



Universidade de Brasília
IE – Departamento de Estatística
Estágio Supervisionado 2

APLICAÇÃO DE TÉCNICAS DE ANÁLISE DE SOBREVIVÊNCIA EM PACIENTES SUBMETIDOS À INTERVENÇÃO CORONÁRIA PERCUTÂNEA

Priscila Nascimento de Alcântara Garcia

Relatório final

Orientador: Prof. Dr. Antônio Eduardo Gomes
Coorientador: Profa. Dra. Juliana Betini Fachini

Brasília
2013

APLICAÇÃO DE TÉCNICAS DE ANÁLISE DE SOBREVIVÊNCIA EM PACIENTES SUBMETIDOS À INTERVENÇÃO CORONÁRIA PERCUTÂNEA

Priscila Nascimento de Alcântara Garcia

Relatório apresentado à disciplina Estágio Supervisionado 2 do curso de graduação em Estatística, Departamento de Estatística, Universidade de Brasília, como parte dos requisitos necessários para o grau de Bacharel em Estatística.

Orientador: Prof. Antônio Eduardo Gomes
Coorientador: Profa. Juliana Betini Fachini

Brasília
2013

Sumário

Lista de Figuras	4
Lista de Tabelas	6
1. Introdução	7
2. Objetivos e Metodologia	8
3. Referencial Teórico	10
4. Conceitos básicos de análise de sobrevivência	13
4.1. Função de sobrevivência	13
4.2. Fração de cura	15
4.3. Função de Taxa de falha ou Função de Risco	16
4.4. Função de risco acumulado	17
4.5. Estimação da Função de sobrevivência	17
4.5.1. Estimador de Kaplan-Meier	18
4.5.2. Quantidades básicas obtidas a partir de $S(t)$	19
4.6. Comparação de Curvas de sobrevivência	19
4.7. Modelos Probabilísticos	20
4.8. Função Taxa de Falha Empírica	22
4.9. Estimação dos Parâmetros dos Modelos	22
4.9.1. Método de Máxima Verossimilhança Clássico	23
4.10. Intervalo de Confiança	24
4.11. Teste de Hipóteses	24
4.11.1. Teste de Wald	24
4.11.2. Teste da Razão de Verossimilhança	24
4.11.3. Teste Escore	25
4.12. Escolha do Modelo Probabilístico	25
4.12.1. Método Gráfico	25
4.12.2. Comparação de modelos	26
4.13. Critérios de seleção de modelos	27
4.14. Modelos de Regressão Paramétricos	27
4.14.1. Reparametrização da distribuição de Probabilidade	27
4.14.2. Modelos de Regressão de locação e escala	27
5. Análise dos dados	29
5.1. Analisando as variáveis	30
5.1.1 Idade	30
5.1.2 Regressão logística gênero e óbito	30

5.1.3 Regressão logística gênero e reintervenção	32
5.1.4 Comorbidades.....	33
5.2. Variável Resposta: Óbito	40
5.3. Variável Resposta: Reintervenção	50
5.4. Variável Creatinina	57
Conclusão.....	59
Referências bibliográficas.....	60

Lista de Figuras

Figura 1 – Forma típica da função de sobrevivência	14
Figura 2 – Função de sobrevivência imprópria	14
Figura 3 – Funções de taxa de falha	17
Figura 4 – Formas que a função TTT pode assumir	22
Figura 5 – Curvas de sobrevivência estimada por Kaplan-Meier e curvas de sobrevivência estimada dos modelos: Weibull, Exponencial, Log Normal e Log Logístico	26
Figura 6 – Histograma da idade	30
Figura 7 – Box-plot comparativo para a variável gênero com relação aos tempos de falha óbito.....	41
Figura 8 – Box-plot comparativo para a variável quantidade de artérias tratadas com relação aos tempos de falha óbito	42
Figura 9 – Curva de sobrevivência Kaplan-Meier considerando o evento de interesse o óbito.....	43
Figura 10 - TTT plot óbito	44
Figura 11 – Estimativas de Kaplan-Meier considerando o evento de interesse o óbito para a variável quantidade de artérias tratadas. Os tempos representados por + mostram onde ocorreram censuras em cada grupo 1	44
Figura 12 – Estimativas de Kaplan-Meier considerando o evento de interesse o óbito para a variável gênero. Os tempos representados por + mostram onde ocorreram censuras em cada gênero	46
Figura 13 – Curvas de sobrevivência estimada por Kaplan-Meier considerando o evento de interesse o óbito e curvas de sobrevivência estimada dos modelos: Weibull, Log-normal e Log-logística.....	47
Figura 14 – Curvas de sobrevivência estimada por Kaplan-Meier considerando o evento de interesse o óbito e curvas de sobrevivência estimada dos modelos: Weibull e Weibull com fração de cura.....	49
Figura 15 – Box-plot comparativo para variável gênero com relação aos tempos de falha da reintervenção.....	50
Figura 16 – Box-plot comparativo para a variável quantidade de artérias tratadas com relação aos tempos de falha reintervenção	51
Figura 17 – Curva de sobrevivência Kaplan-Meier considerando o evento de interesse a reintervenção	51
Figura 18 - TTT plot reintervenção	52
Figura 19– Estimativas de Kaplan-Meier considerando o evento de interesse a reintervenção para a variável quantidade de artérias tratadas. Os tempos representados por + mostram onde ocorreram censuras em cada grupo	53
Figura 20 – Estimativas de Kaplan-Meier considerando o evento de interesse a reintervenção para a variável gênero. Os tempos representados por + mostram onde ocorreram censuras em cada grupo.....	54
Figura 21 – Curvas de sobrevivência estimada por Kaplan-Meier considerando o evento de interesse a reintervenção para a variável resposta reintervenção e curvas de sobrevivência estimada dos modelos: Weibull, Log-normal e Log-logística.....	55

Figura 22 – Curvas de sobrevivência estimada por Kaplan-Meier considerando o evento de interesse a reintervenção para a variável resposta reintervenção e curvas de sobrevivência estimada dos modelos: Weibull e Weibull com fração de cura..... 57

Figura 23 – Box-plot comparativo para a variável creatinina com relação ao gênero .. 58

Lista de Tabelas

Tabela 1: Tabela de contingência gerada no tempo t_j	20
Tabela 2 – Variáveis e seus respectivos rótulos	29
Tabela 3 – Tabela de Contingência: Óbito x Gênero	31
Tabela 4 – Avaliar significância do parâmetro: $\beta=0$: Óbito x Gênero	31
Tabela 5 – Análises de estimação da Máxima verossimilhança: Óbito x Gênero	31
Tabela 6 – Estimação do Odds Ratio: Óbito x Gênero	32
Tabela 7 – Associação entre as probabilidades previstas e as respostas observadas: Óbito x Gênero	32
Tabela 8 – Tabela de Contingência: Reintervenção x Gênero.....	32
Tabela 9 – Avaliar significância do parâmetro: $\beta=0$ 1: Reintervenção x Gênero	33
Tabela 10 – Análises de estimação da Máxima verossimilhança: Reintervenção x Gênero	33
Tabela 11 – Estimação do Odds Ratio: Reintervenção x Gênero	33
Tabela 12 – Associação entre as probabilidades previstas e as respostas observadas: Reintervenção x Gênero	33
Tabela 13 – Frequência dos pacientes portadores de ICC por gênero	34
Tabela 14 – Frequência dos pacientes etílico por gênero.....	35
Tabela 15 – Frequência dos pacientes obesos por gênero	35
Tabela 16 – Frequência dos pacientes tabagistas por gênero	36
Tabela 17 – Frequência dos pacientes diabéticos por gênero	37
Tabela 18 – Frequência dos pacientes sedentários por gênero.....	37
Tabela 19 – Frequência dos pacientes portadores de HAS por gênero	38
Tabela 20 – Frequência dos pacientes dislipidêmicos por gênero.....	39
Tabela 21 – Frequência dos pacientes portadores de IRC por gênero	39
Tabela 22 – Frequência dos pacientes portadores de hipotireoidismo por gênero.....	40
Tabela 23 – Frequência do tipo de stent implantado	40
Tabela 24 - Resultados dos testes de Wilcoxon utilizados para as comparações de grupos, dois a dois, considerando o evento de interesse o óbito	45
Tabela 25 - Testes não-paramétricos para comparação de curvas de sobrevivência obtidas para o gênero feminino e masculino.....	46
Tabela26 - Comparando distribuições através das medidas AIC, AICc e BIC considerando o evento de interesse o óbito.....	47
Tabela 27 – Resultados dos testes logrank utilizados para as comparações de grupos, dois a dois, considerando o evento de interesse a reintervenção.....	53
Tabela 28 - Comparando distribuições através das medidas AIC, AICc e BIC, considerando o evento de interesse a reintervenção.....	55

1. Introdução

A Análise de sobrevivência é uma das áreas da estatística que mais cresceu nas últimas décadas em razão ao desenvolvimento e aprimoramento de técnicas estatísticas combinado com computadores mais velozes. Essa técnica é adequada para situações de estudo em que o tempo de ocorrência de um evento é a variável resposta. A principal característica de dados de sobrevivência é a presença de censura, que é a observação parcial da resposta. (COLOSIMO; GIOLO, 2006)

Nesse trabalho foram analisados pacientes submetidos à angioplastia transluminal coronária, tratamento não cirúrgico, porém invasivo, das obstruções das artérias coronárias por meio de cateter balão, com o objetivo de aumentar o fluxo de sangue para o coração. Foi analisado, separadamente, duas variáveis respostas: o tempo entre a angioplastia e o óbito, e o tempo entre a angioplastia e uma segunda intervenção.

Além de existirem poucas pesquisas analisando a sobrevida dos pacientes submetidos à angioplastia e devido à presença de muitas censuras e, portanto, muitos indivíduos não suscetíveis ao evento de interesse, objetivou-se avaliar a sobrevida desses pacientes utilizando um modelo bem estruturado com fração de cura.

Estudos envolvendo uma fração de sobreviventes de longa duração são de grande importância em medicina, uma vez que é esperado após um tratamento para um determinado tipo de doença, que exista uma alta proporção de indivíduos com certa característica. Nesse sentido, é importante estudar a proporção de pacientes que permanecem imunes e analisar o efeito de covariáveis, tanto sobre o tempo de sobrevivência quanto sobre a taxa (ou fração) de sobreviventes curados.

2. Objetivos e Metodologia

Para esse estudo foram observados os casos de todos os pacientes submetidos à angioplastia em um hospital de referência em Brasília no período de janeiro a dezembro de 2011. Foram incluídos no estudo 97 pacientes com idade média de 63 anos, sendo a maioria do gênero masculino (69,4%). O interesse inicial do estudo era avaliar a evolução clínica e sobrevida dos pacientes por um período de 1 ano após a realização do procedimento. Após a observação de muitas censuras, esse tempo foi estendido até junho de 2013. Para verificar se o paciente sofreu algum dos eventos de interesse (segunda intervenção e/ou óbito) outros dois hospitais de Brasília foram considerados.

Apesar dos tempos de observação de cada paciente não serem iguais, após os 12 meses iniciais, não foram observadas a ocorrência do evento de interesse em muitos pacientes.

Apesar dos dados apresentarem muitas censuras, é possível deduzir que esses pacientes de fato não sofreram nenhuma dos dois eventos de interesse. Isso ocorre devido à “fidelidade” do paciente para com o seu médico. Após sofrer uma angioplastia, o paciente, em geral, demonstra tamanha gratidão que provavelmente não trocaria de médico.

Foi verificado se o paciente também realizou esse procedimento antes do início do estudo. Verificar se o paciente é portador de alguma comorbidade, ou seja, se o paciente é diabético, dislipidêmico e se tem Hipertensão Arterial Sistêmica (HAS), são algumas das principais covariáveis a serem analisadas.

Os dados foram coletados pelo hospital conforme os procedimentos foram realizados e foram registrados acessando os prontuários após a intervenção. O seguimento dos pacientes no estudo foi realizado também acessando o prontuário dos mesmos no hospital de intervenção e em dois outros. A significância de cada variável foi verificada com o objetivo de analisar as maiores correlações e, se possível, viabilizar em estudos futuros outras análises como a insuficiência renal após intervenção (Nefropatia).

O objetivo desse trabalho é basicamente realizar um controle de qualidade do serviço realizado nos pacientes submetidos à angioplastia. Foram comparados os resultados do estudo com os dados da literatura médica. Estudos recentes têm demonstrado que o percentual de mortes após angioplastia primária é de 6,4% em 94 meses, e ocorrem 29% de eventos cardiovasculares maiores em 73 meses. Após um Infarto agudo do miocárdio (IAM), independente do uso de stent, a prevalência de morte aumenta para 8,2%. (TREVISIO et al, 2012)

Os pacientes foram separados em grupos de acordo com a quantidade de artérias tratadas. Foi analisado se a covariável grupo influência no tempo de sobrevida e na reintervenção devido à reestenose.

Foram verificados quais são as principais características em comum entre os pacientes e se elas estão diretamente relacionadas com o Infarto agudo do Miocárdio, seu tratamento e reincidências, e quais fatores de risco são mais recorrentes.

Nesse estudo foram considerados dois eventos de interesse: o tempo até a ocorrência de uma segunda intervenção e o tempo até o óbito. Uma análise de sobrevivência multivariada poderia ser considerada, mas não será abordada nesse trabalho.

Como limitações do estudo destaca-se o pequeno tamanho amostral (97 pacientes), devido ao pouco tempo de implantação do serviço de hemodinâmica no Hospital.

Para a análise de dados foram utilizados os softwares R, SAS e SPSS.

3. Referencial Teórico

O infarto agudo do miocárdio (IAM), conhecido popularmente como infarto do coração, enfarte ou ataque cardíaco, refere-se à morte de parte do músculo cardíaco (miocárdio), que ocorre de forma rápida (ou aguda) devido à obstrução do fluxo sanguíneo das artérias coronárias. O IAM está associado às altas taxas de mortalidade, chegando a aproximadamente 30% no ambiente pré-hospitalar em decorrência principalmente de arritmias ventriculares. (ALBERT EINSTEIN - SOCIEDADE BENEFICENTE ISRAELITA BRASILEIRA, 2013) No Brasil, estima-se a ocorrência de 300 mil a 400 mil casos anuais de infarto, e que a cada 5 a 7 casos, ocorra um óbito. (HOSPITAL DAS CLÍNICAS DA UNIVERSIDADE FEDERAL DE MINAS GERAIS, 2013)

O risco de Doença Arterial Crônica (DAC) aumenta progressivamente após cada década acima de 40 anos de idade, sendo o gênero masculino um fator de risco adicional. Além disso, podem ser considerados como fatores de risco o Tabagismo, a Hipertensão arterial, a Dislipidemia, História familiar de DAC precoce (homem < 55 anos e mulher < 65 anos) e Diabetes. Em pacientes mais jovens ou com poucos fatores de risco, o uso de cocaína ou metanfetaminas deve ser investigado. (HOSPITAL DAS CLÍNICAS DA UNIVERSIDADE FEDERAL DE MINAS GERAIS, 2013)

O Cateterismo Cardíaco - também conhecido como Cinecoronariografia ou Estudo Hemodinâmico - é um exame invasivo que pode ser realizado de forma eletiva (sem caráter de urgência), para confirmar a presença de obstruções nas artérias coronárias em pacientes com suspeita de aterosclerose (formação de placas de gordura nas artérias) ou avaliar o funcionamento das válvulas e do músculo cardíaco. Pode ser aplicado emergencialmente em casos de infarto agudo do miocárdio para determinar a exata localização da obstrução e planejar a melhor estratégia de intervenção. O exame também possibilita o reconhecimento de defeitos congênitos em recém-nascidos e crianças.

Para que o médico localize as áreas de bloqueio nas artérias e saiba exatamente a localização da obstrução, são obtidas imagens radiológicas em tempo real. Este procedimento envolve injeção de contraste iodado, substância radiopaca empregada em exames radiológicos, amplamente utilizada para fins diagnósticos, que “desenha” a parede da artéria, aparecendo o seu contorno quando as imagens são tiradas por meio de um equipamento que emite feixes de raios x. Quando captados formam imagens dinâmicas. (NOGUEIRA, 2013)

Realizada sob anestesia local, a técnica de Intervenção Coronária Percutânea (ICP), angioplastia transluminal coronária, é o tratamento não cirúrgico, porém invasivo, das obstruções das artérias coronárias por meio de cateter balão, com o objetivo de aumentar o fluxo de sangue para o coração. Esse procedimento é semelhante ao cateterismo, porém tendo como principal diferença a função terapêutica.

A ICP é realizada com o auxílio de um delicado e sofisticado cateter-balão, disponíveis em diversos diâmetros e comprimentos, que é posicionado no local da lesão obstrutiva. A seguir este balão é expandido (insuflado) "esmagando" assim a placa contra as paredes do vaso, desobstruindo a artéria e permitindo que o fluxo sanguíneo retorne ao normal. Trata-se de um tratamento extremamente seguro e eficiente, onde as complicações graves ocorrem na ordem de 2 a 3%. (FREITAS (FMUP), 2013).

Após a desobstrução da artéria coronária, por meio da angioplastia, procede-se, ou não, ao implante de uma prótese endovascular (para ser utilizada no interior dos vasos) conhecida como 'stent' - pequeno tubo de metal, semelhante a um pequeníssimo bobo de cabelo, usado para manter a artéria aberta. Atualmente existem dois tipos de stents: os convencionais e os farmacológicos ou recobertos com drogas. (ALBERT EINSTEIN - SOCIEDADE BENEFICENTE ISRAELITA BRASILEIRA, 2013).

