

QUALIDADE DE SEMENTES DE SOJA: UM ESTUDO VIA PREDIÇÃO DE QUANTIDADES ALEATÓRIAS

Rafaela Cardoso Alves¹

RESUMO

A deterioração de sementes de soja durante o armazenamento é um processo inevitável que leva a redução da viabilidade das sementes de um lote. A predição da evolução da deterioração no armazenamento pode possibilitar tomada de decisões importantes sobre o aproveitamento e descarte de lotes de sementes. O objetivo deste trabalho foi estabelecer uma metodologia para avaliar a viabilidade das sementes armazenadas, a partir da predição pontual e intervalar do número de sementes inviáveis, determinadas no teste de tetrazólio num intervalo de tempo futuro. Para realizar este estudo, foram utilizados dois tipos de dados: reais e simulados. Os dados reais foram obtidos em experimento no Laboratório de Análise de Sementes da UFLA e utilizou as cultivares M.Soy e Serra Negra avaliadas semanalmente em sua viabilidade pelo teste de tetrazólio por um período de 10 semanas. O programa aplicado para a realização das simulações foi desenvolvido em linguagem C (KERNIGHAN; RICHTER, 1988). Essa escolha ocorreu devido à fácil adaptação à metodologia na obtenção dos tempos de vida desejados. Com essa metodologia constatou-se que os dados reais e os dados obtidos via simulação seguem uma distribuição Weibull e que é possível conhecer a distribuição dos tempos de vida, quando se trabalha com ensaios destrutivos. O cálculo da probabilidade de cobertura comprovou que os intervalos de predição, para o número de subamostras falhas em um intervalo de tempo futuro contidos nos resultados, são confiáveis.

Palavras-chave: Qualidade de semente de soja. Predição de quantidades aleatórias. Distribuição *Weibull*. Ensaios destrutivos.

SOYBEAN SEED QUALITY: A RONDON QUANTITIES PREDICTION STUDY

ABSTRACT

The soybean seeds deterioration during storage is an inevitable process which leads to reduction of viability of the seed lots. The prediction of the evolution of deterioration in storage can enable important decisions about the use and disposal of seed lots. The aim of this study was to establish a methodology for assessing the viability of stored seeds from the point and interval prediction of the number of non-viable seeds, as determined in tetrazolium test in a future time interval. To conduct this study, we used two types of data: real and simulated. The actual data was obtained in an experiment at the Seed Analysis Laboratory of UFLA. It was used the cultivars, M.Soy and Serra Negra which were evaluated weekly on their viability by the tetrazolium test for a period of 10 weeks. The program used to perform the simulations, was developed in C (KERNIGHAN; RICHTER, 1988). This choice was due to easy adaptation to the methodology in obtaining the desired lifetimes. It was found that the actual data and the data obtained via simulation follow a Weibull distribution and it is possible to see the lifetime distribution, when working with destructive testing. The calculation of the coverage's probability proved that the prediction intervals to the number of sub-samples faults in a range of future time contained in the results are reliable.

¹ Mestre em Estatística e Experimentação Agropecuária pela Universidade Federal de Lavras (UFLA), graduada em Matemática pela Universidade Federal de Uberlândia (UFU) e atualmente é professora do Instituto Federal do Triângulo Mineiro (IFTM). E-mail: rafaela@iftm.edu.br.

Keywords: Soybean seed quality. Prediction of random numbers. Weibull distribution. Destructive testing.

