

**Avaliação do teste logrank
em experimentos
seqüenciais agrupados**

Iracema Hiroko Iramina Arashiro

Dissertação apresentada ao
Instituto de Matemática e Estatística
da Universidade de São Paulo
para obtenção do Título de
Mestre em Estatística

Área de Concentração: **Estatística**

Orientador: **Prof. Dr. Antonio Carlos Pedroso de Lima**

Durante a elaboração deste trabalho a autora recebeu apoio financeiro do CNPq

São Paulo

2002

Avaliação do teste logrank em experimentos seqüenciais agrupados

Este exemplar corresponde à redação
final da tese devidamente corrigida e
defendida por Iracema Hiroko Iramina Arashiro
e aprovada pela comissão julgadora.

São Paulo, 2 de Dezembro de 2002.

Comissão Julgadora:

- Prof. Dr. Antonio Carlos Pedroso de Lima (orientador) - IME/USP
- Prof. Dr. Heleno Bolfarine - IME/USP
- Profa. Dra. Liciania Vaz de Arruda Silveira Chalita - IBB/UNESP

*Ao Renato e
às minhas filhas:
Hitomi, Sayuri
e Kaori*

Agradecimentos

Agradeço em primeiro lugar a Deus por me dar forças, saúde e coragem para iniciar e terminar este trabalho;

Ao Antonio Carlos, pelas orientações, discussões, colaboração no processamento das longas simulações e paciência com as minhas dificuldades;

Ao Renato, que sempre esteve presente e me apoiou nas horas difíceis, e tanto colaborou para que eu alcançasse o meu objetivo;

Às minhas filhas, por me amarem e compreenderem a minha ausência temporária;

Aos meus pais e demais familiares, pelo apoio e disponibilidade em cuidar das minhas filhas durante o período de estudos;

À minha irmã Noriko, que apesar da distância que nos separa, sempre me incentivou com os *gambate's*;

Aos professores, pelas aulas que ajudaram a enriquecer o meu conhecimento;

Aos colegas, Nívea, Regina, Joanlise, Denise, Gisela, Adrilayne, Rosangela, Andréia, Conrad, Paulo, Arturo, Delhi, Audrey, Francisco e tantos outros que estiveram presentes nesta jornada.

Resumo

Testes seqüenciais são usados em estudos de longa duração de modo a possibilitar o término antecipado de um experimento.

Na metodologia desenvolvida para lidar com esquemas seqüenciais, destacam-se os procedimentos de Pocock (1977) e de O'Brien & Fleming (1979) cujos testes são efetuados cada vez que um grupo de tamanho fixo é acrescido na amostra. Pocock sugere que em estudos de análise de sobrevivência, esses testes sejam executados em função do aumento no número de falhas observado, que deve ser fixo e pré-determinado em todas as etapas da análise. Entretanto, é comum que as unidades experimentais sejam avaliadas repetidas vezes, a tempos pré-fixados do calendário, desta forma, em cada etapa podem ser observadas quantidades diferentes de falhas.

Este trabalho apresenta uma avaliação da metodologia baseada nos limites de rejeição de Pocock e O'Brien & Fleming, quando testes seqüenciais são executados a tempos pré-fixados, independentemente do número de falhas observado em cada instante. Resultados teóricos são estudados e simulações realizadas para avaliar o poder do teste logrank seqüencial, considerando-se também, a possibilidade da ocorrência de censuras aleatórias.

As simulações sugerem que os testes seqüencias executados desta forma produzem tamanho e poder próximos a seus níveis nominais, dando respaldo ao uso desta metodologia em problemas práticos.

Abstract

Sequential tests are usually considered in long-term studies, in order to allow for early termination of the trial.

In the methodology developed for the sequential scheme, the most popular procedures are due to Pocock (1977) and O'Brien & Fleming (1979) in which tests are performed every time that a group of fixed size is added to the experiment. Pocock suggests that in the survival analysis framework the procedures should be executed in function of the increment in the number of observed failures, which are fixed and known in all of the steps of the analysis. However, usually the experimental units can be evaluated more than once, at fixed times of the calendar, so that in each stage a different number of failures can be observed.

This work evaluates the methodology based on the Pocock's and O'Brien & Fleming's boundaries, when executed at fixed times, regardless the number of failures that take a place in each evaluation time. Theoretical results are studied and simulations are considered to evaluate the power for the logrank sequential test, in the presence of random censoring.

The simulations suggest that the sequential tests performed in this way tend to produce size and power closed to the nominal levels, giving backrest to the use of this methodology in practical problems.

Índice

1	Introdução	1
2	Os Testes Seqüenciais em Grupo	3
2.1	Introdução	3
2.2	Testes Bilaterais Seqüenciais em Grupo	4
2.2.1	Definições	4
2.2.2	O Teste Bilateral de Pocock	6
2.2.3	O Teste Bilateral de O'Brien & Fleming	9
2.3	Testes Unilaterais Seqüenciais Simétricos em Grupo	11
2.3.1	Teste Unilateral Seqüencial Simétrico de Pocock	12
2.3.2	Teste Unilateral Seqüencial Simétrico de O'Brien & Fleming	14
2.4	Testes Unilaterais Seqüenciais Assimétricos em Grupo	15
2.4.1	Teste da Razão de Probabilidades Seqüencial (TRPS)	16
2.4.2	Teste Unilateral Seqüencial Assimétrico de Pocock	17
2.4.3	Teste Unilateral Seqüencial Assimétrico de O'Brien & Fleming	19
3	Teste Logrank Seqüencial em Grupo	22
3.1	Introdução	22
3.2	O Teste Logrank	23
3.3	Resultados Assintóticos	25
4	Simulações	39
4.1	Introdução	39
4.2	Planejamento	40

4.2.1	Distribuição Assintótica de $\log \hat{\omega}$	40
4.2.2	Tamanho e Poder do Teste	43
4.2.3	Número Esperado de Falhas	44
4.3	Resultados	48
4.3.1	Teste Logrank Bilateral Seqüencial em Grupo	48
4.3.2	Teste Logrank Unilateral Seqüencial Simétrico em Grupo	51
4.3.3	Teste Logrank Unilateral Seqüencial Assimétrico em Grupo	53
5	Discussão	56
A	Resultados da Simulação para os Testes Bilaterais	58
B	Resultados da Simulação para os Testes Unilaterais Simétricos	63
C	Resultados da Simulação para os Testes Unilaterais Assimétricos	68
D	Gráficos dos Resultados da Simulação para os Testes Bilaterais e Unilaterais	73
D.1	Aumento Médio do Número de Falhas por Etapa da Análise	74
D.1.1	Testes Logrank Bilaterais	74
D.1.2	Testes Logrank Unilaterais Simétricos	76
D.1.3	Testes Logrank Unilaterais Assimétricos	78
D.2	Número de Rejeições Encontrado em 500 Simulações, Sob H_0	80
D.3	Poder do Teste Logrank Seqüencial em Grupo	81
D.3.1	Testes Bilaterais	82
D.3.2	Testes Unilaterais Simétricos	83
D.3.3	Testes Unilaterais Assimétricos	84
E	Cálculo dos Valores Críticos dos Testes Seqüenciais em Grupo	85
E.1	Função Densidade Condicional	85
E.2	Método da Constante de Verossimilhança Seqüencial em Grupo	87

Capítulo 1

Introdução

Ensaaios clínicos de longa duração podem ser periodicamente avaliados a fim de se verificar, entre outros aspectos, se existem evidências estatísticas suficientes que permitam encerrar o experimento antecipadamente. Os testes são feitos à medida que a informação fornecida pelas unidades experimentais vai sendo coletada no tempo, sem esperar que todos os dados sejam obtidos, como é o caso dos testes com amostras de tamanhos fixos (amostra fixa). O conjunto de testes executados desta forma é chamado de *análise seqüencial*.

Em geral, o uso desse método permite que os estudos terminem antecipadamente, com baixo tamanho amostral quando comparados com experimentos de amostra fixa. Desta forma, gera-se economia de tempo, recursos financeiros e assegura-se a integridade dos pacientes. Além disso, do ponto de vista ético, o tratamento testado pode ser suspenso imediatamente sempre que forem detectados aspectos nocivos como efeitos colaterais indesejáveis, evitando-se que todos os participantes inscritos no protocolo recebam o tratamento. Administrativamente, os estudos podem ser avaliados do ponto de vista do planejamento checando-se, por exemplo, se as suposições iniciais estão sendo respeitadas, e se estão sendo obedecidas as normas de conduta exigidas no protocolo do estudo. Caso ocorra algum imprevisto nas avaliações iniciais, este pode ser corrigido através de um replanejamento.

Apesar das vantagens acima apresentadas, a execução dos testes seqüenciais envolve várias dificuldades do ponto de vista metodológico. O tamanho amostral, por exemplo,

é aleatório pois a cada análise tem-se diferentes números de elementos na amostra; além disso, a execução de análises repetidas dificulta o controle dos erros associados a decisões incorretas.

Os testes seqüenciais, em geral, podem ser aplicados a qualquer tipo de resposta. Entretanto, em estudos de sobrevivência, a avaliação dos testes estatísticos se torna mais complexa pois o interesse, no geral, reside no tempo de sobrevivência, que é influenciado pela duração do experimento.

Este trabalho tem por objetivo estudar o procedimento de comparação de duas distribuições de sobrevivência através do teste logrank, analisado seqüencialmente, avaliando a variação de seu poder quando são incorporadas censuras aleatórias e os dados são acrescidos de forma desigual, isto é, quando em cada etapa da análise seqüencial tende-se a diferentes incrementos no número de falhas.

No Capítulo 2, serão apresentados, de forma geral, os testes seqüenciais bilaterais de Pocock e O'Brien & Fleming, que formam as bases deste estudo, e suas versões para os testes unilaterais seqüenciais simétricos e assimétricos. No Capítulo 3 serão abordados o teste logrank e a distribuição da sua estatística, adaptada para a análise seqüencial. Finalmente no Capítulo 4 serão apresentados estudos de simulações realizados para se avaliar o poder do teste logrank calculado seqüencialmente. Nos anexos A, B e C são apresentados os resultados encontrados, organizados em tabelas.

Capítulo 2

Os Testes Seqüenciais em Grupo

2.1 Introdução

Teste seqüencial em grupo é um conjunto de testes efetuados em etapas, cujos dados são observados e acrescidos na amostra em grupos. Ele se procede da seguinte forma: coleta-se uma amostra inicial e a partir das observações feitas executa-se o primeiro teste a fim de verificar se há evidências suficientes para aceitar ou rejeitar a hipótese nula. Caso nenhuma decisão possa ser tomada, coleta-se o segundo grupo de observações e refaz-se o teste e assim sucessivamente, até que a decisão de rejeitar ou aceitar a hipótese nula possa ser tomada.

Neste capítulo serão apresentados os testes de Pocock e O'Brien & Fleming, que utilizam essa metodologia agrupada, no contexto da comparação de duas médias populacionais e considerando as abordagens bilateral e unilateral (simétrico e assimétrico).

Os testes seqüenciais em grupo, assim como os testes com amostra fixa, necessitam da especificação de valores críticos para que uma decisão possa ser tomada. Neste caso, considera-se um valor crítico para cada fase da análise, gerando uma seqüência de valores críticos, que formam os limites de parada (*boundaries stopping*). A estatística do teste pode ser escolhida de diversas formas, aqui será usada a estatística padronizada devido a sua facilidade de compreensão e comparação.

Os testes são planejados de forma a possibilitar a parada antecipada de um dado experimento, podendo ocorrer de duas formas: somente ao se rejeitar a hipótese nula (testes bilaterais) e parar sob a hipótese nula ou alternativa (testes unilaterais). Essas duas modalidades estão apresentadas nas seções que se seguem.

2.2 Testes Bilaterais Seqüenciais em Grupo

Pocock (1977) e O'Brien & Fleming (1979) apresentaram testes bilaterais seqüenciais em grupo que são de fácil implementação e podem ser adaptados para uma variedade de distribuições de probabilidade. Estes são usados, em geral, para comparar dois tipos de tratamentos, que serão denominados experimental e controle, sem perda de generalidade.

Na fase de planejamento é necessário fixar, além das taxas de erros α e β , o número máximo de análises K . Assume-se que cada grupo tem mesmo tamanho m e através de um esquema de aleatorização, os indivíduos são alocados para um ou outro tratamento. Os dados são analisados a cada acréscimo de $2m$ respostas *independentes*.

Para cada etapa $k = 1, \dots, K$, uma estatística padronizada, Z_k é calculada considerando a informação acumulada. O teste termina com a rejeição de H_0 se $|Z_k|$ exceder um dado valor crítico c_k . Se o teste continuar até a K -ésima e última etapa e $|Z_K| < c_K$, H_0 é aceita. Como mencionado anteriormente, o teste bilateral seqüencial em grupo, termina antecipadamente somente se ocorre a rejeição da hipótese nula. A seqüência de valores críticos $\{c_1, \dots, c_k\}$ é escolhida de maneira a alcançar a probabilidade de erro tipo I (α global) pré-especificado. Estes cálculos envolvem a estruturação teórica markoviana a ser apresentada na seção seguinte.

2.2.1 Definições

Considere um experimento, com amostra fixa, que envolva dois tratamentos, experimental e controle, cujas respostas de interesse podem ser modeladas por uma distribuição Normal de médias desconhecidas, μ_E e μ_C , respectivamente. Além disso, suponha que a variância σ^2 comum aos dois tratamentos é conhecida. Então, $X_{Ei} \sim N(\mu_E, \sigma^2)$ e $X_{Ci} \sim N(\mu_C, \sigma^2)$,

$i = 1, 2, \dots$, independentes. Seja θ a diferença entre as médias μ_E e μ_C . Se $\theta = 0$, temos a igualdade dos dois tratamentos; se $\theta = \pm\delta$, temos a superioridade de um tratamento em relação ao outro. As hipóteses de interesse podem então ser escritas como:

$$\begin{aligned} H_0: \theta &= \mu_E - \mu_C = 0; \\ H_A: \theta &\neq 0 \quad (\mu_E - \mu_C = \pm\delta). \end{aligned}$$

O teste bilateral para amostra fixa com n indivíduos em cada tratamento é baseado na estatística padronizada Z , dada por:

$$Z = \frac{1}{\sqrt{2n\sigma^2}} \left(\sum_{i=1}^n X_{Ei} - \sum_{i=1}^n X_{Ci} \right) \sim N(\theta\sqrt{I}, 1), \quad (2.1)$$

em que, $I = n/(2\sigma^2)$.

O teste com amostra fixa rejeita H_0 se $|Z| > c$ e c representa o valor crítico para um nível de significância α , ou seja, $c = \Phi^{-1}(1 - \alpha/2)$. O tamanho amostral n é encontrado de modo a se atingir o poder $1 - \beta$, quando $\theta = \pm\delta$.

No caso do teste seqüencial em grupo, tem-se uma estatística padronizada para cada etapa da análise seqüencial, ou seja, tem-se a seqüência $\{Z_1, \dots, Z_K\}$. Essa seqüência de estatísticas tem a distribuição conjunta canônica com os níveis de informação I_1, \dots, I_K para o parâmetro θ se:

$$\begin{aligned} (i) \quad & (Z_1, \dots, Z_K) \text{ tem distribuição Normal multivariada,} \\ (ii) \quad & \mathbf{E}(Z_k) = \theta\sqrt{I_k}, k = 1, \dots, K, \\ (iii) \quad & \text{Cov}(Z_i, Z_j) = \sqrt{I_i I_j}, \quad 1 \leq i, j \leq K. \end{aligned} \quad (2.2)$$

Com esta estrutura, os valores de I_k não são influenciados pelos valores das estatísticas $\{Z_1, \dots, Z_{k-1}\}$ das etapas anteriores, o que simplifica os cálculos das taxas de erros para testes seqüenciais em grupo. Assim, $\{Z_1, \dots, Z_K\}$ é uma seqüência de estatísticas com estrutura de Markov (Jennison & Turnbull (2000), pág. 222).

Os testes de Pocock e O'Brien & Fleming são conhecidos como *testes repetidos de significância*, pois em cada etapa da análise, procuram rejeitar a hipótese nula de acordo

com um nível de significância α'_k , $k = 1, \dots, K$, porém utilizam diferentes seqüências de valores críticos $\{c_1, \dots, c_K\}$, que são escolhidos convenientemente de modo a assegurar um α -global pré-especificado.

Armitage, McPherson & Rowe (1969) mostram que, quando o mesmo nível de significância é usado repetidamente nas etapas da análise, o erro final cometido é maior que seu valor nominal em cada etapa. Desta forma, o cálculo de probabilidade de erro tipo I global deve levar em conta os níveis de significância a serem utilizados em todas as etapas da análise.

A probabilidade de erro tipo I total, isto é, o α global de um teste seqüencial é dado por

$$Pr\{|Z_k| \geq c_k \text{ para algum } k = 1, \dots, K\},$$

quando $\{Z_1, \dots, Z_K\}$ segue a distribuição conjunta canônica (2.2) sob a hipótese nula. O poder do teste, quando $\theta = \pm\delta$, é dado por

$$Pr\left\{\bigcup_{k=1}^K (|Z_j| < c_j \text{ para } j = 1, \dots, k-1 \text{ e } |Z_k| \geq c_k)\right\} = 1 - \beta,$$

avaliados para $\{Z_1, \dots, Z_K\}$ sob a hipótese alternativa.

2.2.2 O Teste Bilateral de Pocock

A teoria moderna em análise seqüencial recebeu grande influência dos trabalhos de Wald (1947), que desenvolveu o *Teste da Razão de Probabilidades Seqüencial (TRPS)*. Este teste apresenta o menor tamanho amostral (Wald & Wolfowitz 1948), porém possui alguns problemas como a possibilidade de um ensaio não terminar se o verdadeiro parâmetro não convergir para o valor suposto pela hipótese nula ou pela alternativa, e necessitar de avaliações contínuas, tornando-se trabalhoso. Muitas propostas surgiram a partir de Wald, na tentativa de que os testes seqüenciais pudessem tornar-se mais práticos, principalmente em experimentos clínicos.

Pocock (1977) está entre os autores que obteve êxito nesta busca. Ele propõe que os acréscimos de dados sejam independentes e em grupos de tamanhos iguais, mantendo o nível de significância constante e igual a α'_k em todas as etapas da análise, com um número máximo de etapas pré-especificado igual a K . Isso facilita a aplicação dos testes seqüenciais, pois controlam-se os erros tipo I e II de modo mais eficiente. Além da praticidade, Pocock mostra que a sua metodologia pode ser aplicada a uma maior gama de respostas incluindo o caso que leva ao teste logrank.

Definição

A estatística padronizada de Pocock, na k -ésima etapa da análise, para o teste de hipóteses apresentado na seção anterior, é dado por:

$$Z_k = \frac{1}{\sqrt{(2m\sigma^2)}\sqrt{k}} \left(\sum_{i=1}^{mk} X_{Ei} - \sum_{i=1}^{mk} X_{Ci} \right), \quad k = 1, \dots, K, \quad (2.3)$$

em que m representa o número de pacientes por tratamento em cada etapa e $Z_k \sim N(\Delta\sqrt{k}, 1)$ com

$$\Delta = (\mu_E - \mu_C) \sqrt{\frac{m}{2\sigma^2}}. \quad (2.4)$$

Note que a estatística Z_k é calculada de forma semelhante a (2.1) de um teste fixo, com a diferença que a cada atualização utilizam-se todos os dados disponíveis e portanto, diferentes tamanhos amostrais ($2mk$) acumulados em cada etapa da análise seqüencial.

Seja $C_P(K, \alpha)$, a constante de Pocock para K análises com nível de significância global α . O teste de Pocock rejeita H_0 , na etapa k se $|Z_k| \geq C_P(K, \alpha)$, $k = 1, \dots, K - 1$; caso contrário, o teste continua. Se na última etapa da análise $|Z_K| \geq C_P(K, \alpha)$, H_0 é rejeitada; se não, ela é aceita.

Os valores críticos, constantes e iguais a $C_P(K, \alpha)$ estão apresentados na Tabela 2.1 para $\alpha = 0,05$ e alguns valores de K .

Tabela 2.1: Valores críticos de Pocock para $\alpha = 0,05$ e alguns valores de K .

K	$C_P(K, \alpha)$
2	2,178
3	2,289
4	2,361
5	2,413
6	2,453
7	2,485
8	2,512
9	2,535
10	2,555

Estes valores são calculados numericamente, usando-se a distribuição conjunta da seqüência de estatísticas Z_1, \dots, Z_K . Detalhes estão apresentados no apêndice E.

A constante $C_P(K, \alpha)$ para outros valores de α e K , assim como os correspondentes Δ podem ser encontrados em Pocock (1977).

O teste de Pocock rejeita H_0 sempre com o mesmo valor de α'_k , independente da etapa da análise e é dado por:

$$\alpha'_k = 2[1 - \Phi\{C_P(K, \alpha)\}].$$

Por exemplo, pela tabela pode-se ver que para $K = 5$ e $\alpha = 0,05$, o valor crítico é igual a 2,413 e neste caso o nível de significância do teste em cada etapa é $\alpha'_k = 0,0158$, com $k = 1, \dots, 5$. É importante notar que α'_k é o valor da probabilidade marginal de Z_k para rejeitar a hipótese nula e não simplesmente α/K , ou seja,

$$\alpha'_k = Pr\{|Z_k| \geq c_k \mid \theta = 0\}, \text{ para } 1 \leq k \leq K,$$

isto é,

$$\alpha'_k = Pr\{\text{rejeitar } H_0 \text{ na análise } k, \text{ dado que } H_0 \text{ é verdadeira}\},$$

enquanto que

$$\alpha = Pr\left\{\bigcup_{k=1}^K (|Z_j| < c_j \text{ para } j = 1, \dots, k-1 \text{ e } |Z_k| \geq c_k) \mid \theta = 0\right\},$$

isto é,

$$\alpha = \Pr\{\text{rejeitar } H_0 \text{ na análise } k = 1 \text{ ou } k = 2 \text{ ou } \dots \text{ ou } k = K \mid \theta = 0\}.$$

Para efeitos de planejamento, especifica-se σ^2 , δ^2 e o tamanho amostral m é dado por $\Delta^2\sigma^2/\delta^2$, em que δ representa a diferença entre as médias μ_E e μ_C , sob a hipótese alternativa, e Δ é dado por (2.4).

O gráfico da Figura 2.1 ilustra os limites de parada para o caso de um teste bilateral de Pocock em 5 etapas, nível de significância, $\alpha = 0,05$ e poder $1 - \beta = 0,90$.

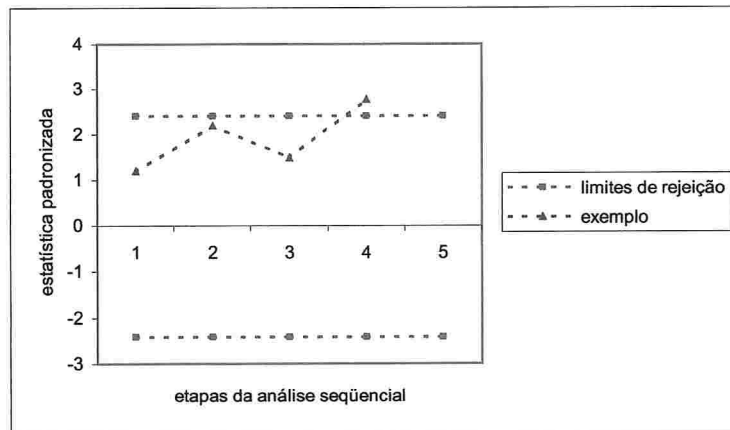


Figura 2.1: Gráfico do teste seqüencial de Pocock com 5 etapas e $\alpha = 0,05$.

Note que os limites superior e inferior são simétricos e se mantêm constantes ao longo do ensaio. O teste rejeita H_0 na etapa em que o valor da estatística cruza o limite superior ou o limite inferior e, enquanto estiver entre os dois limites o experimento continua. Se até a última etapa da análise, o teste não tiver rejeitado a hipótese nula, então ela é aceita. No exemplo do gráfico, o teste rejeita a hipótese nula na quarta etapa da análise seqüencial.

2.2.3 O Teste Bilateral de O'Brien & Fleming

O teste de O'Brien & Fleming é feito da mesma forma que o teste de Pocock. Sujeitos são alocados por tratamento em grupos independentes de tamanhos iguais a m e os dados

são analisados à medida que são disponibilizados. O teste de O'Brien & Fleming também pode ser definido em termos de uma estatística padronizada Z_k como em (2.3).

Seja $C_B(K, \alpha)$ a constante de O'Brien & Fleming para K análises planejadas com nível de significância α . O teste rejeita H_0 na etapa k se $|Z_k| \geq C_B(K, \alpha)\sqrt{(K/k)}$, para $k = 1, \dots, K - 1$, caso contrário o teste continua; se na última etapa $|Z_K| \geq C_B(K, \alpha)$ a hipótese H_0 é rejeitada, se não, ela é aceita.

