

Ahora para:

$$\begin{aligned}
\int_d^{\infty} S(d) &= -\frac{(x^2 + 2x + 2)e^{-2x}}{2} \Big|_d^{\infty} \\
&= -0 + \frac{(d^2 + 2d + 2)}{2} e^{-2d}
\end{aligned}$$

Entonces:

$$\begin{aligned}
\Rightarrow e(d) &= \frac{\frac{(d^2+2d+2)}{2} e^{-2d}}{(d^2 + d + 1)e^{-2d}} \\
\therefore e(d) &= \frac{d^2 + 2d + 2}{2(d^2 + d + 1)}
\end{aligned}$$

■

4) Para los datos A y B realiza lo siguiente:

Datos A

25	457	82	680	115	855	126	877	155	974
161	1193	243	1340	294	1884	340	2558	384	3476

Realiza pruebas de hipótesis para un modelo exponencial y gamma y compara los parámetros estimados.

Datos B

3	11	27	36	47	49	54	77	78	85
104	121	130	138	139	140	143	153	193	195
205	207	216	224	233	237	254	257	259	265
273	275	278	281	396	405	412	423	436	456
473	475	503	510	534	565	656	656	716	734
743	756	784	786	819	826	841	842	853	860
877	942	942	945	998	1029	1066	1101	1128	1167
1194	1209	1223	1283	1288	1296	1310	1320	1367	1369
1373	1382	1383	1395	1436	1470	1512	1607	1699	1720
1772	1780	1858	1922	2042	2247	2348	2377	2418	2795
2964	3156	3858	3872	4084	4620	4901	5021	5331	5771
6240	6385	7089	7482	8059	8079	8316	11453	22274	32043

Ajusta una densidad de acuerdo al análisis descriptivo y tu experiencia.

Solución A:

Primero hagamos un histograma con su respectiva densidad de los datos:

Histograma y densidad tabla A

En el siguiente gráfico se verifica la razón de que las pruebas de hipótesis se vayan a hacer para una gamma y una exponencial:

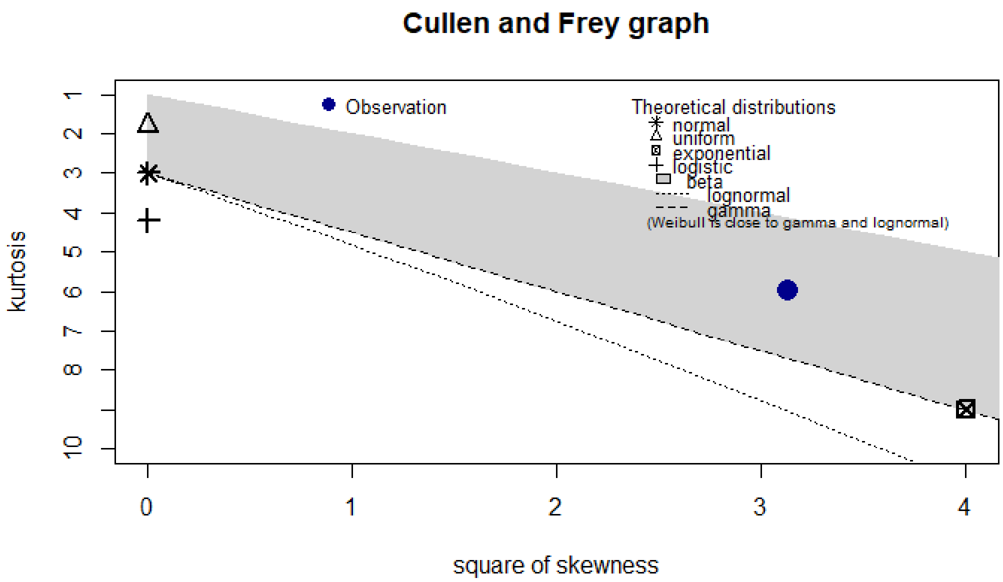


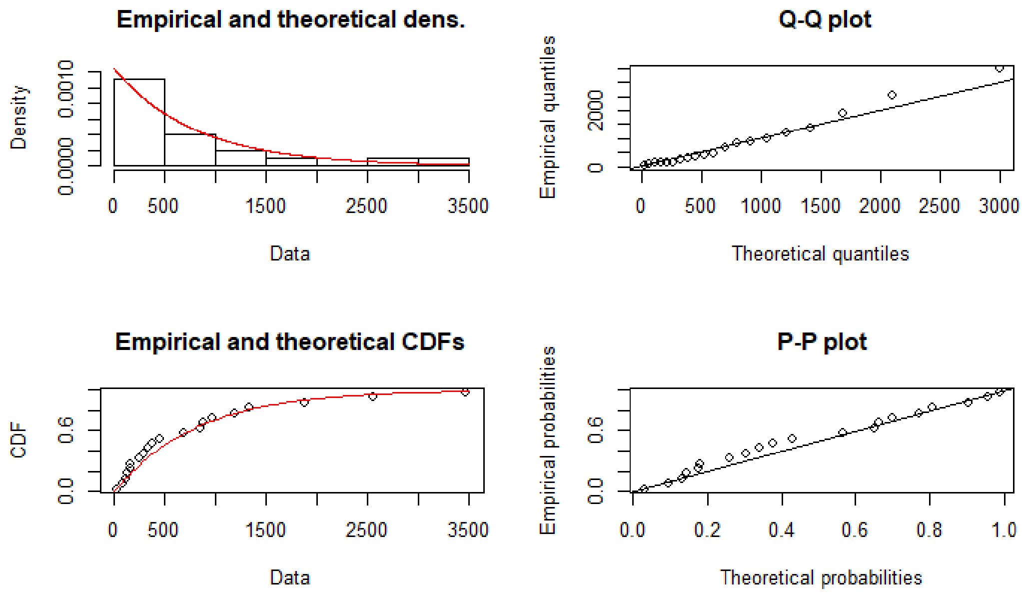
Gráfico Cullen & Fray para tabla A

Como podemos observar debido a la kurtosis (1.5187) y la asimetría (1.5128) nuestra observasión queda dentro del rango de una distribución beta, gamma y exponencial, debido a la naturaleza de nuestros datos descartamos de inmediato beta. Ahora ajustemos ambas distribuciones para poder obtener un estimado de nuestros parametros.

Haciendo uso de la funcion $fitdistr()$ de la paqueteria $fitditrplus$ de R obtuvimos el siguiente parametro para ajustar un modelo exponencial a nuestros datos:

```
Fitting of the distribution ' exp ' by matching moments
Parameters:
        estimate
rate 0.001233122
```

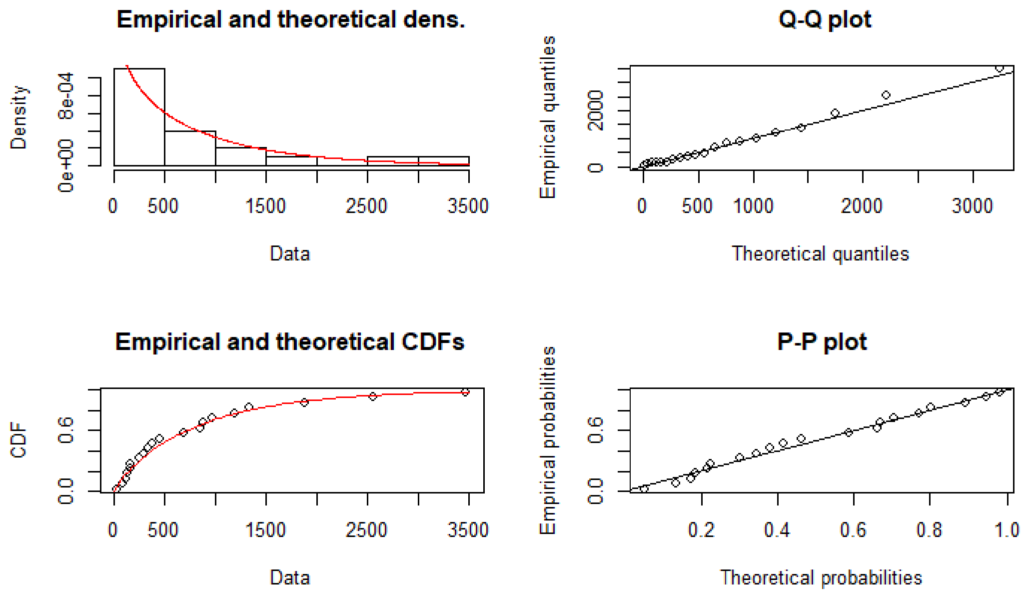
En el siguiente gráfico observamos el ajuste de la muestra a la distribución empírica por el método de igualación de momentos:



Ajuste de los datos a un modelo exponencial para tabla A

Hagamos lo mismo para la gamma:

```
Fitting of the distribution ' gamma ' by matching moments
Parameters:
          estimate
shape 0.829491299
rate  0.001022864
```



Ajuste de los datos a un modelo gamma para tabla A

Ahora que contamos con los parámetros estimados procederemos a hacer nuestras pruebas de hipótesis.

Prueba de Kolmogoróv-Smirnov para la exponencial y la gamma:

Sea $X =$ *la distribución que siguen las observaciones de la tabla A*

$$H_0: \; X \sim exp(0.001233122) \;\; vs \;\; Ha: \; X \nsim exp(0.001233122), \quad \alpha = 0.05$$

```
 One-sample Kolmogorov-Smirnov test

data:  x
D = 0.12281, p-value = 0.8882
alternative hypothesis: two-sided
```

$\rightarrow$ No se rechaza H_0 a un nivel de significancia $\alpha = 0.05$. En otras palabras no existe suficiente información estadísitica para rechazar Ho.

$$\boxed{\therefore X \sim exp(0.001233122)}$$

Sea $X =$ *la distribución que siguen las observaciones de la tabla A*

$$H_0: \; X \sim gamma(0.829491299, 0.001022864) \; vs \; Ha: \; X \nsim gamma(0.829491299, 0.001022864)$$
$$\alpha = 0.05$$

```
	One-sample Kolmogorov-Smirnov test
```

```
data:  x
D = 0.10519, p-value = 0.963
alternative hypothesis: two-sided
```

$\rightarrow$ No se rechaza H_0 a un nivel de significancia $\alpha = 0.05$. En otras palabras no existe suficiente información estadísitica para rechazar Ho.

$$\boxed{\therefore X \sim gamma(0.829491299, 0.001022864)}$$

Prueba de Anderson-Darling para la exponencial y la gamma:

Sea $X =$ *la distribución que siguen las observaciones de la tabla A*

$$H_0:\ X \sim exp(0.001233122)\ \ vs\ \ Ha:\ X \nsim exp(0.001233122), \quad \alpha = 0.05$$

```
 Anderson-Darling GoF Test

data:  x  and  pexp
AD = 0.32473, p-value = 0.9176
alternative hypothesis: NA
```

$\rightarrow$ No se rechaza H_0 a un nivel de significancia $\alpha = 0.05$. En otras palabras no existe suficiente información estadísitica para rechazar Ho.

$$\boxed{\therefore X \sim exp(0.001233122)}$$

Sea $X =$ *la distribución que siguen las observaciones de la tabla A*

$$H_0:\ X \sim gamma(0.829491299, 0.001022864)\ vs\ Ha:\ X \nsim gamma(0.829491299, 0.001022864)$$
$$\alpha = 0.05$$

```
 Anderson-Darling GoF Test

data:  x  and  pgamma
AD = 0.23465, p-value = 0.9779
alternative hypothesis: NA
```

$\rightarrow$ No se rechaza H_0 a un nivel de significancia $\alpha = 0.05$. En otras palabras no existe suficiente información estadísitica para rechazar Ho.

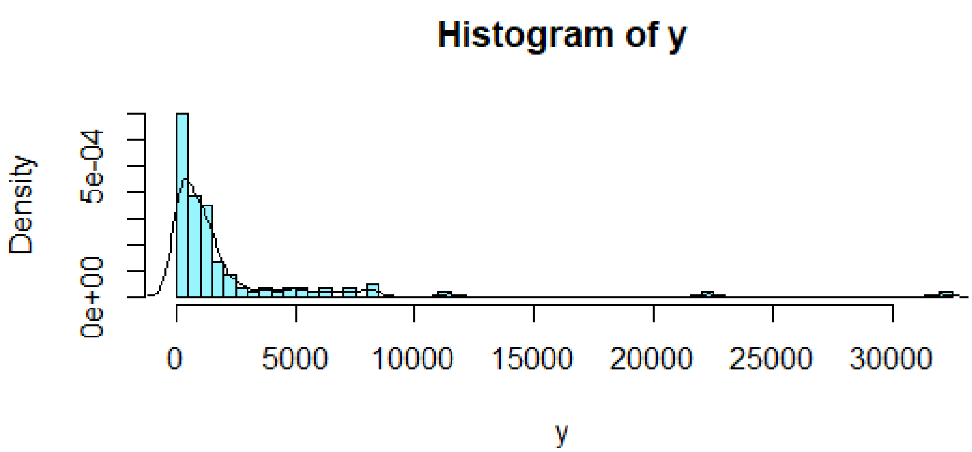
$$\boxed{\therefore X \sim gamma(0.829491299, 0.001022864)}$$

Como pudimos observar nuestra muestras se ajusta a los respectivos modelos gamma y exponencial. EL código de esta sección se encuentra en el Anexo.

