Lista de Exercícios 1

Gabriel D'assumpção de Carvalho

2025-03-16

Contents

1 Bibliotecas

```r
#install.packages('moments')
library(ggplot2)
library(actuar)
library(moments)
```

2 Questão 1

Uma amostra de pagamentos de contratos de seguro saúde produziu uma média amostral anual de \$1, 300 e um desvio padrão de \$400 . Para um próximo ano, espera-se comercializar uma carteira com 2500 clientes. Utilize o Teorema do Limite Central a fimm de estimar a probabilidade de que os pagamentos em benefícios ultrapasse o valor esperado em 1%.

```r
n = 2500
Ex = 1300
dp = 400

Tdp = dp * sqrt(n)
Tloss = n * Ex

b1 = Tloss * 1.01

z = (b1 - Tloss)/Tdp

p1 = 1-pnorm(z)
print(p1) # 0.05208128
```

```
## [1] 0.05208128
```

Como o teorema do limite central afirma que uma amostra tendendo ao infinito segue uma N(0,1), a probabilidade do pagamento do benefício ultrapassar 1% do valor esperado (\$3.282.500) é de aproximadamente 5.21%.

3 Questão 2

Analise as distribuições lognormal, Pareto (Tipo II) e Gama com respeito ao comportamento caudal. Como sugestão, utilize as distribuições Gamma (0.2, 500) , lognormal(3.709290, 1.338566^2) e Pareto (2.5, 150). Primeiro mostre que, assim definidas, estas distribuições possuem médias e desvios padrões iguais. Em seguinda, utilize o computador para avaliar o comprimento caudal destas distribuições, estabelecendo uma ordem, da 'mais densa' para a 'mais leve'. Adicionalmente, esboce os gráficos das funções de sobrevivência e taxa de risco para cada um dos casos analisados.

```r
n = 10000

g_shape = 0.2
g_scale = 500
```

```r
l_mean = 3.709290
l_var = 1.338566**2
l_sd = sqrt(l_var)

p_shape = 2.5
p_scale = 150

gamma_sample = rgamma(n, shape = g_shape, scale = g_scale)
lognormal_sample = rlnorm(n, meanlog = l_mean, sdlog=l_sd)
pareto_sample = rpareto(n, shape = p_shape, scale = p_scale)

df = data.frame(Distribution = c('Gamma', 'Lognormal', 'Pereto'), Mean = c(mean(gamma_sample), mean(log


print(df)
```

```
##   Distribution      Mean StandardDeviation
## 1        Gamma 101.55416          225.7473
## 2    Lognormal 102.90497          244.5400
## 3       Pereto  99.01495          186.4999
```

No Gráfico acima podemos ver as distribuições e suas respectivas médias e desvio padrão, foram geradas 10 mil de amostras aleatórios para cada uma das distribuição, podemos ver que a média se aproxima de 100 e o desvio padrão de 223, podemos ver uma certa diferênça no desvio padrão da lognormal e pareto por serem distribuição mais complexar de se contrar a sua variância.

```r
kurtosi_gamma = kurtosis(gamma_sample)
kurtosi_lognormal= kurtosis(gamma_sample)
kurtosi_pareto = kurtosis(pareto_sample)

df = data.frame(Distribution = c('Gamma', 'Lognormal', 'Pereto'), Kurtosi = c(kurtosi_gamma, kurtosi_lo

print(df)
```