Os stents convencionais podem acarretar um processo cicatricial exacerbado que leva a reestenose (reobstrução) do vaso em 10 a 20% dos casos, ou seja, a obstrução retorna na mesma localização anterior, geralmente no período entre 3 e 6 meses após o procedimento. Para evitar esse processo cicatricial surgiram os stents farmacológicos que são constituídos do mesmo material metálico acrescido de um medicamento de liberação lenta no local de implante, reduzindo-se o processo de cicatrização e evitando-se a reestenose. Portanto, segundo a literatura médica, pacientes que recebem o implante de stent farmacológico tem probabilidade menor de sofrer uma segunda intervenção. (LAVALL, 2010). Os stents farmacológicos podem ser divididos em primeira e segunda geração.

O cateterismo, a angioplastia e o implante de stent são procedimentos invasivos e por isso precisam de uma via de acesso para que o cateter atinja as artérias coronárias. Na maioria dos casos a artéria femoral (região da virilha) é utilizada. Outras opções seriam a artéria braquial (no nível da prega do cotovelo) – não foi observado nenhum caso no estudo em questão – e a artéria Radial (localizada na região do punho). Porém nem todos os pacientes com indicação de angioplastia coronária podem utilizar a artéria radial como via de acesso. Existem alguns critérios para selecionar o paciente como a condição anatômica da artéria radial, do trajeto arterial até as coronárias e o tipo de lesão e coronária a ser tratada. Porém, em torno de 30% a 50% dos casos essa via pode ser utilizada variando conforme as condições de cada paciente e os critérios adotados pelo médico. Os pacientes selecionados para utilização dessa via de acesso contam com algumas vantagens durante e pós-procedimento como menor desconforto no local da punção comparado ao acesso femoral, imediata deambulação após o procedimento, alta hospitalar mais precoce e menor risco de complicações vasculares no local da punção. (GUBOLINO, 2013).

Em um estudo randomizado, envolvendo três centros na Itália e um centro na Holanda que incluiu pacientes com IAM, foi avaliado se a via radial está associada a um melhor desfecho quando comparada à via femoral nos pacientes submetidos à angioplastia. Concluíram que o uso da via radial está associado à diminuição de morbidade e mortalidade cardíaca quando realizada em centros com experiência nessa via. (PINTON, 2013).

Uma das possíveis consequências desses procedimentos é a Nefropatia induzida por contraste (NIC) que é o comprometimento da função renal após o uso endovenoso de contraste. É a terceira causa mais frequente de insuficiência renal em pacientes hospitalizados, com uma incidência média de 10,2%. (NEFROPATIA INDUZIDA POR CONTRASTE: COMO PREVENIR?, 2002).

4. Conceitos básicos de análise de sobrevivência

A análise de sobrevivência tem, geralmente, como variável resposta o tempo até a ocorrência de um evento de interesse. O tempo de início do estudo deve ser precisamente definido. Nesse caso foi utilizada uma coorte aberta, ou seja, os pacientes entrarão no estudo em momentos diferentes (data da realização do procedimento).

O evento de interesse precisa ser definido de forma clara em todos os estudos cujas técnicas de análise de sobrevivência foram utilizadas. Nesse trabalho foram definidos alguns eventos de interesse, como a reintervenção e o óbito. Para ambos os eventos o tempo foi medido em dias.

O evento de interesse pode ocorrer devido a uma única causa ou devido a duas ou mais. Situações em que causas de falhas competem entre si são denominadas na literatura de *riscos competitivos* e, conseqüentemente como riscos múltiplos. (Prentice et al., 1978). Esse caso não será abordado nesse trabalho.

Uma característica decorrente de estudos clínicos é a presença de observações incompletas ou parciais, denominadas censura. No estudo em questão foram predominantemente observadas censuras à direita do tipo 1, ou seja, como o tempo de duração do estudo foi fixado, foram todos os casos os quais não ocorre o evento de interesse. Censura à direita aleatória também foi observada, corresponde a todos os casos onde ocorre a perda de acompanhamento do paciente.

Quando n indivíduos entram em tempos diferentes no estudo e é pré-estabelecido o tempo do término do experimento, então para cada indivíduo é definido um tempo de observação, pois o tempo estabelecido começa a ser contado no momento em que cada indivíduo entra no estudo. Essa situação caracteriza censuras múltiplas do tipo 1.

Mesmo censurados, todos os resultados provenientes de um estudo de sobrevivência devem ser utilizados na análise estatística, pois mesmo sendo incompletas, as observações censuradas fornecem informações sobre o tempo de vida de pacientes e, além disso, a omissão das censuras no cálculo das estatísticas de interesse pode acarretar conclusões viesadas.

4.1. Função de sobrevivência

A função de sobrevivência é definida como a probabilidade de uma observação não falhar até certo tempo t , ou seja, a probabilidade de um indivíduo sobreviver ao tempo t . Em termos probabilísticos é definida como:

$$S(t) = P(T > t) = 1 - F(t),$$

em que $F(t)$ é a função de distribuição acumulada. A função $S(t)$ é uma função monótona decrescente e contínua. (LAWLESS, 2003)

A função de sobrevivência tem como propriedades:

$$\lim_{t \rightarrow 0} S(t) = 1 \text{ e } \lim_{t \rightarrow \infty} S(t) = 0.$$

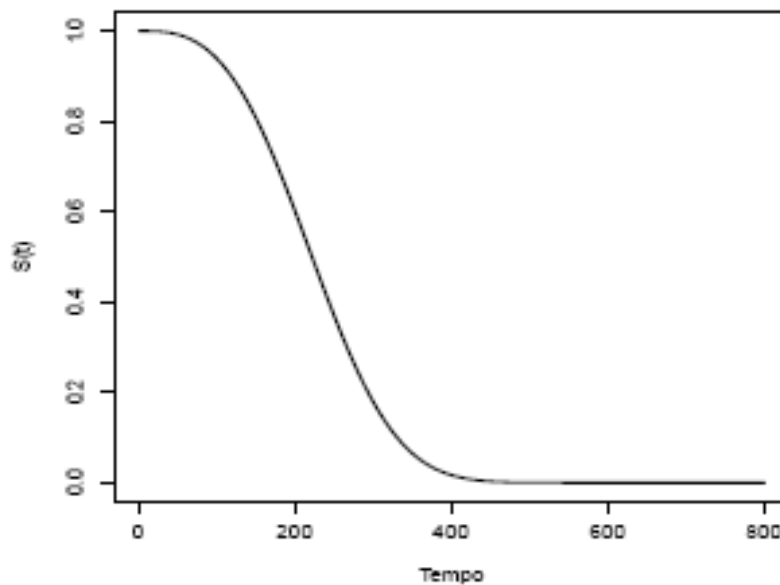


Figura 1 – Forma típica da função de sobrevivência

Quando o $\lim_{t \rightarrow \infty} S(t) = p$, sendo p uma probabilidade qualquer, a função de sobrevivência é uma função imprópria. Nessa situação utilizam-se modelos de sobrevivência com fração de cura para modelar os dados. Um exemplo dessa situação pode ser observado na figura 2.

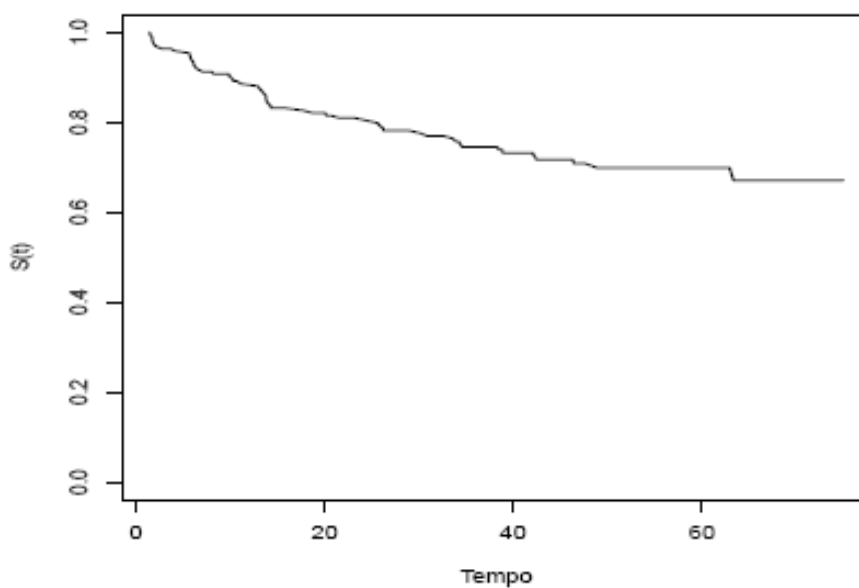


Figura 2 – Função de sobrevivência imprópria

4.2. Fração de cura

Os principais modelos de análise de sobrevivência têm como pressuposto o fato de que os indivíduos do estudo irão, em algum momento, experimentar o evento de interesse definido.

A função de sobrevivência é considerada própria quando todos os indivíduos são suscetíveis ao evento de interesse. Entretanto, existem situações que, para uma proporção de indivíduos, o evento de interesse não ocorrerá. A função é considerada imprópria quando à medida que o tempo tende a infinito a função não tende a zero. Isso pode indicar que existe uma proporção de indivíduos curados ou imunes. Dessa forma, a função de sobrevivência juntamente com suas propriedades são muito importantes para a identificação de dados com a presença de indivíduos curados.

Nesse caso, considerar os modelos de sobrevivência usuais, que assumem que a função de sobrevivência converge para zero quando a variável tempo tende a infinito (função de sobrevivência própria), podem não ser adequados. Para modelar esse tipo de dados, modelos com fração de cura são mais apropriados (Fachini, 2011).

Esses indivíduos não suscetíveis, definidos como curados, aparecem na base de dados como observações censuradas, visto que o evento de interesse não é observado, ou seja, não houve falha.

A presença de indivíduos curados na base de dados é verificada ao construir o gráfico da função de sobrevivência empírica, estimada pelo método de Kaplan-Meier. Verifica-se no gráfico o comportamento da cauda direita e caso ela permaneça de maneira constante em um nível acima de zero por um período grande, conclui-se que há proporção de indivíduos curados (Fachini, 2011). A modelagem de fração de cura para o caso univariado pode ser abordada seguindo a metodologia introduzida por Berkson e Gage, (1952) que considera a construção de uma função de sobrevivência populacional na forma de mistura.

Quando a metodologia de proporção de curados introduzida por Berkson e Gage (1952) é adotada para modelar, utiliza-se a função de sobrevivência populacional em forma de mistura, uma vez que a população é dividida em duas partes: indivíduos suscetíveis, e indivíduos não suscetíveis ao evento de interesse. Assim, a função de sobrevivência populacional em forma de mistura é dada por:

$$S_{pop}(t) = (1 - \phi) + \phi S(t),$$

sendo $S(t)$ uma função de sobrevivência própria, $(1 - \phi)$ a probabilidade de indivíduos serem não suscetíveis (curados), e ϕ a probabilidade de indivíduos serem suscetíveis ao evento de interesse.

Pode-se notar ainda que a probabilidade de cura ($1 - \phi$) pode variar de indivíduo para indivíduo, pois é razoável assumir que esta pode depender de características individuais (covariáveis). Nesses casos, o efeito de covariáveis na probabilidade de cura é comumente formulado por um modelo logístico. Assim, se for definida uma variável indicadora Y_i , tal que $Y_i = 0$ se o indivíduo i for curado, segue que a probabilidade de cura pode ser modelada em termos das covariáveis z_i por meio da função logística, de modo que:

$$p_{0i} = 1 - \phi(z_i) = P(Y_i = 0|z_i) = \frac{\exp(z'_i \beta)}{1 + \exp(z'_i \beta)}$$

para $i = 1, \dots, n$, em que β denota o vetor de coeficientes de regressão.

A literatura sobre modelos de sobrevivência com fração de cura é extensa e está em rápido desenvolvimento. É possível destacar como referências fundamentais os livros de Maller & Zhou (1996) e Ibrahim *et al.* (2001), como também o artigo de Tsodikov *et al.* (2003) e o artigo de Cooner *et al.* (2007).

4.3. Função de Taxa de falha ou Função de Risco

A taxa de falha no intervalo $[t_1, t_2)$ é definida como a probabilidade de que o evento de interesse ocorra neste intervalo, dado que não ocorreu antes de t_1 , dividida pelo comprimento do intervalo. De forma geral, a taxa de falha no intervalo $[t, t + \Delta t)$ é expressa por:

$$\lambda(t) = \frac{S(t) - S(t + \Delta t)}{\Delta t S(t)}$$

Pode, ainda, ser expressa em termos da função densidade de probabilidade e da função de sobrevivência:

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

Assumindo Δt bem pequeno, $\lambda(t)$ representa a taxa de falha instantânea no tempo t condicional à sobrevivência até o tempo t . A taxa de falha é útil para descrever a distribuição do tempo de vida de pacientes. Ela representa a forma em que a taxa instantânea de falha, ou força de mortalidade, muda com o tempo.

A função de risco é mais informativa do que a função de sobrevivência. Diferentes funções de sobrevivência podem ter formas semelhantes, enquanto as respectivas funções de taxa de falha podem diferir drasticamente. Dessa forma, a modelagem da função taxa de falha é um importante método para dados de sobrevivência (COLOSIMO; GIOLO, 2006), pois pode ter forma crescente, decrescente, constante ou não monótona.

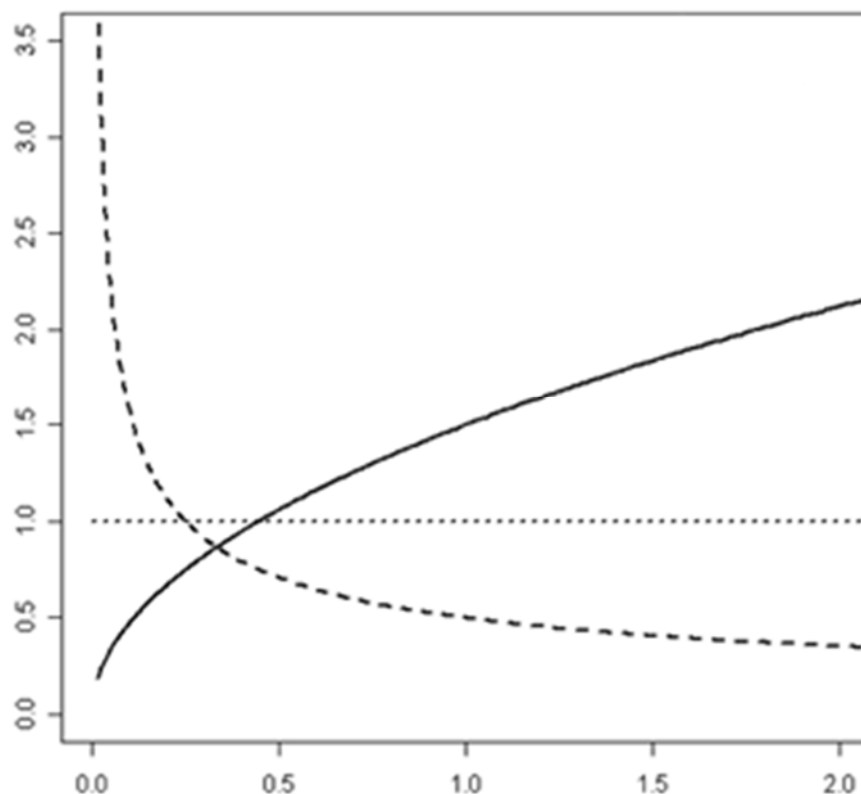


Figura 3 – Funções de taxa de falha

A função crescente indica que a taxa de falha do paciente aumenta com o transcorrer do tempo, mostrando que existe um efeito gradual de envelhecimento. A função constante indica que a taxa de falha não se altera com o passar do tempo. A função decrescente mostra que a taxa de falha diminui à medida que o tempo passa. (COLOSIMO; GIOLO, 2006)

4.4. Função de risco acumulado

A função de risco acumulado fornece a taxa de falha acumulada do indivíduo e é definida por:

$$\Lambda(t) = \int_0^t \lambda(u) du$$

Essa função não tem interpretação direta, mas é útil na avaliação da função de maior interesse que é taxa de falha, $\lambda(t)$. Isto acontece na estimação não-paramétrica em que $\Lambda(t)$ apresenta um estimador com propriedades ótimas e $\lambda(t)$ é difícil de ser estimada.

4.5. Estimação da Função de sobrevivência

O procedimento inicial para dados com censura é encontrar a estimativa para a função de sobrevivência e a partir dela estimar as estatísticas de interesse. Para realizar

essa estimativa podem ser considerados três estimadores: Kaplan-Meier, Nelson Aalen e Tabela de vida. As estimativas obtidas pelo estimador de Nelson Aalen são maiores ou iguais às obtidas pelo estimador de Kaplan-Meier. Nesse estudo foi utilizado o estimador de Kaplan-Meier por ser o estimador de máxima verossimilhança da função de sobrevivência.

4.5.1. Estimador de Kaplan-Meier

O estimador não paramétrico de Kaplan-Meier, também conhecido como estimador limite-produto, é sempre baseado em um número de intervalos igual ao número de falha distintos e se reduz a estimar q_j que é dado por:

$$\hat{q}_j = \frac{\text{n}^\circ \text{ de falhas em } t_{j-1}}{\text{n}^\circ \text{ de observações sob risco em } t_{j-1}}$$

para $j = 1, \dots, k + 1$, em que $t_{k+1} = \infty$

A expressão geral do Kaplan-Meier pode ser apresentada após estas considerações preliminares:

- $t_1 < t_2 < \dots < t_k$, os k tempos distintos e ordenados de falha;
- d_j é o número de falha em t_j , $j = 1, \dots, k$;
- n_j é o número de indivíduos sob risco em t_j , ou seja, os indivíduos que não falham e não foram censurados até o instante imediatamente anterior a t_j .