1 INTRODUÇÃO

Estudos de Predição de Quantidades Aleatórias na obtenção do número provável de falhas em um determinado tempo são de grande importância, podendo auxiliar na solução de problemas futuros (NORDMAN; MEEKER, 2002). Um exemplo da aplicação desses estudos é o caso das sementes em que se deseja avaliar a viabilidade no armazenamento, isto é, número de sementes inviáveis num determinado intervalo de tempo futuro. Tal predição é interessante para empresas que comercializam sementes, agricultores e pecuaristas, os quais freqüentemente se defrontam com problemas na escolha dos lotes de maior qualidade. A avaliação da qualidade fisiológica de sementes de soja pode ser efetuada tradicionalmente pelo teste de germinação, no entanto testes como o de tetrazólio, conforme metodologia descrita por França Neto, Krzyzanowski e Costa (1998), além de permitir uma avaliação mais rápida da viabilidade, também possibilita distinção dos tipos e danos a que as sementes foram submetidas no processo de produção. Os métodos de predição pontual e intervalar aplicados em dados obtidos via experimento e por simulação podem ser utilizados na análise de ensaios não destrutivos (NELSON, 2000). Mas, pouco se sabe sobre a sua aplicação em experimentos destrutivos, como é o caso da aplicação do teste de tetrazólio na avaliação de sementes de soja. A Análise da viabilidade das sementes a partir do conhecimento do número de sementes inviáveis, no intervalo de tempo compreendido entre o início do armazenamento até um tempo censurado, possibilita prever o número de sementes não viáveis, num intervalo de tempo futuro preestabelecido. Um trabalho relacionado com a predição de quantidades aleatórias é o realizado por Nelson (1972), que apresenta intervalos de confiança e testes de hipóteses para a razão de proporções de uma população multinomial, usando amostra aleatória simples de uma população infinita. O autor apresenta três exemplos: estudo da preferência do consumidor, escolha de votos, tempo de vida de uma moeda em circulação, ilustrando o uso de intervalos de confiança para a razão de proporções. Nelson e Hahn (1973) apresentam intervalos de predição de amostras futuras contendo resultados de predição simples (de uma população com N unidades, todas elas com igual probabilidade de seleção) e intervalos de predição simultâneos e suas definições. Estes intervalos de predição para uma amostra futura são apresentados para populações: normal, binomial, poisson e exponencial. Também são discutidos intervalos de predição não paramétricos, que utiliza diferentes exemplos para cada população apresentada. O mesmo autor sugere que usando outras distribuições como *Weibull*, Valor Extremo, Lognormal e Gama também é possível o cálculo do intervalo de predição. Em particular, uma interessante discussão sobre a distribuição *Weibull* pode-se encontrar em Nelson (1985), que apresenta estimadores pontuais e limites de confiança para dados censurados múltiplos e desenvolve a teoria base para

metodologias particularmente usadas, quando se tem pouca ou nenhuma falha. Escobar e Meeker (1999) descrevem métodos para a construção de limites e intervalos de predição assumindo predição de novas amostras e predição dentro de amostras, para resultados futuros, usando dados de vida censurados. Estudos deste tipo são realizados atualmente na engenharia, geralmente quando se quer prever o número de componentes de um determinado equipamento, que falharão num intervalo de tempo futuro.

Portanto, o objetivo desta pesquisa é propor um método estatístico para prever o potencial de armazenamento de lotes de sementes a partir da predição do número de sementes viáveis que são determinadas através do teste de tetrazólio, num determinado intervalo de tempo futuro.

2 MATERIAL E MÉTODOS

Para a aplicação da metodologia proposta neste trabalho foram utilizados dois lotes de sementes de soja das cultivares M-Soy e Serra Negra, com porcentagem inicial de viabilidade 90% e 85%, respectivamente. Armazenadas em condições ambientais com temperatura média de 28°C e umidade relativa do ar 70%. A avaliação da viabilidade das sementes ao longo do armazenamento por 10 semanas foi efetuada no Laboratório de Análises de Sementes da Universidade Federal de Lavras (UFLA). Para conhecer a distribuição probabilística dos dados correspondentes ao tempo até a deterioração da semente (T), a seguinte metodologia foi proposta:

2.1 Obtenção do Tempo até a Deterioração da Semente (T)

O método seguido para obtenção de T está esquematizado no Anexo A e é descrito a seguir. Supondo sementes ensacadas para a comercialização, divide-se o conteúdo de um saco de sementes em “M” amostras com “n” sementes cada uma. De cada amostra retira-se, aleatória e sistematicamente, subamostras de tamanho “n₁”, isto é, ter-se-ia agora “M” subamostras. Em cada subamostra aplica-se o Teste de Tetrazólio para determinar a porcentagem de sementes inviáveis. Se a porcentagem ultrapassar um determinado valor “P”, preestabelecido, então se considera a subamostra como inviável e registra-se o tempo em que foi realizado o teste que, em geral, é dado em semanas. Uma subamostra considerada inviável, implica parar de retirar mais subamostras da amostra respectiva. Uma subamostra considerada viável implica na seleção de uma nova subamostra de tamanho “n₁”, da amostra respectiva. Assim, aplica-se o Teste de Tetrazólio nestas novas subamostras e determina-se a porcentagem de sementes deterioradas que ultrapassam “P”, continua-se subamostrando e aplicando o Teste de Tetrazólio. O procedimento continua até que a porcentagem de sementes deterioradas da subamostra ultrapasse o valor “P” (no qual o dado registrado para a amostra correspondente será o tempo em que o teste foi aplicado) ou até terminar com as sementes (sendo que, neste caso, o tempo a ser registrado será um dado censurado). Dessa forma, é possível obter “M” dados, correspondente ao tempo até a perda da viabilidade de uma subamostra, com os quais foi possível determinar a respectiva

