

Análise de Sobrevida
Teoria e Aplicações em Saúde

Caderno de Respostas

Marilia Sá Carvalho
Valeska Lima Andreozzi
Claudia Torres Codeço
Maria Tereza Serrano Barbosa
Silvia Emiko Shimakura

2

O tempo na análise de sobrevivida

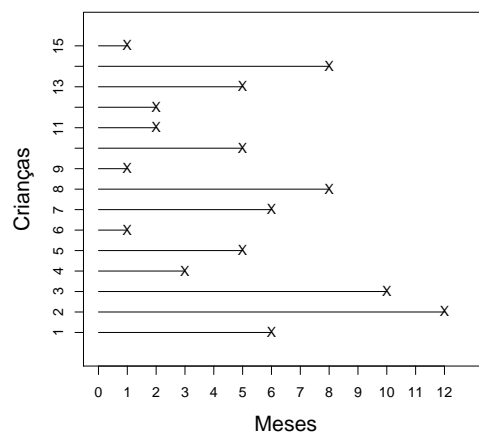
Exercícios

Exercício 2.1: O tempo de aleitamento, isto é, o tempo decorrido desde o nascimento até o desmame, pode ser considerado como uma variável tempo de sobrevivida. Suponha que o tempo até o desmame, em meses, tenha sido registrado para 15 crianças:

