



Análise do tempo até a germinação de sementes de *Jatropha curcas* L.: uma aplicação com uso de técnicas de análise de sobrevivência

Carla Franciele Höring¹, Renan Douglas Floriano Scavazzini¹ e Daniele Cristina Tita Granzotto²

¹Graduando(a) do curso de Estatística, Universidade Estadual de Maringá (UEM)

²Docente, Departamento de Estatística, Universidade Estadual de Maringá (UEM)

RESUMO

Muitas são as espécies que podem ser usadas para a extração de óleos vegetais com a finalidade de gerar recursos energéticos. Neste trabalho estudamos a espécie *Jatropha curcas* L., popularmente conhecida como pinhão-manso, a qual apresenta forte resistência à seca e é economicamente eficiente produzindo, no mínimo, duas toneladas de óleo por hectare com tempo de vida útil de produção em torno de 40 anos. Com o objetivo de estudar os tempos de germinação das sementes desta espécie, tempos estes importantes e diretamente relacionados ao retorno econômico de seu plantio, usamos técnica não paramétricas em análise de sobrevivência, tais como os estimadores de Kaplan-Meier e Nelson-Aalen. Algumas covariáveis tais como, o meio onde a semente foi plantada, a temperatura e o tempo de armazenamento das sementes, foram consideradas no estudo.

Palavras chave: Sobrevivência, Métodos não-paramétricos, Nelson-Aalen, Kaplan Meier, Regressão.

1 INTRODUÇÃO

Uma espécie amplamente cultivada para ser utilizada na extração de óleos vegetais, com a finalidade de gerar recursos energéticos, é a *Jatropha curcas* L., popularmente conhecida como pinhão-manso. Esta é uma planta oleaginosa exigente em insolação, a qual apresenta forte resistência à seca e é economicamente eficiente produzindo, no mínimo, duas toneladas de óleo por hectare com tempo de vida útil de produção em torno de 40 anos.

Originário da América do Sul, esta espécie vegetal foi introduzida no século XVIII, e era usada na época para aumentar a ação purgativa do óleo de ricino, com o qual era misturado. No Brasil, o pinhão-manso ocorre praticamente em todas as regiões, sempre de forma dispersa, adaptando-se em condições edafoclimáticas variáveis. De modo geral, cresce rapidamente em solos pedregosos, de baixa umidade e terrenos abandonados, ver por exemplo Biodieselbr (2007).

A obtenção de sementes de alta qualidade é o requisito principal para que uma espécie possa ser utilizada pelos agricultores no estabelecimento de suas lavouras. Na semente estão contidas todas as informações genéticas necessárias para a geração de uma nova planta contribuindo para a produtividade. A relação entre a germinação obtida no laboratório e a emergência das plântulas no campo é a causa do desenvolvimento do conceito de vigor.

No presente trabalho temos por objetivo de estudar o tempo de germinação das sementes da espécie *Jatropha curcas* L. Para tal, faremos o uso de técnicas de análise de sobrevivência na presença de covariáveis explicativas e tempos censurados. Área estatística em pleno desenvolvimento e expansão, análise de sobrevivência, em sua essência, consiste numa coleção de procedimentos estatísticos utilizados para analisar dados relacionados ao tempo até a ocorrência de um determinado evento de interesse, neste caso, germinação das sementes. O estudo é feito à partir de um tempo inicial pré-estabelecido, (o momento em que a semente foi hidratada) e tem como complicador os casos onde a germinação não ocorre no prazo de estudo determinado, 30 dias.

2 METODOLOGIA

Usando técnicas de análise de sobrevivência, temos o interesse em analisar o tempo até a germinação das sementes da espécie *Jatropha curcas* L., importante na fabricação do biocombustível e com grande potencial econômico. Por apresentar algumas características positivas particulares à espécie, suas sementes são altamente comercializadas e portanto, merecem um estudo mais aprofundado quando a qualidade da germinação.