distribuição de probabilidade. Após a aplicação do método foi possível obter o tempo até a perda da viabilidade da subamostra (T), isto é, em qual semana cada subamostra falhou. Se a amostra não perder a viabilidade até a última semana, T será considerado censurado. Utilizando-se os dados obtidos na descrição acima, foi possível determinar a distribuição de probabilidades de T . Uma vez determinada a distribuição de probabilidades, foi efetuada a predição do número de subamostras deterioradas, por saco, em um intervalo de tempo.

2.2 Predição do número futuro de falhas

Supondo que 1 saco de sementes contenha no total “ N ” subamostras, temos que $X + Y + Z = N$ em que, (X, Y, Z) segue uma distribuição Trinomial com probabilidades: $p =$ probabilidade de uma subamostra falhar em um único experimento no intervalo de tempo $[0, t_c)$, $q =$ probabilidade de uma subamostra falhar em um único experimento no intervalo de tempo $[t_c, t_w)$, $r =$ probabilidade de uma subamostra falhar em um único experimento no intervalo de tempo $[t_w, \infty)$.

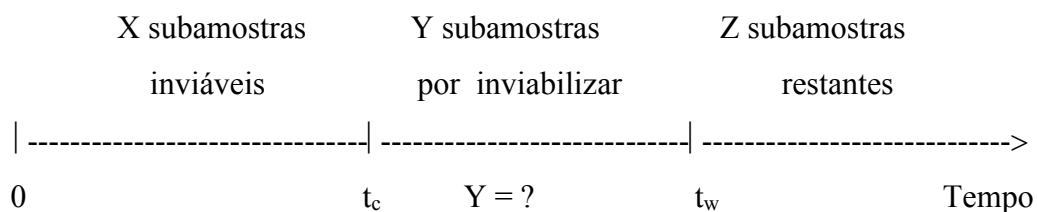


Figura 1 – Esquema de predição do número futuro de subamostras inviáveis
Fonte: Autor.

Os valores de p , q e r dependerão da distribuição dos dados. Obtidos T , distribuição de T e valores de p , q e r as predições pontual e intervalar do número de sementes inviáveis em um intervalo de tempo futuro foram calculadas, segundo Nelson (2000 *apud* NORDMAN; MEEKER, 2002) da seguinte forma:

2.2.1 Predição Pontual

O número de subamostras falhas em um intervalo do tempo futuro $[t_c, t_w]$ será dado por:

$$\hat{Y} = N * \hat{q} \quad (1)$$

em que N é o número total de subamostras e \hat{q} o estimador de máxima verossimilhança de q .

2.2.2 Predição Intervalar

A predição intervalar é feita usando-se o procedimento da razão de verossimilhança. Sendo que (x, y, z) tem distribuição Trinomial com parâmetros p , q e r .

As hipóteses testadas foram:

$$H_0: \theta = \begin{pmatrix} p \\ q \\ r \end{pmatrix} \text{ assume valores relacionados com a distribuição dos dados.}$$

$$H_1: \theta = \begin{pmatrix} p \\ q \\ r \end{pmatrix} \text{ não assume valores relacionados com a distribuição dos dados.}$$