Solución B: Primero obtengamos algunas medidas descriptivas de los datos en cuestión:

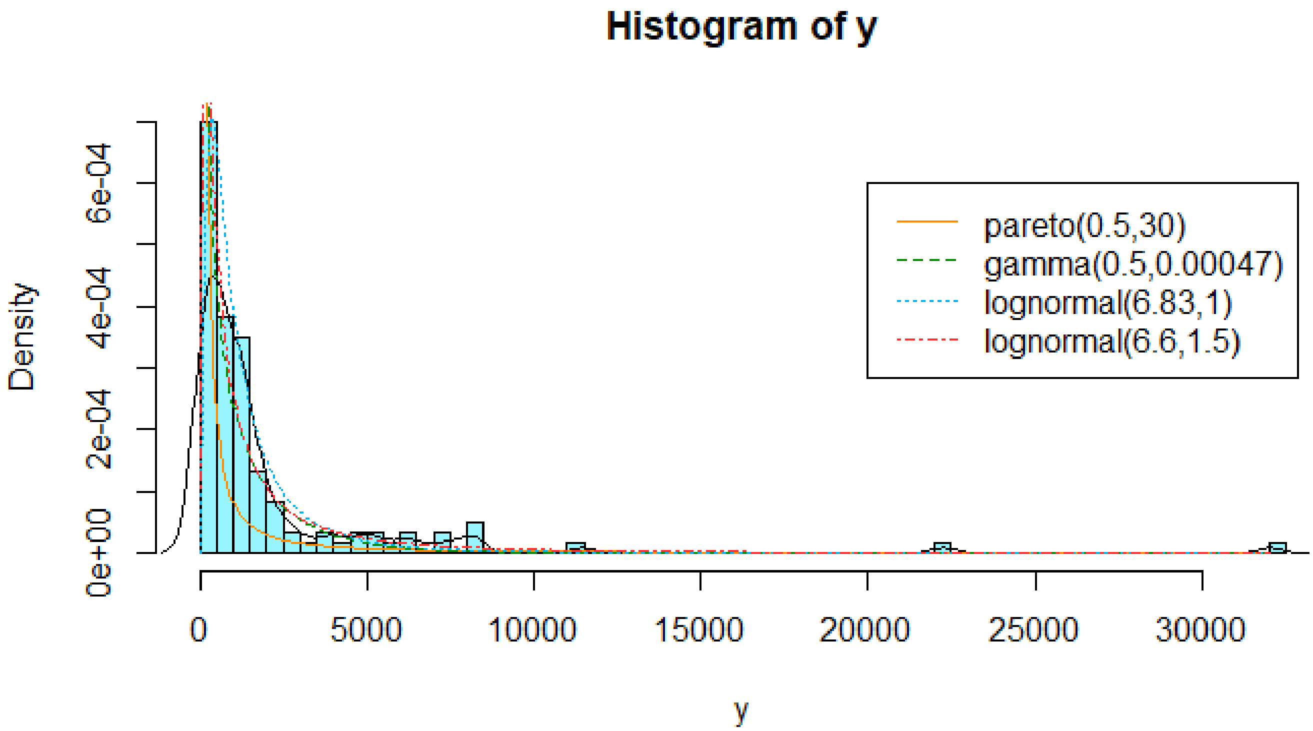
```
> #Medidas descriptivas#
> summary(y)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
    3.0   271.0   868.5  2020.3  1733.0 32043.0 
> tail(y)
[1]  8059  8079  8316 11453 22274 32043
> quantile(y)
     0%     25%     50%     75%    100% 
    3.0   271.0   868.5  1733.0 32043.0 
> var(y)
[1] 15601373
> sd((y))
[1] 3949.857
> range(y)
[1]     3 32043
> median(y)
[1] 868.5
> mfv(y)
[1] 656 942
> #sesgo y curtosis
> skew(y) #Asimétirca hacia la derecha
[1] 5.033963
> kurtosi(y)#Leptocúrtica
[1] 31.10723
```

Lo que rescataremos de los datos obtenidos será que estamos tratando con una muestra de datos que se pueden ajustar a una distribución asimétrica hacia la derecha y leptocúrtica, con media 2020.3, varianza 15601373 y desviación estándar 3949.857. Hagamos un histograma de nuestros datos y grafiquemos su densidad: Ahora tratemos de ajustar algunas distribuciones manualmente:



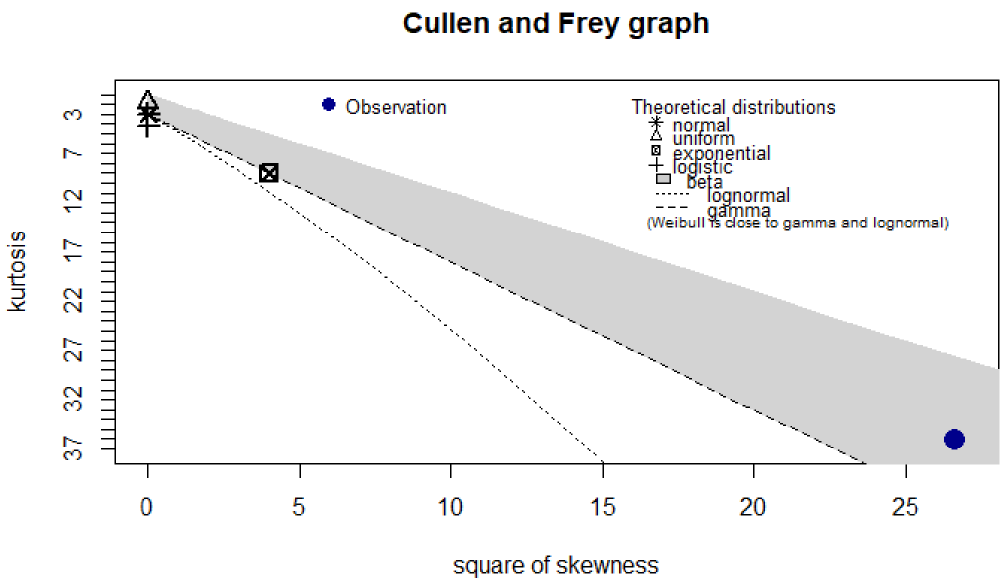
Histograma y densidad para tabla B

Cómo podemos ver tal parece que una lognormal podría ser una distribución adecuada para nuestros datos, por lo que ahora haremos uso del diagrama de Cullen and Fray para darnos una



Ajuste manual de densidades

mejor idea de a que distribución podemos ajustar nuestros datos.

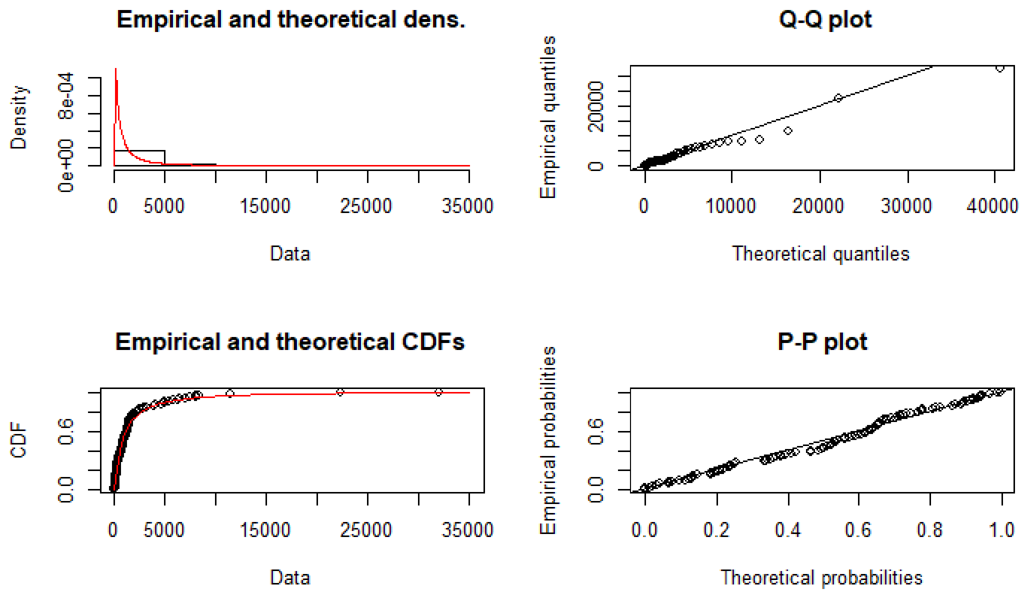


Cullen and Fray para tabla B

EL gráfico nos sugiere que ajustemos una beta, pero la naturaleza de nuestros parámetros no lo permiten, recordemos que el dominio de una beta esta entre 0 y 1, ahora bien, tratamos de ajustar una gamma pero con los parámetros que obtuvimos de R no logró sobrevivir a la bondad de ajuste ni por KS ni AD, por lo que optamos por una lognormal, para ello presentamos los parametros estimados por máxima verosimilitud:

```
        Fitting of the distribution ' lnorm ' by maximum likelihood
Parameters:
        estimate Std. Error
meanlog 6.624172 0.13795726
sdlog   1.511246 0.09755032
```

Los respectivos gráficos del ajuste:



Ajuste lognormal para tabla B

Ahora bien, decidimos ajustar una lognormal(6.624172,1.511246), hagamos las respectivas prueba de hipótesis para verificar esto:

Prueba de Kolmogoróv-Smirnov para ajustar una lognormal

Sea $X =$ *la distribución que siguen las observaciones de la tabla B*

$H_0: \ X \sim lognormal(6.624172, 1.511246) \ vs \ Ha: \ X \nsim lognormal(6.624172, 1.511246) \ \alpha = 0.05$

```
        One-sample Kolmogorov-Smirnov test

data:  y
D = 0.086673, p-value = 0.3282
alternative hypothesis: two-sided
```

$\rightarrow$ No se rechaza H_0 a un nivel de significancia $\alpha = 0.05$. En otras palabras no existe suficiente información estadísitica para rechazar Ho.

$$\boxed{\therefore X \sim lognormal(6.624172, 1.511246)}$$

Prueba de Anderson-Darling para ajustar una lognormal

Sea $X =$ *la distribución que siguen las observaciones de la tabla B*

$H_0: \ X \sim lognormal(6.624172, 1.511246) \ vs \ Ha: \ X \nsim lognormal(6.624172, 1.511246) \ \alpha = 0.05$

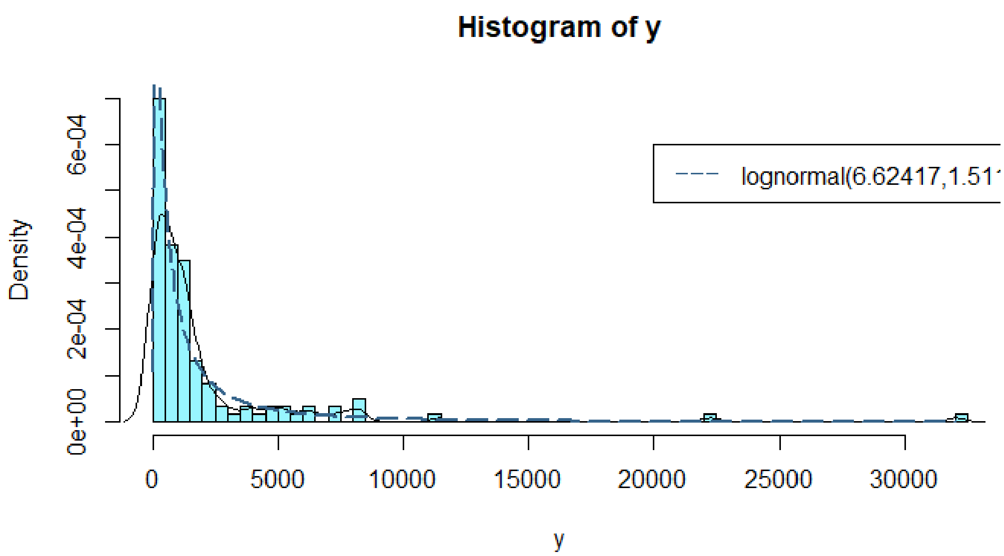
```
        Anderson-Darling GoF Test
```

```
data:  y and  plnorm
AD = 0.69679, p-value = 0.5612
alternative hypothesis: NA
```

$\rightarrow$ No se rechaza H_0 a un nivel de significancia $\alpha = 0.05$. En otras palabras no existe suficiente información estadísitica para rechazar Ho.

$$\boxed{\therefore X \sim lognormal(6.624172, 1.511246)}$$

Por lo tanto una de las distribucione que consideramos como posibles para el ajuste de nuestros datos es la $lognormal(6.624172, 1.511246)$.Por último el gráfico de nuestros datos con la ditribución ajustada: El código de esta sección se encuentra en el Anexo.



Ajuste lognormal para tabla B

5) Si $X \sim Gamma(\alpha = 0.2, \theta = 500\ )$ y $Y \sim ParetoII(\alpha = 2.5, \theta = 150)$, demuestra que tienen la misma esperanza y que la cola de la distribución pareto es más pesada que la de la gamma. Grafica las densidades y compara.

Tenemos que la esperanza de x es $\mathbb{E}[X] = \alpha\theta$. Y la esperanza de Y es $\mathbb{E}[Y] = \frac{\theta}{\alpha-1}$, entonces:

$$\mathbb{E}[X] = \mathbb{E}[Y]$$

$$0.2(500) = \tfrac{150}{1.5}$$

$$100 = 100\ \square$$

Por lo tanto, Y y X tienen la misma esperanza

Para ver cual distribución tiene la cola más pesada tomamos el siguiente criterio de comparación

$\displaystyle\lim_{x\to\infty} \frac{f_1(x)}{f_2(x)} \to \infty$ entonces $f_1(x)$ tiene una cola más pesada

$$\lim_{x\to\infty} \frac{\alpha\theta^{\alpha}x\Gamma(\alpha)}{(\frac{x}{\theta})^{\alpha}e^{-\alpha/\theta}(x+\theta)^{\alpha+1}} = \lim_{x\to\infty} \frac{2.5(150)^{2.5}x\Gamma(0.2)}{(\frac{x}{500})^{0.2}e^{-x/500}(x+150)^{3.5}}$$

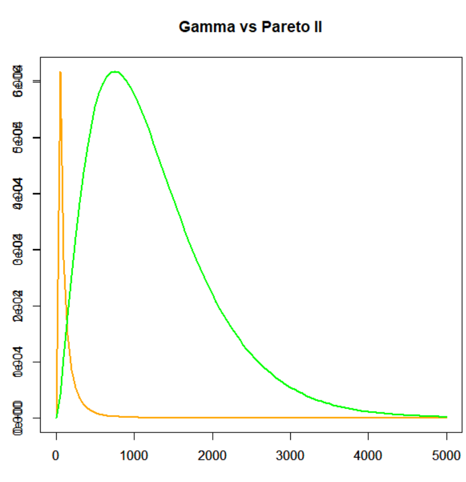
Sea $(500)^{0.2}\Gamma(0.2)(2.5)(150)^{2.5} = k$

$$k \lim_{x\to\infty} \frac{xe^{x/500}}{x^{0.2}(x+150)^{3.5}} = k \lim_{x\to\infty} \frac{xe^{x/500}}{x^{0.2}(x+150)^{3}.5}$$

$$= k \lim_{x\to\infty} \frac{x^{0.8e^{\alpha/500}}}{(x+150)^{3.5}} = \infty$$

Ya que el numerador crece más rápido que el denominador, entonces la *Pareto II* tiene una cola más pesada

Mismo resultado que se puede observar en la siguiente gráfica.



6) Si x tiene una función media de exceso de pérdida e(x) = x + 1, x ¿ 0. Encuentra f(x), F(x) y S(x).