Valores de $C_B(K, \alpha)$ estão apresentadas na Tabela 2.2 para $\alpha = 0,05$ e alguns valores de K .

Tabela 2.2: Valores críticos de O'Brien & Fleming para $\alpha = 0,05$ e alguns valores de K .

K	$C_B(K, \alpha)$
2	1,977
3	2,004
4	2,024
5	2,040
6	2,053
7	2,063
8	2,072
9	2,080
10	2,087

Constantes $C_B(K, \alpha)$ para outros valores de α e K , podem ser encontradas em Jennison & Turnbull (2000).

Em termos de nível de significância nominal, H_0 é rejeitada na análise k , $k = 1, \dots, K$, se o nível de significância bilateral de H_0 está abaixo de $\alpha'_k = 2[1 - \Phi\{C_B(K, \alpha)\sqrt{(K/k)}\}]$, assegurando assim, o erro tipo I global.

A seqüência de valores críticos, neste caso, depende da etapa da análise seqüencial. Eles têm valores maiores no início, dificultando a rejeição da hipótese nula, e valores menores nas análises finais, levando a um teste mais conservativo, quando comparado ao procedimento de Pocock.

O gráfico da Figura 2.2 ilustra os limites de parada para o caso do teste bilateral de O'Brien & Fleming, considerando 5 etapas de análise, nível de significância igual a 0,05 e poder 0,90.

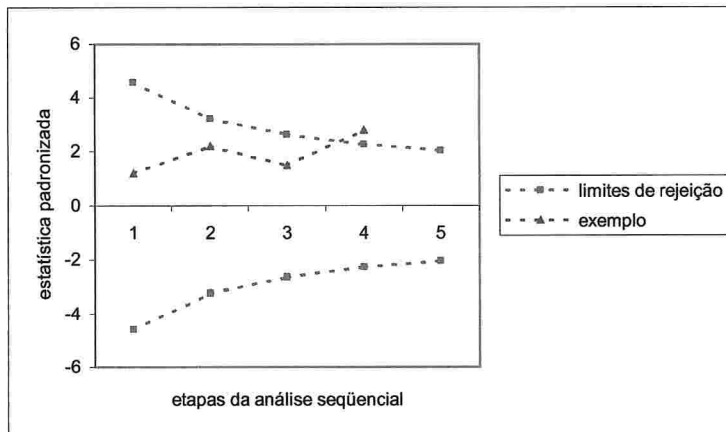


Figura 2.2: Gráfico do teste seqüencial de O'Brien & Fleming com 5 etapas e $\alpha = 0,05$.

O teste rejeita H_0 no momento em que o valor da estatística encontrada cruzar um dos limites (superior ou inferior); caso contrário o experimento continua. Note que os limites são simétricos e mais amplos nas análises iniciais estreitando à medida que o experimento se desenvolve. No exemplo do gráfico, o teste rejeita H_0 na quarta etapa da análise seqüencial.

2.3 Testes Unilaterais Seqüenciais Simétricos em Grupo

Muitos estudos clínicos objetivam verificar se a resposta média de um tratamento experimental é superior a do controle, levando ao uso de testes unilaterais.

DeMets & Ware (1980) adaptaram os testes bilaterais de Pocock e O'Brien & Fleming para o caso unilateral gerando limites de rejeição simétricos. Dois outros testes unilaterais também foram propostos: o primeiro é uma modificação do teste original de Wald (1947) combinado com as idéias de Pocock (1977) e o segundo é uma combinação do teste de Wald (limite inferior) com os de O'Brien & Fleming (limite superior), gerando assim, os testes unilaterais com limites de rejeição assimétricos, que serão tratados com mais detalhes na seção seguinte.

Os testes seqüenciais unilaterais prevêm paradas antecipadas tanto para rejeitar a hipótese nula, como para aceitá-la. Se o valor da estatística de teste Z_k for maior ou igual a c_k ($Z_k \geq c_k$), rejeita-se a hipótese nula; se $Z_k \leq -c_k$ aceita-se H_0 .

Os valores de c_k , $k = 1, \dots, K$ representam a seqüência de valores críticos que definem a regra de parada que levam à rejeição ou aceitação de H_0 .

2.3.1 Teste Unilateral Seqüencial Simétrico de Pocock

Suponha a mesma estrutura apresentada para os testes bilaterais (seção 2.2.2), só que agora, considerando as hipóteses

$$H_0: \theta = \mu_E - \mu_C < 0;$$

$$H_A: \theta \geq 0 \ (\mu_E - \mu_C \geq 0).$$

A estatística padronizada de teste na k -ésima análise Z_k , em um teste com m sujeitos em cada grupo de tratamento e em cada etapa é dada da mesma forma que em (2.3).

A utilização desse procedimento segue o esquema semelhante ao apresentado para o teste bilateral.

Seja $C_{P,1}(K, \alpha)$ a constante de Pocock unilateral, com nível de significância α e K análises, o teste rejeita H_0 , na etapa k se $Z_k \geq C_{P,1}(K, \alpha)$ e aceita H_0 se $Z_k \leq -C_{P,1}(K, \alpha)$, $k = 1, \dots, K - 1$. Na última etapa K , se $Z_K \geq C_{P,1}(K, \alpha)$, o teste rejeita H_0 , caso contrário, aceita-a.

Os valores críticos são constantes e iguais a $C_{P,1}(K, \alpha)$ e estão reproduzidas na Tabela 2.3 para $\alpha = 0,05$, e alguns valores de K . Estes valores também são calculados numericamente usando a distribuição conjunta canônica da seqüência de estatísticas Z_1, \dots, Z_k , de variáveis aleatórias independentes e identicamente distribuídas. As constantes $C_{P,1}(K, \alpha)$ para outros valores de K , α e β podem ser encontrados em DeMets & Ware (1980).

Os valores de α'_k dependem de α e K mas não do número de observações acumuladas entre os testes. Na fase de planejamento, especifica-se α , $1 - \beta$, K e valores de δ ($\mu_E - \mu_C$) sob a hipótese alternativa e calcula-se o tamanho máximo amostral $2mK$, ou equivalentemente, os valores de m necessários para se alcançar as especificações planejadas.

Tabela 2.3: Valores críticos do teste de Pocock unilateral simétrico para $\alpha = 0,05$ e alguns valores de K .

K	$C_{P,1}(K, \alpha)$
2	1,876
3	1,993
4	2,068
5	2,122
6	2,164
7	2,198
8	2,226
9	2,245
10	2,270

O gráfico da Figura 2.3 ilustra os limites de parada para o caso de um teste unilateral simétrico de Pocock com 5 etapas, nível de significância $\alpha = 0,05$ e poder $1 - \beta = 0,90$.

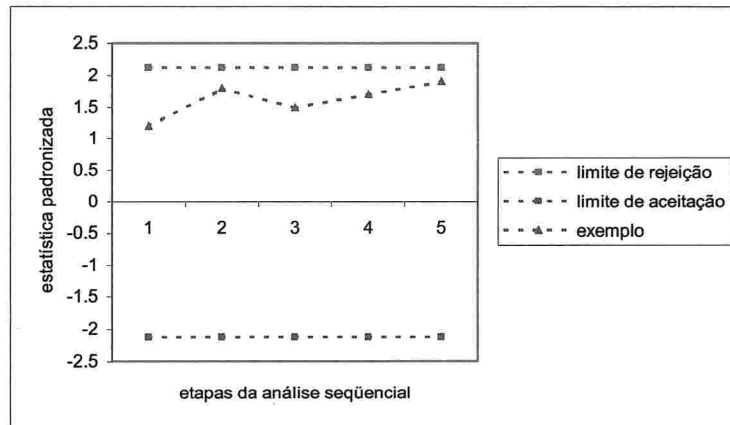


Figura 2.3: Gráfico do teste unilateral simétrico de Pocock com 5 etapas $\alpha = 0,05$

Note que os limites (superior e inferior) são simétricos como no teste bilateral original. O teste rejeita H_0 no momento em que o valor da estatística encontrada cruzar o limite superior e aceita H_0 quando cruza o limite inferior; enquanto estiver entre os dois limites o experimento continua, e se até a última etapa planejada a hipótese nula não tiver sido rejeitada, então ela é aceita. No exemplo, o teste continuou até a última etapa sem rejeitar a hipótese nula, que é então aceita.

2.3.2 Teste Unilateral Seqüencial Simétrico de O'Brien & Fleming

A mesma idéia desenvolvida para o teste unilateral de Pocock pode ser aplicada ao de O'Brien & Fleming (1979).

O teste unilateral simétrico de O'Brien & Fleming rejeita H_0 na etapa k , com $k = 1, \dots, K - 1$, se

$$Z_k \geq C_{B,1}(K, \alpha) \sqrt{(K/k)}$$

e aceita H_0 se

$$Z_k \leq -C_{B,1}(K, \alpha) \sqrt{(K/k)}.$$

Na etapa final K , se $Z_K \geq C_{B,1}(K, \alpha)$, rejeita-se H_0 , caso contrário, aceita-se H_0 .

Os valores críticos de $C_{B,1}(K, \alpha)$ estão reproduzidos na Tabela 2.4 para $\alpha = 0,05$ e alguns valores de K .

Tabela 2.4: Valores críticos de O'Brien & Fleming para teste unilateral simétrico com $\alpha = 0,05$ e alguns valores de K .

K	$C_{B,1}(K, \alpha)$
2	1,678
3	1,710
4	1,733
5	1,751
6	1,765
7	1,776
8	1,786
9	1,794
10	1,801

O gráfico a seguir mostra os limites de parada para o caso de um teste unilateral de O'Brien & Fleming com 5 etapas, nível de significância 0,05 e poder 0,90. O teste rejeita H_0 no momento em que o valor da estatística cruzar o limite superior e aceita quando cruzar o limite inferior. Enquanto estiver entre os dois limites, o experimento continua.

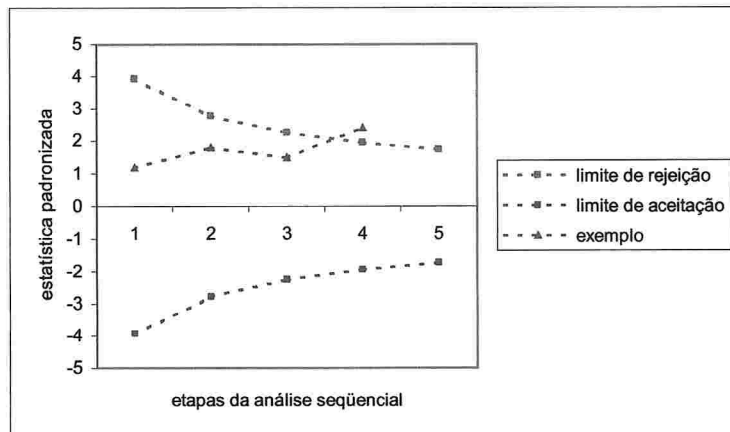


Figura 2.4: Gráfico do teste unilateral seqüencial simétrico de O'Brien & Fleming com 5 etapas e $\alpha = 0,05$

Note que os limites superior e inferior também são simétricos como no teste bilateral correspondente. No exemplo do gráfico, o teste rejeita a hipótese nula na quarta etapa da análise seqüencial.

2.4 Testes Unilaterais Seqüenciais Assimétricos em Grupo

Além das modificações descritas na seção anterior, DeMets & Ware (1980) adaptaram as idéias de Wald (1947), de modo que a aceitação de H_0 seja mais rápida que nos testes anteriores.

No caso em que um tratamento experimental é nocivo ou apresenta efeitos colaterais, o ensaio deve aceitar H_0 assim que possível evitando assim, que ele seja prolongado além do necessário. Para isso, DeMets & Ware utilizam os limites de Wald (1947) que são baseados no teste da razão de probabilidade seqüencial (TRPS), apresentado a seguir.

2.4.1 Teste da Razão de Probabilidades Seqüencial (TRPS)

O teste seqüencial de Wald (1947), foi desenvolvido inicialmente para testar hipóteses simples como

$$H_0: \theta = \mu_0;$$

$$H_A: \theta = \mu_A,$$

sendo depois adaptado para outros tipos de testes.

O procedimento é baseado no teste da razão de probabilidades conjunta de acordo com o esquema a seguir.

Suponha que um conjunto de dados pertença a uma distribuição de probabilidade cujos parâmetros estão em um espaço paramétrico $\Theta = \Theta_0 \cup \Theta_1$ e $\Theta_0 \cap \Theta_1 = \emptyset$.

A cada amostragem seleciona-se um novo elemento e com os dados disponíveis, calcula-se:

- $L(X, \theta_0)$, a função probabilidade conjunta de (X_1, \dots, X_n) sob a hipótese nula ($\theta \in \Theta_0$),
- $L(X, \theta_1)$, a função probabilidade conjunta de (X_1, \dots, X_n) sob a hipótese alternativa ($\theta \in \Theta_1$) e
- a razão, $\frac{L(X, \Theta_1)}{L(X, \Theta_0)}$.

O teste rejeita H_0 se

$$\frac{L(X, \theta_1)}{L(X, \theta_0)} \geq A$$

e aceita H_0 se

$$\frac{L(X, \theta_1)}{L(X, \theta_0)} \leq B;$$

Enquanto

$$B < \frac{L(X, \theta_1)}{L(X, \theta_0)} < A, \tag{2.5}$$

a amostragem contínua.

Os valores aproximados para A e B são dados por $(1 - \beta)/\alpha$ e $\beta/(1 - \alpha)$, respectivamente, de modo que as probabilidades de erros tipo I e II sejam no máximo α e β (para maiores detalhes, ver apêndice E).

2.4.2 Teste Unilateral Seqüencial Assimétrico de Pocock

DeMets & Ware modificaram o teste de Wald, no sentido de que o acréscimo de dados, agora, são feitos em grupo e não mais individualmente. Além disso, fixam o número máximo de análises, K e a estatística de teste contínua sendo a mesma que (2.3), como sugerido por Pocock (1977).

Sejam $C_{P,I}(\alpha; \beta)$ e $C_{P,S}(K, \alpha)$ as constantes que definem os limites inferior e superior, respectivamente, para um teste com nível de significância α , poder $1 - \beta$ e com K análises.

O valor da constante, $C_{P,I}(\alpha; \beta)$ é dado por $\log\{\beta/(1 - \alpha)\}$, como na seção anterior. Os valores de $C_{P,S}(K, \alpha)$ do limite superior e Δ são determinados de modo a satisfazer as taxas de erros α e β .

O teste unilateral assimétrico de Pocock rejeita H_0 na etapa k , $k = 1, \dots, K - 1$, se

$$Z_k \geq \frac{C_{P,S}(K, \alpha)}{\Delta\sqrt{k}} + \frac{\Delta\sqrt{k}}{2} \quad e$$

aceita H_0 se

$$Z_k \leq \frac{C_{P,I}(\alpha; \beta)}{\Delta\sqrt{k}} + \frac{\Delta\sqrt{k}}{2};$$

se na última etapa, K ,

$$Z_K \geq \frac{C_{P,S}(K, \alpha)}{\Delta\sqrt{K}} + \frac{\Delta\sqrt{K}}{2},$$

o teste rejeita H_0 , caso contrário, aceita-a.

Neste teste, os limites de rejeição são mais difíceis de serem encontrados, pois além de K e α existe a influência dos valores dos parâmetros sob a hipótese alternativa. Na fase

de planejamento, especificam-se os valores de δ , K e α para calcular o tamanho amostral que é dado por $m = \Delta^2 \sigma^2 / \delta^2$.

Na Tabela 2.5 estão apresentados os valores das constantes dos limites superior, quando o limite inferior $C_{P,I}(0, 05; 0, 1) = -2, 251$ e para alguns valores de K .

Tabela 2.5: Valores dos limite superior quando o limite inferior $C_{P,I}(0, 05; 0, 1) = -2, 251$ com $\alpha = 0, 05$ e $\beta = 0, 1$, para alguns valores de K

K	$C_{P,S}(K, \alpha)$	Δ
2	1,407	2,280
3	1,738	1,927
4	1,910	1,698
5	2,016	1,535
6	2,089	1,411
7	2,144	1,314
8	2,187	1,234
9	2,223	1,168
10	2,253	1,113

Valores de $C_{P,S}(K, \alpha)$ para outros valores de K e α podem ser encontrados em DeMets & Ware (1980).

O gráfico da Figura 2.5 ilustra os limites de parada para o teste unilateral assimétrico de Pocock para $K = 5$, $\alpha = 0, 05$ e $1 - \beta = 0, 9$.

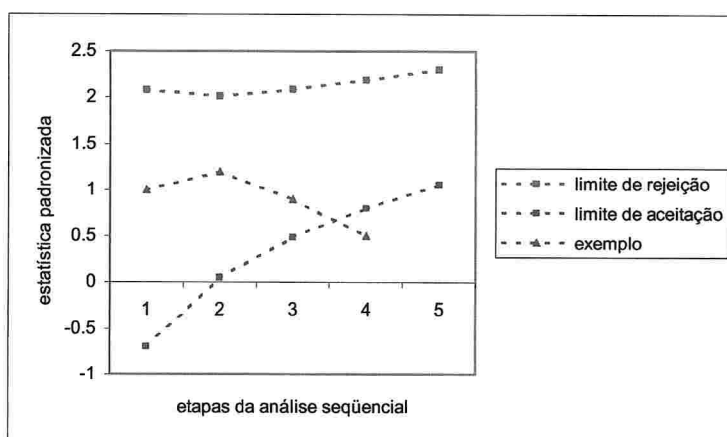


Figura 2.5: Gráfico do teste unilateral seqüencial assimétrico de Pocock com 5 etapas, $\alpha = 0,05$ e $1 - \beta = 0,9$.

Note que neste caso o limite inferior permite a aceitação da hipótese nula com valores maiores que os observados para o teste unilateral seqüencial simétrico de Pocock. No exemplo do gráfico, o teste aceita a hipótese nula na quarta etapa da análise seqüencial.

2.4.3 Teste Unilateral Seqüencial Assimétrico de O'Brien & Fleming

DeMets & Ware (1982) fizeram uma modificação nos testes bilaterais de O'Brien & Fleming (1979) para caso unilateral, de forma análoga aos testes de Pocock (1977). A idéia foi combinar as vantagens do teste de O'Brien & Fleming no limite superior e do teste de DeMets & Ware (1980) no limite inferior.

O teste unilateral assimétrico de O'Brien & Fleming rejeita H_0 na etapa k , com $k = 1, \dots, K - 1$, se

$$Z_k \geq C_{B,S}(K, \alpha) \sqrt{K/k}$$

e aceita H_0 se

$$Z_k \leq \frac{C_{B,I}(\alpha, \beta)}{\Delta \sqrt{k}} + \frac{\Delta \sqrt{k}}{2}.$$

Na última etapa, K , se

$$Z_K \geq C_{B,S}(K, \alpha),$$

o teste rejeita H_0 e caso contrário, aceita H_0 .

O valor crítico $C_{B,I}(\alpha, \beta)$ é dado por $\log(\beta/1 - \alpha)$, como no caso anterior e o valor de $C_{B,S}(K, \alpha)$ e Δ são determinados de modo que o ensaio tenha a taxa de erro tipo I igual a α e poder $1 - \beta$; isso exige cálculos de interpolação iterativa simultânea nos dois parâmetros para se obter os resultados desejados. Maiores detalhes podem ser encontrados em DeMets & Ware (1982). Na fase de planejamento, o tamanho amostral por grupo para cada etapa é determinado a partir de Δ para um dado K, α e $1 - \beta$.

Os valores de $C_{B,S}(K, \alpha)$ estão reproduzidos na Tabela 2.6 quando $C_{B,I}(0, 05; 0, 1) = -2, 251$, $\alpha = 0, 05$ e com $1 - \beta = 0, 9$.

Tabela 2.6: Valores de $C_{B,S}(K, \alpha)$, para $\alpha = 0, 05$, poder $1 - \beta = 0, 90$ e limite inferior $C_{B,I}(0, 05; 0, 1) = -2, 251$.

K	$C_{B,S}$	Δ
2	1,669	2,096
3	1,696	1,730
4	1,717	1,509
5	1,733	1,358
6	1,745	1,245
7	1,753*	1,165*
8	1,762	1,086
9	1,768*	1,031*
10	1,775	0,976

Valores indicados com (*) foram obtidos por interpolação; os demais valores de $C_{B,S}(K, \alpha)$ podem ser encontrados em DeMets & Ware (1982).

O gráfico da Figura 2.6 ilustra os limites de parada para os testes unilaterais assimétricos de O'Brien & Fleming para $K = 5$, $\alpha = 0, 05$ e $\beta = 0, 1$.

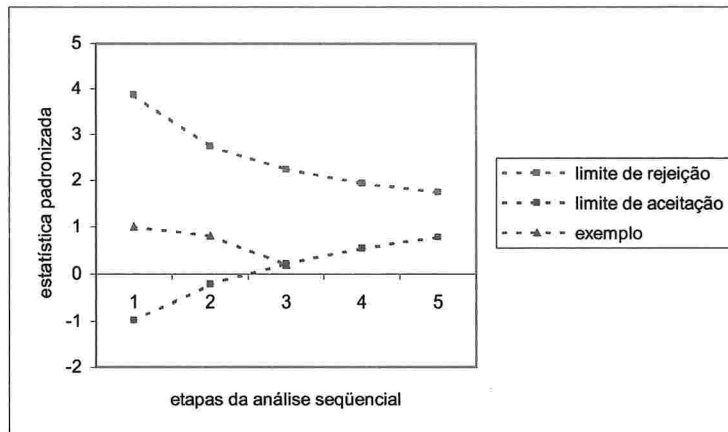


Figura 2.6: Gráfico do teste unilateral assimétrico de O'Brien & Fleming com 5 etapas e $\alpha = 0,05$.

Note que o limite superior é semelhante ao seu correspondente do teste unilateral simétrico enquanto que o limite inferior permite aceitação da hipótese nula com valores maiores, reduzindo assim, o número de ensaios sob H_0 . No exemplo do gráfico, o teste aceita a hipótese nula na terceira etapa da análise seqüencial.

Será visto no próximo capítulo que em estudos de análise de sobrevivência cuja a resposta de interesse é o tempo até a cura ou alívio de sintomas, o procedimento estatístico correspondente gera uma seqüência de estatísticas de teste que tem distribuição conjunta canônica (2.2).

Capítulo 3

Teste Logrank Seqüencial em Grupo

3.1 Introdução

A metodologia desenvolvida para testes seqüenciais tem como base situações em que a resposta de interesse é mensurada instantaneamente. Em estudos de sobrevivência, contudo, a necessidade de acompanhamento até a ocorrência de um evento e a presença de censuras dificultam a aplicação do procedimento seqüencial, uma vez que o tempo de observação correspondente a cada paciente vai se alterando a cada etapa da análise até que ocorra falha ou censura.

Quando o interesse é comparar duas funções de sobrevivência, o teste logrank é o mais eficiente quando as duas distribuições apresentam funções de risco proporcionais (Lawless (1982)). Para auxiliar o monitoramento de testes seqüenciais, Tsiatis (1981) apresenta a distribuição conjunta da estatística logrank calculada seqüencialmente. Seus resultados mostram que essa seqüência converge para um processo com incrementos independentes, e assim, a metodologia seqüencial em grupo originalmente desenvolvida para monitorar testes de médias, também pode ser estendida ao teste logrank.

3.2 O Teste Logrank

Seja T uma variável aleatória contínua, com função densidade de probabilidade $f(t), t \geq 0$.

A função de sobrevivência é definida como

$$S(t) = Pr\{T > t\} = \int_t^{\infty} f(u)du, t \geq 0.$$

A taxa de risco correspondente é dada por:

$$h(t) = \lim_{\Delta t \rightarrow 0^+} \frac{Pr(t < T \leq t + \Delta t | T > t)}{\Delta t} = \frac{f(t)}{S(t)}, t \geq 0.$$

O teste logrank tem poder ótimo para detectar diferenças quando as taxas de risco em dois tipos de tratamentos são proporcionais. Nesta situação, supõe-se que se a taxa de risco é $h(t)$ para indivíduos que recebem o tratamento experimental, então para aqueles que recebem o tratamento controle, o risco será dado por $\omega h(t)$, $\omega > 0$. Desta forma, para testar se existem diferenças entre os tratamentos, as hipóteses de interesse podem ser expressas como:

$$H_0: \omega = 1;$$

$$H_A: \omega \neq 1.$$

Suponha que deseja-se testar a igualdade das funções de sobrevivência dadas por S_E e S_C , para os tratamentos experimental e controle, respectivamente. Assume-se que a alocação dos indivíduos para um tratamento ou outro segue um esquema de aleatorização e que os dois tratamentos têm igual número de indivíduos. Em cada etapa da análise seqüencial observa-se o tempo de falha ou censura dos sujeitos envolvidos no estudo até aquele momento. Seja $d_k, k = 1, \dots, K$, o número total de falhas observadas nos dois grupos na k -ésima etapa da análise seqüencial. Denote os tempos distintos de falhas por $t_{i,k}$ tal que $t_{1,k} < t_{2,k} < \dots < t_{d_k,k}$ e seja $r_{ij,k}$, o número de indivíduos em risco do grupo j no instante imediatamente anterior a $t_{i,k}$. Finalmente, seja $\delta_{Ei,k}$ e $\delta_{Ci,k}$, o total de falhas observados nos grupos experimental e controle, respectivamente, no tempo $t_{i,k}$ e defina $\delta_{i,k} = \delta_{Ei,k} + \delta_{Ci,k}$, para $i = 1, \dots, d_k$.

