



COPPE/UFRJ

CONSIDERAÇÕES SOBRE O EFEITO DE FATORES DE PROGNÓSTICO NA
ALOCÇÃO EM ENSAIOS CLÍNICOS E SOBRE O TESTE DE VUONG

James Dean Oliveira dos Santos Junior

Tese de Doutorado apresentada ao Programa de Pós-graduação em Engenharia de Produção, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Doutor em Engenharia de Produção.

Orientador: Basílio de Bragança Pereira

Rio de Janeiro
Dezembro de 2010

CONSIDERAÇÕES SOBRE O EFEITO DE FATORES DE PROGNÓSTICO NA
ALOCÇÃO EM ENSAIOS CLÍNICOS E SOBRE O TESTE DE VUONG

James Dean Oliveira dos Santos Junior

TESE SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ
COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA (COPPE)
DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS
REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR
EM CIÊNCIAS EM ENGENHARIA DE PRODUÇÃO.

Aprovada por:

Prof. Basílio de Bragança Pereira, Ph.D

Prof. Antônio Carlos M. Ponce de Leon, Ph.D

Prof. Dani Gamerman, Ph.D.

Profa. Laura Letícia Ramos Rifo, D.Sc.

Prof. Ronir Raggio Luiz, D.Sc.

RIO DE JANEIRO, RJ – BRASIL

DEZEMBRO DE 2010

Santos Junior, James Dean Oliveira dos

Considerações sobre o Efeito de Fatores de Prognóstico na Alocação em Ensaios Clínicos e sobre o Teste de Vuong/James Dean Oliveira dos Santos Junior. – Rio de Janeiro: UFRJ/COPPE, 2010.

XII, 83 p.: il.; 29,7cm.

Orientador: Basílio de Bragança Pereira

Tese (doutorado) – UFRJ/COPPE/Programa de Engenharia de Produção, 2010.

Referências Bibliográficas: p. 81–83.

1. Alocação Sequencial. 2. Teste de hipótese para famílias separadas. I. Pereira, Basílio de Bragança. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia de Produção. III. Título.

*À Diana, minha estatística
suficiente e completa.*

Agradecimentos

Agradeço à minha família, que com grande amor e compreensão suportou minhas freqüentes ausências.

Ao Basílio, grande professor, orientador e amigo, que soube ouvir e compartilhar idéias.

Aos meus amigos da COOPE (em ordem alfabética): Aroldo, Emília, Joelma, Laura, Leonardo e Luciana. Meus sinceros agradecimentos por todos os momentos que tivemos.

Aos meus professores da UFAM, UNICAMP e UFRJ, dos quais sempre consegui apoio e exemplos a seguir.

Aos membros da banca examinadora pelos valiosos comentários.

À secretaria da COOPE, em especial à Andréia e ao Pedrinho, pelo auxílio em vários momentos.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico dos Estados Unidos do Brasil (CNPq) pelo suporte financeiro.

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Doutor em Ciências (D.Sc.)

CONSIDERAÇÕES SOBRE O EFEITO DE FATORES DE PROGNÓSTICO NA ALOCAÇÃO EM ENSAIOS CLÍNICOS E SOBRE O TESTE DE VUONG

James Dean Oliveira dos Santos Junior

Dezembro/2010

Orientador: Basílio de Bragança Pereira

Programa: Engenharia de Produção

Esta tese apresenta dois temas. O primeiro trata sobre o efeito dos fatores de prognóstico na alocação de pacientes em um ensaio clínico do tipo caso/controle. Mostramos que desenhos para ensaios clínicos que não utilizam fatores de prognóstico para realizar a alocação, mas que privilegiam o balanço entre os grupos de tratamentos, produzem resultados próximos do ótimo com o aumento do tamanho da amostra. Mostramos também como obter a alocação ótima, com a restrição de balanço, para uma população de pacientes conhecida. O segundo tema trata sobre o desempenho do teste de Vuong e comparações deste com o teste de Cox para discriminar entre modelos pertencentes a famílias de distribuições separadas. Mostramos que o teste de Vuong possui algumas desvantagens dependendo do verdadeiro modelo gerador dos dados. Em particular existem casos em que, mesmo com uma amostra acima de 200, o teste tende a falhar em discriminar entre os dois modelos. Comparações com o teste de Cox mostram que este é mais estável para os modelos testados aqui e sugestões de como se utilizar o teste de Vuong são apresentadas.

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Doctor of Science (D.Sc.)

James Dean Oliveira dos Santos Junior

December/2010

Advisor: Basílio de Bragança Pereira

Department: Production Engineering

This thesis presents two themes. The first deals with the effect of prognostic factors in the sequential allocation of patients in a clinical trial with two groups (case/control). We showed that designs for clinical trials that do not use prognostic factors to accomplish the allocation, but force the balance between treatment groups produce results close to the optimum, with the increasing of the sample size. We also show how to obtain the optimal allocation, with the restriction of balance, for a known population. The second theme deals with the performance of the Vuong's test and comparisons of this with the Cox's test to discriminate between models belonging to separate families of distributions. We showed that the Vuong's test has some disadvantages depending on the true model. In particular, there are cases where even with a large sample size, the Vuong's test tends to fail to discriminate between the two models. Comparisons with the Cox's test showed that it is more stable for the models tested here and suggestions on how to use the Vuong's test are presented.

Sumário

Lista de Figuras	xi
Lista de Tabelas	xii
1 Organização	1
I Sobre o efeito dos fatores de prognóstico na alocação de pacientes em ensaios clínicos quando utilizamos desenhos sem informação das covariáveis	3
2 Introdução	4
2.1 Ensaios clínicos e desenhos de ensaios	4
2.2 Objetivos e organização	6
3 Revisão bibliográfica	8
3.1 Introdução	8
3.2 Medidas de desempenho para desenhos com alocação sequencial	8
3.2.1 Viés de seleção	9
3.2.2 Balanço entre os grupos de tratamentos	10
3.2.3 A eficiência de um desenho	12
3.3 Alguns desenhos para ensaios clínicos	14
3.3.1 Desenhos sistemáticos	15
3.3.2 Desenho aleatório	15
3.3.3 Moeda Viesada	17
4 Resultados	20
4.1 Reescrevendo $\mathbb{V}(\hat{\alpha}_1 - \hat{\alpha}_2)$	21

4.2	Sobre o comportamento assintótico de $\mathbb{V}(\hat{\alpha}_1 - \hat{\alpha}_2)$	22
4.3	Alocação ótima quando a matriz de fatores de prognóstico é conhecida	28
4.4	Simulação	33
4.4.1	Evitando problemas computacionais	33
4.4.2	Resultados de Simulação	34
5	Considerações Finais	40
II	Considerações sobre o teste de Vuong e seu desempenho contra o teste de Cox para discriminar modelos pertencentes à famílias separadas	42
6	Introdução	43
6.1	Um breve histórico sobre testes de hipóteses para discriminar modelos pertencentes a famílias separadas	43
6.2	Objetivos e organização	44
7	Os Testes de Cox e Vuong	46
7.1	A base dos testes de Cox e Vuong	46
7.2	O Teste de Cox	48
7.3	O Teste de Vuong	52
8	Resultados	54
8.1	Notas sobre o teste de Vuong	54
8.1.1	O Caso Lognormal versus Exponencial	56
8.1.2	O Caso Lognormal versus Weibull	58
8.2	Um cenário para simular comparações entre os testes	61
8.3	Simulações	63
9	Considerações Finais	68
III	Apêndice e referências bibliográficas	69
A	Demonstrações	70
A.1	Lognormal	70

A.2	Exponencial	72
A.3	Weibull	75
A.4	O Teste Lognormal versus Exponencial	77
A.5	O Teste Lognormal versus Weibull	79

Lista de Figuras

4.1	Valores de $f_i(n, \Delta)/f^*$ em um estudo de simulação	39
8.1	Gráfico de $1/CV(f, g)$ para o teste H_f : lognormal versus H_g : exponencial, quando a H_f é verdadeira.	57
8.2	Proporção de aceitação de H_f para as hipóteses H_f : lognormal e H_g : exponencial. H_f é o verdadeiro modelo.	58
8.3	Simulações do teste de Vuong para as hipóteses H_f : lognormal e H_g : exponencial, onde H_g é verdadeira. A proporção de aceitação de H_0 , H_f e H_g são, respectivamente, p_0 , p_F e p_G	59
8.4	Resultados de simulação do teste de Vuong para as hipóteses H_f : Lognormal e H_g : Weibull, onde H_f é verdadeira. A proporção de aceitação de H_0 , H_f e H_g foram, respectivamente, p_0 , p_F e p_G	60
8.5	Nível e poder estimados para o teste lognormal versus exponencial, com hipótese nula sendo H_f : lognormal. Linha cheia: teste de Cox; Linha pontilhada: teste de Vuong.	64
8.6	Nível e poder estimados para o teste exponencial versus lognormal, com hipótese nula sendo H_f : exponencial. Linha cheia: teste de Cox; Linha pontilhada: teste de Vuong.	65
8.7	Nível e poder estimados para o teste lognormal versus Weibull, com hipótese nula sendo H_f : lognormal. Linha cheia: teste de Cox; Linha pontilhada: teste de Vuong.	66
8.8	Nível e poder estimados para o teste lognormal versus Weibull, com hipótese nula sendo H_f : Weibull. Linha cheia: teste de Cox; Linha pontilhada: teste de Vuong.	67

Lista de Tabelas

3.1	Probabilidade de balanço para diferentes valores de n , onde a coluna Prob. equivale a $\mathbb{P}(D_1 = 0 \Delta_a)$ se n for par e $\mathbb{P}(D_1 = 1 \Delta_a)$ em caso contrário.	16
3.2	Probabilidade de balanço para o desenho da moeda viesada para diferentes valores de n . A coluna Prob. dá a probabilidade de balanço estimada.	18
3.3	Viés de seleção estimado para o desenho da moeda viesada considerando alguns valores de n	19
4.1	Cenários de simulação	35
4.2	Valores de \bar{v} para o cenário 1.	36
4.3	Valores de \bar{v} para o cenário 2.	36
4.4	Valores de \bar{v} para o cenário 3.	36
4.5	Valores de \bar{v} para o cenário 4.	37
8.1	Estatísticas para o teste de Cox. LN: distribuição lognormal; Exp: distribuição exponencial; Wei: distribuição Weibull.	62
8.2	Distribuições utilizadas na geração das amostras	63
A.1	Alguns valores para a função $\phi^{(k)}(x)$	75

Capítulo 1

Organização

A presente tese foi organizada levando em consideração os trabalhos desenvolvidos pelo autor, sugeridos e orientados pelo professor Basílio de Bragança Pereira. Alguns tópicos foram desenvolvidos ao longo de três anos e, após ponderarmos sobre as sugestões da banca de qualificação, optamos por desenvolver dois temas, apresentados nesta em duas partes. Ambas as partes podem ser lidas em qualquer ordem, uma vez que se tratam de trabalhos distintos.

O primeiro tema refere-se ao peso de fatores de prognóstico na alocação de pacientes em um ensaio clínico com dois tratamentos. Este estudo surgiu inicialmente por outro caminho: o orientador sugeriu um método de alocação de pacientes utilizando a distância de Mahalanobis entre os fatores de prognóstico do novo paciente e os dois grupos de pacientes já existentes. Após analisarmos os resultados, percebemos que este método e outros métodos que não utilizavam a informação dos fatores de prognóstico eram semelhantes. Disto, conjecturamos que os fatores de prognóstico deveriam ter influência limitada na alocação e a partir disto desenvolvemos nosso trabalho.

O segundo tema refere-se ao desempenho do teste de Vuong para discriminar entre modelos pertencentes a famílias de distribuições separadas e seu desempenho contra o teste de Cox para o mesmo fim. O teste de Vuong possui uma vantagem computacional em relação ao teste de Cox que o torna bem mais atraente. Embora estes testes existam desde a década de oitenta, ainda não havia sido realizada uma comparação entre os dois. Um possível motivo para isso pode ser a natureza dos testes: ambos se valem de hipóteses distintas para realizar seus testes, e simulações

se tornam complicadas de avaliar. Notamos que o teste de Vuong possui características que podem torná-lo uma escolha ruim. Este fato foi explorado e discutido, e simulações foram realizadas para ilustrá-lo e para comparar seu desempenho contra o teste de Cox.

Parte I

Sobre o efeito dos fatores de prognóstico na alocação de pacientes em ensaios clínicos quando utilizamos desenhos sem informação das covariáveis

Capítulo 2

Introdução

2.1 Ensaios clínicos e desenhos de ensaios

Um ensaio clínico destina-se a comparar o efeito de um ou mais tratamentos médicos sobre uma variável de interesse (em geral, uma condição do paciente). Para tanto, pacientes são alocados em grupos, onde cada grupo recebe um tratamento e alguma técnica inferencial é utilizada para detectar se existem diferenças nas respostas de cada tratamento.

O ensaio deve ser planejado de modo a evitar vieses que possam gerar conclusões equivocadas. Assim, tornam-se necessários estudos sobre como fatores externos podem influenciar os resultados de um ensaio e como estes devem ser planejados para evitar tal influência. O planejamento de um ensaio é denominado *desenho do ensaio*. Diferentes tipos de desenhos foram introduzidos ao longo das últimas décadas, como o “Sanduíche” de Student, Student (1937), e ainda são sugeridos, como o desenho apresentado em Fossaluzza et al. (2009).

Em muitas situações, os pacientes devem ser tratados o mais rápido possível. Neste caso, assim que o paciente aceita participar do ensaio, ele deverá ser alocado imediatamente em um dos grupos. Esse tipo de situação é denominada *alocação sequencial*. Doravante, focaremos nossa discussão apenas nesse tipo de situação. Nestes casos, deseja-se que o desenho planejado:

- possua grupos balanceados: o número de pacientes em cada grupo deve ser próximo (de preferência igual).

- não produza vieses: a alocação deve ser justa, de modo a não favorecer qualquer um dos grupos.

Diferentes desenhos tem sido propostos para lidar com essas questões. O desenho aleatório, difundido por Fisher (por exemplo, ver Fisher (1956)) é o mais popular. Este desenho visa minimizar o efeito de um agente externo, dando ao paciente uma chance igual de pertencer a qualquer um dos grupos de tratamentos. Entretanto, este desenho é criticado por gerar falta de balanço entre os grupos, especialmente para pequenas amostras. Desenhos que ainda possuem um mecanismo aleatório de alocação mas que forçam o balanço entre os grupos de tratamentos foram introduzidos em Efron (1971) e outros métodos surgiram desde então. O resumo de alguns métodos pode ser visto em Atkinson (2002) e em Antognini and Giovagnoli (2004). Um estudo relatando os métodos mais utilizados na prática pode ser visto em Scott et al. (2002).

Sabe-se que certas condições em que o paciente se encontra podem influenciar o resultado do tratamento aplicado. Estas condições são denominadas fatores de prognóstico. Outra categoria de desenhos foi introduzida em Pocock and Simon (1975), onde o interesse estava em balancear também os fatores de prognóstico em cada grupo. Sua motivação é simples: para evitar vieses, todos os grupos devem possuir pacientes “semelhantes”. Este método é denominado minimização e foi proposto concomitantemente em Taves (1974). A maior crítica deste método é a falta de capacidade de lidar com muitos fatores de prognóstico.

Percebendo que em geral os ensaios clínicos utilizam modelos lineares para analisar as respostas dos tratamentos, Atkinson (1970) introduziu um desenho baseado na variância do melhor preditor linear para a variável resposta do próximo paciente e na variância dos contrastes dos tratamentos.

Após um estudo bibliográfico, percebemos que a maneira de lidar com os fatores de prognóstico na alocação dos pacientes ainda é um tópico bastante discutido. Na prática, desenhos como o aleatório e a moeda viesada são preferíveis a desenhos que utilizam a informação corrente dos fatores de prognóstico. Entretanto, sob o ponto de vista da inferência frequentista, devemos procurar pela alocação que minimize o erro quadrático médio do estimador do efeito dos tratamentos. Sob este critério, qual é o efeito de desenhos que não utilizam qualquer informação sobre os fatores

de prognóstico? Neste trabalho realizamos um estudo sobre esse efeito no caso especial onde existem apenas dois tratamentos a serem avaliados (um ensaio do tipo caso/controle).

2.2 Objetivos e organização

Seja Y_i a resposta do i -ésimo paciente no tratamento ao qual ele foi submetido. Consideremos o seguinte modelo linear:

$$\mathbb{E}(Y_i) = \alpha_1 I(i \in T_1) + \alpha_2 I(i \in T_2) + \mathbf{x}_i^T \beta \quad (2.1)$$

onde $I(\cdot)$ é a função indicadora, T_j é o conjunto de pacientes que receberam o tratamento j ($j = 1, 2$), \mathbf{x}_i é o vetor de fatores de prognóstico para o i -ésimo paciente, α_j é o efeito do j -ésimo tratamento e β é o vetor de parâmetros dos fatores de prognóstico. Consideremos as variáveis Y 's não-correlacionadas.

Considerando o modelo na Equação (2.1), um modo de medir se existe diferença entre os tratamentos é a estimação do contraste $\alpha_1 - \alpha_2$. Em Atkinson (2002) o autor nota, em um estudo de simulação, que as variâncias do contraste de diversos desenhos se aproximam entre si com o aumento do tamanho da amostra. Entretanto, não foi apresentada uma demonstração analítica para o fato, e não encontramos nenhuma referência neste sentido em nosso levantamento bibliográfico.

Essa convergência possui uma implicação interessante: se para um número moderado de pacientes as variâncias dos contrastes geradas pelos desenhos Δ_1 e Δ_2 forem muito próximas, poderemos optar pelo desenho com maior facilidade de implementação, uma vez que as eficiências dos estimadores estarão próximas.

Outra implicação mais sutil se refere aos fatores de prognóstico. Sabemos que desenhos diferentes proporcionam uma distribuição diferente dos fatores de prognóstico nos grupos envolvidos. Se as alocações geradas pelos desenhos Δ_1 e Δ_2 geram variâncias semelhantes, então os fatores de prognósticos devem possuir influência limitada na variância e questões como obter um balanço destes fatores dentro de cada grupo se tornam menos relevantes que questões como minimização de viés de seleção e balanço entre tratamentos (essas questões serão apresentadas e discutidas nas Seções 3.2.1 e 3.2.2).

Levantamos as seguintes questões:

1. É possível caracterizar o efeito dos fatores de prognóstico na variância dos contrastes?
2. Seja n é o tamanho da amostra. Existe $f(n)$ tal que $\mathbb{V}(\hat{\alpha}_1 - \hat{\alpha}_2)/f(n) \rightarrow 1$ quando $n \rightarrow \infty$? Se existir, se não levarmos em consideração os fatores de prognóstico para a alocação, quais são as condições que o desenho deve possuir para que tal convergência seja possível?

Este trabalho responde as duas questões dadas acima, focando apenas o caso em que existem dois tratamentos a serem avaliados (estudos do tipo caso/controle). No Capítulo 3, apresentamos uma revisão bibliográfica sobre os principais métodos de avaliação de desenhos de ensaios clínicos e três desenhos que não utilizam fatores de prognóstico para realizar a alocação dos pacientes. No Capítulo 4, apresentamos soluções para as questões propostas, mostrando que:

1. A variância dos contrastes pode ser decomposta em duas partes, onde a primeira é a variância dos contrastes sem o efeito dos fatores de prognóstico (variância de uma ANOVA com um critério de classificação), e a segunda parte é uma função que depende dos fatores de prognóstico e do desenho adotado.
2. Existem condições brandas para garantir que a segunda parte da variância citada acima seja limitada. Essas condições são compartilhadas pelos desenhos enunciados no Capítulo 3.
3. Supondo a existência de uma lista com todos os fatores de prognóstico dos pacientes antes de começar o ensaio clínico, é possível encontrar a melhor alocação dos pacientes via um problema de programação inteira. Com isto, podemos avaliar a eficiência média de qualquer desenho aplicado na alocação destes pacientes.

Capítulo 3

Revisão bibliográfica

3.1 Introdução

Neste Capítulo, apresentamos as medidas utilizadas na literatura para avaliar um desenho de um ensaio clínico com alocação sequencial. Em seguida, apresentamos três desenhos para alocação sequencial que não utilizam fatores de prognóstico. Estes desenhos serão utilizados no Capítulo 4.

Na Seção 3.2 apresentamos as três medidas para avaliação de um desenho mais empregadas na literatura: o viés de seleção, o balanço entre os grupos de tratamentos e a eficiência de Atkinson. Na Seção 3.3, apresentamos os desenhos aleatório, sistemático e o da moeda viesada, ambos com suas motivações e características.

Assumiremos que existem apenas dois tratamentos, sendo n_1 e n_2 o número de pacientes em cada um deles. Denotaremos o total de pacientes no estudo por $n = n_1 + n_2$. Quando necessário, identificaremos os dois tratamentos por T_1 e T_2 . Utilizaremos apenas a letra T se o índice do tratamento não for relevante.

3.2 Medidas de desempenho para desenhos com alocação sequencial

Uma vez que existem diversas formas de se alocar pacientes em dois grupos, devemos ter medidas para avaliar estes modos de alocação.

Basicamente, a literatura tem avaliado um desenho sob seguintes aspectos:

1. Capacidade de gerar grupos balanceados: gerar grupos balanceados significa ter um número equivalente de pacientes recebendo cada um dos tratamentos do ensaio.
2. Capacidade de minimizar o viés de seleção: se refere a minimização da vulnerabilidade do desenho à decisão sobre a alocação dos pacientes nos tratamentos.
3. Utilizar de modo eficaz a informação dos fatores de prognóstico.

Nesta seção apresentamos as medidas convencionais utilizadas para avaliar estes três aspectos.

3.2.1 Viés de seleção

Um ensaio com o objetivo de comparar dois ou mais tratamentos pode produzir resultados viesados se a alocação dos indivíduos nos diferentes grupos não foi realizada de forma imparcial, independente de ser intencional ou não.

No geral, o viés provocado pela alocação dos pacientes é denominado viés de seleção. Este termo foi introduzido em Blackwell and Hodges Jr. (1957).

Um modo de medir o viés de seleção de um desenho é mensurar com que facilidade alguém pode prever aonde o próximo paciente será alocado. Sob este raciocínio, considere um desenho Δ e dois sujeitos, A e B . Ambos os sujeitos conhecem o desenho Δ e o sujeito A fará a alocação dos pacientes segundo este desenho. O sujeito B tentará adivinhar aonde o sujeito A irá alocar o próximo paciente. Seja G a proporção das vezes em que o sujeito B acertou o grupo no qual o paciente foi alocado. Então, uma medida para o viés de seleção do desenho Δ , representada por $SB(\Delta)$, será dada por

$$SB(\Delta) = \mathbb{E}(G). \quad (3.1)$$

A interpretação para essa medida é bastante intuitiva: se o valor de $SB(\Delta)$ for alto, então o desenho é de fácil predição, o que o deixa vulnerável ao viés de seleção.

Seja $p^*(j)$ a probabilidade do sujeito B acertar a j -ésima alocação. Podemos mostrar que $SB(\Delta)$ está limitada inferiormente por $1/2$.

Proposição 3.1. *Seja $SB(\Delta)$ o viés de seleção de um desenho Δ qualquer. Então*

$$\frac{1}{2} \leq SB(\Delta) \leq 1. \quad (3.2)$$

Demonstração. Como G é uma proporção, teremos que $\mathbb{E}(G) \leq 1$. Seja $p_T(j)$ a probabilidade de que o j -ésimo paciente seja alocado no tratamento T . Como o sujeito B sempre escolherá o tratamento com maior probabilidade de receber o próximo paciente, teremos que $p^*(j) = \max\{p_T(j), 1 - p_T(j)\} \geq 1/2$. Assim,

$$SB(\Delta) = \mathbb{E}(G) = \frac{1}{n} \sum_{j=1}^n p^*(j) \geq \frac{1}{2}. \quad (3.3)$$

□

Notemos que, para todo $j = 1, \dots, n$:

- se $p^*(j) = 1/2$, então teremos que todos os pacientes tem a mesma chance de receber qualquer um dos dois tratamentos e $SB(\Delta) = 1/2$;
- se $p^*(j) = 1$, então teremos que a sequência de alocação já estava pré-determinada e $SB(\Delta) = 1$.

Assim, o viés de seleção é uma medida cujos extremos são o desenho aleatório e os desenhos nos quais a sequência das alocações é pré-determinada (denominados desenhos sistemáticos, apresentados na Seção 3.3).

Uma alternativa (não adotada neste trabalho) representa o viés de seleção por $2SB(\Delta) - 1$, sendo que um resultado igual a zero representa um desenho aleatório e um resultado igual a um representa um desenho sistemático.

3.2.2 Balanço entre os grupos de tratamentos

Dizemos que um ensaio está balanceado se o número de pacientes em cada grupo é igual. Quando o número de pacientes alocados nos tratamentos difere, dizemos que existe uma falta de balanço entre os grupos de tratamento.

Intuitivamente, é desejável que o número de pacientes em cada grupo de tratamento seja próximo pois, o aumento do tamanho da amostra tende a diminuir o erro padrão dos estimadores utilizados e o efeito de um tratamento com poucos pacientes poderia ser subestimado.

Pode-se mostrar que o balanço entre os grupos de tratamento possui boas características. Por exemplo, consideremos um experimento, sem fatores de prognóstico, no qual a resposta do i -ésimo paciente pode ser escrita como

$$Y_i = \alpha_1 I(i \in T_1) + \alpha_2 I(i \in T_2) + \varepsilon_i, \quad (3.4)$$

onde os pacientes são independentes e identicamente distribuídos com $\mathbb{E}(\varepsilon_i) = 0$ e $\mathbb{V}(\varepsilon_i) = \sigma^2$. Utilizando o Método dos Mínimos Quadrados, podemos mostrar que o contraste $\alpha_1 - \alpha_2$ é estimado por

$$\sum_{i \in T_1} \frac{Y_i}{n_1} - \sum_{i \in T_2} \frac{Y_i}{n_2}, \quad (3.5)$$

A variância do contraste é dada por

$$\mathbb{V} \left(\sum_{i \in I_1} \frac{Y_i}{n_1} - \sum_{i \in I_2} \frac{Y_i}{n_2} \right) = \sigma^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right). \quad (3.6)$$

A variância em (3.6) depende de n_1 e n_2 . A seguinte proposição mostra que a melhor estratégia para minimizar esta variância é balancear os tratamentos.