Solución:
PD. $S(x) = \frac{e(0)}{e(x)} e^{-\int_0^x \frac{1}{e(y)} dy}$
Consideremos:

$$e(t) = \frac{\int_t^\infty S(x)dx}{S(t)} \quad \Rightarrow \quad k(t) = \int_t^\infty S(x)dx = S(t)e(t)$$

Sea

$$\int_0^x \frac{1}{e(y)} dy = -\int_0^x d(log(k(t)) = -log(k(t))|_0^x = -log\left(\frac{k(x)}{k(0)}\right) = -log\left(\frac{S(x)e(x)}{e(0)}\right) \quad S(0) = 1$$

$$\therefore S(x) = \frac{e(0)}{e(x)} e^{-\int_0^x \frac{1}{e(y)} dy} \tag{1}$$

Sustituyendo en (1):

$$S(x) = \frac{1}{x+1} e^{-\int_0^x \frac{1}{x+1} dy} = \frac{1}{x+1} e^{-ln(x+1)} = \frac{1}{(x+1)^2}$$

$$\Rightarrow F(x) = 1 - \frac{1}{(x+1)^2}$$

$$\Rightarrow F'(x) = f(x) = \frac{2}{(x+1)^3}$$

$$\therefore S(x) = \frac{1}{(x+1)^2} \qquad F(x) = 1 - \left(\frac{1}{x+1}\right)^2 \qquad f(x) = \frac{2}{(x+1)^3}$$

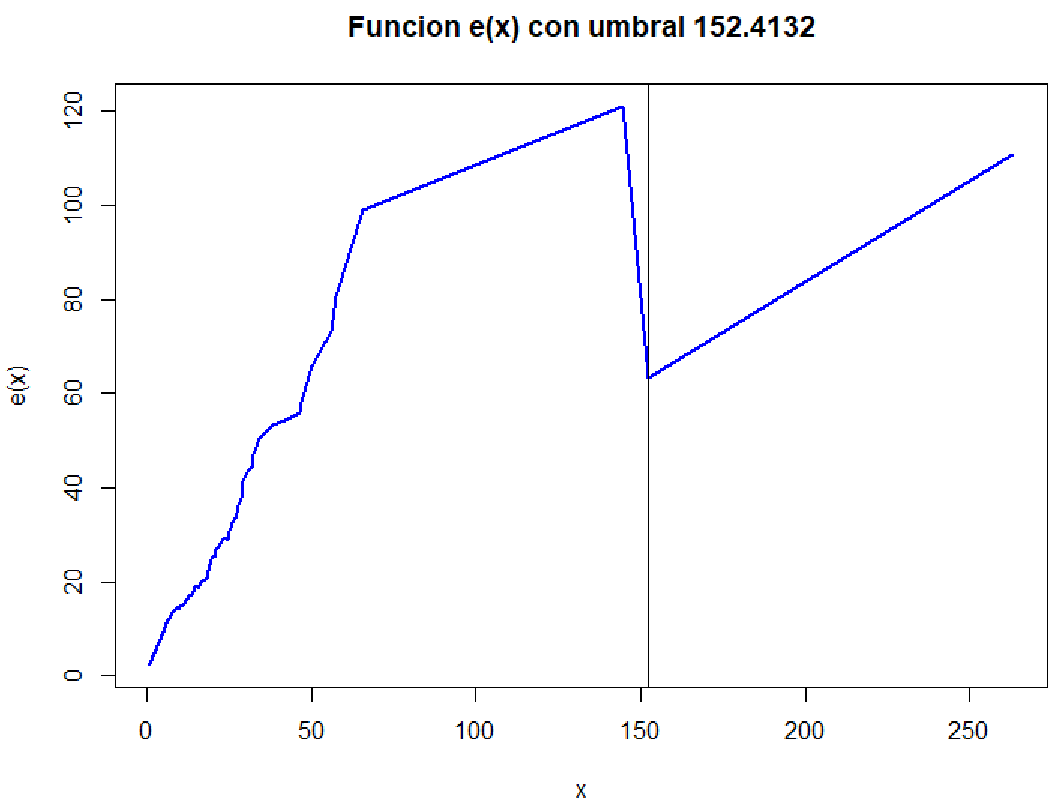
■

7) Para los datos *danishuni* de la paquetería *fitdistrplus*, grafica la función media de exceso de pérdida y elige el umbral adecuado a partir del cual ajustarías un densidad Pareto Generalizada. Extra: Ajusta la Pareto Generalizada y realiza pruebas de hipótesis.

Sea

$$\mathrm{E}\left[X - U | X > U\right] \quad \text{y} \quad \hat{e}_x(x) = \frac{\displaystyle\sum_{i=1}^{n} (X_{i:n} - u)^+}{\displaystyle\sum_{i=1}^{n} 1(X_{i:n} > u)}$$

Ordenando y graficando obtenemos la siguiente gráfica, donde vien podríamos considerar un valor alrededor de 150. En este caso decidimos considerar un umbral de 152. Sin embargo, es importante mencionar que a partir del valor 150 solo se tiene otro dato más.



Umbral al 152.41

8) Para 2 variables aleatorias X, Y , la función media de exceso de pérdida es $e_y(30) = e_x(30) + 4$. Sea $X \sim U(0, 100)$ y $Y \sim U(0, W)$, calcula W.

Sabemos

$$e_x(u) = \mathbb{E}[X - u|X > u] = \frac{\int_u^\infty S(x)dx}{S(u)}$$

$$S(x) = 1 - F(x) \qquad f(x) = \begin{cases} 0 & si \quad x < a \\ \frac{x-a}{b-a} & si \quad a \leq x < b \\ 1 & si \quad x \geq b \end{cases}$$

Para $X \sim U(0, 100)$

$$F(x) = \frac{x-0}{100-0} = \begin{cases} 0 & si \quad x < 30 \\ \frac{x}{100} & si \quad 30 \leq x < 100 \\ 1 & si \quad x \geq 100 \end{cases}$$

Para $Y \sim U(0, W)$

$$F(x) = \frac{y-0}{W-0} = \begin{cases} 0 & si \quad y < 0 \\ \frac{y}{W} & si \quad 0 \leq y < W \\ 1 & si \quad y \geq W \end{cases}$$

Calculemos $e_x(u)$

$$\begin{aligned} e_x(30) =& \frac{\int_{30}^{100} 1 - \frac{x}{100} dx}{1 - \frac{30}{100}} \\ =& \frac{\int_{30}^{100} dx - \frac{1}{100}\int_{30}^{100} xdx}{\frac{100-30}{100}} \\ =& \frac{x|_{30}^{100} - \frac{1}{100}(\frac{x^2}{2}|_{30}^{100})}{\frac{100-30}{100}} \\ =& \frac{x|_{30}^{100} - \frac{1}{100}(\frac{x^2}{2}|_{30}^{100})}{\frac{100-30}{100}} \\ =& \frac{(100-30) - \frac{1}{100(2)}(100^2 - 30^2)}{\frac{100-30}{100}} \\ =& \frac{\frac{2(100^2)-30(100)(2)-100^2+30^2}{2(100)}}{\frac{100-30}{100}} \\ =& \frac{\frac{(100-30)^2}{2}}{\frac{100-30}{1}} \\ =& \frac{100-30}{2} \\ \to e_x(30) =& 35 \end{aligned}$$

$$
\begin{aligned}
e_y(30) =& \frac{\int_{30}^{W} 1 - \frac{y}{W} dy}{1 - \frac{30}{W}} \\
=& \frac{\int_{30}^{W} dy - \frac{1}{W}\int_{30}^{W} ydy}{\frac{W-30}{W}} \\
=& \frac{(W-30) - \frac{1}{W}(\frac{W^2-30^2}{2})}{\frac{W-30}{W}} \\
=& \frac{\frac{2W^2-60W-W^2+30^2}{2W}}{\frac{W-30}{W}} \\
=& \frac{\frac{W^2-60W+30^2}{2}}{\frac{W-30}{1}} \\
=& \frac{\frac{(W-30)^2}{2}}{\frac{W-30}{1}} \\
\rightarrow e_y(30) =& \frac{W-30}{2}
\end{aligned}
$$

Sustituyamos en la ecuación que nos dieron:

$$
\begin{aligned}
e_y(30) =& e_x(30) + 4 \\
\frac{W-30}{2} =& 35 + 4 \\
W - 30 =& 39 * 2 \\
W =& 78 + 30 \\
\therefore W =& 108
\end{aligned}
$$

9) Las pérdidas siguen una distribución $Pareto(\alpha = 0.5, \theta = 10000)$. Evalua la función media de exceso de pérdida en θ.

Para obtener la función media de exceso de pérdida, se utiliza la función de supervivencia dada por:

$$S_X(t) = [\frac{\theta}{\theta + t}]^\alpha I(t > 0); \alpha > 0$$

De donde:

$$e_X(d) = \frac{\int_d^\infty S_X(t)dt}{S_X(d)} = \frac{\theta^\alpha \int_d^\infty (t+\theta)^{-\alpha} dt}{[\theta/(d+\theta)]^\alpha} = (d+\theta)^\alpha \int_d^\infty (t+\theta)^{-\alpha} dt$$

Notemos que la integral converge solo si $\alpha > 1$, por lo que la función media de exceso de pérdida para una Pareto con $\alpha = 0.5$ no existe ya que $\alpha < 1$.

10) Calcula el VaR y T V aR al 90 % para las siguientes distribuciones:

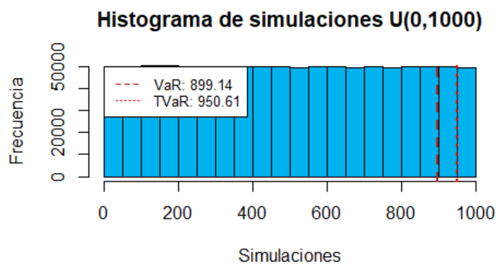
a) U(0, 1000)

b) $Pareto(\alpha = 1.6, \theta = 2)$

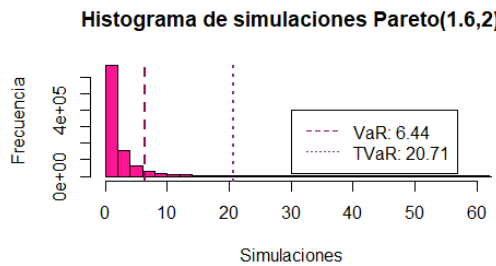
c) Exp(media = 1000)

d) $Lognormal(\mu = 5, \sigma = 2)$

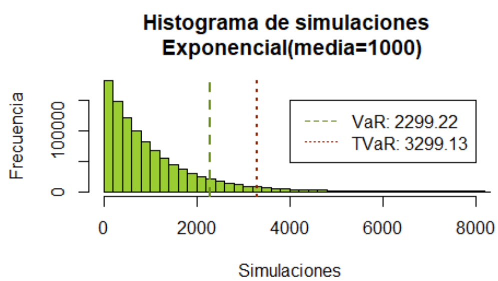
e) $N(\mu = 1000, \sigma = 1000)$



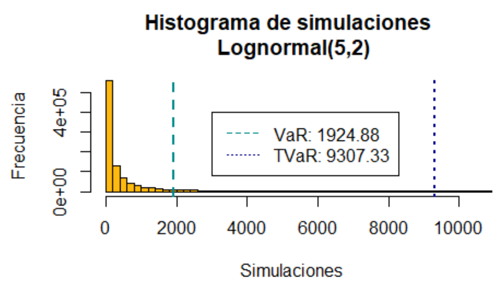
(a) VaR $= 899.1412 \cdot VaR_{exacto} = 900 \cdot$ TVaR = 950.6084



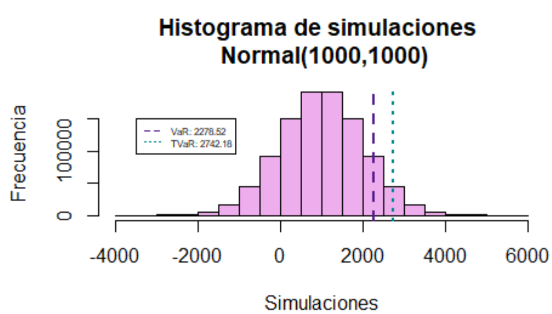
(b) VaR $= 6.437328 \cdot VaR_{exacto} = 6.43393 \cdot$ TVaR = 20.70896



(c) VaR $= 2299.219 \cdot VaR_{exacto} = 2302.585 \cdot$ TVaR = 3299.127



(d) VaR $= 1924.881 \cdot VaR_{exacto} = 1925.812 \cdot$ TVaR = 9307.325



(e) VaR $= 2278.525 \cdot VaR_{exacto} = 2281.552 \cdot$ TVaR = 2742.177

11) Realiza un análisis descriptivo a la variable *Losses.in.Thousands* de la base *RModuleDay.csv*. Ajusta la densidad adecuada, calcula el VaR y $TVaR$ real al 99 %, el estimado (de acuerdo al modelo ajustado) y compara.

Solución:
La base está compuesta por 15,290 observaciones y 7 variables o características: identificador, edad, años de experiencia, numro d vehículos, género, estado civil y pérdidas en miles. En particular será de interés analizar esta última variables, las pérdidas.

Análisis descriptivo:
Consideremos que nuestros datos están en miles.

NA	Mínimo	Mediana	Media	Máximo	Q_{25}	Q_{75}
0	12.54	354.94	389.86	3500.00	226.43	488.68

La pérdida media se encuentra alrededor de 389.36 u.m., alcanzando un monto máximo de hasta 3500 u.m. y un mínimo de 12.54 u.m., es decir, la volatilidad de los datos es alta pues obtenemos un rango igual a 3487.46. Por otro lado, la mediana de las pérdidas se encuentra por dabajo del monto esperado por lo que se debe bensar en una distribución sesgada a la derecha. Finalmente, cabe señalar que la falta de datos no seráun problema pues contamos con una base completa.

A continuación se muestra el histograma de nuestra variable de interés:

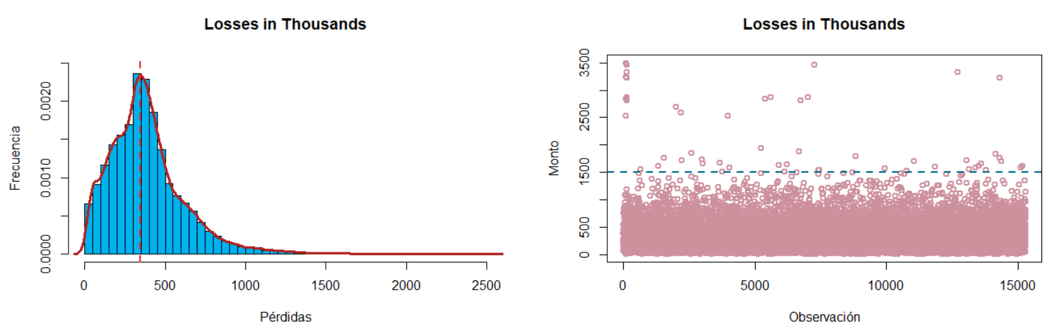
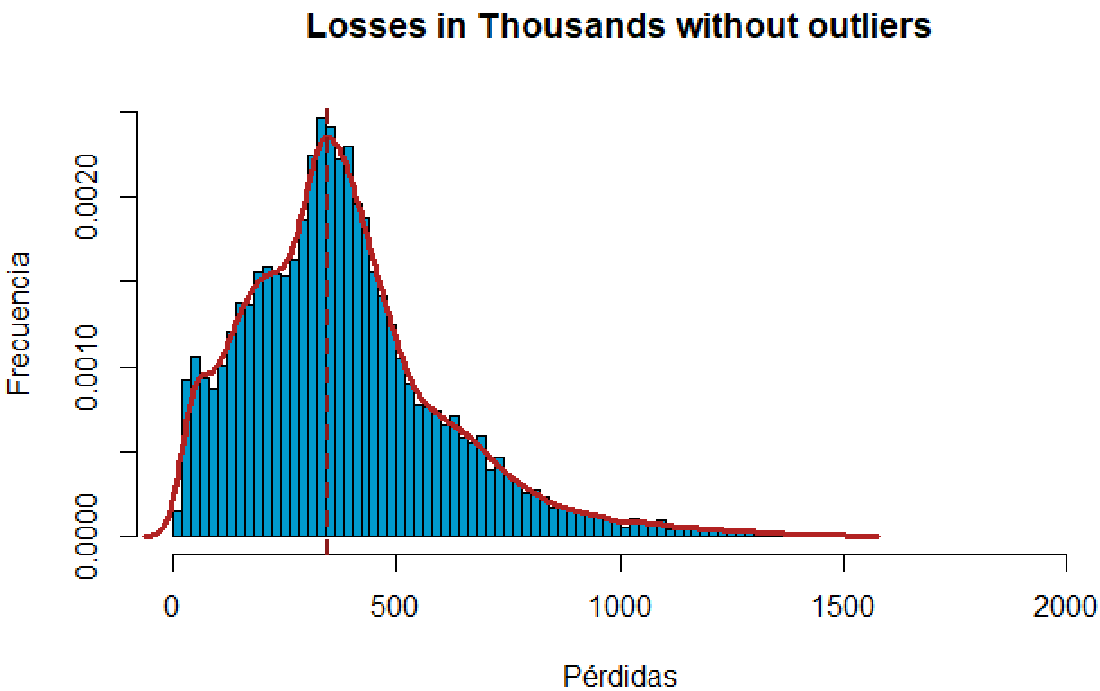


Figura 2: Losses in thousands

En el gráfico a la izquierda podemos observar el histograma de los montos de pérdida, como se esperaba notamos un sesgo a la derecha igual a 1.095839, con moda igual al monto de 345.5913 u.m. En el lado derecho encontramos un gráfico de puntos donde se ilustra el monto de cada observación, los cuales se encuentran por debajo de 1500 u.m. monetarias, superando este monto, del total de la cartera, solo el 0.3675988 % de las pérdidas. Para obtener un mejor ajuste, y dado que los datos atípicos no son significativos respecto a nuestra información total, retirarémos 56 observaciones outliers. Cabe señalar que se le podría ajustar un modelo generalizado pareto a todas aquellas observaciones por arriba de nuestro umbral de 1500 u.m., sin embargo, este análisis se enfocará en los datos más significativos.