Em cada tempo de falha $t_{i,k}$, os dados podem ser dispostos em forma de uma tabela de contingência 2 x 2 como segue:

Tabela 3.1: Tabela de contingência no instante $t_{i,k}$

	grupo		total
	controle	experimental	
falha	$\delta_{Ci,k}$	$\delta_{Ei,k}$	$\delta_{i,k}$
não falha	$r_{Ci,k} - \delta_{Ci,k}$	$r_{Ei,k} - \delta_{Ei,k}$	$d_k - \delta_{i,k}$
total	$r_{Ci,k}$	$r_{Ei,k}$	d_k

Note que para os casos em que não ocorrem empates entre os tempos de falha, quando $\delta_{i,k} = 1$ tem-se $\delta_{Ei,k} = 1$ e $\delta_{Ci,k} = 0$ ou $\delta_{Ei,k} = 0$ e $\delta_{Ci,k} = 1$ para cada par (i, k) .

Como discutido em Kalbfleisch & Prentice (1980), sob a hipótese nula, dado $r_{Ei,k}$, $r_{Ci,k}$ e $\delta_{i,k}$, a variável aleatória $\delta_{Ci,k}$ tem distribuição hipergeométrica com função de probabilidade

$$\frac{\binom{r_{Ci,k}}{\delta_{Ci,k}} \binom{r_{Ei,k}}{\delta_{Ei,k}}}{\binom{d_k}{\delta_{i,k}}},$$

sendo

$$\mathbf{E}_{i,k} = \mathbf{E}(\delta_{Ci,k}) = \frac{r_{Ci,k} \delta_{i,k}}{r_{Ei,k} + r_{Ci,k}},$$

e variância

$$v_{i,k} = \text{var}(\delta_{Ci,k}) = \frac{r_{Ei,k} r_{Ci,k} \delta_{i,k} (r_{Ei,k} + r_{Ci,k} - \delta_{i,k})}{(r_{Ei,k} + r_{Ci,k} - 1)(r_{Ei,k} + r_{Ci,k})^2}.$$

A estatística logrank padronizada na etapa k da análise seqüencial é dada por

$$Z_k = \frac{\sum_{i=1}^{d_k} (\delta_{C_{i,k}} - \mathbf{E}_{i,k})}{\sqrt{\sum_{i=1}^{d_k} v_{i,k}}}. \quad (3.1)$$

Na ausência de empates, o numerador em (3.1) pode ser escrito como:

$$\sum_{i=1}^{d_k} \left\{ \delta_{C_{i,k}} - \frac{r_{C_{i,k}}}{r_{E_{i,k}} + r_{C_{i,k}}} \right\},$$

e o denominador por

$$\left(\sum_{i=1}^{d_k} \frac{r_{E_{i,k}} r_{C_{i,k}}}{(r_{E_{i,k}} + r_{C_{i,k}})^2} \right)^{1/2}.$$

A estatística de teste (3.1) é usada para testar a hipótese nula $H_0 : \theta = \log \omega = 0$ no lugar de $H_0 : \omega = 1$, pois dado I_1, \dots, I_K , $\log \omega$ tem, assintoticamente, a distribuição conjunta canônica (2.2). A informação I_k para o log do risco é

$$I_k = \sum_{i=1}^{d_k} v_{i,k}.$$

A distribuição conjunta da seqüência da estatística logrank é estabelecida assintoticamente por Harrington, Fleming & Green (1982) e Tsiatis (1981), e para pequenas amostras seu comportamento foi estudado através de simulação por Gail, DeMets & Slud (1982); Jennison & Turnbull (1984); DeMets & Gail (1985) e Jennison (1992).

Na seção seguinte será apresentado o resultado encontrado por Tsiatis (1981) que mostra que os incrementos na estrutura de covariância entre as etapas da análise seqüencial do teste logrank, ao longo do tempo, são assintoticamente independentes.

3.3 Resultados Assintóticos

Na comparação de dois grupos, é sabido que a estatística *score* obtida da verossimilhança parcial do modelo de riscos proporcionais de Cox (1972) corresponde ao teste logrank bilateral, sob a hipótese de que os riscos são iguais (Kalbfleisch & Prentice (1980)).

Tsiatis (1981) mostra que essa estatística *score* calculada em diferentes pontos de tempo, converge assintoticamente para uma distribuição normal multivariada, com incrementos independentes, quando os indivíduos entram de forma aleatória ao longo do estudo. Conseqüentemente o teste logrank pode ser executado seqüencialmente, uma vez que esta é a condição necessária para que o teste seqüencial seja aplicado.

Seja T a variável aleatória que representa o tempo de falha e Y o tempo de entrada do indivíduo no estudo. Suponha que a taxa de risco esteja relacionada com a covariável X em um modelo log linear na forma

$$h(t | X = x) = h_0(t)e^{\beta x},$$

em que $h(t | X = x)$, denota a taxa de risco no tempo t , dado que a covariável $X = x$ e $h_0(t)$ é a função de risco baseline, desconhecida.

Resultados teóricos gerais são desenvolvidos assumindo que X é uma variável aleatória com média μ_X e variância σ_X^2 . No caso apresentado aqui, X é definida como 1 ou 0 para indivíduos no grupo controle ou experimental.

Inicialmente os indivíduos são recrutados no intervalo de tempo $[0, t_{rec}]$ e o instante de entrada, Y , é suposto ser uma variável aleatória positiva com densidade $g(y) > 0$ quando $0 \leq y \leq t_{rec}$, e 0 caso contrário, para algum $t_{rec} < \infty$. Supõe-se também que Y é independente do vetor aleatório (T, X) .

Supondo que a hipótese nula seja verdadeira, para os dados avaliados no tempo t^* , devem ser observadas as variáveis $T(t^*)$ e $\delta_i(t^*)$ tais que:

1. $T(t^*) = \max\{\min(T, t^* - Y), 0\}$, que pode representar o tempo de falha ($T(t^*) = T$) ou censura administrativa ($T(t^*) = t^* - Y$), e para aqueles indivíduos que entrarão depois do instante de avaliação t^* tem-se $T(t^*) = 0$.

2. $\delta_i(t^*)$, a variável aleatória indicadora de falha, tal que $\delta_i(t^*) = 1$ se $T \leq t^* - Y$ e $\delta_i(t^*) = 0$, caso contrário.

No instante de avaliação t^* , os dados podem ser representados como n vetores independentes $\{T_i(t^*), \delta_i(t^*), X_i\}$, para $i = 1, \dots, n$.

Na derivação de resultados assintóticos será necessário aumentar o número de indivíduos que entram no intervalo de tempo fixo $[0, t_{rec}]$, fazendo $n \rightarrow \infty$.

A verossimilhança parcial produzida por Cox (1975) quando avaliadas no tempo t^* é expressa por:

$$L(\beta, t^*) = \prod_{i=1}^n \left\{ \frac{e^{\beta X_i}}{\sum_{j \in R_i} e^{\beta X_j}} \right\}^{\delta_i(t^*)},$$

em que R_i denota o conjunto de índices j tais que $\{T_j(t^*) \geq T_i(t^*)\}$, isto é, o conjunto dos indivíduos em risco. O teste *score*, sob H_0 , tem a forma:

$$D(t^*) = [d \log L(\beta, t^*) / d\beta]_{\beta=0} = \sum_{i=1}^n X_i \delta_i(t^*) - \sum_{i=1}^n \delta_i(t^*) \left(\sum_{j \in R_i} X_j / r_i \right), \quad (3.2)$$

sendo r_i o total de indivíduos em R_i .

As idéias de Tsiatis (1981) são baseadas nos resultados do Lema 1 e Teorema 1 que estão apresentados e demonstrados a seguir.

Lema 1: A estatística $n^{-\frac{1}{2}} \{D(t^*) - \bar{D}(t^*)\}$ converge em probabilidade para zero, em que

$$\bar{D}(t^*) = \sum_{i=1}^n (X_i - \mu_X) [\delta_i(t^*) - H\{T_i(t^*)\}], \quad (3.3)$$

e $H(s)$ denota a função de risco acumulada definida como

$$H(s) = \begin{cases} \int_0^s h(y) dy & , s \geq 0 \\ 0 & , s < 0 \end{cases}.$$

Demonstração do Lema 1

A demonstração será composta de 3 partes: a primeira parte apresenta os cálculos da esperança de $N_i(t^*, y)$, a segunda uma equivalência de duas expressões e a terceira uma convergência em probabilidade:

Parte I - Cálculos da Esperança de $N_i(t^*, y)$

Sejam $f(t)$ a função densidade de probabilidade de t (tempo de falha do indivíduo) e $g(y)$ a função densidade de y (instante de entrada), além disso tem-se que $\delta(t^*) = 1$ se $T < t^* - y$ e $\delta_i(t^*) = 0$ caso contrário e, por definição, $T(t^*) = \min(T, t^* - y)$.

Seja:

$$N_i(t^*, t) = \mathbb{I}\{T_i(t^*) \leq t, \delta_i(t^*) = 1\},$$

o processo pontual indicando a ocorrência de falha em t ou antes para o indivíduo i , no instante t^* .

Desta forma,

$$\mathbf{E}[N_i(t^*, y)] = \mathbf{E}[\delta_i(t^*)] = Pr\{\delta_i(t^*) = 1\} = \int_0^{t^*} \int_0^{t^*-y} f(t)g(y)dt dy. \quad (3.4)$$

Sabendo que $f(t) = h(t)\overline{F}(t) = h(t)(1 - F(t))$, vem que

$$\mathbf{E}[N_i(t^*, y)] = \int_0^{t^*} \int_0^{t^*-y} h(t)g(y)dt dy - \int_0^{t^*} \int_0^{t^*-y} h(t)F(t)g(y)dt dy.$$

Integrando por partes o segundo termo, temos:

$$\mathbf{E}[N_i(t^*, y)] = \int_0^{t^*} [H(t)]_0^{t^*-y} g(y)dy - \int_0^{t^*} [H(t)F(t)]_0^{t^*-y} g(y)dy + \int_0^{t^*} \int_0^{t^*-y} H(t)f(t)g(y)dt dy.$$

$$\begin{aligned} \mathbf{E}[N_i(t^*, y)] &= \int_0^{t^*} H(t^* - y)g(y)dy - \int_0^{t^*} H(t^* - y)F(t^* - y)g(y)dy + \\ &\quad + \int_0^{t^*} \int_0^{t^*-y} H(t)f(t)g(y)dt dy \\ &= \int_0^{t^*} [1 - F(t^* - y)]H(t^* - y)g(y)dy + \int_0^{t^*} \int_0^{t^*-y} H(t)f(t)g(y)dt dy \\ &= \int_0^{t^*} \int_0^{t^*-y} H(t)f(t)g(y)dt dy + \int_0^{t^*} \int_{t^*-y}^{\infty} H(t^* - y)f(t)g(y)dt dy \\ &= \mathbf{E}[H\{T_i(t^*)\}]. \end{aligned}$$

Pode-se concluir que

$$\mathbf{E}[N_i(t^*, y)] = \mathbf{E}[H(T_i(t^*))]. \quad (3.5)$$

Parte II - Equivalência

A estatística *escore* (3.2) pode ser escrita como $D(t^*) = \bar{D}(t^*) + E(t^*)$ em que

$$E(t^*) = \int_0^{t^*} (\mu_X - \sum_{j \in R(t)} X_j/r(t)) \sum_{i=1}^n [dN_i(t^*, t) - h(t)\mathbb{I}\{T_i(t^*) \geq t\}] dt \quad e$$

$$r(t) = \sum_{j=1}^n \mathbb{I}\{T_j(t^*) \geq t\}, \quad \text{é o número de indivíduos em } R(t).$$

Logo,

$$D(t^*) = \sum_{i=1}^n X_i \delta_i(t^*) - \sum_{i=1}^n \int_0^{t^*} (\sum_{j \in R(t)} X_j/r(t)) dN_i(t^*, t).$$

Adicionando e subtraindo $dH\{T(t^*)\}$ no segundo termo da expressão anterior tem-se

$$\begin{aligned} D(t^*) &= \sum_{i=1}^n X_i \delta_i(t^*) - \sum_{i=1}^n \int_0^{t^*} (\sum_{j \in R(t)} X_j/r(t)) (dN_i(t^*, t) - dH\{T(t^*)\} + dH\{T(t^*)\}) \\ &= \sum_{i=1}^n X_i \delta_i(t^*) - \sum_{i=1}^n \int_0^{t^*} (\sum_{j \in R(t)} X_j/r(t)) (dN_i(t^*, t) - dH\{T(t^*)\}) - \\ &\quad - \sum_{i=1}^n \int_0^{t^*} (\sum_{j \in R(t)} X_j/r(t)) dH\{T(t^*)\}. \end{aligned}$$

Agora, adicionado e subtraindo μ_X ao segundo termo

$$\begin{aligned} D(t^*) &= \sum_{i=1}^n X_i \delta_i(t^*) - \underbrace{\sum_{i=1}^n \int_0^{t^*} (\sum_{j \in R(t)} X_j/r(t) - \mu_X) (dN_i(t^*, t) - dH\{T(t^*)\})}_{E(t^*)} - \\ &\quad - \sum_{i=1}^n \int_0^{t^*} (\sum_{j \in R(t)} X_j/r(t)) dH\{T(t^*)\} - \sum_{i=1}^n \int_0^{t^*} \mu_X (dN_i(t^*, t) - dH\{T(t^*)\}). \\ D(t^*) &= E(t^*) + \sum_{i=1}^n X_i \delta_i(t^*) - \sum_{i=1}^n \int_0^{t^*} (\sum_{j \in R(t)} X_j/r(t)) dH\{T(t^*)\} - \\ &\quad - \sum_{i=1}^n \int_0^{t^*} \mu_X (dN_i(t^*, t) - dH\{T(t^*)\}). \end{aligned}$$

Observe que o terceiro termo do lado direito da igualdade acima pode ser escrito como

$$\begin{aligned}
\sum_{i=1}^n \int_0^{t^*} \sum_{j \in r(t)} \frac{X_j}{r(t)} dH\{T_i(t^*)\} &= \sum_{i=1}^n \int_0^{t^*} \sum_{l=1}^n \frac{X_l}{r(t)} \mathbb{I}\{T_l \geq t\} h(t^*) \mathbb{I}\{T_i(t^*) \geq t\} dt \\
&= \int_0^{t^*} \sum_{l=1}^n \frac{X_l}{r(t)} \mathbb{I}\{T_l \geq t\} h(t^*) \underbrace{\sum_{i=1}^n \mathbb{I}\{T_i(t^*) \geq t\}}_{r(t)} dt \\
&= \int_0^{t^*} \sum_{l=1}^n X_l \mathbb{I}\{T_l \geq t\} h(t^*) dt = \sum_{i=1}^n X_i H\{T_i(t^*)\}.
\end{aligned}$$

Assim,

$$\begin{aligned}
D(t^*) &= E(t^*) + \sum_{i=1}^n X_i \delta_i(t^*) - \sum_{i=1}^n X_i H\{T_i(t^*)\} - \sum_{i=1}^n \int_0^{t^*} \mu_X (dN_i(t^*, t) - dH\{T_i(t^*)\}) \\
&= E(t^*) + \sum_{i=1}^n X_i \delta_i(t^*) - \sum_{i=1}^n X_i H\{T_i(t^*)\} - \sum_{i=1}^n \mu_X (\delta(t^*) - H\{T_i(t^*)\}) \\
&= E(t^*) + \underbrace{\sum_{i=1}^n (X_i - \mu_X) (\delta_i(t^*) - H\{T_i(t^*)\})}_{\bar{D}(t^*)};
\end{aligned}$$

ou seja, $D(t^*) = E(t^*) + \bar{D}(t^*)$.

Parte III - Convergência - Avaliando $E(t^*)$

Baseando-se na lei fraca dos grandes números,

$$\sum_{j \in R(t)} \frac{X_j}{n} = \sum_{i=1}^n \frac{X_i}{n} \mathbb{I}\{T_i(t^*) \geq t\} \xrightarrow{p} \mathbf{E}[X_i \mathbb{I}\{T_i(t^*) \geq t\}] = \mu_X Pr\{T_i(t^*) \geq t\}. \quad (a)$$

$$\frac{r(t)}{n} = \sum_{i=1}^n \frac{\mathbb{I}\{T_i(t^*) \geq t\}}{n} \xrightarrow{p} \mathbf{E}[\mathbb{I}\{T_i(t^*) \geq t\}] = Pr\{T_i(t^*) \geq t\}. \quad (b)$$

Por (a) e (b) tem-se que

$$\sum_{j \in R(t)} \frac{X_j}{r(t)} = \sum_{j \in R(t)} \frac{X_j}{n} \frac{n}{r(t)} \xrightarrow{p} \mu_X.$$

Desta forma $(\mu_X - \sum_{j \in R(t)} \frac{X_j}{r(t)})$ da expressão $E(t^*)$ converge para zero em probabilidade, enquanto que

$$J^n(y) = n^{-1/2} \sum_{i=1}^n [N_i(t^*, y) - \int_0^y h(u) \mathbb{I}\{T_i(t^*) \geq u\} du]$$

converge para um processo gaussiano de média zero, pois

$$n^{-1/2} \sum_{i=1}^n [N_i(t^*, y) - \int_0^y h(u) \mathbb{I}\{T_i(t^*) \geq u\} du]$$

converge para

$$\mathbf{E}[N_i(t^*, y) - \int_0^y h(u) \mathbb{I}\{T_i(t^*) \geq u\} du].$$

Além disso,

$$\begin{aligned} & \mathbf{E}[N_i(t^*, y) - \int_0^y h(u) \mathbb{I}\{T_i(t^*) \geq u\} du] \\ &= \mathbf{E}[N_i(t^*, y)] - \mathbf{E}[\int_0^y h(u) \mathbb{I}\{T_i(t^*) \geq u\} du] \\ &= \mathbf{E}[N_i(t^*, y)] - \mathbf{E}[H\{T_i(t^*)\}] \stackrel{(3.5)}{=} 0 \end{aligned}$$

Assim, a estatística $n^{-1/2}E(t^*)$ pode ser expressa como

$$\int_0^{t^*} \left\{ \mu_X - \sum_{j \in R(t)} \frac{X_j}{r(t)} \right\} dJ^n(t),$$

que converge para um processo Gaussiano de média zero, de forma similar ao demonstrado em Breslow & Crowley (1974). Pode-se dizer então, que a distribuição assintótica de $D(t^*)$ converge para a mesma distribuição de $\bar{D}(t^*)$.

A distribuição assintótica da estatística *score* ao longo do tempo é dada pelo teorema a seguir.

Teorema 1: O vetor aleatório $n^{-1/2}\{D(t^*), D(t^{*'})\}$ para $t^{*'} > t^*$, converge em distribuição para uma distribuição normal bivariada de média zero e matriz de covariância

$$\Omega = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix},$$

em que

$$\sigma_{11} = \sigma_X^2 Pr\{\delta(t^*) = 1\} = \sigma_X^2 \int_0^{t^*} \int_0^{t^*-y} f(t)g(y) dt dy,$$

$$\sigma_{12} = \sigma_{11}$$

$$\sigma_{22} = \sigma_X^2 Pr\{\delta(t^{*'}) = 1\} = \sigma_X^2 \int_0^{t^{*'}} \int_0^{t^{*'}-y} f(t)g(y)dt dy.$$

Demonstração do Teorema 1

A demonstração do teorema será composta, também, por duas partes, a primeira referente ao cálculo da variância e a segunda ao cálculo da covariância.

Parte I - Variância

Pelo Lema 1, a distribuição assintótica de $n^{-1/2}\{D(t^*), D(t^{*'})\}$, para $t^{*'} > t^*$ é a mesma de $n^{-1/2}\{\bar{D}(t^*), \bar{D}(t^{*'})\}$, ou seja,

$$n^{-1/2} \left[\sum_{i=1}^n (X_i - \mu_X)(\delta_i(t^*) - H\{T_i(t^*)\}), \sum_{i=1}^n (X_i - \mu_X)(\delta_i(t^{*'}) - H\{T_i(t^{*'})\}) \right]$$

Assim,

$$\begin{aligned} \sigma_{11} &= var[n^{-1/2}\bar{D}(t^*)] = n^{-1}var\left[\sum_{i=1}^n (X_i - \mu_X)(\delta_i(t^*) - H\{T_i(t^*)\})\right] \\ &\stackrel{ind.}{=} n^{-1}n\{var[(X_i - \mu_X)(\delta_i(t^*) - H\{T_i(t^*)\})]\} \\ &= var[(X_i - \mu_X)(\delta_i(t^*) - H\{T_i(t^*)\})] \\ &= \mathbf{E}[(X_i - \mu_X)(\delta_i(t^*) - H\{T_i(t^*)\})^2] - \mathbf{E}^2[(X_i - \mu_X)(\delta_i(t^*) - H\{T_i(t^*)\})]. \end{aligned}$$

Como X_i 's são independentes de $\delta_i(t^*)$ e de $H\{T_i(t^*)\}$ tem-se:

$$\begin{aligned} \sigma_{11} &= \underbrace{\mathbf{E}[(X_i - \mu_X)^2]}_{\sigma_X^2} \mathbf{E}[(\delta_i(t^*) - H\{T_i(t^*)\})^2] - \\ &\quad - \underbrace{\mathbf{E}^2[(X_i - \mu_X)]}_{0 \text{ (lema 1)}} \mathbf{E}^2[(\delta_i(t^*) - H\{T_i(t^*)\})] \\ &= \sigma_X^2 \mathbf{E}[(\delta_i(t^*) - H\{T_i(t^*)\})^2] \end{aligned}$$

Então,

$$\sigma_{11} = \sigma_X^2 \{\mathbf{E}[\delta_i(t^*)^2] - 2\mathbf{E}[\delta_i(t^*)H\{T_i(t^*)\}] + \mathbf{E}[H\{T_i(t^*)\}]^2\}. \quad (3.6)$$

Mas sabe-se que:

$$\mathbf{E}[\delta_i(t^*)^2] = \mathbf{E}[\delta_i(t^*)] = Pr\{\delta_i(t^*) = 1\}, \quad (3.7)$$

$$\begin{aligned} \mathbf{E}[\delta_i(t^*)H\{T_i(t^*)\}] &= \int_0^{t^*} \int_0^\infty \delta_i(t^*)H\{T_i(t^*)\}f(t)g(y)dtdy \\ &= \int_0^{t^*} \int_0^{t^*-y} \mathbb{I}\{t \leq t^* - y\}H(t)f(t)g(y)dtdy + \\ &\quad + \int_0^{t^*} \int_{t^*-y}^\infty \mathbb{I}\{t \leq t^* - y\}H(t^* - y)f(t)g(y)dtdy \\ &= \int_0^{t^*} \int_0^{t^*-y} H(t)f(t)g(y)dtdy. \end{aligned} \quad (3.8)$$

Além disso, tem-se que,

$$\mathbf{E}[H\{T_i(t^*)\}^2] = \int_0^{t^*} \int_0^{t^*-y} H(t)^2 f(t)g(y)dtdy + \int_0^{t^*} \int_{t^*-y}^\infty H(t^* - y)^2 f(t)g(y)dtdy.$$

Integrando o primeiro termo por partes,

$$\begin{aligned} \mathbf{E}[H\{T_i(t^*)\}^2] &= \int_0^{t^*} H(t^* - y)^2 F(t^* - y)g(y)dy - \int_0^{t^*} \int_0^{t^*-y} 2H(t)h(t)F(t)g(y)dtdy + \\ &\quad + \int_0^{t^*} H(t^* - y)^2(1 - F(t^* - y))g(y)dy. \end{aligned}$$

Cancelando os termos semelhantes,

$$\mathbf{E}[H\{T_i(t^*)\}^2] = - \int_0^{t^*} \int_0^{t^*-y} 2H(t)h(t)F(t)g(y)dtdy + \int_0^{t^*} H(t^* - y)^2 g(y)dy.$$

Lembrando que $h(t) = f(t)/S(t)$ e que $F(t) = 1 - S(t)$, tem-se

$$\begin{aligned} \mathbf{E}[H\{T_i(t^*)\}^2] &= - \int_0^{t^*} \int_0^{t^*-y} 2H(t)\frac{f(t)}{S(t)}(1 - S(t))g(y)dtdy + \int_0^{t^*} H(t^* - y)^2 g(y)dy \\ &= - \int_0^{t^*} \int_0^{t^*-y} 2H(t)h(t)g(y)dy + \int_0^{t^*} \int_0^{t^*-y} 2H(t)f(t)g(y)dtdy + \\ &\quad + \int_0^{t^*} H(t^* - y)^2 g(y)dy. \end{aligned}$$

Integrando o primeiro termo

$$\begin{aligned}
\mathbf{E}[H\{T_i(t^*)\}^2] &= - \int_0^{t^*} H(t^* - y)^2 g(y) dy + \int_0^{t^*} \int_0^{t^* - y} 2H(t)f(t)g(y) dt dy + \\
&\quad + \int_0^{t^*} H(t^* - y)^2 g(y) dy \\
&= \int_0^{t^*} \int_0^{t^* - y} 2H(t)f(t)g(y) dt dy.
\end{aligned}$$

Logo,

$$\mathbf{E}[H\{T_i(t^*)\}^2] = 2 \int_0^{t^*} \int_0^{t^* - y} H(t)f(t)g(y) dt dy. \quad (3.9)$$

Substituindo (3.7), (3.8) e (3.9) em (3.6) tem-se que,

$$\sigma_{11} = \sigma_X^2 Pr\{\delta_i(t^*) = 1\}.$$

De forma análoga

$$\sigma_{22} = \sigma_X^2 Pr\{\delta_i(t^{*'}) = 1\}.$$

A seguir será mostrado o cálculo da covariância σ_{12} e que $\sigma_{11} = \sigma_{12}$.