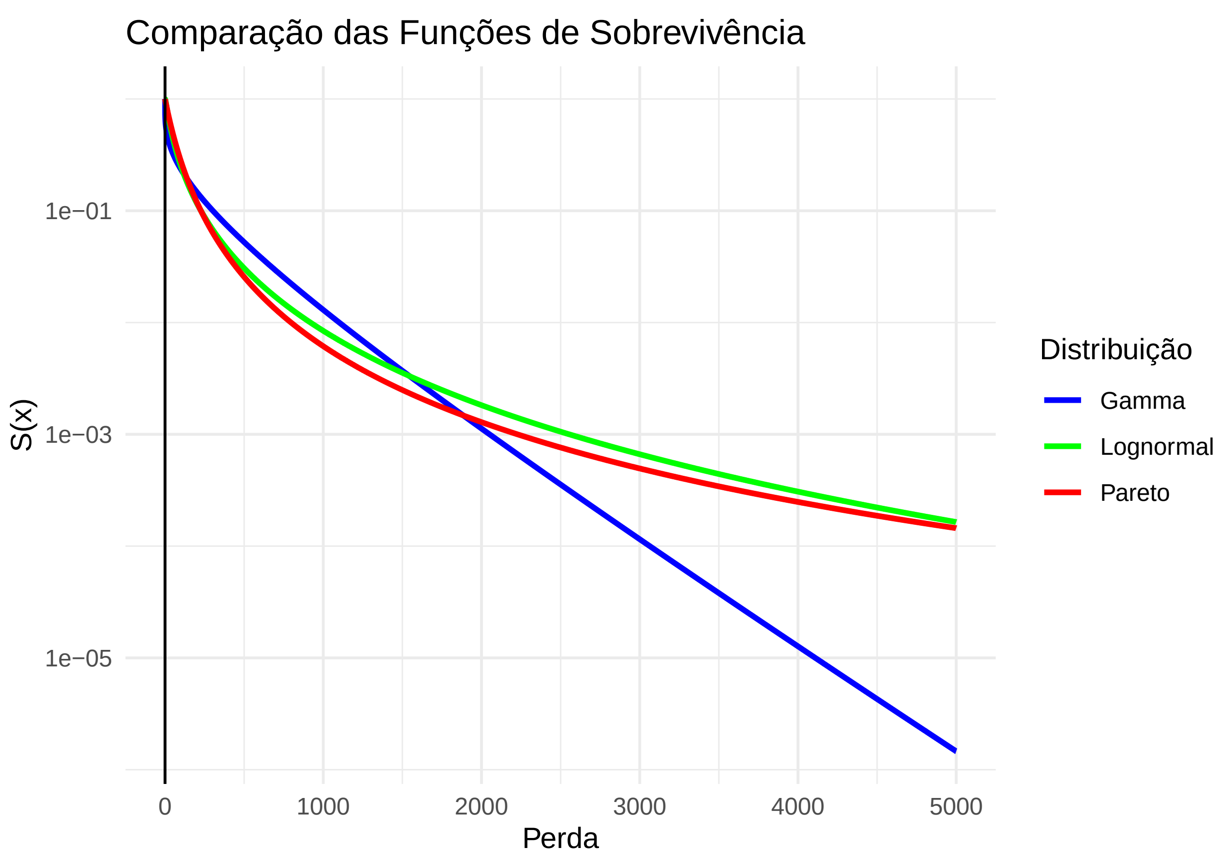
```
##   Distribution   Kurtosi
## 1        Gamma  30.94557
## 2    Lognormal  30.94557
## 3       Pereto 195.08316
```

Como pode ser visto na tabela acima, foi utilizado a kurtosi que á a razão entre o quarto momento da média e do desvio padrão, ela mede o achatamento da distribuição, portanto quanto maior essa estatistica maior vai ser as probabilidade das extremidades.