Si bien, es de suma importancia considerar aquellas severidades altas pues podrían desestabilizar la liquidez de la empresa, al remover los datos atípicos notamos que la moda de las pérdidas se sigue conservado alrededor del monto previo, en este caso, $moda = 345.5099$.

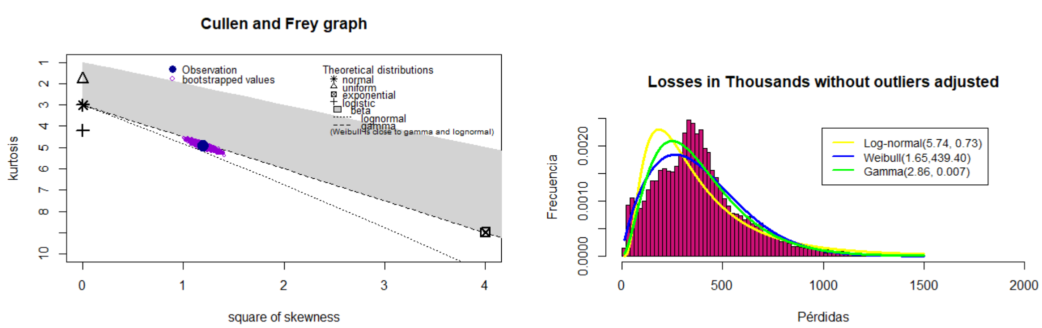


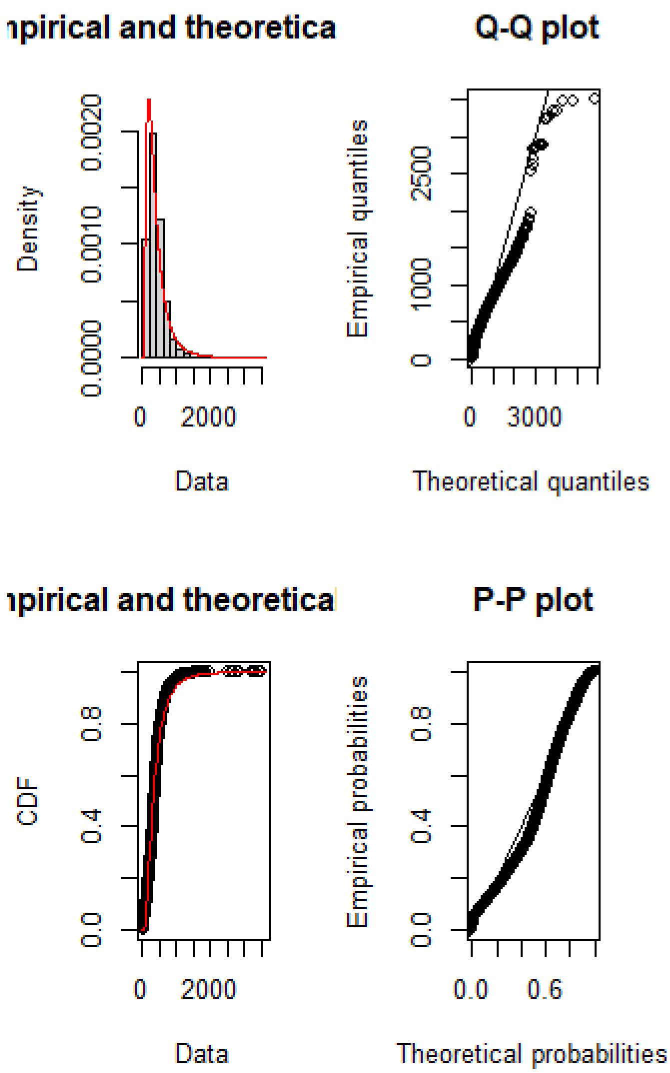
Figura 3: Losses in thousands

Considerando la estimación del gráfico de Cullen a la izquierda, y que es la que se ajusta, físicamente, mejor a nuestros datos. Se propone un modelo lognormal.

Aplicando el método de momentos tenemos:

$\mu = 5.789713$

$\sigma^2 = 0.3531482$



Calculamos el VaR y TVaR con nuetros datos:

- VaR = 1027.005
- TVaR = 1177.734

VaR y TVaR estimados:

- VaR = 1025.4458
- TVaR = 1447.970

Podemos observar que mediante ambos caminos se llega a una pérdida máxima muy similiar, sin embargo, las pérdidas medias esperadas estimmadas están por arriba del dato real.

Nota. Es claro que se debe mejorar el modelo propuesto.

■

12) Para los siguientes conjuntos de datos ajusta un modelo discreto $(a, b, 0)$

Datos C

Reclamaciones	Frecuencia
0	65623
2	12571
2	1644
3	148
4	13
5	1
6	0
Total	80000

Datos D

Reclamaciones	Pólizas
0	861
1	121
2	13
3	3
4	1
5	0
6	1
7+	0
Total	1000

Solución datos C:

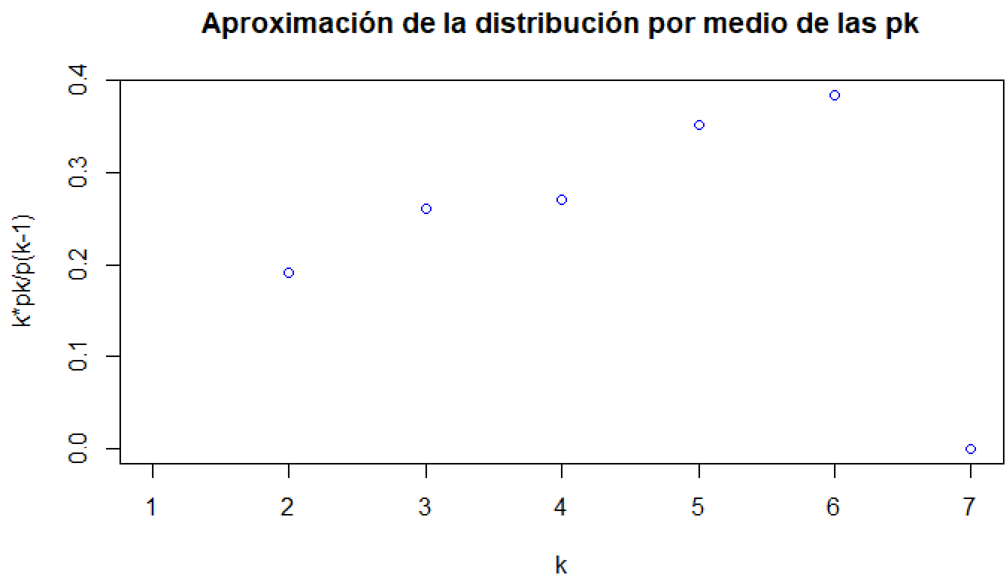
Primero utilicemos la siguiente ecuación para crear un gráfico que como vimos en clase nos ayudará a determinar que tipo de modelos podemos ajustar por medio de la pendiente de la ecuación lineal:

$$\frac{kp_k}{p_{k-1}} = ak + b, \qquad k = 1, 2, 3, ...$$

Para nuestros datos obtenemos los siguientes resutlados

```
> #Utilizamos la formula k*p_k/p_k_-1
> pt1<-c()
> for(i in 2:7){
+   pt1[i]<-T1[i,1]*T1[i,2]/T1[i-1,2]
+ }
> pt1
[1]        NA 0.1915639 0.2615544 0.2700730 0.3513514 0.3846154 0.0000000
```

Graficamos nuestros resultados



Aproximación de un modelo discreto Datos C

Lo que nos sugiere un modelo binomial negativo, calulemos la esperanza y la varianza de nuestros datos para proponerlos cómo parametros de nuestra binomial negativa:

$$\mathbb{E}[X] = 0.2045 \quad \& \quad \mathbb{V}[X] = 0.2589$$

Utilizamos la siguiente formula para obtener nuestra β y k:

$$\beta = \frac{\sigma^2}{\mu} - 1 = \frac{\mathbb{E}[X^2] - \mathbb{E}[X]^2}{\mathbb{E}[X]} - 1 = 0.0615 \quad \& \quad k = \frac{\mu^2}{\sigma^2 - \mu} = \frac{\mathbb{E}[X]^2}{\mathbb{E}[X^2] - \mathbb{E}[X] - \mathbb{E}[X]^2} = 3.3244$$

Ahora nuestra a y b:

$$a = \frac{\beta}{1+\beta} = 0.0579 \quad \& \quad b = (k-1)\frac{\beta}{1+\beta} = 0.1347$$

Ahora utilizaremos la sig formula para obtener los resutlados con nuestros paramteros calculados y compararlos con los origniales:

$$p_k = p_{k-1}(a + b\frac{1}{i}), \quad k = 1, 2, 3, ...$$

No olvidemos que para la BN p_0 se calcula de la siguiente forma:

$$p_0 = (1+\beta)^{-1}$$

Entonces comparemos los valores estimados vs los originales:

Como podemos ver los valores son muy similares, por lo que decimos que el modelo elegido fue el correcto.

reclamaciones	frecuencia
0	65599.62
1	12637.72
2	1583.50
3	162.86
4	14.92
5	1.27
6	0.10

(a) Valores p_k estimados para BN

reclamaciones	frecuencias
0	65623
1	12571
2	1644
3	148
4	13
5	1
6	0

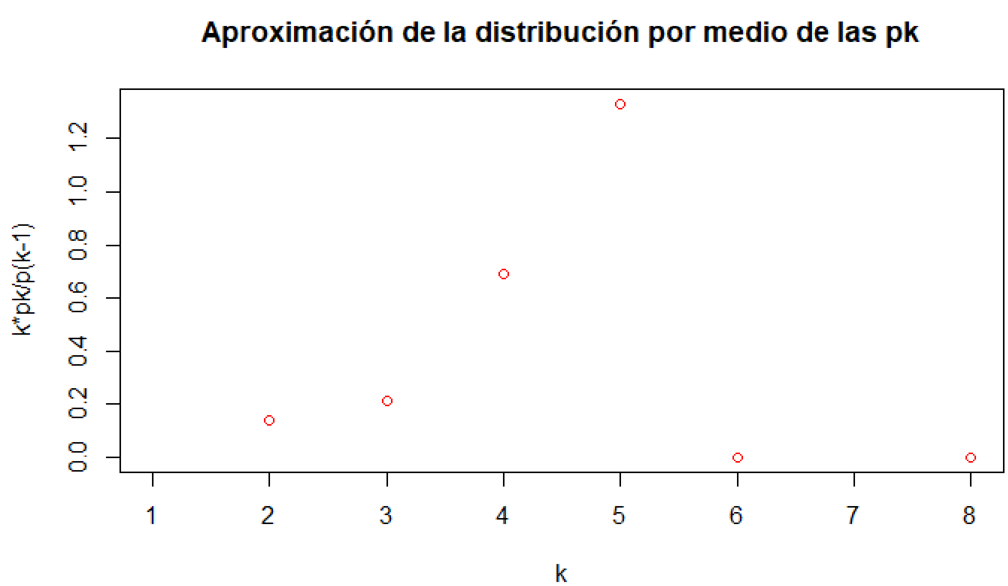
(b) Valores originales

Solución datos D

Obtengamos nuestras respectivas $p_k' s$

```
> #Utilizamos la formula k*p_k/p_k_-1
> pt2<-c()
> for(i in 2:8){
+   pt2[i]<-T2[i,1]*T2[i,2]/T2[i-1,2]
+ }
> pt2
[1]        NA 0.1405343 0.2148760 0.6923077 1.3333333 0.0000000        Inf 0.0000000
```

Graficamos nuestros resultados



Aproximación de un modelo discreto Datos D

Lo que nos sugiere un modelo binomial negativo, calulemos la esperanza y la varianza de nuestros datos para proponerlos cómo parametros de nuestra binomial negativa:

$$\mathbb{E}[X] = 0.166 \quad \& \quad \mathbb{V}[X] = 0.252$$

Utilizamos la siguiente formula para obtener nuestra β y k:

$$\beta = \frac{\sigma^2}{\mu} - 1 = \frac{\mathbb{E}[X^2] - \mathbb{E}[X]^2}{\mathbb{E}[X]} - 1 = 0.3521 \quad \& \quad k = \frac{\mu^2}{\sigma^2 - \mu} = \frac{\mathbb{E}[X]^2}{\mathbb{E}[X^2] - \mathbb{E}[X] - \mathbb{E}[X]^2} = 0.4715$$

Ahora nuestra a y b:

$$a = \frac{\beta}{1+\beta} = 0.2604 \quad \& \quad b = (k-1)\frac{\beta}{1+\beta} = -0.1376$$

Ahora utilizaremos la sig formula para obtener los resutlados con nuestros paramteros calculados y compararlos con los origniales:

$$p_k = p_{k-1}(a + b\frac{1}{i}), \quad k = 1, 2, 3, ...$$

No olvidemos que para la BN p_0 se calcula de la siguiente forma:

$$p_0 = (1+\beta)^{-1}$$

Entonces comparemos los valores estimados vs los originales:

reclamaciones	frecuencia
0	867.43
1	106.50
2	20.40
3	4.38
4	0.99
5	0.23
6	0.05
7	0.01

(a) Valores p_k estimados para BN

reclamaciones	frecuencias
0	861
1	121
2	13
3	3
4	1
5	0
6	1
7	0

(b) Valores originales

Como podemos ver los valores son muy similares, por lo que decimos que el modelo elegido fue el correcto.

El código de este ejercicio se encuentra en el anexo

13) Tenemos que $p(k) = \frac{2p(k-1)}{k}$, $k = 1, 2, 3, ...$

Suponemos que es de clase $(a, b, 0)$ ya que la podemos reescribir como:

$$\frac{p(k)}{p(k-1)} = \frac{2}{k} = a + \frac{b}{k} \text{ con } a = 0 \text{ y } b = 2$$

Como se tiene que $a = 0$ y $b = 0$, tomando como referencia la tabla de distribuciones de clase $(a, b, 0)$ (pág 55) vemos que se trata de una $Poisson(\lambda)$ con $\lambda = 2$ (ya que $b = \lambda$)

Entonces $P_0 = e^{-2}$

$P_1 = \frac{2(P_0)}{1} = 2e^{-2}$

$P_2 = \frac{2(P_1)}{2} = 2e^{-2}$

$P_3 = \frac{2(_2)}{3} = \frac{2(2e^{-2})}{3} = \frac{4}{3}e^{-2}$

$P_4 = \frac{2(P_3)}{4} = \frac{2(\frac{4}{3}e^{-2})}{4} = \frac{2}{3}e^{-2}$

14) Si N∼Geometrica con media 2, describe las funciones de densidad, distribución y esperanza Zero Truncada y Zero Modificada si $P_0^M = \frac{1}{6}$. Calcula los primeros 5 valores para las 3 variables y compara.

Solución:
Sabemos que $E[X] = \beta$ dado que $X \sim Geom(\beta)$, entonces $\beta = 2$.