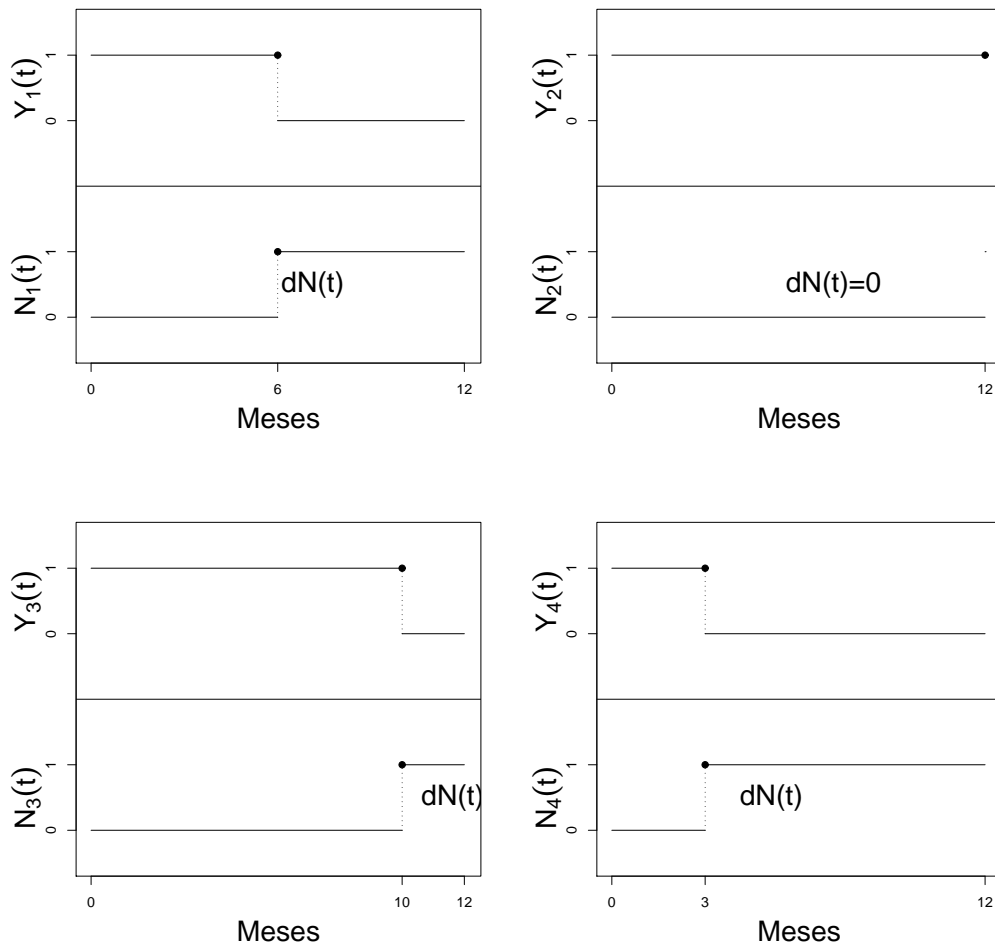
6 12 10 3 5 1 6 8 1 5 2 2 5 8 1

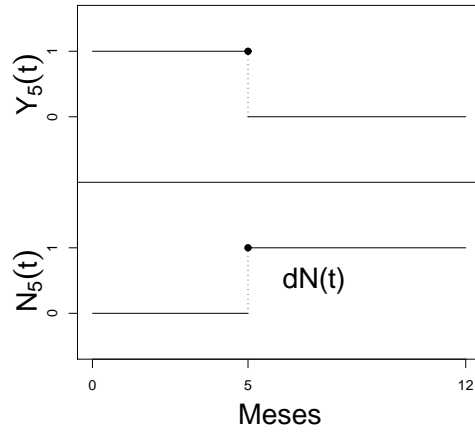
Considerando que não houve censura:

1. Represente graficamente os tempos de observação das 15 crianças.



2. Represente as trajetórias dos primeiros cinco indivíduos utilizando as variáveis $N(t)$ e $Y(t)$ do processo de contagem.





3. Como você construiria um banco de dados para analisar estes dados pelo processo clássico?

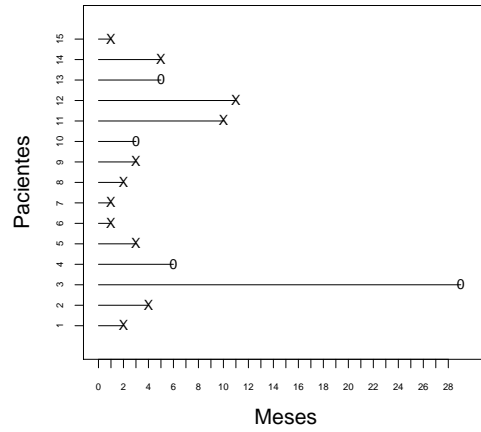
crianca	tempo	status
1	6	1
2	12	1
3	10	1
4	3	1
5	5	1
6	1	1
7	6	1
8	8	1
9	1	1
10	5	1
11	2	1
12	2	1
13	5	1
14	8	1
15	1	1

Exercício 2.2: Considere agora o tempo de sobrevivência de 15 pacientes submetidos a hemodiálise. Neste caso, a variável de interesse é o tempo desde a primeira diálise até o óbito.