Proposição 3.2. *Se $n_1 < n_2$, então alocar o próximo paciente no tratamento T_1 minimiza a variância em (3.6).*

Demonstração. Temos duas opções: alocar o próximo paciente no tratamento T_1 ou no tratamento T_2 . Como $n_1 < n_2$:

- Se alocarmos o paciente no grupo T_1 ,

$$\mathbb{V}(\hat{\alpha}_1 - \hat{\alpha}_2) = \sigma^2 \left(\frac{1}{n_1 + 1} + \frac{1}{n_2} \right) := \text{diff}_1.$$

- Se alocarmos o paciente no grupo T_2 ,

$$\mathbb{V}(\hat{\alpha}_1 - \hat{\alpha}_2) = \sigma^2 \left(\frac{1}{n_1} + \frac{1}{n_2 + 1} \right) := \text{diff}_2.$$

Agora, basta notarmos que

$$\begin{aligned} \frac{\text{diff}_1}{\text{diff}_2} &= \frac{n_1 + n_2 + 1}{(n_1 + 1)n_2} \times \frac{(n_2 + 1)n_1}{n_1 + n_2 + 1} \\ &= \left(1 + \frac{1}{n_2} \right) \frac{n_1}{n_1 + 1}, \text{ mas como } n_2 > n_1 \\ &< 1. \end{aligned} \quad (3.7)$$

Portanto, teremos que a opção que minimiza a variância do contraste é alocar o indivíduo no tratamento com o menor número de pacientes. \square

Entretanto, em um ensaio clínico com alocação sequencial não é possível obter um desenho balanceado sem aumentar o viés de seleção. Portanto, a falta de balanço se torna algo comum em um desenho e, por isso, existe a necessidade de medi-la. Existem algumas medidas na literatura como:

- $D_1 = |n_1 - n_2|$.
- $D_2 = |n_1 - n_2|/n$.

Se ambas as medidas forem iguais a zero, teremos que o ensaio clínico estará balanceado. Em caso contrário, haverá falta de balanço.

A diferença entre as medidas apresentadas acima é que D_1 não leva em consideração o tamanho da amostra. Por exemplo, considere dois grupos, um com 10 pacientes e outro com 2. Teremos $D_1 = 8$ e $D_2 = 0,8$. Agora, considere dois grupos, um com 100 pacientes e outro com 92. Neste caso, também teremos $D_1 = 8$, mas $D_2 = 0,08$, indicando que a falta de balanço no primeiro caso possui maior gravidade que no segundo caso.

Uma medida de qualidade do desenho é a probabilidade de balanço, na entrada do n -ésimo paciente, entre os tratamentos, dada por $\mathbb{P}(D_1(n) = 0|\Delta)$ se n for par e $\mathbb{P}(D_1(n) = 1|\Delta)$ se n for ímpar.

3.2.3 A eficiência de um desenho

Em ensaios clínicos, é comum termos o conhecimento de possíveis fatores de prognóstico que podem influenciar o tratamento. Alguns desenhos lidam com a alocação levando em consideração estes fatores, como Pocock and Simon (1975), Taves (1974), Atkinson (1982) e Fossaluza et al. (2009). Tais propostas foram dadas seguindo duas motivações base: balanço do número de pacientes em cada grupo com determinado fator de prognóstico e minimização da variância do estimador da diferença entre os tratamentos. Em Atkinson (2002), o autor introduz uma metodologia para avaliar a eficácia da alocação gerada por um desenho considerando os fatores de prognóstico envolvidos. Discutimos aqui essa abordagem em detalhes.

Em ensaios clínicos, é comum estimarmos a diferença entre as médias dos tratamentos via um modelo linear. Portanto, consideremos o seguinte modelo

$$\mathbb{E}(Y_i) = \alpha_1 I(i \in T_1) + \alpha_2 I(i \in T_2) + \mathbf{x}_i^T \beta, \quad (3.8)$$

onde α_j é o efeito do j -ésimo tratamento ($j = 1, 2$), \mathbf{x}_i^T é o vetor de fatores de prognóstico associados ao i -ésimo paciente e β é um vetor (de dimensão d) de

parâmetros associados ao efeito dos fatores de prognóstico. Podemos reescrever o modelo na Equação (3.8) como segue:

$$\mathbb{E}(\mathbf{y}) = \mathbf{H}\alpha + \mathbf{X}\beta = \mathbf{G}\theta, \quad (3.9)$$

onde $\theta = (\alpha, \beta)$, $\mathbf{y} = \{Y_1, \dots, Y_n\}$, $\mathbf{G} = (\mathbf{H}, \mathbf{X})$ e \mathbf{X} é a matriz de fatores de prognóstico. A matriz \mathbf{H} é uma matriz de variáveis indicadoras onde, para a j -ésima linha, a primeira coluna vale 1 se o j -ésimo paciente for alocado no tratamento 1 (e 0 em caso contrário) e a segunda coluna vale 1 se o j -ésimo paciente for alocado no tratamento 2 (e 0 em caso contrário). O estimador de mínimos quadrados para θ será

$$\hat{\theta} = (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \mathbf{y}. \quad (3.10)$$

Como discutido, estamos interessados em detectar se existem diferenças entre os tratamentos. Consideremos o contraste $\hat{\alpha}_1 - \hat{\alpha}_2$ (na forma matricial, escrito $\mathbf{A}^T \hat{\theta}$ onde $\mathbf{A}^T = (\mathbf{L}^T, \mathbf{0}_d)$, com $\mathbf{L}^T = (1, -1)$). Teremos

$$\mathbb{E}(\mathbf{A}^T \hat{\theta}) = \alpha_1 - \alpha_2. \quad (3.11)$$

Notemos que o valor esperado do contraste não depende da alocação nem das covariáveis. Nesse sentido, qualquer desenho produz um estimador não viesado para o contraste e, portanto, teremos que o melhor desenho será aquele que minimizar a variância do constraste. Notemos que

$$\begin{aligned} \mathbb{V}(\hat{\alpha}_1 - \hat{\alpha}_2) &= \mathbb{V}(\mathbf{A}^T \hat{\theta}) = \mathbf{A}^T \mathbb{V}(\hat{\theta}) \mathbf{A} = \sigma^2 \mathbf{A}^T (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{A} \\ &= \sigma^2 \mathbf{A}^T \begin{pmatrix} \mathbf{H}^T \mathbf{H} & \mathbf{H}^T \mathbf{X} \\ \mathbf{X}^T \mathbf{H} & \mathbf{X}^T \mathbf{X} \end{pmatrix}^{-1} \mathbf{A} \\ &= \sigma^2 \mathbf{L}^T (\mathbf{H}^T \mathbf{H} - \mathbf{H}^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{H})^{-1} \mathbf{L}. \end{aligned} \quad (3.12)$$

A variância acima deixa claro que sua minimização depende da alocação dos pacientes em conjunto com a informação dos fatores de prognósticos. Consideremos o seguinte lema:

Lema 3.3. *Temos que*

$$\mathbb{V}(\hat{\alpha}_1 - \hat{\alpha}_2) \geq \mathbf{L}^T (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{L} \quad (3.13)$$

onde a igualdade ocorre se $\mathbf{H}^T \mathbf{X} = \mathbf{0}$.

Com isto, temos que a menor variância é atingida quando não são utilizados fatores de prognóstico. É claro que estes não podem ser descartados, pois prejudicaria o ajuste do modelo. Entretanto, este resultado é interessante pois serve como cota inferior para a variância do contraste. Isto sugere a seguinte definição.

Definição 3.4 (Eficiência de Atkinson). *A eficiência do desenho será definida por*

$$\varepsilon_n = \frac{4/n}{\mathbf{A}^T(\mathbf{G}^T\mathbf{G})^{-1}\mathbf{A}}. \quad (3.14)$$

A eficiência definida acima foi construída com a razão entre $\mathbf{L}^T(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{L}$ e $\mathbf{A}^T(\mathbf{G}^T\mathbf{G})^{-1}\mathbf{A}$. O numerador foi calculado considerando um desenho balanceado ($n_1 = n_2$). O denominador na Equação (3.14) é sempre maior ou igual que seu numerador, logo o desenho será dito eficiente se $\varepsilon_n = 1$.

O leitor interessado no artigo Atkinson (2002) notará que o autor comenta que o desenho balanceará os fatores de prognóstico se $\mathbf{H}^T\mathbf{X} = \mathbf{0}$. Entretanto, essa medida não condiz com a modelagem usada, pois esta implicaria em $\sum_{i \in T} X_{ik} = 0$, $k = 1, \dots, d$ (o que inclusive não condiz com o senso comum de balanço). Na verdade, este comentário surge no momento errado do artigo. O autor posteriormente define outro modelo,

$$\mathbb{E}(Y_i) = \alpha_1 I(i \in T) - \alpha_2 I(i \in T^c) + \mathbf{x}_i^T \beta,$$

com $\mathbb{V}(\mathbf{y}) = \sigma^2 \mathbf{I}_n$. Neste caso, a matriz de alocação assumirá valores em $\{-1, 1\}$ e $\mathbf{H}^T\mathbf{X} = \mathbf{0}$ implicará em $\sum_{i \in T} X_{ik} = \sum_{i \in T^c} X_{ik}$, $k = 1, \dots, d$, gerando a noção de balanço desejada.

3.3 Alguns desenhos para ensaios clínicos

Apresentamos a seguir três desenhos: aleatório, sistemático e da moeda viesada.

O desenho aleatório ainda é o método mais difundido, sendo de fácil implementação e de motivação simples. O desenho sistemático é utilizado aqui como extremo ao desenho aleatório. Embora os desenhos sistemáticos sejam completamente determinísticos, alguns autores os defendem, alegando que a real fonte de aleatorização está no paciente em si. Por último, o desenho da moeda viesada trata-se de um desenho com uma motivação simples porém com muitas discussões estatísticas e um tratamento teórico sério que levou o artigo original (Efron (1971)) a ser uma das principais referências sobre o assunto.

3.3.1 Desenhos sistemáticos

Os desenhos sistemáticos (Δ_s), consistem em desenhos nos quais a sequência da alocação está pré-estabelecida antes da entrada dos pacientes no experimento.

O desenho sistemático mais conhecido é o alternado. Nele, os pacientes são alocados na ordem $T_1T_2T_1T_2T_1\dots$, ou $T_2T_1T_2T_1T_2\dots$. Considerando este desenho, teremos:

- Se n for par, teremos que $D_1 = 0$. Em caso contrário, teremos que $D_1 = 1$. Portanto, a probabilidade de falta de balanço será zero.
- Seu viés de seleção será um, pois sempre saberemos qual será o próximo tratamento. Uma alternativa seria selecionar uma das duas sequências com probabilidade p , o que resultaria no viés de seleção $(p + n - 1)/n$. Entretanto, essa opção traz pouca melhoria.

3.3.2 Desenho aleatório

No desenho aleatório, representado por Δ_a , os pacientes devem ser alocados em qualquer um dos dois tratamentos com probabilidade $1/2$.

Em relação a falta de balanço entre os grupos de tratamentos, notemos que para qualquer n finito, teremos $(n_1|n) \sim \text{Binomial}(n, 1/2)$. Assim,

$$\begin{aligned} \mathbb{P}(D_1 = d|n) &= \mathbb{P}(|2n_1 - n| = d) = \mathbb{P}\left(n_1 = \frac{n+d}{2}\right) + \mathbb{P}\left(n_1 = \frac{n-d}{2}\right) \\ &= \binom{n}{\frac{n+d}{2}} \frac{1}{2^n} + \binom{n}{\frac{n-d}{2}} \frac{1}{2^n}, \end{aligned} \quad (3.15)$$

assim, se n for par,

$$\mathbb{P}(D_1 = d|n) = \begin{cases} \binom{n}{\frac{n}{2}} \frac{1}{2^n}, & d = 0 \\ \binom{n}{\frac{n+d}{2}} \frac{1}{2^{n-1}}, & d = 2, 4, 8, \dots, n \end{cases}$$

e, se n for ímpar,

$$\mathbb{P}(D_1 = d|n) = \begin{cases} \binom{n}{\frac{n+d}{2}} \frac{1}{2^{n-1}}, & d = 1, 3, 5, \dots, n. \end{cases}$$

A Tabela 3.1 mostra a probabilidade de balanço para alguns valores de n . Note-mos que, com o aumento de n , é cada vez menos provável que o experimento esteja balanceado.

n	Prob.	n	Prob.	n	Prob.
1	1.00	11	0.45	21	0.34
2	0.50	12	0.23	22	0.17
3	0.75	13	0.42	23	0.32
4	0.38	14	0.21	24	0.16
5	0.62	15	0.39	25	0.31
6	0.31	16	0.20	26	0.15
7	0.55	17	0.37	27	0.30
8	0.27	18	0.19	28	0.15
9	0.49	19	0.35	29	0.29
10	0.25	20	0.18	30	0.14

Tabela 3.1: Probabilidade de balanço para diferentes valores de n , onde a coluna Prob. equivale a $\mathbb{P}(D_1 = 0|\Delta_a)$ se n for par e $\mathbb{P}(D_1 = 1|\Delta_a)$ em caso contrário.

É natural que esse desenho apresente o menor viés de seleção. De fato, como a probabilidade de alocar o j -ésimo paciente em qualquer um dos dois grupos é a mesma, teremos

$$SB(\Delta) = \mathbb{E}(G) = \frac{1}{n} \sum_{j=1}^n p^*(j) = \frac{1}{2}.$$

3.3.3 Moeda Viesada

Em Efron (1971), o autor sugere um desenho que se comporta como o aleatório se o ensaio estiver balanceado, e força o balanço em caso contrário. Intuitivamente, este desenho se utiliza de duas moedas, uma honesta para o caso de o ensaio estar balanceado e outra viesada em caso contrário. Por esse motivo, este foi denominado desenho da moeda viesada. Denotaremos este desenho por Δ_b .

Formalmente, com a chegada de um novo paciente, as seguintes regras para alocação são adotadas por este desenho:

- se $n_1 - n_2 > 0$, aloque o paciente no tratamento T_1 com probabilidade q e no tratamento T_2 com probabilidade p ;
- se $n_1 - n_2 = 0$, aloque o paciente no tratamento T_1 com probabilidade $1/2$ e no tratamento T_2 com probabilidade $1/2$;
- se $n_1 - n_2 < 0$, aloque o paciente no tratamento T_1 com probabilidade p e no tratamento T_2 com probabilidade q ,

onde $p \geq q$, $p+q = 1$. O desenho tende para o balanceamento, sendo que a tendência é mais fraca se $p = 1/2$ e fica mais forte a medida que p se aproxima de 1. A escolha “favorita” do autor é $p = 2/3$.

Notemos que o desenho Δ_b depende da configuração atual do ensaio para realizar a próxima alocação. A probabilidade estimada de balanço para alguns valores de n é dada na Tabela 3.2. Podemos notar que este desenho possui probabilidade de balanço superior ao desenho aleatório.

Em vez de calcular D_1 para todo n , descobriremos qual a proporção do tempo em que o desenho fica balanceado quando $n \rightarrow \infty$. Seja $D_1(n) = |n_1 - n_2|$ a quantidade de falta de balanço do processo no instante n . A probabilidade de $D_1(n+1)$ depende unicamente do valor de $D_1(n)$. Teremos que $\{D_1(n), n = 0, 1, 2, \dots\}$, será uma cadeia de Markov com probabilidades de transição dadas por

$$\mathbb{P}(D_1(n+1) = d+1 | D_1(n) = d) = q,$$

$$\mathbb{P}(D_1(n+1) = d-1 | D_1(n) = d) = p,$$

$$\mathbb{P}(D_1(n+1) = 1 | D_1(n) = 0) = 1.$$

n	Prob.	n	Prob.	n	Prob.
1	1,000	11	0,788	21	0,766
2	0,660	12	0,514	22	0,512
3	0,880	13	0,779	23	0,766
4	0,592	14	0,526	24	0,516
5	0,843	15	0,763	25	0,744
6	0,569	16	0,538	26	0,482
7	0,816	17	0,766	27	0,748
8	0,547	18	0,520	28	0,519
9	0,816	19	0,797	29	0,741
10	0,524	20	0,498	30	0,523

Tabela 3.2: Probabilidade de balanço para para o desenho da moeda viesada para diferentes valores de n . A coluna Prob. dá a probabilidade de balanço estimada.

Notemos que $\{D_1(n), n = 0, 1, \dots\}$ é um passeio aleatório com barreira, recorrente positivo e de período 2, e, portanto, sua distribuição estacionária é dada por

$$\pi_j = \begin{cases} \frac{p-q}{2p}, & j = 0, \\ \frac{(p-q)q^{j-1}}{2p^j}, & j > 0. \end{cases}$$

Como $D_1(n)$ só assume valores pares ou ímpares se n for par ou ímpar (periodicidade 2), as probabilidades limite devem ser dobradas, resultando em

$$\lim_{n \rightarrow \infty} \mathbb{P}(D_1(2n) = 0) = 2\pi_0 = \frac{p-q}{p}, \quad (3.16)$$

$$\lim_{n \rightarrow \infty} \mathbb{P}(D_1(2n+1) = 1) = 2\pi_1 = \frac{q(p-q)}{p}. \quad (3.17)$$

Para o cálculo do viés de seleção, se $D_1(j) = 0$, então $p^*(j) = 1/2$. Em caso contrário, o palpite será dado a favor do tratamento com menor número de pacientes, com probabilidade p . Assim,

$$SB(\Delta_b) = \mathbb{E}(G) = \frac{1}{2}\mathbb{P}(D_1(n) = 0) + p\mathbb{P}(D_1(n) > 0).$$

A Tabela 3.3 mostra o viés de seleção estimado de Δ_b para $n = 1, 2, \dots, 30$ utilizando $p = 2/3$.

n	Prob.	n	Prob.	n	Prob.
1	0,5	11	0,54	21	0,54
2	0,56	12	0,58	22	0,58
3	0,52	13	0,54	23	0,54
4	0,57	14	0,58	24	0,58
5	0,53	15	0,54	25	0,54
6	0,57	16	0,58	26	0,59
7	0,53	17	0,54	27	0,54
8	0,58	18	0,58	28	0,58
9	0,53	19	0,53	29	0,54
10	0,58	20	0,58	30	0,58

Tabela 3.3: Viés de seleção estimado para o desenho da moeda viesada considerando alguns valores de n

Capítulo 4

Resultados

No Capítulo 3, discutimos que, em geral, modelos lineares são empregados para analisar se existem diferenças entre os tratamentos envolvidos. Neste caso, estamos interessados em estimar o contraste dos tratamentos. Uma vez que o estimador para os contrastes é um estimador não viesado, independente do desenho escolhido, gostaríamos de determinar qual o efeito do desenho na variância do contraste, dada por

$$\mathbb{V}(\hat{\alpha}_1 - \hat{\alpha}_2) \propto \mathbf{A}^T (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{A}. \quad (4.1)$$

Neste capítulo apresentamos os resultados deste trabalho. Na Seção 4.1 apresentamos uma forma alternativa de visualizar a variância do contraste dada acima. Esta decomposição será fundamental para o desenvolvimento dos resultados posteriores. Na Seção 4.2 decomparamos a variância do contraste em dois termos, sendo que um deles caracteriza o efeito dos fatores de prognóstico. Em seguida, mostramos que sob determinadas condições fracas, a variância do contraste se aproximará da variância de uma ANOVA com um critério de classificação, com o aumento do tamanho da amostra, o que implica que os fatores de prognóstico possuem cada vez menos influência. Na Seção 4.3 mostramos como obter a melhor alocação quando todos os fatores de prognóstico são conhecidos. Disto, podemos obter a eficiência média de qualquer desenho via simulação.

4.1 Reescrevendo $\mathbb{V}(\hat{\alpha}_1 - \hat{\alpha}_2)$

Sabemos que, independente do desenho escolhido, uma vez que os pacientes foram alocados em seus respectivos tratamentos, sempre conseguiremos um estimador não viesado para o contraste. Contudo, notemos a partir da Equação (4.1) que a variância do constraste será influenciada pelas matrizes de alocação e pelos fatores de prognóstico envolvidos. Nesta seção, escreveremos esta variância de forma alternativa.

Lembremos que

$$\mathbb{V}(\hat{\alpha}_1 - \hat{\alpha}_2) = \sigma^2 \mathbf{L}^T (\mathbf{H}^T \mathbf{H} - \mathbf{H}^T \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T)^{-1} \mathbf{L}. \quad (4.2)$$

Seja \mathbf{h}^T um vetor de zeros e uns, onde $h_i = 1$ se o i -ésimo paciente pertence ao grupo T_1 e $h_i = 0$ em caso contrário. Naturalmente $\mathbf{H} = (\mathbf{h}, \mathbf{1}_n - \mathbf{h})$.

Seja $\mathbf{D} = \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T$. Então,

$$\begin{aligned} \mathbf{H}^T \mathbf{D} \mathbf{H} &= \begin{pmatrix} \mathbf{h}^T \\ \mathbf{1}_n^T - \mathbf{h}^T \end{pmatrix} \mathbf{D} \begin{pmatrix} \mathbf{h} & \mathbf{1}_n - \mathbf{h} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{h}^T \mathbf{D} \mathbf{h} & \mathbf{h}^T \mathbf{D} (\mathbf{1}_n - \mathbf{h}) \\ (\mathbf{1}_n - \mathbf{h})^T \mathbf{D} \mathbf{h} & (\mathbf{1}_n - \mathbf{h})^T \mathbf{D} (\mathbf{1}_n - \mathbf{h}) \end{pmatrix}. \end{aligned}$$

Como,

$$\mathbf{H}^T \mathbf{H} = \begin{pmatrix} \mathbf{h}^T \mathbf{h} & \mathbf{h}^T (\mathbf{1}_n - \mathbf{h}) \\ (\mathbf{1}_n - \mathbf{h})^T \mathbf{h} & (\mathbf{1}_n - \mathbf{h})^T (\mathbf{1}_n - \mathbf{h}) \end{pmatrix} = \begin{pmatrix} n_1 & 0 \\ 0 & n_2 \end{pmatrix},$$

teremos

$$\mathbf{H}^T \mathbf{H} - \mathbf{H}^T \mathbf{D} \mathbf{H} = \begin{pmatrix} n_1 - \mathbf{h}^T \mathbf{D} \mathbf{h} & -\mathbf{h}^T \mathbf{D} (\mathbf{1}_n - \mathbf{h}) \\ -(\mathbf{1}_n - \mathbf{h})^T \mathbf{D} \mathbf{h} & n_2 - (\mathbf{1}_n - \mathbf{h})^T \mathbf{D} (\mathbf{1}_n - \mathbf{h}) \end{pmatrix}.$$

Considere o seguinte lema, de demonstração trivial.

Lema 4.1. *Seja \mathbf{W} uma matriz de ordem dois, simétrica e invertível. Então*

$$\mathbf{L}^T \mathbf{W} \mathbf{L} = \frac{w_{11} + 2w_{12} + w_{22}}{w_{11}w_{22} - w_{12}^2}. \quad (4.3)$$

Identificando a matriz $\mathbf{H}^T \mathbf{H} - \mathbf{H}^T \mathbf{D} \mathbf{H}$ com a matriz \mathbf{W} do Lema 4.1, teremos

$$\mathbb{V}(\hat{\alpha}_1 - \hat{\alpha}_2) = \sigma^2 \frac{n - \mathbf{1}_n^T \mathbf{D} \mathbf{1}_n}{(n_1 - \mathbf{h}^T \mathbf{D} \mathbf{h})(n_2 - (\mathbf{1}_n - \mathbf{h})^T \mathbf{D} (\mathbf{1}_n - \mathbf{h})) - [\mathbf{h}^T \mathbf{D} (\mathbf{1}_n - \mathbf{h})]^2}. \quad (4.4)$$

Notemos que o numerador da Equação (4.4) não depende de \mathbf{h} . Logo, encontrar a alocação que minimiza a variância do contraste implica em encontrar a alocação que maximiza o denominador desta equação. Seja R tal denominador. Temos que

$$\begin{aligned}
R &= (n_1 - \mathbf{h}^T \mathbf{D} \mathbf{h})(n_2 - (\mathbf{1}_n - \mathbf{h})^T \mathbf{D} (\mathbf{1}_n - \mathbf{h})) - [\mathbf{h}^T \mathbf{D} (\mathbf{1}_n - \mathbf{h})]^2 \\
&= n_1 n_2 - n \mathbf{h}^T \mathbf{D} \mathbf{h} - n_1 \mathbf{1}_n^T \mathbf{D} \mathbf{1}_n + 2n_1 \mathbf{1}_n^T \mathbf{D} \mathbf{h} + \mathbf{h}^T \mathbf{D} \mathbf{h} \mathbf{1}_n^T \mathbf{D} \mathbf{1}_n - (\mathbf{1}_n^T \mathbf{D} \mathbf{h})^2 \\
&= n_1 n - n \mathbf{h}^T \mathbf{D} \mathbf{h} - n_1 \mathbf{1}_n^T \mathbf{D} \mathbf{1}_n + \mathbf{h}^T \mathbf{D} \mathbf{h} \mathbf{1}_n^T \mathbf{D} \mathbf{1}_n - (n_1 - \mathbf{1}_n^T \mathbf{D} \mathbf{h})^2 \\
&= (n - \mathbf{1}_n^T \mathbf{D} \mathbf{1}_n)(n_1 - \mathbf{h}^T \mathbf{D} \mathbf{h}) - (n_1 - \mathbf{1}_n^T \mathbf{D} \mathbf{h})^2.
\end{aligned}$$

e, notando que $n = \mathbf{1}_n^T \mathbf{1}_n$ e $n_1 = \mathbf{1}_n^T \mathbf{h} = \mathbf{h}^T \mathbf{h}$, teremos

$$R = \mathbf{1}_n^T \mathbf{Q} \mathbf{1}_n \mathbf{h}^T \mathbf{Q} \mathbf{h} - (\mathbf{1}_n^T \mathbf{Q} \mathbf{h})^2.$$

Assim, a Equação (4.4) pode ser reescrita como

$$\mathbb{V}(\hat{\alpha}_1 - \hat{\alpha}_2) = \sigma^2 \frac{\mathbf{1}_n^T \mathbf{Q} \mathbf{1}_n}{\mathbf{1}_n^T \mathbf{Q} \mathbf{1}_n \mathbf{h}^T \mathbf{Q} \mathbf{h} - [\mathbf{1}_n^T \mathbf{Q} \mathbf{h}]^2}. \quad (4.5)$$

4.2 Sobre o comportamento assintótico de $\mathbb{V}(\hat{\alpha}_1 - \hat{\alpha}_2)$

Considerando um modelo linear para analisar o resultado do ensaio clínico, sabemos que tanto a matriz de alocação \mathbf{H} quanto a matriz de fatores de prognóstico \mathbf{X} possuem papel importante para a determinação da variância do contraste desejado.