Nos basaremos en las siguientes funciones:

- Función de desidad:
$$P[N = k] = \left(\frac{\beta}{1+\beta}\right)^k \frac{1}{1+\beta} \quad k = 0, 1, 2, ...$$

- Función de distribución de la clase (a,b,0):
$$\frac{kP_k}{P_{k-1}} = ak + b \quad k = 1, 2, ... \quad a = \frac{\beta}{1+\beta} \quad b = 0 \quad P_0 = (1+\beta)^{-1}$$

- Función de distribución truncada en cero:
$$P_k^T = \frac{P_k}{1 - P_0} \quad k = 1, 2, ...$$

- Función de distribución modificada en cero:
$$P_k^M = (1 - P_0^M) P_k^T \quad k = 1, 2, ...$$

- Esperanza cero truncada:
$$E[N^T] = \sum_{k=1}^{\infty} kP_k^T = \sum_{k=1}^{\infty} k \frac{\left(\frac{2}{3}\right)^k \frac{1}{3}}{1 - \frac{1}{3}}$$
$$= \frac{1}{2} \sum_{k=1}^{\infty} k \left(\frac{2}{3}\right)^k = \frac{1}{2}\left(\frac{2/3}{(1/3)^2}\right)$$

- Esperanza cero modificada:
$$E[N^M] = 0P_0^M + \sum_{k=1}^{\infty} kP_k^M = \sum_{k=1}^{\infty} k(1 - P_0^M) P_k^T$$
$$= \sum_{k=1}^{\infty} k \left(1 - \frac{1}{6}\right) \frac{\left(\frac{2}{3}\right)^k \frac{1}{3}}{1 - \frac{1}{3}} = \left(\frac{5}{6}\right) E[N^T]$$

■

k	P[N=k]	P_k	P_k^T	P_k^M
0	0.3333	0.3333	0	0.1667
1	0.2222	0.2222	0.3333	0.2778
2	0.1481	0.1481	0.2222	0.1852
3	0.0988	0.0987	0.1481	0.1234
4	0.0658	0.0658	0.0987	0.0823
5	0.0439	0.0439	0.0658	0.0549
E[N]	2	2	3	2.5

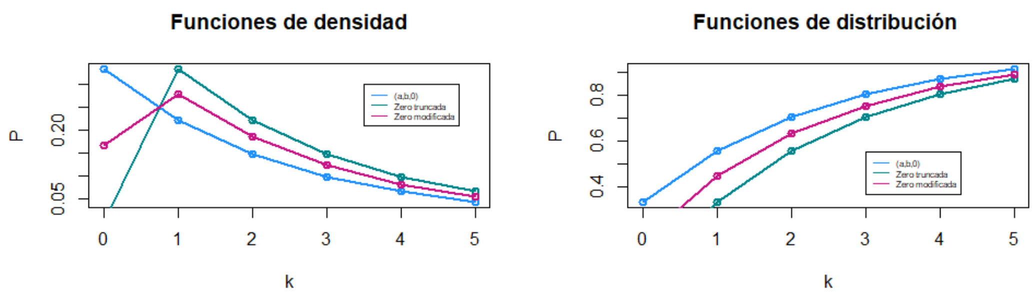


Figura 6: Geométrica(2)

15) Para la distribución Poisson(λ) con función de densidad. Si $P_o^M = 0.5$

Solución:
Sabemos que $E[X] = \lambda$=1 dado que $X \sim Poisson(\lambda)$, entonces $\lambda = 1$.

Nos basaremos en las siguientes funciones:

- Función de desidad:

$$\mathbb{P}(N = K) = \frac{\lambda^k e^{-\lambda}}{k!};\ k = 0, 1, 2, ..;\quad \lambda \geq 0$$

- Función de distribución de la clase (a,b,0):

$$\frac{kP_k}{P_{k-1}} = ak + b \quad k = 1, 2, ... \quad a = 0 \quad b = \lambda = 1 \quad P_0 = e^{-1}$$

- Función de distribución truncada en cero:

$$P_k^T = \frac{P_k}{1 - P_0} \quad k = 1, 2, ...$$

- Función de distribución modificada en cero:

$$P_k^M = (1 - P_0^M)P_k^T \quad k = 1, 2, ...$$

- Esperanza cero truncada:

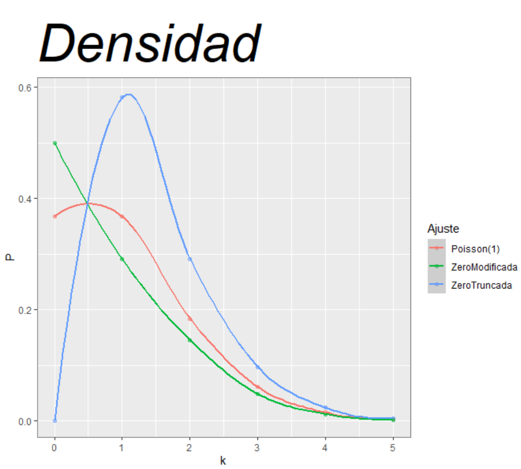
$$E[N^T] = \sum_{k=1}^{\infty} kP_k^T = \frac{1}{1-P_0}\sum_{k=0}^{\infty}(k)P_k$$
$$\frac{\lambda}{1-P_0} = \frac{\lambda}{1-e^{-\lambda}}$$
$$\therefore \mathbb{E}\left[P_{T_k}\right] \simeq 1.581976707$$

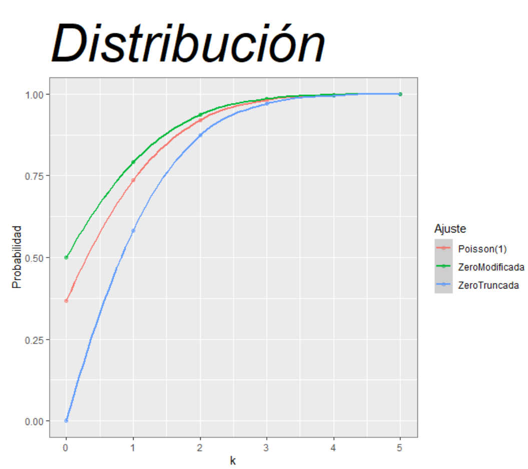
- Esperanza cero modificada:

$$\begin{aligned}
\mathbb{E}\left[N^M\right] &= \sum_{k=0}^{\infty} k(P_k^M) \\
&= \sum_{k=0}^{\infty} k\frac{1-P_0^M}{1-P_0}P_k \\
&= \frac{1-P_0^M}{1-P_0}\sum_{k=0}^{\infty} k(P_k) \\
&= \frac{1-0.5}{1-e^{-\lambda}} = \frac{0.5\lambda}{1-e^{-\lambda}} \\
\therefore \mathbb{E}\left[N^M\right] &\simeq \boxed{0.7909883534}
\end{aligned}$$

$\blacksquare$

k	Poisson(1)	Distribución Poisson(1)	Zero Truncada	Distribución Zero Truncada	Zero Modificada	Distribución Zero Modificada
0	0.367879441	0.367879441	0	0	0.5	0.5
1	0.367879441	0.735758882	0.581976707	0.581976707	0.290988353	0.790988353
2	0.183939721	0.919698603	0.290988353	0.87296506	0.145494177	0.93648253
3	0.06131324	0.981011843	0.096996118	0.969961178	0.048498059	0.984980589
4	0.01532831	0.996340153	0.024249029	0.994210208	0.012124515	0.997105104
5	0.003065662	0.999405815	0.004849806	0.999060013	0.002424903	0.999530007





16) *a*) $X_1 \sim U(0,5000)$ y $X_2 \sim exp(5000)$, modificamos la cobertura agregando un límite de póliza $u = 4500$. Encuentra la densidad, distribución y esperanza para las variables modificadas Y_L y Y_P en cada uno de los dos modelos, grafica la densidad original vs la modificada y compara.

b) $X \sim U(0,100)$, si se agrega un deducible d:
Calcula d si $E(Y_L) = 32$
Calcula d si $E(Y_P) = 32$

Solución: Primero para la $U(0,5000)$.

$$f_x(x) = \frac{1}{5000} \qquad \& \qquad F_x(x) = \frac{x}{5000}, \quad con \;\; 0 \leq x \leq 5000$$

Así entonces calculemos para Y_L

$$f_{y_L}(y) = \begin{cases} F_x(\frac{d}{1+r}) = F_x(0) = \boxed{0} & si \quad y = 0 \\ \frac{1}{\alpha(1+r)} f_x(\frac{y+\alpha d}{\alpha(1+r)}) = f_x(y) = \boxed{\dfrac{1}{5000}} & si \quad 0 < y < 4500 \\ 1 - F_x(\frac{u}{1+r}) = 1 - F_x(4500) = 1 - \frac{4500}{5000} = \boxed{\dfrac{1}{10}} & si \quad y = 4500 \end{cases}$$

$$F_{y_L}(y) = \begin{cases} F_x(\frac{d}{1+r}) = F_x(0) = \boxed{0} & si \quad y = 0 \\ F_x(\frac{y+\alpha d}{\alpha(1+r)}) = F_x(y) = \boxed{\dfrac{y}{5000}} & si \quad 0 < y < 4500 \\ \boxed{1} & si \quad y \geq 4500 \end{cases}$$

$$\begin{aligned} \mathbb{E}[y_L] =& \alpha(1+r)[\mathbb{E}(x \wedge \frac{u}{1+r}) - \mathbb{E}(x \wedge = \frac{d}{1+r})] \\ =& 1(1+0)[\mathbb{E}(x \wedge 4500) - \mathbb{E}(x \wedge = 0)] = \mathbb{E}(x \wedge 4500) \\ =& \int_0^u S(x)dx = \int_0^u 1 - \frac{x}{5000} dx = \int_0^u dx - \int_0^u \frac{x}{5000} dx \\ =& x|_0^{4500} - \frac{x^2}{10000}|_0^{4500} = 4500 - \frac{4500^2}{10000} = 2475 \\ \therefore \mathbb{E}[y_L] =& \boxed{2475} \end{aligned}$$

Ahora calculemos para Y_p

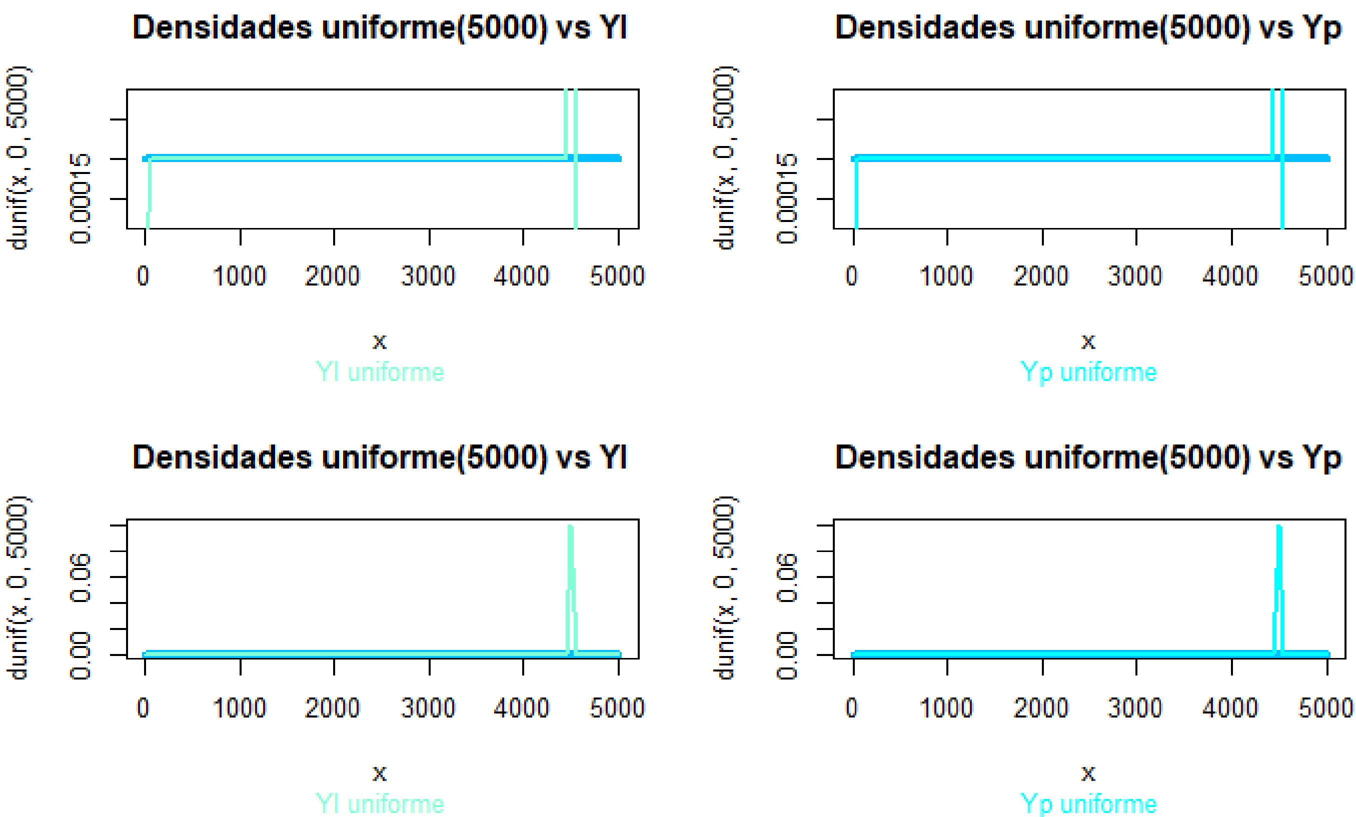
$$f_{y_p}(y) = \begin{cases} \boxed{0} & si \quad y = 0 \\ \frac{1}{\alpha(1+r)} \frac{f_x(\frac{y+\alpha d}{\alpha(1+r)})}{1 - F_x(\frac{d}{1+r})} = \frac{f_x(y)}{1 - F_x(0)} = \boxed{\dfrac{1}{5000}} & si \quad 0 < y < 4500 \\ \frac{1 - F_x(\frac{u}{1+r})}{1 - F_x(\frac{d}{1+r})} = 1 - F_x(4500) = 1 - \frac{4500}{5000} = \boxed{\dfrac{1}{10}} & si \quad x = 4500 \end{cases}$$

$$F_{y_p}(y) = \begin{cases} \boxed{0} & si \quad y = 0 \\ \frac{F_x\left(\frac{y+\alpha d}{\alpha(1+r)}\right) - F_x\left(\frac{d}{1+r}\right)}{1 - F_x\left(\frac{d}{1+r}\right)} = F_x(y) = \boxed{\frac{y}{5000}} & si \quad 0 < y < 4500 \\ \boxed{1} & si \quad y \geq 4500 \end{cases}$$

$$\mathbb{E}[y_p] = \frac{\mathbb{E}[y_L]}{S_x(0)} = \frac{2475}{1 - F_x(0)} = \frac{2475}{1}$$

$$\therefore \mathbb{E}[y_p] = \boxed{2475}$$

Grafiquemos las densidades para Y_L y Y_p



Densidades $U(0, 5000)$ vs Y_L Y_p

Ahora para $exp(5000)$

$$f_x(x) = 5000e^{-5000x} \qquad \& \qquad F_x(x) = 1 - e^{-5000x}, \quad con \ \ x \geq 0$$