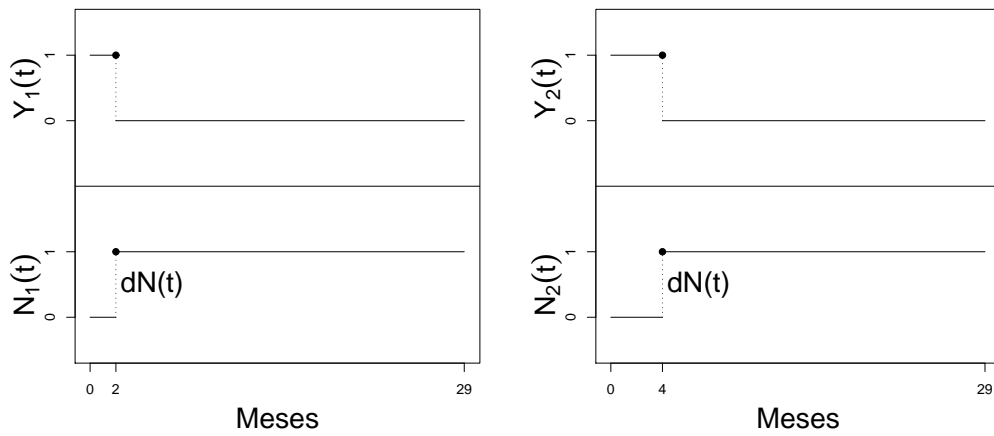
2 4 29+ 6+ 3 1 1 2 3 9+ 10 11 5+ 5 1

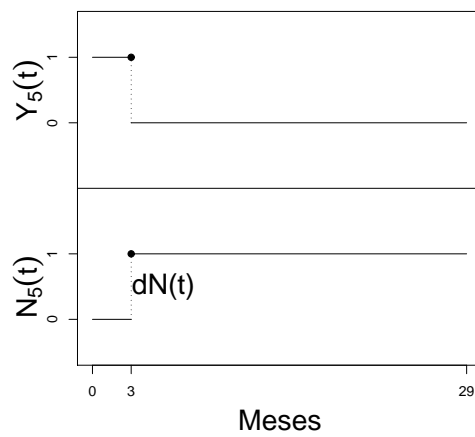
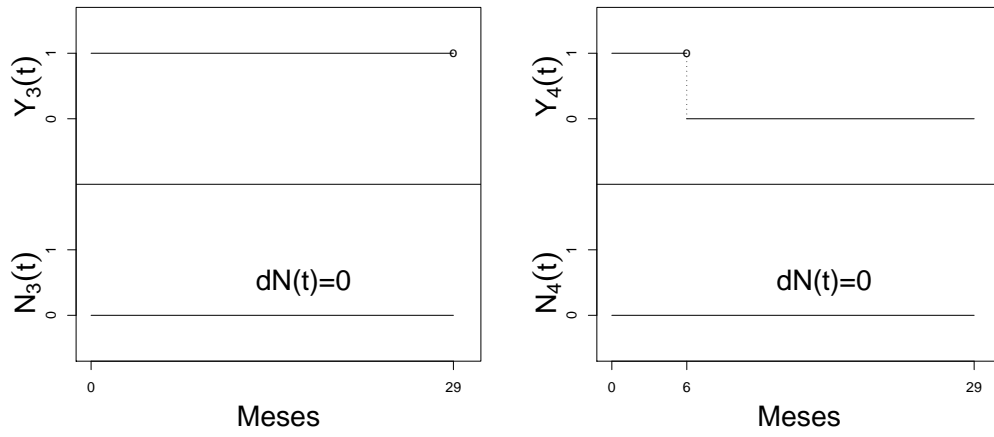
Os tempos censurados (censura à direita) estão indicados pelo sinal de +. Considere que todos os pacientes entraram juntos no início do estudo.

1. Represente graficamente os tempos de observação dos pacientes, utilizando abordagem clássica.



2. Represente as trajetórias dos primeiros cinco indivíduos utilizando as variáveis $N(t)$ e $Y(t)$ do processo de contagem.



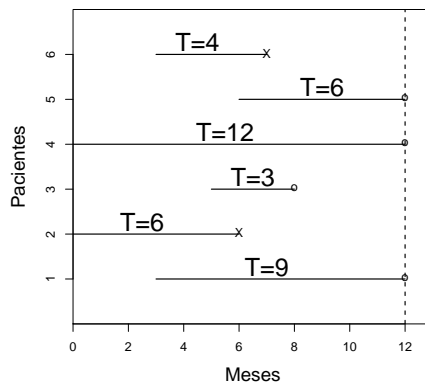


3. Como você construiria um banco de dados para analisar esses dados pelo processo clássico?

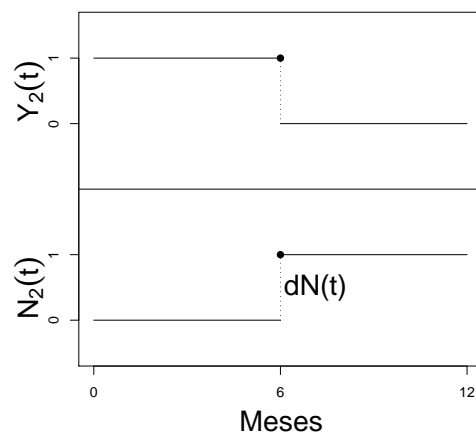
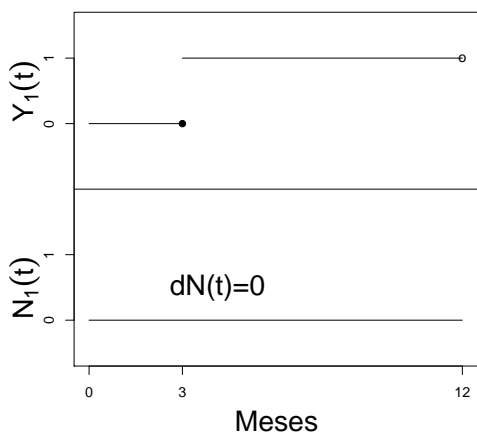
paciente	tempo	status
1	2	1
2	4	1
3	29	0
4	6	0
5	3	1
6	1	1
7	1	1
8	2	1
9	3	1
10	9	0