O estimador de Kaplan-Meier foi proposto por Kaplan e Meier (1958) (também chamado de estimador produto-limite, ver por exemplo [4]). Por se tratar de um estimador não-viciado, fracamente consistente, com distribuição assintótica normal e ser um estimador de máxima verossimilhança de uma das principais funções, a função de sobrevivência $S(t)$, este estimado é amplamente usado na área de sobrevivência e confiabilidade. O procedimento do método é o de encontrar uma estimativa para a função de sobrevivência e, a partir dela, estimar estatísticas de interesse tais como, dentre outras, o tempo médio ou mediano e alguns percentis, considerando para construção um número de intervalos de tempo igual ao número de falhas distintas sendo, os limites dos intervalos, os tempos de falha da amostra.

Suponha que existem n sementes em estudo e $k(\leq n)$ falhas (sementes sem germinação) nos tempos $t_1 < t_2 < \dots < t_k$. Considerando $S(t)$ como uma função discreta com probabilidade maior que zero somente nos tempos de falha t_i , $i = 1, \dots, k$, tem-se que $S(t_i) = (1 - q_1)(1 - q_2) \dots (1 - q_i)$, onde q_i é a probabilidade de uma semente germinar no intervalo $[t_{i-1}, t_i)$ sabendo que ele não germinou até t_{i-1} e considerando $t_0 = 0$.

Ou seja, $q_i = P(T \in [t_{i-1}, t_i) | T \geq t_{i-1})$ e o estimador de Kaplan-Meier reduz-se então, a estimar q_i que é dado por, \hat{q}_i que é o número de germinações no intervalo $[t_{i-1}, t_i)$ dividido pelo número de sementes sob risco em t_{i-1} para $i = 1, \dots, k$.

Então o estimador de Kaplan-Meier, é então, dado por:

$$\hat{S}(t) = \prod_{j:t_j \leq t} \left(\frac{n_j - d_j}{n_j} \right) = \prod_{j:t_j \leq t} \left(1 - \frac{d_j}{n_j} \right) \quad (1)$$

para d_j o número de falhas em t_j , $j = 1, 2, \dots, k$, n_j o número de indivíduos em risco em t_j , ou seja, os indivíduos que não falharam e não censuraram até o instante imediatamente anterior a t_j .

Outro estimador usado neste trabalho é o de Nelson-Aalen, estimador não paramétrico da função de risco, ver por exemplo [3]. Para tal vamos definir $A(t) = \int_0^t \alpha(x) dx$, sendo $\alpha(t) = \frac{f(t)}{S(t)}$ a função taxa de falhas. Assim $\alpha(t) = \lim_{h \rightarrow 0} \frac{A(t+h) - A(t)}{h} \xrightarrow{h \rightarrow 0} A(t+h) - A(t) \approx h\alpha(t)$. Desta forma $h\alpha(t) \approx P(\text{ocorrer a germinação em } (t, t+h] | \text{Não ocorreu ainda})$ e no instante de falha t_i essa quantidade pode ser estimada por d_i/n_i . O estimador de Nelson-Aalen para $A(t)$ é dado por

$$\hat{A}(t) = \sum_{i:t_i < t} \frac{d_i}{n_i}.$$

3 RESULTADOS

Para este estudo estamos considerando os dados relacionados ao tempo de falha, neste caso tempo de germinação, das sementes da espécie *Jatropha curcas* L., coletados no Laboratório de Tecnologia de Sementes e Mudanças da Universidade Estadual do Oeste do Paraná (UNIOESTE) no ano de 2007.

Para os testes de germinação foram utilizadas quatro repetições de 25 sementes em dois ambientes (areia e rolo de papel) e duas temperaturas constantes (25°C e 30°C), considerando cinco períodos (em dias) de armazenamento em local não controlado: 17, 40, 66, 108 e 164 dias. As contagens do número de plântulas germinadas foram realizadas diariamente durante 30 dias. Considerou-se semente germinada a protrusão da radícula e a emissão dos cotilédones no substrato rolo de papel, e a presença dos cotilédones acima da superfície no substrato areia.

Com relação aos tempos de germinação quando desconsideramos as possíveis influências das variáveis, observamos que tempo mínimo foi de 3 dias com uma mediana de 9 dias, se considerarmos os tempos censurados, ou uma mediana de tempo de 9 dias quando desconsideramos as censuras. Além disso os tempos médios observados foram de 16.25 e 6.83 dias, com e sem censuras respectivamente, com desvio padrão de 11.59 e 2.83. Figura 1, painel à esquerda apresenta a distribuição dos tempos observados, ou seja, dos tempos de germinação de todas as sementes que de fato experimentaram o evento de interesse.