Da seção anterior, sabemos que

$$\mathbb{V}(\hat{\alpha}_1 - \hat{\alpha}_2) = \sigma^2 \frac{\mathbf{1}_n^T \mathbf{Q} \mathbf{1}_n}{\mathbf{1}_n^T \mathbf{Q} \mathbf{1}_n \mathbf{h}^T \mathbf{Q} \mathbf{h} - [\mathbf{1}_n^T \mathbf{Q} \mathbf{h}]^2},$$

onde $\mathbf{Q} = \mathbf{I}_n - \mathbf{X} \mathbf{C} \mathbf{X}^T$, com $\mathbf{C} = (\mathbf{X}^T \mathbf{X})^{-1}$.

Agora, seja $\bar{\mathbf{x}}^T = \mathbf{1}_n^T \mathbf{X} / n$ o vetor de médias dos fatores de prognóstico. Assim,

$$\mathbf{1}_n^T \mathbf{Q} \mathbf{1}_n = \mathbf{1}_n^T \mathbf{1}_n - \mathbf{1}_n^T \mathbf{X} \mathbf{C} \mathbf{X}^T \mathbf{1}_n = n (1 - n \bar{\mathbf{x}}^T \mathbf{C} \bar{\mathbf{x}}). \quad (4.6)$$

Seja $\bar{\mathbf{x}}_{(1)}^T$ o vetor de médias dos fatores de prognóstico dentro do grupo T_1 . Note-mos que $\bar{\mathbf{x}}_{(1)}^T = \mathbf{h}^T \mathbf{X} / n_1$, logo

$$\mathbf{h}^T \mathbf{Q} \mathbf{h} = \mathbf{h}^T \mathbf{h} - \mathbf{h}^T \mathbf{X} \mathbf{C} \mathbf{X}^T \mathbf{h} = n_1 (1 - n_1 \bar{\mathbf{x}}_{(1)}^T \mathbf{C} \bar{\mathbf{x}}_{(1)}). \quad (4.7)$$

Por último, notemos que

$$\mathbf{1}_n^T \mathbf{Q} \mathbf{h} = \mathbf{1}_n^T \mathbf{h} - \mathbf{1}_n^T \mathbf{X} \mathbf{C} \mathbf{X}^T \mathbf{h} = n_1 (1 - n \bar{\mathbf{x}}^T \mathbf{C} \bar{\mathbf{x}}_{(1)}). \quad (4.8)$$

Pelas Equações (4.6), (4.7) e (4.8) teremos a expressão alternativa

$$\mathbb{V}(\hat{\alpha}_1 - \hat{\alpha}_2) = \sigma^2 \frac{n (1 - n \bar{\mathbf{x}}^T \mathbf{C} \bar{\mathbf{x}})}{nn_1 (1 - n \bar{\mathbf{x}}^T \mathbf{C} \bar{\mathbf{x}}) \left(1 - n_1 \bar{\mathbf{x}}_{(1)}^T \mathbf{C} \bar{\mathbf{x}}_{(1)}\right) - n_1^2 (1 - n \bar{\mathbf{x}}^T \mathbf{C} \bar{\mathbf{x}}_{(1)})^2}. \quad (4.9)$$

Podemos ainda notar que $\mathbb{V}(\hat{\alpha}_1 - \hat{\alpha}_2)$ é igual a

$$\sigma^2 \frac{n}{n_1 n_2} \left\{ \frac{1 - n \bar{\mathbf{x}}^T \mathbf{C} \bar{\mathbf{x}}}{\frac{n}{n_2} (1 - n \bar{\mathbf{x}}^T \mathbf{C} \bar{\mathbf{x}}) \left(1 - n_1 \bar{\mathbf{x}}_{(1)}^T \mathbf{C} \bar{\mathbf{x}}_{(1)}\right) - \frac{n_1}{n_2} (1 - n \bar{\mathbf{x}}^T \mathbf{C} \bar{\mathbf{x}}_{(1)})^2} \right\}. \quad (4.10)$$

Notemos pela equação acima que a variância do contraste pode ser decomposta em duas partes. A parte fora das chaves é equivalente a variância do contraste na ausência de fatores de prognóstico, já apresentada na Equação (3.6) e reproduzida a seguir

$$\mathbb{V}^*(\hat{\alpha}_1 - \hat{\alpha}_2) = \sigma^2 \frac{n}{n_1 n_2}. \quad (4.11)$$

A parte dentro das chaves na Equação (4.10) mostra o efeito das covariáveis e da alocação na determinação da variância. Discutiremos em quais condições as variâncias em (4.10) e (4.11) são assintoticamente equivalentes, ou seja,

$$\frac{\mathbb{V}^*(\hat{\alpha}_1 - \hat{\alpha}_2)}{\mathbb{V}(\hat{\alpha}_1 - \hat{\alpha}_2)} \rightarrow 1, \quad n \rightarrow \infty. \quad (4.12)$$

Para tanto, consideremos primeiro a Definição 4.2 e as Proposições 4.3, 4.4, 4.5 e 4.6, enunciadas e demonstradas a seguir.

Definição 4.2. *Sob um desenho Δ , denotaremos a probabilidade de que o j -ésimo paciente pertença ao grupo T por*

$$p_T(j) = \mathbb{P}(j \in T | \Delta). \quad (4.13)$$

Proposição 4.3. *Considere que cada paciente é independente dos demais e seja \mathbf{x}_i^T o vetor de fatores de prognóstico do i -ésimo paciente. Sejam $\mathbb{E}(\mathbf{x}) = \mu$ e $\mathbb{V}(\mathbf{x}) = \Sigma$, com Σ não singular. Então*

$$n \bar{\mathbf{x}}^T \mathbf{C} \bar{\mathbf{x}} \xrightarrow{q.c.} \mu^T (\Sigma + \mu \mu^T)^{-1} \mu, \quad (4.14)$$

quando $n \rightarrow \infty$.

Demonstração. Pela Lei Forte dos Grandes Números, sabemos que

$$\frac{\mathbf{X}^T \mathbf{X}}{n} \xrightarrow{q.c.} \Sigma + \mu\mu^T,$$

quando $n \rightarrow \infty$. Assim, teremos que

$$n\mathbf{C} = n(\mathbf{X}^T \mathbf{X})^{-1} = \left(\frac{\mathbf{X}^T \mathbf{X}}{n} \right)^{-1} \xrightarrow{q.c.} (\Sigma + \mu\mu^T)^{-1},$$

onde é bem conhecido que tal inversa existe pela não singularidade de Σ . Assim, novamente pela Lei Forte dos Grandes Números, teremos que

$$n\bar{\mathbf{x}}^T \mathbf{C} \bar{\mathbf{x}} \xrightarrow{q.c.} \mu^T (\Sigma + \mu\mu^T)^{-1} \mu, \quad n \rightarrow \infty.$$

□

Proposição 4.4. *Se existir*

$$\pi_T = \lim_{n \rightarrow \infty} \sum_{i=1}^n \frac{p_T(i)}{n}$$

teremos que

$$\frac{n_1}{n} = \sum_{i=1}^n \frac{I(i \in T)}{n} \xrightarrow{P} \pi_T, \quad n \rightarrow \infty. \quad (4.15)$$

Demonstração. Seja $p_T(i, j) = \mathbb{P}((i, j) \in T | \Delta)$. Notemos que

$$p_T(i, j) = \mathbb{P}(j \in T | i, \Delta) p_T(i) \leq p_T(i).$$

Utilizando a desigualdade de Tchebychev, teremos

$$\begin{aligned}
\mathbb{P} \left(\left| \frac{\sum_{i=1}^n I(i \in T)}{n} - \pi_T \right| \geq \varepsilon \right) &\leq \frac{1}{\varepsilon^2} \mathbb{E} \left(\sum_{i=1}^n \frac{I(i \in T)}{n} - \pi_T \right)^2 \\
&= \frac{1}{\varepsilon^2} \left[\mathbb{E} \left(\sum_{i=1}^n \frac{I(i \in T)}{n} \right)^2 + \pi_T^2 - 2\pi_T \mathbb{E} \left(\sum_{i=1}^n \frac{I(i \in T)}{n} \right) \right] \\
&= \frac{1}{\varepsilon^2} \left[\mathbb{E} \left(\sum_{i=1}^n \frac{I(i \in T)}{n} \right)^2 + \pi_T^2 - 2\pi_T \sum_{i=1}^n \frac{p_T(i)}{n} \right] \\
&= \frac{1}{\varepsilon^2} \left[\mathbb{E} \left(\sum_{i=1}^n \frac{I(i \in T)}{n^2} + \sum_{i \neq j} \frac{I(i \in T)I(j \in T)}{n^2} \right) \right. \\
&\quad \left. + \pi_T^2 - 2\pi_T \sum_{i=1}^n \frac{p_T(i)}{n} \right] \\
&= \frac{1}{\varepsilon^2} \left[\frac{1}{n} \sum_{i=1}^n \frac{p_T(i)}{n} + \frac{1}{n} \sum_{i \neq j} \frac{p_T(i, j)}{n} + \pi_T^2 - 2\pi_T \sum_{i=1}^n \frac{p_T(i)}{n} \right] \\
&\leq \frac{1}{\varepsilon^2} \left[\frac{1}{n} \sum_{i=1}^n \frac{p_T(i)}{n} + \frac{1}{n} \sum_{i \neq j} \frac{p_T(i)}{n} + \pi_T^2 - 2\pi_T \sum_{i=1}^n \frac{p_T(i)}{n} \right] \\
&\rightarrow 0, \quad n \rightarrow \infty,
\end{aligned}$$

logo $n_1/n \xrightarrow{P} \pi_T$. □

Proposição 4.5. *Considere as mesmas condições das Proposições 4.3 e 4.4. Então,*

$$\bar{x}_{(1)k} = \sum_{i=1}^n \frac{X_{ik} I(i \in T)}{n_1} \xrightarrow{P} \mu_k, \quad (4.16)$$

quando $n \rightarrow \infty$.

Demonstração. Primeiro, notemos que

$$\bar{x}_{(1)k} = \frac{\sum_{i=1}^n X_{ik} I(i \in T)}{n_1} = \frac{\sum_{i=1}^n X_{ik} I(i \in T)/n}{n_1/n}.$$

O denominador da equação acima converge em probabilidade para π_T , pela Proposição 4.4. Mostraremos então que o numerador converge em probabilidade para $\pi_T \mu_k$, demonstrando assim o resultado. Notando que

$$0 < \sum_{i=1}^n \frac{p_T(i)^2}{n^2} \leq \frac{1}{n} \sum_{i=1}^n \frac{p_T(i)}{n} \rightarrow 0,$$

quando $n \rightarrow \infty$ teremos, pelo Teorema do Confronto, que $\sum_{i=1}^n p_T(i)^2/n^2 \xrightarrow{P} 0$. Pela desigualdade de Tchebychev,

$$\begin{aligned}
& \mathbb{P} \left(\left| \sum_{i=1}^n \frac{I(i \in T)X_{ik}}{n} - \mu_k \pi_T \right| \geq \varepsilon \right) \leq \frac{1}{\varepsilon^2} \mathbb{E} \left(\sum_{i=1}^n \frac{I(i \in T)X_{ik}}{n} - \mu_k \pi_T \right)^2 \\
&= \frac{1}{\varepsilon^2} \mathbb{E} \left(\mu_k^2 \pi_T^2 - 2\mu_k \pi_T \sum_{i=1}^n \frac{I(i \in T)X_{ik}}{n} + \sum_{i=1}^n \sum_{j=1}^n \frac{I(i \in T)I(j \in T)X_{ik}X_{jk}}{n^2} \right) \\
&= \frac{1}{\varepsilon^2} \left(\mu_k^2 \pi_T^2 - 2\mu_k \pi_T \sum_{i=1}^n \frac{p_T(i)\mu_k}{n} + \sum_{i=1}^n \frac{(\mu_k^2 + \sigma^2)p_T(i)}{n^2} + \sum_{i \neq j} \mu_k^2 \frac{p_T(i)p_T(j)}{n^2} \right) \\
&= \frac{1}{\varepsilon^2} \left(\mu_k^2 \pi_T^2 - 2\mu_k \pi_T \sum_{i=1}^n \frac{p_T(i)\mu_k}{n} + \sum_{i=1}^n \frac{(\mu_k^2 + \sigma^2)p_T(i)}{n^2} \right. \\
&\quad \left. + \sum_{i=1}^n \sum_{j=1}^n \mu_k^2 \frac{p_T(i)p_T(j)}{n^2} - \mu_k^2 \sum_{i=1}^n \frac{p_T(i)^2}{n^2} \right) \\
&\rightarrow 0, \quad \text{quando } n \rightarrow \infty.
\end{aligned}$$

Portanto, teremos que $\sum_{i=1}^n \frac{I(i \in T)X_{ik}}{n} \xrightarrow{P} \pi_T \mu_k$, demonstrando o resultado. \square

Proposição 4.6. *Seja $\bar{\mathbf{x}}_{(1)}^T = \mathbf{h}^T \mathbf{X}/n_1$ e considere que as condições das Proposições 4.3 e 4.4 estão satisfeitas. Por último, considere que o desenho Δ aloca os pacientes independente dos fatores de prognóstico. Então, se $n_1 \rightarrow \infty$ e $n_2 \rightarrow \infty$ quando $n \rightarrow \infty$, teremos*

$$n_1 \bar{\mathbf{x}}_{(1)}^T \mathbf{C} \bar{\mathbf{x}}_{(1)} \xrightarrow{P} \mu^T (\Sigma + \mu \mu^T)^{-1} \mu \pi_T. \quad (4.17)$$

Demonstração. Da Proposição 4.3, sabemos que $n\mathbf{C} \xrightarrow{P} (\Sigma + \mu \mu^T)^{-1}$ (pois convergência quase certa implica em convergência em probabilidade). Da Proposição 4.4, sabemos que $n_1/n \xrightarrow{P} \pi_T$. Da Proposição 4.5 sabemos que $\mathbf{x}_{(1)}^T \xrightarrow{P} \mu^T$. Assim

$$n_1 \bar{\mathbf{x}}_{(1)}^T \mathbf{C} \bar{\mathbf{x}}_{(1)} = \frac{n_1}{n} \bar{\mathbf{x}}_{(1)}^T n\mathbf{C} \bar{\mathbf{x}}_{(1)} \xrightarrow{P} \mu^T (\Sigma + \mu \mu^T)^{-1} \mu \pi_T.$$

\square

As proposições acima são suficientes para garantir o seguinte teorema.

Teorema 4.7. *Considere um ensaio clínico onde os pacientes são independentes entre si. Seja Δ um desenho com alocação sequencial com as seguintes propriedades:*

1. *O desenho não utiliza a informação dos fatores de prognóstico para realizar a alocação.*

2. Existe $\lim_{n \rightarrow \infty} \sum_{i=1}^n p_T(i)/n$.

3. n_1 e n_2 divergem quando $n \rightarrow \infty$.

Então, se \mathbf{X} possui matriz de covariâncias não-singular, para este desenho

$$\frac{\mathbb{V}^*(\hat{\alpha}_1 - \hat{\alpha}_2)}{\mathbb{V}(\hat{\alpha}_1 - \hat{\alpha}_2)} \xrightarrow{P} 1, \quad n \rightarrow \infty.$$

Demonstração. Segue imediatamente das Proposições 4.3, 4.4, 4.5 e 4.6. \square

Na literatura, muitas propostas de desenhos que não utilizam fatores de prognóstico possuem o limite acima. Essa é uma característica esperada em desenhos que privilegiam o balanço entre os tratamentos, uma vez que a frequência relativa de alocações em cada um dos grupos deve convergir para $1/2$. A seguinte proposição mostra que os desenhos apresentados no Capítulo 3 satisfazem as condições do Teorema 4.7.

Proposição 4.8. *Os desenhos aleatório, sistemático alternado e da moeda viesada satisfazem as condições do Teorema 4.7.*

Demonstração. As condições (1) e (3) são imediatas. Verifiquemos então a condição (2) para cada um dos desenhos. Para o desenho aleatório teremos

$$\sum_{i=1}^n \frac{p_T(i)}{n} = \sum_{i=1}^n \frac{1}{2n} = \frac{1}{2}$$

logo, $\pi_T = 1/2$.

Para o desenho sistemático alternado, sabemos que $p_T(i) = 1$ se i for ímpar e $p_T(i) = 0$ em caso contrário. Seja $\Gamma_1 = \{n = 1, 2, \dots : n \text{ é ímpar}\}$ e $\Gamma_2 = \{n = 1, 2, \dots : n \text{ é par}\}$.

$$\sum_{i=1}^n \frac{p_T(i)}{n} = \sum_{i \in \Gamma_1} \frac{p_T(i)}{n} + \sum_{i \in \Gamma_2} \frac{p_T(i)}{n} = \sum_{i \in \Gamma_1} \frac{1}{n} = \frac{[n/2]}{n} \rightarrow \frac{1}{2} \quad n \rightarrow \infty$$

logo $\pi_T = 1/2$.

Para o desenho da moeda viesada, notemos que

$$\begin{aligned}
p_T(i) &= \sum_{d=-\infty}^{\infty} \mathbb{P}(i \in T | n_1 - n_2 = d) \mathbb{P}(n_1 - n_2 = d) \\
&= \sum_{d=-\infty}^1 \mathbb{P}(i \in T | n_1 - n_2 = d) \mathbb{P}(n_1 - n_2 = d) \\
&\quad + \sum_{d=1}^{\infty} \mathbb{P}(i \in T | n_1 - n_2 = d) \mathbb{P}(n_1 - n_2 = d) + \frac{1}{2} \mathbb{P}(n_1 - n_2 = 0) \\
&= p \sum_{d=-\infty}^1 \mathbb{P}(n_1 - n_2 = d) + q \sum_{d=1}^{\infty} \mathbb{P}(n_1 - n_2 = d) + \frac{1}{2} \mathbb{P}(n_1 - n_2 = 0) \\
&= \sum_{d=1}^{\infty} \mathbb{P}(n_1 - n_2 = d) + \frac{1}{2} \mathbb{P}(n_1 - n_2 = 0) \\
&= \frac{1}{2} (1 - \mathbb{P}(n_1 - n_2 = 0)) + \frac{1}{2} \mathbb{P}(n_1 - n_2 = 0) = \frac{1}{2}.
\end{aligned}$$

□

4.3 Alocação ótima quando a matriz de fatores de prognóstico é conhecida

A eficiência de Atkinson compara a variância do contraste observada com uma cota inferior para a variância. Uma eficiência de 100% implica que a variância do modelo é idêntica à variância do modelo sem os fatores de prognóstico que utiliza um desenho que privilegia o balanço. Entretanto, vale notar que este mínimo dificilmente será atingido na prática, uma vez que ele será atingido somente se $\mathbf{H}^T \mathbf{X} = \mathbf{0}$.

A melhor forma de comparar a eficácia dos desenhos com relação a alocação dos pacientes considerando os fatores de prognóstico seria comparar os resultados da alocação com o melhor resultado que poderia ser obtido para a população em estudo. Considerando ainda que a minimização do contraste e o balanço entre os grupos de tratamentos são as melhores estratégias para alocar os pacientes, nesta seção mostraremos como obter a melhor alocação de uma população conhecida. Este tipo de estratégia é útil para verificar a eficácia de um desenho considerando uma população fictícia, ou mesmo para obter a melhor alocação no caso em que uma lista dos pacientes com seus respectivos fatores de prognóstico está disponível, sem a necessidade de utilizar um método de alocação sequencial.

Suponha que um observador externo sabe exatamente quais serão os pacientes que deverão participar do ensaio, incluindo seus fatores de prognóstico. Neste caso, ele tem o privilégio de saber qual será o peso, no fim do ensaio, de alocar um paciente em um dos dois grupos. Assim, este observador alocará os pacientes de modo a obter a menor variância do contraste, respeitando restrições de balanço se ele julgar necessário. A tarefa de encontrar a melhor alocação é árdua: por exemplo, existem 68.923.264.410 possibilidades distintas de alocar quarenta pacientes de modo balanceado, o que torna a busca pelo ótimo por força bruta uma tarefa com elevado custo computacional e métodos de otimização devem ser explorados.

Mostraremos que a alocação ótima pode ser obtida via um problema de programação interia. Para tanto, lembremos que a variância do constraste é dada por

$$\mathbb{V}(\hat{\alpha}_1 - \hat{\alpha}_2) = \frac{\mathbf{1}_n^T \mathbf{Q} \mathbf{1}_n}{\mathbf{1}_n^T \mathbf{Q} \mathbf{1}_n \mathbf{h}^T \mathbf{Q} \mathbf{h} - [\mathbf{1}_n^T \mathbf{Q} \mathbf{h}]^2} = \frac{\mathbf{1}_n^T \mathbf{Q} \mathbf{1}_n}{R}. \quad (4.18)$$

Conforme dito anteriormente, o numerador da variância acima não depende da matriz de alocação. Portanto, o efeito da alocação na variância do contraste pode ser medido através de R . Em particular, minimizar a variância do contraste para uma matriz de fatores de prognóstico fixa \mathbf{X} implica em maximizar R . Podemos notar que

$$R = \mathbf{1}_n^T \mathbf{Q} \mathbf{1}_n \mathbf{h}^T \mathbf{Q} \mathbf{h} - (\mathbf{1}_n^T \mathbf{Q} \mathbf{h})^2 = \mathbf{1}_n^T \mathbf{Q} [\mathbf{1}_n \mathbf{h}^T - \mathbf{h} \mathbf{1}_n^T] \mathbf{Q} \mathbf{h} = \mathbf{1}_n^T \mathbf{Q} \mathbf{B} \mathbf{Q} \mathbf{h},$$

onde $\mathbf{B} = \mathbf{1}_n \mathbf{h}^T - \mathbf{h} \mathbf{1}_n^T$.

Proposição 4.9. *Os elementos da matriz \mathbf{B} são $b_{ij} = I(i \in T_1) - I(j \in T_1)$.*

Demonstração. Seja $\mathbf{E} = \mathbf{1}_n \mathbf{h}^T$. A i -ésima linha de \mathbf{E} será o próprio vetor \mathbf{h}^T , logo $e_{ij} = I(j \in T_1)$.

Agora, seja $\mathbf{F} = \mathbf{h} \mathbf{1}_n^T$. A i -ésima linha de \mathbf{F} será $\mathbf{1}_n^T$ se o i -ésimo indivíduo for alocado no grupo tratamento e $\mathbf{0}_n^T$ em caso contrário. Assim $f_{ij} = I(i \in T_1)$.

Portanto

$$b_{ij} = I(j \in T_1) - I(i \in T_1).$$

□

A matriz \mathbf{B} indica se dois pacientes foram alocados em grupos diferentes. Um resultado $b_{ij} = 0$ implica que os pacientes i e j foram alocados no mesmo grupo; um

resultado $b_{ij} = 1$ indica que o paciente j foi alocado no grupo T_1 e o paciente i no grupo T_2 ; $b_{ij} = -1$ indica a alocação oposta.

Sejam $q_{ij}, i, j = 1, \dots, n$ os elementos da matrix $\mathbf{Q} = \mathbf{I}_n - \mathbf{X}^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}$. Fazendo $q_{.j} = \sum_{i=1}^n q_{ij}$, teremos

$$\begin{aligned}
R &= \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n q_{.i} q_{kj} h_k (h_j - h_i) = \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n q_{.i} q_{kj} h_k h_j - \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n q_{.i} q_{kj} h_k h_i \\
&= q_{.} \sum_{j=1}^n \sum_{k=1}^n q_{kj} h_k h_j - \sum_{i=1}^n \sum_{k=1}^n q_{.i} q_{.k} h_k h_i \\
&= q_{.} \sum_{i=1}^n \sum_{j=1}^n q_{ij} h_i h_j - \sum_{i=1}^n \sum_{j=1}^n q_{.i} q_{.j} h_i h_j \\
&= \sum_{i=1}^n \sum_{j=1}^n (q_{.} q_{ij} - q_{.i} q_{.j}) h_i h_j \tag{4.19}
\end{aligned}$$

onde $h_i = I(i \in T_1)$ e $q_{.i} = q_{.}$ pela simetria de \mathbf{Q} . Pela equação acima podemos perceber como os elementos da matrix \mathbf{Q} funcionam como pesos para a alocação, indicando qual o custo de alocar os pacientes i e j no mesmo grupo.

Basicamente, o problema de encontrar a melhor alocação, com restrição de balanço entre os grupos de tratamento, é equivalente a encontrar a solução do problema

$$\max \sum_{i=1}^n \sum_{j=1}^n (q_{.} q_{ij} - q_{.i} q_{.j}) h_i h_j$$

s.a.

$$\sum_{i=1}^n h_i = n_1$$

$$h_i \in \{0, 1\} \forall i = 1, \dots, n.$$

Podemos linearizar este problema criando a variável

$$u_{ij} = h_i h_j = I(i \in T) I(j \in T). \tag{4.20}$$

Assim, basta encontrarmos as restrições necessárias para que a matrix $\mathbf{U} = \{u_{ij}\}$ seja equivalente a matrix

$$\mathbf{T} = \begin{pmatrix} h_1^2 & h_1 h_2 & \dots & h_1 h_n \\ h_2 h_1 & h_2^2 & \dots & h_2 h_n \\ \vdots & \vdots & \ddots & \vdots \\ h_n h_1 & h_n h_2 & \dots & h_n^2 \end{pmatrix} \tag{4.21}$$

onde $\sum_{i=1}^n h_i = n_1$.