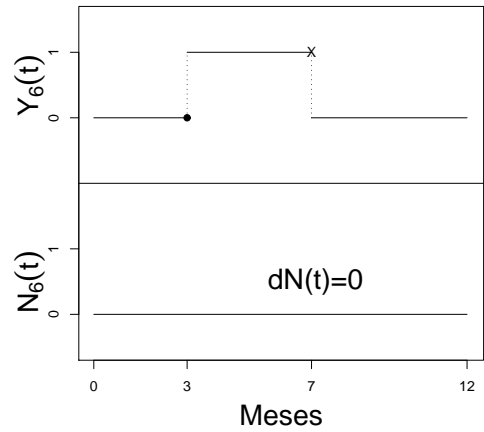
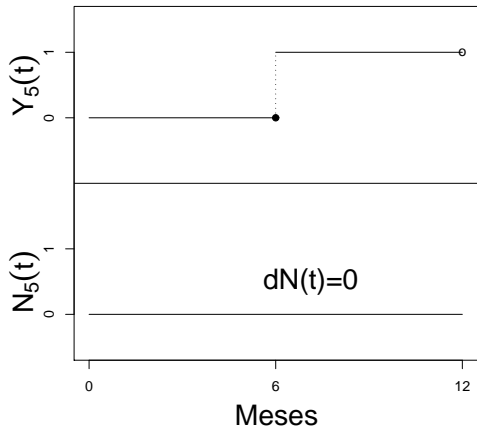
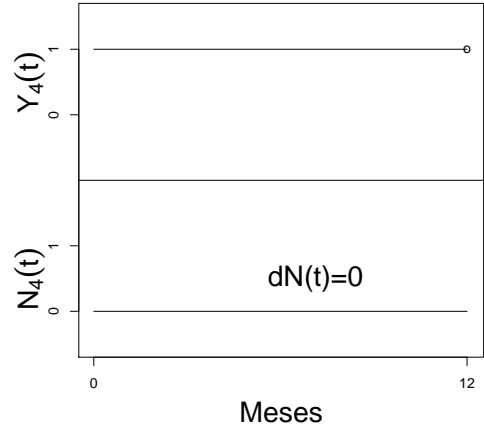
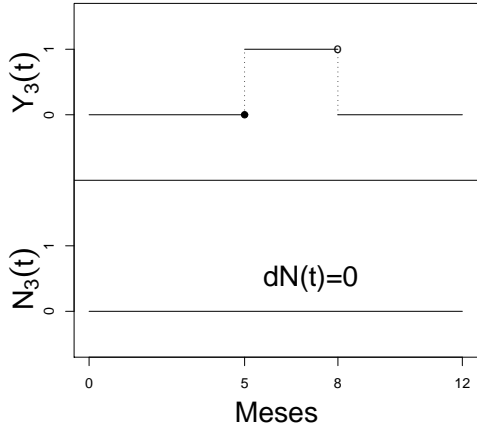
11	10	1
12	11	1
13	5	0
14	5	1
15	1	1

Exercício 2.3: Suponha que, em um hospital, 6 pacientes HIV positivo são acompanhados ao longo de um ano. No gráfico abaixo, as linhas horizontais representam o tempo de acompanhamento de cada paciente (linhas terminadas em \times indicam a ocorrência do desfecho (óbito), linhas terminadas com \circ indicam censuras). Represente, utilizando as variáveis $N(t)$ e $Y(t)$, as trajetórias de cada um dos 6 pacientes.