Parte II - Covariância

Tem-se que,

$$\begin{aligned}
\sigma_{12} &= cov[n^{-1/2}D(t^*), n^{-1/2}D(t^{*'})] \\
&= n^{-1} cov\left[\sum_{i=1}^n [(X_i - \mu_X)(\delta(t^*) - H\{T_i(t^*)\})], \sum_{i=1}^n (X_i - \mu_X)(\delta(t^{*'}) - H\{T_i(t^{*'})\})\right] \\
&\stackrel{ind.}{=} cov[(X_i - \mu_X)(\delta(t^*) - H\{T_i(t^*)\}), (X_i - \mu_X)(\delta(t^{*'}) - H\{T_i(t^{*'})\})] \\
&= \mathbf{E}[(X_i - \mu_X)(\delta_i(t^*) - H\{T_i(t^*)\})(X_i - \mu_X)(\delta_i(t^{*'}) - H\{T_i(t^{*'})\})] - \\
&\quad - \underbrace{\mathbf{E}[(X_i - \mu_X)]}_{o(\text{lema 1})} \mathbf{E}[(\delta_i(t^*) - H\{T_i(t^*)\})] \underbrace{\mathbf{E}[(X_i - \mu_X)]}_{o(\text{lema 1})} \mathbf{E}[(\delta_i(t^{*'}) - H\{T_i(t^{*'})\})] \\
&= \underbrace{\mathbf{E}[(X_i - \mu_X)^2]}_{\sigma_X^2} \mathbf{E}[(\delta_i(t^*) - H\{T_i(t^*)\})(\delta_i(t^{*'}) - H\{T_i(t^{*'})\})] \\
&= \sigma_X^2 \mathbf{E}[(\delta_i(t^*) - H\{T_i(t^*)\})(\delta_i(t^{*'}) - H\{T_i(t^{*'})\})].
\end{aligned}$$

Assim,

$$\begin{aligned}\sigma_{12} &= \sigma_X^2 \{ \mathbf{E}[\delta_i(t^*)\delta_i(t^{*'})] - \mathbf{E}[\delta_i(t^*)H\{T_i(t^{*'})\}] - \mathbf{E}[\delta_i(t^{*'})H\{T_i(t^*)\}] + \\ &\quad + \mathbf{E}[H\{T_i(t^*)\}H\{T_i(t^{*'})\}] \}. \end{aligned} \quad (3.10)$$

Sabe-se que,

•

$$\mathbf{E}[\delta_i(t^*)\delta_i(t^{*'})] = \mathbf{E}[\delta_i(t^*)] = Pr\{\delta_i(t^*) = 1\}, \quad (3.11)$$

•

$$\begin{aligned}\mathbf{E}[\delta_i(t^*)H\{T_i(t^{*'})\}] &= \int_0^{t^{*'}} \int_0^\infty \mathbb{I}\{t \leq t^* - y\} H\{T_i(t^{*'})\} f(t)g(y) dt dy \\ &= \int_0^{t^{*'}} \int_0^{t^*-y} \mathbb{I}\{t \leq t^* - y\} H\{T_i(t^{*'})\} f(t)g(y) dt dy + \\ &\quad + \int_0^{t^{*'}} \int_{t^*-y}^{t^{*'}-y} \mathbb{I}\{t \leq t^* - y\} H\{T_i(t^{*'})\} f(t)g(y) dt dy + \\ &\quad + \int_0^{t^{*'}} \int_{t^{*'}-y}^\infty \mathbb{I}\{t \leq t^* - y\} H\{T_i(t^{*'})\} f(t)g(y) dt dy \\ &= \int_0^{t^{*'}} \int_0^{t^*-y} H\{t\} f(t)g(y) dt dy, \end{aligned} \quad (3.12)$$

•

$$\begin{aligned}\mathbf{E}[\delta_i(t^{*'})H\{T_i(t^*)\}] &= \int_0^{t^{*'}} \int_0^\infty \mathbb{I}\{t \leq t^{*'} - y\} H\{T_i(t^*)\} f(t)g(y) dt dy \\ &= \int_0^{t^{*'}} \int_0^{t^*-y} \mathbb{I}\{t \leq t^{*'} - y\} H\{T_i(t^*)\} f(t)g(y) dt dy + \\ &\quad + \int_0^{t^{*'}} \int_{t^*-y}^{t^{*'}-y} \mathbb{I}\{t \leq t^{*'} - y\} H\{T_i(t^*)\} f(t)g(y) dt dy + \\ &\quad + \int_0^{t^{*'}} \int_{t^{*'}-y}^\infty \mathbb{I}\{t \leq t^{*'} - y\} H\{T_i(t^*)\} f(t)g(y) dt dy \\ &= \int_0^{t^{*'}} \int_0^{t^*-y} H\{t\} f(t)g(y) dt dy + \\ &\quad + \int_0^{t^{*'}} \int_{t^*-y}^{t^{*'}-y} H\{t^* - y\} f(t)g(y) dt dy. \end{aligned} \quad (3.13)$$

Além disso,

$$\begin{aligned}
\mathbf{E}[H\{T_i(t^*)\}H\{T_i(t^{*'})\}] &= \int_0^{t^{*'}} \int_0^{t^*-y} H\{t\}H\{t\}f(t)g(y)dtdy + \\
&+ \int_0^{t^{*'}} \int_{t^*-y}^{t^{*'}-y} H\{t^*-y\}H\{t\}f(t)g(y)dtdy + \\
&+ \int_0^{t^{*'}} \int_{t^{*'}-y}^{\infty} H\{t^*-y\}H\{t^{*'}-y\}f(t)g(y)dtdy.
\end{aligned}$$

Integrando o primeiro e o segundo termo por partes,

$$\begin{aligned}
\mathbf{E}[H\{T_i(t^*)\}H\{T_i(t^{*'})\}] &= \int_0^{t^{*'}} H(t^*-y)^2 F(t^*-y)g(y)dy - \\
&- 2 \int_0^{t^{*'}} \int_0^{t^*-y} H\{t\}h\{t\}F(t)g(y)dtdy + \\
&+ \int_0^{t^{*'}} H(t^*-y)H(t^{*'}-y)F(t^{*'}-y)g(y)dy - \\
&- \int_0^{t^{*'}} H(t^*-y)H(t^*-y)F(t^*-y)g(y)dy - \\
&- \int_0^{t^{*'}} \int_{t^*-y}^{t^{*'}-y} H\{t^*-y\}h\{t\}F(t)g(y)dtdy + \\
&+ \int_0^{t^{*'}} H(t^*-y)H(t^{*'}-y)g(y)dy - \\
&- \int_0^{t^{*'}} H(t^*-y)H(t^{*'}-y)F(t^{*'}-y)g(y)dy.
\end{aligned}$$

Cancelando os termos semelhantes, tem-se

$$\begin{aligned}
\mathbf{E}[H\{T_i(t^*)\}H\{T_i(t^{*'})\}] &= -2 \int_0^{t^{*'}} \int_0^{t^*-y} H\{t\}h\{t\}F(t)g(y)dtdy - \\
&- \int_0^{t^{*'}} \int_{t^*-y}^{t^{*'}-y} H\{t^*-y\}h\{t\}F(t)g(y)dtdy + \\
&+ \int_0^{t^{*'}} H(t^*-y)H(t^{*'}-y)g(y)dy.
\end{aligned}$$

usando novamente as equivalências, $h(t) = f(t)/S(t)$ e $F(t) = 1 - S(t)$,

$$\begin{aligned} \mathbf{E}[H\{T_i(t^*)\}H\{T_i(t^{*'})\}] &= -2 \int_0^{t^{*'}} \int_0^{t^*-y} H\{t\} \frac{f(t)}{S(t)} (1 - S(t)) g(y) dt dy - \\ &\quad - \int_0^{t^{*'}} \int_{t^*-y}^{t^{*'}-y} H\{t^* - y\} \frac{f(t)}{S(t)} (1 - S(t)) g(y) dt dy + \\ &\quad + \int_0^{t^{*'}} H(t^* - y) H(t^{*'} - y) g(y) dy. \end{aligned}$$

$$\begin{aligned} \mathbf{E}[H\{T_i(t^*)\}H\{T_i(t^{*'})\}] &= -2 \int_0^{t^{*'}} \int_0^{t^*-y} H\{t\} h(t) g(y) dt dy + \\ &\quad + 2 \int_0^{t^{*'}} \int_0^{t^*-y} H\{t\} f(t) g(y) dt dy - \\ &\quad - \int_0^{t^{*'}} \int_{t^*-y}^{t^{*'}-y} H\{t^* - y\} h(t) g(y) dt dy + \\ &\quad + \int_0^{t^{*'}} \int_{t^*-y}^{t^{*'}-y} H\{t^* - y\} f(t) g(y) dt dy + \\ &\quad + \int_0^{t^{*'}} H(t^* - y) H(t^{*'} - y) g(y) dy. \end{aligned}$$

Integrando o primeiro e o terceiro termos

$$\begin{aligned} \mathbf{E}[H\{T_i(t^*)\}H\{T_i(t^{*'})\}] &= - \int_0^{t^{*'}} H\{t^* - y\}^2 g(y) dy + \\ &\quad + 2 \int_0^{t^{*'}} \int_0^{t^*-y} H\{t\} f(t) g(y) dt dy - \\ &\quad - \int_0^{t^{*'}} [H\{t^* - y\} H(t^{*'} - y) - H\{t^* - y\} H(t^* - y)] g(y) dt dy + \\ &\quad + \int_0^{t^{*'}} \int_{t^*-y}^{t^{*'}-y} H\{t^* - y\} f(t) g(y) dy + \\ &\quad + \int_0^{t^{*'}} H(t^* - y) H(t^{*'} - y) g(y) dy. \end{aligned}$$

Cancelando os termos semelhantes,

$$\begin{aligned} \mathbf{E}[H\{T_i(t^*)\}H\{T_i(t^{*'})\}] &= 2 \int_0^{t^{*'}} \int_0^{t^*-y} H\{t\}f(t)g(y)dt dy + \\ &+ \int_0^{t^{*'}} \int_{t^*-y}^{t^{*'}-y} H\{t^*-y\}f(t)g(y)dt dy. \end{aligned} \quad (3.14)$$

Substituindo (3.11), (3.12), (3.13) e (3.14) em (3.10) tem-se:

$$\sigma_{12} = \sigma_X^2 \mathbf{E}[\delta_i(t^*)] = \sigma_X^2 Pr\{\delta_i(t^*) = 1\} = \sigma_{11}.$$

Assim fica demonstrado o Teorema 1.

A conseqüência desse teorema é que a covariância assintótica entre

$$n^{-1/2}D(t^*) \quad \text{e} \quad n^{-1/2}\{D(t^{*'}) - D(t^*)\} \quad \text{é zero, pois,}$$

$$\begin{aligned} cov[D(t^*), D(t^{*'}) - D(t^*)] &= \mathbf{E}[D(t^*)(D(t^{*'}) - D(t^*))] - \mathbf{E}[D(t^*)]\mathbf{E}[D(t^{*'}) - D(t^*)] \\ &= \mathbf{E}[D(t^*)D(t^{*'})] - \mathbf{E}[D(t^*)^2] - \mathbf{E}[D(t^*)]\mathbf{E}[D(t^{*'})] + \mathbf{E}^2[D(t^*)] \\ &= \underbrace{\mathbf{E}[D(t^*)D(t^{*'})] - \mathbf{E}[D(t^*)]\mathbf{E}[D(t^{*'})]}_{\sigma_{12}} - \underbrace{\mathbf{E}[D(t^*)^2] + \mathbf{E}^2[D(t^*)]}_{-\sigma_{11}}. \end{aligned}$$

Por (c), tem-se que $cov[D(t^*), D(t^{*'}) - D(t^*)] = 0$.

Como mencionado anteriormente, os testes seqüenciais são aplicados sob a suposição de independência entre as informações acrescidas a cada etapa da análise. Pelos resultados apresentados acima, pode-se concluir que $n^{-1/2}D(t^*)$ converge para um processo cujo os incrementos são independentes, permitindo assim, que o teste logrank seja analisado seqüencialmente.

Capítulo 4

Simulações

4.1 Introdução

O monitoramento dos experimentos seqüenciais é feito, em geral, a intervalos de tempos pré-fixados. Essa regularidade na inspeção leva à observação de aumentos desiguais no número de falhas em cada etapa da análise. Os testes seqüenciais bilaterais de Pocock (1977) e O'Brien & Fleming (1979) foram desenvolvidos para análises com aumentos iguais de informação, isto é, supondo igual acréscimo no número de falhas entre as etapas de análise. Sendo assim, essas metodologias podem se tornar inadequadas para executar testes em tempos pré-fixados. DeMets & Gail (1985) mostraram, através de simulações, que o teste logrank seqüencial bilateral ainda pode ser aplicado, mesmo com aumentos irregulares. Usando essas idéias, neste capítulo será avaliado o comportamento do poder do teste logrank seqüencial para os casos unilateral e bilateral, na presença de censuras aleatórias.

Para melhor compreensão do procedimento de simulação desenvolvido neste trabalho, será apresentado a seguir, aspectos do planejamento de experimentos com amostra fixa em análise de sobrevivência que serão utilizados para simular os testes seqüenciais.

4.2 Planejamento

Existem várias propostas de planejamento em análise de sobrevivência para executar o teste logrank com amostra fixa, entre os quais pode ser destacado o trabalho de Rubinstein, Gail & Santner (1981) em que são consideradas características semelhantes aos experimentos reais.

Rubinstein et al. (1981) apresentam um planejamento para um experimento que objetiva executar o teste logrank unilateral em amostra fixa. Sua proposta é inserir um período adicional de acompanhamento (τ), além do tempo de recrutamento (t_{rec}), a fim de diminuir o número de indivíduos na amostra. Para calcular o número esperado de falhas, eles utilizam a distribuição assintótica de $\log \hat{\omega}$, em que $\omega = \lambda_C/\lambda_E$, λ_C e λ_E são parâmetros de distribuições exponenciais para os tempos de falhas dos indivíduos que pertencem ao grupo controle e experimental, respectivamente.

As hipóteses a serem testadas são do tipo

$$H_0 : \omega = 1;$$

$$H_A : \omega > 1,$$

equivalentemente,

$$H_0 : \log \omega = 0;$$

$$H_A : \log \omega > 0.$$

É importante observar que é comum o uso da estatística $\log \hat{\omega}$ no lugar de $\hat{\omega}$, pois sua distribuição assintótica é menos viesada e apresenta melhor taxa de convergência (Rubinstein et al. (1981)). A distribuição assintótica de $\log \hat{\omega}$ está apresentada na seção a seguir.

4.2.1 Distribuição Assintótica de $\log \hat{\omega}$

As seguintes suposições serão consideradas:

1. O tempo de recrutamento dos indivíduos ocorrem no intervalo $[0, t_{rec}]$, sendo que a entrada no estudo se dá segundo um processo de Poisson de média R . Desta forma, o número médio do total de indivíduos no estudo será $N = Rt_{rec}$;

O tempo observado t_{ij} dos N_j indivíduos do grupo j é dado por, $t_{ij} = \min\{T_{ij}, C_{ij}\}$ sendo que T_{ij} representa o tempo de falha do indivíduo i do grupo j e C_{ij} o tempo de censura do mesmo indivíduo. Desta forma, se $T_{ij} \leq C_{ij}$, uma falha é observada; caso contrário, uma censura é observada. Assim,

$$T_j = \sum_{i=1}^{N_j} t_{ij},$$

representa o tempo total de teste dos indivíduos do grupo j ;

O número de falhas no grupo j condicional a N , o total de indivíduos na amostra, será denotado por D_j^N . Esta quantidade é dada por

$$D_j^N = \sum_{i=1}^{N_j} \delta_{ij};$$

2. Existe um período adicional de acompanhamento denotado por τ ;

3. Os tempos de falha, censura e entrada dos indivíduos são independentes entre os grupos e entre si.

Supondo, para o grupo j , tempos de falha exponenciais de média $1/\lambda_j$, tem-se que a função de verossimilhança das N_j observações é dada por

$$L(t_1, \dots, t_{N_j}; \lambda_j | N) = \prod_{i=1}^{N_j} \lambda_j^{\delta_{ij}} \exp\{-\lambda_j t_{ij}\} = \lambda_j^{\sum_{i=1}^{N_j} \delta_{ij}} \exp\{-\sum_{i=1}^{N_j} \lambda_j t_{ij}\} = \lambda_j^{D_j^N} \exp\{-\lambda_j T_j\},$$

com $j = C, E$ e o log da verossimilhança por

$$\log(L(t_1, \dots, t_{N_j}; \lambda_j | N)) = l(t_1, \dots, t_{N_j}; \lambda_j | N) = D_j^N \log \lambda_j - \lambda_j T_j.$$

A estimativa de máxima verossimilhança de λ_j , denotada por $\hat{\lambda}_j$, é obtida derivando-se $l(t_1, \dots, t_{N_j}; \lambda_j | N)$ em relação ao parâmetro λ_j e igualando-o a zero, assim,

$$\frac{\partial \log(L(t_1, \dots, t_{N_j}; \lambda_j | N))}{\partial \lambda_j} = \frac{D_j^N}{\lambda_j} - T_j = 0.$$

Logo,

$$\widehat{\lambda}_j = \frac{D_j^N}{T_j}, \quad \text{com } j = C, E.$$

A Informação de Fisher condicional ao total de indivíduos na amostra, $\mathbf{I}(\lambda_j | N)$ é dada por

$$\mathbf{I}(\lambda_j | N) = \mathbf{E}\left[-\frac{\partial^2 \log(L(t_1, \dots, t_{N_j}; \lambda_j | N))}{\partial \lambda_j^2}\right].$$

Assim,

$$\mathbf{I}(\lambda_j | N) = \mathbf{E}\left[\frac{D_j^N}{\lambda_j^2}\right] = \frac{1}{\lambda_j^2} \mathbf{E}[D_j^N] = \frac{1}{\lambda_j^2} \mathbf{E}[D_j | N],$$

e, usando a distribuição assintótica do estimador de máxima verossimilhança,

$$\widehat{\lambda}_j \stackrel{a}{\sim} N\left(\lambda_j; \frac{\lambda_j^2}{\mathbf{E}[D_j | N]}\right), \quad \text{com } j = C, E.$$

A distribuição assintótica de uma função $g(\widehat{\lambda}_j)$ pode ser obtida através do método delta, levando a:

$$g(\widehat{\lambda}_j) \stackrel{a}{\sim} N(g(\lambda_j), [g'(\lambda_j)]^2 \mathbf{I}^{-1}(\lambda_j | N)).$$

Seja $g(\lambda_j) = \log(\lambda_j)$, então,

$$g'(\lambda_j) = 1/\lambda_j \quad \text{e} \quad [g'(\lambda_j)]^2 \mathbf{I}^{-1}(\lambda_j | N) = (1/\lambda_j)^2 \frac{\lambda_j^2}{\mathbf{E}[D_j | N]}.$$

Logo,

$$\log \widehat{\lambda}_j \stackrel{a}{\sim} N(\log \lambda_j, 1/\mathbf{E}[D_j | N]), \quad \text{com } j = C, E.$$

Desta forma, dada a independência entre $\widehat{\lambda}_C$ e $\widehat{\lambda}_E$,

$$\log \widehat{\lambda}_C - \log \widehat{\lambda}_E \stackrel{a}{\sim} N(\log \lambda_C - \log \lambda_E, \mathbf{E}^{-1}[D_C | N] + \mathbf{E}^{-1}[D_E | N]).$$

Isto equivale a

$$\log \frac{\widehat{\lambda}_C}{\widehat{\lambda}_E} \stackrel{a}{\sim} N\left(\log \frac{\lambda_C}{\lambda_E}, \mathbf{E}^{-1}[D_C | N] + \mathbf{E}^{-1}[D_E | N]\right).$$

4.2.2 Tamanho e Poder do Teste

Pela seção anterior, pode-se afirmar que:

$$\frac{\log \hat{\omega} - \log \omega}{\sqrt{\sigma^2}} \stackrel{a}{\sim} N(0, 1),$$

em que

$$\sigma^2 = \mathbf{E}^{-1}[D_C | N] + \mathbf{E}^{-1}[D_E | N].$$

Sob H_0 , $\log \omega = 0$, e assim a estatística de teste a ser utilizada é:

$$\frac{\log \hat{\omega}}{\sqrt{\sigma^2}}.$$

Logo,

$$\alpha = Pr\{\text{rejeitar } H_0 \mid H_0 \text{ é verdadeira}\} = Pr\{\log \hat{\omega}/\sqrt{\sigma^2} \geq Z_\alpha \mid H_0 \text{ é verdadeira}\},$$

em que $Z_\alpha = \Phi^{-1}(1 - \alpha)$ com $\Phi(\cdot)$ a função densidade acumulada da Normal $(0, 1)$. O teste rejeita H_0 se

$$\frac{\log \hat{\omega}}{\sqrt{\sigma^2}} \geq Z_\alpha.$$

Por outro lado,

$$\begin{aligned} \beta &= Pr\{\text{rejeitar } H_A \mid H_A \text{ é verdadeira}\} = \\ &= Pr\{\log \hat{\omega}/\sqrt{\sigma^2} < Z_\alpha \mid H_A \text{ é verdadeira}\} = \\ &= Pr\{\log \hat{\omega}/\sqrt{\sigma^2} - \mathbf{E}[\log \hat{\omega}/\sqrt{\sigma^2}] < Z_\alpha - \mathbf{E}[\log \hat{\omega}/\sqrt{\sigma^2}] \mid H_A \text{ é verdadeira}\} \end{aligned}$$

Como

$$\log \hat{\omega}/\sqrt{\sigma^2} - \mathbf{E}[\log \hat{\omega}/\sqrt{\sigma^2}] \sim N(0, 1),$$

temos que

$$Z_\alpha - \mathbf{E}\left[\frac{\log \hat{\omega}}{\sqrt{\sigma^2}}\right] = -Z_\beta \Leftrightarrow Z_\alpha + Z_\beta = \mathbf{E}\left[\frac{\log \hat{\omega}}{\sqrt{\sigma^2}}\right].$$

Elevando ao quadrado, segue que

$$(Z_\alpha + Z_\beta)^2 = \left(\mathbf{E}\left[\frac{\log \hat{\omega}}{\sqrt{\sigma^2}}\right]\right)^2 \Leftrightarrow (Z_\alpha + Z_\beta)^2 = \left(\frac{\log \omega}{\sqrt{\sigma^2}}\right)^2$$

e assim,

$$\frac{(\log \omega)^2}{(Z_\alpha + Z_\beta)^2} = \frac{1}{\mathbf{E}[D_C | N]} + \frac{1}{\mathbf{E}[D_E | N]}. \quad (4.1)$$

Através de (4.1) pode-se estabelecer o poder $1 - \beta$ e nível de significância α em um teste com amostra fixa.

4.2.3 Número Esperado de Falhas

Como os indivíduos entram na amostra segundo um processo com Poisson de taxa R , o total de indivíduos na amostra, $N = N_C + N_E$, será uma variável aleatória com média Rt_{rec} . Conhecido o valor de N , os pacientes entram no estudo uniformemente no intervalo de tempo $[0, t_{rec}]$.

Além disso, os indivíduos são alocados aleatoriamente para um ou outro tratamento, isto significa que, dado N , N_j tem distribuição binomial de parâmetros N e $1/2$.

