MI616 - Análide de Sobrevivência
Profa.: Hildete Prisco Pinheiro
1^a Lista de Exercícios - Censuras, Função de sobrevivência e Função risco

1. Considere os dados da tabela abaixo. Calcule e faça o gráfico das estimativas da função de sobrevivência, função densidade de probabilidade e da função risco.

Anos de observação	Número de vivos no início do intervalo	Número de mortos no intervalo
0-1	1100	240
1-2	860	180
2-3	680	184
3-4	496	138
4-5	358	118
5-6	240	60
6-7	180	52
7-8	128	44
8-9	84	32
≥ 9	52	28

2. Dada a função risco $\lambda(t) = c$, derive a função de sobrevivência e a função densidade de probabilidade.

3. Dada a função de sobrevivência $S(t) = exp(-t^\gamma)$, derive a função densidade de probabilidade e a função risco.

4. Utilizando a definição abaixo:

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{P\{\text{indivíduo na idade t falhar no intervalo}(t, t + \Delta t)\}}{\Delta t}$$

ache que

$$\lambda(t) = \frac{f(t)}{S(t)}$$

utilizando definições básicas de probabilidade condicional.

5. Seja T uma variável aleatória não negativa com função densidade de probabilidade $f_T(t)$, função de sobrevivência $S_T(t)$ e função de risco $\lambda_T(t)$. Mostre que

$$S_T(t) = exp\left\{-\int_0^t \lambda_T(s)ds\right\}.$$

6. Suponha que a taxa de falha da variável aleatória tempo de sobrevivência T seja expressa pela função linear $\lambda(t) = \beta_0 + \beta_1 t$, com $\beta_0 > 0$ e $\beta_1 \geq 0$. Obtenha $S(t)$ e $f(t)$.

7. Considere a distribuição de tempos de falha com função de sobrevivência dada por

$$S(t) = exp\{1 - exp[(\lambda t)^\alpha]\}.$$

(a) Se $\alpha = 1/2$, mostre que a função de taxa de falha (função risco) é decrescente até um certo ponto e depois torna-se crescente. Encontre o ponto de mínimo (analiticamente).
(b) Para $\alpha = 2$, mostre que a função de taxa de falha é crescente.

8. O tempo, em dias, para o desenvolvimento de tumor em ratos expostos a uma substância cancerígena segue uma distribuição Weibull com função de sobrevivência dada por

$$S(t) = exp\{-\rho t^\kappa\}.$$

Assuma que foi possível estimar os parâmetros ρ e κ, sendo encontrados $\hat{\rho} = 0,0001$ e $\hat{\kappa} = 2$.
(a) Qual é a probabilidade de um rato sobreviver sem tumor aos primeiros 30 dias? E aos primeiros 45?
(b) Qual é o tempo médio até o aparecimento de tumor? E o tempo mediano?
(c) Encontre a taxa de falha de aparecimento de tumor aos 30, 45 e 60 dias. Interprete os resultados (como você explicaria os resultados para o pesquisador?).

9. Considere a distribuição Weibull com parâmetros ρ e κ (i.e., função de sobrevivência dada pelo exercício anterior). Utilizando qualquer *software* ou pacote estatístico de sua preferência, construa gráficos da função de taxa de falha (ou função risco) da distribuição Weibull, variando-se os valores dos parâmetros ρ e κ (não utilize $\kappa = 1$). Construa também gráficos das respectivas funções de sobrevivência. Considere pelo menos 6 combinações diferentes de valores de ρ e κ.

10. Mostre que, se T tem função de distribuição acumulada contínua F com função risco acumlada $\Lambda(\cdot)$, e se

$$X = min(T, c) = \begin{cases} T & \text{se } T \leq c \\ c & \text{se } T > c \end{cases}$$

então $E(\Lambda(X)) = F(c)$.

11. Suponha que um estudo é dividido em duas etapas. Na primeira, 80 doentes satisfazendo um certo protocolo são selecionados e, num dado instante, todos são acompanhados até que 35 pacientes estejam curados, quando então a primeira etapa é dada por encerrada. Numa segunda etapa, com duração pré-fixada de 30 meses, 120 pacientes são estudados. Destes, 57 iniciam o período de acompanhamento num mesmo instante, 50 apresentam a cura antes do término do estudo, 5 são retirados do estudo por motivos alheios aos objetivos dos pesquisadores e 2 ainda apresentavam o sintoma da doença ao final dos 30 meses. Os 63 pacientes restantes para a segunda etapa entraram no estudo em instantes diferentes e aleatórios, durante os 10 primeiros meses. Ao final desta etapa, 55 pacientes já estavam curados, 7 estavam com os sintomas da doença, e um paciente teve seus dados de instante de entrada e acompanhamento perdidos por problemas no sistema informatizado do hospital. Com base nessas informações, discuta o que ocorre para cada grupo de pacientes descritos, qual o tipo de censura que apresentam, se alguma. Sua discussão deve contemplar todos os 200 pacientes.

12. Sejam $T_1, \ldots T_n$ variáveis aleatórias independentes contínuas não negativas com funções risco $h_1(\cdot) \ldots h_n(\cdot)$. Prove que $T = min(T_1, \ldots, T_n)$ tem função risco $\sum h_j(\cdot)$.

13. Sejam $T_1, \ldots T_n$ variáveis aleatórias independentes com distribuição Weibull de parâmetros $\rho_1, \ldots, \rho_n$ e índice comum κ. Mostre que $T =$

$min(T_1, \ldots, T_n)$ também tem distribuição Weibull com índice *kappa*.

14. Suponha uma distribuição exponencial composta com taxa representada pela variável aleatória P, isto é, para cada indivíduo o tempo de sobrevivência é distribuído exponencialmente mas a taxa varia aleatoriamente entre os indivíduos. A variável aleatória P representa essa taxa com densidade $f_P(\cdot)$ e a distribuição condicional de T dado $P = \rho$ é

$$f_{T|P}(t|\rho) = \rho e^{-\rho t}.$$

Prove que

$$\begin{aligned} \mathrm{E}(T) &= \mathrm{E}(1/P) \\ \mathrm{Var}(T) &= 2\mathrm{E}(1/P^2) - [\mathrm{E}(1/P)]^2. \end{aligned}$$