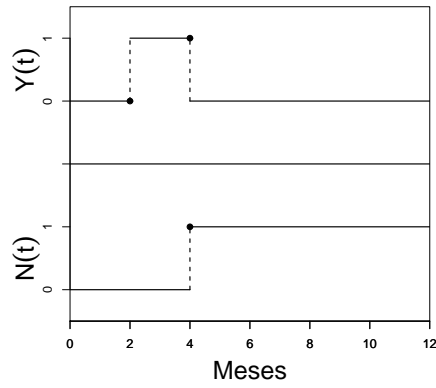


A seguir os gráficos dos 6 pacientes utilizando o processo de contagem





Exercício 2.4: Um paciente tem a seguinte trajetória de observação, segundo o processo de contagem:



Com base neste gráfico, responda:

1. Qual foi o mês de entrada do paciente no estudo?

Resposta: Mês dois

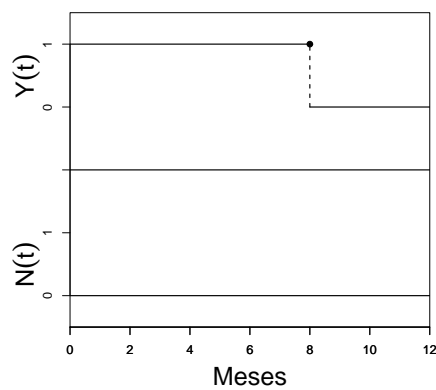
2. Em que mês ocorreu o desfecho?

Resposta: Mês quatro

3. Quais os meses em que o paciente estava sob risco de óbito?

Resposta: Meses dois, três e quatro.

Exercício 2.5: Um outro paciente tem a seguinte trajetória de observação:



Com base neste gráfico, responda:

1. Qual foi a data de entrada do paciente no estudo?

Resposta: Mês zero

2. Em que data ocorreu o desfecho?

Resposta: Mês oito

Exercício 2.6: No R, abra o banco de dados dos pacientes com Aids, atendidos no Ipec (no Apêndice C1 você encontra uma descrição deste banco de dados). Este banco está em formato *csv* e também pode ser visualizado em um programa de planilha. Lembre-se de indicar ao R em que diretório está o arquivo com o banco de dados usando o menu *File > Change Dir*.

Os dados estão disponíveis em <http://dengue.procc.fiocruz.br/~sobrevida/dados/>. Copie o banco *ipec.csv* para este mesmo diretório de trabalho.

```
> ipec <- read.table("ipec.csv", header = T, sep = ";")
> ipec[1:5, ]
```

	id	ini	fim	tempo	status	sexo	escola	idade	risco	acompan	obito	anotrat
1	1	1243	2095	852	1	M	3	34	0	1	S	1991
2	2	2800	2923	123	1	M	2	38	6	1	S	9
3	3	1250	2395	1145	1	M	NA	32	0	1	S	1992
4	4	1915	4670	2755	0	M	NA	43	6	0	N	1992
5	5	2653	4770	2117	0	M	NA	40	0	1	N	1992

	tratam	doenca	propcp
1	1	4	3
2	0	7	4
3	1	3	4
4	1	10	4
5	1	5	4

Crie a variável resposta da análise de sobrevivência, combinando as informações de tempo e status no formato clássico e no formato de processo de contagem. Não esqueça de chamar a biblioteca *survival*.

```
> require(survival)
```