Além dos tempos de falha, Rubinstein et al. (1981) consideram a possibilidade de ocorrência de censuras aleatórias. Neste caso, é suposto que a sua distribuição segue uma exponencial de parâmetros ϕ_j , com $j = C, E$. Desta forma, dado um intervalo de tempo z a partir da entrada até o final do estudo, a probabilidade de que uma observação do grupo j resultará em uma falha é:

$$\begin{aligned}
Pr\{T < C; T < z\} &= \mathbf{E}[Pr\{T < C; T < z\} | T] \\
&= \int_0^z [Pr\{T < C; T < z\} | T = t] f_T(t) dt \\
&= \int_0^z [Pr\{T < C | T = t\} f_T(t) dt \\
&= \int_0^z \lambda_j \exp\{-\lambda_j t\} \exp\{-\phi_j t\} dt \\
&= \int_0^z \lambda_j \exp\{-(\lambda_j + \phi_j)t\} dt \\
&= -\lambda_j \frac{\exp\{-(\lambda_j + \phi_j)t\}}{\lambda_j + \phi_j} \Big|_0^z \\
&= \frac{\lambda_j}{\lambda_j + \phi_j} (1 - \exp\{-(\lambda_j + \phi_j)z\}). \tag{4.2}
\end{aligned}$$

Dado o número total de indivíduos na amostra (N), a probabilidade de $\delta_{ji} = 1$ pode ser dada por

$$\begin{aligned}
\mathbf{E}[\delta_{ji} | N] &= Pr\{\delta_{ji} = 1 | N\} \\
&= Pr\{T < C; T < Z | N\} = \mathbf{E}[Pr\{T < C; T < Z | Z, N\}] \\
&= \int_{\tau}^{\tau+t_{rec}} Pr\{T < C; T < Z | Z = z, N\} f_{Z|N}(z) dz \\
&= \int_{\tau}^{\tau+t_{rec}} Pr\{T < C; T < Z | N\} \frac{1}{t_{rec}} dz.
\end{aligned}$$

E o número total de falhas do grupo j (D_j) por:

$$\begin{aligned}
\mathbf{E}[D_j | N] &= \mathbf{E}\left[\sum_{i=1}^{N_j} \delta_{ji} | N\right] \\
&= \mathbf{E}\left\{\left[\sum_{i=1}^{N_j} \delta_{ji} | N\right]\right\}. \tag{4.3}
\end{aligned}$$

Lembrando que $\mathbf{E}[N_j | N] = N/2$

$$\begin{aligned}
\mathbf{E}[D_j | N] &= \mathbf{E}\left\{\left[\sum_{i=1}^{N/2} \delta_{ji}\right]\right\} \\
&= \mathbf{E}\left\{\frac{N}{2} \mathbf{E}[\delta_{ji}]\right\} \\
&= \mathbf{E}\left\{\frac{N}{2} \int_{\tau}^{\tau+t_{rec}} Pr\{T < C; T < Z\} \frac{1}{t_{rec}} dz\right\} \\
&= \mathbf{E}\left\{\frac{N}{2}\right\} \int_{\tau}^{\tau+t_{rec}} Pr\{T < C; T < Z\} \frac{1}{t_{rec}} dz \\
&= \frac{Rt_{rec}}{2} \int_{\tau}^{\tau+t_{rec}} Pr\{T < C; T < Z\} \frac{1}{t_{rec}} dz. \tag{4.4}
\end{aligned}$$

Substituindo (4.2) em (4.4) tem-se

$$\begin{aligned}
\mathbf{E}[D_j | N] &= \frac{Rt_{rec}}{2} \int_{\tau}^{\tau+t_{rec}} \frac{\lambda_j}{\lambda_j + \phi_j} (1 - \exp\{-(\lambda_j + \phi_j)z\}) \frac{1}{t_{rec}} dz \\
&= \frac{Rt_{rec}}{2} \frac{\lambda_j}{\lambda_j + \phi_j} \int_{\tau}^{\tau+t_{rec}} (1 - \exp\{-(\lambda_j + \phi_j)z\}) \frac{1}{t_{rec}} dz \\
&= \frac{Rt_{rec}}{2} \frac{\lambda_j}{\lambda_j + \phi_j} \left\{ \frac{1}{t_{rec}} \left[\int_{\tau}^{\tau+t_{rec}} 1 dz - \int_{\tau}^{\tau+t_{rec}} \exp\{-(\lambda_j + \phi_j)z\} dz \right] \right\} \\
&= \frac{Rt_{rec}}{2} \frac{\lambda_j}{\lambda_j + \phi_j} \left\{ \frac{1}{t_{rec}} \left[z \Big|_{\tau}^{\tau+t_{rec}} + \frac{\exp\{-(\lambda_j + \phi_j)z\}}{\lambda_j + \phi_j} \Big|_{\tau}^{\tau+t_{rec}} \right] \right\} \\
&= \frac{Rt_{rec}}{2} \frac{\lambda_j}{\lambda_j + \phi_j} \left\{ \frac{1}{t_{rec}} t_{rec} + \frac{1}{t_{rec}} \frac{\exp\{-(\lambda_j + \phi_j)(\tau + t_{rec})\} - \exp\{-(\lambda_j + \phi_j)\tau\}}{\lambda_j + \phi_j} \right\} \\
&= \frac{Rt_{rec}}{2} \frac{\lambda_j}{\lambda_j + \phi_j} \left\{ 1 - \frac{\exp\{-(\lambda_j + \phi_j)\tau\}}{t_{rec}(\lambda_j + \phi_j)} \{1 - \exp\{-(\lambda_j + \phi_j)t_{rec}\}\} \right\}.
\end{aligned}$$

Denominando-se $\lambda_j + \phi_j$ por λ_j^* ,

$$\mathbf{E}[D_j | N] = \frac{Rt_{rec}}{2} \frac{\lambda_j}{\lambda_j^*} \left\{ 1 - \frac{\exp\{-\lambda_j^* \tau\}}{t_{rec} \lambda_j^*} \{1 - \exp\{-\lambda_j^* t_{rec}\}\} \right\}. \tag{4.5}$$

Fixado o nível de significância, o poder do teste e os valores de λ_i , λ_i^* e τ , pode-se calcular o tempo de recrutamento de pacientes t_{rec} substituindo-se (4.5) em (4.1), tal que

$$\sum_{i=C,E} \left[\frac{2(\lambda_i^*)^2}{n \lambda_i} \right] [\lambda_i^* t_{rec} - e^{-\lambda_i^* \tau} (1 - e^{-\lambda_i^* t_{rec}})]^{-1} = \frac{(\log \omega)^2}{(Z_{\alpha} + Z_{\beta})^2}. \tag{4.6}$$

Para a resolução desta equação foi utilizado o método iterativo da bissecção, encontrado em Atkinson (1993), fornecendo assim os valores de t_{rec} .

A simulação consiste em usar esse planejamento para amostra fixa, dividindo o tempo total do experimento em um número de intervalos de tempos igual ao número de análises interinas planejadas e assim executar o teste no final de cada um dos intervalos de tempo.

O uso desse procedimento (entradas aleatorizadas ao longo do experimento e análise em várias etapas), leva cada uma das observações a poder apresentar respostas diferentes em cada etapa da avaliação, dependendo de a falha ter ocorrido ou não. Este fato é exemplificado na Figura 4.1.

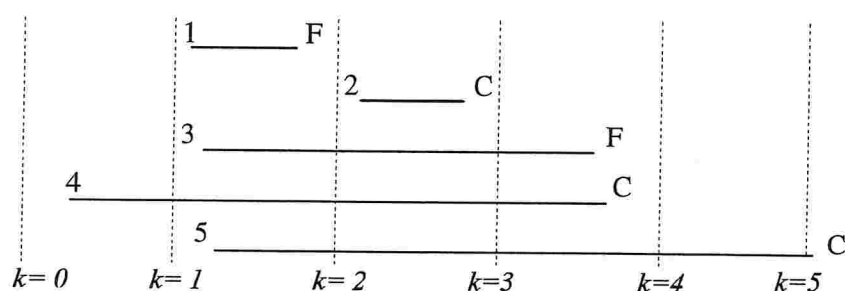


Figura 4.1: Exemplo de possíveis respostas dos indivíduos em cada etapa da análise.

O esquema apresentado mostra como as respostas dos indivíduos variam em cada etapa da análise. Os indivíduos 1 e 2 são exemplos de casos simples, pois apresentam a falha ou censura (por perda de acompanhamento) logo na sua primeira análise. Assim, o tempo observado se mantém constante nas análises subsequentes até o final do experimento.

Os indivíduos 3 e 4 estão censurados (administrativamente) em algumas etapas da análise, mas apresentam falha ou censura antes do final do experimento. Desta forma, os valores de tempo observados variam até a etapa em que apresentam a falha ou censura por perda de acompanhamento.

O indivíduo 5 não apresentou falha até o final do experimento. Neste caso, os tempos observados variam em todas as etapas da análise.

Para obter os resultados apresentados nas seções a seguir, foram consideradas todas as situações descritas anteriormente.

4.3 Resultados

4.3.1 Teste Logrank Bilateral Seqüencial em Grupo

Como mencionado, Pocock (1977) sugere que seus testes seqüenciais bilaterais podem ser aplicados ao teste logrank quando o aumento do número de falhas em cada etapa da análise é igual a um valor constante d . Neste caso, o poder do teste pode ser calculado usando a distribuição normal de média $\Delta = \theta(d/4)^{1/2}$, em que θ representa $\log \omega$ (Gail, DeMets & Slud (1982)). Nos casos em que ocorrem aumentos diferentes de falhas em cada análise, esse poder não pode ser calculado diretamente. Dada a falta de resultados analíticos, serão apresentados resultados de simulações baseados em 500 amostras para cada situação.

DeMets & Gail (1985) apresentam este estudo para o caso do teste logrank bilateral e sem censura, no contexto descrito na seção anterior. Eles foram aqui reproduzidos com a finalidade de compará-los com aqueles que apresentam censuras aleatórias ao longo do experimento.

Os seguintes parâmetros foram fixados para a simulação:

1. $K = 5$ estágios;
2. nível de significância $\alpha = 0,05$;
3. as taxas de entrada, R , utilizadas foram 10, 100, 250 e 1000, por unidade de tempo;
4. falhas seguindo o modelo exponencial, de parâmetros $\lambda_C = 1$ e $\lambda_E = 0,5$, para altas taxas de risco; $\lambda_C = 0,04$ e $\lambda_E = 0,02$, para baixas taxas de risco, o que implicará em risco relativo $\omega = 2$;
5. censuras seguindo o modelo exponencial, de parâmetros $\phi_C = \lambda_C \epsilon$ e $\phi_E = \lambda_E \epsilon$, em que ϵ é a constante que irá gerar as censuras aleatórias tal que

$$\frac{\lambda_j}{\lambda_j + \lambda_j \epsilon} = \text{taxa de falha do grupo } j .$$

Os valores críticos para rejeitar a hipótese nula, obtidas das Tabelas 2.1 e 2.2 do Capítulo 2, são $C_P(5; 0, 05) = 2, 413$ para o caso do teste de Pocock (1977) e $C_B(5; 0, 05) = 2, 04\sqrt{5/k}$, $k = 1, \dots, 5$ para o teste de O'Brien & Fleming (1979).

Espera-se que a ocorrência de censuras aleatórias diminua o poder do teste logrank seqüencial, devido ao menor número de falhas e conseqüentemente perda de informação. Para verificar isto foram geradas 500 simulações para cada situação. As taxas de censuras aleatórias utilizadas foram 20%, 50% e 80%.

A Tabela A.1, mostra que em todas as situações o total das médias de falhas encontrado está próximo de seu valor esperado calculado através de (4.5). Além disso, o número médio de indivíduos recrutados está de acordo com a taxa média da distribuição de Poisson especificada.

O aumento médio de falhas entre as etapas da análise são desiguais, com exceção dos casos com taxa de entrada $R = 10$ e taxas de censuras de 50% e 80% (gráficos das Figuras D.3 e D.4) que apresentam aumentos regulares entre as etapas 2 e 5, nos casos sem o período adicional de acompanhamento e entre as etapas 2 e 4 nos casos em que houve o período adicional.

Motivada pelo trabalho de DeMets & Gail (1985) e com a finalidade de se controlar a simulação será definido um *intervalo de tolerância* para o nível de significância α dado por

$$r \left(\alpha \pm 1, 96 \sqrt{\frac{\alpha(1 - \alpha)}{r}} \right),$$

em que r representa o número de réplicas efetuadas nas simulações, e α é o nível de significância fixado. No caso aqui estudado, esse intervalo é dado, aproximadamente, por [15,35].

A Tabela A.2, mostra o número encontrado de rejeições em 500 simulações. O nível de significância nominal ficou controlado na maioria dos casos, ou seja, esteve dentro do intervalo de tolerância definido anteriormente, como mostra o gráfico da Figuras D.13. Nove exceções ocorreram dentre as 96 simulações executadas, sendo 6 delas com 36 re-

jeições e o pior caso com 41 rejeições, que ocorreu no caso do teste de Pocock, sem censura aleatória, com a taxa de entrada $R=10$ e período adicional de acompanhamento $\tau = 1$.

Sob a hipótese alternativa, no caso sem censuras, pode-se notar que os números de rejeições ficam próximos aos valores planejados para os teste de O'Brien & Fleming e teste com amostra fixa (gráficos das Figuras D.16 e D.17), enquanto que o teste de Pocock se mostra um pouco inferior a ambos.

Nos casos com censura, pode-se observar que o poder do teste vai diminuindo conforme aumenta a taxa de censura aleatória. O poder diminui sensivelmente quando a taxa de entrada ($R=10$ e $R=100$) é baixa e assim o número de entradas (recrutamentos) também é baixo, gerando poucas censuras administrativas (tipo I). No caso da taxa de entrada alta ($R=250$ e $R=1000$) o poder também diminui, porém com menor intensidade devido a alta taxa de entrada e censuras do tipo I, além da taxa de risco ser baixa.

Em geral, pode-se dizer que o poder do teste seqüencial de O'Brien & Fleming está próximo ao poder do teste com amostra fixa; já o teste de Pocock sempre apresenta o poder um pouco inferior a ambos.

Por outro lado, a Tabela A.3, mostra que o teste de Pocock apresenta sempre o menor número médio de falhas e o menor número médio de ensaios, seja sob H_0 ou sob H_1 , ocorrendo apenas uma igualdade quando $R=250$, $\tau = 0$ e censura aleatória de 50%.

Sob a hipótese nula, pode-se dizer que a análise seqüencial não mostrou muitas vantagens em relação à amostra fixa, pois em todas as situações, o número médio de falhas apresentou valores próximos nos três casos (Pocock, O'Brien & Fleming e amostra fixa) enquanto que sob a hipótese alternativa, ambos os testes seqüenciais apresentaram sempre menor número médio de falhas, sendo o de Pocock o menor dos dois. As mesmas conclusões podem ser obtidas observando-se a Tabela A.4, em que o número de ensaios é praticamente o mesmo que o planejado. Desta forma, sob H_0 , o teste seqüencial gastaria o mesmo tempo que um teste com amostra fixa. Sob H_1 , os testes seqüenciais apresentam o número de ensaios menor que o planejado, indicando que a análise seqüencial apresenta vantagens nesta situação.

Os resultados encontrados para o caso sem as censuras aleatórias foram semelhantes

àqueles apresentados por DeMets & Gail (1985).

4.3.2 Teste Logrank Unilateral Seqüencial Simétrico em Grupo

Conforme mencionado anteriormente, muitos ensaios clínicos objetivam testar não a igualdade, mas a superioridade de um tratamento em relação a outro. Nestes casos, os testes unilaterais são mais apropriados. As hipóteses do teste logrank podem então ser expressas por

$$H_0: \log \omega = \log(\lambda_C/\lambda_E) \leq 0;$$

$$H_A: \log \omega = \log(\lambda_C/\lambda_E) > 0.$$

Usando as mesmas especificações encontradas em DeMets & Gail (1985) verifica-se que os poderes dos testes logrank unilaterais seqüenciais simétricos de Pocock e O'Brien & Fleming permanece em níveis aceitáveis, mesmo que as avaliações sejam executadas com diferentes incrementos a cada etapa da análise.

A estatística logrank padronizada usada é exatamente a mesma apresentada em (3.1).

Os valores críticos para rejeitar a hipótese nula é $C_P(5; 0, 05) = 2,122$ para o caso do teste de Pocock e $C_B(5; 0, 05) = 1,751\sqrt{5/k}$, $k = 1, \dots, 5$ para o teste de O'Brien & Fleming e os respectivos valores simétricos para aceitar a hipótese nula, como visto nas Tabelas 2.3 e 2.4 do Capítulo 2.

A Tabela B.1 mostra, como esperado, aumentos desiguais no número médio de falhas em cada etapa da análise, sob as diversas combinações de parâmetros, exceto quando a censura aleatória é 50% e 80% e a taxa de entrada $R = 10$ (gráficos das Figuras D.7 e D.8). Nestes casos, pode ser observado que nas etapas 2-5 ($\tau = 0$) e 2-4 ($\tau = 1$), os aumentos permaneceram praticamente constantes.

Como no caso dos testes bilaterais, pode-se notar que o total das médias de falhas encontrado é próximo daquele previsto inicialmente por (4.5) e o número de recrutamentos está de acordo com a taxa de Poisson fixada.

A Tabela B.2 mostra o número encontrado de rejeições em 500 simulações, podendo-se notar que o nível de significância estimado esteve dentro do intervalo de tolerância [15,35],

exceto em 5 casos (gráfico da Figura D.14), sendo dois com amostra fixa e sem censura aleatória (13 e 36 rejeições), 2 com o teste de Pocock (38 e 36 rejeições) com as taxas de 20% e 50% de censuras aleatórias e o último ocorrendo no caso do teste seqüencial de O'Brien & Fleming (38 rejeições) com a taxa de censura de 50%.

Sob a hipótese alternativa, no caso sem censura (gráficos das Figuras D.18 e D.19), pode-se notar que o número de rejeições fica próximo aos valores planejados, tanto para os testes de O'Brien & Fleming como para os testes com amostra fixa, enquanto que o teste de Pocock se mostra um pouco inferior a ambos.

Nos casos com censura, observa-se que o poder do teste vai diminuindo conforme aumenta a taxa de censura aleatória. O poder cai sensivelmente quando se tem baixa taxa de entrada ($R=10$ e $R=100$), como no caso do teste bilateral, pois o número de falhas diminui visivelmente. Quando a taxa de entrada é alta ($R=250$ e $R=1000$), o poder também diminui, mas com menor intensidade, devido ao grande número de indivíduos recrutados, alta taxa de censura administrativa e baixa taxa de risco.

Em geral, pode-se dizer que o poder do teste seqüencial de O'Brien & Fleming está próximo do poder para o teste com amostra fixa e o de Pocock sempre um pouco inferior a ambos. Por outro lado, ao observar a Tabela B.3, verifica-se que o teste de Pocock apresenta sempre o menor número médio de falhas e o menor número médio de ensaios, seja sob H_0 ou sob H_1 .

Sob a hipótese alternativa, ambos os testes seqüenciais apresentam o menor número médio de falhas, sendo o de Pocock o menor dos dois.

Como no caso dos testes bilaterais, sob a hipótese nula a análise seqüencial não apresenta grandes vantagens em relação à amostra fixa, uma vez que o número médio de falhas é equivalente nos três casos e o tempo gasto para se executar o teste seqüencial é praticamente o mesmo. Entretanto, sob a hipótese alternativa, os testes seqüenciais apresentam o menor número de falhas e menor tempo para o término do experimento.

4.3.3 Teste Logrank Unilateral Seqüencial Assimétrico em Grupo

Na seção anterior vimos que, sob a hipótese nula, a análise seqüencial não apresenta muitas vantagens. DeMets & Ware (1981, 1982) apresentam uma proposta de modo que os ensaios possam parar antecipadamente, com um menor número de ensaios, tanto sob a hipótese nula como sob a hipótese alternativa.

O teste unilateral assimétrico de Pocock rejeita a hipótese nula para valores maiores ou iguais a

$$\frac{2,016}{1,535\sqrt{k}} + \frac{1,535\sqrt{k}}{2}, \quad k = 1, \dots, K,$$

e aceita H_0 para valores menores que

$$\frac{-2,251}{1,535\sqrt{k}} + \frac{1,535\sqrt{k}}{2}.$$

O teste unilateral de O'Brien & Fleming rejeita a hipótese nula para valores maiores ou iguais a

$$1,733\sqrt{5/k}, \quad k = 1, \dots, K,$$

e aceita H_0 para valores menores que

$$\frac{-2,251}{1,358\sqrt{k}} + \frac{1,358\sqrt{k}}{2}.$$

como visto nas Tabelas 2.5 e 2.6 do Capítulo 2.

Os resultados de 500 simulações estão apresentados nas tabelas do apêndice C.

Assim como nos casos anteriores, a Tabela C.1 mostra que o total de médias de falhas encontrado está de acordo com seu valor esperado obtido através de (4.5) e o número médio de entradas está de acordo com a taxa média da distribuição de Poisson planejada.

O aumento médio do número de falhas por etapa é desigual em todas as situações, com exceção aos casos com $R = 10$, $\tau = 0$ e taxas de censuras de 50% e 80%, situações em que os aumentos ocorreram de forma uniforme entre as etapas 2-5 da análise (gráficos das Figuras D.11 e D.12).

A Tabela C.2 mostra o número encontrado de rejeições em 500 simulações. Pode-se observar que o nível de significância calculado esteve dentro do intervalo de tolerância [15,35] exceto em cinco situações (gráfico da Figura D.15): duas quando $R = 10$ e $\tau = 0$ com taxa de censura aleatória de 50% (Pocock 44 rejeições e O'Brien & Fleming 41), outras duas quando $R = 10$, $\tau = 1$ e 20% de censura aleatória (O'Brien & Fleming - 36 rejeições e com amostra fixa 36) e uma última para testes com amostra fixa com 80% de censura aleatória (37 rejeições).

Nas simulações feitas, sob a hipótese alternativa e sem censura, o poder dos testes seqüenciais mostrou-se um pouco inferior ao teste com amostra fixa (gráficos das Figuras D.20 e D.21). Dentre os dois testes seqüenciais, o de Pocock apresentou sempre o menor poder.

Nos casos com censura aleatória, observa-se uma redução no poder, tanto para os testes seqüenciais como para o teste com amostra fixa, sendo este último o mais poderoso.

Apesar do poder dos testes seqüenciais se mostrar um pouco inferior ao do teste com amostra fixa, pode-se ver que a análise seqüencial apresenta a vantagem esperada deste tipo de análise, ou seja, apresenta um menor número médio de falhas, como visto na Tabela C.3. Além disso, nota-se que o número médio de ensaios (Tabela C.4), neste caso, é inferior ao do teste unilateral simétrico apresentado na seção anterior, tanto sob a hipótese alternativa como sob a hipótese nula, estando de acordo com as idéias iniciais de DeMets & Ware (1980, 1982).

Capítulo 5

Discussão

Esta dissertação apresentou uma avaliação do poder do teste logrank seqüencial (unilateral e bilateral), na presença de censuras aleatórias, através de simulações realizadas no programa Splus, supondo que a distribuição de probabilidade do tempo de falha segue um modelo exponencial.

DeMets & Gail (1985) comentam que a prática de se executar testes em tempos fixos de calendário no lugar de aumentos iguais de falhas, é baseado no desejo de se atualizar os dados a cada avaliação. Princípios teóricos para definir os limites de parada que levam em conta aumentos desiguais já foram discutidos por Slud & Wei (1982), Tsiatis (1982) e Lan & DeMets (1983). Entretanto, muitos pesquisadores ainda preferem usar os limites de Pocock (1977) e de O'Brien & Fleming (1979) por não exigirem nenhum cálculo especial.

Os resultados das simulações indicaram que esses limites têm tamanho e poder empíricos próximos a seus níveis nominais, para ensaios com alta ($\lambda_C = 1$) ou baixa ($\lambda_C = 0,04$) taxa de risco, mesmo que o aumento de informações, entre as etapas da análise, não ocorram de forma homogênea.

Estudos similares podem ser feitos para outros padrões de entrada de indivíduos como, por exemplo, a uniforme; ou ainda usando outras distribuições de tempo de falha, como a Weibull. Além dessas modificações, estudos seqüenciais em um teste mais geral como o de Tarone & Ware (1977) pode ser avaliado similar ao apresentado por Tsiatis (1982).

O teste log-rank seqüencial foi estudado, também, por Bereta (2002) em seu traba-

lho de dissertação em que é comparado a outro teste não-paramétrico, o Kaplan-Meier ponderado, para testar a igualdade de duas curvas de sobrevivência.

Outros trabalhos como os de Lan & Lachin (1990), Lan & DeMets (1983), Kim, Boucher & Tsiatis (1995) e Lan, Rosenberger & Lachin (1995) merecem ser destacados. Eles apresentam discussões sobre os *ensaios de máxima informação* (os pacientes são acrescidos no ensaio até que se encontre a quantidade de informação pré-determinada) e os *ensaios de máxima duração* (o ensaio termina quando vence o tempo de calendário estabelecido). São estudos bem convenientes para os dados de sobrevivência, porém são mais complexos no sentido de calcular os erros de tipo I que os discutidos neste trabalho.

Apesar dos resultados apresentados nesta dissertação ajudarem a elucidar abordagens seqüenciais usualmente empregados, seria interessante estudar os métodos para dados mais heterogêneos, através da modelagem de regressão. Para o caso do modelo de Cox, pode-se tomar como ponto de partida o trabalho de Tsiatis, Rosner & Tritchler (1985).