Así entonces calculemos para Y_L

$$f_{y_L}(y) = \begin{cases} F_x(\frac{d}{1+r}) = F_x(0) = \boxed{0} & si \qquad y = 0 \\ \\ \frac{1}{\alpha(1+r)} f_x(\frac{y+\alpha d}{\alpha(1+r)}) = f_x(y) = \boxed{5000e^{-5000y}} & si \quad 0 < y < 4500 \\ \\ 1 - F_x(\frac{u}{1+r}) = 1 - F_x(4500) = 1 - (1 - e^{-5000*4500}) = \boxed{e^{-\frac{9}{10}}} & si \qquad y = 4500 \end{cases}$$

$$F_{y_L}(y) = \begin{cases} F_x(\frac{d}{1+r}) = F_x(0) = \boxed{0} & si \qquad y = 0 \\ \\ F_x(\frac{y+\alpha d}{\alpha(1+r)}) = F_x(y) = \boxed{1 - e^{-5000y}} & si \quad 0 < y < 4500 \\ \\ \boxed{1} & si \qquad y \geq 4500 \end{cases}$$

$$\begin{aligned} \mathbb{E}[y_L] =& \alpha(1+r)[\mathbb{E}(x \wedge \frac{u}{1+r}) - \mathbb{E}(x \wedge \frac{d}{1+r})] \\ =& 1(1+0)[\mathbb{E}(x \wedge 4500) - \mathbb{E}(x \wedge 0)] = \mathbb{E}(x \wedge 4500) \\ =& \int_0^u S(x)dx = \int_0^u 1 - (1 - e^{-5000x})dx = \int_0^u e^{-5000x}dx \\ =& -\frac{e^{-5000x}}{5000}|_0^{4500} \\ =& -\frac{e^{-5000(4500)} + 1}{5000} = \frac{1}{5000} \\ \therefore \mathbb{E}[y_L] =& \boxed{0.0002} \end{aligned}$$

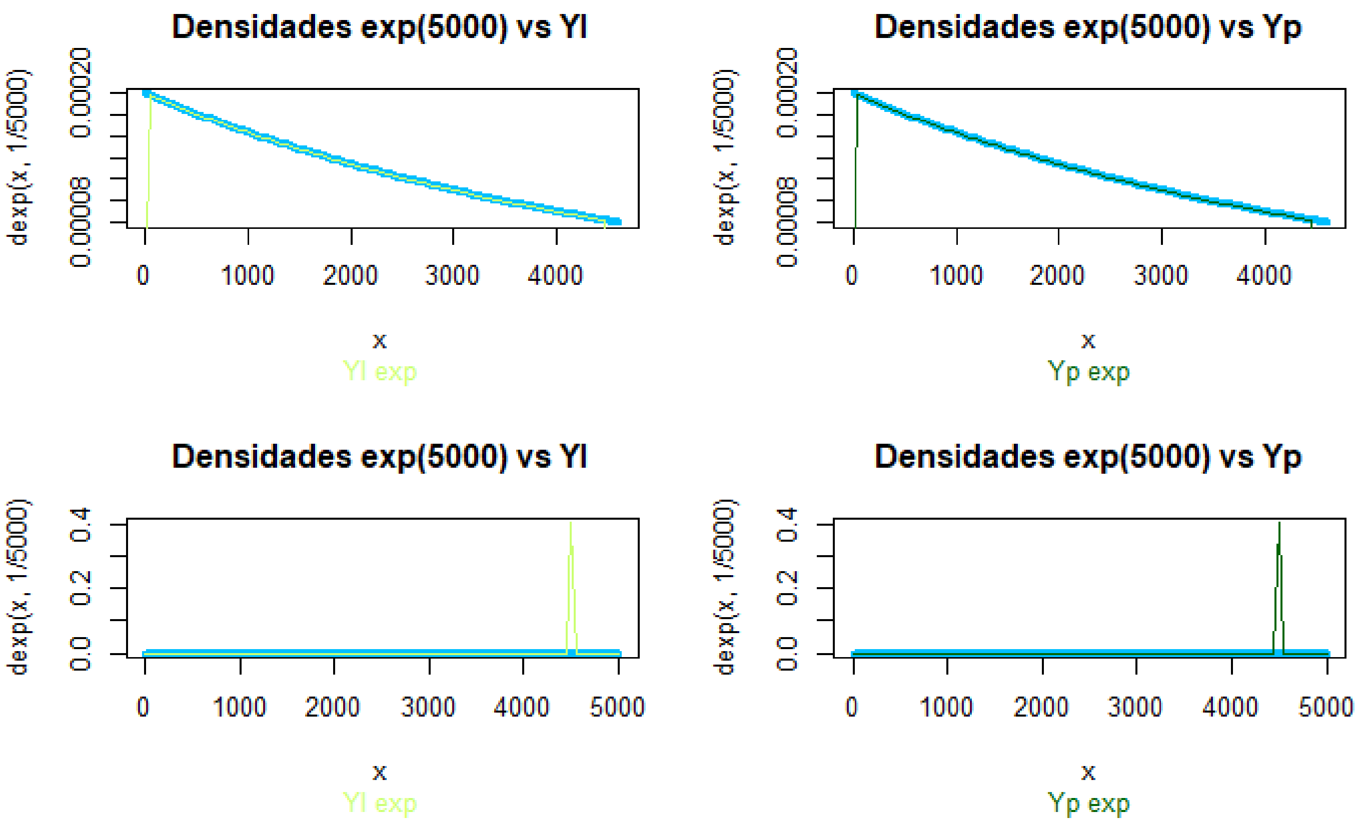
Ahora calculemos para Y_p

$$f_{y_p}(y) = \begin{cases} \boxed{0} & si \qquad y = 0 \\ \\ \frac{1}{\alpha(1+r)} \frac{f_x(\frac{y+\alpha d}{\alpha(1+r)})}{1 - F_x(\frac{d}{1+r})} = \frac{f_x(y)}{1 - F_x(0)} = \boxed{5000e^{-5000y}} & si \quad 0 < y < 4500 \\ \\ \frac{1 - F_x(\frac{u}{1+r})}{1 - F_x(\frac{d}{1+r})} = 1 - F_x(4500) = 1 - (1 - e^{-5000*4500}) = \boxed{e^{-\frac{9}{10}}} & si \qquad y = 45000 \end{cases}$$

$$F_{y_p}(y) = \begin{cases} \boxed{0} & si \qquad y = 0 \\ \\ \frac{F_x(\frac{y+\alpha d}{\alpha(1+r)}) - F_x(\frac{d}{1+r})}{1 - F_x(\frac{d}{1+r})} = F_x(y) = \boxed{1 - e^{-5000y}} & si \quad 0 < y < 4500 \\ \\ \boxed{1} & si \qquad y \geq 4500 \end{cases}$$

$$\mathbb{E}[y_p] = \frac{\mathbb{E}[y_L]}{S_x(0)} = \frac{\frac{1}{5000}}{1 - F_x(0)} = \frac{1}{5000} = \boxed{0.0002}$$

Grafiquemos las densidades para Y_L y Y_p



Densidades $exp(5000)$ vs $Y_L \;\; Y_p$

b

Calcule d si $\mathbb{E}[Y_L] = 32$

$$
\begin{aligned}
\mathbb{E}[y_L] =& \alpha(1+r)[\mathbb{E}(x \wedge \frac{u}{1+r}) - \mathbb{E}(x\wedge = \frac{d}{1+r})] \\
=& 1(1+0)[\mathbb{E}(x \wedge 4500) - \mathbb{E}(x \wedge d)] \\
=& \int_0^{4500} 1 - \frac{x}{100} dx - \int_0^d 1 - \frac{x}{100} dx \\
=& 4500 - [\frac{x^2}{200}|_0^{4500}] - [d - (\frac{x^2}{200}|_0^d)] \\
=& 4500 - \frac{4500^2}{200} - d + \frac{d^2}{200} \\
32 =& \frac{d^2}{200} - d - 96750 \\
=& d^2 - 200d - 200(96782) = d^2 - 200d - 19356400 \\
=& \frac{200\sqrt{200^2 - 4(1)(-19356400}}{2(1)} \\
& d1 \approx 4500.727213 \quad \& \quad d2 \approx -4300.72 \\
\therefore d \approx& 4500.727213 \;\; para \;\; \mathbb{E}[Y_L] = 32
\end{aligned}
$$

Calcule d si $\mathbb{E}[Y_p] = 32$

$$\begin{aligned}
\mathbb{E}[Y_p] =& \frac{\mathbb{E}[y_L]}{S_x(\frac{d}{1+r})} = \frac{\mathbb{E}[y_L]}{S_x(d)} = \frac{32}{1-\frac{d}{100}} = \frac{32(100)}{100-d} \\
32 =& \frac{3200}{100-d} \\
3200 - 32d =& 3200 \\
-32d =& 0 \\
d =& 0 \\
\therefore d =& 0 \;\; para \;\; \mathbb{E}[Y_p] = 32
\end{aligned}$$

El código del ejercicio 16 se encuentra en el anexo.

17) La v. a. continua X tiene la siguiente densidad:

$f(x) = \begin{cases} 0.05 & si \quad 0 \le x \le 20 \\ 0 & \qquad eoc \end{cases}$

Encuentra la densidad, distribución y esperanza para Y_L y Y_P si $d = 3.5$.

1. Para Y_p:

$$f_{Y^P} = \frac{\frac{1}{20}}{1 - \frac{3.5}{20}} = \frac{2}{33}$$

Entonces

$$F_{Y^P} = \frac{\frac{y+3.5}{20} - \frac{3.5}{20}}{1 - \frac{3.5}{20}} = \frac{\frac{y+3.5-3.5}{20}}{1 - \frac{3.5}{20}} = \frac{\frac{y}{20}}{\frac{33}{40}} = \frac{40y}{20 * 33} = \frac{2y}{33}$$

$$\mathbb{E}[Y_P] = \frac{\mathbb{E}[Y_L]}{\delta(\alpha)} = \frac{5.8875}{1 - \frac{3.5}{20}} = \frac{157}{22}$$

2. Para Y_L:

$f_{Y_L} = \begin{cases} \frac{3.5}{20} \quad o \quad \frac{7}{40} & si \quad y = 0 \\ \frac{1}{20} & si \quad y > 0 \end{cases}$ $\qquad F_{Y_L} = \begin{cases} \frac{7}{40} & si \quad y = 0 \\ \frac{y+3.5}{20} & si \quad y > 0 \end{cases}$

$\mathbb{E}[xu] = \int_0^u \delta(x)dx = \int_0^u 1 - \frac{x}{20}dx = x|_0^u - \frac{x^2}{20}|_0^u = u - 0 - \frac{u^2}{20} = u - \frac{u^2}{20}$

$\mathbb{E}[Y_L] = \mathbb{E}[x] = \mathbb{E}[xu] = 10 - (3.5 - \frac{3.5^2}{20}) = 10 - 3.5 + \frac{3.5^2}{20} = \frac{437}{80} = 5.8875$

18) X $\sim$ exp con media 900, d = 500. El asegurador quiere el doble de la tasa de eliminación de pérdida que se tiene actualmente. Determina el nuevo deducible.

Solución:
Consideremos:

$$\text{Tasa de eliminación de pérdida } = \frac{E[X \wedge d]}{E[X]} \qquad E[X \wedge d] = \int_0^d [1 - F(x)]dx$$

Dado que X$\sim$ $Exponencial(\frac{1}{900})$

$$F(x) = 1 - e^{\frac{x}{900}}$$

Primero calculemos la tasa de eliminación de pérdida actual:

$$\begin{aligned} E[X \wedge d] &= \int_0^{500} e^{\frac{x}{900}} dx \qquad Sea \;\; a = -\frac{1}{900}x \quad da = -\frac{1}{900}dx \\ &= 900 \int_{-\frac{5}{9}}^{0} e^a da \qquad\qquad a_1 = -\frac{500}{900} \quad a_2 = 0 \\ &= 900[1 - e^{-\frac{5}{9}}] = 383.621912 \end{aligned}$$

$$\therefore \text{Tasa de eliminación de pérdida actual} = \tfrac{383.62192134}{900} = 0.4262466$$

Calculemos el nuevo deducible que garantice el doble de la tasa de eliminación de pérdida actual:

$$\begin{aligned} \frac{\int_0^d e^{\frac{x}{900}} dx}{900} &= 2(0.4262466) \\ \frac{900 \int_{-\frac{d}{900}}^0 e^a da}{900} &= 0.8524932 \\ 1 - e^{-\frac{d}{900}} &= 0.8524932 \\ e^{-\frac{d}{900}} &= 0.1475068 \\ -\frac{d}{900} &= ln(0.1475068) \end{aligned}$$

$\therefore$ **El nuevo deducibles es: d=1722.492649**

■

19) La v. a. X tiene las siguientes características:

x	$F(x)$	$E(x \wedge x)$
0	0	0
100	0.2	91
200	0.6	153
1000	1.0	331

Calcula $e(100)$.

Solución:

$$\mathbb{E}[X] = e(d)S(d) + \mathbb{E}(x \wedge d)$$

Entonces tenemos:

$$\rightarrow e(d) = \frac{\mathbb{E}[x] - \mathbb{E}[x \wedge d]}{S(d)}$$

d	F(d)	E(X^d)	f(d)	d*f(x)
0	0	0	0	0
100	0.2	91	0.2	20
200	0.6	153	0.4	80
1000	1	331	0.4	400

$$\mathbb{E}[X] = 500$$

$$\therefore e(100) = \frac{\mathbb{E}[x] - \mathbb{E}[x \wedge 100]}{1 - F(100)} = \boxed{511.25}$$

■

20) Para una v. a. $Pareto(\alpha = .78, \theta = 100)$ encuentra la densidad, distribución y esperanza para Y_L y Y_P. Compara las densidades y distribuciones en R.