Formato clássico:

```
> Surv(ipec$tempo, ipec$status)
```

```

[1] 852 123 1145 2755+ 2117+ 329+ 60 151 1563 1247 84 214
[13] 25+ 1348 158 555 408 1116 998 1125 944+ 54 151 855
[25] 116 80+ 1757+ 194 183 37 237+ 1506 168+ 134 803+ 18
[37] 371 173 688 163 3178+ 29 50+ 887 516 645 310 204
[49] 1344+ 1261 285 83 150 1307+ 1076+ 1226 865+ 811 2898 80
[61] 967 618 235 2236+ 152 892 81+ 1085 1073+ 1615+ 35 290
[73] 1780+ 3228+ 52 733 3213+ 1983 2304+ 572 21 1272+ 1646+ 304
[85] 418 854 2973+ 40 850 1139 323 1507+ 2717+ 1735+ 388+ 145
[97] 905 927 1027+ 631 2495+ 1331+ 623 2568+ 2013+ 721 1952+ 397
[109] 254 1630+ 1523+ 146+ 108 1835+ 499 333 202+ 2437+ 1015 2138+
[121] 22 2090+ 179 2439+ 1063+ 85+ 343+ 2215+ 259 2258+ 1371 39
[133] 2371+ 975+ 952 2492+ 1478+ 295+ 992 1011+ 644 426 537+ 1454+
[145] 1869+ 714+ 1310+ 2084+ 1918+ 1649+ 290+ 1685+ 1348+ 652+ 1384+ 1471+
[157] 1512+ 378+ 1352+ 419 1426+ 1488+ 1315+ 643+ 1197+ 1343+ 1176+ 944
[169] 340 881+ 915+ 948+ 985+ 1242+ 955+ 987+ 899+ 1056+ 775 785+
[181] 731+ 16 680+ 21+ 444+ 524+ 217+ 440+ 470+ 390+ 344+ 578+
[193] 504+

```

Formato de processo de contagem:

```
> Surv(ipec$ini, ipec$fim, ipec$status)
```

```

[1] (1243,2095 ] (2800,2923 ] (1250,2395 ] (1915,4670+] (2653,4770+]
[6] ( 3, 332+] ( 36, 96 ] ( 1, 152 ] ( 544,2107 ] ( 71,1318 ]
[11] ( 946,1030 ] ( 802,1016 ] ( 266, 291+] (1544,2892 ] ( 57, 215 ]
[16] (1270,1825 ] (2753,3161 ] ( 940,2056 ] ( 393,1391 ] (1000,2125 ]
[21] ( 238,1182+] ( 423, 477 ] ( 206, 357 ] ( 480,1335 ] ( 226, 342 ]
[26] ( 249, 329+] (3052,4809+] (1802,1996 ] (1395,1578 ] ( 354, 391 ]
[31] ( 493, 730+] (1113,2619 ] ( 638, 806+] ( 655, 789 ] (1189,1992+]
[36] ( 943, 961 ] (1715,2086 ] ( 792, 965 ] (1037,1725 ] ( 820, 983 ]
[41] ( 884,4062+] (2262,2291 ] (1121,1171+] (1131,2018 ] ( 878,1394 ]
[46] (1316,1961 ] (1107,1417 ] (1190,1394 ] ( 393,1737+] (1274,2535 ]
[51] (1172,1457 ] (2360,2443 ] (2074,2224 ] (1019,2326+] ( 605,1681+]
[56] (1915,3141 ] (3948,4813+] (1314,2125 ] (1502,4400 ] (1347,1427 ]
[61] (1379,2346 ] (2352,2970 ] (2625,2860 ] (2586,4822+] (1406,1558 ]
[66] (1466,2358 ] (3314,3395+] (3413,4498 ] (3712,4785+] (3207,4822+]
[71] (1592,1627 ] (1537,1827 ] (3018,4798+] (1555,4783+] (1541,1593 ]
[76] (1589,2322 ] (1609,4822+] (1682,3665 ] (2465,4769+] (1243,1815 ]
[81] (1667,1688 ] (1605,2877+] (3157,4803+] (2066,2370 ] (1929,2347 ]
[86] (2216,3070 ] (1809,4782+] (1670,1710 ] (1983,2833 ] (2883,4022 ]
[91] (1766,2089 ] (3313,4820+] (1977,4694+] (3087,4822+] (2286,2674+]
[96] (1877,2022 ] (1852,2757 ] (1549,2476 ] (3795,4822+] (2475,3106 ]
[101] (2310,4805+] (2870,4201+] (1935,2558 ] (2199,4767+] (2800,4813+]
[106] (2990,3711 ] (2857,4809+] (3586,3983 ] (2143,2397 ] (3124,4754+]
[111] (3276,4799+] (2208,2354+] (2209,2317 ] (2976,4811+] (2626,3125 ]
[116] (3838,4171 ] (2314,2516+] (2311,4748+] (2280,3295 ] (2684,4822+]