Com o desenvolvimento tecnológico, as informações sobre os indivíduos envolvidos no estudo podem ser obtidas instantaneamente. Essa forma de coleta de dados, possibilita que as observações possam ser analisadas assim que chegam aos terminais de coleta e análise. Desta forma, modelos baseados em processo estocásticos de contagem a tempo contínuo ganham importância e estudos similares aos desenvolvidos neste trabalho apresentariam grande relevância prática.

Outro aspecto importante que mereceria atenção refere-se a experimentos seqüenciais para dados de sobrevivência multivariados com estrutura de dependência. Este problema ainda está totalmente em aberto, com poucas indicações na literatura.

Apêndice A

Resultados da Simulação para os Testes Bilaterais

Tabela A.1: Aumento médio no número de falhas, por etapas da análise, dos testes bilaterais e para várias taxas de entradas, λ_C , λ_E , τ e taxas de censuras de 0%, 20%, 50% e 80%

tx.		etapas da análise					falhas		Entra- das			
cens.	parâm.	R	τ	t_{rec}	1	2	3	4		5	Total	Esper.
0%	$\lambda_C = 1,0$ $\lambda_E = 0,5$	10	0	10,27	9,5	17,6	19,5	20,2	20,6	87,3	87,7	102,5
			1	9,55	10,3	18,2	20,4	2,8	18,4	88,3	87,6	96,0
		100	0	1,97	5,1	13,4	19,9	24,6	27,7	90,7	91,3	196,5
			1	1,33	7,3	18,4	26,0	22,7	15,9	90,2	90,0	132,7
	$\lambda_C = 0,04$ $\lambda_E = 0,02$	250	0	5,24	4,1	12,0	19,7	27,2	34,6	97,7	97,2	1310,8
			2	3,63	4,6	13,9	22,4	28,5	27,6	97,0	97,1	906,7
		1000	0	2,59	4,1	11,9	19,8	27,2	34,3	97,3	97,8	2586,5
			2	1,29	6,4	18,8	24,6	24,1	23,6	97,7	97,9	1289,4
20%	$\lambda_C = 1,0$ $\lambda_E = 0,5$	10	0	10,27	8,8	15,0	16,0	16,3	16,3	72,5	72,6	102,7
			1	9,55	9,0	15,1	16,6	16,7	14,4	71,9	72,1	95,1
		100	0	1,97	5,1	13,1	18,1	22,1	24,6	83,1	83,0	196,4
			1	1,33	6,9	17,3	23,3	19,4	12,6	79,5	79,6	132,7
	$\lambda_C = 0,04$ $\lambda_E = 0,02$	250	0	5,24	4,1	12,2	19,5	26,8	33,2	95,5	95,9	1311,1
			2	3,63	4,7	13,6	22,5	27,8	27,0	95,7	95,5	907,2
		1000	0	2,59	4,1	11,6	19,6	27,3	34,3	97,0	97,1	2591,0
			2	1,29	6,4	19,1	24,7	24,3	23,2	97,8	96,8	1290,2
50%	$\lambda_C = 1,0$ $\lambda_E = 0,5$	10	0	10,27	6,8	10,0	10,2	10,3	10,4	47,7	47,6	102,6
			1	9,55	7,1	10,4	10,7	10,7	8,1	46,9	46,7	95,8
		100	0	1,97	4,7	10,9	14,5	16,5	18,0	64,6	64,7	196,4
			1	1,33	6,3	14,5	18,3	12,6	6,3	58,1	58,2	133,2
	$\lambda_C = 0,04$ $\lambda_E = 0,02$	250	0	5,24	4,1	11,8	19,0	25,6	32,0	92,6	92,0	1313,9
			2	3,63	4,7	13,2	21,9	26,4	24,7	90,8	91,0	906,4
		1000	0	2,59	3,8	11,9	19,1	26,3	33,2	94,5	95,1	2589,8
			2	1,29	6,2	18,9	23,8	22,9	21,9	93,7	93,7	1288,2
80%	$\lambda_C = 1,0$ $\lambda_E = 0,5$	10	0	10,27	3,5	4,0	3,9	4,0	4,1	19,5	19,9	102,1
			1	9,55	3,5	4,3	4,3	4,3	2,8	19,2	19,1	95,6
		100	0	1,97	3,5	6,7	7,7	7,8	7,8	33,5	33,4	196,7
			1	1,33	4,7	8,4	8,9	3,5	0,8	26,3	26,3	132,6
	$\lambda_C = 0,04$ $\lambda_E = 0,02$	250	0	5,24	4,0	10,7	16,6	21,7	25,6	78,6	78,9	1309,8
			2	3,63	4,5	12,3	19,0	22,0	18,4	76,3	75,9	907,5
		1000	0	2,59	4,0	11,2	18,2	24,2	30,3	87,9	87,7	2591,2
			2	1,29	6,48	18,1	21,6	19,8	17,5	83,5	82,7	1292,2

Tabela A.2: Número de rejeições dos testes bilaterais simétricos por 500 simulações sob H_0 e H_1 e diversas taxas de censuras

hipó- teses	taxa de censura			0%			20%			50%			80%			
	parâmetros	R	t_{rec} τ	Poc	OBF	Fixa	Poc	OBF	Fixa	Poc	OBF	Fixa	Poc	OBF	Fixa	
H_0	$\lambda_C = 1, 0$	10	10,27	0	37	36	32	30	29	23	26	25	20	21	28	29
			9,55	1	41	36	36	36	38	35	23	26	28	16	15	20
	$\lambda_E = 1, 0$	100	1,97	0	31	26	25	34	25	29	26	21	21	26	18	21
			1,33	1	27	24	24	24	24	20	24	20	21	18	18	25
	$\lambda_C = 0, 04$	250	5,24	0	34	32	28	36	34	30	12	23	17	26	22	20
			3,63	2	25	25	23	31	23	23	19	28	22	28	16	18
$\lambda_E = 0, 04$	1000	2,59	0	18	26	26	30	28	24	23	33	25	18	30	30	
		1,29	2	17	16	18	36	32	32	29	24	19	23	30	31	
H_1	$\lambda_C = 1, 0$	10	10,27	0	424	447	445	358	402	407	281	322	327	115	141	
			9,55	1	407	438	438	364	404	408	245	294	296	120	142	146
	$\lambda_E = 0, 5$	100	1,97	0	426	458	462	385	422	423	355	404	404	213	262	
			1,33	1	426	456	459	404	429	433	338	371	382	159	190	
	$\lambda_C = 0, 04$	250	5,24	0	427	457	459	426	456	452	410	457	459	382	414	
			3,63	2	422	453	455	426	454	458	422	455	462	391	416	
$\lambda_E = 0, 02$	1000	2,59	0	414	441	443	417	449	460	408	444	451	410	449		
		1,29	2	419	449	456	423	456	460	412	449	452	395	426		

obs.: Poc : Pocock; OBF: O'Brien & Fleming

Tabela A.3: Número médio de falhas dos testes bilaterais sob H_0 e H_1 e diversas taxas de censuras

hipó-teses	taxa de censura		0%			20%			50%			80%				
	parâmetros	R	τ	Poc	OBF	Fixa	Poc	OBF	Fixa	Poc	OBF	Fixa	Poc	OBF	Fixa	
H_0	$\lambda_C = 1, 0$	10	0	90,6	92,8	93,9	74,3	75,7	76,5	47,3	48,1	48,6	19,9	20,0	20,1	
			1	87,0	89,9	91,1	71,8	73,4	74,4	46,5	47,3	47,7	18,9	19,1	19,2	
	$\lambda_E = 1, 0$	100	0	107,0	109,7	110,8	94,7	97,5	98,6	72,7	73,7	74,3	34,4	35,1	35,4	
			1	102,4	105,2	105,8	89,4	91,9	91,5	62,0	63,2	63,4	26,0	26,5	26,5	
				0	123,1	125,5	127,3	120,3	123,9	125,8	119,5	119,4	120,4	96,5	99,1	99,9
				2	124,8	127,8	129,0	121,4	125,0	125,7	116,4	117,7	118,6	91,7	94,6	95,2
			0	127,0	128,4	129,4	123,9	127,0	128,5	122,5	124,1	125,6	111,9	112,9	113,7	
			2	127,1	128,9	129,8	123,2	126,3	127,5	118,8	121,7	122,7	103,4	105,0	105,9	
H_1	$\lambda_C = 1, 0$	10	0	55,1	62,7	87,3	50,9	56,0	72,5	37,9	40,8	47,7	17,9	18,4	19,5	
			1	57,7	66,5	88,3	49,7	56,1	71,9	38,5	41,5	46,9	17,2	18,3	19,2	
	$\lambda_E = 0, 5$	100	0	58,8	62,6	90,7	57,5	62,0	83,1	46,3	49,6	64,6	28,5	29,9	33,5	
			1	56,8	65,4	90,2	54,4	62,1	79,5	44,1	49,4	58,1	23,4	25,2	26,3	
				0	62,7	67,5	97,7	62,1	66,8	95,9	60,6	64,5	92,6	53,8	57,6	78,6
				2	63,8	69,3	97,0	63,3	68,7	95,7	58,9	64,1	90,8	53,0	58,5	76,3
			0	62,6	67,0	97,3	63,3	67,0	97,0	63,6	67,7	94,5	55,7	59,8	87,9	
			2	62,3	70,4	97,7	59,9	69,6	97,8	61,0	68,3	93,7	57,1	63,1	83,5	

obs.: Poc : Pocock; OBF: O'Brien & Fleming

Tabela A.4: Número médio de ensaios dos testes bilaterais sob H_0 e H_1 e diversas taxas de censuras

hipó-teses	taxa de censura	R	τ	0%		20%		50%		80%	
				Poc	OBF	Poc	OBF	Poc	OBF	Poc	OBF
H_0	$\lambda_C = 1, 0$	10	0	4,84	4,95	4,86	4,92	4,86	4,95	4,94	4,97
			1	4,80	4,94	4,83	4,93	4,88	4,97	4,92	4,99
	$\lambda_E = 1, 0$	100	0	4,87	4,97	4,34	4,96	4,90	4,97	4,88	4,97
			1	4,85	4,97	4,88	4,96	4,87	4,97	4,89	4,98
	$\lambda_C = 0, 04$	250	0	4,88	4,96	4,85	4,95	4,98	4,98	4,87	4,97
			2	4,87	4,97	4,87	4,98	4,93	4,97	4,85	4,97
$\lambda_E = 0, 04$	1000	0	4,94	4,98	4,87	4,97	4,91	4,96	4,94	4,98	
		2	4,91	4,97	4,86	4,96	4,86	4,96	4,90	4,96	
H_1	$\lambda_C = 1, 0$	10	0	3,41	3,81	3,69	4,00	4,05	4,34	4,58	4,72
			1	3,45	3,88	3,58	3,97	4,22	4,40	4,48	4,75
	$\lambda_E = 0, 5$	100	0	3,73	3,94	3,84	4,10	3,91	4,15	4,34	4,53
			1	3,32	3,67	3,48	3,85	3,68	4,01	4,24	4,52
	$\lambda_C = 0, 04$	250	0	3,84	4,06	3,84	4,06	3,87	4,06	3,94	4,14
			2	3,77	4,00	3,78	4,00	3,70	3,94	3,80	4,09
$\lambda_E = 0, 02$	1000	0	3,84	4,05	3,88	4,06	3,94	4,14	3,78	4,00	
		2	3,50	3,86	3,41	3,81	3,53	3,85	3,61	3,90	

obs.: Poc : Pocock; OBF: O'Brien & Fleming

Apêndice B

Resultados da Simulação para os Testes Unilaterais Simétricos

Tabela B.1: Aumento médio no número de falhas, por etapas da análise, dos testes unilaterais simétricos e para várias taxas de entradas, λ_C , λ_E , τ e taxas de censuras de 0%, 20%, 50% e 80%

tx.					etapas da análise					falhas		Entra- das	
cens.	parâm.	R	τ	t_{rec}	1	2	3	4	5	Total	Esper.		
0%	$\lambda_C = 1$ $\lambda_E = 0,5$	10	0	8,65	7,7	14,4	16,0	16,8	16,8	71,7	71,6	86,4	
			1	7,93	7,6	14,9	16,4	17,2	15,0	71,1	71,5	78,7	
		100	0	1,75	4,2	11,0	16,3	20,3	23,1	75,0	75,4	174,2	
			1	1,12	6,1	15,7	21,7	17,5	12,8	73,8	73,6	111,9	
	$\lambda_C = 0,04$ $\lambda_E = 0,02$	250	0	4,72	3,3	9,7	15,9	22,1	28,6	79,6	79,3	1179,3	
			2	3,1	3,9	11,5	19,2	22,9	22,0	79,6	79,3	787,5	
		1000	0	2,34	3,3	9,6	16,1	22,3	28,2	79,6	80,0	2338,5	
			2	1,09	5,5	16,4	19,8	19,1	18,8	79,6	79,7	1089,4	
	20%	$\lambda_C = 1$ $\lambda_E = 0,5$	10	0	8,65	6,9	12,1	13,1	13,5	13,8	59,5	59,6	86,0
				1	7,93	7,1	12,0	13,0	13,2	10,3	55,5	59,1	79,7
			100	0	1,75	4,0	10,3	15,2	18,2	20,4	68,0	69,0	175,8
				1	1,12	5,9	14,6	19,5	14,6	9,1	63,6	65,5	112,6
$\lambda_C = 0,04$ $\lambda_E = 0,02$		250	0	4,72	3,3	10,1	16,1	21,4	27,5	78,4	78,3	1179,6	
			2	3,15	4,1	11,6	18,3	22,1	21,4	77,6	78,1	787,3	
		1000	0	2,34	3,1	9,8	16,1	22,3	28,1	79,5	79,5	2339,8	
			2	1,09	5,7	16,4	19,5	19,3	18,4	79,3	78,8	1091,7	
50%		$\lambda_C = 1$ $\lambda_E = 0,5$	10	0	8,65	5,4	8,1	8,6	8,7	8,6	39,4	39,5	86,2
				1	7,93	5,4	8,5	8,7	8,7	6,1	37,5	38,6	77,2
			100	0	1,75	3,9	9,2	12,3	14,2	15,2	54,8	54,7	175,5
				1	1,12	5,6	12,4	15,3	9,7	5,2	48,2	48,3	112,0
	$\lambda_C = 0,04$ $\lambda_E = 0,02$	250	0	4,72	3,4	9,7	15,8	20,8	26,2	75,9	75,5	1181,0	
			2	3,15	3,9	11,3	18,2	21,0	19,9	74,4	74,7	788,3	
		1000	0	2,34	3,4	9,5	15,5	21,9	27,7	78,0	78,0	2340,0	
			2	1,09	5,6	16,1	18,8	18,2	17,6	76,2	76,5	1089,4	
	80%	$\lambda_C = 1$ $\lambda_E = 0,5$	10	0	8,65	2,9	3,5	3,4	3,3	3,4	16,6	16,7	86,8
				1	7,93	3,0	3,6	3,6	3,6	2,2	15,9	15,8	79,1
			100	0	1,75	3,0	5,5	6,7	6,7	6,7	28,7	29,0	174,6
				1	1,12	4,2	7,3	7,4	2,5	0,6	22,0	22,1	111,9
$\lambda_C = 0,04$ $\lambda_E = 0,02$		250	0	4,72	3,7	9,1	13,8	17,9	21,5	65,6	65,6	1178,5	
			2	3,15	3,8	10,5	15,9	17,2	15,2	62,7	63,0	786,9	
		1000	0	2,34	3,2	9,2	15,1	20,4	25,0	72,9	72,5	2338,0	
			2	1,09	5,6	15,3	17,2	15,5	13,9	67,5	67,8	1090,3	

Tabela B.2: Número de rejeições dos testes unilaterais simétricos por 500 simulações sob H_0 e H_1 e diversas taxas de censuras

hipó- teses	taxa de censura			0%			20%			50%			80%				
	parâmetros	R	t_{rec}	τ	Poc		OBF		Fixa		Poc		OBF		Fixa		
					Poc	OBF	Fixa	Poc	OBF	Fixa	Poc	OBF	Fixa	Poc	OBF	Fixa	
H_0	$\lambda_C = 1$	10	8,65	0	35	22	18	31	22	22	27	24	20	19	21	21	
			7,93	1	28	30	36	30	24	24	26	22	22	25	25	30	30
	$\lambda_E = 1$	100	1,75	0	32	29	24	26	28	27	33	31	26	24	27	23	
			1,12	1	25	15	13	31	31	28	24	18	23	20	21	27	
	$\lambda_C = 0,04$	250	4,72	0	31	27	21	23	28	24	26	25	24	22	22	26	29
			3,15	2	29	31	24	38	32	31	36	38	33	25	30	27	27
$\lambda_E = 0,04$	1000	2,34	0	30	30	32	27	29	21	24	31	24	24	33	34	25	
		1,09	2	24	21	19	31	30	25	30	24	22	30	20	23	23	
H_1	$\lambda_C = 1$	10	8,65	0	394	437	436	364	394	400	305	344	345	164	191	186	
			7,93	1	401	432	430	386	384	393	279	327	339	154	192	189	
	$\lambda_E = 0,5$	100	1,75	0	417	452	450	389	430	434	359	404	413	243	289	287	
			1,12	1	414	439	446	395	431	434	339	381	388	204	230	244	
	$\lambda_C = 0,04$	250	4,72	0	418	453	453	436	464	464	426	454	455	376	424	426	
			3,15	2	428	462	463	420	445	448	409	443	440	394	425	429	
$\lambda_E = 0,02$	1000	2,34	0	431	458	466	430	463	462	429	455	461	408	450	452		
		1,09	2	419	440	448	433	452	458	416	451	450	408	431	431		

obs.: Poc : Pocock; OBF: O'Brien & Fleming

Tabela B.3: Número médio de falhas dos testes unilaterais simétricos sob H_0 e H_1 e diversas taxas de censuras

hipó- teses	taxa de censura			0%			20%			50%			80%			
	parâmetros	R	τ	Poc	OBF	Fixa	Poc	OBF	Fixa	Poc	OBF	Fixa	Poc	OBF	Fixa	
H_0	$\lambda_C = 1$	10	0	70,3	74,6	76,5	58,5	61,5	62,9	37,8	39,6	40,4	16,3	16,7	17,0	
			1	70,8	73,9	75,2	57,8	60,9	62,0	36,5	38,3	39,0	15,1	15,7	15,9	
	$\lambda_E = 1$	100	0	87,3	90,6	92,7	78,3	81,5	83,9	58,7	61,8	63,3	29,4	30,3	31,1	
			1	82,2	95,9	87,3	71,7	74,9	76,0	51,2	53,1	53,6	21,9	22,6	22,6	
	H_1	$\lambda_C = 0,04$	250	0	98,8	101,7	104,4	96,5	100,6	103,3	91,3	95,2	98,6	79,7	81,2	83,7
				2	96,7	101,8	104,8	94,6	99,2	102,4	91,0	94,2	97,3	74,8	78,3	80,0
$\lambda_E = 0,04$		1000	0	98,1	102,4	106,4	97,9	101,9	106,0	96,5	99,7	102,8	87,0	90,5	93,6	
			2	98,8	103,5	105,6	98,0	101,6	104,0	93,2	97,6	99,6	80,8	84,8	86,0	
H_1	$\lambda_C = 1$	10	0	44,5	55,0	71,7	40,3	44,9	59,5	28,7	31,2	39,4	14,5	15,2	16,6	
			1	43,9	50,8	71,1	39,4	44,1	55,5	28,8	31,2	37,5	14,2	14,8	15,9	
	$\lambda_E = 0,5$	100	0	45,6	50,4	75,0	42,9	46,8	68,0	37,5	41,2	54,8	23,2	24,6	28,7	
			1	44,6	51,6	73,8	41,9	48,1	63,6	34,6	39,3	48,2	18,9	20,8	22,0	
$\lambda_C = 0,04$	250	0	47,8	51,2	79,6	47,3	50,6	78,4	46,1	49,6	75,9	44,3	47,1	65,6		
		2	48,1	52,9	79,6	47,3	51,6	77,6	46,7	51,6	74,4	41,5	45,6	62,7		
$\lambda_E = 0,02$	1000	0	47,6	51,6	79,6	48,7	51,2	79,5	48,7	51,4	78,0	46,1	45,5	72,9		
		2	49,0	54,2	79,6	47,4	53,7	79,3	47,0	52,5	76,2	43,3	49,0	67,5		

obs.: Poc : Pocock; OBF: O'Brien & Fleming

Tabela B.4: Número médio de ensaios dos testes unilaterais simétricos sob H_0 e H_1 e diversas taxas de censuras

hipó-teses	taxa de censura		0%		20%		50%		80%	
	parâmetros	R	Poc	OBF	Poc	OBF	Poc	OBF	Poc	OBF
H_0	$\lambda_C = 1$	0	4,63	4,88	4,68	4,9	4,70	4,91	4,79	4,92
		10	4,72	4,91	4,69	4,91	4,68	4,90	4,72	4,92
	$\lambda_E = 1$	0	4,78	4,93	4,74	4,70	4,70	4,91	4,76	4,90
		100	4,73	4,91	4,73	4,92	4,76	4,92	4,77	4,92
		0	4,81	4,72	4,77	4,92	4,73	4,89	4,83	4,90
		250	4,70	4,89	4,70	4,89	4,76	4,88	4,75	4,91
H_1	$\lambda_C = 0,04$	0	4,73	4,89	4,73	4,88	4,78	4,90	4,75	4,90
		1000	4,72	4,91	4,76	4,90	4,72	4,91	4,73	4,93
	$\lambda_E = 0,5$	0	3,36	3,71	3,60	3,95	3,78	4,08	4,37	4,58
		10	3,30	3,73	3,65	4,01	3,88	4,15	4,42	4,60
		100	3,57	3,87	3,64	3,91	3,79	4,09	4,16	4,39
		1000	3,19	3,56	3,35	3,70	3,48	3,83	3,98	4,36
	$\lambda_C = 0,04$	0	3,70	3,92	3,67	3,89	3,66	3,89	3,87	4,07
		250	3,53	3,78	3,54	3,79	3,61	3,87	3,64	3,91
	$\lambda_E = 0,02$	0	3,68	3,92	3,72	3,90	3,78	3,97	3,78	3,96
		1000	3,39	3,66	3,30	3,63	3,36	3,67	3,38	3,73

obs.: Poc : Pocock; OBF: O'Brien & Fleming

Apêndice C

Resultados da Simulação para os Testes Unilaterais Assimétricos

Tabela C.1: Aumento médio no número de falhas, por etapas da análise, dos testes unilaterais assimétricos e para várias taxas de entradas, λ_C , λ_E , τ e taxas de censuras de 0%, 20%, 50% e 80%

tx. cens.	parâm.	R	τ	t_{rec}	etapas da análise					falhas		Entra- das	
					1	2	3	4	5	Total	Esper.		
0%	$\lambda_C = 1$	10	0	8,65	7,3	13,9	16,2	16,7	17,1	71,3	71,6	86,0	
			1	7,93	7,7	14,7	17,0	17,5	14,6	71,6	71,5	79,6	
	$\lambda_E = 0,5$	100	0	1,75	4,1	10,9	16,4	20,4	23,4	74,3	75,4	174,4	
			1	1,12	5,9	15,7	21,7	17,4	12,7	67,5	73,6	112,0	
	$\lambda_C = 0,04$	250	0	4,72	3,4	9,7	15,8	22,0	28,0	78,8	79,3	1180,0	
			2	3,15	4,0	11,5	19,3	23,2	22,1	80,1	79,3	788,3	
			0	2,34	3,3	9,71	2,0	22,3	28,9	80,2	80,0	2338,6	
			2	1,09	5,7	16,3	19,6	18,9	19,0	79,5	79,7	1089,9	
	20%	$\lambda_C = 1$	10	0	8,65	6,9	12,0	13,4	13,7	13,8	52,9	59,6	86,7
				1	7,93	7,1	12,6	13,9	14,1	11,7	52,3	59,1	79,5
$\lambda_E = 0,5$		100	0	1,97	4,0	10,4	15,4	18,6	20,7	69,1	69,0	175,8	
			1	1,12	5,9	14,7	19,6	14,9	10,0	65,3	65,4	111,7	
$\lambda_C = 0,04$		250	0	4,72	3,3	9,8	16,2	21,8	27,8	78,9	78,3	1180,2	
			2	3,15	4,2	11,7	18,8	22,4	21,3	78,4	78,1	787,9	
			0	2,34	3,2	9,8	15,9	22,2	28,0	79,1	79,5	2339,5	
			2	1,09	5,7	16,2	19,4	19,1	18,4	78,8	78,9	1089,8	
50%		$\lambda_C = 1$	10	0	8,65	5,4	8,1	8,5	8,8	8,5	34,	39,5	86,1
				1	7,93	5,6	8,4	8,9	8,7	6,5	32,5	38,6	79,3
	$\lambda_E = 0,5$	100	0	1,75	3,8	9,2	12,1	14,3	15,3	54,7	54,7	174,7	
			1	1,12	5,4	12,6	15,4	9,7	5,45	4,2	48,3	112,3	
	$\lambda_C = 0,04$	250	0	4,72	3,3	9,8	15,6	21,0	26,2	75,9	75,5	1179,0	
			2	3,15	3,9	11,2	18,5	21,6	19,6	74,8	74,7	787,6	
			0	2,34	3,2	9,6	15,8	21,8	27,3	77,7	78,0	2342,9	
			2	1,09	5,8	16,3	18,8	18,3	17,6	76,8	76,5	1090,5	
	80%	$\lambda_C = 1$	10	0	8,65	2,8	3,6	3,4	3,4	3,5	16,7	16,7	86,6
				1	7,9	3,0	3,7	3,6	3,6	2,1	16,1	15,8	79,6
$\lambda_E = 0,5$		100	0	1,75	3,1	5,8	6,5	6,7	7,0	29,1	29,0	175,0	
			1	1,12	4,1	7,5	7,2	2,6	0,7	22,1	22,1	112,4	
$\lambda_C = 0,04$		250	0	4,72	3,2	8,9	13,9	18,0	21,8	65,9	65,6	1182,5	
			2	3,15	3,8	10,5	16,1	17,5	15,0	62,9	63,0	788,2	
			0	2,34	3,2	9,3	14,9	20,4	25,0	72,9	72,5	2339,0	
			2	1,09	5,6	15,5	16,9	15,4	14,0	67,4	67,8	1090,3	