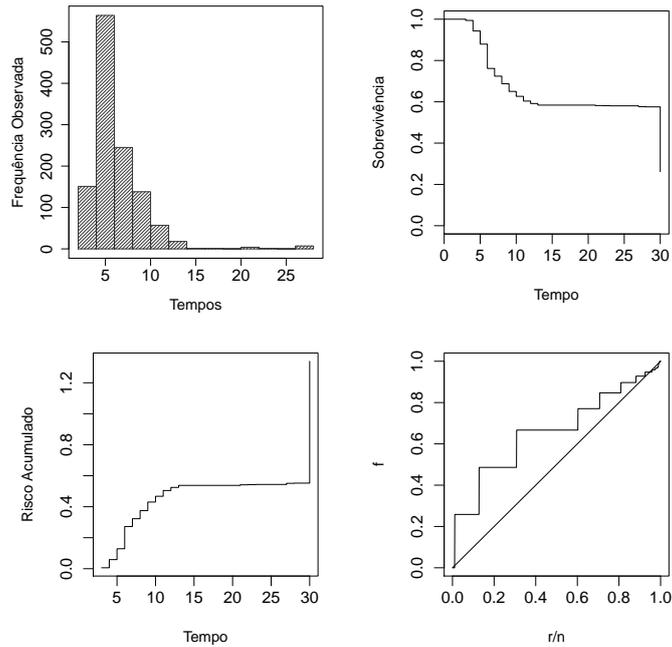


Figura 1: Histograma dos tempos de germinação observados, curva de sobrevivência estimada por Kaplan-Meier, curva de risco acumulada e TTTPlot dos tempos observados, respectivamente.

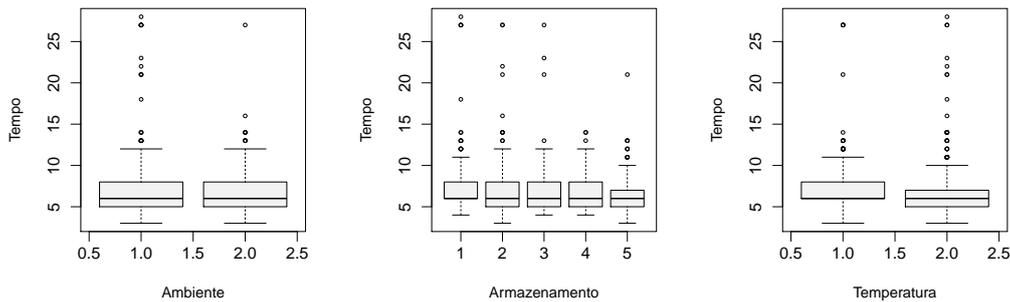


Figura 2: Análise dos tempos observados com relação ao ambiente, tempo de armazenamento e temperaturas.

Já as Figuras 1, painéis à direita acima e à esquerda abaixo, respectivamente, apresentam a curva de sobrevivência estimada por Kaplan-Meier e de risco acumulada estimada. Através da curva de sobrevivência é possível observar que aproximadamente 55% das sementes não germinam durante o experimento, percentual alto e de alto impacto econômico. Podemos ainda observar que o tempo mediano não parametricamente estimado está entre 5 e 6 dias.

Ainda com relação aos tempos globalmente observados, podemos observar à partir do TTTPlot (*Total Time on Test*, ver por exemplo [1]), apresentado na Figura 1, painel à direita abaixo, um possível comportamento unimodal da função de taxa de falha. Ou seja, o risco de germinação nos primeiros dias é crescente, atingindo um pico (moda) em algum tempo x . Após este tempo, a taxa de falha muda de forma, tornando-se decrescente com o passar do tempo. Este tipo de taxa de falha pode ser modelada por distribuições do tipo log-logística, ou transmutada log-logística, por exemplo.