$d = 120$
$u = 1000$
$\alpha = 80\,\%$
$r = 5\,\%$

Solución

Para una $Pareto(0.78, 100)$ tenemos:

$$f_x(x) = \frac{0.78(100)^{0.78}}{x^{0.78+1}} \quad \& \quad F_x(x) = 1 - (\frac{100}{x})^{0.78}, \quad x \geq 100$$

Así entonces calculemos para Y_L

$$f_{y_L}(y) = \begin{cases} F_x(\frac{d}{1+r}) = F_x(\frac{120}{1.05}) = 1 - (\frac{100}{\frac{120}{1.05}})^{0.78} = \boxed{0.098913987} \;\; si \;\; y = 0 \\ \frac{1}{0.8(1.05)} f_x(\frac{y+0.8(120)}{0.8(1.05)}) = \frac{1}{0.8(1.05)}(\frac{0.78(100)^{0.78}}{(\frac{y+0.8(120)}{0.8(1.05)})^{0.78+1}}) = \frac{0.8(1.05)^0.78+1}{0.8(1.05)} \frac{28.32008}{(y+96)^{1.78}} = \boxed{\frac{24.7191}{(y+96)^{1.78}}} \frac{si}{0<y<704} \\ 1 - F_x(\frac{u}{1+r}) = 1 - F_x(\frac{1000}{1.05}) = 1 - 1 + (\frac{100}{(\frac{1000}{1.05})})^{0.78} = (\frac{1*1.05*100}{10*100})^{0.78} = \boxed{0.1724} \;\; si \; y = 704 \end{cases}$$

$$F_{y_L}(y) = \begin{cases} F_x(\frac{d}{1+r}) = F_x(\frac{120}{1.05}) = 1 - (\frac{100}{\frac{120}{1.05}})^{0.78} = \boxed{0.098913987} \;\; si \;\; y = 0 \\ F_x(\frac{y+\alpha d}{\alpha(1+r)}) = 1 - (\frac{y+0.8(120)}{0.8(1.05)})^{0.78} = 1 - (\frac{0.8*1.05*100}{y+96})^{0.78} = \boxed{1 - (\frac{84}{y+96})^{0.78}} \;\; si \;\; 0 < y < 704 \\ \boxed{1} \;\; si \;\; y \geq 704 \end{cases}$$

$$\begin{aligned} \mathbb{E}[y_L] =& \alpha(1+r)[\mathbb{E}(x \wedge \frac{u}{1+r}) - \mathbb{E}(x \wedge = \frac{d}{1+r})] \\ =& 0.84[\mathbb{E}(x \wedge \frac{1000}{1.05}) - \mathbb{E}(x \wedge \frac{120}{1.05})] \\ =& 0.84[\int_0^{\frac{1000}{1.05}} (\frac{100}{x})^{0.78} dx - \int_0^{\frac{120}{1.05}} (\frac{100}{x})^{0.78} dx] \\ =& 0.84 * 100^{0.78}[\int_0^{\frac{1000}{1.05}} \frac{1}{x^{0.78}} dx - \int_0^{\frac{120}{1.05}} \frac{1}{x^{0.78}} dx] \\ =& 30[20.55491444 - 12.8924527] \\ \therefore \mathbb{E}[y_L] =& \boxed{233.6940} \end{aligned}$$

Ahora calculemos para Y_p

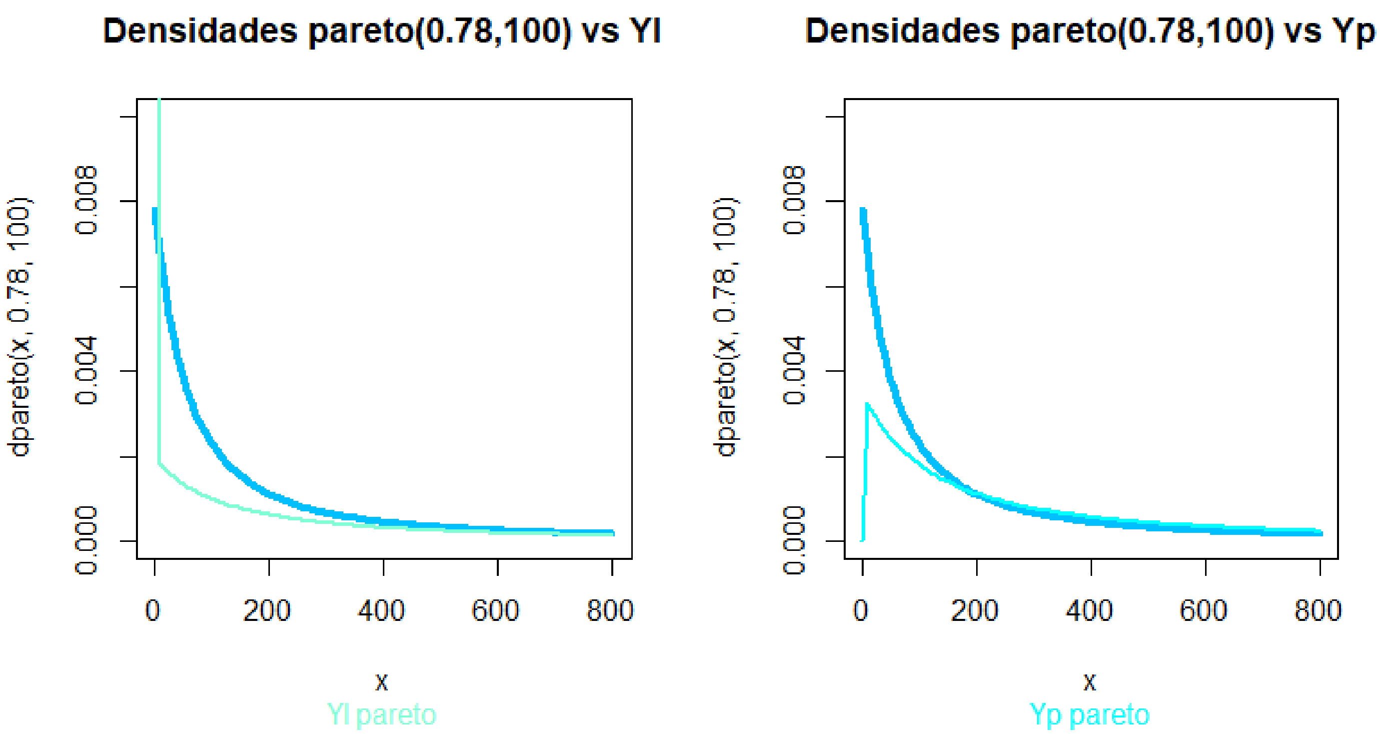
$$f_{y_p}(y) = \begin{cases} \boxed{0} \;\; si \;\; y = 0 \\ \frac{1}{\alpha(1+r)} \frac{f_x(\frac{y+\alpha d}{\alpha(1+r)})}{1-F_x(\frac{d}{1+r})} = \frac{1}{0.8*1.05} \frac{\frac{0.78*100^{0.78}}{(\frac{y+0.8*120^{0.78}}{0.8*1.05})^{0.78}}}{1-(1-(\frac{100}{\frac{120}{1.05}})^0.78)} = \frac{0.78*120^{0.78}*108^{0.78}}{(y+96)^{0.78}} = \boxed{\frac{27.4325}{(y+96)^{0.78}}} \;\; si \;\; 0 < y < 704 \\ \frac{1-F_x(\frac{u}{1+r})}{1-F_x(\frac{d}{1+r})} = \frac{1-[1-(\frac{100}{\frac{1000}{1.05}})^{0.78}]}{1-[1-(\frac{100}{\frac{120}{1.05}})^{0.78}]} = \frac{(\frac{120}{1.05})^{0.78}}{(\frac{1000}{1.05})^{0.78}} = (\frac{120}{1000})^{0.78} = \boxed{0.1913} \;\; si \;\; y = 704 \end{cases}$$

$$F_{y_p}(y) = \begin{cases} \boxed{0} \quad si \quad y = 0 \\ \frac{F_x(\frac{y+\alpha d}{\alpha(1+r)}) - F_x(\frac{d}{1+r})}{1-F_x(\frac{d}{1+r})} = \frac{1-(\frac{100}{\frac{y+96}{0.8*1.05}})^{0.78} - 1 + (\frac{100}{\frac{120}{1.05}})^{0.78}}{1-1+(\frac{100}{\frac{120}{1.05}})^{0.78}} = 1 - \frac{(\frac{120}{1.05})^{0.78}}{(\frac{y+96}{0.8*1.05})^{0.78}} = \boxed{1-(\frac{96}{y+96})^{0.78}} \; si \; 0 < y < 704 \\ \boxed{1} \quad si \; y \geq 704 \end{cases}$$

$$\mathbb{E}[y_p] = \frac{\mathbb{E}[y_L]}{S_x(\frac{d}{1+r})} = \frac{233.6940}{(\frac{100}{\frac{120}{1.05}})^0.78} =$$

$$\therefore \mathbb{E}[y_p] = \boxed{259.3471}$$

Grafiquemos las densidades para Y_L y Y_p



Densidades $Pareto(0.78, 100)$ vs Y_L Y_p

El código correspondiente al ejercicio 20 se encuentra en el anexo.

ANEXO
Ejercicio 4a

```
#####################################################
### Tarea Examen 1. Teoría del riesgo. Ejercicio 4A ###
#####################################################

#Vector de observaciones
x<-c(25, 457, 82, 680, 115, 855, 126, 877, 155, 974, 161, 1193, 243, 1340, 294, 1884, 340, 2558, 384,
#Medidas descriptivas#
summary(x)
#Min. 1st Qu.  Median    Mean 3rd Qu.    Max. #
#25.0   159.5   420.5   811.0  1028.8  3476.0 #
tail(x)
quantile(x)
var(x)
sd((x))
range(x)
median(x)
library(modeest)
mfv(x)
#sesgo y curtosis, moda
library(psych)
skew(x) #Asimétirca hacia la derecha
kurtosi(x)
#histograma
hist(x,freq=F,breaks = 25,col="deeppink",ylim=)#¿de qué le veo cara?
d<-density(x)
lines(d)

#Veamos a que podemos ajustar
library(fitdistrplus)
descdist(x)

#Cómo podemos ver se encuentra cerca de una exponencial y una gamma, hagamos estimación de paramteros
#Por igualación de momentos
f_exp<-fitdist(x,"exp","mme")
plot(f_exp)
f_exp

f_gamma<-fitdist(x,"gamma","mme")
plot(f_gamma)
f_gamma

#Por máxima verosimilitud
f_exp_ml<-fitdistr(x,"exponential")
f_exp_ml

f_gamma_ml<-fitdistr(x,"gamma")
f_gamma_ml

#Pruebas de Hipótesis
```

```
#####################
##Kolmogorov-Smirnof##
#####################
ks.test(x,pexp,f_exp_ml$estimate[1])
ks.test(x,pgamma,f_gamma_ml$estimate[1],f_gamma_ml$estimate[2])
#####################
###Anderson-Darling###
#####################
ADGofTest::ad.test(x,pexp,f_exp_ml$estimate[1])
ADGofTest::ad.test(x,pgamma,f_gamma_ml$estimate[1],f_gamma_ml$estimate[2])
```

Ejercicio 4b

```
######################################################
### Tarea Examen 1. Teoría del riesgo. Ejercicio 4B ###
######################################################
library(psych)
library(modeest)
library(actuar)
library(fitdistrplus)
#Vector de observaciones
y<-c(3 ,11, 27, 36, 47, 49, 54, 77, 78, 85,
     104,121, 130, 138, 139, 140 ,143,153, 193, 195,
     205, 207, 216, 224, 233, 237, 254, 257 ,259, 265,
     273 ,275 ,278, 281, 396, 405, 412, 423, 436, 456,
     473,475, 503, 510, 534, 565, 656, 656, 716, 734,
     743, 756, 784, 786, 819 ,826 ,841 ,842, 853, 860,
     877, 942, 942 ,945 ,998 ,1029, 1066, 1101, 1128, 1167,
     1194, 1209, 1223, 1283, 1288, 1296, 1310, 1320,1367 ,1369,
     1373 ,1382 ,1383 ,1395, 1436 ,1470 ,1512 ,1607 ,1699 ,1720,
     1772 ,1780 ,1858,1922,2042,2247, 2348,2377, 2418, 2795,
     2964 ,3156 ,3858, 3872, 4084 ,4620 ,4901, 5021, 5331, 5771,
     6240 ,6385 ,7089, 7482,8059 ,8079, 8316, 11453, 22274 ,32043)
#yi<-y/100000


#Medidas descriptivas#
summary(y)
tail(y)
quantile(y)
var(y)
sd((y))
range(y)
median(y)
mfv(y)
#sesgo y curtosis
skew(y) #Asimétirca hacia la derecha
kurtosi(y)#Leptocúrtica

#histograma
hist(y,freq=F,breaks = 50,col= "cadetblue1")#¿de qué le veo cara?
```

```
d<-density(y)
lines(d)

u<-seq(min(y),max(y),by=100)

#una pareto?
lines(u,dpareto(u,0.5,30),col="darkorange",lwd=1,lty=1)
#una gamma?
lines(u,dgamma(u,0.5,0.00047),col="green4",lwd=1,lty=2)
#una lognormal?
lines(u,dlnorm(u,6.83,1),col="deepskyblue2",lwd=1,lty=3)
lines(u,dlnorm(u,6.6,1.5),col="firebrick1",lwd=1,lty=4)
# leyenda
legend( x=20000, y=0.0006,
        legend = c("pareto(0.5,30)","gamma(0.5,0.00047)","lognormal(6.83,1)", "lognormal(6.6,1.5)"),
        col = c("darkorange", "green4","deepskyblue2","firebrick1"),
        lty=1:4, horiz = FALSE)

#Veamos a que podemos ajustar
descdist(y)

#Ajustamos una log-normal

f5p<-fitdist(y,"lnorm")
plot(f5p)
f5p
f5<-fitdistr(y,"lognormal")
f5
#pruebas de bondad f5p
ks.test(y,plnorm,f5p$estimate[1],f5p$estimate[2])
ADGofTest::ad.test(y,plnorm,f5p$estimate[1],f5p$estimate[2])
#Por lo tanto el modelo que se ajusta a nuestros dato
#es una Log-normal con media 6.62417 y var 1.51124

hist(y,freq=F,breaks = 50,col= "cadetblue1")#¿de qué le veo cara?
d<-density(y)
lines(d)
lines(u,dlnorm(u,6.62417,1.51124),col="steelblue4",lwd=2,lty=5)
# leyenda
legend( x=20000, y=0.0006,
        legend = c("lognormal(6.62417,1.51124"),
        col = c("steelblue4"),
        lty=5, horiz = FALSE)
```

Ejercicio 10

```
#"""Calcula el V aR y T V aR al 90% para las siguientes distribuciones:"""

n<-1000
m<-1000

CondVaR<-function(x,n,m,VaR){
```

```
  aux<-c()
  for(i in 1:n){
    for(j in 1:m){
      if(x[i,j]>VaR){
        aux[j]<-x[i,j]
      }
    }
  }
  return(mean(aux))
}

### U(0,1000)
# VaR
set.seed(123)
y1<-rep(0,n)
x1<-matrix(0,n,m)

for(i in 1:n){
  x1[i,]<-runif(m, 0, 1000) #muestra 'i' de valores aleatorios de
  una U(0,1000)
  y1[i]<-quantile(x1[i,],probs=0.9) #VaR al 90% de cada muestra
  de tamaño  'm'
}

VaR1<-mean(y1)

# TVaR
TVaR1<-CondVaR(x1,n,m,VaR1)

hist(x1, main= "Histograma de simulaciones U(0,1000)", col =
"deepskyblue2", xlab = "Simulaciones", ylab="Frecuencia")
abline(v=VaR1, col = "firebrick", lty = 2, lwd=2)
abline(v=TVaR1, col = "red", lty =3,, lwd=2)
legend(0, 50000, c(paste("VaR:",round(VaR1,2)), paste("TVaR:",
round(TVaR1,2))), col = c("firebrick", "red"), lty =c(2,3), cex
=.8)


### Parteo(1.6,2)
# VaR
library(actuar)

y2<-rep(0,n)
x2<-matrix(0,n,m)

for(i in 1:n){
  x2[i,]<-rpareto(m,shape=1.6,scale=2)
  y2[i]<-quantile(x2[i,],probs=0.9)
}

VaR2<-mean(y2)
```

```
# TVaR

TVaR2<-CondVaR(x2,n,m,VaR2)

hist(x2, main= "Histograma de simulaciones Pareto(1.6,2)", col =
"deeppink", xlab = "Simulaciones", ylab="Frecuencia", breaks
=10000,xlim= c(0,60))
abline(v=VaR2, col = "deeppink4", lty = 2, lwd=2)
abline(v=TVaR2, col = "mediumorchid4", lty =3, lwd=2)
legend(30, 4e+05, c(paste("VaR:",round(VaR2,2)), paste("TVaR:",
round(TVaR2,2))),
       col = c("deeppink4", "mediumorchid4"), lty =c(2,3))


# Exp(media = 1000)
## VaR
y3<-rep(0,n)
x3<-matrix(0,n,m)

for(i in 1:n){
  x3[i,]<-rexp(m,1/1000)
  y3[i]<-quantile(x3[i,],probs=0.9)
}

VaR3<-mean(y3)

# TVaR

TVaR3<-CondVaR(x3, n,m, VaR3)

hist(x3, main= "Histograma de simulaciones \n
Exponencial(media=1000)", col = "olivedrab3", xlab =
"Simulaciones", ylab="Frecuencia", breaks =100, xlim= c(0,8000))
abline(v=VaR3, col = "olivedrab4", lty = 2, lwd=2)
abline(v=TVaR3, col = "orangered4", lty =3, lwd=2)
legend(4000, 150000, c(paste("VaR:",round(VaR3,2)),
paste("TVaR:", round(TVaR3,2))),
       col = c("olivedrab4", "orangered4"), lty =c(2,3))


### Lognormal( 5,  2)
# VaR
y4<-rep(0,n)
x4<-matrix(0,n,m)

for(i in 1:n){
  x4[i,]<-rlnorm(m,5,2)
  y4[i]<-quantile(x4[i,],probs=0.9)
}

VaR4<-mean(y4)
```

```
# TVaR
TVaR4<-CondVaR(x4,n,m,VaR4)

hist(x4, main= "Histograma de simulaciones \n Lognormal(5,2)",
col = "darkgoldenrod1",
     xlab = "Simulaciones", ylab="Frecuencia", breaks =10000,
     xlim= c(0,10500))
abline(v=VaR4, col = "darkcyan", lty = 2, lwd=2)
abline(v=TVaR4, col = "darkblue", lty =3, lwd=2)
legend(3000, 4e+05, c(paste("VaR:",round(VaR4,2)), paste("TVaR:",
round(TVaR4,2))), col = c("darkcyan", "darkblue"), lty =c(2,3))

###N( 1000, 1000^2)
# VaR
y5<-rep(0,n)
x5<-matrix(0,n,m)

for(i in 1:n){
  x5[i,]<-rnorm(m,1000,1000)
  y5[i]<-quantile(x5[i,],probs=0.9)
}

VaR5<-mean(y5)

# TVaR
TVaR5<-CondVaR(x5,n,m,VaR5)

hist(x5, main= "Histograma de simulaciones \n Normal(1000,1000)", col = "plum2", xlab = "Simulaciones
abline(v=VaR5, col = "purple4", lty = 2, lwd=2)
abline(v=TVaR5, col = "turquoise4", lty =3, lwd=2)
legend(-3500, 150000, c(paste("VaR:",round(VaR5,2)),
paste("TVaR:", round(TVaR5,2))), col = c("purple4", "turquoise4"), lty =c(2,3), cex=.5)
```