Proposição 4.10. *As seguintes restrições são necessárias para que \mathbf{U} seja igual a \mathbf{T} :*

$$(1) u_{ij} = u_{ji}, \quad \forall i, j = 1, \dots, n.$$

$$(2) \sum_{i=1}^n u_{ii} = n_1.$$

$$(3) \sum_{i=1}^n u_{ij} = n_1 u_{jj}, \quad \forall j \in 1, \dots, n.$$

$$(4) u_{ij} \in \{0, 1\} \forall i, j = 1, \dots, n.$$

Demonstração. Mostraremos que as restrições dadas são necessárias para estabelecer a igualdade entre as matrizes, o que equivale a mostrar que $u_{ij} = u_{ii}u_{jj}$. Assim, devemos mostrar que:

$$(a) u_{ii} = 0 \text{ ou } u_{jj} = 0 \Rightarrow u_{ij} = 0$$

$$(b) u_{ij} = 1 \Rightarrow u_{ii} = 1 \text{ e } u_{jj} = 1$$

$$(c) u_{ii} = 1 \text{ e } u_{jj} = 1 \Rightarrow u_{ij} = 1$$

$$(d) u_{ij} = 0 \Rightarrow u_{ii} = 0 \text{ ou } u_{jj} = 0$$

As demonstrações seguem:

(a) Pela restrição (3), se $u_{jj} = 0$, teremos

$$\sum_{k=1}^n u_{kj} = 0 \Rightarrow u_{kj} = 0, \quad \forall k = 1, \dots, n,$$

pois todos os elementos são não negativos pela restrição (4). Agora, se $u_{ii} = 0$, pela restrição (1) teremos

$$\sum_{k=1}^n u_{ik} = \sum_{k=1}^n u_{ki} = 0 \Rightarrow u_{ik} = 0, \quad \forall k = 1, \dots, n.$$

(b) Se $u_{ij} = 1$ então, pela restrição (3),

$$\sum_{k=1}^n u_{kj} = n_1 u_{jj} \Rightarrow u_{ij} + \sum_{k \neq i} u_{kj} = n_1 u_{jj},$$

como $u_{ij} = 1$ e $\sum_{k \neq i} u_{kj} \geq 0$, teremos que $u_{jj} = 1$. Analogamente, teremos

$$\sum_{k=1}^n u_{ki} = n_1 u_{ii} \Rightarrow u_{ij} + \sum_{k \neq j} u_{ki} = n_1 u_{ii},$$

o que implica em $u_{ii} = 1$.

(c) Se $u_{ii} = 1$ e $u_{jj} = 1$, teremos pela restrição (3) que

$$u_{ij} + \sum_{k \neq i} u_{kj} = n_1.$$

Suponha por absurdo que $u_{ij} = 0$. Então existem n_1 elementos no conjunto $\Gamma = \{s : u_{sj} = 1, s \neq i\}$. Assim, pelo item (b), $u_{ss} = 1$ para todo $s \in \Gamma$. Logo, teremos

$$\sum_{k=1}^n u_{kk} = u_{ii} + \sum_{k \in \Gamma} u_{kk} = n_1 + 1,$$

o que contraria a restrição (2). Portanto, $u_{ij} = 1$.

(d) Segue imediatamente de (c), pois se $u_{ii} = 1 = u_{jj}$, deveríamos ter $u_{ij} = 1$.

Portanto, se $u_{ij} = 0$ devemos ter que $\min\{u_{ii}, u_{jj}\} = 0$.

□

A partir desta proposição, podemos enunciar o seguinte teorema.

Teorema 4.11. *O problema de programação inteira*

$$\max \sum_{i=1}^n \sum_{j=1}^n (q_{.i}q_{ij} - q_{.i}q_{.j}) u_{ij}$$

s.a.

$$u_{ij} = u_{ji}, \quad \forall i, j = 1, \dots, n.$$

$$\sum_{i=1}^n u_{ii} = n_1.$$

$$\sum_{i=1}^n u_{ij} = n_1 u_{jj}, \quad \forall j \in 1, \dots, n.$$

$$u_{ij} \in \{0, 1\} \forall i, j = 1, \dots, n.$$

é equivalente ao problema

$$\max \sum_{i=1}^n \sum_{j=1}^n (q_{.i}q_{ij} - q_{.i}q_{.j}) h_i h_j$$

s.a.

$$\sum_{i=1}^n h_i = n_1$$

$$h_i \in \{0, 1\} \forall i = 1, \dots, n.$$

4.4 Simulação

Vimos anteriormente que, sob determinadas condições, desenhos que não se utilizam da informação dos fatores de prognóstico para realizar a alocação dos pacientes produzem, assintoticamente, a mesma variância do contraste de uma ANOVA com um critério de classificação. Mostramos ainda que é possível analisar o desempenho de qualquer desenho com relação aos fatores de prognóstico através de um problema de programação inteira. Nesta seção, mostraremos estes resultados em um estudo de simulação.

4.4.1 Evitando problemas computacionais

Até o presente momento, nós assumimos que a variância do contraste podia ser calculada. Entretanto, não podemos calculá-la em alguns casos. Mais especificamente, não poderemos computá-la se

- (1) o número de pacientes for menor que $2+r(\mathbf{X})$, onde $r(\mathbf{X})$ é o posto da matriz \mathbf{X} .
- (2) \mathbf{h}^T é um vetor que contém somente zeros ou somente uns.

Para evitar parcialmente o problema em (1), optamos por utilizar apenas $n \geq 2 + r(\mathbf{X})$. Esta medida não evita a inexistência de $(\mathbf{H}\mathbf{Q}\mathbf{H})^{-1}$, uma vez que ainda existe a possibilidade de que $r((\mathbf{H}\mathbf{Q}\mathbf{H})^{-1}) = 1$ ou 0 . Para fatores de prognóstico contínuos, tomar o valor de n a partir de $2 + r(\mathbf{X})$ se mostrou uma estratégia satisfatória, mas observamos problemas para fatores de prognóstico discretos, o que nos fez optar por começar a simulação com tamanhos maiores de amostra nestes casos.

Em relação ao problema (2), sabemos que existirá $(\mathbf{H}^T\mathbf{H})^{-1}$ se existir pelo menos um indivíduo alocado em cada grupo. A probabilidade de não conseguirmos tal inversa é baixa em desenhos que procuram balancear o ensaio, mas é possível que isto ocorra quando o tamanho da amostra é pequeno. Assim, caso os $n - 1$ últimos pacientes tenham sido alocados em um único grupo, optamos por alocar o último paciente no outro grupo. Isto possui influência direta no viés de seleção. Seja $SB(\Delta)'$ o viés de seleção ao adotarmos esta estratégia e seja $w(n)$ a probabilidade de que os

$n - 1$ primeiros pacientes tenham sido alocados em um único grupo. Teremos

$$\begin{aligned} SB(\Delta)' &= w(n) \frac{\sum_{j=1}^{n-1} p^*(j) + 1}{n} + (1 - w(n)) \sum_{j=1}^n \frac{p^*(j)}{n} \\ &= SB(\Delta) + \frac{w(n)}{n} (1 - p^*(n)). \end{aligned} \quad (4.22)$$

Assim, teremos que $SB(\Delta)' \rightarrow SB(\Delta)$ quando $n \rightarrow \infty$. Além disso, é esperado de $w(n)/n$ convergir para zero rapidamente, uma vez que $w(n)$ deve diminuir com o aumento do número de pacientes no ensaio.

Para evitar a inexistência de $(\mathbf{X}^T \mathbf{X})^{-1}$, utilizaremos no lugar da inversa usual, a inversa generalizada de Moore-Penrose $(\mathbf{X}^T \mathbf{X})^-$, uma vez que $\mathbf{X}(\mathbf{X}^T \mathbf{X})^- \mathbf{X}^T$ sempre existe.

4.4.2 Resultados de Simulação

O efeito estimado dos fatores de prognósticos

Discutimos na Seção 4.2 que, sob determinadas condições da população em estudo e do desenho adotado, a variância do contraste dos tratamentos de um ensaio clínico do tipo caso-controle será assintoticamente equivalente à variância do contraste sem levar em consideração os fatores de prognóstico (ver o Teorema 4.7). Discutiremos aqui um estudo de simulação mostrando esta convergência, utilizando os desenhos apresentados na Seção 3.3

Fixamos um número n de pacientes. Para cada um dos pacientes, foram gerados q fatores de prognóstico segundo os cenários apresentados na Tabela 4.1. Para estes n pacientes, foram realizadas as alocações via os desenhos aleatório, sistemático alternado e da moeda viesada. Após os pacientes serem alocados, computamos e guardamos a seguinte estatística:

$$v(n, \Delta) = \frac{1 - n\bar{\mathbf{x}}^T \mathbf{C}\bar{\mathbf{x}}}{\frac{n}{n_2} (1 - n\bar{\mathbf{x}}^T \mathbf{C}\bar{\mathbf{x}}) \left(1 - n_1 \bar{\mathbf{x}}_{(1)}^T \mathbf{C}\bar{\mathbf{x}}_{(1)}\right) - \frac{n_1}{n_2} (1 - n\bar{\mathbf{x}}^T \mathbf{C}\bar{\mathbf{x}}_{(1)})^2}.$$

A estatística $v(n, \Delta)$ representa o efeito dos fatores de prognóstico na variância após a alocação sob o desenho Δ . O Teorema 4.7 diz que com o aumento de n , $v(n, \Delta)$ deve ser aproximar cada vez mais de um.

Cenário	Qtd. de Fatores	Distribuição dos Fatores	No. de Pacientes
1	4	Normal(0,1) independentes	$n = 10, 11, \dots, 500$
2	4	Normal(0,1) dependentes (com correlação 0,8)	$n = 10, 11, \dots, 500$
3	8	Normal(0,1) independentes	$n = 10, 11, \dots, 100$
4	4	Lognormal(0,1), Bernoulli(0,5) Normal ₂ ($\mathbf{0}, 0, 8\mathbf{J}_2 + 0, 2\mathbf{I}_2$)	$n = 20, 21, \dots, 100$

Tabela 4.1: Cenários de simulação

O processo de alocar os pacientes foi repetido 500 vezes, sendo que para cada alocação dos n pacientes, novos fatores de prognósticos foram gerados para cada cenário e cada valor de n fixado. Para a i -ésima repetição, consideramos $v_i(n, \Delta)$ o valor de $v(n, \Delta)$ observado para o tamanho de amostra n sob o desenho Δ . Com isto, para cada cenário e cada n fixado, calculamos a estatística

$$\bar{v}(n, \Delta) = \frac{1}{500} \sum_{i=1}^{500} v_i(n, \Delta),$$

utilizada como estimador de $\mathbb{E}[v(n, \Delta)]$. A coleção $\{\bar{v}(n, \Delta), n > n_0\}$ nos dará informações sobre a convergência desejada.

A Tabela 4.2 mostra os valores obtidos de \bar{v} para alguns valores de n no cenário 1. Podemos observar que o decaimento de \bar{v} para um é lento, embora a partir de $n = 15$ pareça ser razoável supor que v é menor que 2. A Tabela 4.3 mostra o cenário 2, no qual os fatores de prognóstico foram obtidos segundo uma distribuição normal multivariada com média zero, variâncias marginais iguais a um e correlação igual a 0,8. Podemos notar que os resultados das duas tabelas são semelhantes.

A Tabela 4.4 mostra os valores de $\bar{v}(n, \Delta)$ obtidos sob cenário 3. O alto valor de \bar{v} para os valores iniciais de n se deve ao elevado número de fatores de prognóstico. Podemos notar que a convergência (estimada) para o valor um é mais lenta que a convergência observada nos dois cenários anteriores.

A Tabela 4.5 mostra os valores de $\bar{v}(n, \Delta)$ para o cenário 4. Um dos fatores de prognóstico envolvidos possui distribuição Bernoulli. Por esse motivo, para n pequeno encontramos dificuldades para inverter as matrizes envolvidas. Conseguimos bons resultados para valores de n partir de 20. Para n maior que 30, parece ser

n	Aleatório	Sistemático	Moeda Viesada
10	2,381	2,422	2,326
15	1,530	1,512	1,493
20	1,313	1,283	1,317
40	1,123	1,122	1,122
60	1,055	1,057	1,054
100	1,045	1,043	1,044
200	1,019	1,020	1,021
500	1,007	1,008	1,008

Tabela 4.2: Valores de \bar{v} para o cenário 1.

n	Aleatório	Sistemático	Moeda Viesada
10	2,499	2,421	2,320
15	1,504	1,513	1,505
20	1,299	1,312	1,314
40	1,123	1,117	1,125
60	1,076	1,075	1,077
100	1,046	1,043	1,045
200	1,020	1,021	1,019
500	1,008	1,007	1,008

Tabela 4.3: Valores de \bar{v} para o cenário 2.

n	Aleatório	Sistemático	Moeda Viesada
10	4614,96	625,77	153772,84
15	2,92	3,05	3,15
20	1,87	1,89	1,88
40	1,28	1,27	1,27
60	1,16	1,17	1,17
100	1,09	1,09	1,09

Tabela 4.4: Valores de \bar{v} para o cenário 3.

n	Aleatório	Sistemático	Moeda Viesada
20	1.307	1.300	1.310
30	1.175	1.172	1.169
40	1.126	1.120	1.122
50	1.091	1.095	1.095
60	1.076	1.072	1.074
100	1.044	1.043	1.044

Tabela 4.5: Valores de \bar{v} para o cenário 4.

razoável supor que o $\mathbb{E}[v(n, \Delta)]$ é menor que 1,2.

Avaliação do desempenho dos desenhos aleatório, sistemático e da moeda viesada para uma população conhecida

Estamos interessados em avaliar o quão bem os desenhos aleatório, sistemático alternado e da moeda viesada minimizam a variância do estimador da diferença entre tratamentos. Já sabemos que, com o aumento do tamanho da amostra, ambos os desenhos terão esta variância próxima de $4/n$. Entretanto, conforme discutido na Seção 4.3, podemos avaliar a eficiência destes desenhos em uma população conhecida.

Consideramos uma população de pacientes fixa, com seus respectivos fatores de prognóstico representada pela matriz \mathbf{X} . Esta população foi composta de 24 pacientes com um fator de prognóstico, simulado a partir da distribuição normal padrão. Um ensaio clínico é realizado visando alocar estes pacientes em dois grupos, sendo que o balanço entre os grupos de tratamentos é desejado. O software R foi utilizado para fazer esta simulação, e a semente utilizada foi 123.

Neste caso, um observador externo possui acesso a todas as informações destes indivíduos e pode, antes de começar o experimento, encontrar a alocação ótima de cada um dos pacientes, como mostrado no Teorema 4.11, com a restrição de que cada grupo tenha exatamente 12 pacientes. Esta alocação será nosso referencial.

Para avaliar os desenhos, propomos o seguinte algoritmo:

1. Calculemos a alocação ótima dos n pacientes. Denotemos este valor por f^* .
2. Para cada iteração:

- (a) Escolha ao acaso uma permutação dos n pacientes.
- (b) Faça a alocação sequencial destes pacientes segundo os três desenhos e calcule o valor da função objetivo destas alocações. Faça $f_i(\Delta)$ o valor da função objetivo na i -ésima iteração da alocação utilizando o desenho Δ .
- (c) Guarde o valor $f_i(\Delta)/f^*$

A Figura 4.1 mostra os valores f_i/f^* para este estudo. As regiões em cinza na figura destacam os valores observados na simulação que possuem um valor de f_i/f^* superior ou igual a 0,9. A área destas regiões correspondem a, respectivamente, 66%, 86% e 83% dos dados. Notemos que, mais uma vez, os desenhos que forçam o balanço tendem a atingir resultados mais próximos do ótimo.

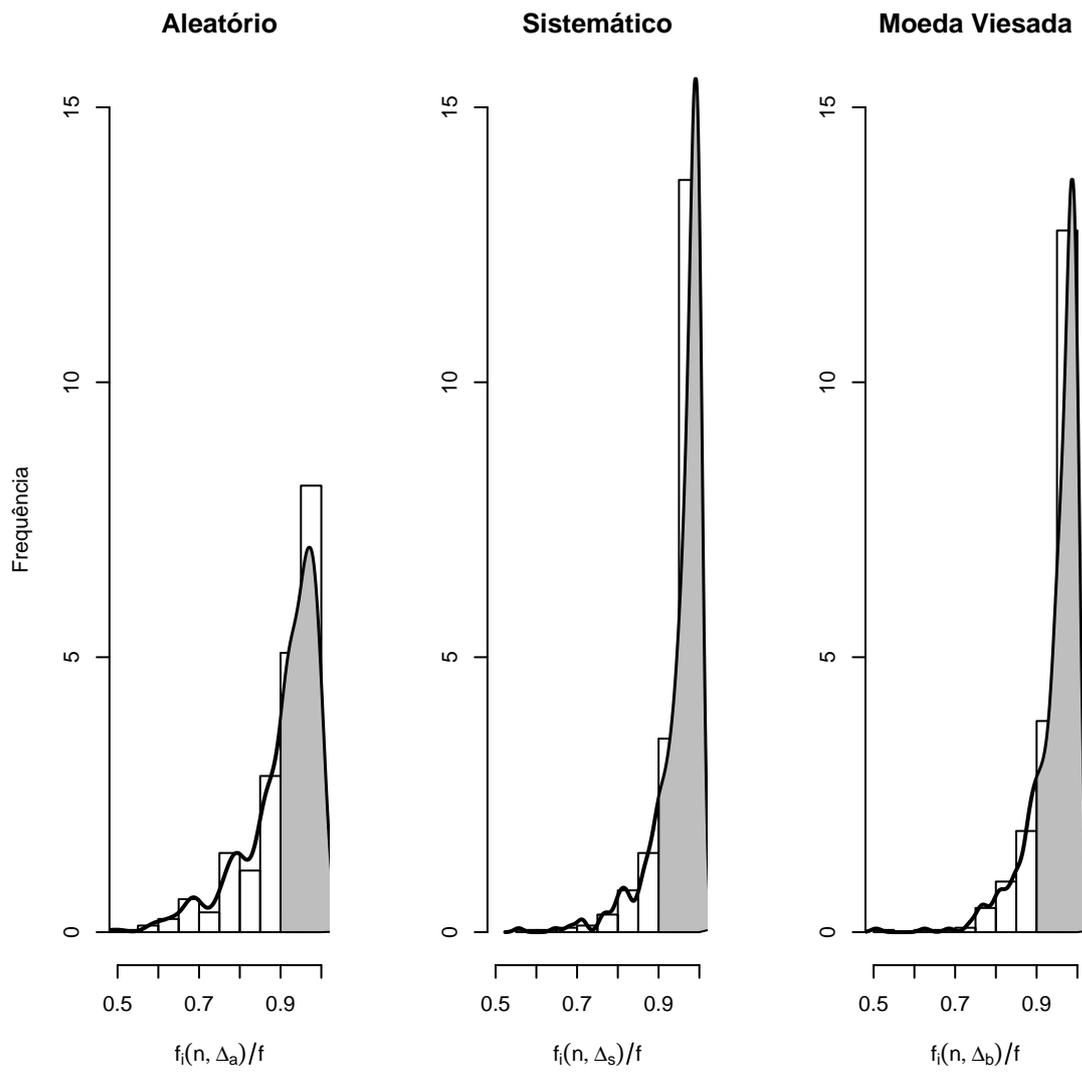


Figura 4.1: Valores de $f_i(n, \Delta)/f^*$ em um estudo de simulação

Capítulo 5

Considerações Finais

Em ensaios clínicos, estamos interessados em determinar o efeito de certo tratamento sob uma variável resposta. Sabemos que existem covariáveis (fatores de prognóstico) que podem influenciar no resultado, e devemos levá-los em conta na análise, mesmo que seja para constar posteriormente a sua baixa influência.

A questão principal tratada aqui está em como alocar os pacientes nos grupos caso/controle sabendo que existem tais fatores. Se o objetivo final do ensaio está em obter uma estimativa da diferença entre os efeitos dos dois tratamentos, então sob o ponto de vista da inferência clássica, devemos procurar por um desenho que nos forneça estimativas não viesadas e com menor variância para estes efeitos.

Se um modelo linear for utilizado para analisar o efeito dos tratamentos, então podemos mostrar que o estimador para a diferença entre tratamentos é não viesado, independente da escolha do desenho. Mostramos neste trabalho que a variância deste estimador é proporcional a

$$\frac{n}{n_1 n_2} v(n, \Delta)$$

onde n_1 e n_2 são o número de pacientes em cada um dos grupos, $n = n_1 + n_2$ e $v(n, \Delta)$ é o efeito dos fatores de prognóstico em conjunto com a alocação.

Do que foi exposto até o momento, a escolha do desenho deve ser feita de modo a minimizar a variância do estimador da diferença entre tratamentos. Consideremos então as seguintes propriedades:

1. Os pacientes são independentes entre si;
2. $\lim_{n \rightarrow \infty} p_T(i) = \pi_T$;

3. O desenho Δ não depende de covariáveis.

Se um desenho possui estas propriedades, teremos que $v(n, \Delta) \rightarrow 1$ quando $n \rightarrow \infty$. Deste modo, existe n_0 tal que para todo $n > n_0$ a variância do estimador da diferença entre os tratamentos será assintoticamente equivalente a n/n_1n_2 e, deste modo, sua minimização implica em fazer $n_1 = n_2$.

Portanto, desenhos que não se utilizam de fatores de prognóstico e que são construídos para forçar o balanço possuem a propriedade de depender cada vez menos dos fatores de prognóstico na variância da diferença entre os efeitos dos tratamentos a medida que o tamanho da amostra cresce.

A decisão de escolher desenhos que possuem as propriedades enumeradas acima não implica que não existem desenhos que produzem variâncias menores utilizando a informação dos fatores de prognóstico. Entretanto, neste trabalho mostramos que é possível encontrar a alocação ótima para qualquer matriz de fatores de prognóstico fixada. Com isto, pode-se avaliar a eficácia de qualquer desenho face a um banco de dados específico. Em nosso estudo de simulação, mostramos que para um único fator de prognóstico com distribuição normal padrão, o desenho da moeda viesada produz uma eficiência acima de 90% com uma probabilidade aproximada de 0,83. Um resultado análogo pode ser obtido com o desenho sistemático. Entretanto, o desenho da moeda viesada possui a vantagem de minimizar o viés de seleção e deveria ser preferível ao desenho sistemático neste sentido.

Do que foi exposto até então, parece razoável utilizar desenhos mais simples, que não levam em consideração os fatores de prognóstico. A praticidade de implementação e seus bons resultados assintóticos ao lidar com fatores de prognósticos fazem com que desenhos como o da moeda viesada sejam boas escolhas para alocar sequencialmente os pacientes em ensaios clínicos.

Por último, vale ressaltar que o problema de programação inteira em 4.11 cresce polinomialmente. Alternativas considerando a relaxação quadrática do problema original estão sendo exploradas e serão abordadas em trabalhos futuros.

Parte II

Considerações sobre o teste de
Vuong e seu desempenho contra o
teste de Cox para discriminar
modelos pertencentes à famílias
separadas

Capítulo 6

Introdução

6.1 Um breve histórico sobre testes de hipóteses para discriminar modelos pertencentes a famílias separadas

Consideremos as famílias de distribuições $\mathcal{F} = \{f(y|\alpha) : \alpha \in A\}$ e $\mathcal{G} = \{g(y|\beta) : \beta \in B\}$. Estas famílias serão ditas separadas se um membro arbitrário de uma destas não puder ser obtido como limite de um membro da outra. Por exemplo, a família de distribuições lognormais é separada da família de distribuições Weibull, pois estas não possuem nenhum membro em comum. Já a família de distribuições gama não é separada da família de distribuições exponenciais, pois esta última está contida na primeira.

A primeira discussão sobre testes para a seleção entre modelos pertencentes à famílias separadas foi apresentada em Cox (1961) e Cox (1962). O teste apresentado nestes trabalhos consiste em estudar o comportamento assintótico da razão de verossimilhanças de dois modelos sob a hipótese de que um deles é verdadeiro.

Os trabalhos pioneiros de Cox deram origem aos estudos de testes de hipóteses para discriminar entre modelos pertencentes a famílias separadas. Análises empíricas do poder e nível do teste de Cox para alguns modelos são dadas em Pereira (1976). Em Pesaran (1982) são discutidas aproximações assintóticas de Bahadur's para estimar o poder do teste. Aplicações em modelos lineares, modelos não-lineares multivariados e modelos de sobrevivência podem ser vistos em Pesaran (1974), Pereira

(1976) e Pesaran and Deaton (1978).

A maior crítica em relação ao teste de Cox está na dificuldade de se obter as expressões analíticas necessárias para a realização do teste. Uma alternativa computacional para evitar o cálculo das estatísticas do teste de Cox pode ser vista em Wei-Yin (1985), onde o autor sugere o uso do método bootstrap. Outro procedimento baseado em simulação é dado em Lu et al. (2008). Outros trabalhos seguiram uma abordagem semelhante (M. H. Pesaran (1993), Coulibaly and Brorsen (1999)). Um exemplo de aplicação, utilizando essa abordagem computacional pode ser visto em Dameus et al. (2002).

Outros meios para selecionar modelos pertencentes à famílias separadas surgiram. Um estudo sobre o uso de fatores de Bayes para este fim pode ser visto em Araujo and Pereira (2007). Uma abordagem não-paramétrica é dada em Clarke (2003), onde o autor utiliza para criar um teste do sinal a partir da razão de verossimilhança para cada observação.

As condições de regularidade para o teste de Cox foram apresentadas, formalmente, em White (1982). Este trabalho serviu de base para que Vuong construísse um teste assintótico para famílias separadas (Vuong (1989)). Este teste possui um artifício elegante, a criação de uma hipótese que diz que ambos os modelos são verossímeis para os dados. Sob esta hipótese, a criação de uma estatística de teste com distribuição conhecida se torna mais simples e de fácil implementação em softwares estatísticos. Algumas aplicações deste teste podem ser vistas em Hu et al. (2010) e Clarke and Signorino (2010).

6.2 Objetivos e organização

O desempenho do teste de Cox já foi discutido em uma série de estudos de simulação, como Pereira (1978), Pereira (1979) e Roja et al. (2004). Entretanto, em nossa pesquisa bibliográfica encontramos poucas comparações entre o desempenho do teste de Cox e o de Vuong. Talvez isto tenha ocorrido pela natureza dos testes: a hipótese nula do teste de Vuong, sob qual as decisões são tomadas, não coincide com nenhuma das duas hipóteses do teste de Cox. Além disso, se os dados forem realmente provenientes de um dos dois modelos, o teste de Vuong não pode ser

avaliado tradicionalmente (discutiremos este tópico na Seção 8.1).

Neste trabalho mostramos o desempenho dos dois testes para algumas famílias de distribuições separadas. No Capítulo 7 apresentamos os dois testes. No Capítulo 8 apresentamos os nossos resultados. Primeiro, discutimos que o desempenho do teste de Vuong pode ser variável, apresentando alguns exemplos. Em seguida, apresentamos um estudo de simulação para comparar o desempenho dos dois testes.