```

```

[121] (2454,2476 ] (2713,4803+] (2311,2490 ] (2370,4809+] (3756,4819+]
[126] (2565,2650+] (2599,2942+] (2553,4768+] (2601,2860 ] (2553,4811+]
[131] (2726,4097 ] (2739,2778 ] (2447,4818+] (3830,4805+] (2429,3381 ]
[136] (2311,4803+] (3299,4777+] (4510,4805+] (2384,3376 ] (3749,4760+]
[141] (2676,3320 ] (2985,3411 ] (4192,4729+] (3159,4613+] (2921,4790+]
[146] (4078,4792+] (2934,4244+] (2645,4729+] (2857,4775+] (3173,4822+]
[151] (4509,4799+] (3082,4767+] (3465,4813+] (3188,3840+] (3271,4655+]
[156] (3276,4747+] (3287,4799+] (4439,4817+] (3446,4798+] (3305,3724 ]
[161] (3391,4817+] (3307,4795+] (3425,4740+] (4117,4760+] (3612,4809+]
[166] (3479,4822+] (3572,4748+] (3796,4740 ] (3527,3867 ] (3921,4802+]
[171] (3798,4713+] (3808,4756+] (3772,4757+] (3557,4799+] (3867,4822+]
[176] (3594,4581+] (3923,4822+] (3733,4789+] (4019,4794 ] (4033,4818+]
[181] (4040,4771+] (4053,4069 ] (4137,4817+] (4208,4229+] (4362,4806+]
[186] (4279,4803+] (4593,4810+] (4320,4760+] (4343,4813+] (4419,4809+]
[191] (4406,4750+] (4199,4777+] (4301,4805+]

```

Exercício 2.7: Para ir se familiarizando com o R e com o banco de dados do Ipec, que será utilizado nos próximos capítulos, faça uma análise exploratória das variáveis de interesse. A seguir estão alguns comandos necessários para arrumar o banco de dados e em seguida explorá-lo.

Lendo o banco de dados:

```
> ipec <- read.table("ipec.csv", header = T, sep = ";")
```

Listando os primeiros cinco registros:

```
> ipec[1:5, ]
```

```

  id ini  fim tempo status sexo escola idade risco acompan obito anostat
1  1 1243 2095   852     1    M     3    34     0     1     S    1991
2  2  2800 2923   123     1    M     2    38     6     1     S     9
3  3  1250 2395  1145     1    M    NA    32     0     1     S   1992
4  4  1915 4670  2755     0    M    NA    43     6     0     N   1992
5  5  2653 4770  2117     0    M    NA    40     0     1     N   1992
  tratam doenca propcp
1     1     4     3
2     0     7     4
3     1     3     4
4     1    10     4
5     1     5     4

```

Tamanho do banco de dados (número de registros e número de variáveis, respectivamente):

```
> dim(ipec)
```

```
[1] 193 15
```

Listando o nome das variáveis contidas no banco:

```
> names(ipec)
```

```
[1] "id"      "ini"     "fim"     "tempo"   "status"  "sexo"   "escola"  
[8] "idade"  "risco"   "acompan" "obito"   "anotrat" "tratam" "doenca"  
[15] "propcp"
```

No R, o código para informação ignorada é *NA*, logo precisamos substituir os códigos 9, 99 ou I por *NA*

```
> ipec$anotrat[ipec$anotrat == 9] <- NA  
> ipec$doenca[ipec$doenca == 99] <- NA  
> ipec$obito <- factor(ipec$obito, levels = c("N", "S"))
```

É preciso indicar ao R quais são as variáveis cujos valores numéricos representam categorias. Se você quiser indicar um *label* para estes fatores utilize o comando *factor* com a descrição dos *labels*, acompanhando a ordem numérica do registro. A descrição está na Tabela C.1.

```
> ipec$escola <- factor(ipec$escola, labels = c("sem", "fundam",  
+ "medio", "sup"))  
> ipec$risco <- factor(ipec$risco)  
> ipec$acompan <- factor(ipec$acompan)  
> ipec$anotrat <- factor(ipec$anotrat)  
> ipec$tratam <- factor(ipec$tratam)  
> ipec$doenca <- factor(ipec$doenca)  
> ipec$propcp <- factor(ipec$propcp)
```