Tabela C.2: Número de rejeições dos testes unilaterais assimétricos por 500 simulações sob H_0 e H_1 e diversas taxas de censuras

hipó- teses	taxa de censura			0%			20%			50%			80%																		
	parâmetros	R	t_{rec}	τ	Poc		OBF		Fixa		Poc		OBF		Fixa																
					Poc	OBF	Poc	OBF	Poc	OBF	Poc	OBF	Poc	OBF	Poc	OBF	Poc	OBF	Poc	OBF	Poc	OBF	Poc	OBF	Poc	OBF	Poc	OBF	Poc	OBF	Poc
H_0	$\lambda_C = 1$	10	8,65	0	33	25	24	29	24	21	44	41	33	16	21	33	16	21	27												
					22	21	23	35	36	36	24	23	23	28	30	26															
					24	21	21	28	25	23	26	18	20	16	17	21															
	$\lambda_E = 1$	100	1,12	1	20	17	20	31	28	28	25	26	33	19	20	33	19	20	22												
					25	21	25	28	33	19	18	22	28	27	25	25															
					23	23	25	27	25	27	19	23	28	28	28	33															
$\lambda_C = 0,04$	250	3,15	2	19	16	21	15	16	25	30	30	30	28	30	30	28	30	37													
				24	24	26	18	21	19	19	21	20	23	25	28																
				24	24	26	18	21	19	19	21	23	25	28																	
H_1	$\lambda_C = 1$	10	8,65	0	359	403	443	324	377	406	252	304	332	149	187	202	149	187	202												
					370	422	446	336	392	409	243	290	304	127	161	170															
					359	410	454	346	397	439	287	348	401	194	247	275															
	$\lambda_E = 0,5$	100	1,75	1	384	432	458	346	399	424	303	347	375	169	214	240	169	214	240												
					346	410	456	346	403	457	335	382	451	306	368	430															
					348	411	465	358	414	454	343	408	445	320	380	419															
$\lambda_C = 0,04$	250	3,15	2	328	393	454	352	397	459	341	391	464	339	397	454	339	397	454													
				384	439	467	357	415	449	384	428	453	361	409	434																
				384	439	467	357	415	449	384	428	453	361	409	434																

obs.: Poc : Pocock; OBF: O'Brien & Fleming

Tabela C.3: Número médio de falhas dos testes unilaterais assimétricos sob H_0 e H_1 e diversas taxas de censuras

hipó- teses	taxa de censura			0%			20%			50%			80%			
	parâmetros	R	τ	Poc	OBF	Fixa	Poc	OBF	Fixa	Poc	OBF	Fixa	Poc	OBF	Fixa	
H_0	$\lambda_C = 1$	10	0	32,5	39,8	76,2	27,3	33,7	62,8	19,3	23,4	40,9	8,6	10,0	16,8	
			1	35,1	41,9	75,2	29,5	35,8	61,9	19,5	23,0	39,1	7,7	9,0	15,7	
	$\lambda_E = 1$	100	0	29,6	40,2	92,0	28,4	36,7	82,2	23,2	29,4	63,0	13,1	16,1	31,0	
			1	38,7	48,7	87,1	36,6	43,9	75,9	27,2	32,3	52,8	13,7	15,8	21,9	
			250	0	33,4	43,1	105,0	32,9	40,4	103,7	30,0	38,6	98,5	25,9	33,7	83,3
				2	34,2	45,4	105,0	34,3	45,5	103,0	33,3	44,1	97,0	30,6	38,6	79,5
		1000	0	32,3	42,6	105,9	31,0	40,3	105,3	30,0	39,0	103,0	29,3	36,7	93,7	
			2	46,0	56,6	104,7	43,3	53,0	104,5	39,8	50,9	100,7	36,0	45,6	87,2	
H_1	$\lambda_C = 1$	10	0	34,6	44,3	71,3	32,7	40,0	59,8	22,8	28,2	39,3	9,9	12,1	16,7	
			1	36,9	47,4	71,6	33,7	40,9	59,4	23,1	28,3	38,1	9,9	12,1	16,1	
	$\lambda_E = 0,5$	100	0	33,9	42,5	75,3	29,0	39,8	69,1	26,6	33,8	54,7	16,6	20,7	29,1	
			1	37,4	48,2	73,5	35,7	45,5	65,2	28,6	36,4	48,6	16,2	19,3	22,1	
			250	0	35,5	43,9	78,8	36,0	43,0	78,9	30,9	39,5	75,9	30,1	37,3	65,9
				2	36,5	45,6	80,1	38,6	46,7	78,4	36,1	44,7	74,8	32,1	39,9	62,9
		1000	0	34,8	43,3	80,2	34,8	42,8	79,1	34,5	42,3	77,7	33,4	40,3	72,9	
			2	40,6	50,3	79,5	40,8	50,1	78,8	37,4	47,7	76,8	36,4	44,4	67,4	

obs.: Poc : Pocock; OBF: O'Brien & Fleming

Tabela C.4: Número médio de ensaios dos testes unilaterais assimétricos sob H_0 e H_1 e diversas taxas de censuras

hipó-teses	taxa de censura		0%		20%		50%		80%	
	R	τ	Poc	OBf	Poc	OBf	Poc	OBf	Poc	OBf
H_0	$\lambda_C = 1$	0	2,44	2,87	2,38	2,87	2,53	3,01	2,58	3,00
		10	2,51	2,92	2,54	2,99	2,51	2,94	2,38	2,80
	$\lambda_E = 1$	0	2,34	2,86	2,42	2,86	2,39	2,81	2,43	2,86
		100	2,46	2,94	2,51	2,93	2,46	2,88	2,44	2,92
	$\lambda_C = 0,04$	0	2,46	2,87	2,47	2,79	2,42	2,79	2,38	2,77
		250	2,38	2,83	2,39	2,84	2,42	2,88	2,49	2,90
$\lambda_E = 0,04$	0	2,44	2,85	2,40	2,78	2,38	2,75	2,44	2,78	
	1000	2,61	3,04	2,49	2,88	2,41	2,89	2,39	2,86	
H_1	$\lambda_C = 1$	0	2,78	3,38	3,00	3,57	3,10	3,73	3,04	3,68
		10	2,84	3,47	3,05	3,57	3,10	3,71	3,00	3,65
	$\lambda_E = 0,5$	0	2,92	3,45	2,77	3,46	2,96	3,54	3,16	3,75
		100	2,79	3,35	2,89	3,48	2,95	3,55	3,23	3,85
	$\lambda_C = 0,04$	0	3,05	3,53	3,07	3,49	2,85	3,36	3,01	3,51
		250	2,93	3,41	3,07	3,50	2,99	3,46	3,04	3,54
$\lambda_E = 0,02$	0	3,00	3,49	3,03	3,48	3,04	3,48	3,09	3,52	
	1000	2,93	3,46	2,97	3,45	2,80	3,38	2,96	3,46	

obs.: Poc : Pocock; OBf: O'Brien & Fleming

Apêndice D

Gráficos dos Resultados da Simulação para os Testes Bilaterais e Unilaterais

D.1 Aumento Médio do Número de Falhas por Etapa da Análise

D.1.1 Testes Logrank Bilaterais

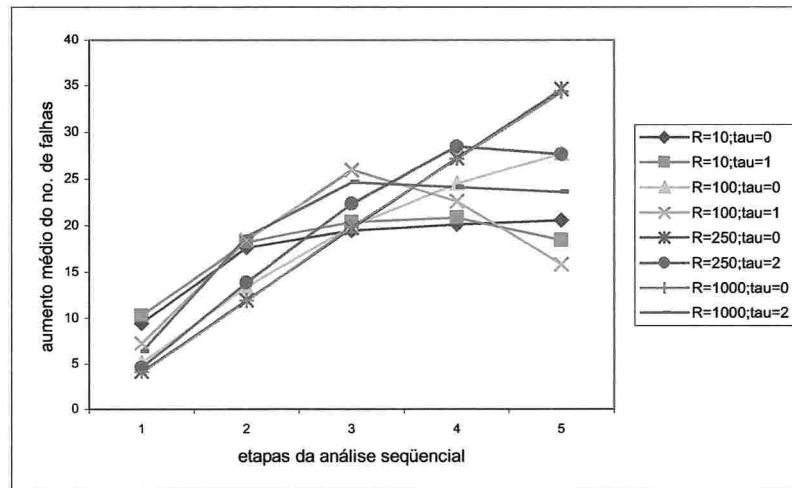


Figura D.1: Aumento médio do número de falhas por etapa e sem censura.

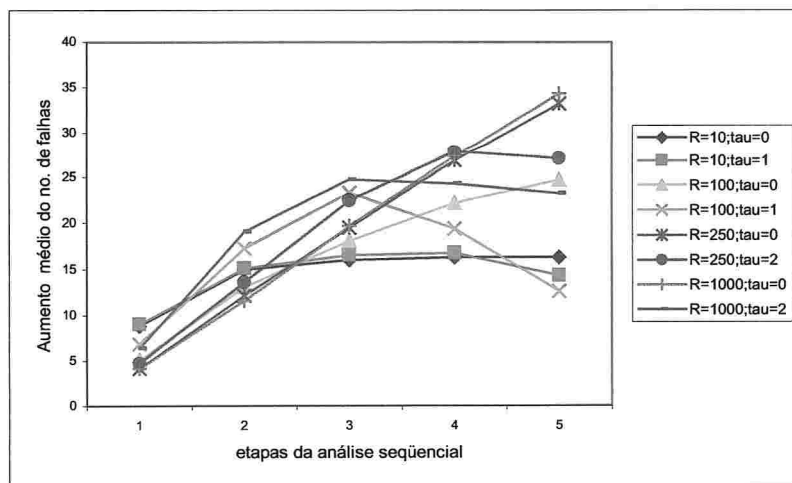


Figura D.2: Aumento médio do número de falhas por etapa com 20% de censura aleatórias.

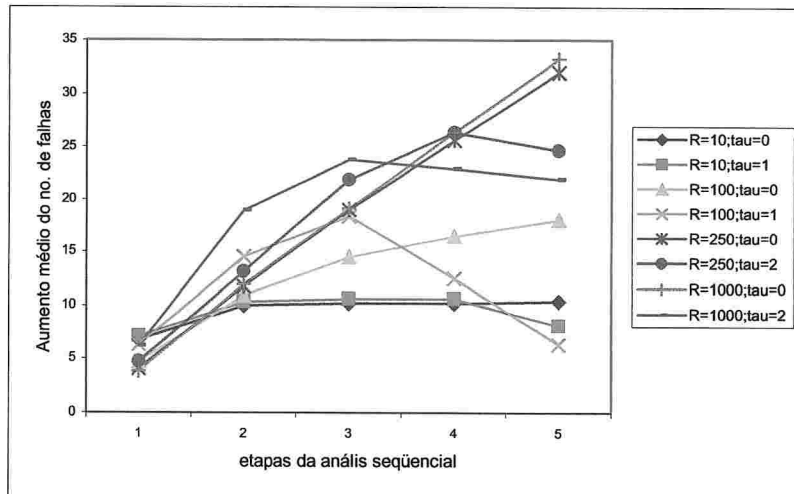


Figura D.3: Aumento médio do número de falhas por etapa com 50% de censuras aleatórias.

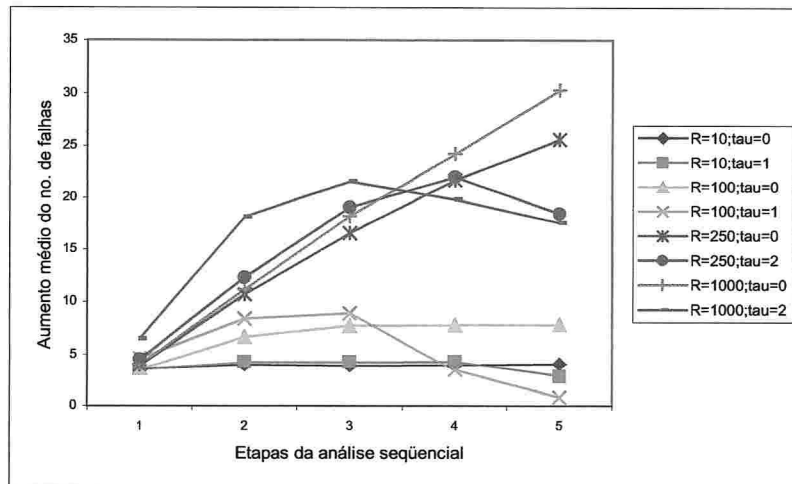


Figura D.4: Aumento médio do número de falhas por etapa com 80% de censuras aleatórias.

D.1.2 Testes Logrank Unilaterais Simétricos

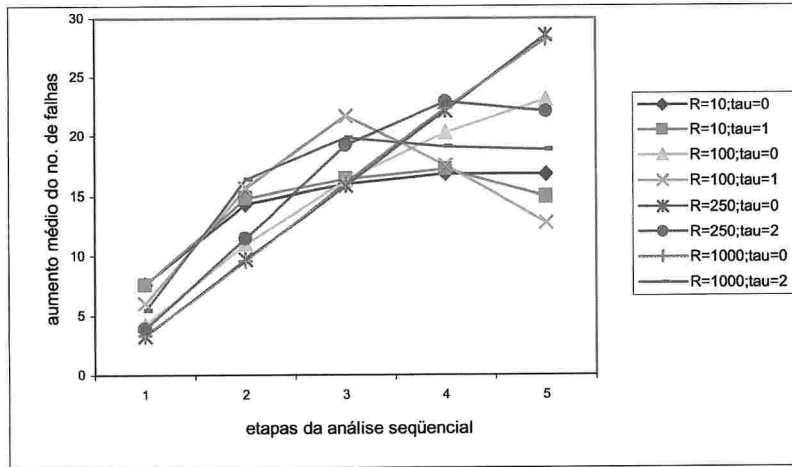


Figura D.5: Aumento médio do número de falhas por etapa e sem censura.

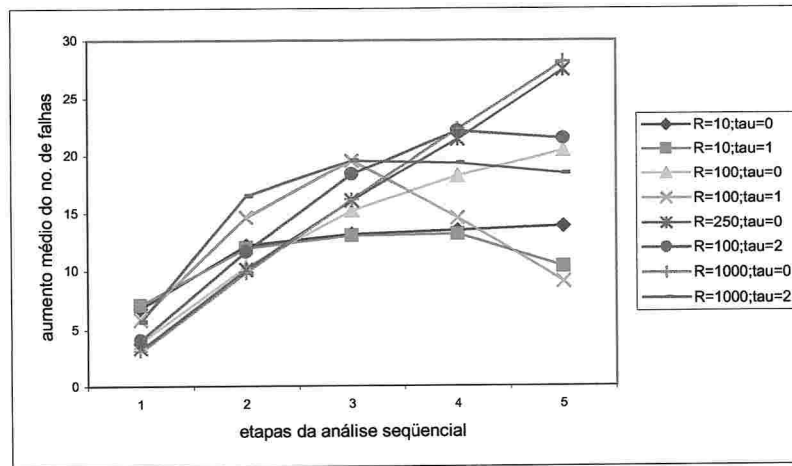


Figura D.6: Aumento médio do número de falhas por etapa com 20% de censura aleatórias.

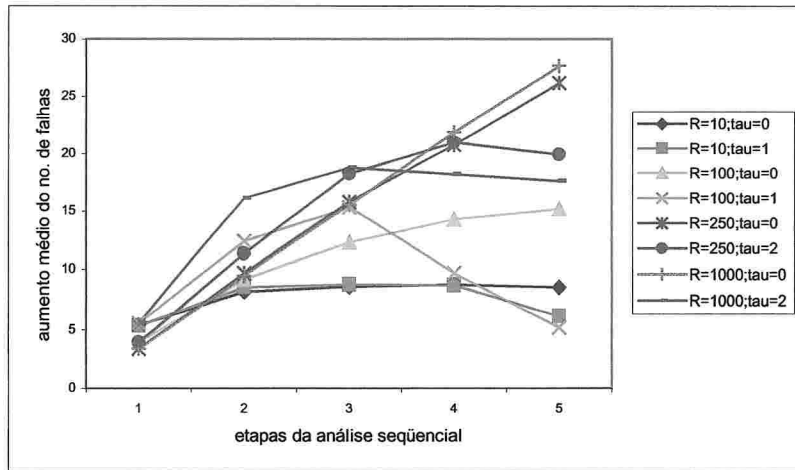


Figura D.7: Aumento médio do número de falhas por etapa com 50% de censuras aleatórias.

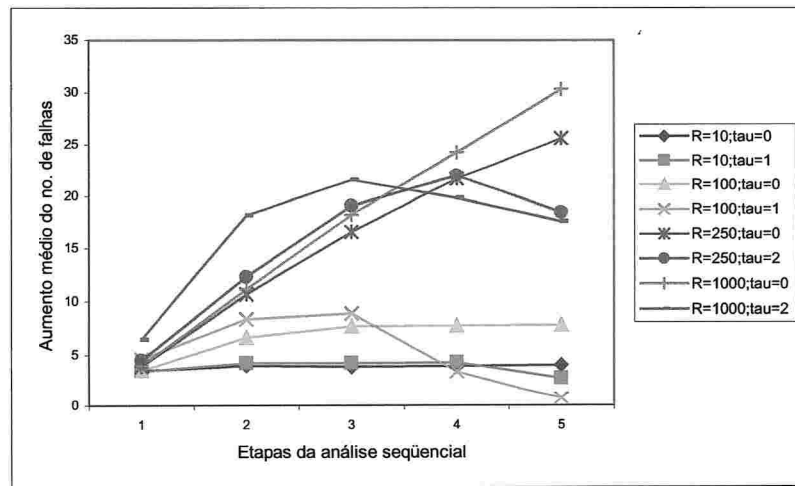


Figura D.8: Aumento médio do número de falhas por etapa com 80% de censuras aleatórias.

D.1.3 Testes Logrank Unilaterais Assimétricos

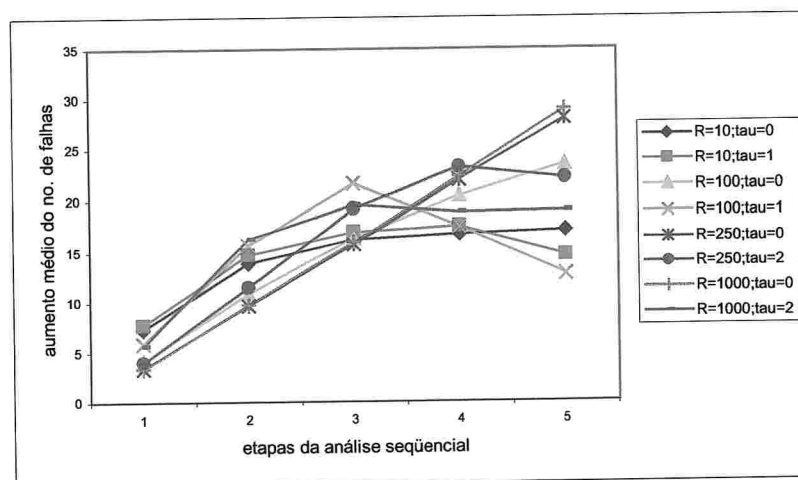


Figura D.9: Aumento médio do número de falhas por etapa e sem censura.

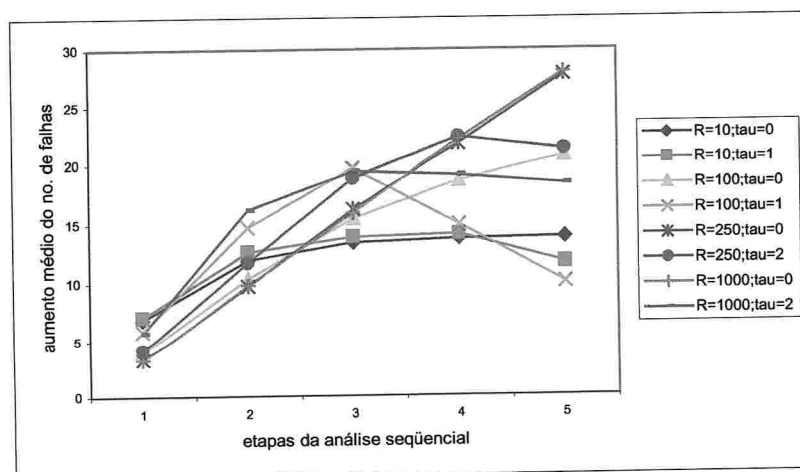


Figura D.10: Aumento médio do número de falhas por etapa com 20% de censura aleatórias.

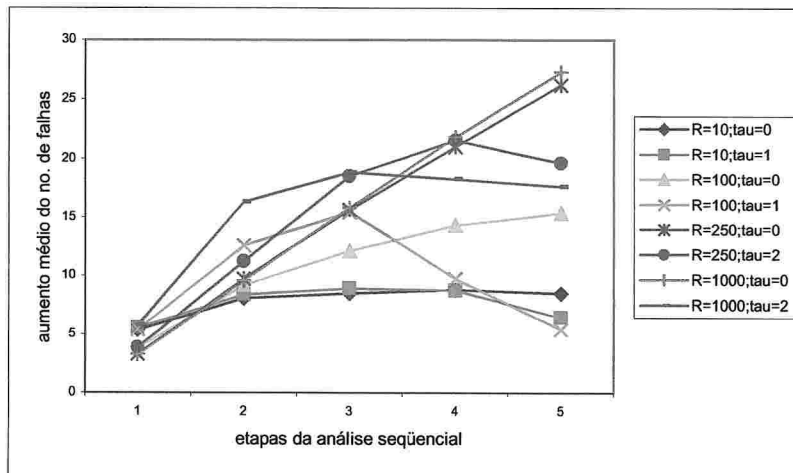


Figura D.11: Aumento médio do número de falhas por etapa com 50% de censuras aleatórias.

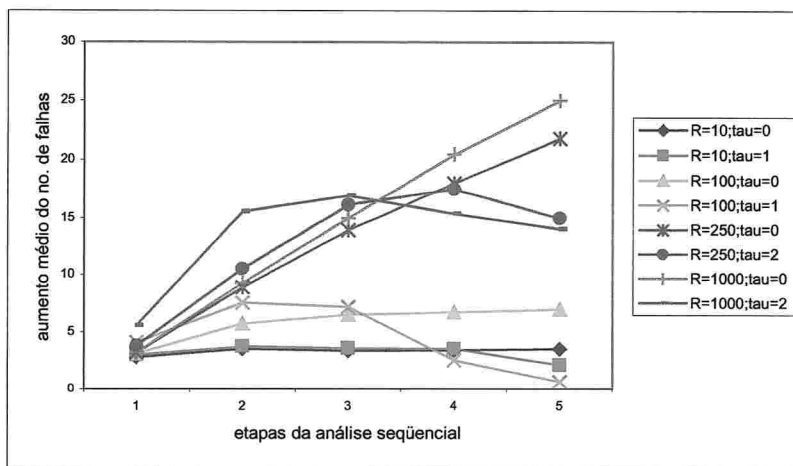


Figura D.12: Aumento médio do número de falhas por etapa com 80% de censuras aleatórias.