O estimador Kaplan-Meier é definido como:

$$\hat{S}(t) = \prod_{j:t_j < t} \left(\frac{n_j - d_j}{n_j} \right)$$

Por ser um estimador de máxima verossimilhança, $\hat{S}(t)$ é um estimador fracamente consistente e não viciado para amostras grandes. Breslow e Crowley (1974) mostram a consistência e a normalidade assintótica de $\hat{S}(t)$. Logo, $\hat{S}(t) \sim N\left(S(t), \text{Var}\left(\hat{S}(t)\right)\right)$.

Para que seja possível construir intervalos de confiança e realizar testes de hipóteses para $S(t)$, é necessário avaliar a precisão do estimador. A variância assintótica desse estimador é estimada pela fórmula de Greenwood expressa por:

$$\widehat{\text{var}}\left(\hat{S}(t)\right) = \left(\hat{S}(t)\right)^2 \sum_{j:t_j < t} \frac{d_j}{n_j(n_j - d_j)}$$

Um intervalo de confiança para $S(t)$ considerando $100(1 - \alpha)\%$ de confiança é dado por:

$$\hat{S}(t) \pm z_{\alpha/2} \sqrt{\widehat{Var}[\hat{S}(t)]}$$

onde $\alpha/2$ denota o $\alpha/2$ -percentil da distribuição Normal padrão, $\hat{S}(t)$ é o estimador de Kaplan-Meier e $\widehat{Var}[\hat{S}(t)]$ é o estimador da variância do estimador de Kaplan-Meier.

Deve-se observar que o intervalo de confiança de $S(t)$ não é muito indicado para amostras pequenas e, se t é um valor extremo, pode incluir valores fora do intervalo $(0,1)$. Uma alternativa é utilizar uma transformação para $S(t)$, como por exemplo, $g(\hat{S}(t)) = \log[-\log g(\hat{S}(t))]$, sugerida por Kalbfleish e Prentice (1978).

4.5.2. Quantidades básicas obtidas a partir de $\hat{S}(t)$

Tempo médio de vida: dado pela área sob a função de sobrevivência, sob a curva de Kaplan-Meier estimada. Não carrega a informação das censuras. Por isso, para casos com muitas censuras não funcionariam muito bem. Se o maior tempo observado for um tempo de censura, o tempo médio de vida é subestimado.

Tempo Mediano: obtido por meio da interpolação linear. Ou seja, utilizar retas nas estimativas de Kaplan-Meier, em vez de utilizar $\hat{S}(t)$ na forma escada. Como a censura dificulta o cálculo da média, a mediana recebe a preferência.

4.6. Comparação de Curvas de sobrevivência

Quando existem variáveis regressoras categóricas, o objetivo é comparar o efeito das categorias no tempo de sobrevivência. Para comparar grupos que sofreram tratamentos distintos, um procedimento natural seria usar os resultados assintóticos de $\hat{S}(t)$ para testar a igualdade das funções de sobrevivência em um determinado tempo t . Porém, estatísticas comumente usadas podem ser vistas como generalizações para dados censurados de testes não paramétricos. O teste de *Logrank* (Mantel, 1966) é o mais usado em análise de sobrevivência. Ele é apropriado quando a razão das funções de taxa de falha dos grupos a serem comparados é aproximadamente constante, ou seja, as populações tem a propriedade de riscos proporcionais.

Ao fazer um teste de igualdade de duas funções de sobrevivência, $S_1(t)$ e $S_2(t)$, suponha que no tempo t_j aconteça d_j falhas e que n_j indivíduos estejam suscetíveis a falhar em um tempo imediatamente inferior a t_j na amostra combinada. Assim, para cada t_j fixo, é possível construir a seguinte tabela de contingência 2×2 .

Tabela 1: Tabela de contingência gerada no tempo t_j

	Grupos		Totais
	1	2	
Ocorrer o evento de interesse	d_{1j}	d_{2j}	d_j
Não ocorrer o evento de interesse	$n_{1j} - d_{1j}$	$n_{2j} - d_{2j}$	$n_j - d_j$
Totais	n_{1j}	n_{2j}	n_j

Condicional à experiência de falha e censura até o tempo t_j (fixando as marginais de coluna) e ao número de falhas no tempo t_j (fixando as marginais de linha), a distribuição de d_{1j} é uma Hipergeométrica com média $E(d_{1j})$ e variância $V(d_{1j})$ dadas por:

$$E(d_{1j}) = d_j \frac{n_{1j}}{n_j} = w_{1j}$$

e

$$V(d_{1j}) = d_j \frac{n_{1j}}{n_j} \left(1 - \frac{n_{1j}}{n_j}\right) \frac{(n_j - d_j)}{(n_j - 1)} = V_{1j}.$$

A estatística $(d_{1j} - w_{1j})$ tem média zero e variância V_{1j} . Se as k tabelas de contingência forem independentes, a estatística do teste Logrank é definida por:

$$T = \frac{[\sum_{j=1}^k (d_{2j} - w_{2j})]^2}{\sum_{j=1}^k (V_j)_2}.$$

e tem aproximadamente distribuição Qui-Quadrado com 1 grau de liberdade.

Quando não é possível observar a proporcionalidade dos riscos, o teste de Wilcoxon é o mais adequado. Esse teste utiliza peso igual ao número de indivíduos suscetíveis a falha e coloca mais peso na proporção inicial do eixo do tempo.

O teste de Wilcoxon, em particular, pondera pelo número de indivíduos suscetíveis a falhar que depende da experiência de sobrevivência, assim como da de censura. Se o padrão de censura é nitidamente diferente nos grupos, então o teste pode rejeitar, não somente com base nas diferenças das sobrevivências entre os grupos mas devido ao padrão de censura.

4.7. Modelos Probabilísticos

Embora exista uma série de modelos probabilísticos utilizados em análise de dados de sobrevivência, alguns deles ocupam uma posição de destaque por sua

comprovada adequação a várias situações práticas. Entre estes modelos, é possível citar o exponencial, o Weibull e o log-normal.

Distribuição Exponencial: Caso particular da distribuição Weibull, apresenta um único parâmetro e é a única distribuição que se caracteriza por ter uma função de taxa de falha constante. Por isso é aplicável quando o tempo é curto. Quanto maior o risco, menor o tempo médio de sobrevivência e menor a variabilidade deste em torno da média.

Distribuição Weibull: Popular em aplicações práticas, a distribuição Weibull apresenta um grande variedade de formas, a sua função de taxa de falha é monótona, isto é, ela é crescente, decrescente ou constante. Sabendo que γ é o parâmetro de forma, a função de risco $\lambda(t)$ é estritamente crescente para $\gamma > 1$, estritamente decrescente para $\gamma < 1$ e constante para $\gamma = 1$. Para $\gamma = 1$, tem-se a função de taxa de falha da distribuição exponencial que é um caso particular da Weibull.

Distribuição Log-Normal: Se T tem distribuição Log-Normal com parâmetros μ e σ , então a variável $Y = \log(T)$ tem distribuição normal com parâmetros μ e σ . As funções de taxa de falha não são monótonas como as da distribuição Weibull. Elas crescem, atingem um valor máximo e depois decrescem. (Unimodais)

Distribuição Log-Logística: Se T tem distribuição Log-Logística com α (parâmetro de escala) e γ (parâmetro de forma), então a variável $Y = \log(T)$ tem distribuição logística com parâmetros $-\infty < \mu < \infty$ e $\sigma > 0$, em que $\gamma = 1/\sigma$ e $\alpha = \exp(\mu)$. Apresenta expressões explícitas para as funções de sobrevivência e risco. Essa última apresenta riscos unimodais.

Distribuição Gama: Possui uma função risco crescente ou decrescente convergindo para um valor constante quando t cresce. A função densidade da distribuição gama, que é caracterizada por dois parâmetros, k e α , em que $k > 0$ é chamado parâmetro de forma e $\alpha > 0$ de escala, é expressa por:

Distribuição Gama Generalizada: Esta distribuição é de grande utilidade, pois representa uma família paramétrica de distribuições. Com isso, ela é muito utilizada para discriminar modelos probabilísticos. A família Gama Generalizada possui as seguintes distribuições como casos particulares:

$$\begin{aligned} \text{Gama: } & f(t; \alpha, 1, k) \\ \text{Weibull: } & f(t; \alpha, \gamma, 1) \\ \text{Exponencial: } & f(t; \alpha, 1, 1) \\ \text{Lognormal: } & \text{quando } k \rightarrow \infty \\ \text{Qui - Quadrado: } & f(t; 2, n/2, 1) \end{aligned}$$

4.8. Função Taxa de Falha Empírica

Existem várias formas para o gráfico da função de taxa de falha da variável T assumir. Dessa forma, é importante utilizar uma metodologia adequada para identificar o modelo mais apropriado para esta variável. Uma técnica de verificação gráfica para ajuste do modelo é conhecida como TTT plot, proposto por Aarset (1987).

A curva TTT é obtida construindo um gráfico de

$$G(r/n) = \frac{[(\sum_{i=1}^r T_{i:n}) + (n-r)T_{r:n}]}{(\sum_{i=1}^n T_{i:n})},$$

por r/n , sendo que $r=1, \dots, n$, e $T_{i:n}$, $i=1, \dots, n$ são as estatísticas de ordem da amostra.

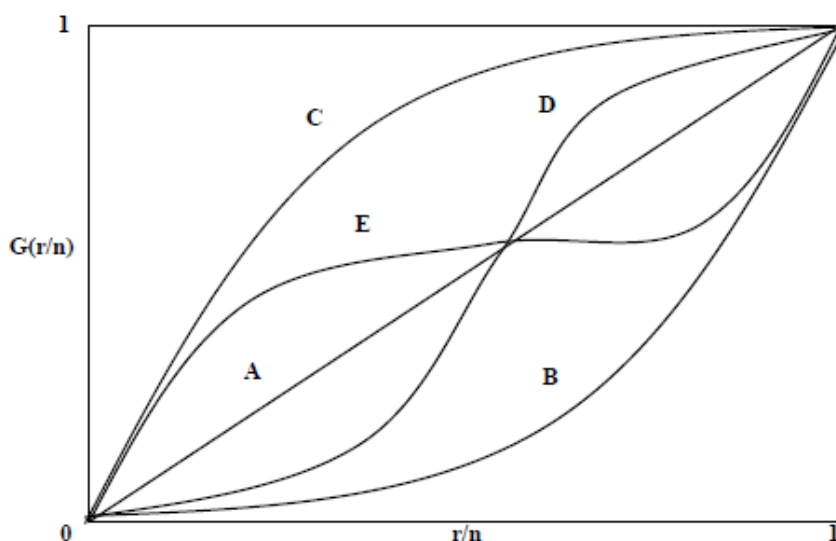


Figura 4 – Formas que a função TTT pode assumir

Uma reta diagonal (A) como resultado do TTT plot indica uma função de risco constante, enquanto que se a curva obtida for côncava (B) ou convexa (C), para um risco monotonamente decrescente ou crescente, respectivamente, nestes casos um candidato ao ajuste desses dados é o modelo Weibull simples. Se primeiramente a função é convexa e depois se torna côncava (D), a função de risco tem forma de “U”, e se a função é inicialmente côncava e depois convexa (E), então a taxa de falha é unimodal.

4.9. Estimação dos Parâmetros dos Modelos

Em cada estudo envolvendo tempos de falha, os parâmetros devem ser estimados a partir das observações amostrais, para que o modelo fique determinado e, assim, seja possível responder às perguntas de interesse.

É possível citar dois importantes métodos de estimação:

- **Método dos Mínimos Quadrados:** inapropriado para estudos de tempos de vida, pois não é capaz de incorporar a informação das observações censuradas no seu processo de estimação.
- **Método de Máxima Verossimilhança Clássico:** escolhe o parâmetro que melhor explique a amostra observada.

4.9.1. Método de Máxima Verossimilhança Clássico

Usualmente a função de máxima verossimilhança é dada pela seguinte equação:

$$L(\theta) = \prod_{i=1}^n f(t_i, \theta).$$

No entanto, na equação acima o tempo de censura é desconsiderado. Portanto é necessário utilizar uma variável indicadora.

$$\delta_i = \begin{cases} 1, & \text{se } t_i \text{ é tempo de falha} \\ 0, & \text{se } t_i \text{ é tempo de censura} \end{cases}$$

A contribuição de cada elemento da amostra para a função de verossimilhança é dada por:

$$\begin{cases} f(t_i, \theta), & \text{se } t_i \text{ é tempo de falha} \\ S(t_i, \theta), & \text{se } t_i \text{ é tempo de censura} \end{cases}$$

Considerando todos os mecanismos de censura, sob a suposição de que eles são não informativos (não carregam informação sobre os parâmetros), a função de verossimilhança, a menos de uma constante, é representada por:

$$L(\theta) = \prod_{i=1}^r [f(t, \theta)] \prod_{i=1}^n [S(t_i, \theta)],$$

ou, equivalentemente, por:

$$L(\theta) = \prod_{i=1}^n [f(t, \theta)]^{\delta_i} [S(t_i, \theta)]^{1-\delta_i} = \prod_{i=1}^n [\lambda(t, \theta)]^{\delta_i} [S(t_i, \theta)].$$

Os estimadores de máxima verossimilhança são os valores de θ que maximizam $L(\theta)$.

4.10. Intervalo de Confiança

Para a construção do intervalo de confiança é utilizada a distribuição assintótica do estimador de máxima verossimilhança de $\hat{\theta}$. Para grandes amostras, essa propriedade estabelece, sob certas condições de regularidade, que $\hat{\theta}$ tem distribuição Normal multivariada de média θ e matriz de variância-covariância $V(\hat{\theta})$, isto é,

$$\hat{\theta} \sim N_k(\theta, Var(\hat{\theta})),$$

sendo k a dimensão de $\hat{\theta}$.

Sob certas condições de regularidade, a matriz de variância-covariância é aproximadamente o negativo da inversa da informação de Fisher:

$$Var(\hat{\theta}) \approx - [I_F(\theta)]^{-1},$$

em que $I_F(\theta) = E \left[\left(\frac{\partial \log L(\theta)}{\partial \theta} \right)^2 \right]$.

4.11. Teste de Hipóteses

Para testar hipóteses relacionadas ao vetor $\theta = (\theta_1, \dots, \theta_p)^T$ de p parâmetros, sendo $H_0: \theta = \theta_0$, três testes são em geral utilizados: Teste de Wald, Teste da Razão de Verossimilhança e Teste Escore.

4.11.1. Teste de Wald

Baseado na distribuição assintótica de $\hat{\theta}$, o Teste de Wald é uma generalização do Teste t de Student e é geralmente o mais utilizado para testar hipóteses relativas a um único parâmetro θ_j . Considerando a hipótese nula, a estatística de teste é dada por

$$W = \frac{(\hat{\theta} - \theta_0)^2}{\widehat{Var}(\hat{\theta})}.$$

Sob H_0 , W tem aproximadamente distribuição Qui-quadrado com p graus de liberdade. A um nível de 100% de significância, valores de W superiores ao valor tabelado da distribuição $X_{p,1-\alpha}^2$ indicam a rejeição de H_0 .

4.11.2. Teste da Razão de Verossimilhança

Baseado na função de verossimilhança, esse teste envolve a comparação dos valores do logaritmo da função de verossimilhança maximizada sem restrição e sob H_0 . A estatística é dada por:

$$TRV = -2 \log \left[\frac{L(\hat{\theta}_0)}{L(\hat{\theta})} \right] = 2 [\log L(\hat{\theta}) - \log L(\hat{\theta}_0)].$$

Sob H_0 , TRV tem aproximadamente distribuição Qui-quadrado com p graus de liberdade. A um nível de 100% de significância, valores do TRV superiores ao valor tabelado da distribuição $X_{p,1-\alpha}^2$ indicam a rejeição de H_0 .

4.11.3. Teste Escore

Obtido a partir da função escore $U(\theta) = \frac{\partial \log L(\theta)}{\partial \theta}$ e da matriz de variância-covariância de $U(\theta)$, esse teste tem a estatística dada por:

$$S = U'(\theta_0)[-F(\theta_0)]^{-1}U(\theta_0).$$

Portanto H_0 é rejeitada, a um nível de 100% de significância, se os valores de S forem superiores ao valor tabelado da distribuição $X_{p,1-\alpha}^2$.

4.12. Escolha do Modelo Probabilístico

A escolha do modelo é extremamente importante, pois o método de máxima verossimilhança somente pode ser aplicado após a definição de um modelo probabilístico adequado para os dados.

A proposta empírica consiste em ajustar os modelos probabilísticos conhecidos e, com base na comparação entre valores estimados e observados, decidir qual deles “melhor” explica os dados amostrais. A forma mais simples e eficiente de selecionar um modelo é por meio de técnicas gráficas. Entretanto teste de hipóteses com modelos encaixados (Cox e Hinkley, 1974) também podem ser utilizados.

4.12.1. Método Gráfico

O primeiro método gráfico consiste na comparação da função de sobrevivência do modelo proposto com a função de sobrevivência estimada pelo método de Kaplan-Meier. Para realizar este procedimento é necessário ajustar os modelos propostos ao conjunto de dados para obter as estimativas dos parâmetros e então estimar a função de sobrevivência dos respectivos modelos. O modelo adequado é aquele em que sua curva de sobrevivência mais se aproximar da curva de sobrevivência estimada pelo método de Kaplan-Meier.

Para uma melhor visualização durante o processo de decisão, a curva de sobrevivência de Kaplan-Meier e a curva de sobrevivência estimada pelo modelo podem ser plotadas no mesmo gráfico como ilustra o exemplo da figura 5.