ANÁLISE EXPLORATÓRIA

No R o nome das variáveis deve ser precedido do nome do objeto seguido do símbolo \$, como vínhamos usando anteriormente. Podemos encurtar essa sintaxe utilizando-se a função *attach()* permitindo, desta forma, que os nomes das variáveis do objeto *ipec* sejam usados diretamente.

```
> attach(ipec)
```

Cálculo das medidas resumo:

```
> summary(ipecc)
```

```
      id      ini      fim      tempo      status
Min.   : 1   Min.   : 1   Min.   : 96   Min.   : 16.0   Min.   :0.0000
1st Qu.: 49  1st Qu.:1406  1st Qu.:2095  1st Qu.: 290.0  1st Qu.:0.0000
Median : 97  Median :2454  Median :3711  Median : 852.0  Median :0.0000
Mean   : 97  Mean   :2397  Mean   :3335  Mean   : 938.2  Mean   :0.4663
3rd Qu.:145  3rd Qu.:3314  3rd Qu.:4790  3rd Qu.:1348.0  3rd Qu.:1.0000
Max.   :193  Max.   :4593  Max.   :4822  Max.   :3228.0  Max.   :1.0000
```

```
sexo      escola      idade      risco      acompan      obito      anotrat
F: 49     sem      :59   Min.   :20.00  0      :87   0:57     N      :80   1992   :22
M:144     fundam:44  1st Qu.:30.00  1      : 9   1:99     S      :92   1995   :20
          medio :55   Median :35.00  2      : 7   2:37     NA's:21  1993   :19
          sup   :24   Mean   :36.55  3      :30   1996   :18
          NA's  :11   3rd Qu.:43.00  5      :16   1994   :16
          Max.  :68.00  6      : 7   (Other):54
          NA's  :37   NA's   :44
```

```
tratam      doenca      propcp
0: 44     3      :31   0: 38
1:100    10      :25   2: 24
2: 35     7      :17   3:  3
3: 14     1      :12   4:128
          8      :12
          (Other):29
          NA's   :67
```

Para as variáveis categóricas podemos também usar a função `table()` para calcular as frequências de cada categoria:

```
> table(escola)
```

```
escola
  sem fundam medio  sup
   59    44    55   24
```

```
> table(tratam)
```

```
tratam
 0  1  2  3
44 100 35 14
```

Distribuição das idades dos pacientes por gênero:

```
> boxplot(idade ~ sexo, main = "Idade por Gênero", ylab = "Idade",  
+         xlab = "Sexo")
```

Número de pacientes por grupo de categoria provável de transmissão e por sexo:

```
> table(risco, sexo)
```

	sexo	
risco	F	M
0	0	87
1	2	7
2	2	5
3	25	5
5	2	14
6	1	6

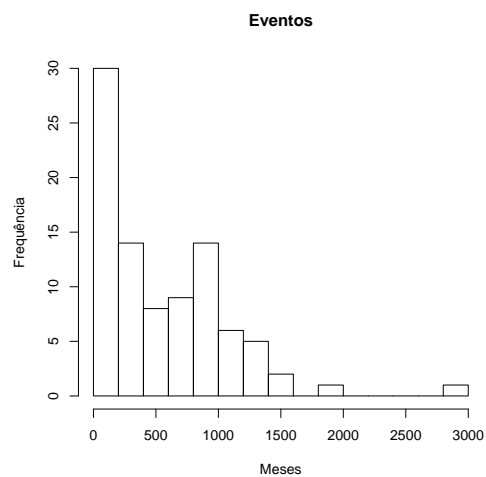
Número de eventos e censuras observadas:

```
> table(status)
```

status	
0	1
103	90

Distribuição dos tempos de sobrevivência:

```
> hist(tempo[status == 1], breaks = 12, main = "Eventos", ylab = "Frequência",  
+       xlab = "Meses")
```



Distribuição dos tempos de censura:

```
> hist(tempo[status == 0], breaks = 12, main = "Censurados", ylab = "Frequência",  
+       xlab = "Meses")
```