D.2 Número de Rejeições Encontrado em 500 Simulações, Sob H_0

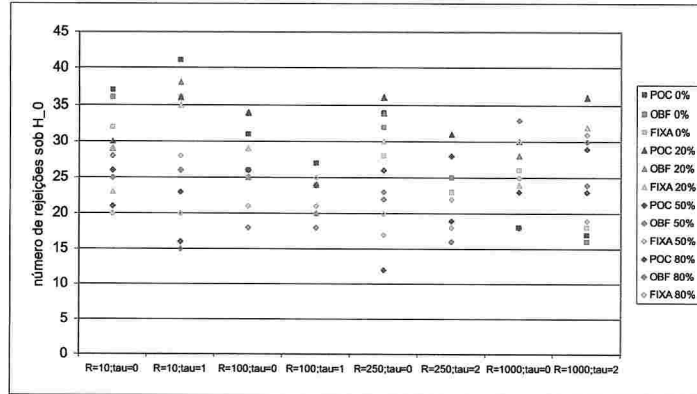


Figura D.13: Número de rejeições sob H_0 dos testes logrank bilaterais seqüenciais.

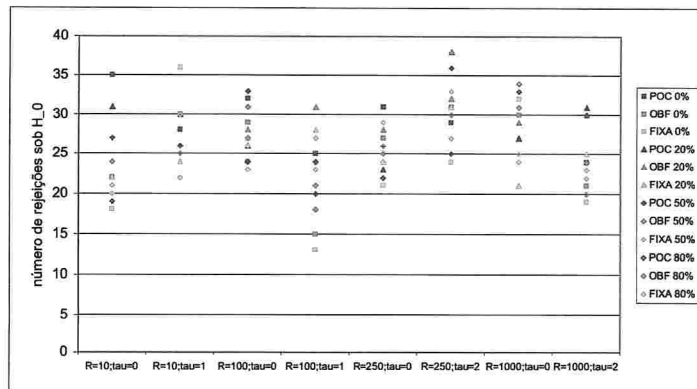


Figura D.14: Número de rejeições sob H_0 dos testes logrank unilaterais seqüenciais simétricos.

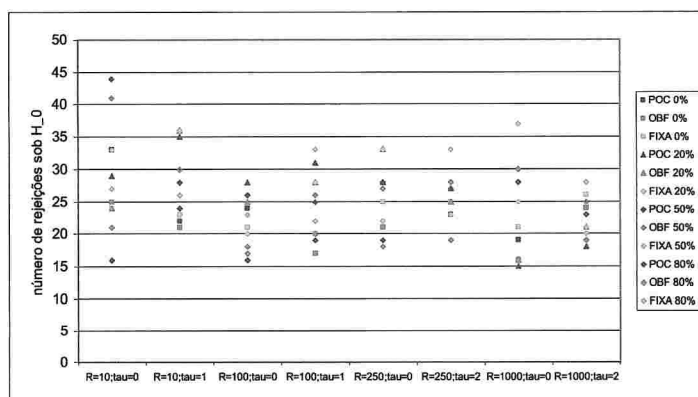


Figura D.15: Número de rejeições sob H_0 dos testes logrank unilaterais seqüenciais assimétricos.

D.3 Poder do Teste Logrank Seqüencial em Grupo

Nas tabelas que se seguem foram utilizados os seguintes códigos para definir os parâmetros de cada simulação.

- A: $\lambda_C = 1$, $\lambda_E = 0,5$, $R = 10$ e $\tau = 0$;
- B: $\lambda_C = 1$, $\lambda_E = 0,5$, $R = 10$ e $\tau = 1$;
- C: $\lambda_C = 1$, $\lambda_E = 0,5$, $R = 100$ e $\tau = 0$;
- D: $\lambda_C = 1$, $\lambda_E = 0,5$, $R = 100$ e $\tau = 1$;
- E: $\lambda_C = 0,04$, $\lambda_E = 0,02$, $R = 250$ e $\tau = 0$;
- F: $\lambda_C = 0,04$, $\lambda_E = 0,02$, $R = 250$ e $\tau = 2$;
- G: $\lambda_C = 0,04$, $\lambda_E = 0,02$, $R = 1000$ e $\tau = 0$;
- H: $\lambda_C = 0,04$, $\lambda_E = 0,02$, $R = 1000$ e $\tau = 2$.

D.3.1 Testes Bilaterais

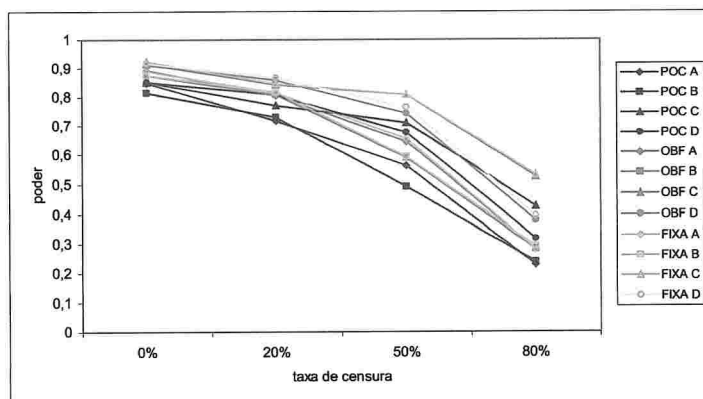


Figura D.16: Poder dos testes logrank bilaterais seqüenciais, com $\lambda_C = 1$ e $\lambda_E = 0,5$.

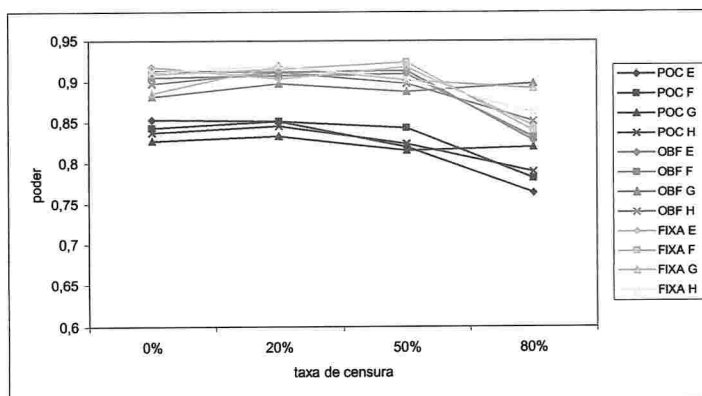


Figura D.17: Poder dos testes logrank bilaterais seqüenciais, com $\lambda_C = 0,04$ e $\lambda_E = 0,02$.

D.3.2 Testes Unilaterais Simétricos

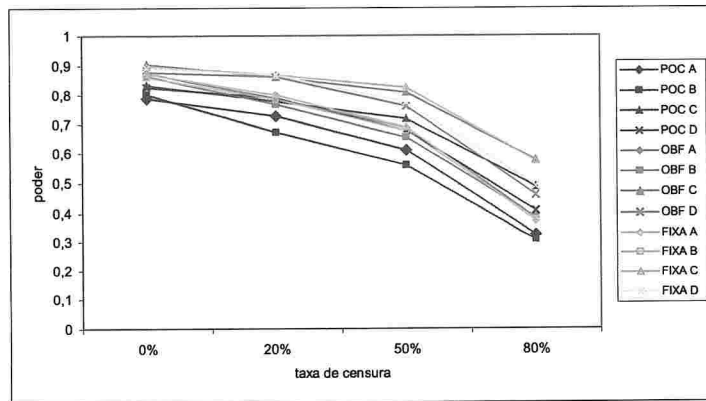


Figura D.18: Poder dos testes logrank unilaterais seqüenciais simétricos, com $\lambda_C = 1$ e $\lambda_E = 0,5$.

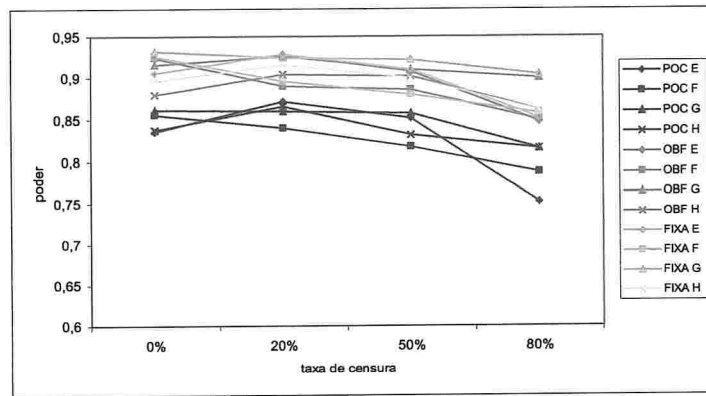


Figura D.19: Poder dos testes logrank unilaterais seqüenciais simétricos, com $\lambda_C = 0,04$ e $\lambda_E = 0,02$.

D.3.3 Testes Unilaterais Assimétricos

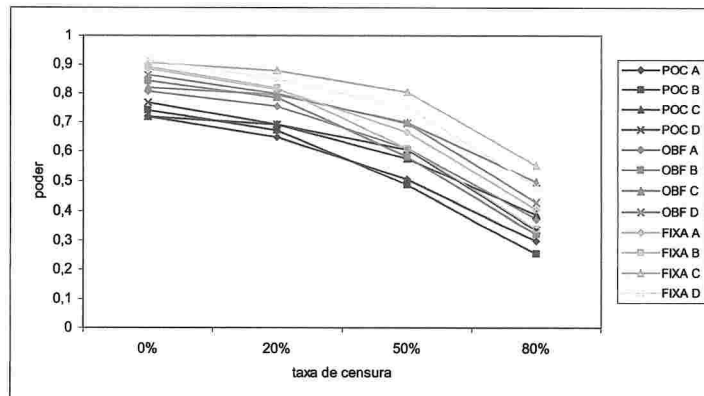


Figura D.20: Poder dos testes logrank unilaterais seqüenciais assimétricos, com $\lambda_C = 1$ e $\lambda_E = 0,5$.

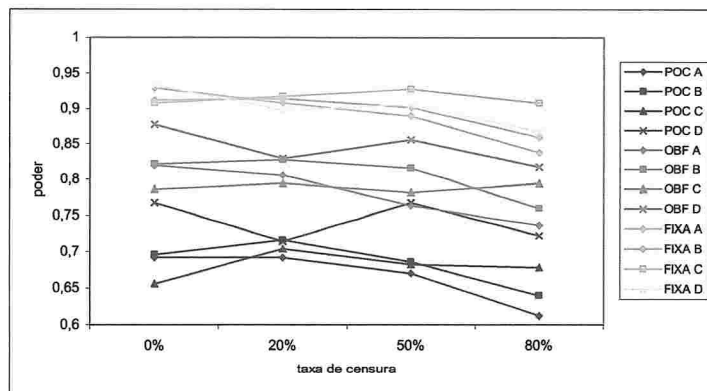


Figura D.21: Poder dos testes logrank unilaterais seqüenciais assimétricos, com $\lambda_C = 0,04$ e $\lambda_E = 0,02$.

Apêndice E

Cálculo dos Valores Críticos dos Testes Seqüenciais em Grupo

E.1 Função Densidade Condicional

Jennison & Turnbull (2000) apresentam alguns detalhes para o cálculo dos valores críticos de testes seqüenciais em grupo. Um resumo deste cálculo é apresentado para melhor entendimento da definição dos limites de rejeição e aceitação.

Seja R_i o evento “rejeitar H_0 na i -ésima etapa”. Então,

$Pr\{R_i\}$ = probabilidade de rejeitar a hipótese nula na i -ésima etapa e

$Pr\{\bar{R}_i\}$ = probabilidade de não rejeitar a hipótese nula na i -ésima etapa.

Como

$\alpha = Pr\{\text{rejeitar } H_0 \text{ na análise } k=1 \text{ ou } k=2 \text{ ou } \dots \text{ ou } k=K \mid H_0 \text{ é verdadeira}\},$

α pode ser definido da seguinte forma:

$$\begin{aligned}\alpha &= Pr\{R_1\} + Pr\{\bar{R}_1 R_2\} + Pr\{\bar{R}_1 \bar{R}_2 R_3\} + \dots + Pr\{\bar{R}_1 \bar{R}_2 \dots \bar{R}_{K-1} R_K\} \\ &= Pr\{R_1\} + Pr\{R_2 \mid \bar{R}_1\} Pr\{\bar{R}_1\} + Pr\{R_3 \mid \bar{R}_1 \bar{R}_2\} Pr\{\bar{R}_1 \bar{R}_2\} + \\ &\quad + \dots + Pr\{R_K \mid \bar{R}_1 \bar{R}_2 \dots \bar{R}_{K-1}\} Pr\{\bar{R}_1 \bar{R}_2 \dots \bar{R}_{K-1}\}.\end{aligned}$$

Definindo-se

$$\begin{aligned}\psi_k(a_1, b_1, \dots, a_k, b_k; \theta) &= Pr_{\theta=0}\{a_1 < Z_1 < b_1, \dots, a_{k-1} < Z_{k-1} < b_{k-1}, Z_k > b_k\}, \\ k &= 1, \dots, K,\end{aligned}$$

tem-se que $\psi_k(a_1, b_1, \dots, a_k, b_k; \theta)$ é a probabilidade conjunta de não rejeitar a hipótese nula até a $(k - 1)$ -ésima etapa da análise, rejeitando na etapa k . Isso significa que

$$\alpha = \psi_1 + \psi_2 + \dots + \psi_{K-1} + \psi_K.$$

Cada ψ_k pode ser definida como integrais $k - upla$ de funções condicionais de uma Normal.

Lembrando que Z_1, \dots, Z_K é uma seqüência de estatísticas padronizadas que segue a distribuição conjunta canônica (2.2). Para uma dada seqüência com níveis de informações I_1, \dots, I_K , tem-se:

1. $Z_1 \sim N(\theta\sqrt{I_1}, 1)$, ou seja, Z_1 tem a função densidade dada por

$$\begin{aligned}f(z_1; \theta) &= \Phi(z_1 - \theta\sqrt{I_1}) \text{ em que,} \\ \Phi(x) &= \frac{\exp(-x^2/2)}{\sqrt{2\pi}}.\end{aligned}$$

2. $\Delta_k = I_k - I_{k-1}$ com $k = 2, \dots, K$, em que para cada k tem-se:

$$Z_k\sqrt{I_k} - Z_{k-1}\sqrt{I_{k-1}} \sim N(\theta\Delta_k, \Delta_k),$$

independentemente de Z_1, \dots, Z_K .

Desta forma, a função densidade de Z_k , dado $Z_1 = z_1, \dots, Z_{k-1} = z_{k-1}$, depende somente de Z_{k-1} , isto é:

$$f(z_k; \theta/z_{k-1}) = \frac{\sqrt{I_k}}{\sqrt{\Delta_k}} \phi\left(\frac{z_k\sqrt{I_k} - z_{k-1}\sqrt{I_{k-1}} - \theta\Delta_k}{\sqrt{\Delta_k}}\right)$$

A probabilidade condicional $Pr\{Z_k | Z_{k-1}, \dots, Z_1\}$ é igual a $Pr\{Z_k | Z_{k-1}\}$ devido à propriedade markoviana da distribuição de Z_k , com $k = 1, \dots, K$.

Desta forma, cada ψ_k pode ser escrita como:

$$\psi_k(a_1, b_1, \dots, a_k, b_k; \theta) = \int_{a_1}^{b_1} \dots \int_{a_{k-1}}^{b_{k-1}} \int_{b_k}^{\infty} f(z_1; \theta) f(z_2; \theta | z_1) \dots f(z_k; \theta | z_{k-1}) dz_k \dots dz_1.$$

Em um teste bilateral de Pocock, por exemplo, os limites inferior e superior são iguais em módulo, isto é, $a_k = -b_k$. Desta forma, uma análise com $K = 5$ etapas e $\alpha = 0,05$, tem-se $a_k = 2,413$, assim:

$$\begin{aligned} \alpha'_1 &= P(|Z_1| \geq 2,413) = 0,0158 \\ \alpha'_2 &= P(|Z_2| \geq 2,413) = 0,0158 \\ &\vdots \\ \alpha'_5 &= P(|Z_k| \geq 2,413) = 0,0158 \end{aligned}$$

Note que neste caso o nível de significância global de 0,05 é diferente de $5 \times 0,0158$.

E.2 Método da Constante de Verossimilhança Seqüencial em Grupo

Os valores de A e B em (2.5) podem ser aproximados por $(1 - \beta)/\alpha$ e $\beta/(1 - \alpha)$, respectivamente. Estes valores foram encontrados conforme segue (Siegmund (1985), pág. 10-11).

Considere X_1, \dots, X_n i.i.d. $X_i \sim N(\mu, \sigma^2)$.

As hipóteses testadas são

$$\begin{aligned} H_0: X &\sim N(\mu_0, \sigma^2); \\ H_A: X &\sim N(\mu_1, \sigma^2). \end{aligned}$$

O limite superior A é tal que:

$$\alpha = Pr_{H_0}(l_N \geq A).$$

O limite inferior B é tal que:

$$\beta = Pr_{H_1}(l_N \leq B),$$

em que N é uma variável aleatória representando o número de observações até o término do experimento e $l_N = \frac{L(X, \theta_1)}{L(X, \theta_0)}$.

Seja \mathcal{A}_n um sub-conjunto do espaço n -dimensional em que $B < l_k(\xi_1, \dots, \xi_k) < A$, para $k = 1, \dots, n-1$ e $l_n(\xi_1, \dots, \xi_n) \geq A$. Desta forma, tem-se que ;

$$\{N = n; l_n \geq A\} = \{(x_1, \dots, x_n) \in \mathcal{A}_n\}.$$

Por definição, $\alpha = Pr\{l_N \geq A \mid H_0 \text{ é verdadeira}\}$. Assim,

$$\begin{aligned} \alpha = Pr_0\{l_N \geq A\} &= \sum_{n=1}^{\infty} Pr_0\{N = n, l_n \geq A\} \\ &= \sum_{n=1}^{\infty} \int_{\mathcal{A}_n} f_{0n} d\xi_1 \dots d\xi_n \\ &= \sum_{n=1}^{\infty} \int_{\mathcal{A}_n} \frac{f_{0n}}{f_{1n}} f_{1n} d\xi_1 \dots d\xi_n \\ &= \sum_{n=1}^{\infty} \mathbf{E}_1[l_n^{-1} / N = n; l_n \geq A] Pr_1\{l_n \geq A; N = n\} \\ &\leq \frac{1}{A} \sum_{n=1}^{\infty} Pr_1\{l_n \geq A; N = n\} \\ &= \frac{1}{A} Pr_1\{l_N \geq A\} \\ &= \frac{1}{A}(1 - \beta). \end{aligned}$$

Logo,

$$\alpha \leq \frac{1}{A}(1 - \beta),$$

levando a

$$A \leq \frac{1 - \beta}{\alpha}.$$

Usando o mesmo raciocínio pode-se obter o limite inferior B .

Seja \mathcal{B}_n , o conjunto do espaço n -dimensional em que $B < l_k(\xi_1, \dots, \xi_k) < A$, para $k = 1, \dots, n-1$ e $l_n(\xi_1, \dots, \xi_n) \leq B$. Desta forma, tem-se $\{N = n; l_n \leq B\} = \{(x_1, \dots, x_n) \in \mathcal{B}_n\}$.

$$\begin{aligned}
 \beta &= P_1\{l_N \leq B\} = \sum_{n=1}^{\infty} P_1\{N = n, l_n \leq B\} = \sum_{n=1}^{\infty} \int_{\mathcal{B}_n} f_{1n} d\xi_1 \dots d\xi_n \\
 &= \sum_{n=1}^{\infty} \int_{\mathcal{B}_n} \frac{f_{1n}}{f_{0n}} f_{0n} d\xi_1 \dots d\xi_n \\
 &= \sum_{n=1}^{\infty} \mathbf{E}_0[l_n/N = n; l_n \leq B] Pr_0\{N = n; l_n \leq B\} \\
 &\leq B \sum_{n=1}^{\infty} Pr_0\{N = n; l_n \leq B\} \\
 &= B Pr_0\{l_N \leq B\} \\
 &= B(1 - \alpha).
 \end{aligned}$$

Assim,

$$\beta \leq B(1 - \alpha).$$

Desta forma obtém-se o limite inferior B , isto é,

$$B \geq \frac{\beta}{1 - \alpha}.$$

Assim, (2.5) pode ser escrita como

$$\frac{\beta}{1 - \alpha} \leq \frac{\mathbf{L}(X; \theta_1)}{\mathbf{L}(X; \theta_0)} \leq \frac{1 - \beta}{\alpha}.$$

Referências Bibliográficas

- Armitage, P., McPherson, C. K. & Rowe, B. C. (1969). Repeated significance tests on accumulating data, *Journal of the Royal Statistical Society A* **132**: 235–244.
- Atkinson, K. E. (1993). *Elementary Numerical Analysis*, New York: Wiley.
- Bereta, E. M. P. (2002). *Testes de significância repetidos em análise de sobrevivência*. Dissertação de Mestrado, Universidade Estadual de Campinas-Imecc, Campinas, SP.
- Breslow, N. & Crowley, J. (1974). A large sample study of the life table and product limit estimates under random censorship, *Ann. Statist.* **2**: 437–453.
- Cox, D. R. (1972). Regression models and life-tables(with discussion), *Journal Royal Statist. Soc. B.* **34**: 187–220.
- Cox, D. R. (1975). Partial likelihood, *Biometrika* **62**: 269–276.
- DeMets, D. L. & Gail, M. H. (1985). Use of logrank tests and group sequential methods at fixed calendar times, *Biometrics* **41**: 1039–1044.
- DeMets, D. L. & Ware, J. H. (1980). Group sequential methods for clinical trials with one-sided hypothesis, *Biometrika* **67**: 651–660.
- DeMets, D. L. & Ware, J. H. (1982). Asymmetric group sequential boundaries for monitoring clinical trials, *Biometrika* **69**: 661–663.
- Gail, M. H., DeMets, D. L. & Slud, E. V. (1982). Simulation studies on increments of the two-sample logrank score test for survival data, with application to group sequential

- boundaries, In *Survival Analysis* . (Eds., J. Crowley and R.A. Johnson), Hayward, California: Institute of Mathematical Statistics, 287-301.
- Harrington, D. P., Fleming, T. R. & Green, S. J. (1982). Procedures for serial testing in censored survival data, In *Survival Analysis* . (Eds., J. Crowley and R.A. Johnson), Hayward, California: Institute of Mathematical Statistics, 269-286.
- Jennison, C. (1992). Bootstrap test and confidence intervals for a hazard ratio when the number of observed failures is small, with application to group sequential survival studies, In *Computing Science and Statistics* **23**: 89–97.
- Jennison, C. & Turnbull, B. W. (1984). Repetead confidence intervals for group sequential clinical trials, *Control Clinical Trials* **5**: 33–45.
- Jennison, C. & Turnbull, B. W. (2000). *Group Sequential Methods With Applications to Clinical Trials*, Chapman and Hall/CRC: New York.
- Kalbfleisch, J. D. & Prentice, R. L. (1980). *The Statistical Analysis of Failure Time Data*, New York: Wiley.
- Kim, K., Boucher, H. & Tsiatis, A. A. (1995). Design and analysis of group sequential logrank tests in maximum duration versus information trials, *Biometrics* **51**: 988–1000.
- Lan, K. K. G. & DeMets, D. L. (1983). Discrete sequential boundaries for clinical trials, *Biometrika* **70**: 659–663.
- Lan, K. K. G. & Lachin, J. M. (1990). Implementation of group sequential logrank tests in a maximum duration trial, *Biometrics* **46**: 759–770.
- Lan, K. K. G., Rosenberger, W. F. & Lachin, J. M. (1995). Sequential monitoring of survival data with the wilcoxon statistic, *Biometrics* **51**: 1175–1183.
- Lawless, J. (1982). *Statistical Models and Methods for Lifetime Data*, John Wiley & Sons: New York.

- O'Brien, P. C. & Fleming, T. R. (1979). A multiple testing procedure for clinical trials, *Biometrics* **35**: 549–556.
- Pocock, S. J. (1977). Group sequential methods in the design and analysis of clinical trials, *Biometrika* **64**: 191–199.
- Rubinstein, L. V., Gail, M. H. & Santner, T. J. (1981). Planning the duration of a comparative clinical trial with loss to follow-up and a period of continued observation, *Journal of chronic diseases* **34**: 469–479.
- Siegmund, D. (1985). *Sequential Analysis*, Springer-Verlag: New York.
- Slud, E. V. & Wei, L.-J. (1982). Two sample repeated significance tests based on the modified wilcoxon statistic, *Journal of the American Statistical Association* **77**: 862–868.
- Tarone, R. E. & Ware, J. (1977). On distribution free tests for equality of survival distributions, *Biometrika* **64**: 156–160.
- Tsiatis, A. A. (1981). The asymptotic joint distribution of the efficient scores test for the proportional hazards model calculated over time, *Biometrika* **68**: 311–315.
- Tsiatis, A. A. (1982). Repetead significance testing for a general class of statisics used in censored survival analysis, *Journal of the American Statistical Association* **77**: 855–861.
- Tsiatis, A. A., Rosner, G. L. & Tritchler, D. L. (1985). Group sequential tests with censored survival data adjusting for covariates, *Biometrika* **72**: 365–373.
- Wald, A. (1947). *Sequential Analysis*, New York: Wiley.
- Wald, A. & Wolfowitz, J. (1948). Ontimum character of the sequential probability ratio test, *Ann. Math. Statist.* **19**: 326–339.