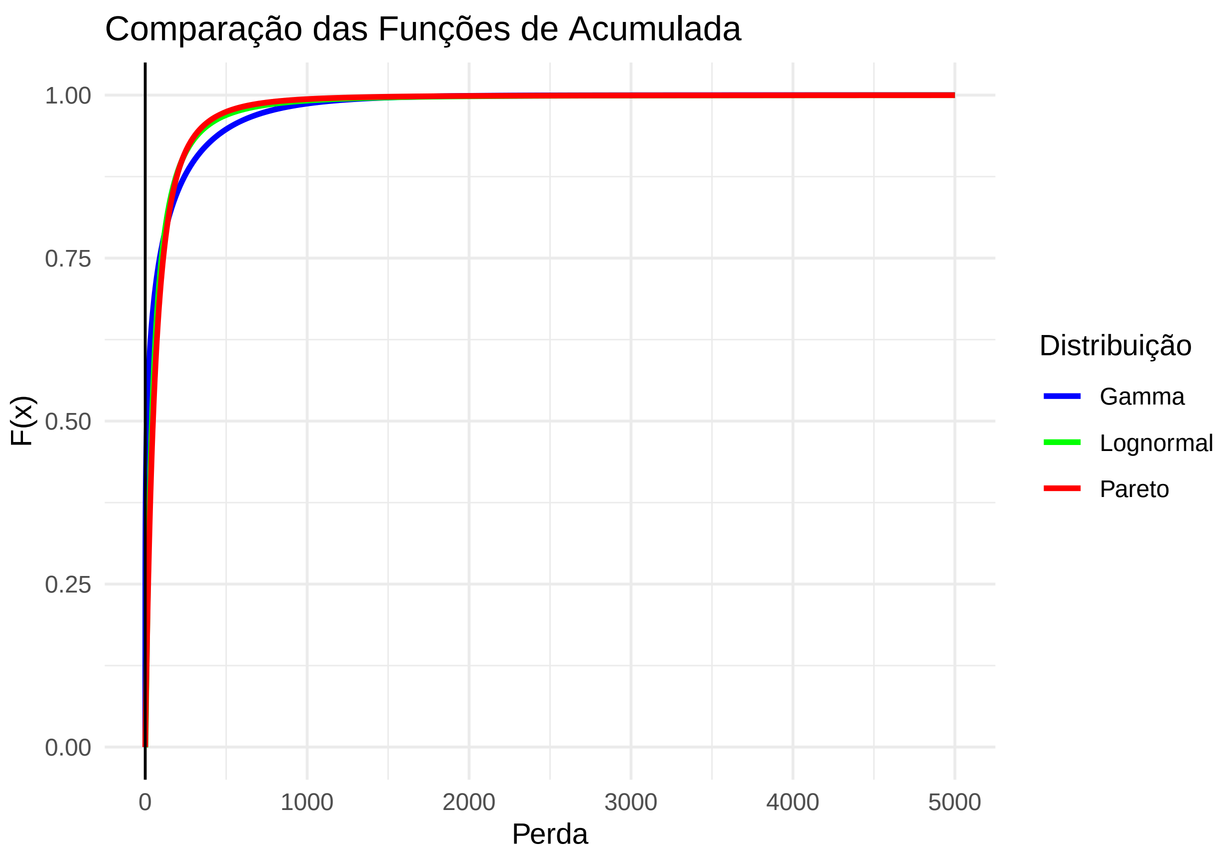
```r
q = seq(0, 5000, length.out = n)

S_gamma =  1 - pgamma(q, shape = g_shape, scale = g_scale)
F_gamma = 1 - S_gamma

S_lognormal = 1 - plnorm(q, meanlog = l_mean, sdlog = l_sd)
F_lognormal = 1 - S_lognormal

S_pareto = 1 - ppareto(q, shape = p_shape, scale = p_scale)
```

```r
F_pareto = 1 - S_pareto

# Criar um dataframe para evitar erro de dimensionamento
df = data.frame(
  q = rep(q, 3),
  S = c(S_gamma, S_pareto, S_lognormal),
  F = c(F_gamma, F_pareto, F_lognormal),
  Distribuição = rep(c("Gamma", "Pareto", "Lognormal"), each = n)
)
```

```r
ggplot(df, aes(x = q, y = S, color = Distribuição)) +
  geom_line(size = 1) +
  geom_vline(xintercept = 0, color = "black", size = 0.5, linetype = "solid") +
  labs(
    title = "Comparação das Funções de Sobrevivência",
    x = "Perda",
    y = "S(x)",
    color = "Distribuição"
  ) +
  scale_color_manual(values = c("Gamma" = "blue", "Pareto" = "red", "Lognormal" = "green")) +
  scale_y_log10() +  # Escala log para destacar diferenças nas caudas
  theme_minimal()
```



```r
ggplot(df, aes(x = q, y = F, color = Distribuição)) +
  geom_line(size = 1) +
  geom_vline(xintercept = 0, color = "black", size = 0.5, linetype = "solid") +
  labs(
    title = "Comparação das Funções de Acumulada",
    x = "Perda",
    y = "F(x)",
    color = "Distribuição"
  ) +
  scale_color_manual(values = c("Gamma" = "blue", "Pareto" = "red", "Lognormal" = "green")) +
  #scale_y_log10() +  # Escala log para destacar diferenças nas caudas
  theme_minimal()
```



4 Questão 3

Considere a variável aleatória X tenho função densidade de probabilidade defnida por $f(x) = (1+2x^2)e^{-2x}, \quad x \leq 0$.

(i) Determine a função de sobrevivência S(x).

Resposta:

$$S(x) = 1 - \int_0^x f(t)dt = 1 - \int_0^x (1+2t^2)e^{-2t}dt. \tag{1}$$

(ii) Determine a função taxa de risco h(x).

Resposta:

$$h(x) = \frac{f(x)}{S(x)} = \frac{S'(x)}{S(x)} = -\frac{d}{dx}\log(S(x)). \tag{2}$$

(iii) Determine a função $S_e(x)$ associada a 'distribuição de equilíbrio'.

Resposta:

$$S_e(x) = \frac{\int_x^\infty S(t), dt}{E[X]} = \frac{\int_x^\infty S(t), dt}{e(0)}. \tag{3}$$

(iv) Determine a função de vida média residual e(x).

Resposta:

$$e(x) = \frac{\int_x^\infty tf(t), dt}{S(x)}. \tag{4}$$

(v) Determine os limites $\lim_{x\to\infty}$ h(x) e $\lim_{x\to\infty} e(x)$.

Resposta:

(vi) Mostre que $e(x)$ é estritamente decrescente, mas $h(x)$ não é estritamente crescente.

Resposta:

(vii) Esboce graficamente os resultados encontrados nos itens (i) a (vi).