Sendo que $(X, Y, Z) \sim \text{Trinomial}(p, q, r)$, então a função densidade é dada por:

$$f_{X,Y,Z}(x,y,z) = \frac{N!}{x!y!z!} p^x q^y r^z \quad (2)$$

A verossimilhança multinomial irrestrita dos parâmetros p e q é dada por:

$$L(p,q;x,y) = \frac{N!}{x!y!(N-x-y)!} p^x q^y (1-p-q)^{N-x-y} \quad (3)$$

Maximizando, tem-se:

$$L^*(x,y) \quad (4)$$

A função de verossimilhança restrita, em que p , q e r dependem da distribuição dos dados, é determinada sob H_0 e denotado por:

$$K(\Theta;x,y) \quad (5)$$

Maximizando (5), tem-se:

$$K^*(x,y) \quad (6)$$

Assim, usando (4) e (6) a razão de verossimilhança será dada por:

$$\lambda = \frac{K^*(x,y)}{L^*(x,y)} \quad (7)$$

Aplicando-se $-2 \log$ na razão de verossimilhança λ , tem-se:

$$Q(x,y) = -2 \log (K^*(x, y) - L^*(x, y)) \quad (8)$$

em que $Q(x,y)$ segue aproximadamente uma distribuição Qui-quadrado com 1 grau de liberdade.

Logo,

$$\Pr (Q(x, y) \leq \chi^2_{(1-\alpha; 1)}) \cong 1 - \alpha \quad (9)$$

Dado o número de falhas X no intervalo de tempo $[0, t_c]$, ao nível $100(1 - \alpha)\%$. O menor valor de $y(\underline{Y})$ que satisfaz a equação (9) será considerado o limite inferior para o intervalo de predição para Y . O maior valor de $y(\tilde{Y})$ que satisfaz a equação (9) será considerado o limite superior para o intervalo de predição para Y .

2.3 Cálculo da Probabilidade de Cobertura para Intervalos de Predição Bilateral

Para a obtenção da probabilidade de cobertura do preditor intervalar, para comprovar que o método usado para determinar intervalos de predição é confiável mediante simulação foram considerados os seguintes parâmetros: número de amostras = 120; tamanho da amostra = 1000;

tamanho da subamostra = 10; pdi (probabilidade da inviabilidade inicial) = 0,10 e $P = 0,15$. Com esses parâmetros determina-se os intervalos de predição $[Y; \tilde{Y}]$ $[Y$, com nível de confiança igual a 95% para cada simulação realizada.

Sendo $X = x$ como número de subamostras falhas no intervalo $[0, t_c]$, então, Y tem distribuição binomial com parâmetros $N-x$ e π , em que π é dado pela seguinte expressão:

$$\pi = \frac{q}{q-r}$$

Logo, π é a probabilidade condicional em que a subamostra seja considerada inviável no intervalo de tempo $[t_c, t_w]$, dado que não perdeu a viabilidade até t_c .

Conhecendo x e obtendo $[Y; \tilde{Y}]$ a probabilidade de cobertura condicional (PC) ao valor de x para o intervalo de predição (IP) $[Y; \tilde{Y}]$, com nível nominal de $100(1-\alpha)\%$ será dada por:

$$PC[IP(1-\alpha) | x] = \Pr[Y < Y < \tilde{Y} | x]$$

A probabilidade de cobertura não condicional (ao valor de x) será dada pela seguinte expressão:

$$PC[IP(1-\alpha)] = E_x \{PC[IP(1-\alpha) | x]\}$$

Esta expressão representa a média das probabilidades de cobertura obtidas com os dados de cada simulação.

3 RESULTADOS E DISCUSSÃO

Serão apresentados tabelas e gráficos dos resultados do número de subamostras futuras que seriam consideradas inviáveis no intervalo de tempo $[4,16]$ semanas. Em virtude de serem inúmeros os resultados provenientes de dados obtidos via simulação e experimento real, foram selecionados alguns para serem apresentados. Os gráficos mostram qual distribuição probabilística melhor se adequa aos dados e as tabelas contem os valores de \hat{b} (estimativa do parâmetro, forma da distribuição *weibull*), $\hat{\eta}$ (estimativa do parâmetro, escala da distribuição *weibull*), \hat{q} (estimativa de máxima verossimilhança de q), x (número de subamostras falhas em $[0,4]$ semanas, preditor pontual (\hat{Y}) e preditor intervalar $[Y; \tilde{Y}]$.