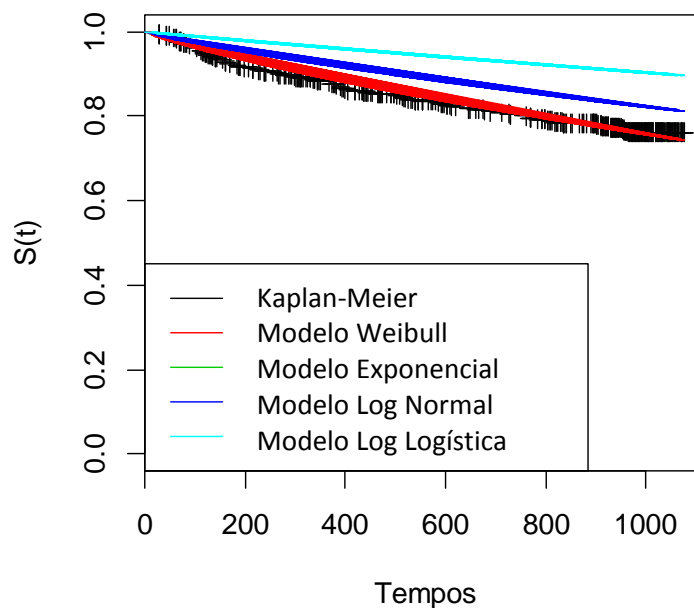


Figura 5 – Curvas de sobrevivência estimada por Kaplan-Meier e curvas de sobrevivência estimada dos modelos: Weibull, Exponencial, Log Normal e Log Logístico

O segundo método gráfico consiste na linearização da função de sobrevivência tendo como ideia básica a construção de gráficos que sejam aproximadamente lineares, caso o modelo proposto seja apropriado. Violações da linearidade podem ser rapidamente verificadas visualmente.

4.12.2. Comparação de modelos

Como as conclusões a partir de técnicas gráficas podem diferir para diferentes analistas, ou seja, existe um componente subjetivo na sua interpretação, os testes de hipóteses são muito utilizados. Desde que o modelo completo seja adequado, a hipótese nula a ser testada é se o modelo de interesse deve ser utilizado.

O teste é realizado a partir dos seguintes ajustes: (1) modelo generalizado (modelo completo) e obtenção do valor do logaritmo de sua função de verossimilhança $\log L(\hat{\theta}_G)$ e (2) modelo de interesse (caso particular do modelo completo) e obtenção do valor do logaritmo da sua função de verossimilhança $\log L(\hat{\theta}_M)$. A partir desses valores é possível calcular a estatística da razão de verossimilhança:

$$TRV = -2 \log \left[\frac{L(\hat{\theta}_M)}{L(\hat{\theta}_G)} \right] = 2 [\log L(\hat{\theta}_G) - \log L(\hat{\theta}_M)],$$

que sob H_0 , tem aproximadamente uma distribuição Qui-quadrado com número de graus de liberdade igual a diferença do número de parâmetros ($\hat{\theta}_G$ e $\hat{\theta}_M$) dos modelos comparados.

4.13. Critérios de seleção de modelos

Akaike's Information Criterion (AIC): Critério de informação que considera o número de parâmetros do modelo e a logverossimilhança. Quanto menor, melhor o ajuste do modelo.

Bayesian Information Criterion (BIC): Critério de informação que considera o número de parâmetros do modelo, a logverossimilhança e o número de observações. Da mesma forma, quanto menor melhor o ajuste do modelo.

4.14. Modelos de Regressão Paramétricos

Na prática a maioria dos estudos envolve a presença de covariáveis que podem estar relacionadas com o tempo de sobrevivência. A covariável pode gerar estratos, se ela for categórica isto é feito automaticamente; caso contrário, a covariável pode ser categorizada quando for necessário.

O objetivo da análise de sobrevivência está centrado na relação entre os tempos de sobrevivência e as variáveis explicativas de interesse. Dessa forma, modelos de regressão podem ser formulados segundo diversos caminhos.

4.14.1. Reparametrização da distribuição de Probabilidade

Consiste em considerar uma reparametrização da distribuição de probabilidade dos tempos para se obter um modelo de regressão. Seja um vetor $\mathbf{x} = (x_0, \dots, x_p)^T$ formado por observações das $(p+1)$ variáveis regressoras. Deve-se escolher a reparametrização que dê a melhor interpretação prática. As seguintes reparametrizações garantem que o parâmetro seja positivo, por isso são bastante utilizadas.

$$\alpha = \exp(x^T \beta) \text{ ou } \alpha = \exp(-x^T \beta)$$

4.14.2. Modelos de Regressão de localização e escala

Seja T uma variável aleatória representando o tempo até a ocorrência do evento de interesse e seja $\mathbf{x} = (x_0, \dots, x_p)^T$ um vetor formado por observações das $(p+1)$ variáveis regressoras. A classe ou família de Modelos de localização e escala se caracteriza pelo fato de $Y = \log(T)$ ter uma distribuição de parâmetros de localização $\mu(\mathbf{x})$ dependendo das variáveis regressoras e um parâmetro de escala σ . O modelo pode ser escrito na forma

$$Y = \mu(\mathbf{x}) + \sigma w,$$

em que $\sigma > 0$ e w tem distribuição conhecida que é independente de \mathbf{x} .

Uma função densidade de probabilidade para o modelo de localização e escala tem a seguinte forma:

$$f(y|x) = \frac{1}{\sigma} g\left(\frac{y - \mu(x)}{\sigma}\right), -\infty < y < \infty$$

em que $-\infty < \mu(x) < \infty$ é o parâmetro de locação, $\sigma > 0$ é o parâmetro de escala e $g(\cdot)$ é uma função densidade de probabilidade totalmente especificada definida em $(-\infty, \infty)$.

A função de sobrevivência correspondente é $G\left(\frac{y - \mu(x)}{\sigma}\right)$, em que $G(\cdot)$ é função de sobrevivência de W . Uma propriedade importante deste modelo é que as covariáveis atuam linearmente em y e multiplicativamente em T .

Pode-se assumir várias formas para $\mu(x)$ dependendo do conjunto de dados. Em geral, a maneira mais usada é $\mu(x) = x^T \beta$, em que $\beta = (\beta_0, \dots, \beta_p)$ é o vetor de parâmetros desconhecidos associado às covariáveis.

5. Análise dos dados

O banco de dados utilizado para esse estudo possui 97 pacientes. A censura foi observada em 90% dos casos.

As variáveis que foram analisadas são:

Tabela 2 – Variáveis e seus respectivos rótulos

GÊNERO	GÊNERO DO PACIENTE (MASCULINO, FEMININO)
EST. CIVIL	ESTADO CIVIL DO PACIENTE (CASADO, DIVORCIADO, SOLTEIRO, UNIÃO CONCENSUAL, VIUVO)
BAIRRO	BAIRRO ONDE RESIDE O PACIENTE
DIAGNOSTICO	DIAGNÓSTICO NA ADMISSÃO DO PACIENTE NO HOSPITAL
HIPERSENSIBILIDADE	ALERGIAS
ANTECEDENTES FAMILIARES	HISTÓRICO FAMILIAR DE DOENÇAS
USO DE MEDICAMENTOS	USO DE MEDICAMENTO REGULAR (SIM, NÃO)
HIPERURICEMIA	PORTADOR DESSA COMORBIDADE (SIM, NÃO)
DAOP	PORTADOR DESSA COMORBIDADE (SIM, NÃO)
AVE	PORTADOR DESSA COMORBIDADE (SIM, NÃO)
ICC	PORTADOR DESSA COMORBIDADE (SIM, NÃO)
DPOC	PORTADOR DESSA COMORBIDADE (SIM, NÃO)
ETILICO	PORTADOR DESSE HÁBITO (SIM, NÃO, EX)
OBESIDADE	PORTADOR DESSA COMORBIDADE (SIM, NÃO)
TABAGISTA	PORTADOR DESSE HÁBITO (SIM, NÃO, EX)
DIABETICO	PORTADOR DESSA COMORBIDADE (SIM, NÃO)
SEDENTARIO	PORTADOR DESSA COMORBIDADE (SIM, NÃO)
HAS	PORTADOR DESSA COMORBIDADE (SIM, NÃO)
DISLIPIDÊMICO	PORTADOR DESSA COMORBIDADE (SIM, NÃO)
IRC	PORTADOR DESSA COMORBIDADE (SIM, NÃO)
HIPOTIROIDISMO	PORTADOR DESSA COMORBIDADE (SIM, NÃO)
DT PROCED.	DATA DO PROCEDIMENTO
ICP PREVIA	REALIZOU ICP ANTERIORMENTE (SIM, NÃO)
REESTENOSE ICP PREVIA	TEVE REESTENOSE DECORRENTE DE ICP PRÉVIA (SIM, NÃO)
QND DE ARTÉRIAS TRATADAS	QUANTIDADE DE ARTÉRIAS TRATADAS
STENT	TIPO DO STENT (CONVENCIONAL, FARMACOLÓGICO)
MARCA	MARCA DO STENT
CREATININA POS ICP	VALOR DA CREATININA (OBTIDA ATRAVÉS DE EXAME LABORATORIAL) APÓS ICP
OBITO	DATA DO ÓBITO
REINTERVENÇÃO	DATA DA REINTERVENÇÃO
ULTIMO CONTATO	ÚLTIMO CONTATO COM O PACIENTE

5.1. Analisando as variáveis

A seguir serão apresentadas algumas análises para verificar o perfil dos pacientes observados.

Realizou-se uma análise exploratória dos dados considerando principalmente as variáveis referentes aos tempos até a ocorrência dos dois eventos de interesse. Foi calculado o coeficiente de correlação τ de Kendall, resultando em $\tau = 0,00891$, indicando uma ausência de correlação entre os tempos dos eventos de interesse (óbito e reintervenção)

5.1.1 Idade

A variável idade é frequentemente utilizada para uma análise descritiva dos dados. A idade média dos pacientes presentes no estudo é de 63,67 anos. Sendo que a idade média das mulheres é 66 enquanto que para os homens observou-se 62,67. A moda de idade para mulheres foi igual a 57 e para homens 46. Considerando a moda geral, observa-se que a angioplastia ocorreu mais em pacientes com 62 anos.

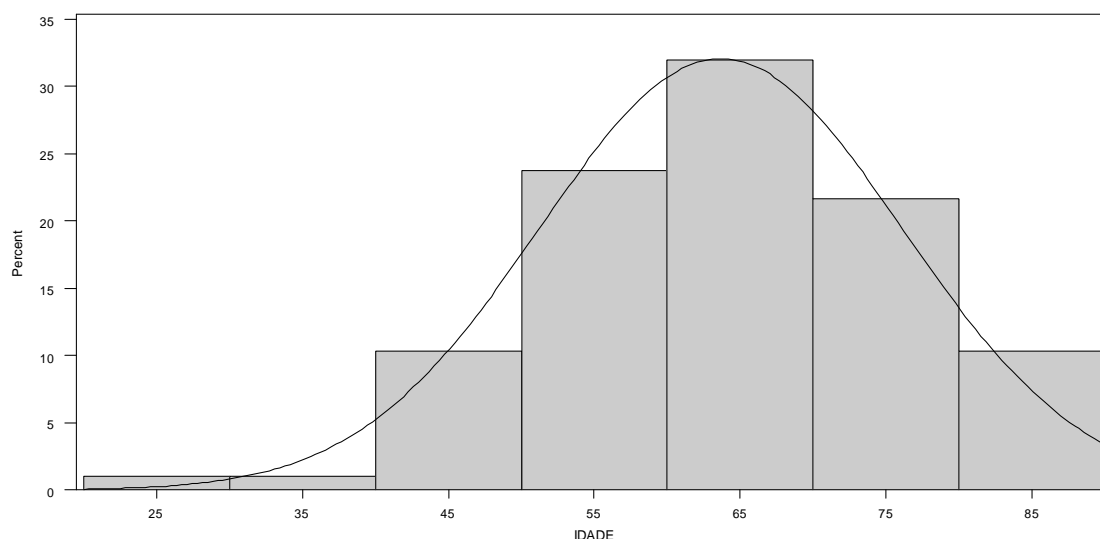


Figura 6 – Histograma da idade

5.1.2 Regressão logística gênero e óbito

Fazendo o teste de verossimilhança, é possível concluir que existe regressão logística entre as variáveis gênero e óbito. Apesar dos dados apresentarem muitas censuras, é possível deduzir que esses pacientes de fato não sofreram nenhuma dos dois eventos de interesse. Isso ocorre devido à “fidelidade” do paciente para com o seu médico. Após sofrer uma angioplastia, o paciente, em geral, demonstra muita gratidão que provavelmente não trocaria de médico. Dessa forma, caso o paciente tenha

necessitado passar por outra angioplastia, ele recorreria ao mesmo médico. Portanto, é possível considerar que as censuras representam muito mais que a perda de acompanhamento do paciente, mas sim a não ocorrência do evento de interesse.

Através do método de regressão logística, considerando que o critério de convergência foi satisfeito, de acordo com o teste de Wald é possível observar um *Odds Ratio (OR)* com valor igual de 0,381. Isso significa que a chance do óbito ocorrer é menor em homens do que em mulheres. Com um $\alpha = 0,05$ observou-se um intervalo de confiança que varia entre 0,101 e 1,431. O que pode contrariar a conclusão apresentada acima. Caso o *OR* seja igual a 1,431, é possível dizer que mulheres morrem 43,1% mais que homens.

Tabela 3 – Tabela de Contingência: Óbito x Gênero

GÊNERO	ÓBITO		Total
	SIM	NÃO	
Feminino	5	24	29
	5,15%	24,74%	29,90%
	17,24%	82,76%	
	50,00%	27,59%	
Masculino	5	63	68
	5,15%	64,95%	70,10%
	7,35%	92,65%	
	50,00%	72,41%	
Total	10 10,31%	87 89,69%	97 100,00%

Tabela 4 – Avaliar significância do parâmetro: Beta=0: Óbito x Gênero

Teste	Qui-quadrado	GL	P-valor
<i>Likelihood Ratio</i>	1,9883	1	0,1585
<i>Score</i>	2,1499	1	0,1426
<i>Wald</i>	2,0357	1	0,1536

Tabela 5 – Análises de estimação da Máxima verossimilhança: Óbito x Gênero

Parâmetro	DF	Valor estimado	Erro Padrão	Qui-quadrado de Wald	P-valor
<i>Intercept</i>	1	-0,6035	1,0874	0,3080	0,5789
GÊNERO	1	-0,9651	0,6764	2,0357	0,1536

Tabela 6 – Estimação do Odds Ratio: Óbito x Gênero

Efeito	Valor Estimado	95% Intervalo de Confiança de Wald	
GÊNERO	0,381	0,101	1,434

Tabela 7 – Associação entre as probabilidades previstas e as respostas observadas: Óbito x Gênero

<i>Percentual Concordante</i>	36,2	<i>D de Somers</i>	0,224
<i>Percentual Discordante</i>	13,8	<i>Gamma</i>	0,448
<i>Percentual Tied</i>	50,0	<i>Tau-a</i>	0,042
<i>Pares</i>	870	<i>c</i>	0,612

5.1.3 Regressão logística gênero e reintervenção

Fazendo o teste de verossimilhança, é possível concluir que existe regressão logística entre as duas variáveis.

Através do método de regressão logística, considerando que o critério de convergência foi satisfeito, de acordo com o teste de Wald é possível observar um *Odds Ratio (OR)* com valor igual a 1,549. Isso significa que a reintervenção ocorre 54,9% mais em homens do que em mulheres. Com um $\alpha = 0,05$ observou-se um intervalo de confiança que varia entre 0,302 e 7,95. O que pode contrariar a conclusão apresentada acima. Caso o *OR* seja igual a 0,302, é possível dizer que a chance de ocorrer reintervenção em mulheres é menor que em homens.

Tabela 8 – Tabela de Contingência: Reintervenção x Gênero

GÊNERO	REINTERVENÇÃO		
	<i>SIM</i>	<i>NÃO</i>	<i>Total</i>
<i>Feminino</i>	2	27	29
	2,06%	27,84%	29,90%
	6,90%	93,10%	
	22,22%	30,68%	
<i>Masculino</i>	7	61	68
	7,22%	62,89%	70,10%
	10,29%	89,71%	
	77,78%	69,32%	
Total	9	88	97
	9,28%	90,72%	100,00%

Tabela 9 – Avaliar significância do parâmetro: Beta=0 1: Reintervenção x Gênero

Teste	Qui-quadrado	GL	P-valor
<i>Likelihood Ratio</i>	0,2935	1	0,5880
<i>Score</i>	0,2788	1	0,5975
<i>Wald</i>	0,2752	1	0,5999

Tabela 10 – Análises de estimação da Máxima verossimilhança: Reintervenção x Gênero

Parâmetro	DF	Valor estimado	Erro Padrão	Qui-quadrado de Wald	P-valor
<i>Intercepto</i>	1	-3,0404	1,5190	4,0063	0,0453
GÊNERO	1	0,4377	0,8344	0,2752	0,5999

Tabela 11 – Estimação do Odds Ratio: Reintervenção x Gênero

Efeito	Valor Estimado	95% Intervalo de Confiança de Wald
GÊNERO	1,549	0,302 7,950

Tabela 12 – Associação entre as probabilidades previstas e as respostas observadas: Reintervenção x Gênero

<i>Percentual Concordante</i>	23,9	<i>D de Somers</i>	0,085
<i>Percentual Discordante</i>	15,4	<i>Gamma</i>	0,215
<i>Percentual Tied</i>	60,7	<i>Tau-a</i>	0,014
<i>Pares</i>	792	<i>C</i>	0,542

5.1.4 Comorbidades

A presença de comorbidade, presença de doença(s), além da doença cardíaca, que pode(m) alterar o efeito de interesse no estudo, é um fator de muita relevância.