Capítulo 7

Os Testes de Cox e Vuong

7.1 A base dos testes de Cox e Vuong

Sejam $f(\cdot|\alpha)$ e $g(\cdot|\beta)$ dois modelos pertencentes à famílias separadas. Para uma amostra $\mathbf{Y} = \{Y_1, \dots, Y_n\}$ de variáveis aleatórias independentes e identicamente distribuídas, gostaríamos de testar as seguintes hipóteses:

- H_f : os dados são provenientes do modelo $f(\cdot|\alpha)$.
- H_g : os dados são provenientes do modelo $g(\cdot|\beta)$.

Naturalmente, pode ocorrer que nenhuma das duas hipóteses seja verdadeira. Portanto, seja h o modelo gerador de \mathbf{Y} . Neste caso, estamos interessados em saber qual dos dois modelos se aproxima mais do modelo h . A adequácia do modelo f pode ser medida apropriadamente pelo Critério de Informação de Kullback-Leibler (KLIC),

$$KLIC(f(\cdot|\alpha); h) = \mathbb{E}_H \left[\log \frac{h(\mathbf{Y})}{f(\mathbf{Y}|\alpha)} \right] = \int \log \frac{h(\mathbf{Y})}{f(\mathbf{Y}|\alpha)} dH(y). \quad (7.1)$$

É conhecido que $KLIC(f; h) \geq 0$, com a igualdade ocorrendo somente quando $F(\mathbf{Y}|\alpha) = H(\mathbf{Y})$ em quase toda a parte no suporte de \mathbf{Y} . Assim, um modelo f com um valor de $KLIC(f; h)$ próximo de zero seria um candidato adequado para representar o modelo gerador dos dados. Notemos ainda que

$$KLIC(f(\cdot|\alpha); h) = \int \log h(\mathbf{Y}) dH(\mathbf{Y}) - \int \log f(\mathbf{Y}|\alpha) dH(\mathbf{Y}). \quad (7.2)$$

Deste modo, minimizar $KLIC(f(\cdot|\alpha); h)$ sob α , implica em encontrar α_0 tal que

$$\mathbb{E}_H [\log f(\mathbf{Y}|\alpha_0)] = \sup_{\alpha} \mathbb{E}_H [\log f(\mathbf{Y}|\alpha)]. \quad (7.3)$$

O valor α_0 é denominado *parâmetro pseudo-verdadeiro*. Toda a discussão acima poderia ser feita para o modelo $g(\cdot|\beta)$, onde o menor valor para o critério de informação de Kullback-Leibler seria obtido por

$$KLIC(g(\cdot|\beta_0); h) = \mathbb{E}_H [\log h(\mathbf{Y})] - \sup_{\beta} \mathbb{E}_H [\log g(\mathbf{Y}|\beta)]. \quad (7.4)$$

Assim, o modelo f será preferível ao modelo g se

$$KLIC(f(\mathbf{Y}|\alpha_0); h) \leq KLIC(g(\mathbf{Y}|\beta_0); h), \quad (7.5)$$

o que é equivalente a

$$\mathbb{E}_H [\log f(\mathbf{Y}|\alpha_0)] \geq \mathbb{E}_H [\log g(\mathbf{Y}|\beta_0)] \Leftrightarrow \mathbb{E}_H \left[\log \frac{f(\mathbf{Y}|\alpha_0)}{g(\mathbf{Y}|\beta_0)} \right] \geq 0. \quad (7.6)$$

Portanto, as hipóteses H_f e H_g podem ser avaliadas pela quantidade

$$\mathbb{E}_H \left[\log \frac{f(\mathbf{Y}|\alpha_0)}{g(\mathbf{Y}|\beta_0)} \right]. \quad (7.7)$$

Supondo inicialmente que o modelo h é conhecido, a estatística natural para discriminar entre os dois modelos competidores seria,

$$\log \frac{f(\mathbf{Y}|\alpha_0)}{g(\mathbf{Y}|\beta_0)} = \sum_{i=1}^n \log f(Y_i|\alpha_0) - \sum_{i=1}^n \log g(Y_i|\beta_0). \quad (7.8)$$

Entretanto, o modelo h é desconhecido, o que implica que não conhecemos o valor de α_0 e β_0 . Contudo, sob certas condições de regularidade (Vuong (1989)), teremos que os estimadores de máxima verossimilhança (EMV) de α e β , representados por $\hat{\alpha}$ e $\hat{\beta}$ respectivamente, convergem em probabilidade para α_0 e β_0 . Assim, podemos construir a seguinte estatística de teste para a discriminação entre os dois modelos:

$$\log \frac{f(\mathbf{Y}|\hat{\alpha})}{g(\mathbf{Y}|\hat{\beta})} = \sum_{i=1}^n \log f(Y_i|\hat{\alpha}) - \sum_{i=1}^n \log g(Y_i|\hat{\beta}) = \sum_{i=1}^n \log \frac{f(Y_i|\hat{\alpha})}{g(Y_i|\hat{\beta})}. \quad (7.9)$$

Ambos os testes de Cox e Vuong utilizam a estatística da Equação (7.9) em seus testes, embora de modo diferente. Em ambos os casos, a distribuição da estatística de teste é dada pelo Teorema 7.1, enunciado abaixo.

Teorema 7.1. *Sob certas condições de regularidade, se $f(\cdot|\alpha_0) \neq g(\cdot|\beta_0)$, teremos que*

$$\sqrt{n} \left\{ \frac{1}{n} \sum_{i=1}^n \log \frac{f(Y_i|\hat{\alpha})}{g(Y_i|\hat{\beta})} - \mathbb{E}_H \left[\log \frac{f(Y_1|\alpha_0)}{g(Y_1|\beta_0)} \right] \right\} \xrightarrow{D} N(0, \omega^2),$$

onde α_0 e β_0 são os limites em probabilidade de $\hat{\alpha}$ e $\hat{\beta}$ considerando o verdadeiro modelo $h(\cdot)$ e

$$\omega^2 = \mathbb{E}_H \left[\log \frac{f(Y_1|\alpha_0)}{g(Y_1|\beta_0)} \right]^2 - \left[\mathbb{E}_H \left(\log \frac{f(Y_1|\alpha_0)}{g(Y_1|\beta_0)} \right) \right]^2$$

Demonstração. Ver Vuong (1989). □

7.2 O Teste de Cox

Seja $\mathbf{Y} = \{Y_1, \dots, Y_n\}$ uma amostra de variáveis aleatórias independentes e identicamente distribuídas. Sejam $f(\cdot|\alpha)$ e $g(\cdot|\beta)$ dois modelos pertencentes à famílias de distribuições separadas. Sejam H_f e H_g as hipóteses nula e alternativa, respectivamente, onde H_f afirma que o modelo $f(\cdot|\alpha)$ é verdadeiro e H_g afirma que o modelo $g(\cdot|\beta)$ é verdadeiro. Considere que apenas um dos dois modelos deve ser o verdadeiro modelo gerador de Y . Começaremos a construção do teste de Cox com a seguinte estatística para testar H_f :

$$T_f^*(C) = \sum_{i=1}^n \log \frac{f(Y_i|\hat{\alpha})}{g(Y_i|\hat{\beta})} - n \mathbb{E}_F \left[\log \frac{f(Y_1|\alpha)}{g(Y_1|\beta_0)} \right]_{\alpha=\hat{\alpha}}, \quad (7.10)$$

onde β_0 é o limite em probabilidade do estimador $\hat{\beta}$ sob a hipótese H_f (o último termo da expressão acima diz que primeiro devemos calcular o valor esperado e depois substituir os parâmetros pelos seus respectivos estimadores de máxima verossimilhança). Vamos mostrar que essa estatística possui distribuição assintótica normal. Como $\hat{\alpha}$ é o EMV para α teremos que, sob certas condições de regularidade,

$$\mathbb{E}_F \left[\log \frac{f(Y_1|\alpha)}{g(Y_1|\beta_0)} \right]_{\alpha=\hat{\alpha}} \xrightarrow{P} \mathbb{E}_F \left[\log \frac{f(Y_1|\alpha)}{g(Y_1|\beta_0)} \right], \quad (7.11)$$

quando $n \rightarrow \infty$, onde a convergência é da ordem de $1/n$. Assim,

$$\begin{aligned}
\frac{T_f^*(C)}{\sqrt{n}} &= \frac{1}{\sqrt{n}} \left(\sum_{i=1}^n \log \frac{f(Y_i|\hat{\alpha})}{g(Y_i|\hat{\beta})} - n\mathbb{E}_F \left[\log \frac{f(Y_1|\alpha)}{g(Y_1|\beta_0)} \right]_{\alpha=\hat{\alpha}} \right) \\
&= \frac{1}{\sqrt{n}} \left(\frac{1}{n} \sum_{i=1}^n \log \frac{f(Y_i|\hat{\alpha})}{g(Y_i|\hat{\beta})} - n\mathbb{E}_F \left[\log \frac{f(Y_1|\alpha)}{g(Y_1|\beta_0)} \right]_{\alpha=\hat{\alpha}} \pm n\mathbb{E}_F \left[\log \frac{f(Y_1|\alpha)}{g(Y_1|\beta_0)} \right] \right) \\
&= \sqrt{n} \left\{ \frac{1}{n} \sum_{i=1}^n \log \frac{f(Y_i|\hat{\alpha})}{g(Y_i|\hat{\beta})} - \mathbb{E}_F \left[\log \frac{f(Y_1|\alpha)}{g(Y_1|\beta_0)} \right] \right\} \\
&\quad + \sqrt{n} \left\{ \mathbb{E}_F \left[\log \frac{f(Y_1|\alpha)}{g(Y_1|\beta_0)} \right]_{\alpha=\hat{\alpha}} - \mathbb{E}_F \left[\log \frac{f(Y_1|\alpha)}{g(Y_1|\beta_0)} \right] \right\}. \tag{7.12}
\end{aligned}$$

O primeiro termo entre chaves converge em distribuição para uma distribuição normal, pelo Teorema 7.1. O segundo termo entre chaves converge para zero em probabilidade sob H_f . Portanto, pelo Teorema de Slutsky teremos que

$$\frac{T_f^*(C)}{\sqrt{n}} \xrightarrow{D} N(0, \omega^2). \tag{7.13}$$

Para podermos padronizar a estatística $T_f^*(C)$ para obter a distribuição normal padrão, precisamos encontrar um estimador consistente para ω^2 . Para tanto, consideremos os seguintes lemas:

Lema 7.2. *Seja $L_f(\alpha) = \log f(\mathbf{Y}|\alpha)$. Seja $U(\alpha) = \frac{\partial}{\partial \alpha} L_f(\alpha) = L'_f(\alpha)$. Então*

1. $U(\hat{\alpha}) = 0$
2. $\mathbb{E}[U(\alpha)] = 0$
3. $\frac{\partial}{\partial \alpha} U(\alpha) = -nI(\alpha)$, onde $I(\alpha)$ é Informação de Fisher.

Lema 7.3. *Seja $\mathbb{E}(Z) = \beta^T W$, onde Z é um vetor de variáveis aleatórias independentes e identicamente distribuídas. Então,*

1. O melhor preditor linear para Z será $\hat{Z} = \mathbb{C}(Z, W)[\mathbb{V}(W)]^{-1}W$.
2. $\mathbb{V}(\hat{Z}) = \mathbb{C}(Z, W)[\mathbb{V}(W)]^{-1}\mathbb{C}(Z, W)$
3. $\mathbb{V}(Z - \mathbb{C}(Z, W)[\mathbb{V}(W)]^{-1}W) = \mathbb{V}(Z) - \mathbb{V}(\hat{Z})$.

Antes de encontrar um estimador para ω^2 , encontraremos uma aproximação para a estatística $T_f^*(C)$. Consideremos a expansão de $L_f(\alpha)$ em Série de Taylor em torno de $\hat{\alpha}$

$$L_f(\alpha) = L_f(\hat{\alpha}) + (\hat{\alpha} - \alpha)L'_f(\hat{\alpha}) \approx L_f(\hat{\alpha}). \tag{7.14}$$

Podemos aproximar $\mathbb{E}_F [L_f(\alpha)]|_{\alpha=\hat{\alpha}}$ por um valor de α arbitrário em sua vizinhança como segue

$$\mathbb{E}_F [L_f(\alpha)]|_{\alpha=\hat{\alpha}} = \mathbb{E}_F [L_f(\alpha)] + (\hat{\alpha} - \alpha) \frac{\partial}{\partial \alpha} \mathbb{E}_F [L_f(\alpha)]. \quad (7.15)$$

Agora, notemos que

$$\begin{aligned} \frac{\partial}{\partial \alpha} \mathbb{E}_F [L_f(\alpha)] &= \int \frac{\partial}{\partial \alpha} [f(\mathbf{Y}|\alpha) \log f(\mathbf{Y}|\alpha)] d\mathbf{Y} \\ &= \int \left[\frac{\partial}{\partial \alpha} f(\mathbf{Y}|\alpha) + \log f(\mathbf{Y}|\alpha) \frac{\partial}{\partial \alpha} f(\mathbf{Y}|\alpha) \right] d\mathbf{Y} \\ &= \int \left[\frac{\partial}{\partial \alpha} \log f(\mathbf{Y}|\alpha) + \log f(\mathbf{Y}|\alpha) \frac{\partial}{\partial \alpha} \log f(\mathbf{Y}|\alpha) \right] f(\mathbf{Y}|\alpha) d\mathbf{Y} \\ &= \mathbb{E}_F \left[\log f(\mathbf{Y}|\alpha) \frac{\partial}{\partial \alpha} \log f(\mathbf{Y}|\alpha) \right] + \mathbb{E}_F [U(\alpha)] \\ &= \mathbb{E}_F \left[\log f(\mathbf{Y}|\alpha) \frac{\partial}{\partial \alpha} \log f(\mathbf{Y}|\alpha) \right] = \mathbb{C}_F \left(L_f(\alpha), \frac{\partial}{\partial \alpha} \log f(\mathbf{Y}|\alpha) \right) \\ &= \mathbb{C}_F \left(L_f(\alpha), L'_f(\alpha) \right) \end{aligned} \quad (7.16)$$

e que

$$L'_f(\alpha) \approx L'_f(\hat{\alpha}) + (\hat{\alpha} - \alpha) L''_f(\hat{\alpha}) \Rightarrow \hat{\alpha} - \alpha \approx \frac{L'(\alpha)}{L''(\hat{\alpha})} = -\frac{L'(\alpha)}{nJ(\hat{\alpha})}. \quad (7.17)$$

Utilizando as Equações 7.14, 7.15, 7.16 e 7.17, teremos

$$L_f(\hat{\alpha}) - \mathbb{E}_F [L_f(\hat{\alpha})]|_{\hat{\alpha}=\alpha} = L_f(\alpha) - \mathbb{E}_F [L_f(\alpha)] + \frac{L'_f(\alpha)}{nJ(\hat{\alpha})} \mathbb{C} \left(L_f(\alpha), L'_f(\alpha) \right). \quad (7.18)$$

O mesmo raciocínio pode ser aplicado para $L_g(\beta) = \log g(\mathbf{Y}|\beta)$, obtendo:

$$L_g(\beta_0) \approx L_g(\hat{\beta}) \quad (7.19)$$

$$\begin{aligned} \mathbb{E}_F [L_g(\beta_0)]|_{\alpha=\hat{\alpha}} &= \mathbb{E}_F [L_g(\beta_0)] + (\hat{\alpha} - \alpha) \frac{\partial}{\partial \alpha} \mathbb{E}_F [L_g(\beta_0)] \\ &= \mathbb{E}_F [L_g(\beta_0)] - \frac{L'_f(\alpha)}{nJ(\alpha)} \frac{\partial}{\partial \alpha} \mathbb{E}_F [L_g(\beta_0)]. \end{aligned} \quad (7.20)$$

Uma vez que

$$\begin{aligned} \frac{\partial}{\partial \alpha} \mathbb{E}_F [L_g(\beta_\alpha)] &= \int \frac{\partial}{\partial \alpha} [\log g(\mathbf{Y}|\beta_\alpha) f(\mathbf{Y}|\alpha)] d\mathbf{Y} \\ &= \int f(\mathbf{Y}|\alpha) \frac{\partial}{\partial \alpha} g(\mathbf{Y}|\beta_\alpha) + \log g(\mathbf{Y}|\beta_\alpha) \frac{\partial}{\partial \alpha} f(\mathbf{Y}|\alpha) d\mathbf{Y} \\ &= \int \left[\frac{\partial}{\partial \alpha} g(\mathbf{Y}|\beta_\alpha) + \log g(\mathbf{Y}|\beta_\alpha) \frac{\partial}{\partial \alpha} \log f(\mathbf{Y}|\alpha) \right] f(\mathbf{Y}|\alpha) d\mathbf{Y} \\ &= \mathbb{E}_F \left[\log g(\mathbf{Y}|\beta_\alpha) \frac{\partial}{\partial \alpha} \log f(\mathbf{Y}|\alpha) \right] = \mathbb{C}_F \left(L_g(\beta_\alpha), L'_f(\alpha) \right), \end{aligned} \quad (7.21)$$

teremos

$$L_g(\hat{\beta}) - \mathbb{E}_F [L_g(\beta_\alpha)]_{\alpha=\hat{\alpha}} = L_g(\beta_\alpha) - \mathbb{E}_F [L_g(\beta_\alpha)] + \frac{L'_f(\alpha)}{nJ(\hat{\alpha})} \mathbb{C}_F \left(L_g(\beta_\alpha), L'_f(\alpha) \right). \quad (7.22)$$

Teremos que a estatística de teste $T_f^*(C)$ poderá ser reescrita a partir das Equações (7.18) e (7.22), conforme segue:

$$\begin{aligned} T_f^*(C) &= \log \frac{f(\mathbf{Y}|\alpha)}{g(\mathbf{Y}|\beta_\alpha)} - n\mathbb{E}_F \left[\frac{f(Y|\alpha)}{g(Y_1|\beta_\alpha)} \right] \\ &\quad + \frac{L'_f(\alpha)}{nJ(\hat{\alpha})} \left[\mathbb{C}_F \left(L_f(\alpha), L'_f(\alpha) \right) - \mathbb{C}_F \left(L_g(\beta_\alpha), L'_f(\alpha) \right) \right] \\ &= \log \frac{f(\mathbf{Y}|\alpha)}{g(\mathbf{Y}|\beta_\alpha)} - n\mathbb{E}_F \left[\frac{f(Y|\alpha)}{g(Y_1|\beta_\alpha)} \right] \\ &\quad + \frac{1}{n} \mathbb{V}_F(L'_f(\alpha))^{-1} \mathbb{C}_F \left(L_f(\alpha) - L_g(\beta_\alpha), L'_f(\alpha) \right) L'_f(\alpha) \end{aligned}$$

pois $J(\hat{\alpha}) \approx \mathbb{V}(L'_f(\alpha))$. Agora, notando que o último termo na expressão acima é o melhor preditor linear para $L_f(\alpha) - L_g(\beta_\alpha)$, podemos utilizar o Lema 7.3 para mostrar que

$$\mathbb{V} [T_f^*(C)] = \mathbb{V}_F \left[\log \frac{f(\mathbf{Y}|\alpha)}{g(\mathbf{Y}|\beta_\alpha)} \right] - \mathbb{C}_F \left(\log \frac{f(\mathbf{Y}|\alpha)}{g(\mathbf{Y}|\beta_\alpha)}, \frac{\partial}{\partial \alpha} \log f(\mathbf{Y}|\alpha) \right) \frac{\partial}{\partial \alpha} \log f(\mathbf{Y}|\alpha) \quad (7.23)$$

O estimador para ω^2 será dado por $\mathbb{V}[T_f^*(C)]$, substituindo o parâmetro α por $\hat{\alpha}$. Denotemos esta estatística por $\hat{\omega}^2$. Assim, a estatística para o teste de Cox será dada por

$$T_f(C) = \frac{T_f^*(C)}{\sqrt{\mathbb{V}[T_f^*(C)]}}. \quad (7.24)$$

Se H_f for verdadeira, então $T_f(C)$ terá distribuição normal padrão. Pode-se mostrar que sob H_g teremos que $T_f(C)$ terá média negativa (basta notar que a última chave na Equação 7.12 será negativa). Assim, fixado um nível de significância (assintótico) γ , teremos que a região de rejeição do teste de Cox será dada por $R = \{\mathbf{Y} : T_f(C) < z_\gamma\}$.

Quando H_g é a hipótese nula e H_f a alternativa, resultados análogos são obtidos produzindo a estatística $T_g(C)$. Uma abordagem alternativa para realização do teste é calcular $T_f(C)$ e $T_g(C)$, e decidir sobre a rejeição ou não das hipóteses segundo os possíveis pares de resultados (ver Pereira (1998)).

7.3 O Teste de Vuong

Seja $h(\cdot)$ o verdadeiro (e desconhecido) modelo gerador de \mathbf{Y} e sejam $f(\cdot|\alpha)$ e $g(\cdot|\beta)$ os dois modelos competidores pertencentes à famílias separadas. As hipóteses do teste de Vuong são:

$$H_0 : \mathbb{E}_H \left[\log \frac{f(Y_1|\alpha_0)}{g(Y_1|\beta_0)} \right] = 0, \quad (7.25)$$

implicando que f e g são equivalentes, contra

$$H_f : \mathbb{E}_H \left[\log \frac{f(Y_1|\alpha_0)}{g(Y_1|\beta_0)} \right] > 0, \quad (7.26)$$

implicando que f é melhor que g , ou

$$H_g : \mathbb{E}_H \left[\log \frac{f(Y_1|\alpha_0)}{g(Y_1|\beta_0)} \right] < 0, \quad (7.27)$$

implicando que f é pior que g .

As hipóteses do teste de Vuong são baseadas no Critério de Informação de Kullback-Leibler. Notemos que, sob a hipótese H_0 , os modelos f e g possuem a mesma aproximação do modelo h , carregando portanto a mesma informação.

A estatística para o teste assintótico de Vuong é

$$T(V) = \frac{1}{\hat{\omega}\sqrt{n}} \log \frac{f(\mathbf{Y}|\hat{\alpha})}{g(\mathbf{Y}|\hat{\beta})}, \quad (7.28)$$

onde

$$\hat{\omega}^2 = \frac{1}{n} \sum_{i=1}^n \left[\log \frac{f(Y_i|\hat{\alpha})}{g(Y_i|\hat{\beta})} \right]^2 - \left[\frac{1}{n} \sum_{i=1}^n \log \frac{f(Y_i|\hat{\alpha})}{g(Y_i|\hat{\beta})} \right]^2. \quad (7.29)$$

A estatística de teste é o valor observado da razão de verossimilhanças, com os parâmetros substituídos por seus respectivos estimadores de máxima verossimilhança. O seguinte teorema nos dá condições para a criação de um teste direcional para decidir entre as hipóteses.

Teorema 7.4. *Sob certas condições de regularidade, quando $n \rightarrow \infty$,*

1. *Sob H_0 : $T(V) \xrightarrow{D} N(0, 1)$;*
2. *Sob H_f : $T(V) \xrightarrow{q.c.} +\infty$;*
3. *Sob H_g : $T(V) \xrightarrow{q.c.} -\infty$.*

O Teorema 7.4 é a aplicação direta do Teorema 7.1. Sob H_0 teremos que $T(V) \sim \mathcal{N}(0, 1)$. Entretanto, como a estatística de teste não está padronizada, teremos que $T(V)$ deverá divergir sob as outras duas hipóteses pois, para qualquer n finito suficientemente grande,

$$\mathbb{E}[T(V)] \approx \sqrt{n} \mathbb{E}_H \left[\log \frac{f(Y_1|\alpha_0)}{g(Y_1|\beta_0)} \right]. \quad (7.30)$$

Como o valor esperado acima será positivo se H_f for verdadeira e negativo se H_g for verdadeira, temos que o sinal da estatística de teste nos dá evidências sobre um dos modelos alternativos.

O teste é realizado sob a hipótese H_0 , uma vez que esta possui distribuição assintótica bem definida. O teste é executado do modo convencional, sendo a sua região crítica de nível γ dada por $R = \{\mathbf{Y} : |T(V)| > z_{1-\gamma/2}\}$. Assim, se a estatística estiver dentro da região R^c , aceitamos H_0 ao nível γ fixado, o que implica em dizer que não conseguimos discriminar entre os dois modelos com o conjunto de dados disponível. Se não, se $T(V) > z_{1-\gamma/2}$ aceitaremos a hipótese H_f e se $T(V) < z_{\gamma/2}$ aceitaremos a hipótese H_g .

Capítulo 8

Resultados

Neste capítulo apresentamos nossa discussão sobre o desempenho do teste de Vuong quando um dos dois modelos alternativos é verdadeiro. Na Seção 8.1 mostramos que a opção por analisar a estatística do teste de Vuong sob H_0 pode trazer vantagens e desvantagens e ilustramos analiticamente esta particularidade em dois exemplos. Na Seção 8.2 mostramos um estudo de simulação comparando os testes de Cox e Vuong.

8.1 Notas sobre o teste de Vuong

No capítulo anterior vimos que o teste de Vuong possui mais facilidade computacional que o teste de Cox, uma vez que apenas os estimadores de máxima verossimilhança são necessários para a realização do teste, enquanto que o teste de Cox demanda também os limites em probabilidade destes estimadores sob a hipótese nula e o cálculo da variância. Tal facilidade é possível graças a inclusão de uma hipótese adicional, sob a qual a estatística é testada.

O fato de termos que decidir o teste de Vuong sob a hipótese H_0 pode gerar complicações. Primeiro, não saberemos o nível do teste se um dos modelos competidores for o verdadeiro modelo gerador dos dados. Por exemplo, assumindo que f é o verdadeiro modelo, pela desigualdade de Gibbs teremos

$$\mathbb{E}_F \left[\log \frac{f(x)}{g(x)} \right] = 0 \Leftrightarrow F(x) = G(x) \text{ em quase toda parte,} \quad (8.1)$$

o que é impossível, pois os modelos pertencem à famílias de distribuições separadas. Assim, a hipótese H_0 é inadmissível e, portanto, não podemos calcular o erro tipo I

do teste. Além disso, não temos uma distribuição assintótica sob H_f ou sob H_g , o que impede o cálculo do poder do teste.