3.1 Adequação do modelo (dados obtidos via simulação utilizando para a realização das simulações um programa desenvolvido em linguagem C)

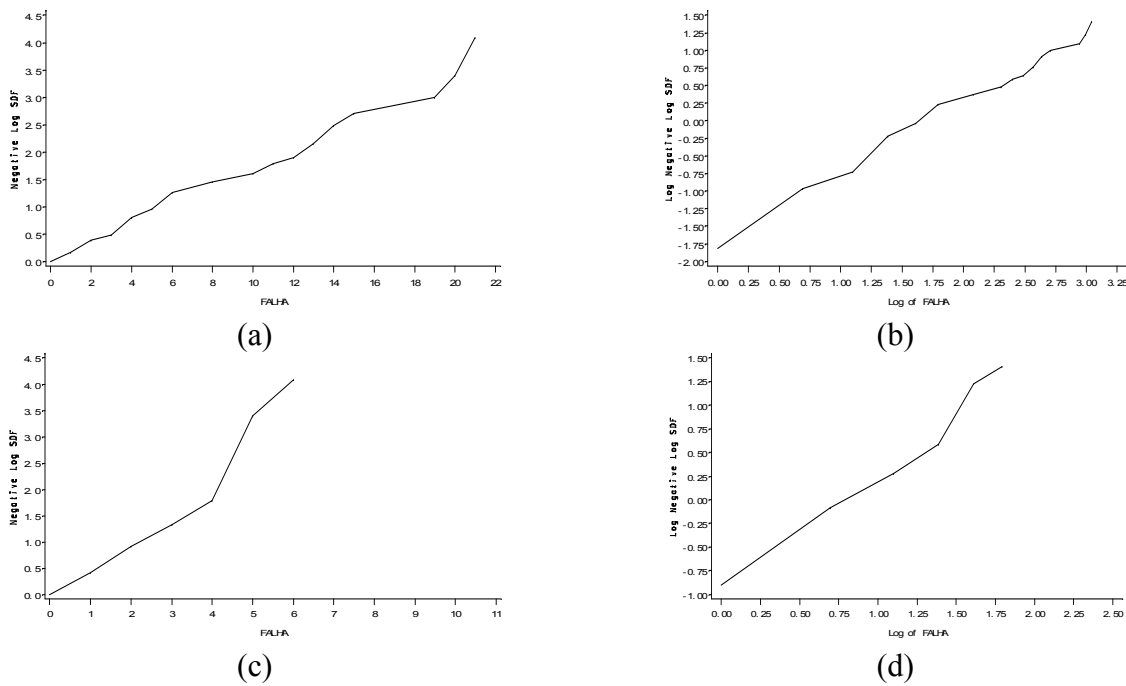


Gráfico 1 – Adequação de probabilidade normal (a e c) e weibull (b e d) para $n_a=60$, $n_1=15$, $p_{di}=0.10$ (a e b), $p_{di}=0.15$ (c e d) e $p=0.15$
 Fonte: Autor.