Em um estudo semelhante realizado em Porto Alegre, variáveis como gênero, obesidade, hipertensão arterial sistêmica, dislipidemia, tabagismo, história familiar, sedentarismo e diabetes não apresentaram significância estatística quando analisadas isoladamente em relação à mortalidade. No entanto, é importante salientar que estes fatores de risco já são conhecidos há décadas e vêm sempre se reafirmando,

demonstrando a importância em manter o foco na prevenção primária destas comorbidades. (TREVISO et al, 2012)

As principais comorbidade observadas foram:

Insuficiência Cardíaca Congestiva (ICC): sobrecarga do coração na tentativa de compensar a falta de oxigênio tecidual devido à condição fisiopatológica na qual o coração encontra-se incapacitado de bombear sangue a uma taxa ideal, com consequente redução da oxigenação dos tecidos, causando falhas em diversos processos, como deficiente remoção de água, sal e impurezas da corrente sanguínea.

Na tabela 13 é possível perceber que não foram observadas mulheres com Insuficiência Cardíaca Congestiva, além de uma ocorrência muito pequena em homens.

Tabela 13 – Frequência dos pacientes portadores de ICC por gênero

GÊNERO	ICC		
	<i>SIM</i>	<i>NÃO</i>	<i>Total</i>
<i>Feminino</i>	0	29	29
	0,00%	29,90%	29,90%
	0,00%	100,00%	
	0,00%	31,52%	
<i>Masculino</i>	5	63	68
	5,15%	64,95%	70,10%
	7,35%	92,65%	
	100,00%	68,48%	
Total	5 5,15%	92 94,85%	97 100,00%

Etílico: paciente que possui/possuía o hábito de ingerir bebidas alcoólicas sem moderação.

Na tabela 14 foram observados apenas mulheres ex etílicas e homens etílicos.

Tabela 14 – Frequência dos pacientes etílico por gênero

GÊNERO(GÊNERO)	ETILICO(ETILICO)			Total
	<i>SIM</i>	<i>EX</i>	<i>NÃO</i>	
<i>Feminino</i>	0	2	27	29
	0,00%	2,06%	27,84%	29,90%
	0,00%	6,90%	93,10%	
	0,00%	100,00%	31,76%	
<i>Masculino</i>	10	0	58	68
	10,31%	0,00%	59,79%	70,10%
	14,71%	0,00%	85,29%	
	100,00%	0,00%	68,24%	
Total	10	2	85	97
	10,31%	2,06%	87,63%	100,00%

Obesidade: excessivo acúmulo de gordura corporal. O diagnóstico é feito através do cálculo de índice de massa corporal (IMC) que consiste em dividir o peso do indivíduo (em quilogramas) pelo quadrado de sua altura (em metros). IMC menor a 18,5 corresponde a pessoas com peso abaixo do normal, entre 18,5 e 24,9 é tido como peso normal, entre 25 e 29,9 representa pessoas com peso acima do normal, entre 30 e 30,9 a pessoa está obesa e quando o IMC é maior do que 40 considera-se a pessoa portadora de obesidade mórbida. (PACIEVITCH, 2013) No caso, esse “diagnóstico” foi única e exclusivamente visual. O caso de ex obeso, foi informado pelo paciente.

A tabela 15 demonstra que apesar da baixa frequência de obesos na amostra, proporcionalmente existem mais mulheres (17,24%) do que homens (11,76%).

Tabela 15 – Frequência dos pacientes obesos por gênero

GÊNERO	OBESIDADE			Total
	<i>SIM</i>	<i>EX</i>	<i>NÃO</i>	
<i>Feminino</i>	5	0	24	29
	5,15%	0,00%	24,74%	29,90%
	17,24%	0,00%	82,76%	
	38,46%	0,00%	28,92%	
<i>Masculino</i>	8	1	59	68
	8,25%	1,03%	60,82%	70,10%
	11,76%	1,47%	86,76%	
	61,54%	100,00%	71,08%	
Total	13	1	83	97
	13,40%	1,03%	85,57%	100,00%

Tabagismo: reconhecidamente, uma doença crônica, resultante da dependência à droga nicotina, é um fator de risco para cerca de 50 doenças. (COORDENAÇÃO NACIONAL DE CONTROLE DO TABAGISMO (INCA/SAS/MS), 2013) O tabagismo é cumulativo e aumenta sua gravidade conforme o número de anos em que a pessoa fumou. Entretanto, foi realizado um estudo de coorte na Holanda, com 21.662 pessoas, que concluiu que ex-fumantes que pararam de fumar por cinco anos apresentavam o mesmo índice de risco cardiovascular que as pessoas que nunca fumaram. (KASTRATI, SCHÖMIG, ELEZI et al.,1997)

A tabela 16 destaca a prevalência de homens fumantes (30,88%) na amostra com relação as mulheres (17,24%).

Tabela 16 – Frequência dos pacientes tabagistas por gênero

GÊNERO	TABAGISTA			Total
	SIM	EX	NÃO	
<i>Feminino</i>	5	5	19	29
	5,15%	5,15%	19,59%	29,90%
	17,24%	17,24%	65,52%	
	19,23%	27,78%	35,85%	
<i>Masculino</i>	21	13	34	68
	21,65%	13,40%	35,05%	70,10%
	30,88%	19,12%	50,00%	
	80,77%	72,22%	64,15%	
Total	26	18	53	97
	26,80%	18,56%	54,64%	100,00%

Diabético: a elevação da Glicose no sangue: hiperglicemia. Há três tipos de diabetes: diabetes tipo 1 - a produção de insulina do pâncreas é insuficiente necessitando de injeções diárias de insulina para manterem a glicose no sangue em valores normais -, diabetes tipo 2 - resistência insulínica corresponde a 90% dos casos de diabetes - e diabetes gestacional - glicose elevada no sangue durante a gravidez, geralmente normalizada após o parto.

A tabela 17 destaca a prevalência de homens diabéticos (25,77%) na amostra com relação as mulheres (6,19%).

Tabela 17 – Frequência dos pacientes diabéticos por gênero

GÊNERO	DIABÉTICO		
	<i>SIM</i>	<i>NÃO</i>	<i>Total</i>
<i>Feminino</i>	6	23	29
	6,19%	23,71%	29,90%
	20,69%	79,31%	
<i>Masculino</i>	19,35%	34,85%	
	25	43	68
	25,77%	44,33%	70,10%
<i>Masculino</i>	36,76%	63,24%	
	80,65%	65,15%	
	31	66	97
Total	31,96%	68,04%	100,00%

Sedentário: falta de atividade física suficiente. O sedentarismo acontece quando a pessoa gasta poucas calorias diárias com atividades físicas. Pesquisas mostram que, o risco de se ter uma doença cardiovascular aumenta uma vez e meia nas pessoas que não têm uma vida ativa. (SABRINA, 2009)

A tabela 18 apresenta a frequência de homens e mulheres sedentários. A coleta desse dado não envolveu nenhum diagnóstico, mas apenas o relato do paciente, o que não torna o dado tão verídico. Apesar disso, foram observados 6 mulheres e 16 homens sedentários.

Tabela 18 – Frequência dos pacientes sedentários por gênero

GÊNERO	SEDENTARIO		
	<i>SIM</i>	<i>NÃO</i>	<i>Total</i>
<i>Feminino</i>	6	23	29
	6,19%	23,71%	29,90%
	20,69%	79,31%	
<i>Masculino</i>	27,27%	30,67%	
	16	52	68
	16,49%	53,61%	70,10%
<i>Masculino</i>	23,53%	76,47%	
	72,73%	69,33%	
	22	75	97
Total	22,68%	77,32%	100,00%

Hipertensão Arterial Sistêmica (HAS): doença grave caracterizada pela elevação da pressão arterial. Não apresenta sintomatologia alguma e mesmo quando há algum sintoma, estes são comuns também a outras doenças.

A tabela 19 apresenta que a maioria dos pacientes (78,35%) da amostra possuem HAS. Essa é uma das comorbidades mais relevantes.

Tabela 19 – Frequência dos pacientes portadores de HAS por gênero

GÊNERO	HAS		Total
	SIM	NÃO	
<i>Feminino</i>	23	6	29
	23,71%	6,19%	29,90%
	79,31%	20,69%	
<i>Masculino</i>	30,26%	28,57%	
	53	15	68
	54,64%	15,46%	70,10%
	77,94%	22,06%	
	69,74%	71,43%	
Total	76	21	97
	78,35%	21,65%	100,00%

Dislipidêmico: altos níveis de gorduras circulando no sangue. Essas gorduras incluem colesterol e triglicérides. A dislipidemia é causada quando se ingere uma dieta rica em colesterol e gorduras, quando o organismo produz colesterol e triglicérides demais ou ambas as situações.

Aproximadamente a metade das mulheres são dislipidêmicas, enquanto 61,76% dos homens possuem altos níveis de gorduras.

Tabela 20 – Frequência dos pacientes dislipidêmicos por gênero

GÊNERO	DISLIPIDÊMICO		
	<i>SIM</i>	<i>NÃO</i>	<i>Total</i>
<i>Feminino</i>	14	15	29
	14,43%	15,46%	29,90%
	48,28%	51,72%	
	25,00%	36,59%	
<i>Masculino</i>	42	26	68
	43,30%	26,80%	70,10%
	61,76%	38,24%	
	75,00%	63,41%	
Total	56	41	97
	57,73%	42,27%	100,00%

Insuficiência Renal Crônica (IRC): lesão renal e geralmente perda progressiva e irreversível da função dos rins. A principal função dos rins é remover os resíduos e o excesso de água do organismo.

Essa comorbidade não apresenta muita relevância na ocorrência de angioplastia, inclusive não foram apresentados muitos casos nessa amostra conforme apresentado na tabela 21. Essa comorbidade foi considerada para uma possível análise junto com as taxas de creatinina.

Tabela 21 – Frequência dos pacientes portadores de IRC por gênero

GÊNERO	IRC		
	<i>SIM</i>	<i>NÃO</i>	<i>Total</i>
<i>Feminino</i>	1	28	29
	1,03%	28,87%	29,90%
	3,45%	96,55%	
	14,29%	31,11%	
<i>Masculino</i>	6	62	68
	6,19%	63,92%	70,10%
	8,82%	91,18%	
	85,71%	68,89%	
Total	7	90	97
	7,22%	92,78%	100,00%

Hipotireoidismo: caracterizado pela diminuição ou pela baixa produção dos hormônios T3 (triiodotironina) e T4 (levotiroxina). As causas mais comuns do hipotireoidismo são a inflamação crônica da tireoide (chamada tireoidite ou doença de Hashimoto), as

manifestações pós-cirúrgicas (retirada parcial ou total da glândula) e as de decorrência de tratamentos prévios de glândula hiperativa. (LABORATÓRIO OSWALDO CRUZ, 2003)

A comorbidade hipotireoidismo não apareceu com muita frequência nos pacientes da amostra, como demonstra a tabela abaixo.

Tabela 22 – Frequência dos pacientes portadores de hipotireoidismo por gênero

GÊNERO	HIPOTIREOIDISMO		
	<i>SIM</i>	<i>NÃO</i>	<i>Total</i>
<i>Feminino</i>	2	27	29
	2,06%	27,84%	29,90%
	6,90%	93,10%	
	40,00%	29,35%	
<i>Masculino</i>	3	65	68
	3,09%	67,01%	70,10%
	4,41%	95,59%	
	60,00%	70,65%	
Total	5	92	97
	5,15%	94,85%	100,00%

Estudo holandês comparando stents convencionais e farmacológicos em pacientes pós-angioplastia por IAM demonstrou que a sobrevida livre de qualquer evento com uso de stents farmacológicos foi significativamente maior ($p < 0,02$), mas não relacionou o tipo de stent à mortalidade. (TREVISO et al, 2012)

No estudo em questão, de acordo com a tabela abaixo, 4,1% realizaram angioplastia sem implante de stent (balão) e 1% implantou dois stents, sendo um convencional e o outro farmacológico.

Tabela 23 – Frequência do tipo de stent implantado

	Frequência	Porcentual
BALÃO	4	4,1%
STENT CONVENCIONAL	41	41,8%
STENT FARMACOLOGICO	51	52,0%
STENT FARMACOLOGICO E CONVENCIONAL	1	1,0%
Total	98	100%

5.2. Variável Resposta: Óbito

A primeira etapa de qualquer análise estatística de dados consiste de uma análise descritiva das variáveis em estudo. A análise de sobrevivência muitas vezes leva em

consideração a presença de covariáveis ou variáveis explicativas que podem influenciar o tempo de sobrevivência ou de censura. É importante detectar se as covariáveis estão relacionadas ao tempo de sobrevivência ou de censura, ou mesmo se estão relacionadas entre si. Inicialmente serão apresentados os gráficos box-plot para as variáveis gênero e quantidade de artérias tratadas.

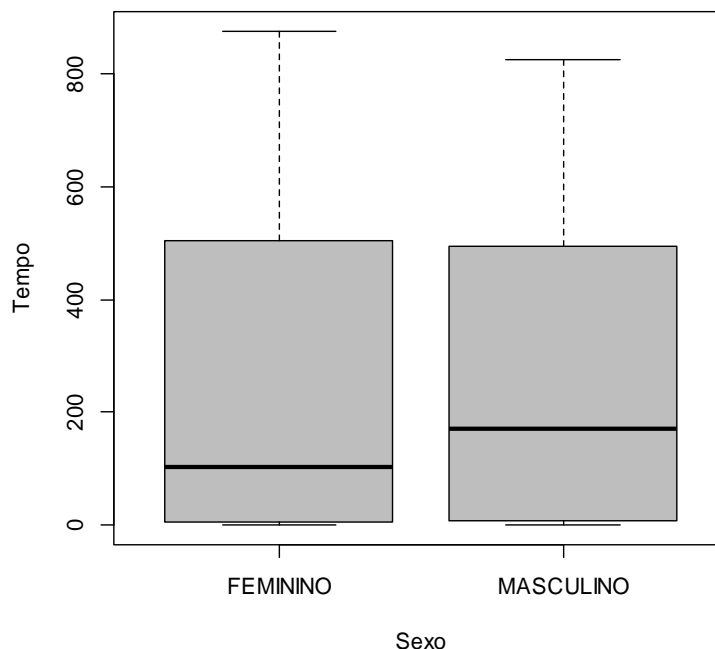


Figura 7 – Box-plot comparativo para a variável gênero com relação aos tempos de falha óbito

Observando o gráfico acima, é possível deduzir que a variável gênero não é significativa com relação a variável resposta óbito. É importante ressaltar, que foram considerados os tempos de falha e censura, e por isso, uma análise mais adequada para dados com censuras foi utilizada na sequência.

O gráfico abaixo apresenta o mesmo caso com relação às censuras. Além disso, também é possível deduzir que a variável quantidade de artérias tratadas não é significativa. No entanto, clinicamente falando, essa variável é importante e por isso é uma forte candidata a compor um modelo de regressão que descreve a função de sobrevivência.

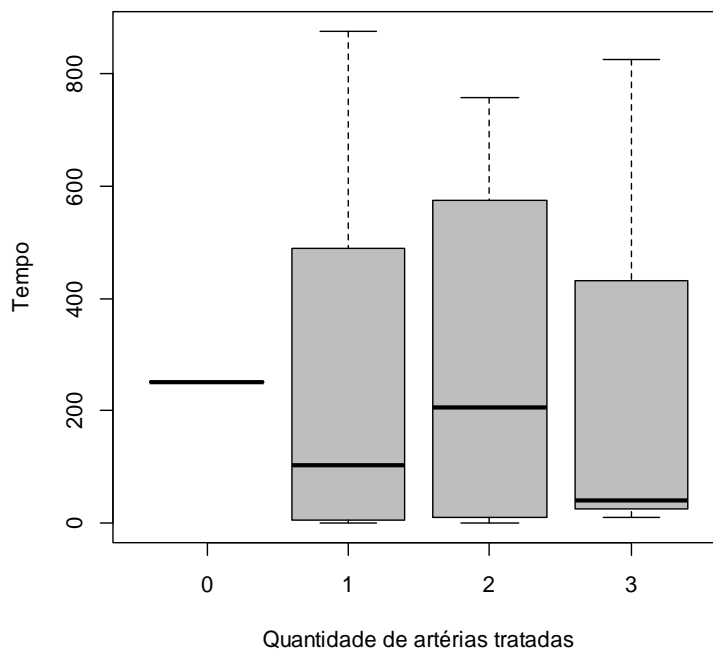


Figura 8 – Box-plot comparativo para a variável quantidade de artérias tratadas com relação aos tempos de falha óbito

Abaixo serão apresentados gráficos com a função de sobrevivência estimada por Kaplan-Meier. Através desses gráficos é possível observar um grande número de censuras. Como a curva de sobrevivência estimada não tende a zero, mais a diante, para modelar os dados foi utilizado um modelo com fração de cura.

É possível observar que a maior parte dos óbitos que ocorreram no estudo, aconteceram no primeiro dia. Uma sugestão para trabalhos futuros é a construção de um modelo que descreve esse fato.