Esperamos que um dos dois modelos propostos seja próximo o suficiente do modelo gerador dos dados, o que nos leva a considerar a hipótese H_0 apenas como um artifício, esperando rejeitá-la em favor do modelo verdadeiro. Contudo, o comprimento do intervalo da região de aceitação de H_0 pode ser grande, dificultando sua rejeição. Para compreender este fato, lembremos que (Teorema 7.1)

$$\frac{1}{\sqrt{n}} \log \frac{f(\mathbf{Y}|\hat{\alpha})}{g(\mathbf{Y}|\hat{\beta})} - \sqrt{n} \mathbb{E}_H \left(\log \frac{f(Y_1|\alpha_0)}{g(Y_1|\beta_0)} \right) \xrightarrow{D} N(0, \omega^2), \quad (8.2)$$

quando $n \rightarrow \infty$. Disto, a partir de certo n_0 , teremos que para todo $n > n_0$,

$$T(V) \approx \frac{\sqrt{n}}{\hat{\omega}} \mathbb{E}_H \left(\log \frac{f(Y_1|\hat{\alpha})}{g(Y_1|\hat{\beta})} \right). \quad (8.3)$$

Em geral, estimadores de máxima verossimilhança convergem para seus valores esperados na ordem de $1/n$. Assim, teremos que

$$T(V) \approx \frac{\sqrt{n}}{\hat{\omega}} \mathbb{E}_H \left(\log \frac{f(Y_1|\hat{\alpha})}{g(Y_1|\hat{\beta})} \right) \approx \frac{\sqrt{n}}{\omega} \mathbb{E}_H \left(\log \frac{f(Y_1|\alpha_0)}{g(Y_1|\beta_0)} \right). \quad (8.4)$$

Seja $CV(f, g)$ o coeficiente de variação de $\log f(Y_1|\alpha_0) - \log g(Y_1|\beta_0)$. Teremos que

$$T(V) \approx \frac{\sqrt{n}}{CV(f, g)}. \quad (8.5)$$

A relação acima traz duas informações:

1. Como a convergência do coeficiente de variação amostral para $CV(f, g)$ é em geral rápida, o teste de Vuong consegue identificar a direção da hipótese a ser aceita com tamanhos de amostra razoavelmente pequenos.
2. O valor absoluto de $CV(f, g)$ possui papel fundamental na rejeição de H_0 : se o valor de $CV(f, g)$ for muito elevado, então o teste precisará de uma amostra maior para poder rejeitar H_0 ; em caso contrário, ele rejeitará H_0 rapidamente.

Em resumo, um tamanho de amostra com o qual os estimadores de máxima verossimilhança se aproximam de seus valores esperados deveria ser razoável para que o teste de Vuong rejeitasse o modelo menos favorecido. Entretanto, a rejeição

de H_0 em favor do modelo mais verossímil dependerá, além do tamanho da amostra, do valor de $CV(f, g)$.

Nas subseções seguintes analisaremos dois casos nos quais $CV(f, g)$ influencia muito na escolha do modelo mais verossímil. Nestes casos, teremos que um dos dois modelos é verdadeiro, logo a hipótese H_0 não é admissível.

8.1.1 O Caso Lognormal versus Exponencial

Sejam $f(\cdot|\alpha)$ e $g(\cdot|\beta)$ as densidades dos modelos lognormal e exponencial, dadas por

$$f(y|\alpha) = \frac{1}{y\sqrt{2\pi\alpha_2}} e^{-\frac{(\log y)^2}{2\alpha_2}}, \quad \alpha_2 > 0. \quad (8.6)$$

$$g(y|\beta) = \frac{1}{\beta} e^{-\frac{y}{\beta}}, \quad \beta > 0. \quad (8.7)$$

Considere as hipóteses H_f e H_g como sendo verdadeiros os modelos f e g respectivamente. Sem perda de generalidade, assuma que $f(\cdot|\alpha)$ é o verdadeiro modelo gerador dos dados. Disto, teremos que

$$\mathbb{E}_F \left[\log \frac{f(\mathbf{Y}|\alpha)}{g(\mathbf{Y}|\beta_0)} \right] = \frac{1}{2} (-\log(2\pi\alpha_2) + \alpha_2 + 1) \quad (8.8)$$

$$\omega^2 = e^{\alpha_2} - 2\alpha_2 - \frac{1}{2}. \quad (8.9)$$

Conforme discutido, teremos que

$$T(V) \approx \sqrt{n} \left(\frac{\frac{1}{2} (-\log(2\pi\alpha_2) + \alpha_2 + 1)}{\sqrt{e^{\alpha_2} - 2\alpha_2 - \frac{1}{2}}} \right) := \frac{\sqrt{n}}{CV(f, g)} \quad (8.10)$$

O gráfico de $1/CV(f, g)$ para o intervalo $(0, 5)$ é dado na Figura 8.1. Podemos notar que, neste intervalo, esta decresce até certo ponto e depois começa a crescer lentamente. Valores de α_2 próximos do mínimo local da função precisarão de um tamanho de amostra maior para rejeitar H_0 em favor do verdadeiro modelo do que os demais valores do intervalo. Notemos ainda que para valores de α_2 muito pequenos, o teste deverá detectar a verdadeira hipótese mesmo com tamanhos de amostra pequenos.

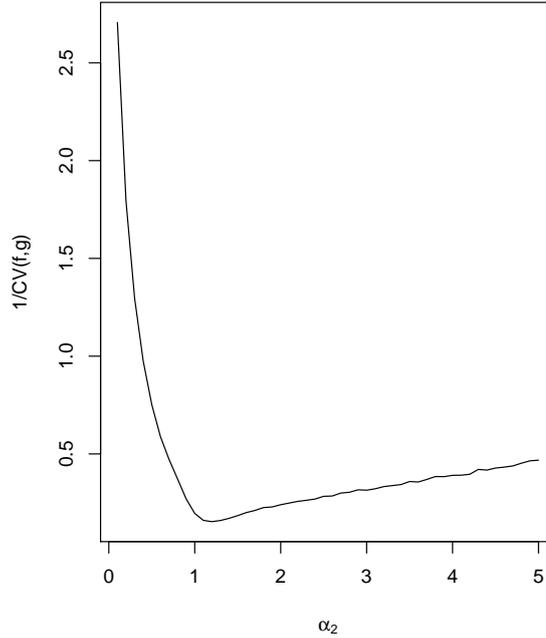


Figura 8.1: Gráfico de $1/CV(f, g)$ para o teste H_f : lognormal versus H_g : exponencial, quando a H_f é verdadeira.

Para ilustrar o que foi dito no parágrafo anterior, simulamos 1000 amostras de tamanhos $n = 10, 20, \dots, 300$ da distribuição $\text{lognormal}(0, \alpha_2)$ com $\alpha_2 = 0, 1; 0, 5; 1; 2; 4$. Para cada amostra realizamos o teste de Vuong e registramos o resultado. Para cada tamanho de amostra, registramos a proporção de aceitação H_f (a proporção de aceitação de H_g foi baixa em todas as simulações). Estes resultados estão apresentados na Figura 8.2. Podemos verificar como a região do espaço paramétrico está interferindo na proporção de aceitação conforme o comportamento de $CV(f, g)$.

Agora, considere que $g(\cdot|\beta)$ é o verdadeiro modelo gerador dos dados. Então,

$$\mathbb{E}_G \left[\log \frac{f(\mathbf{Y}|\hat{\alpha})}{g(\mathbf{Y}|\hat{\beta})} \right] = -0,09057302 \quad (8.11)$$

$$\omega^2 = 0,283408 \quad (8.12)$$

Para este caso, teremos que

$$T(V) \approx -\frac{\sqrt{n}}{5,877697}. \quad (8.13)$$

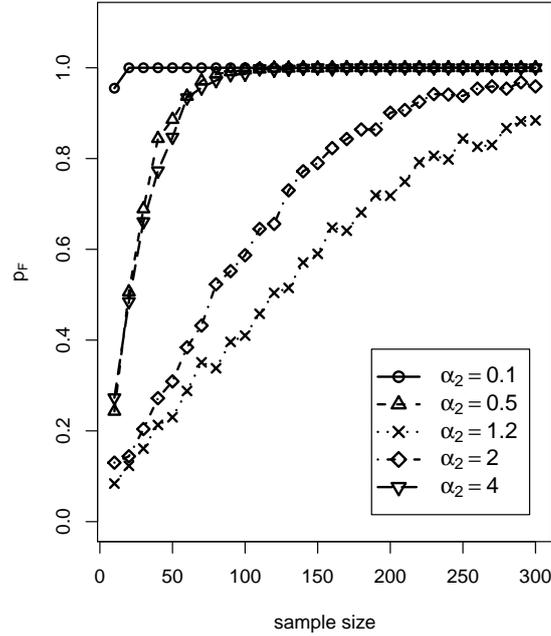


Figura 8.2: Proporção de aceitação de H_f para as hipóteses H_f : lognormal e H_g : exponencial. H_f é o verdadeiro modelo.

Portanto, se a hipótese H_g for verdadeira, teremos que $CV(f, g)$ convergirá para uma constante. O valor de β será irrelevante e o teste precisará de um tamanho de amostra elevado para rejeitar H_0 em favor de H_g . A Figura 8.3 mostra a proporção de aceitação, por tamanho de amostra, das hipóteses H_0 , H_f e H_g para o teste de Vuong em um estudo de simulações no qual 1000 amostras de tamanhos $n = 10, 20, \dots, 300$ foram retiradas da distribuição exponencial com taxa 3. Note que, como o esperado, a proporção de aceitação de H_0 é alta mesmo para valores como $n = 200$.

8.1.2 O Caso Lognormal versus Weibull

Sejam $f(\cdot|\alpha)$ e $g(\cdot|\beta)$ as densidades das distribuições lognormal e Weibull, dadas por

$$f(y|\alpha) = \frac{1}{y\sqrt{2\pi\alpha_2}} e^{-\frac{(\log y - \alpha_1)^2}{2\alpha_2}}, \quad \alpha_1 \in \mathbb{R}, \quad \alpha_2 > 0 \quad (8.14)$$

$$g(y|\beta) = \frac{\beta_2}{y} \left(\frac{y}{\beta_1}\right)^{\beta_2} e^{-\left(\frac{y}{\beta_1}\right)^{\beta_2}}, \quad \beta_1 > 0, \quad \beta_2 > 0. \quad (8.15)$$

Sejam H_f e H_g as hipóteses de que os dados foram gerados pela distribuição

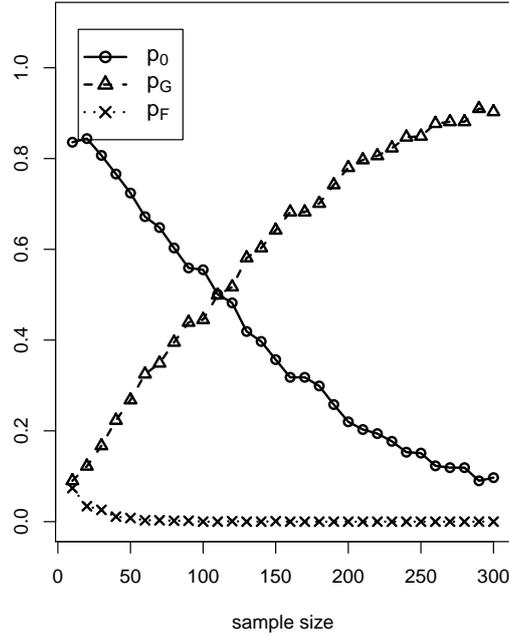


Figura 8.3: Simulações do teste de Vuong para as hipóteses H_f : lognormal e H_g : exponencial, onde H_g é verdadeira. A proporção de aceitação de H_0 , H_f e H_g são, respectivamente, p_0 , p_F e p_G .

lognormal e Weibull, respectivamente, e consideremos H_f verdadeira. Então

$$\mathbb{E}_F \left[\log \frac{f(\mathbf{Y}|\hat{\alpha})}{g(\mathbf{Y}|\hat{\beta})} \right] = 0,08106147 \quad (8.16)$$

$$\omega^2 = 1,218282 \quad (8.17)$$

$$CV(f, g) = 13,61631. \quad (8.18)$$

Notemos que o valor esperado da razão de verossimilhanças é constante (Equação 8.16). Além disso, o alto valor de $CV(f, g)$ indica que será necessário um tamanho de amostra elevado para rejeitarmos H_0 em favor de H_f . A Figura 8.4, mostra a proporção de aceitação das hipóteses H_0 , H_f e H_g para alguns tamanhos de amostra. Para cada tamanho de amostra 1000 testes foram realizados, nos quais as observações foram amostradas da distribuição lognormal(0, 1).

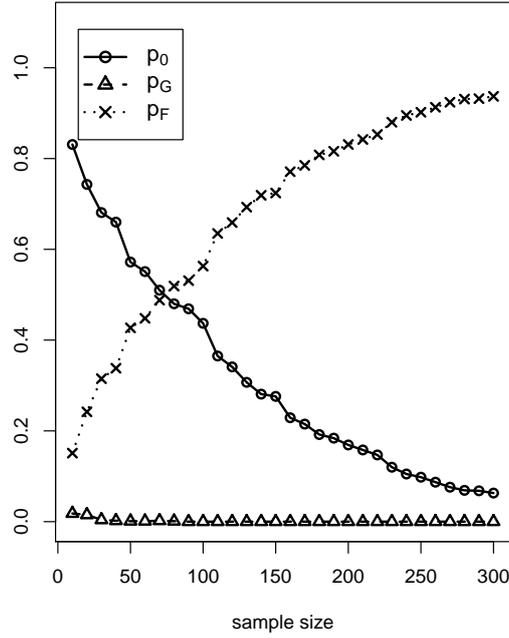


Figura 8.4: Resultados de simulação do teste de Vuong para as hipóteses H_f : Lognormal e H_g : Weibull, onde H_f é verdadeira. A proporção de aceitação de H_0 , H_f e H_g foram, respectivamente, p_0 , p_F e p_G .

Agora, consideremos que $g(\cdot|\beta)$ é o verdadeiro modelo gerador dos dados. Então,

$$\mathbb{E}_G \left[\log \frac{f(\mathbf{Y}|\hat{\alpha})}{g(\mathbf{Y}|\hat{\beta})} \right] = -0,09057302 \quad (8.19)$$

$$\omega^2 = 0,283408 \quad (8.20)$$

$$CV(f, g) = -5,877697 \quad (8.21)$$

Notemos que estes resultados são os mesmo obtidos do teste H_f versus H_g , onde f é o modelo lognormal, g o modelo exponencial e H_g é o verdadeiro modelo. Assim, teremos conclusões análogas aquelas já discutidas.

O teste de Vuong é desfavorável para esses conjuntos de hipóteses: independente do modelo adequado e do valor dos verdadeiros parâmetros, o teste precisará de um tamanho de amostra grande para rejeitar H_0 em favor do verdadeiro modelo.

8.2 Um cenário para simular comparações entre os testes

Uma vez que identificamos algumas características no teste de Vuong que podem interferir na escolha do modelo mais verossímil, se torna interessante mostrar esse efeito em um estudo de simulação, ao mesmo tempo em que comparamos o seu desempenho contra o teste de Cox para os mesmos modelos competidores. Entretanto, para compararmos estes testes, devemos levar em consideração alguns fatores.

Consideremos que estamos interessados em um dos modelos $f(\cdot|\alpha)$ ou $g(\cdot|\beta)$, sendo que necessariamente um deles é o verdadeiro modelo gerador dos dados. Como o teste de Vuong independe de qual é o verdadeiro modelo, ambos os testes podem ser utilizados para essa situação. Entretanto, sabemos que nesta situação a hipótese H_0 , do teste de Vuong, é inadmissível.

Testaremos os seguintes grupos de hipóteses:

$$\begin{cases} H_f : & Y|\alpha \sim \text{Lognormal}(0, \alpha_2) \\ H_g : & Y|\beta \sim \text{Exponencial}(\beta) \end{cases}$$

e

$$\begin{cases} H_f : & Y|\alpha \sim \text{Lognormal}(\alpha_1, \alpha_2) \\ H_g : & Y|\beta \sim \text{Weibull}(\beta_1, \beta_2) \end{cases}$$

onde as densidades destas distribuições são as mesmas mostradas na Seção 8.1. Ambos os teste terão nível 0,05.

Em relação ao teste de Cox optamos por utilizar o teste de significância que consiste em calcular $T_f(C)$ (ou $T_g(C)$), verificar se a estatística de teste está ou não na região de rejeição e decidir por rejeitar ou não $H_f(H_g)$. Lembremos que o teste de Cox é um teste unilateral. As estatísticas para o teste de Cox, considerando as hipóteses acima, são dadas na Tabela 8.1.

Existem algumas dificuldades para avaliar o teste de Vuong via simulação. Primeiramente, não podemos realizar simulações sob H_0 , o que impede, por exemplo, a estimação do nível do teste. Por isso utilizamos uma abordagem alternativa. Como queremos rejeitar H_0 na direção de uma das hipóteses, optamos por realizar o teste sob as hipóteses H_f e H_g . Por exemplo, para a hipótese H_f , podemos utilizar a região crítica $R = \{\mathbf{Y} : T(V) < z_{1-\alpha/2}\}$ (o que inclui a região de aceitação de H_0 e H_g). Podemos proceder de forma análoga com H_g , definindo a região de rejeição $R = \{\mathbf{Y} : T(V) > z_{\alpha/2}\}$. O nível do teste para qualquer tamanho de amostra finito poderá ser estimado via simulação a partir da proporção de vezes que o teste rejeitou H_f (ou H_g) quando o modelo H_f (ou H_g) for utilizado para gerar os dados. De modo análogo, o poder do teste de Vuong foi estimado via simulação como sendo a proporção das vezes que H_f (ou H_g) foi rejeitada dado que H_g (ou H_f) é verdadeira.

Vale lembrar que o teste de Vuong pode precisar de um grande tamanho de amostra para rejeitar H_0 , o que fará com que ele tenha mais poder que o teste de Cox. Pelo menos motivo, o teste de Vuong poderá apresentar um nível superior ao fixado.

Hipóteses	Estatísticas	Variâncias
$H_f : Y \alpha \sim \text{LN}(\alpha_1, \alpha_2)$	$n \log(\hat{\beta}/\hat{\beta}_\alpha)$	$n(e^{\hat{\alpha}_2} - 1 - \frac{3}{2}\hat{\alpha}_2)$
$H_g : Y \beta \sim \text{Exp}(\beta)$	$n(\hat{\alpha}_1 - \hat{\alpha}_{1\beta} + 0.5 \log(\hat{\alpha}_2/\hat{\alpha}_{2\beta}))$	$0.2834n$
$H_f : Y \alpha \sim \text{LN}(\alpha_1, \alpha_2)$	$n(\hat{\beta}_2 \log \hat{\beta}_1 - \hat{\beta}_{2\alpha} \log \hat{\beta}_{1\alpha} - \log \hat{\beta}_2 + \log \hat{\beta}_{2\alpha} - \hat{\alpha}_1(\hat{\beta}_2 - \hat{\beta}_{2\alpha}))$	$0.2183n$
$H_g : Y \beta \sim \text{Wei}(\beta_1, \beta_2)$	$n(\hat{\beta}_2(\hat{\alpha}_1 - \hat{\alpha}_{1\beta}) + 0.5 \log(\hat{\alpha}_2/\hat{\alpha}_{2\beta}))$	$V_g = 0.2834n$

Tabela 8.1: Estatísticas para o teste de Cox. LN: distribuição lognormal; Exp: distribuição exponencial; Wei: distribuição Weibull.

8.3 Simulações

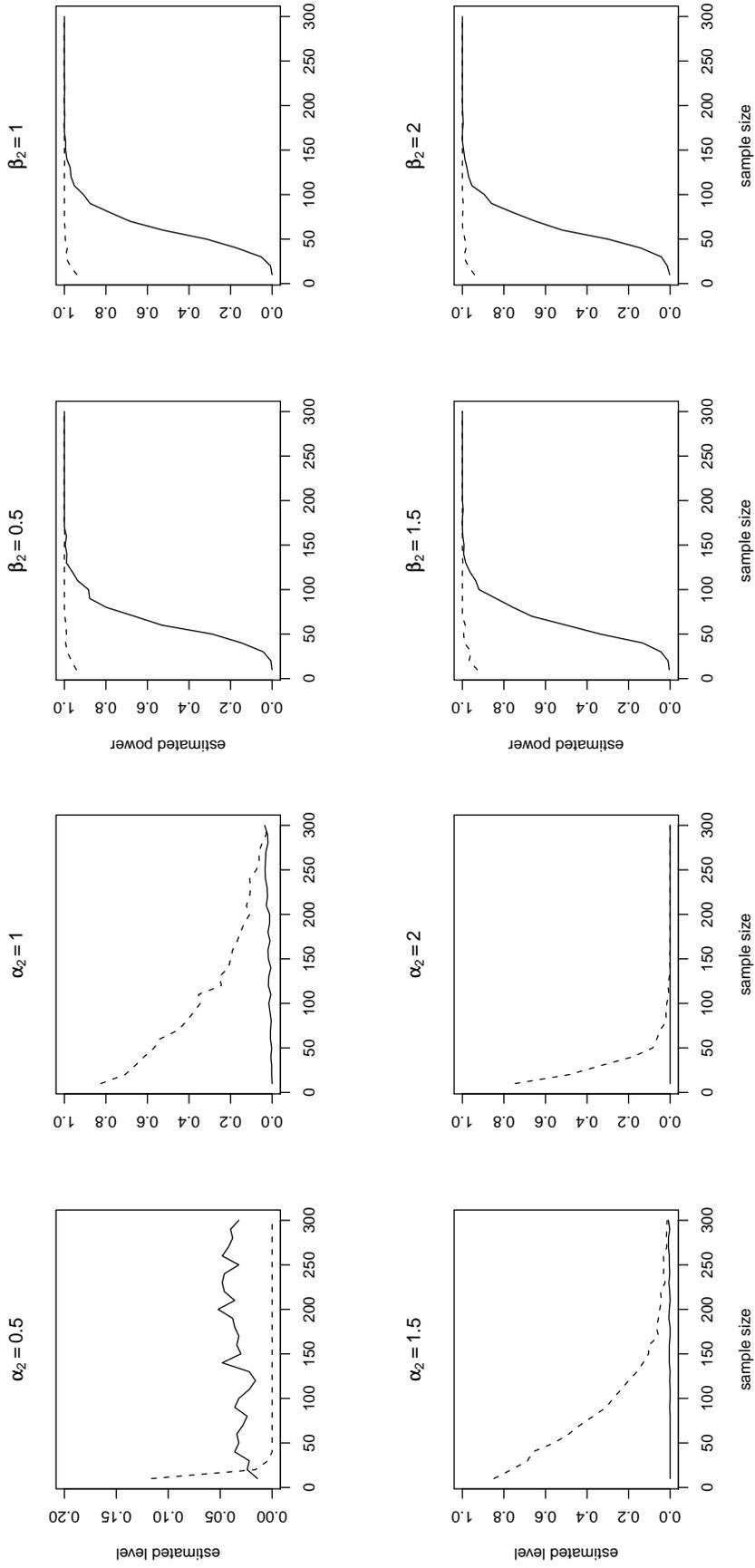
Para cada conjunto de hipóteses, simulamos 1.000 amostras de tamanho $n = 10, 20, \dots, 300$, sob H_f e sob H_g . Para cada amostra realizamos os dois testes, ao nível de 5% de significância. Os parâmetros utilizados em cada simulação estão dados na Tabela 8.2. O cálculo dos estimadores de máxima verossimilhança foram realizados via função `fitdistr`, do pacote `MASS` do software `R`, versão 2.11.

Parâmetros utilizados
Lognormal(0,0.2)(0,1)(0,1.5)(0,2)
Exponencial(0.2)(1)(1.5)(2)
Weibull(1,0.2)(2,1)(2,1.5)(2,2)

Tabela 8.2: Distribuições utilizadas na geração das amostras

A Figura 8.5 mostra o poder e o nível estimados dos dois testes para a hipótese H_f : lognormal versus H_g : exponencial. Na Figura 8.5(a) podemos notar que o teste de Vuong atinge o nível de 5% mesmo para pequenas amostras quando o valor do parâmetro é $\alpha_2 = 0,5$. O mesmo não acontece para os outros valores de α_2 . O teste de Cox manteve em média o nível de 5%. Na Figura 8.5 o poder do teste de Vuong se mostrou mais elevado que o de Cox para pequenas amostras, sendo que o poder dos dois teste se tornaram aproximadamente equivalentes para tamanhos de amostra acima de $n = 100$.