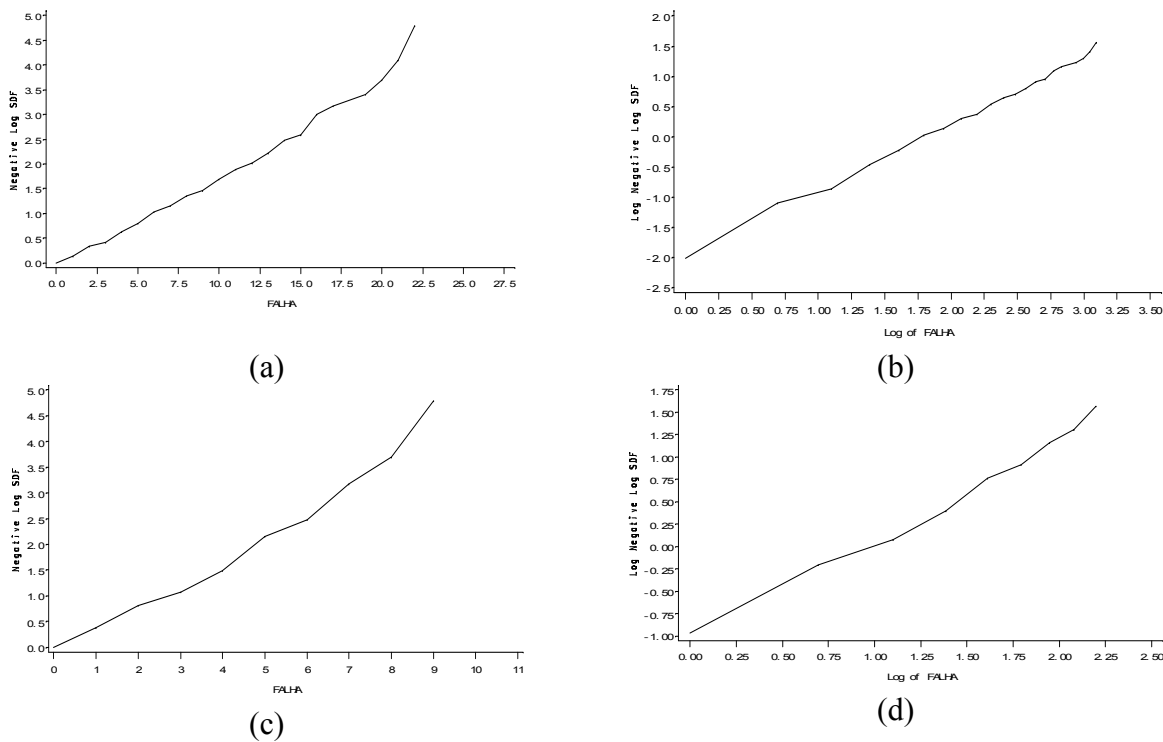


Gráfico 2 – Adequação de normal (a e c) e weibull (b e d) para $n_a=120$, $n_1=15$, $p_{di}=0.10$ (a e b), $p_{di}=0.15$ (c e d) e $p=0.15$
 Fonte: Autor.

Nos gráficos 1 e 2, foi possível perceber que o tempo até a perda de viabilidade de uma subamostra, nos dados simulados, segue o modelo probabilístico *Weibull*, pois, quanto mais o gráfico se aproximar de uma linha reta, melhor ajustados estarão os dados, este é o método dos gráficos de linearização. Nos gráficos 1 e 2 denominados por (a) e (c) tem-se na abcissa o número de falhas em cada semana e na ordenada os dados ajustados à distribuição normal e nos gráficos 1 e 2 denominados por (b) e (d) tem-se na abcissa o logaritmo do número de falhas em cada semana e na ordenada os dados ajustados à distribuição *weibull*. Os dados utilizados para o ajuste foram: Número de amostras (n_a) = 60 e 120, tamanho da amostra(n) = 2000 e 1000, respectivamente nos gráficos 1 e 2. Utilizando para cada caso a probabilidade da inviabilidade inicial(p_{di}) = 0,10 e 0,15; probabilidade da inviabilidade (p) = 0,15 e 0,15 respectivamente; tamanho da subamostra(n_1) = 3, 5, 7, 10, 15, 17, 20. Ao avaliar os dados com $n_a = 60$ amostras, observou-se que não foi possível afirmar que a distribuição normal se ajusta bem ao conjunto de dados de tempo de vida de subamostras de sementes, como o faz Ellis (1984). Porém, ao analisar os gráficos com $n_a = 120$ amostras, a tendência à distribuição *Weibull* é muito melhor que a distribuição normal, pois o gráfico da *Weibull* é o que mais se aproxima de uma reta.

3.2 Predição Pontual e Intervalar para número futuro de sementes inviáveis utilizando dados simulados

Sendo a distribuição *Weibull* a que melhor se ajusta aos dados é possível obter os valores de p , q e r que serão dados pelas seguintes expressões:

$$p = 1 - \exp\left[-\left(\frac{t_c}{\eta}\right)^3\right] \quad (10)$$

$$q = \exp\left[-\left(\frac{t_c}{\eta}\right)^3\right] - \exp\left[-\left(\frac{t_w}{\eta}\right)^3\right] \quad (11)$$

$$r = \exp\left[-\left(\frac{t_w}{\eta}\right)^3\right] \quad (12)$$

Como η e β são os parâmetros de escala e de forma, respectivamente, da distribuição *Weibull*, é preciso estimá-los para se obter os preditores pontual e intervalar, assim: β foi estimado usando o *proc lifereg* do SAS® (SAS INSTITUTE, 1990). Após estimar β , o valor de η foi estimado a partir de p conhecendo-se o estimador de máxima verossimilhança de p .