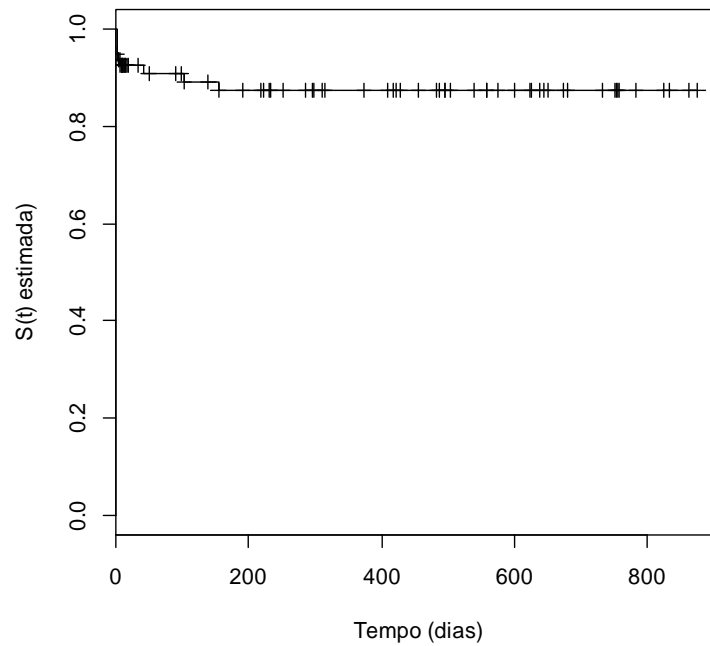


Figura 9 – Curva de sobrevivência Kaplan-Meier considerando o evento de interesse o óbito

O gráfico abaixo é chamado de TTT plot (Tempo total em Teste). Ele representa o comportamento da Função Taxa de Falha empírica. É possível observar que o gráfico apresenta uma curva convexa e depois côncava, o que significa que a função taxa de falha tem forma de U. Para modelar dados dessa natureza, na última década, foram propostas novas distribuições que são baseadas em modificações da distribuição Weibull. Entre elas, destacam-se as distribuições Weibull exponencializada, Weibull Modificada, XTG (chamada anteriormente de Weibull estendida) e Weibull aditiva.

No entanto, como a parte côncava do gráfico é bastante sutil, é possível considerar que a curva é predominantemente convexa. Sendo assim, a função taxa de falha é monotonamente crescente. Para modelar dados dessa natureza é possível utilizar a distribuição Weibull.

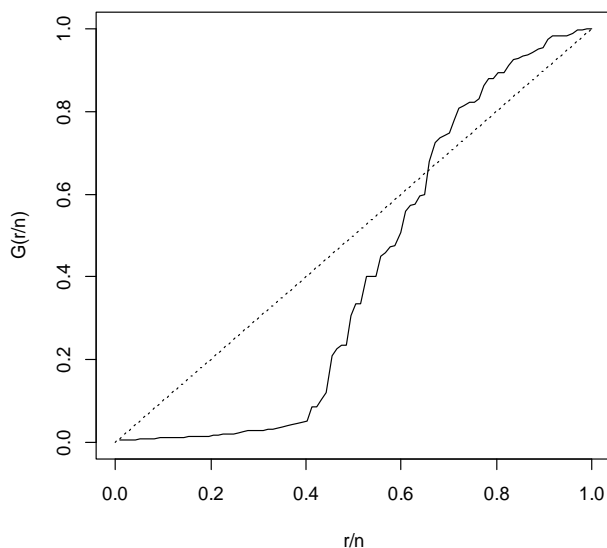


Figura 10 - TTT plot óbito

Observando a função de sobrevivência separada pela quantidade de artérias tratadas no gráfico abaixo, talvez seja possível dizer que os riscos não são proporcionais, pois as funções de sobrevivência plotadas se cruzam no início dos tempos. No entanto, como esse cruzamento só ocorre no começo, o teste de *logrank* para verificar as funções de taxa de falha dos grupos são iguais, pode ser considerado.

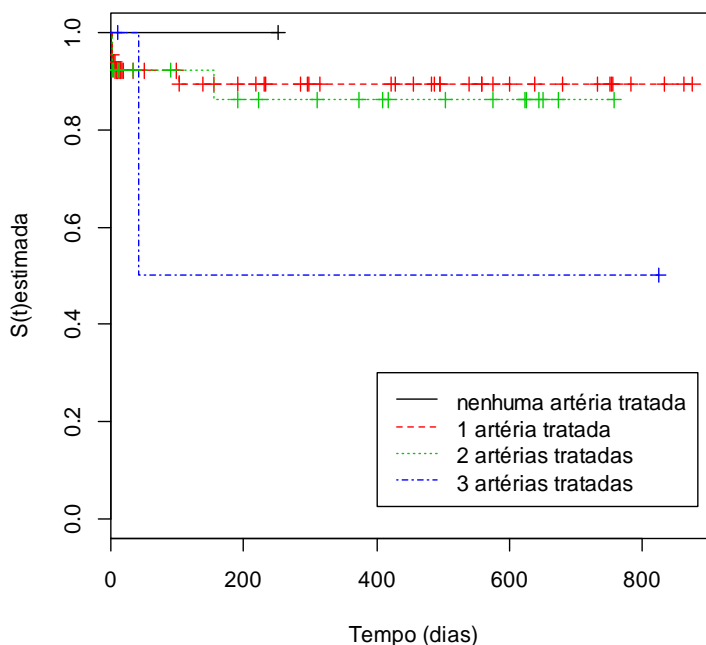


Figura 11 – Estimativas de Kaplan-Meier considerando o evento de interesse o óbito para a variável quantidade de artérias tratadas. Os tempos representados por + mostram onde ocorreram censuras em cada grupo 1

O valor do teste de *logrank* para a comparação entre os grupos resultou em um p-valor = 0,6, indicando que não existe diferença entre as curvas de sobrevivência. Para o teste de *Wilcoxon*, indicado para casos cujo os riscos não são proporcionais, foi obtido um p-valor = 0,63, resultado não muito diferente do teste de *logrank*. Como a variável quantidade de artérias tratadas possui 4 categorias, é interessante testar as categorias duas a duas.

Dos resultados apresentados na tabela abaixo, consequência do teste de *Wilcoxon* (que utiliza peso igual ao número de indivíduos sob risco, coloca mais peso na porção inicial do eixo do tempo) para comparações dos grupos dois a dois, não foram encontradas evidências de diferenças entre nenhum par de grupos.

Tabela 24 - Resultados dos testes de Wilcoxon utilizados para as comparações de grupos, dois a dois, considerando o evento de interesse o óbito

Grupos comparados	Estatística de Teste (T)	p-valor
0 x 1	0,110	0,74
0 x 2	0,145	0,704
0 x 3	0,500	0,48
1 x 2	0,092	0,761
1 x 3	1,470	0,226
2 x 3	1,020	0,312

Observando o gráfico abaixo que mostra as curvas de sobrevivência estimadas por Kaplan-Meier separadas por gênero, é possível afirmar que a razão das funções de taxa de falha dos grupos é aproximadamente constante. Isto é, as populações tem propriedade de taxas de falha proporcionais. Dessa forma, o teste *logrank* é particularmente apropriado para verificar a igualdade entre as funções de sobrevivência.

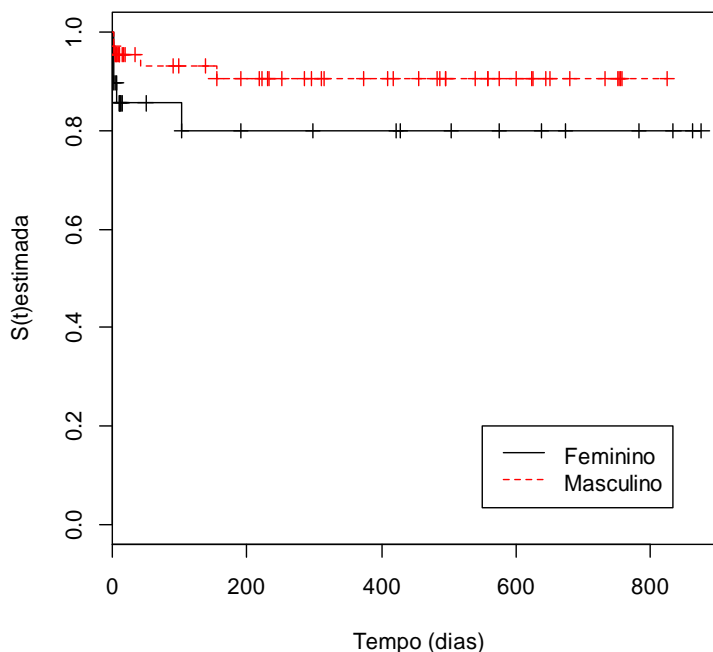


Figura 12 – Estimativas de Kaplan-Meier considerando o evento de interesse o óbito para a variável gênero. Os tempos representados por + mostram onde ocorreram censuras em cada gênero

Os dois testes apresentaram resultados semelhantes, indicando que não existe diferença entre as duas funções de sobrevivência apresentadas no gráfico acima.

Tabela 25 - Testes não-paramétricos para comparação de curvas de sobrevivência obtidas para o gênero feminino e masculino.

Teste	Estatística de Teste (T)	p-valor
Logrank	2,21	0,137
Wilcoxon	2,24	0,134

Tentando ajustar os dados a modelos convencionais, é possível verificar no gráfico abaixo a qualidade do ajuste às distribuições Weibull, Log-normal e Log-logística e que a calda das distribuições comparadas à curva de sobrevivência não se ajustam muito bem.

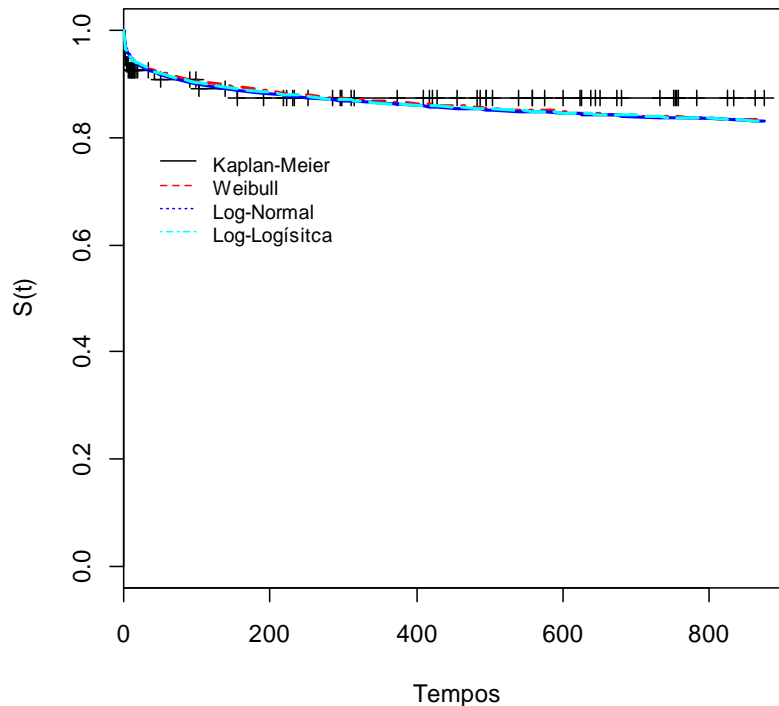


Figura 13 – Curvas de sobrevivência estimada por Kaplan-Meier considerando o evento de interesse o óbito e curvas de sobrevivência estimada dos modelos: Weibull, Log-normal e Log-logística

Além disso, é possível verificar que as distribuições Weibull, Log-normal e Log-Logística apresentaram comportamento muito semelhante, sendo quase impossível visualizar as curvas individualmente. No entanto, o Teste da Razão de verossimilhança não é adequado para esse caso pois os modelos não são encaixados. Ou seja, um modelo não é caso particular do outro. Neste caso o indicado é usar as medidas AIC, AICc e BIC.

Tabela 26 - Comparando distribuições através das medidas AIC, AICc e BIC considerando o evento de interesse o óbito

Modelo	AIC	AICc	BIC
Weibull	143,8019	143,8440	146,3767
Log-Normal	142,1000	142,1421	144,6747
Log-Logística	143.5181	143,5602	146,0928

Segundo Klein e Moeschberger (1997), essas técnicas devem ser usadas como um meio de rejeitar modelos claramente inapropriados e não para provar que um particular modelo está correto. Portanto, apesar da distribuição Log-normal apresentar os menores valores para os critérios apresentados, nesse trabalho foi utilizada a distribuição Weibull devido a sua flexibilidade. As principais vantagens da utilização da distribuição de Weibull para análise de sobrevivência é que através da estimação de apenas dois parâmetros (α e β) são obtidas informações tanto de longevidade média quanto do tipo de curva de sobrevivência. Outra vantagem é que as observações não

necessitam ser realizadas em intervalos constantes, como por exemplo, as tabelas de esperança de vida.

Trinta e três covariáveis potencialmente importantes para descrever o comportamento da resposta foram selecionadas para serem incluídas no modelo. Existem inúmeros possíveis modelos formados pela combinação de todas estas covariáveis. É certamente impraticável ajustar todos estes modelos possíveis a fim de selecionar o que melhor explique a resposta. Isso implica que covariáveis importantes em termos clínicos devem ser incluídas independente de sua significância estatística, assim como a importância clínica deve ser considerada em cada passo de inclusão ou exclusão no processo de seleção de covariáveis.

Para seleção das variáveis, o modelo Weibull foi ajustado contendo uma única variável. Através do Teste da razão de verossimilhança, as variáveis consideradas significativas foram ajustadas conjuntamente. Na presença de certas covariáveis, outras podem deixar de ser significativas. Conseqüentemente, ajustam-se modelos reduzidos, excluindo-se uma covariável de cada vez. Verificam-se as covariáveis que provocam um aumento estatisticamente significativo na estatística da razão de verossimilhança. Somente aquelas que atingirem a significância ao nível de 0,10 permanecem no modelo.

Após verificar a possibilidade de inclusão de termos de interação dupla entre as covariáveis incluídas no modelo, verificou-se que as interações não eram significativas.

Verificou-se que as variáveis gênero e creatinina não eram significativas. No entanto, gênero é uma variável muito visada, e sempre é interessante testá-la. Já a variável creatinina foi incluída no modelo devido a alteração para maior no AIC que sua ausência causava.

O modelo final ficou composto pelas covariáveis e seus respectivos coeficientes: quantidade de artérias tratadas (grupo 1 = 484,6775; grupo 2 = 878,2755; grupo 3 = -291,8132), gênero (masculino = 358,0326), etílico (não = -2350,0169; sim = 793,8603), obesidade (não = 332,5915; sim = 484,0695), tabagista (não = -6556,4933; sim = -6224,8379) e creatinina (-338,0033). Nenhum termo de interação dupla foi significativo.

Os coeficientes estimados estão expressos na escala logarítmica dos tempos, isto é, para $Y = \log(T) = x'\beta + \sigma v$.

É importante ressaltar que para dados como os que estão sendo estudados, cujo a função de sobrevivência não tendem a zero com o passar do tempo, é interessante admitir um modelo de fração de cura, que é representado por

$$S_{pop}(t) = (1 - \phi) + \phi S(t),$$

onde $S(t)$ foi modelado pelo modelo weibull. Dessa forma, utilizando uma metodologia de máxima verossimilhança foram estimados os seguintes parâmetros:

$$\hat{\alpha} = 0,4166181; \hat{\beta} = 99,9741096 \text{ e } \hat{\phi} = 0,1649477.$$

A metodologia de estimação foi implementada no software R, utilizando diversas funções já existentes no mesmo, assim como novas funções com os cálculos das derivadas parciais da função de verossimilhança foram incluídas na programação

utilizada. No processo de otimização, os valores iniciais para os parâmetros da distribuição weibull foram obtidos supondo distribuição do Valor extremo.

Em estudos futuros, será possível utilizar um modelo com fração de cura considerando as covariáveis. O gráfico abaixo mostra como o modelo weibull com fração de cura (sem covariáveis) modela bem melhor os dados, ajustando de maneira mais suave à cauda.

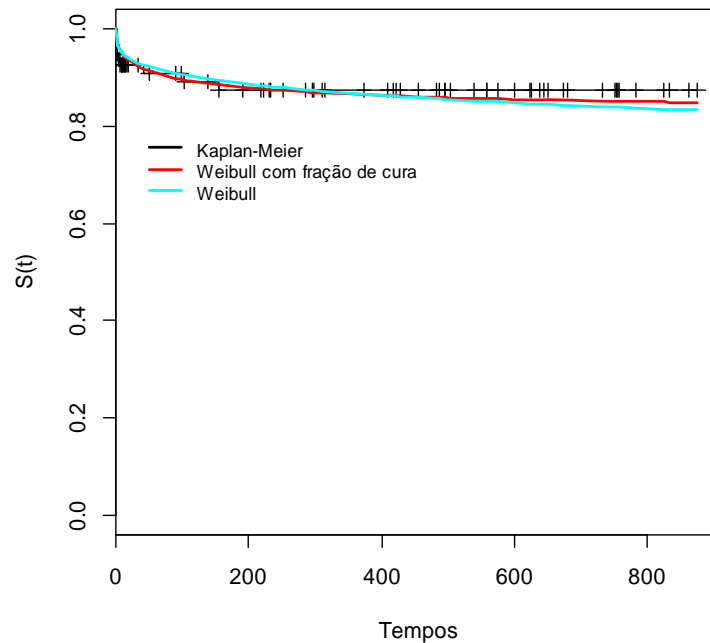


Figura 14 – Curvas de sobrevivência estimada por Kaplan-Meier considerando o evento de interesse o óbito e curvas de sobrevivência estimada dos modelos: Weibull e Weibull com fração de cura

O próximo passo a ser executado é a análise de resíduos.

Em trabalhos futuros serão considerados um modelo weibull com fração de cura e presença de covariáveis para a modelagem dos dados. Além disso, outras modificações da Weibull podem ser utilizadas para descrever os dados.

5.3. Variável Resposta: Reintervenção

Na análise da reintervenção, o óbito foi considerado como censura apenas quando o paciente falecia antes de uma segunda intervenção ocorrer. Na análise do óbito, a reintervenção não foi considerada, pois apenas dois pacientes morreram após uma segunda intervenção.

Como a análise descritiva é a primeira etapa da análise de dados, serão apresentados alguns gráficos para observar o comportamento de algumas covariáveis com relação ao tempo de intervenção.