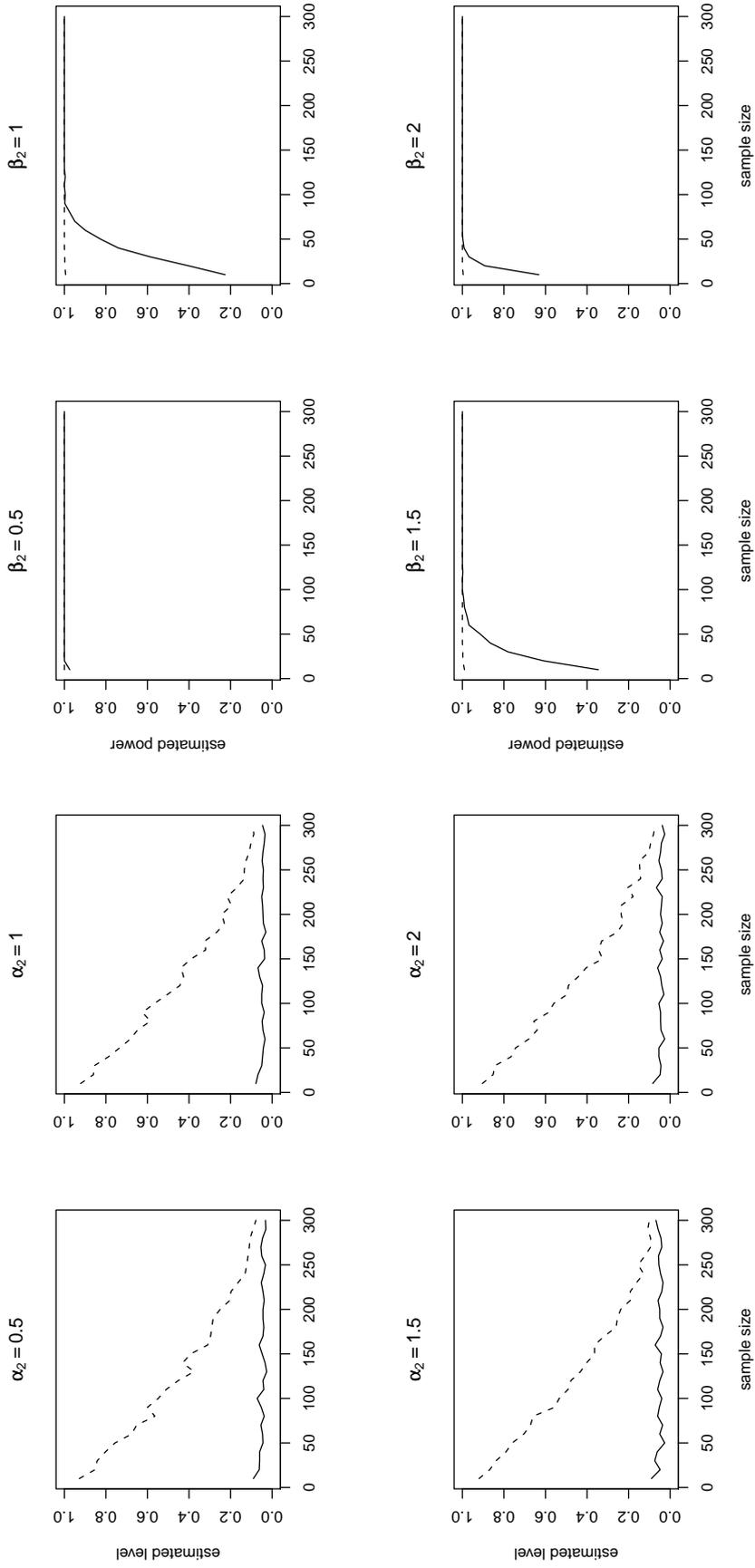
As Figuras 8.6 e 8.7 mostram o nível e poder estimados para os outros testes da Tabela 8.2. Ambas mostram características comuns: o teste de Cox atinge o nível de 5% mais rápido que o de Vuong, mas possui o poder mais baixo, conforme o esperado, devido a alta proporção de aceitação de H_0 para o teste de Vuong nestes casos.



(a) Level

(b) Power

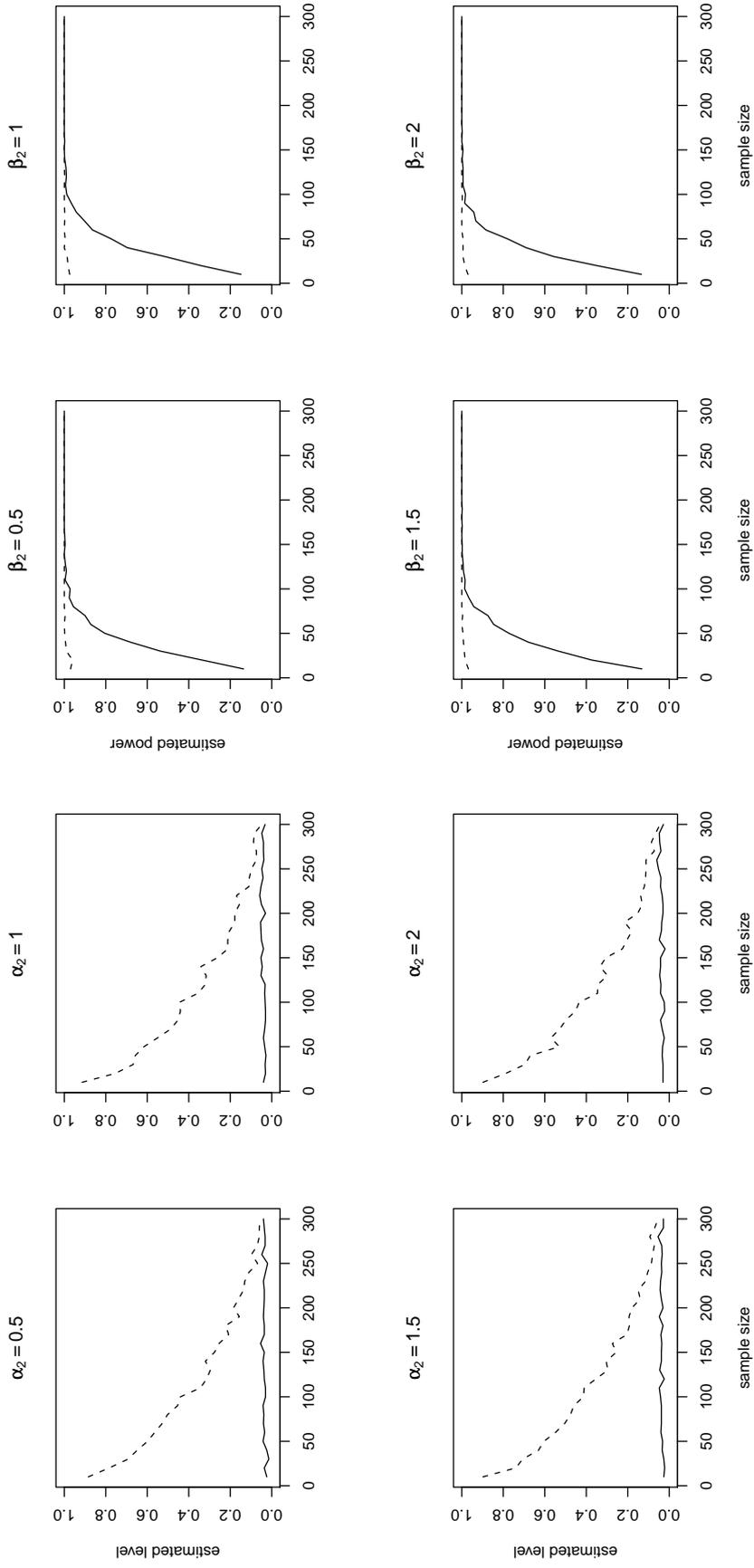
Figura 8.5: Nível e poder estimados para o teste lognormal versus exponencial, com hipótese nula sendo H_f : lognormal. Linha cheia: teste de Cox; Linha pontilhada: teste de Vuong.



(a) Level

(b) Power

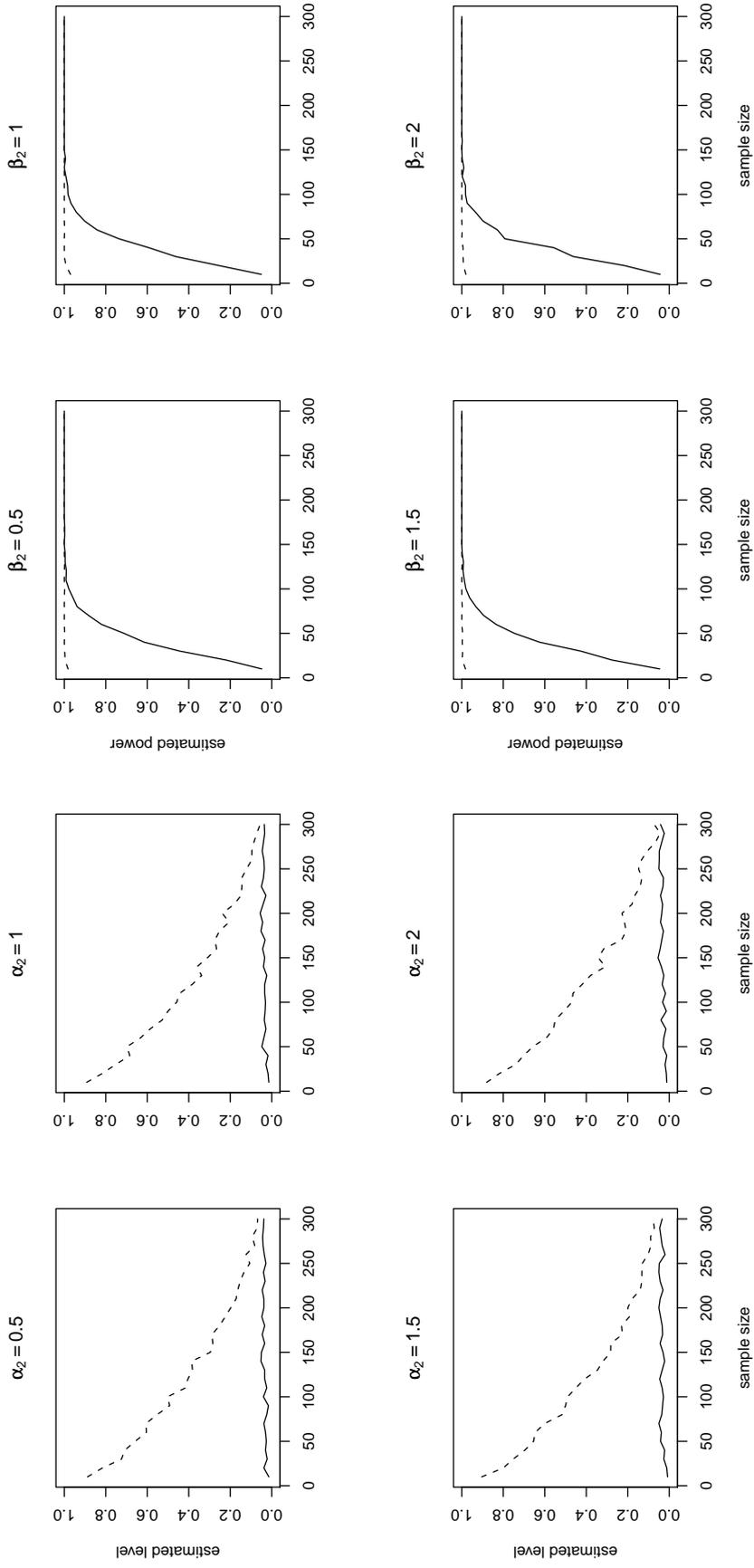
Figura 8.6: Nível e poder estimados para o teste exponencial versus lognormal, com hipótese nula sendo H_f : exponencial. Linha cheia: teste de Cox; Linha pontilhada: teste de Vuong.



(a) Level

(b) Power

Figura 8.7: Nível e poder estimados para o teste lognormal versus Weibull, com hipótese nula sendo H_f : lognormal. Linha cheia: teste de Cox; Linha pontilhada: teste de Vuong.



(a) Level

(b) Power

Figura 8.8: Nível e poder estimados para o teste lognormal versus Weibull, com hipótese nula sendo H_f : Weibull. Linha cheia: teste de Cox; Linha pontilhada: teste de Vuong.

Capítulo 9

Considerações Finais

Embora os testes gerais de seleção de modelos sejam assintóticos, é desejável que eles sejam válidos para tamanhos de amostras que possam ser obtidos na prática. Levando em consideração esta necessidade, mostramos neste trabalho que o teste de Vuong não possui um tamanho de amostra universal recomendado.

Ilustramos estas situações com dois exemplos. Em uma das situações, vimos que o teste de Vuong dependia do valor do parâmetro do verdadeiro modelo e, dependendo deste valor, o tamanho de amostra necessário poderia ser pequeno. Na outra situação vimos que independente do valor do parâmetro do verdadeiro modelo, o teste de Vuong precisaria de um tamanho de amostra elevado para aceitar o verdadeiro modelo.

A diferença fundamental na avaliação dos dois testes é a distribuição da estatística de teste. A estatística do teste de Cox possui distribuição sob a hipótese verdadeira, sendo esta aproximadamente normal. Já o teste de Vuong possui estatística de teste da por

$$T(V) \approx \frac{\sqrt{n}}{CV(f, g)}, \quad (9.1)$$

logo, precisamos de um valor de $CV(f, g)$ pequeno, ou um valor de n grande, para sair da região de aceitação da hipótese H_0 . Isto pode se agravar no caso em que um dos dois modelos é verdadeiro, pois neste caso H_0 não pode ocorrer. Temos que $T(V)$ não possui distribuição assintótica, embora esta se aproxime da distribuição normal para n finito, com média $\sqrt{n}\mathbb{E}(T(V))$. Assim, sugerimos que o teste de Vuong seja utilizado somente após um estudo sobre o valor de $CV(f, g)$ sob ambas as hipóteses.

Parte III

**Apêndice e referências
bibliográficas**

Apêndice A

Demonstrações

A.1 Lognormal

Considere $Y \sim \text{Lognormal}(\theta_1, \theta_2)$, com densidade dada por

$$f(y|\theta) = \frac{1}{y\sqrt{2\pi\theta_2}} \exp \left\{ -\frac{1}{2\theta_2} (\log y - \theta_1)^2 \right\}.$$

O logaritmo da função de verossimilhança é dado por

$$l(\theta) = -\frac{1}{2} \log(2\pi\theta_2) - \log y - \frac{1}{2\theta_2} (\log y - \theta_1)^2$$

Lema A.1. *Se $Y \sim \text{Lognormal}(\theta_1, \theta_2)$, então $\log Y \sim \mathcal{N}(\theta_1, \theta_2)$. Disto, teremos que $(\log Y - \theta_1)^2/\theta_2 \sim \chi_1^2$.*

Lema A.2 (Momentos da Lognormal). *Para qualquer $k > 0$, temos*

$$\mathbb{E}(Y^k) = e^{k\theta_1 + \frac{k^2\theta_2}{2}} \tag{A.1}$$

Do Lema A.2, temos que

$$\mathbb{E}(Y) = e^{\theta_1 + \frac{\theta_2}{2}} \tag{A.2}$$

$$\mathbb{E}(Y^2) = e^{2\theta_1 + 2\theta_2} \tag{A.3}$$

$$\mathbb{V}(Y) = e^{2\theta_1 + 2\theta_2} - e^{2\theta_1 + \theta_2} \tag{A.4}$$

Do Lema A.1, temos que

$$\mathbb{E}(\log Y) = \theta_1 \tag{A.5}$$

$$\mathbb{V}(\log Y) = \theta_2 \tag{A.6}$$

$$\mathbb{E}\left(\frac{(\log Y - \theta_1)^2}{\theta_2}\right) = 1 \tag{A.7}$$

$$\mathbb{V}\left(\frac{(\log Y - \theta_1)^2}{\theta_2}\right) = 2 \tag{A.8}$$

As derivadas parciais de $l(\theta)$ são

$$\begin{aligned}\frac{\partial}{\partial \theta_1} l(\theta) &= \frac{1}{\theta_2} (\log y - \theta_1) \\ \frac{\partial}{\partial \theta_2} l(\theta) &= -\frac{1}{2\theta_2} + \frac{1}{2\theta_2^2} (\log y - \theta_1)^2.\end{aligned}$$

Consideremos outro modelo H , que não seja a distribuição lognormal. Então,

$$\mathbb{E} \left(\frac{\partial}{\partial \theta_1} l(\theta) \right) = 0 \Rightarrow \theta_1^* = \mathbb{E}_H(\log Y) \quad (\text{A.9})$$

$$\mathbb{E} \left(\frac{\partial}{\partial \theta_2} l(\theta) \right) = 0 \Rightarrow \theta_2^* = \mathbb{E}_H(\log Y - \theta_1^*)^2 \quad (\text{A.10})$$

Também teremos que

$$\begin{aligned}\mathbb{E}(Y^k \log Y) &= \int_0^\infty y^{k-1} \frac{\log y}{\sqrt{2\pi\theta_2}} e^{-\frac{1}{2\theta_2}(\log y - \theta_1)^2} dy = \int_{\mathbb{R}} \frac{ue^{uk}}{\sqrt{2\pi\theta_2}} e^{-\frac{1}{2\theta_2}(u-\theta_1)^2} du \\ &= e^{k\theta_1 + k^2 \frac{\theta_2}{2}} \int_{\mathbb{R}} \frac{u}{\sqrt{2\pi\theta_2}} e^{-\frac{1}{2\theta_2}(u-\theta_1 - k\theta_2)^2} du \\ &= (\theta_1 + k\theta_2) e^{k\theta_1 + k^2 \frac{\theta_2}{2}}, \quad k = 1, 2, \dots,\end{aligned} \quad (\text{A.11})$$

$$\begin{aligned}\mathbb{E} \left(Y^k (\log Y - \theta_1)^2 \right) &= \int_0^\infty y^{k-1} \frac{(\log y - \theta_1)^2}{\sqrt{2\pi\theta_2}} e^{-\frac{1}{2\theta_2}(\log y - \theta_1)^2} dy \\ &= \int_{\mathbb{R}} \frac{e^{uk}(u - \theta_1)^2}{\sqrt{2\pi\theta_2}} e^{-\frac{1}{2\theta_2}(u-\theta_1)^2} du \\ &= e^{k\theta_1 + k^2 \frac{\theta_2}{2}} \int_{\mathbb{R}} \frac{(u - \theta_1)^2}{\sqrt{2\pi\theta_2}} e^{-\frac{1}{2\theta_2}(u-\theta_1 - k\theta_2)^2} du \\ &= \theta_2(1 + k^2\theta_2) e^{k\theta_1 + k^2 \frac{\theta_2}{2}}, \quad k = 1, 2, \dots,\end{aligned} \quad (\text{A.12})$$

$$\begin{aligned}\mathbb{E} \left(\log Y (\log Y - \theta_1)^2 \right) &= \int_0^\infty \frac{\log y (\log y - \theta_1)^2}{y\sqrt{2\pi\theta_2}} e^{-\frac{1}{2\theta_2}(\log y - \theta_1)^2} dy \\ &= \int_{\mathbb{R}} \frac{u(u - \theta_1)^2}{\sqrt{2\pi\theta_2}} e^{-\frac{1}{2\theta_2}(u-\theta_1)^2} du \\ &= \theta_1\theta_2\end{aligned} \quad (\text{A.13})$$

e disso,

$$\begin{aligned}\mathbb{C}(Y^k, \log Y) &= (\theta_1 + k\theta_2) e^{k\theta_1 + k^2 \frac{\theta_2}{2}} - \theta_1 e^{k\theta_1 + k^2 \frac{\theta_2}{2}} \\ &= k\theta_2 e^{k\theta_1 + k^2 \frac{\theta_2}{2}}, \quad k = 1, 2, \dots,\end{aligned} \quad (\text{A.14})$$

$$\begin{aligned}\mathbb{C}(Y^k, (\log Y - \theta_1)^2) &= \theta_2(1 + k^2\theta_2) e^{k\theta_1 + k^2 \frac{\theta_2}{2}} - \theta_2 e^{k\theta_1 + k^2 \frac{\theta_2}{2}} \\ &= k^2\theta_2^2 e^{k\theta_1 + k^2 \frac{\theta_2}{2}}, \quad k = 1, 2, \dots,\end{aligned} \quad (\text{A.15})$$

$$\mathbb{C}(\log Y, (\log Y - \theta_1)^2) = 0 \quad (\text{A.16})$$

Por último, teremos que

$$\mathbb{E}(l(\theta)) = -\frac{1}{2} \log(2\pi\theta_2) - \mathbb{E}(\log Y) - \frac{1}{2\theta_2} \mathbb{E}(\log Y - \theta_1)^2 = -\frac{1}{2} \log(2\pi\theta_2) - \theta_1 - \frac{1}{2} \quad (\text{A.17})$$

e que

$$\begin{aligned} \mathbb{V}(l(\theta)) &= \mathbb{V}(\log Y) + \frac{1}{4} \mathbb{V} \left(\frac{(\log Y - \theta_1)^2}{\theta_2} \right) + 2C \left(\log Y, \frac{(\log Y - \theta_1)^2}{\theta_2} \right) \\ &= \theta_2 + \frac{1}{2} \end{aligned} \quad (\text{A.18})$$

A.2 Exponencial

Considere $Y \sim \text{Exponencial}(\theta)$, com densidade dada por

$$f(y|\theta) = \frac{1}{\theta} e^{-\frac{y}{\theta}}.$$

O logaritmo da função de verossimilhança será dado por

$$l(\theta) = \log f(y|\theta) = -\log \theta - \frac{y}{\theta},$$

e a derivada do logaritmo da verossimilhança será

$$\frac{d}{d\theta} l(\theta) = -\frac{1}{\theta} + \frac{y}{\theta^2}$$

Lema A.3. *Os momentos da exponencial são dados por*

$$\mathbb{E}(Y^k) = \theta^k \Gamma(k + 1).$$

Do Lema A.3 teremos $\mathbb{E}(Y) = \theta$ e $\mathbb{V}(Y) = \theta^2$. Agora, seja H um modelo qualquer que satisfaça as condições de regularidade. Então, teremos que

$$\mathbb{E}_H \left(\frac{d}{d\theta} l(\theta) \right) = 0 \Rightarrow \theta^* = \mathbb{E}(Y) \quad (\text{A.19})$$

Antes de encontrar as outras esperanças necessárias, consideremos a seguinte definição.

Definição A.4. *Para $n = 1, 2, \dots$ e $x > 0$,*

$$\phi^{(n)}(x) := \frac{d^n}{dx^n} \Gamma(x) = \int_0^\infty t^{x-1} (\log t)^n e^{-t} dt.$$

e,

$$\psi^{(n)}(x) := \frac{d}{dx} \psi^{(0)}(x),$$

onde

$$\psi^{(0)}(x) := \frac{d}{dx} \log(\Gamma(x)).$$

Da Definição A.4, teremos

$$\phi^{(1)}(x) = \frac{d}{dx}\Gamma(x) = \Gamma(x)\psi^{(0)}(x)$$

$$\phi^{(2)}(x) = \Gamma(x) \left[(\psi^{(0)}(x))^2 + \psi^{(1)}(x) \right]$$

$$\phi^{(3)}(x) = \psi^{(0)}(x)\phi^{(2)}(x) + \Gamma(x) \left(2\psi^{(0)}(x)\psi^{(1)}(x) + \psi^{(2)}(x) \right)$$

$$\begin{aligned} \phi^{(4)}(x) &= \psi^{(1)}(x)\phi^{(2)}(x) + \psi^{(0)}(x)\phi^{(3)}(x) + \psi^{(0)}(x)\Gamma(x) \left(2\psi^{(0)}(x)\psi^{(1)}(x) + \psi^{(2)}(x) \right) \\ &\quad + \Gamma(x) \left(2(\psi^{(1)}(x))^2 + 2\psi^{(0)}(x)\psi^{(2)}(x) + \psi^{(3)}(x) \right) \end{aligned}$$

A Tabela A.2 mostra alguns valores da função ϕ que serão utilizados. A partir da Definição A.4, teremos que

$$\mathbb{E}(\log Y) = \int_0^\infty \frac{\log y}{\theta} e^{-\frac{y}{\theta}} dy = \int_0^\infty \log(u\theta) e^{-u} du = \log \theta + \phi^{(1)}(1), \quad (\text{A.20})$$

$$\begin{aligned} \mathbb{E}(\log Y)^2 &= \int_0^\infty \frac{(\log y)^2}{\theta} e^{-\frac{y}{\theta}} dy = \int_0^\infty (\log(u\theta))^2 e^{-u} du \\ &= \phi^{(2)}(1) + (\log \theta)^2 + 2\phi^{(1)}(1) \log \theta, \end{aligned} \quad (\text{A.21})$$

$$\mathbb{V}(\log Y) = \phi^{(2)}(1) - [\phi^{(1)}(1)]^2 \quad (\text{A.22})$$

$$(\text{A.23})$$

e que

$$\begin{aligned}
\mathbb{E}(\log Y - k)^2 &= \mathbb{E} \left[\left(\log \frac{Y}{\theta} + \log \theta - k \right)^2 \right] \\
&= \mathbb{E} \left(\log \frac{Y}{\theta} \right)^2 + (\log \theta - k)^2 + 2(\log \theta - k) \mathbb{E} \left(\log \frac{Y}{\theta} \right) \\
&= \phi^{(2)}(1) + (\log \theta - k)^2 + 2(\log \theta - k) \phi^{(1)}(1)
\end{aligned} \tag{A.24}$$

$$\begin{aligned}
\mathbb{E}(\log Y - k)^4 &= \mathbb{E} \left[\left(\log \frac{Y}{\theta} + \log \theta - k \right)^4 \right] = \mathbb{E} \left(\log \frac{Y}{\theta} \right)^4 + 4(\log \theta - k) \mathbb{E} \left(\log \frac{Y}{\theta} \right)^3 \\
&\quad + 6(\log \theta - k)^2 \mathbb{E} \left(\log \frac{Y}{\theta} \right)^2 + 4(\log \theta - k)^3 \mathbb{E} \left(\log \frac{Y}{\theta} \right) + (\log \theta - k)^4
\end{aligned} \tag{A.25}$$

$$\begin{aligned}
&= \phi^{(4)}(1) + 4(\log \theta - k) \phi^{(3)}(1) + 6(\log \theta - k)^2 \phi^{(2)}(1) \\
&\quad + 4(\log \theta - k)^3 \phi^{(1)}(1) + (\log \theta - k)^4
\end{aligned} \tag{A.26}$$

$$\begin{aligned}
\mathbb{V}[(\log Y - k)^2] &= \phi^{(4)}(1) - [\phi^{(2)}(1)]^2 + 4(\log \theta - k)^2 [\phi^{(2)}(1) - (\phi^{(1)}(1))^2] \\
&\quad + 4(\log \theta - k) [\phi^{(3)}(1) - \phi^{(1)}(1) \phi^{(2)}(1)]
\end{aligned} \tag{A.27}$$

$$\begin{aligned}
\mathbb{E}(Y \log Y) &= \int_0^\infty \frac{y \log y}{\theta} e^{-\frac{y}{\theta}} dy \\
&= \theta \int_0^\infty u \log(u\theta) e^{-u} du = \theta \log \theta + \theta \int_0^\infty u(\log u) e^{-u} du \\
&= \theta \log \theta + \theta \phi^{(1)}(2)
\end{aligned} \tag{A.28}$$

$$\begin{aligned}
\mathbb{E}(Y(\log Y - k)^2) &= \int_0^\infty \frac{y(\log y - k)^2}{\theta} e^{-\frac{y}{\theta}} dy = \theta \int_0^\infty u(\log(u\theta) - k)^2 e^{-u} du \\
&= \theta \left[\int_0^\infty (u(\log u)^2 + 2(\log \theta - k) \log u + (\log \theta - k)^2) e^{-u} du \right] \\
&= \theta \left(\phi^{(2)}(2) + 2\phi^{(1)}(2)(\log \theta - k) + (\log \theta - k)^2 \right)
\end{aligned} \tag{A.29}$$

$$\begin{aligned}
\mathbb{E} \left[\log Y (\log Y - k)^2 \right] &= \mathbb{E} \left[\log \frac{Y}{\theta} (\log Y - k)^2 \right] + \log \theta \mathbb{E}(\log Y - k)^2 \\
&= \phi^{(3)}(1) + (\log \theta - k)^2 \phi^{(1)}(1) + 2(\log \theta - k) \phi^{(2)}(1) \\
&\quad + \log \theta \mathbb{E}(\log Y - k)^2.
\end{aligned} \tag{A.30}$$

Assim, teremos

$$\mathbb{C}(Y, \log Y) = \theta \left(\phi^{(1)}(2) - \phi^{(1)}(1) \right) = \theta \quad (\text{A.31})$$

$$\begin{aligned} \mathbb{C}(Y, (\log Y - k)^2) &= \theta \left(\phi^{(2)}(2) + 2\phi^{(1)}(2)(\log \theta - k) + (\log \theta - k)^2 \right) \\ &\quad - \theta \left(\phi^{(2)}(1) + (\log \theta - k)^2 + 2(\log \theta - k)\phi^{(1)}(1) \right) \end{aligned} \quad (\text{A.32})$$

$$\begin{aligned} &= \theta \left(\phi^{(2)}(2) - \phi^{(2)}(1) \right) + 2\theta(\log \theta - k) \left(\phi^{(1)}(2) - \phi^{(1)}(1) \right) \\ &= \theta \left(\phi^{(2)}(2) - \phi^{(2)}(1) \right) + 2\theta(\log \theta - k) \end{aligned} \quad (\text{A.33})$$

$$\begin{aligned} \mathbb{C}(\log Y, (\log Y - k)^2) &= \phi^{(3)}(1) + 2(\log \theta - k)\phi^{(2)}(1) \\ &\quad + \phi^{(1)}(1) [(\log \theta - k)^2 - \mathbb{E}(\log Y - k)^2] \end{aligned} \quad (\text{A.34})$$

Por último, teremos

$$\mathbb{E}(l(\theta)) = -\log \theta - 1 \quad (\text{A.35})$$

e

$$\mathbb{V}(l(\theta)) = 1 \quad (\text{A.36})$$

	$x = 1$	$x = 2$
$\phi^{(1)}(x)$	-0,5772157	0,4227843
$\phi^{(2)}(x)$	1,978112	0,8236807
$\phi^{(3)}(x)$	-5,444874	0,4894615
$\phi^{(4)}(x)$	23,56147	1,781976

Tabela A.1: Alguns valores para a função $\phi^{(k)}(x)$.