Ao observarmos o comportamento do tempo de germinação segundo as covariáveis apresentadas, verificamos que o período de armazenamento em local não controlado, bem como o tipo de ambiente e a temperatura onde a semente é colocada para germinar, são fatores determinantes para caracterizar a viabilidade e o vigor da espécie. Observe por exemplo a Figura 2, primeiro boxplot acima. Apesar do tempo mediano permanecer próximo quando apenas consideramos os tempos observados, este se torna um fator altamente significativo quando incluímos na análise os tempos censurados. Isso ocorre para todas as covariáveis aqui apresentadas, ver Figuras 3 para comparação.

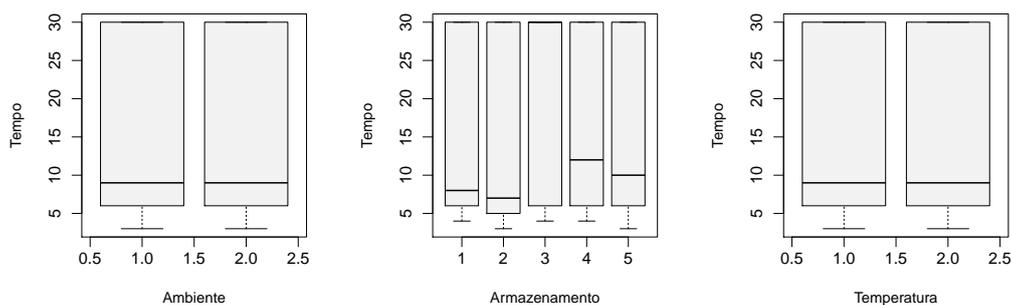


Figura 3: Análise dos tempos observados e censurados com relação ao ambiente, tempo de armazenamento e temperaturas.

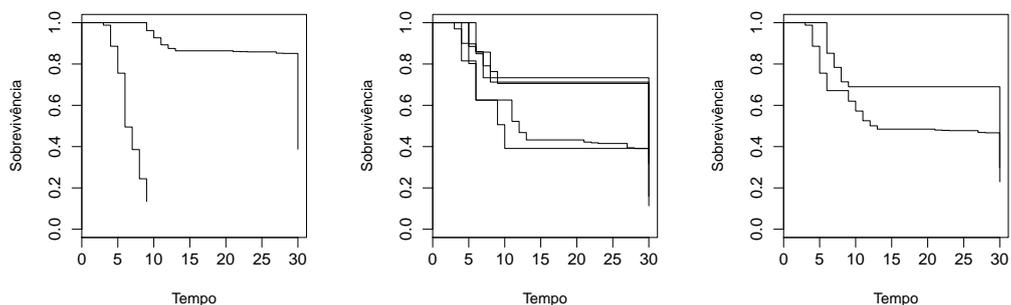


Figura 4: Curvas de sobrevivência quando consideramos as covariáveis: ambiente, tempo de armazenamento e temperaturas.

As Figuras 4 apresentam as curvas de sobrevivência quando consideramos as covariáveis: ambiente, período de armazenamento e temperaturas, respectivamente. Podemos observar que a característica de longa-duração é observada em todos os casos uma vez que a curva de sobrevivência não converge para zero como esperado. Desta forma, não importa a covariável e nem o seu nível, sempre teremos um percentual de pelo menos 30% de sementes que não vão germinar em momento algum, o que indica uma possível modelagem dos tempos por meio de modelos de sobrevivência com fração de cura.

4 CONCLUSÃO

Neste trabalho buscamos estudar os tempos de germinação das sementes da espécie *Jatropha curcas* L. Para tal técnicas não paramétricas em análise de sobrevivência foram utilizadas, tais como o estimador de Kaplan-Meier e de Nelson-Aalen. Por se tratar de um trabalho de iniciação científica ainda no início, uma possível modelagem paramétrica dos tempos esta sendo estudada pelos alunos envolvidos.

Referências

- [1] BERGMAN, Bo; KLEFSJÖ, Bengt. Total Time on Test Plots. Encyclopedia of Statistics in Quality and Reliability, 2007.
- [2] GIOLO, Suely Ruiz; COLOSIMO, Enrico Antônio. Análise de sobrevivência aplicada. Edgard Blucher, 2006.
- [3] MARSHALL, Albert W.; OLKIN, Ingram. Life distributions. Springer, New York, 2007.
- [4] LAWLESS, J. F. Parametric models in survival analysis. Encyclopedia of Biostatistics, 1998.