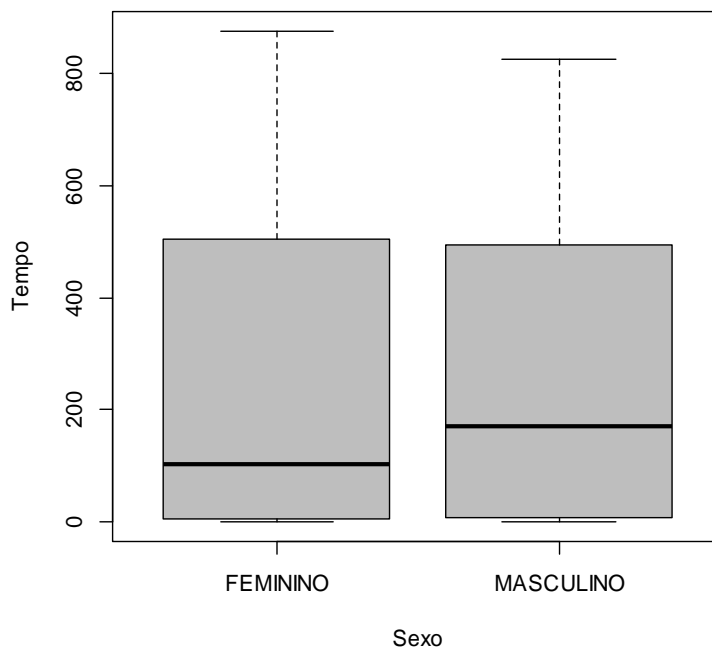


Figura 15 – Box-plot comparativo para variável gênero com relação aos tempos de falha da reintervenção

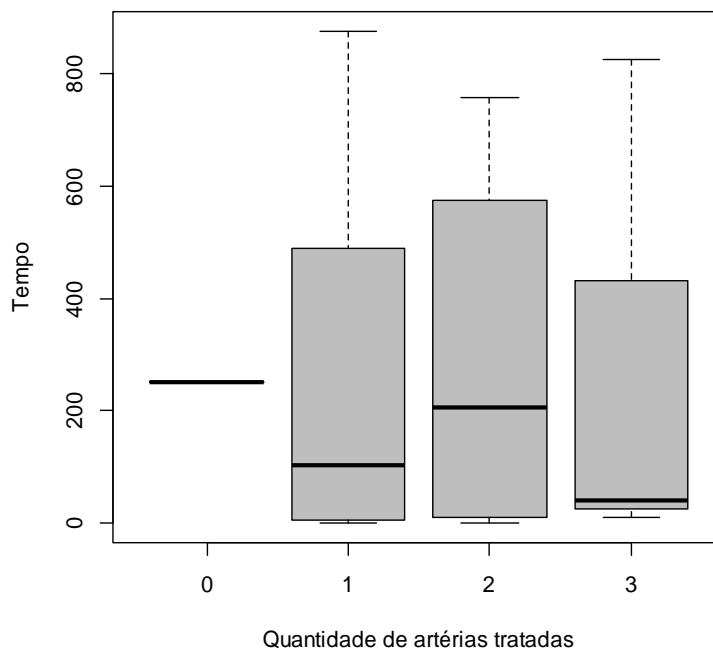


Figura 16 – Box-plot comparativo para a variável quantidade de artérias tratadas com relação aos tempos de falha reintervenção

Quando comparados com os mesmos gráficos para a variável resposta óbito, é possível perceber que eles são muito parecidos. Isso ocorre devido ao óbito antes da reintervenção ser considerado como censura e a amostra ser pequena. Portanto, para ambas as variáveis é possível deduzir uma não significância.

Abaixo serão apresentados gráficos com a função de sobrevivência estimada por Kaplan Meier. Através desses gráficos é possível observar um grande número de censuras. Como a curva de sobrevivência estimada não tende a zero, mais a diante, para modelar os dados foi utilizado um modelo com fração de cura.

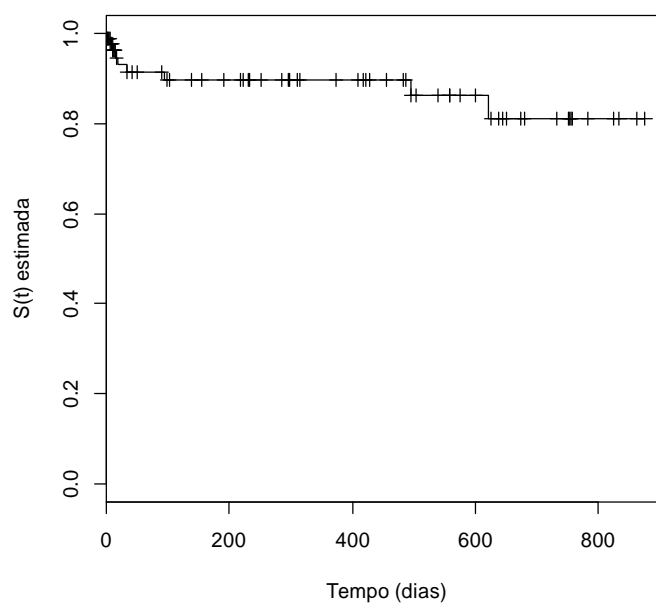


Figura 17 – Curva de sobrevivência Kaplan-Meier considerando o evento de interesse a reintervenção

O estimador Kaplan-Meier fornece um valor de 304,5 dias como uma estimativa do tempo mediano e não permite a obtenção de uma estimativa adequada para o tempo médio de vida devido à grande quantidade de censuras.

O gráfico abaixo é chamado de TTT plot (Tempo total em Teste). Ele representa a o comportamento da Função Taxa de Falha empírica. É possível observar que o gráfico apresenta uma curva convexa e depois côncava, o que significa que a função taxa de falha tem forma de U. Para modelar dados dessa natureza, na última década, foram propostas novas distribuições que são baseadas em modificações da distribuição Weibull. Entre elas, destacam-se as distribuições Weibull exponencializada, Weibull Modificada, XTG (chamada anteriormente de Weibull estendida) e Weibull aditiva.

No entanto, como a parte côncava do gráfico é bastante sutil, é possível considerar que a curva é predominantemente convexa. Sendo assim, a função taxa de falha é monotonamente crescente. Para modelar dados dessa natureza é possível utilizar a distribuição Weibull.

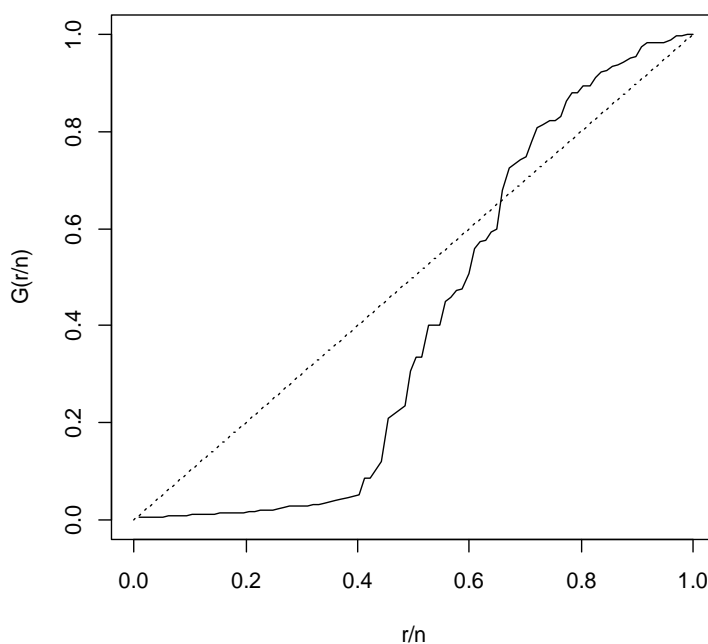


Figura 18 - TTT plot reintervenção

Observando a função de sobrevivência separada pela quantidade de artérias tratadas no gráfico abaixo, talvez seja possível dizer que os riscos são proporcionais, portanto, o teste de *logrank* para verificar se as funções de taxa de falha dos grupos são iguais, pode ser considerado.

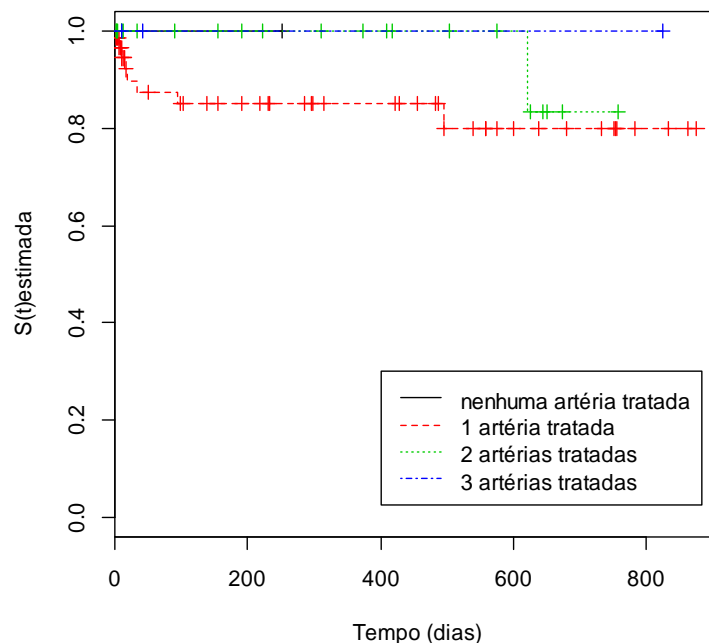


Figura 19– Estimativas de Kaplan-Meier considerando o evento de interesse a reintervenção para a variável quantidade de artérias tratadas. Os tempos representados por + mostram onde ocorreram censuras em cada grupo

O valor do teste de *logrank* para a comparação entre os grupos resultou em um p-valor = 0,545, indicando que não existe diferença entre as curvas de sobrevivência. Para o teste de *Wilcoxon*, indicado para casos cujo os riscos não são proporcionais, foi obtido um p-valor = 0,516, resultado não muito diferente do teste de *logrank*. Como a variável quantidade de artérias tratadas possui 4 categorias, é interessante testar as categorias duas a duas.

Dos resultados apresentados na tabela abaixo, consequência do teste *logrank* (coloca mesmo peso para todo o eixo do tempo, o que reforça o enfoque nos tempos maiores quando comparado ao teste de *Wilcoxon*) para comparações dos grupos dois a dois, não foram encontradas evidências de diferenças entre nenhum par de grupos.

Tabela 27 – Resultados dos testes *logrank* utilizados para as comparações de grupos, dois a dois, considerando o evento de interesse a reintervenção

Grupos comparados	Estatística de Teste (T)	p-valor
0 x 1	0,161	0,689
0 x 2	-	1
0 x 3	-	-
1 x 2	1,61	0,205
1 x 3	0,386	0,534
2 x 3	0,167	0,683

Observando o gráfico abaixo que mostra as curvas de sobrevivência estimadas por Kaplan-Meier separadas por gênero, é possível afirmar que a razão das funções de taxa de falha dos grupos é não é aproximadamente constante. Isto é, as populações não tem propriedade de taxas de falha proporcionais. Dessa forma, o teste *logrank* não é apropriado para verificar a igualdade entre as funções de sobrevivência.

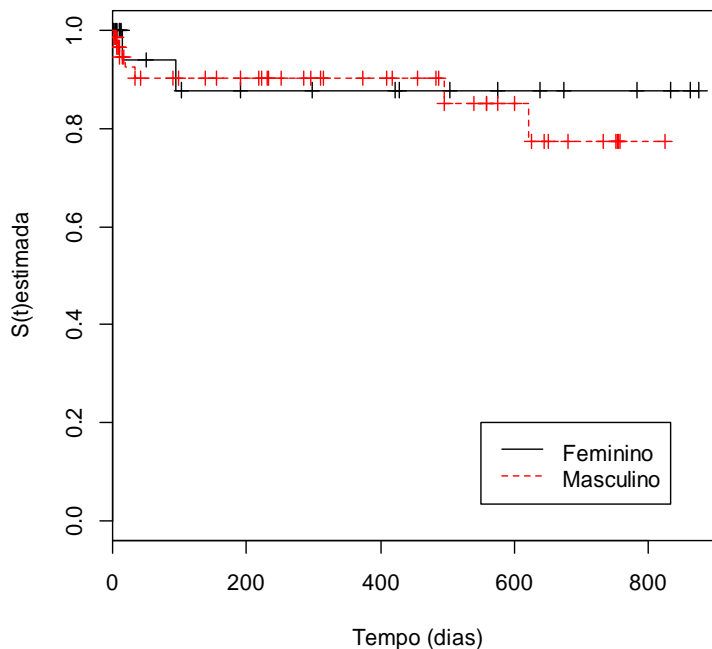


Figura 20 – Estimativas de Kaplan-Meier considerando o evento de interesse a reintervenção para a variável gênero. Os tempos representados por + mostram onde ocorreram censuras em cada grupo

Portanto, o teste *de Wilcoxon* demonstrou, com $p\text{-valor}=0,629$, que não existe diferença entre as curvas de sobrevivência apresentadas no gráfico anterior.

Tentando ajustar modelos convencionais à função de sobrevivência, foi plotado o gráfico abaixo.

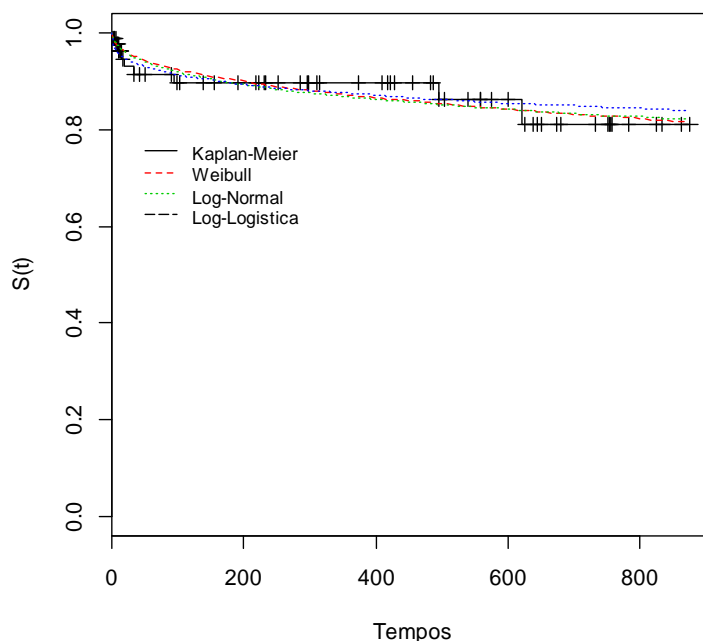


Figura 21 – Curvas de sobrevivência estimada por Kaplan-Meier considerando o evento de interesse a reintervenção para a variável resposta reintervenção e curvas de sobrevivência estimada dos modelos: Weibull, Log-normal e Log-logística

Além disso, é possível verificar que as distribuições Weibull, Log-normal e Log-Logística apresentaram comportamento muito semelhante, sendo a Log-Logística a mais diferente, visualmente falando. No entanto, o Teste da Razão de verossimilhança não é adequado para esse caso pois os modelos não são encaixados. Ou seja, um modelo não é caso particular do outro. Neste caso o indicado é usar as medidas AIC, AICc e BIC.

Tabela 28 - Comparando distribuições através das medidas AIC, AICc e BIC, considerando o evento de interesse a reintervenção

Modelo	AIC	AICc	BIC
Weibull	151,1542	151,1963	146,3767
Log-Normal	150,2544	150,2965	144,6747
Log-Logística	136,5191	136,5612	146,0928

Nesse trabalho foi utilizada a distribuição Weibull devido a sua flexibilidade. As principais vantagens da utilização da distribuição de Weibull para análise de sobrevivência é que através da estimação de apenas dois parâmetros (α e β) são obtidas informações tanto de longevidade média quanto do tipo de curva de sobrevivência. Além disso, como os valores são muito próximos, acredita-se que esse critério não é muito interessante nesse caso.

Trinta e três covariáveis potencialmente importantes para descrever o comportamento da resposta foram selecionadas para serem incluídas no modelo. Existem inúmeros possíveis modelos formados pela combinação de todas estas covariáveis. É certamente impraticável ajustar todos estes modelos possíveis a fim de

selecionar o que melhor explique a resposta. Isso implica que covariáveis importantes em termos clínicos devem ser incluídas independente de sua significância estatística, assim como a importância clínica deve ser considerada em cada passo de inclusão ou exclusão no processo de seleção de covariáveis.

Para seleção das variáveis, o modelo Weibull foi ajustado contendo uma única variável. Através do Teste da razão de verossimilhança, as variáveis consideradas significativas foram ajustadas conjuntamente. Na presença de certas covariáveis, outras podem deixar de ser significativas. Conseqüentemente, ajustam-se modelos reduzidos, excluindo-se uma covariável de cada vez. Verificam-se as covariáveis que provocam um aumento estatisticamente significativo na estatística da razão de verossimilhança. Somente aquelas que atingirem a significância ao nível de 0,10 permanecem no modelo.

Após verificar a possibilidade de inclusão de termos de interação dupla entre as covariáveis incluídas no modelo, verificou-se que as interações não eram significativas.

No modelo cujo o gênero é a única variável, essa variável não é significativa. No entanto, quando as variáveis isoladamente significativas são incluídas em um modelo junto com a variável gênero, essa última se torna significativa. Porém, a variável gênero foi retirada do modelo final após o AIC ser utilizado para comparar os dois modelos.

Sendo assim, o modelo final ficou composto pelas covariáveis e seus respectivos coeficientes: quantidade de artérias tratadas (grupo 1 = -2456,16637; grupo 2 = -1854,82245; grupo 3 = 943,54864), etílico (não = -1338,01254; sim = -1575,46479), obesidade (não = -1733,40049; sim = -2069,99726) e hipotireoidismo (sim = 126,5127). Nenhum termo de interação dupla foi significativo.