A.3 Weibull

Seja Y uma v.a. com distribuição Weibull, com denisdade

$$f(y|\theta_1, \theta_2) = \frac{\theta_2}{y} \left(\frac{y}{\theta_1} \right)^{\theta_2} e^{-\left(\frac{y}{\theta_1} \right)^{\theta_2}}.$$

O logaritmo da verossimilhança é

$$l(\theta) = \log \theta_2 - \log y + \theta_2 (\log y - \log \theta_1) - \left(\frac{y}{\theta_1} \right)^{\theta_2},$$

e suas derivadas parciais são

$$\begin{aligned}\frac{\partial}{\partial \theta_1} l_f(\theta) &= -\frac{\theta_2}{\theta_1} + \frac{\theta_2}{\theta_1^{\theta_2+1}} y^{\theta_2} \\ \frac{\partial}{\partial \theta_2} l_f(\theta) &= \frac{1}{\theta_2} + \log y - \log \theta_1 - \frac{y^{\theta_2}}{\theta_1^{\theta_2}} (\log y - \log \theta_1)\end{aligned}$$

Dado outro modelo H , o pseudo-verdadeiro valor de θ será dado pela solução das seguintes equações:

$$\theta_1^* = \left(\mathbb{E}_H(Y)^{\theta_2^*} \right)^{\frac{1}{\theta_2^*}} \quad (\text{A.37})$$

$$\frac{1}{\theta_2^*} + \mathbb{E}_H(\log Y) - \frac{1}{\theta_2^*} \log \mathbb{E}_H(Y^{\theta_2^*}) - \frac{\mathbb{E}_H(Y^{\theta_2^*} \log Y) - \log \theta_1^* \mathbb{E}_H(Y^{\theta_2^*})}{\mathbb{E}_H(Y^{\theta_2^*})} = 0 \quad (\text{A.38})$$

Proposição A.5. *Se $Y \sim Weibull(\theta_1, \theta_2)$, então*

$$\mathbb{E}[t(Y)] = \mathbb{E}_U[t(\theta_1 U^{1/\theta_2})],$$

onde $U \sim Exp(1)$.

Assim, poderemos utilizar os resultados obtidos para a distribuição exponencial. Teremos

$$\mathbb{E}(Y^k) = \theta_1^k \mathbb{E}_U[U^{k/\theta_2}] = \theta_1^k \Gamma\left(1 + \frac{k}{\theta_2}\right) \quad (\text{A.39})$$

$$\mathbb{V}(Y^{\theta_2}) = \theta_1^{2\theta_2} \Gamma(3) - \theta_1^{2\theta_2} \Gamma(2)^2 = \theta_1^{2\theta_2} \quad (\text{A.40})$$

$$\mathbb{E}(\log Y) = \mathbb{E}_U[\log(\theta_1 U^{1/\theta_2})] = \log \theta_1 + \frac{1}{\theta_2} \mathbb{E}_U(\log U) = \log \theta_1 + \frac{\phi^{(1)}(1)}{\theta_2} \quad (\text{A.41})$$

$$\begin{aligned}\mathbb{E}[(\log Y)^2] &= (\log \theta_1)^2 + \frac{1}{\theta_2^2} \mathbb{E}_U[(\log U)^2] + \frac{2}{\theta_2} \log \theta_1 \mathbb{E}_U(\log U) \\ &= (\log \theta_1)^2 + \frac{\phi^{(2)}(1)}{\theta_2^2} + \frac{2}{\theta_2} \phi^{(1)}(1) \log \theta_1\end{aligned} \quad (\text{A.42})$$

$$\mathbb{V}(\log Y) = \frac{1}{\theta_2^2} \left(\phi^{(2)}(1) - [\phi^{(1)}(1)]^2 \right) = \frac{\psi^{(1)}(1)}{\theta_2^2} \quad (\text{A.43})$$

$$\begin{aligned}\mathbb{E}[(\log Y - k)^2] &= \mathbb{E}_U[(\log(\theta_1 U^{1/\theta_2}) - k)^2] \\ &= \frac{1}{\theta_2^2} \mathbb{E}_U [(\log U + \theta_2(\log \theta_1 - k))^2]\end{aligned} \quad (\text{A.44})$$

$$\begin{aligned}\mathbb{E}[(\log Y - k)^4] &= \mathbb{E}_U[(\log(\theta_1 U^{1/\theta_2}) - k)^4] \\ &= \frac{1}{\theta_2^4} \mathbb{E}_U [(\log U + \theta_2(\log \theta_1 - k))^4]\end{aligned} \quad (\text{A.45})$$

$$\mathbb{V}[(\log Y - k)^2] = \frac{1}{\theta_2^4} \mathbb{V}_U [(\log U + \theta_2(\log \theta_1 - k))^2] \quad (\text{A.46})$$

$$\mathbb{E}[Y^{\theta_2} \log Y] = \theta_1^{\theta_2} \mathbb{E}_U[U(\log \theta_1 + \frac{1}{\theta_2} \log U)] = \theta_1^{\theta_2} \left(\log \theta_1 + \frac{1}{\theta_2} \phi^{(1)}(2) \right) \quad (\text{A.47})$$

$$\mathbb{E}[Y^{\theta_2} (\log Y - k)^2] = \frac{\theta_1^{\theta_2}}{\theta_2^2} \mathbb{E}_U[U(\log U + \theta_2(\log \theta_1 - k))^2] \quad (\text{A.48})$$

$$\begin{aligned} \mathbb{E}[\log Y (\log Y - k)^2] &= \mathbb{E}_U \left[\log(\theta_1 U^{1/\theta_2}) \left(\log(\theta_1 U^{1/\theta_2}) - k \right)^2 \right] \\ &= \frac{\log \theta_1}{\theta_2^2} \mathbb{E}_U \left[(\log U + \theta_2(\log \theta_1 - k))^2 \right] \\ &\quad + \frac{1}{\theta_2^2} \mathbb{E}_U \left[\log U (\log U + \theta_2(\log \theta_1 - k))^2 \right] \end{aligned} \quad (\text{A.49})$$

Covariâncias:

$$\begin{aligned} \mathbb{C}(Y^{\theta_2}, \log Y) &= \theta_1^{\theta_2} \left(\log \theta_1 + \frac{1}{\theta_2} \phi^{(1)}(2) \right) - \theta_1^{\theta_2} \left(\log \theta_1 + \frac{1}{\theta_2} \phi^{(1)}(1) \right) \\ &= \frac{\theta_1^{\theta_2}}{\theta_2} \left(\phi^{(1)}(2) - \phi^{(1)}(1) \right) = \frac{\theta_1^{\theta_2}}{\theta_2} \end{aligned} \quad (\text{A.50})$$

$$\mathbb{C}(Y^{\theta_2}, (\log Y - k)^2) = \frac{\theta_1^{\theta_2}}{\theta_2^2} \mathbb{E}_U \left[(U - 1)(\log U + \theta_2(\log \theta_1 - k))^2 \right] \quad (\text{A.51})$$

$$\mathbb{C}(\log Y, (\log Y - k)^2) = \frac{1}{\theta_2^2} \mathbb{E}_U \left[\left(\log U - \frac{\phi^{(1)}(1)}{\theta_2} \right) (\log U + \theta_2(\log \theta_1 - k))^2 \right] \quad (\text{A.52})$$

Assim, das equações A.39 e A.41 teremos

$$\mathbb{E}[l(\theta)] = \log \theta_2 - \log \theta_1 + \phi^{(1)}(1) - \frac{\phi^{(1)}(1)}{\theta_2} - 1 \quad (\text{A.53})$$

e das equações A.40, A.43 e A.50 teremos

$$\begin{aligned} \mathbb{V}[l(\theta)] &= (\theta_2 - 1)^2 \mathbb{V}(\log Y) + \frac{1}{\theta_1^{2\theta_2}} \mathbb{V}(Y^{\theta_2}) - 2 \frac{(\theta_2 - 1)}{\theta_1^{\theta_2}} \mathbb{C}[\log Y, Y^{\theta_2}] \\ &= (\theta_2 - 1)^2 \frac{\psi^{(1)}(1)}{\theta_2^2} + \frac{1}{\theta_1^{2\theta_2}} \frac{\theta_1^{2\theta_2}}{\theta_1} - 2 \frac{(\theta_2 - 1)}{\theta_1^{\theta_2}} \frac{\theta_1^{\theta_2}}{\theta_2} \\ &= \left(1 - \frac{1}{\theta_2} \right)^2 \psi^{(1)}(1) - 1 + \frac{2}{\theta_2} \end{aligned} \quad (\text{A.54})$$

A.4 O Teste Lognormal versus Exponencial

Consideremos a hipótese nula como sendo a distribuição lognormal, versus a alternativa exponencial. Sob a hipótese nula, pela equação A.19 e pelo Lema ?? teremos que

$$\beta_* = \mathbb{E}_f(Y) = e^{\alpha_1 + \frac{\alpha_2}{2}}.$$

Assim, pelas equações A.5, A.7 e A.2, teremos

$$\begin{aligned}
\mathbb{E}_f(l(\alpha) - l(\beta^*)) &= -\frac{1}{2} \log(2\pi\alpha_2) - \mathbb{E}_f(\log Y) - \frac{1}{2\alpha_2} \mathbb{E}_f(\log Y - \alpha_1)^2 + \log \beta^* \\
&\quad + \frac{1}{\beta^*} \mathbb{E}_f(Y) \\
&= \frac{1}{2} (-\log(2\pi\alpha_2) + \alpha_2 + 1)
\end{aligned} \tag{A.55}$$

Pela equação A.18, temos que

$$\mathbb{V}_f(l_f(\alpha)) = \alpha_2 + \frac{1}{2}.$$

Pela equação A.4, temos que

$$\mathbb{V}_f(l_g(\beta^*)) = \frac{1}{\beta^{*2}} \mathbb{V}_f(Y) = e^{\alpha_2} - 1.$$

Pelas equações A.14 e A.15, teremos

$$\begin{aligned}
\mathbb{C}_f(l_f(\alpha), l_g(\beta^*)) &= \mathbb{C}_f\left(\log Y + \frac{(\log Y - \alpha_1)^2}{2\alpha_2}, \frac{Y}{\beta^*}\right) = \mathbb{C}_f\left(\log Y, \frac{Y}{\beta^*}\right) \\
&\quad + \mathbb{C}_f\left(\frac{(\log Y - \alpha_1)^2}{2\alpha_2}, \frac{Y}{\beta^*}\right) \\
&= \frac{3}{2} \alpha_2,
\end{aligned} \tag{A.56}$$

logo,

$$\begin{aligned}
\mathbb{V}_f(l_f(\alpha) - l_g(\beta^*)) &= \mathbb{V}_f(l_f(\alpha)) + \mathbb{V}_f(l_g(\beta^*)) - 2\mathbb{C}_f(l_f(\alpha), l_g(\beta^*)) \\
&= e^{\alpha_2} - 2\alpha_2 - \frac{1}{2}
\end{aligned}$$

Assim, teremos que o coeficiente de variação de $l_f(\alpha) - l_g(\beta^*)$ será

$$CV(f, g) = 2 \left(\frac{\sqrt{e^{\alpha_2} - 2\alpha_2 - \frac{1}{2}}}{\alpha_2 + 1 - \log(2\pi\alpha_2)} \right)$$

Agora, consideremos que hipótese nula é a de distribuição exponencial, versus a alternativa lognormal. Usando as equações A.9, A.10, A.20 e A.24, teremos

$$\beta_1^* = \mathbb{E}_f(\log Y) = \log \alpha + \phi^{(1)}(1) \tag{A.57}$$

$$\beta_2^* = \mathbb{E}_f[(\log Y - \beta_1^*)^2] = \phi^{(2)}(1) - (\phi^{(1)}(1))^2 = 1,644934 \tag{A.58}$$

Assim,

$$\mathbb{E}_f(l_f(\alpha) - l_g(\beta^*)) = +\frac{1}{2} \log(2\pi\beta_2^*) + \phi^{(1)}(1) - \frac{1}{2} = 0,09057302 \tag{A.59}$$

As variâncias do logaritmo das verossimilhanças serão

$$\mathbb{V}_f(l_f(\alpha)) = \frac{1}{\alpha^2} \mathbb{V}_f(Y) = 1, \quad (\text{A.60})$$

$$\mathbb{V}_f(l_g(\beta^*)) = \mathbb{V}_f(\log Y) + \frac{1}{4\beta_2^{*2}} \mathbb{V}_f[(\log Y - \beta_1^*)^2] + \frac{1}{\beta_2^*} \mathbb{C}_f(\log Y, (\log Y - \beta_1^*)^2) \quad (\text{A.61})$$

$$\begin{aligned} &= \beta_2^* + \frac{1}{4\beta_2^{*2}} \left[\phi^{(4)}(1) - [\phi^{(2)}(1)]^2 + 4(\phi^{(1)}(1))^2 \beta_2^* - 4\phi^{(1)}(1) \left(\phi^{(3)}(1) - \phi^{(1)}(1)\phi^{(2)}(1) \right) \right] \\ &+ \frac{1}{\beta_2^*} \left(\phi^{(3)}(1) + (\phi^{(1)}(1))^3 - 2\phi^{(1)}(1)\phi^{(2)}(1) - \phi^{(1)}(1)\beta_2^* \right) \\ &= 1,283408 \end{aligned} \quad (\text{A.62})$$

Pelas equações A.31 e A.33, teremos

$$\begin{aligned} \mathbb{C}_f(l_f(\alpha), l_g(\beta^*)) &= \mathbb{C}_f \left(\log Y + \frac{(\log Y - \beta_1^*)^2}{2\beta_2^*}, \frac{Y}{\alpha} \right) \\ &= \mathbb{C}_f \left(\log Y, \frac{Y}{\alpha} \right) \\ &+ \mathbb{C}_f \left(\frac{(\log Y - \beta_1^*)^2}{2\beta_2^*}, \frac{Y}{\alpha} \right) \\ &= 1 + \frac{1}{2\beta_2^*} \left(\phi^{(2)}(2) - \phi^{(2)}(1) - 2\phi^{(1)}(1) \right) = 1. \end{aligned} \quad (\text{A.63})$$

Logo,

$$\mathbb{V}_f(l(\alpha), l(\beta^*)) = 0,283408 \quad (\text{A.64})$$

Assim, teremos que o coeficiente de variação para $l_f(\alpha) - l_g(\beta^*)$ será

$$CV(f, g) = \frac{\sqrt{0,283408}}{0,09057302} = 5,877697 \quad (\text{A.65})$$

A.5 O Teste Lognormal versus Weibull

Consideremos H_f como sendo a distribuição lognormal e H_g como sendo a Weibull. Sob H_f , teremos

$$\beta_1^* = [\mathbb{E}_f(Y^{\beta_2^*})]^{1/\beta_2^*} = e^{\alpha_1 + \frac{\alpha_2 \beta_2^*}{2}} \quad (\text{A.66})$$

$$\beta_2^* = \sqrt{\frac{1}{\alpha_2}} \quad (\text{A.67})$$

$$\begin{aligned} \mathbb{E}_f(l_f(\alpha) - l_g(\beta^*)) &= -\frac{1}{2} \log(2\pi\alpha_2) - \frac{1}{2\alpha_2} \mathbb{E}_f[(\log Y - \alpha_1)^2] \\ &- \log \beta_2^* - \beta_2 \mathbb{E}_f(\log Y) + \beta_2 \log \beta_1 + \mathbb{E}_f \left(\frac{Y}{\beta_1} \right)^{\beta_2} \\ &= -\frac{1}{2} \log(2\pi\alpha_2) + \frac{1}{2} - \log \beta_2^* + \frac{\alpha_2 \beta_2^{*2}}{2} \end{aligned} \quad (\text{A.68})$$

$$= -\frac{1}{2} \log(2\pi) + 1 = 0,08106147 \quad (\text{A.69})$$

e

$$\begin{aligned}
\mathbb{V}_f(l_f(\alpha) - l_g(\beta^*)) &= \mathbb{V}_f \left(-\frac{1}{2\alpha_2}(\log Y - \alpha_1)^2 - \beta_2^* \log Y + \frac{Y^{\beta_2^*}}{\beta_1^* \beta_2^*} \right) \\
&= \mathbb{V}_f \left(\frac{1}{2\alpha_2}(\log Y - \alpha_1)^2 \right) + \mathbb{V}_f(\beta_2^* \log Y) + \mathbb{V}_f \left(\frac{Y^{\beta_2^*}}{\beta_1^* \beta_2^*} \right) \\
&+ \frac{\beta_2^*}{\alpha_2} \mathbb{C}_f(\log Y, (\log Y - \alpha_1)^2) - \frac{1}{\alpha_2(\beta_1^*)^{\beta_2^*}} \mathbb{C}_f \left((\log Y - \alpha_1)^2, Y^{\beta_2^*} \right) \\
&- \frac{2\beta_2^*}{(\beta_1^*)^{\beta_2^*}} \mathbb{C}_f(\log Y, Y^{\beta_2^*}) \\
&= e^{\beta_2^{*2}\alpha_2} - \frac{1}{2} - 2\alpha_2\beta_2^{*2} = e - \frac{3}{2} = 1,218282 \tag{A.70}
\end{aligned}$$

Portanto, o coeficiente de variação de $l_f(\alpha) - l_g(\beta^*)$ será

$$CV(f, g) = \frac{\sqrt{1,218282}}{0,08106147} = 13.61631 \tag{A.71}$$

Agora, consideremos H_f como tendo distribuição Weibull e H_g tendo distribuição lognormal. Sob H_f , aplicando as equações A.50 e A.44 nas equações dos pseudo-verdadeiros valores para a lognormal (equações A.9 e A.10), teremos

$$\beta_1^* = \log \alpha_1 + \frac{\phi^{(1)}(1)}{\alpha_2} \tag{A.72}$$

$$\begin{aligned}
\beta_2^* &= \frac{1}{\alpha_2^2} \mathbb{E}_U \left[(\log U - \phi^{(1)}(1))^2 \right] \\
&= \frac{1}{\alpha_2^2} \left[\phi^{(2)}(1) - \phi^{(1)}(1)^2 \right] = \frac{\psi^{(1)}(1)}{\alpha_2^2} \tag{A.73}
\end{aligned}$$

Utilizando as equações A.53, A.44 e A.41 teremos

$$\begin{aligned}
\mathbb{E}_f(l_f(\alpha) - l_g(\beta^*)) &= \mathbb{E}_f(l_f(\alpha)) - \mathbb{E}(l_g(\beta^*)) \\
&= \log \alpha_2 - \log \alpha_1 + \phi^{(1)}(1) \left(1 - \frac{1}{\alpha_2} \right) - 1 \\
&+ \mathbb{E}_f \left[\frac{1}{2} \log(2\pi\beta_2^*) + \log Y + \frac{1}{2\beta_2^*} (\log Y - \beta_1^*)^2 \right] \\
&= \phi^{(1)}(1) + \frac{1}{2} \log(2\pi) + \frac{1}{2} \log \psi^{(1)}(1) = 0,09057302. \tag{A.74}
\end{aligned}$$

Referências Bibliográficas

- A. B. Antognini and A. Giovagnoli. A new 'biased coin design' for the sequential allocation of two treatments. *Appl. Statist.*, 54(4):651–664, 2004.
- M. I. Araujo and B. B. Pereira. A comparison of bayes factors for separated models: Some simulation results. *Communications in Statistics - Simulation and Computation*, 36(2):297–309, 2007.
- A. C. Atkinson. A method for discriminating between models (with discussion). *J. R. Statist. Soc. B*, 32:323–353, 1970.
- A. C. Atkinson. Optimum biased coin designs for sequential clinical trials with prognostic factors. *Biometrika*, 69(4):61–67, 1982.
- A. C. Atkinson. The comparison of designs for sequential clinical trial with covariate information. *J. R. Statist. Soc. A*, 165:349–373, 2002.
- D. Blackwell and J. L. Hodges Jr. Design for the control of selection bias. *The Annals of Mathematical Statistics*, 28(2):449–460, 1957.
- K. A. Clarke. Nonparametric model discrimination in international relations. *The Journal of Conflict Resolution*, 47(1):72–93, 2003.
- K. A. Clarke and C. S. Signorino. Discriminating methods: Tests for non-nested discrete choice models. *Political Studies*, 58:368–388, 2010.
- N. Coulibaly and B. W. Brorsen. Monte carlo sampling approach to testing nonnested hypotheses: Monte carlo results. *Econometric Review*, 18(2):195–209, 1999.
- D. R. Cox. Tests of separate families of hypotheses. In *Proc. 4th Berkeley Symp.*, 1961.
- D. R. Cox. Further results on tests of separate families of hypotheses. *Journal of the Royal Statistical Society. Series B*, 24(2):406–424, 1962.

- A. Dameus, F. R. G-C, W. Brorsen, and K. P. Sukhdial. Aids versus rotterdam demand system: A cox test with parametric bootstrap. *Journal of Agricultural and Resource Economics*, 27(2):335–347, 2002.
- B. Efron. Forcing sequential experiments to be balanced. *Biometrika*, 58:403–417, 1971.
- R. A. Fisher. *The design of experiments*. Oliver and Boyd, 1956.
- V. Fossaluzza, Diniz J. B., B. P. Pereira, Miguel E. C., and C. A. B. Pereira. Sequential allocations to balance prognostic factors in psychiatric clinical trial. *Clinical Science*, 64(6):511–519, 2009.
- W. Hu, K. Mengersen, S. Y. Fu, and S. Tong. The use of zip and cart to model cryptosporidiosis in relation to climatic variables. *Int J Biometeorol*, 54:433–440, 2010.
- M. Lu, G. E. Mizon, and C. Monfardini. Simulation encompassing: testing non-nested hypotheses. *Oxford Bulletin of Economics and Statistics*, (70):781–806, 2008.
- B. Pesaran M. H. Pesaran. A simulation approach to the problem of computing cox’s statistic fo testing nonnested models. *J. Econometrics*, (57):377–392, 1993.
- B. B. Pereira. *Some Results on Tests of Separate Families of Hypotheses*. PhD thesis, University of London, 1976.
- B. B. Pereira. Empirical comparisons of some tests of separate families of hypotheses. *Metrika*, 25:219–234, 1978.
- B. B. Pereira. Testes para discriminar entre as distribuições lognormal, gama e weibull. *Inter-American Statistical Institute*, 33(120):41–46, 1979.
- B. B. Pereira. *Encyclopedia of Biostatistics*, volume 5, chapter Separate Families of Hypotheses, pages 4069–4074. Wiley, 1998.
- M. H. Pesaran. On the general problem of model selection. *The Review of Economic Studies*, 41(2):153–171, 1974.
- M. H. Pesaran. Comparison of local power of alternative tests of non-nested regression models. *Econometrica*, 50(5):1287–1305, 1982.
- M. H. Pesaran and A. S. Deaton. Testing non-nested nonlinear regression models. *Econometrica*, 46(3):677–694, 1978.

- S. J. Pocock and R. Simon. Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. *Biometrics*, (31):103–105, 1975.
- F. A. R. Roja, S. R. Lancheros, and F. L. Neto. Teste de cox para hipóteses separadas com os modelos weibull e bi-weibull. *49a Reunião da RBRAS*, 2004.
- N. W. Scott, G. C. McPherson, C. R. Ramsay, and M. K. Campbell. The method of minimization for allocation to clinical trials: a review. *Controlled Clinical Trials*, (23):667–674, 2002.
- Student. Comparison between balanced and random arrangements of field plots. *Biometrika*, 29:363–379, 1937.
- D. R. Taves. Minimization: a new method of assigning patients to treatment and control groups. *Clin. Pharm. Ther.*, (15):443–453, 1974.
- Q. H. Vuong. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica*, 57(2):307–333, 1989.
- L. Wei-Yin. A new method for testing separate families of hypotheses. *Journal of the American Statistical Association*, 80(390):362–368, 1985.
- H. White. Regularity conditions for cox’s test of non-nested hypotheses. *Journal of Econometrics*, (19):301–318, 1982.