Ejercicio 12

```
######################################################
### Tarea Examen 1. Teoría del riesgo. Ejercicio 12 ###
######################################################

################################################
############### TABLA 1 ######################
################################################
T1<- data.frame(
        "reclamaciones"=0:6,
        "frecuencias"=c(65623,12571,1644,148,13,1,0)
              )

#Utilizamos la formula k*p_k/p_k_-1
pt1<-c()
for(i in 2:7){
  pt1[i]<-T1[i,1]*T1[i,2]/T1[i-1,2]
}
```

```
pt1
#Con los datos que tenemos grafiquemos y tratemos de ajustar alguna distirbución
plot(pt1,col="blue",ann=FALSE)
title(main = "Aproximación de la distribución por medio de las pk",
              ylab = "k*pk/p(k-1)",
              xlab = "k")

#Por el tipo de pendiente asumiremos Binomial negativa,estimemos sus parametros
x_aux=0
for (i in 1:7) {
  x_aux=x_aux+(T1[i,1]*T1[i,2])
}
x_esp=x_aux/80000

x_aux2=0
for (i in 1:7) {
  x_aux2=x_aux2+((T1[i,1]^2)*T1[i,2])
}
x_2m=x_aux2/80000
#Utlizamos la formula recursiva
beta_estimada<-((x_2m-x_esp^2)/x_esp)-1
k_estimada<-x_esp^2/(x_2m-x_esp-x_esp^2)

#Calculemos nuestra a & b
a<-beta_estimada/(1+beta_estimada)
b<-(k_estimada-1)*(beta_estimada/(1+beta_estimada))

#Ahora nuestras respectivas probabilidades
p<-matrix(nrow = 7,ncol = 1)
p[1,1]<-(1+beta_estimada)^-k_estimada
for(i in 1:6){
  p[1+i,1]<-p[i,1]*(a+b*1/i)
}
#Hacemos el producto de las probabilidades por el total y comparamos con vlaores originales
valores_estimados<-data.frame("reclamaciones"=0:6,"frecuencia"=p*80000)
valores_estimados[,-1] <-round(valores_estimados[,-1],2)


################################################
################ TABLA 2 #######################
################################################

T2<- data.frame(
  "reclamaciones"=0:7,
  "frecuencias"=c(861,121,13,3,1,0,1,0)
)

#Utilizamos la formula k*p_k/p_k_-1
pt2<-c()
for(i in 2:8){
  pt2[i]<-T2[i,1]*T2[i,2]/T2[i-1,2]
}
```

```
pt2
#Con los datos que tenemos grafiquemos y tratemos de ajustar alguna distirbución
plot(pt2,col="red",ann=FALSE)
title(main = "Aproximación de la distribución por medio de las pk",
      ylab = "k*pk/p(k-1)",
      xlab = "k")

#Por el tipo de pendiente asumiremos Binomial negativa,estimemos sus parametros
y_aux=0
for (i in 1:7) {
  y_aux=y_aux+(T2[i,1]*T2[i,2])
}
y_esp=y_aux/1000

y_aux2=0
for (i in 1:7) {
  y_aux2=y_aux2+((T2[i,1]^2)*T2[i,2])
}
y_2m=y_aux2/1000
#Utlizamos la formula recursiva
beta2_estimada<-((y_2m-y_esp^2)/y_esp)-1
k2_estimada<-y_esp^2/(y_2m-y_esp-y_esp^2)

#Calculemos nuestra a & b
a2<-beta2_estimada/(1+beta2_estimada)
b2<-(k2_estimada-1)*(beta2_estimada/(1+beta2_estimada))
#Ahora nuestras respectivas probabilidades
p2<-matrix(nrow = 8,ncol = 1)
p2[1,1]<-(1+beta2_estimada)^-k2_estimada
for(i in 1:7){
  p2[1+i,1]<-p2[i,1]*(a2+b2*1/i)
}
#Hacemos el producto de las probabilidades por el total y comparamos con valores originales
valores_estimados2<-data.frame("reclamaciones"=0:7,"frecuencia"=p2*1000)
valores_estimados2[,-1] <-round(valores_estimados2[,-1],2)
```

Ejercicio 14

```
#"""Si N  Geometrica con media 2, describe las funciones de densidad,
distribución y esperanza Zero Truncada y ZeroModificada si PM0 =16. Calcula los primeros 5 valores pa


# Función de densidad
f<-dgeom(0:5, 1/3)
# E[X]
sum(c(0:5)*f)

# Función de distribución
f2<-pgeom(0:5, 1/3)

#Funcion de distribuci on de la clase (a,b,0):
beta<-2; r<-1
a<-beta/(1+beta);b<-0
```

```
P<-numeric()
P[1]<-(1+beta)^(-r)
for(k in 1:5){P[k+1]<-P[k]*(a+b*(1/k))}
# E[X]
sum(c(0:5)*P)

### Distribucion truncada en cero. (PT0=0)
PT<-numeric()
PT[1]<-P[2]/(1-P[1])
for(k in 2:5){PT[k]<-(a+b/k)*PT[k-1]}
PT<-c(0,PT)
# E[X]
sum(c(0:5)*PT)

#### Distribucion modificada en cero (PM0=0.6)
PM<-numeric()
PM[1]<-(1-(1/6))*P[2]/(1-P[1])
for(k in 2:5){PM[k]<-(a+b/k)*PM[k-1]}
PM<-c(1/6,PM)
# E[X]
sum(c(0:5)*PM)

plot(0:5, P, type="o", col ="dodgerblue1", lwd =2,
     main = "Funciones de densidad", xlab= "k", ylab ="P")
lines(0:5, PT, type="o", col ="turquoise4", lwd =2)
lines(0:5, PM, type="o", col ="mediumvioletred", lwd =2)
legend(3.5, .3, c("(a,b,0)", "Zero truncada", "Zero modificada" ),
       col = c("dodgerblue1", "turquoise4","mediumvioletred"), lty =1, cex=.5)

plot(0:5, f2, type="o", col ="dodgerblue1", lwd =2,
     main = "Funciones de distribución ", xlab= "k", ylab ="P")
lines(0:5, cumsum(PT), type="o", col ="turquoise4", lwd =2)
lines(0:5, cumsum(PM), type="o", col ="mediumvioletred", lwd =2)
legend(3., .55, c("(a,b,0)", "Zero truncada", "Zero modificada" ),
       col = c("dodgerblue1", "turquoise4","mediumvioletred"), lty =1, cex=.5)
```

Ejercicio 16

```
######################################################
### Tarea Examen 1. Teoría del riesgo. Ejercicio 16 ###
######################################################

library(actuar)
par(mfrow=c(2,2))
#Para la U(5000) con limte=4500
#Funcion para la densidad de Yl
Yl_u<-coverage(dunif,punif,limit=4500,per.loss=T)

#Graficamos la Uniforme(0,5000)
curve(dunif(x,0,5000),xlim=c(0,5000),col="deepskyblue",lwd=4)
#Graficamos la densidad de Yl
curve(Yl_u(x,0,5000),add=T, col="aquamarine",lwd=2)
```

```
title(main="Densidades uniforme(5000) vs Yl",sub="Yl uniforme",col.sub="aquamarine" )

##Funcion para la densidad de Yp#
Yp_u<-coverage(dunif,punif,limit=4500)
#Graficamos la Uniforme(0,5000)#
curve(dunif(x,0,5000),xlim=c(0,5000),col="deepskyblue",lwd=4)
#Graficamos la densidad de Yp
curve(Yp_u(x,0,5000),add=T, col="cyan1",lwd=2)
title(main="Densidades uniforme(5000) vs Yp",sub="Yp uniforme",col.sub="cyan1" )

#Funcion para la densidad de Yl
Yl_u<-coverage(dunif,punif,limit=4500,per.loss=T)

#Graficamos la Uniforme(0,5000)
curve(dunif(x,0,5000),xlim=c(0,5000),ylim=c(0,0.1),col="deepskyblue",lwd=4)
#Graficamos la densidad de Yl
curve(Yl_u(x,0,5000),add=T, col="aquamarine",lwd=2)
title(main="Densidades uniforme(5000) vs Yl",sub="Yl uniforme",col.sub="aquamarine" )

##Funcion para la densidad de Yp#
Yp_u<-coverage(dunif,punif,limit=4500)
#Graficamos la Uniforme(0,5000)#
curve(dunif(x,0,5000),xlim=c(0,5000),ylim=c(0,0.1),col="deepskyblue",lwd=4)
#Graficamos la densidad de Yp
curve(Yp_u(x,0,5000),add=T, col="cyan1",lwd=2)
title(main="Densidades uniforme(5000) vs Yp",sub="Yp uniforme",col.sub="cyan1" )



##Para la exp(media=5000) con limite=45000##
par(mfrow=c(2,2))
#Funcion para la densidad de Yl
Yl_e<-coverage(dexp,pexp,limit=4500, per.loss=T)
#Graficamos exp(5000)
curve(dexp(x,1/5000),xlim=c(0,4600),col="deepskyblue",lwd=4)
#Graficamos la densidad de Yl
curve(Yl_e(x,1/5000),add=T,col="darkolivegreen1")
title(main="Densidades exp(5000) vs Yl",sub="Yl exp",col.sub="darkolivegreen1" )

#Funcion para la densidad de Yp
Yp_e<-coverage(dexp,pexp,limit=4500)
#Graficamos exp(5000)
curve(dexp(x,1/5000),xlim=c(0,4600),col="deepskyblue",lwd=4)
#Graficamos la densidad de Yp
curve(Yp_e(x,1/5000),add=T, col="darkgreen")
title(main="Densidades exp(5000) vs Yp",sub="Yp exp",col.sub="darkgreen")


#Funcion para la densidad de Yl
Yl_e<-coverage(dexp,pexp,limit=4500, per.loss=T)
#Graficamos exp(5000)
```

```
curve(dexp(x,1/5000),xlim=c(0,5000),ylim=c(0,0.4),col="deepskyblue",lwd=4)
#Graficamos la densidad de Yl
curve(Yl_e(x,1/5000),add=T,col="darkolivegreen1")
title(main="Densidades exp(5000) vs Yl",sub="Yl exp",col.sub="darkolivegreen1" )

#Funcion para la densidad de Yp
Yp_e<-coverage(dexp,pexp,limit=4500)
#Graficamos exp(5000)
curve(dexp(x,1/5000),xlim=c(0,5000),ylim=c(0,0.4),col="deepskyblue",lwd=4)
#Graficamos la densidad de Yp
curve(Yp_e(x,1/5000),add=T, col="darkgreen")
title(main="Densidades exp(5000) vs Yp",sub="Yp exp",col.sub="darkgreen")
```

Ejercicio 20

```
######################################################
### Tarea Examen 1. Teoría del riesgo. Ejercicio 20 ###
######################################################

library(actuar)
par(mfrow=c(1,2))
#Para la Pareto(0.78,100) con d=120,u=1000,alpha=0.8, r=0.05
#Funcion para la densidad de Yl
Yl_p<-coverage(dpareto,ppareto,deductible = 120, limit=1000, inflation = 0.05 ,per.loss=T)

#Graficamos la Pareto(0.78,100)
curve(dpareto(x,0.78,100),xlim=c(0,800),ylim=c(0,0.01),col="deepskyblue",lwd=4)
#Graficamos la densidad de Yl
curve(Yl_p(x,0.78,100),add=T, col="aquamarine",lwd=2)
title(main="Densidades pareto(0.78,100) vs Yl",sub="Yl pareto",col.sub="aquamarine" )

##Funcion para la densidad de Yp
Yp_p<-coverage(dpareto,ppareto,deductible = 120, limit=1000, inflation = 0.05)
#Graficamos la Pareto(0.78,100)
curve(dpareto(x,0.78,100),xlim=c(0,800),ylim=c(0,0.01),col="deepskyblue",lwd=4)
#Graficamos la densidad de Yp
curve(Yp_p(x,0.78,100),add=T, col="cyan1",lwd=2)
title(main="Densidades pareto(0.78,100) vs Yp",sub="Yp pareto",col.sub="cyan1" )


    ## Ejercicio 11

library(readxl)
library(ggplot2)
library(fitdistrplus)

## Se cargan los archivos
setwd(dirname(rstudioapi::getActiveDocumentContext()$path))
rmd<-read.csv("./RModuleDay.csv")
## se verifican las medidas descripivas
str(rmd)
summary(rmd)
```

```
summary(rmd$Losses_in_Thousands)

## Ajuste de encabezados
colnames(rmd)<-c("Ac_No","Age","Years_of_Experience","Number_of_Vehicles","Gender","Married", "Losses


## Metodo de momentos
m1<-sum(rmd$Losses_in_Thousands)/length(rmd$Losses_in_Thousands)
m2<-sum(rmd$Losses_in_Thousands*rmd$Losses_in_Thousands)/length(rmd$Losses_in_Thousands)

meanl<-log(m1)-(log(m2/(m1*m1))/2 )
sdl<-log(m2/(m1*m1))

## Con la libreria fitdistrplus
fitdist(rmd$Losses_in_Thousands,"lnorm","mle")

hist(rmd$Losses_in_Thousands,breaks=50,freq=F,col="green")
n<-seq(0,max(rmd$Losses_in_Thousands),by=.2)
## Calculo de probabilidades
lines(n,dlnorm(n,meanlog = meanl,sdlog = sqrt(sdl)),col="purple",lwd=2.5)

## Para el calculo de QQPlot
QQ<-fitdist(rmd$Losses_in_Thousands,"lnorm")
plot(QQ)
QQ
```