Os coeficientes estimados estão expressos na escala logarítmica dos tempos, isto é, para $Y = \log(T) = x'\beta + \sigma v$.

É importante ressaltar que para dados como os que estão sendo estudados, cujo a função de sobrevivência não tendem a zero com o passar do tempo, é interessante admitir um modelo de fração de cura, que é representado por

$$S_{pop}(t) = (1 - \phi) + \phi S(t),$$

onde $S(t)$ foi modelado pelo modelo weibull. Dessa forma, utilizando uma metodologia de máxima verossimilhança foram estimados os seguintes parâmetros: $\hat{\alpha} = 0,5112003$; $\hat{\beta} = 500,0000884$ e $\hat{\phi} = 0,2306978$. O gráfico abaixo mostra como o modelo Weibull modela os dados em comparação ao modelo Weibull com fração de cura. É possível dizer que o modelo Weibull se ajusta melhor nos tempo iniciais e que o modelo com fração de cura se ajusta melhor na cauda da distribuição. Para um melhor ajuste, poderá ser utilizado em trabalhos futuros um modelo com fração de cura com inclusão de covariáveis.

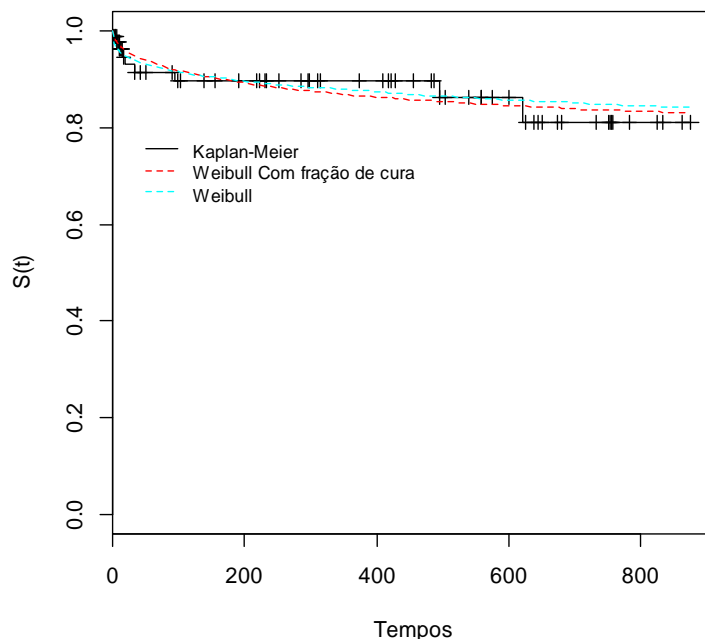


Figura 22 – Curvas de sobrevivência estimada por Kaplan-Meier considerando o evento de interesse a reintervenção para a variável resposta reintervenção e curvas de sobrevivência estimada dos modelos: Weibull e Weibull com fração de cura

Em trabalhos futuros poderá ser considerado um modelo bivariado, com duas variáveis respostas: óbito e reintervenção. Lembrando que a ocorrência do óbito é caracterizado como censura para o evento de interesse reintervenção, e o contrário não ocorre.

5.4. Variável Creatinina

Creatinina, substância que resulta do metabolismo da creatina, reconhecida como essencial para a energia muscular. (SAÚDE IG, 2013) A creatina é produzida pelo fígado, rins e pâncreas e então transportada até os músculos, onde é degradada por meio da catalisação pela enzima creatina-quinase. Essa reação se dá através da contração muscular, ou seja, toda vez que um músculo é contraído tem-se a quebra de creatina produzindo creatinina, que é posteriormente lançada na corrente sanguínea e eliminada através da urina. A creatinina é um importante indicador das condições dos rins, sobretudo para avaliar sua capacidade de filtração. Valores aumentados apontam diminuição da função renal, desidratação, obstrução do trato urinário ou doenças musculares. Os níveis diminuídos podem indicar gravidez e massa muscular reduzida. (CARDOSO, 2013)

Os níveis de creatinina no sangue são estabelecidos através de um exame bioquímico laboratorial, feito com sangue ou urina. Níveis altos de creatinina no sangue e baixos na urina são indicadores de insuficiência renal.

Valores de referência

Crianças até 12 anos = 0,3 a 0,8 mg/dL

Mulheres adultas = 0,6 a 1,0 mg/ dL

Homens adultos = 0,8 a 1,2 mg/dL

Para esse trabalho, foi utilizado um teste de médias para verificar se a média da creatinina após ICP é igual a mediana do intervalo do valor de referência. Foi realizado um teste para o gênero feminino para verificar se a média desse grupo é igual a 0,9; e para o gênero masculino para verificar se a média desse grupo é igual a 1,0.

Para o gênero masculino, obtém-se um p-valor igual a 0,003439. Portanto, deve-se rejeitar H_0 , ou seja, a média da creatinina após a ICP é maior que a mediana do valor de referência para níveis normais. É possível afirmar isso pois a estatística do teste (3,04) encontrada é maior que o limite superior do intervalo de confiança (1,52).

Para o gênero feminino, obtém-se um p-valor igual a 0,01753. Portanto, deve-se rejeitar H_0 , ou seja, a média da creatinina após a ICP é maior que a mediana do valor de referência para níveis normais. É possível afirmar isso pois a estatística do teste (2,54) encontrada é maior que o limite superior do intervalo de confiança (1,19).

Segue abaixo um box-plot para ilustrar os valores para a creatinina por sexo.

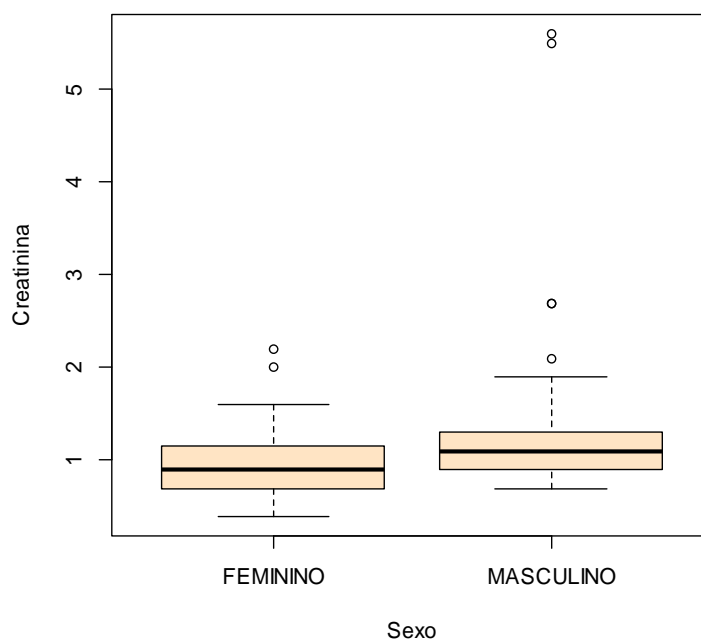


Figura 23 – Box-plot comparativo para a variável creatinina com relação ao gênero

Um estudo mais aprofundado pode ser realizado futuramente.

Conclusão

Em um estudo realizado anteriormente no Hospital Nossa Senhora da Conceição, concluíram que na análise da sobrevida dos pacientes submetidos à angioplastia coronariana com uso de stent 94% das intervenções evoluíram com sucesso, sendo que a complicação mais frequente foi a reestenose do stent. Quanto maior o número de coronárias acometidas, menor foi o tempo de sobrevida. Infarto prévio não teve relação com a sobrevida. (TREVISIO et al, 2012)

No presente estudo, inicialmente observando apenas a análise descritiva, é possível afirmar que 90% das intervenções evoluíram com sucesso. Os pacientes que realizaram uma reintervenção no período de 12 meses após o início do estudo correspondem a 11%.

Com os estudos desenvolvidos ao longo desse texto, observou-se que o modelo Weibull é adequado para estimar os tempos de vida, além de ser um modelo flexível. Porém, os dados, para ambas os eventos de interesse (óbito e reintervenção), necessitam ser modelados por um modelo que assuma fração de cura, pois a função de sobrevivência não tende a zero ao longo do tempo.

Além disso, uma análise bivariada, que assume dois eventos de interesse, pode ser utilizada em trabalhos futuros.

Após a utilização de contraste durante a ICP, o valor da creatinina no sangue é alterado para maior em ambos os gêneros (feminino, masculino). Isso foi comprovado através de uma teste de médias. Uma análise mais aprofundada pode ser realizada em estudos futuros.

Referências bibliográficas

ALBERT EINSTEIN - SOCIEDADE BENEFICENTE ISRAELITA BRASILEIRA. Angioplastia coronária ou intervenção coronária percutânea. Disponível em: <<http://www.einstein.br/Hospital/cardiologia/exames-e-testes-diagnosticos/Paginas/angioplastia-coronaria-ou-intervencao-coronaria-percutanea.aspx>>. Acesso em: 01 mai. 2013.

BANCO DE SAÚDE (Org.). Infarto Agudo do Miocárdio. Disponível em: <<http://www.bancodesaude.com.br/infarto-miocardio/infarto-agudo-miocardio>>. Acesso em: 02 maio 2013.

BERKSON, J.; GAGE, R.P. Survival curve for cancer patients following treatment. *Journal of the American Statistical Association*, Alexandria, v. 47, p. 501-515, 1952.

BRAGA, Juarez Rosso, Registros de síndromes coronarianas agudas: O que são, quais suas lições e como interpretá-los corretamente?, Reseach fellow, Division of Cardiology, Toronto General Hospital, Canada.

CARDOSO, Mayara Lopes. Creatinina. Disponível em: <<http://www.infoescola.com/bioquimica/creatinina/>>. Acesso em: 25 nov. 2013.

CARVALHO, M. S. et al. Análise de sobrevivência: Teoria e aplicações em saúde. FIOCRUZ, RIO DE JANEIRO, 2011.

COLOSIMO, Enrico Antônio; GIOLO, Suely Ruiz. Análise de Sobrevivência. São Paulo: Abe - Projeto Fisher, 2006.

COONER, F., Banerjee, S., Carlin, B., Sinha, D. (2007). Flexible cure rate modeling under latent activation schemes. *Journal of The American Statistical Association*, 102, 560-572.

COORDENAÇÃO NACIONAL DE CONTROLE DO TABAGISMO (INCA/SAS/MS). Programa Nacional de Controle do Tabagismo (PNCT) – Tratamento do Tabagismo. Disponível em: <http://189.28.128.100/dab/docs/portaldab/documentos/pnct_tabagismo.pdf>. Acesso em: 01 jul. 2013.

CORDEIRO, G. M.; ORTEGA, E.M.M.; SILVA, G.O. Modelos de Regressão Estendidos em Análise de Sobrevivência. XII Escola de Modelos de Regressão, 2011.

FACHINI, J. B. Análise de influência local nos modelos de riscos múltiplos. Dissertação (Mestrado em Agronomia) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba 2006.

FACHINI, J. B. Introdução à Análise de Sobrevivência - Notas de Aula. Departamento de Estatística, Universidade de Brasília, Brasília, 2013.

FACHINI, J. B. Modelos de regressão com e sem fração de cura para dados bivariados em análise de sobrevivência. Tese (Doutorado em Ciências) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba 2011.

FREITAS, Professor Doutor Falcão de; (FMUP), Faculdade de Medicina do Porto. ANGIOPLASTIA CORONÁRIA. Disponível em: <<http://medicina2.med.up.pt/cefa/angioplastia.htm>>. Acesso em: 28 jun. 2013.

GUBOLINO, Dr. Luiz Antônio. Angioplastia Coronária com uso da artéria radial como via de acesso. Disponível em: <http://www.institutodocoracao.com.br/noticias/noticia.asp?noticia_id=38>. Acesso em: 28 jun. 2013.

HOSPITAL DAS CLÍNICAS DA UNIVERSIDADE FEDERAL DE MINAS GERAIS (Org.). Protocolo Clínico de Síndromes Coronarianas Agudas. Disponível em: <http://portal.saude.gov.br/portal/arquivos/pdf/protocolo_clinico_sindromes_coronarianas_agudas.pdf>. Acesso em: 28 jun. 2013.

IBRAHIM, J., Chen, M., Sinha, D. (2001) Bayesian survival analysis. Springer Verlag.

INSTITUTO PROCARDÍACO. Dislipidemia – Colesterol e Triglicérides. Disponível em: <<http://www.institutoprocadiaco.com.br/dislipidemia.htm>>. Acesso em: 02 jul. 2013.

KASTRATI A, SCHÖMIG A, ELEZI S, et al. Predictive factors of restenosis after coronary stent placement. J Am Coll Cardiol. 1997;30:1428-36.

KLEIN, J.P; MOESCHBERGER, M.L. Survival analysis techniques for censored and truncated data. 1nd ed. New York: Springer-Verlag, 1997. 357p.

LABORATÓRIO OSWALDO CRUZ. Hipotireoidismo - Perfil e Tratamento. Disponível em: <<http://www.minhavidacom.br/saude/temas/insuficiencia-renal-cronica>>. Acesso em: 18 ago. 2003.

LAVALL, Guilherme. Membro Titular da SBHCI. In: SESSÃO ORDINÁRIA 18.11.10, 20., 2010, Rio de Janeiro. Balão Farmacológico: seu papel na reestenose dos stents. Rio de Janeiro: Academia Nacional de Medicina, 2010. p. 10 - 14.

LAWLESS, J. F. Statistical Methods and Models for Lifetime Data. John Wiley & Sons, New York, 1982.

MALLER, R. A., Zhou, X. (1996). Survival Analysis with Long-Term Survivors. Wiley, New York.

MELDAU, Débora Carvalho. Insuficiência Cardíaca Congestiva. Disponível em: <<http://www.infoescola.com/doencas/insuficiencia-cardiaca-congestiva/>>. Acesso em: 01 jul. 2013.

MONTEIRO, Gustavo P.; COSTA, Ricardo A.; ABIZAID, Alexandre. Intervenção coronária percutânea primária. Sca revista. São Paulo, Brasil, ano 1, v. 1, 01 jul. 2011.

MOTA, Thiago Santos; SILVEIRA, Lician Vaz de Arruda; ANTUNES, Aline Araujo. Modelo De Cox Para Eventos Cardiovasculares Recorrentes Em Pacientes Sob Dialise Com Covariáveis Medidas No Tempo. Revista Brasileira de Biometria, São Paulo, v. 1, n. 30, p.150-159, 01 jan. 2012.

NEFROPATIA INDUZIDA POR CONTRASTE: COMO PREVENIR? São Paulo: Revista da Associação Médica Brasileira, v. 48, n. 1, 01 jan. 2002. Trimestral. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-42302002000100011>. Acesso em: 20 jun. 2013.

NOGUEIRA, Dr. Marcelo C. Angioplastia. Disponível em: <<http://www.drmarcelonogueira.med.br/manuais/Angioplastia.htm>>. Acesso em: 20 jun. 2013.

O QUE é Insuficiência renal crônica? Disponível em: <<http://www.minhavidacom.br/saude/temas/insuficiencia-renal-cronica>>. Acesso em: 02 jul. 2013.

PACIEVITCH, Thais. Obesidade. Disponível em: <<http://www.infoescola.com/doencas/obesidade/>>. Acesso em: 01 jul. 2013

PINTON, Fábio A.; RIBEIRO, Prof. Dr. Expedito E. Acesso Radial versus Femoral em Síndrome Coronariana Aguda com Supra desnivelamento do Segmento ST. Disponível em: <http://cientifico.cardiol.br/cardiosource2/cardio-interv/int_artigo112.asp?cod=537>. Acesso em: 28 jun. 2013.

PINTON, Fábio A.; RIBEIRO, Prof. Dr. Expedito E. Baixo risco de trombose de stent e de reestenose com o uso irrestrito de stents farmacológicos de “nova geração”: Registro SCAAR. Disponível em: <http://cientifico.cardiol.br/cardiosource2/cardio-interv/int_artigo97.asp?cod=454>. Acesso em: 27 jun. 2013.

PRENTICE, R. L., Kalbfleisch J. D., Peterson A. V. Jr., Flournoy N., Farewell V. T., Breslow N. E. (1978). The Analysis of Failure Times in the Presence of Competing Risks. *Biometrics*.

SABRINA. O que é sedentarismo? Disponível em: <<http://www.nutricaoempratica.com.br/alimentacao-correta/o-que-e-sedentarismo>>. 09 jun. 2009.

SAÚDE IG. Creatinina. Disponível em: <<http://saude.ig.com.br/minhasaude/exames/creatinina/ref1237835469635.html>>. Acesso em: 25 nov. 2013.

SGRILLO, R. B., A distribuição de Weibull como modelo de sobrevivência de insetos, Centro de Energia Nuclear na Agricultura – CENA/USP, Piracicaba, 1982.

SILVA, Enf.º Alisson Daniel Fernandes da. Hipertensão Arterial Sistêmica: Uma doença silenciosa. Disponível em:
<<http://www.portaleducacao.com.br/enfermagem/artigos/9731/hipertensao-arterial-sistemica-uma-doenca-silenciosa>>. Acesso em: 30 jun. 2010.

SOCIEDADE BRASILEIRA DE DIABETES (SBD). Tipos de diabetes. Disponível em:
<<http://www.diabetes.org.br/tipos-de-diabetes>>. Acesso em: 02 jul. 2013.

TREVISIO, Daisson José et al. Análise de sobrevida em pacientes submetidos à angioplastia coronariana com stent em um hospital da região sul de Santa Catarina. *Scientia Medi*, Porto Alegre, v. 22, n. 2, p.91-96, 01 dez. 2012.

TSODIKOV, A. D., Ibrahim, J. G., Yakovlev, A. Y. (2003). Estimating cure rates from survival data: an alternative to two-component mixture models. *Journal of the American Statistical Association*, 98, 1063-1078.