Temos,

$$\hat{\eta} = \frac{t_c}{\left[-\log\left(1 - \frac{X}{N}\right)\right]^{1/\beta}} \quad (13)$$

Estimando-se também q , por máxima verossimilhança, obtém-se:

$$\hat{q} = \exp\left[-\left(\frac{t_c}{\hat{\eta}}\right)^3\right] - \exp\left[-\left(\frac{t_w}{\hat{\eta}}\right)^3\right]$$

Substituindo (13) tem-se:

$$\hat{q} = \left[1 - \frac{x}{N}\right] - \left[1 - \frac{x}{N}\right] \left(\frac{t_w}{t_c}\right)^3 \tag{14}$$

Como foi mencionado anteriormente, segundo Nelson (2000), o preditor pontual é dado pela seguinte equação:

$$\hat{Y} = N * \hat{q}$$

Como na predição intervalar utilizou-se o procedimento da Razão de Verossimilhança temos: A função de verossimilhança multinomial irrestrita será dada por:

$$L^*(x, y) = \frac{N!}{x!y!(N-x-y)!} \left(\frac{x}{N}\right)^x \left(\frac{y}{N}\right)^y \left(1 - \frac{x}{N} - \frac{y}{N}\right)^{N-x-y} \tag{15}$$

A função de verossimilhança restrita, sob o modelo *Weibull*, será dada por:

$$K^*(x, y) = \frac{N!}{x!y!(N-x-y)!} \left[1 - e^{-\left(\frac{t_c}{\eta}\right)^\beta}\right]^x * \left[e^{-\left(\frac{t_c}{\eta}\right)^\beta} - e^{-\left(\frac{t_w}{\eta}\right)^\beta}\right]^y * \left[e^{-\left(\frac{t_w}{\eta}\right)^\beta}\right]^{N-x-y} \tag{16}$$

Sendo x um valor conhecido, denominou-se a Razão de verossimilhança por Q (x,y), então:

$$Q(x, y) = -2 \left[\log(K^*(y)) - \log(L^*(y)) \right] \leq \chi^2_{(0.95;1)} (= 3,841) \tag{17}$$

Os valores de y, que satisfizerem a inequação acima, formarão o intervalo de predição. As tabelas 1 e 2 apresentam os valores de \hat{b} , $\hat{\eta}$, \hat{q} , x, predição pontual (\hat{Y}) e predição intervalar $[Y; \tilde{Y}]$ variando no número de amostras (na), tamanho da amostra (n), probabilidade de deterioração inicial(pdi), probabilidade de deterioração(p) e tamanho da subamostra (n₁). Neste caso, x é o número de subamostra falhas em [0, t_c] semanas e y é o número de subamostras que falharão em [t_c, t_w] semanas. Considerando [0,t_c]=[0,4], isto é, as 4 primeiras semanas de armazenamento da semente(1 mês) e [t_c,t_w]=[4,16], isto é, as 12 semanas seguintes (3 meses).

Tabela 1- Valores das predições pontual (\hat{Y}) e intervalar $[Y; \tilde{Y}]$ para na = 60, n = 2000, pdi = 0.10 e p = 0.15

	X	\hat{b}	$\hat{\eta}$	\hat{q}	\hat{Y}	$[Y; \tilde{Y}]$
n ₁ =15	32	0.8	5.6	0.37	21	[18;25]

Fonte: Autor.

Tabela 2- Valores das predições pontual (\hat{Y}) e intervalar $[Y; \tilde{Y}]$ para na = 120, n = 1000, pdi = 0.15 e p = 0.15

	X	\hat{b}	$\hat{\eta}$	\hat{q}	\hat{Y}	$[Y; \tilde{Y}]$
n ₁ =15	94	0.67	2.12	0.2	23	[21;25]

Fonte: Autor.

Ao variar o valor da probabilidade de inviabilidade inicial, existe uma influência significativa no número de subamostras falhas em [0,4] semanas (x) que, conseqüentemente,

afetará \hat{Y} e $[\underline{Y}; \tilde{Y}]$. Então, conclui-se que os 5% de diferença na probabilidade de inviabilidade inicial influenciam significativamente no tempo de viabilidade da semente. Portanto, os dados obtidos nos levam a acreditar que $n = 120$ seja mais apropriada para a realização de estudos deste tipo. Utilizando a probabilidade de cobertura proposta na metodologia, é possível decidir com maior precisão ou obter a confirmação sobre qual o número de amostras mais apropriado.

3.3 Adequação do modelo (dados obtidos via experimento)

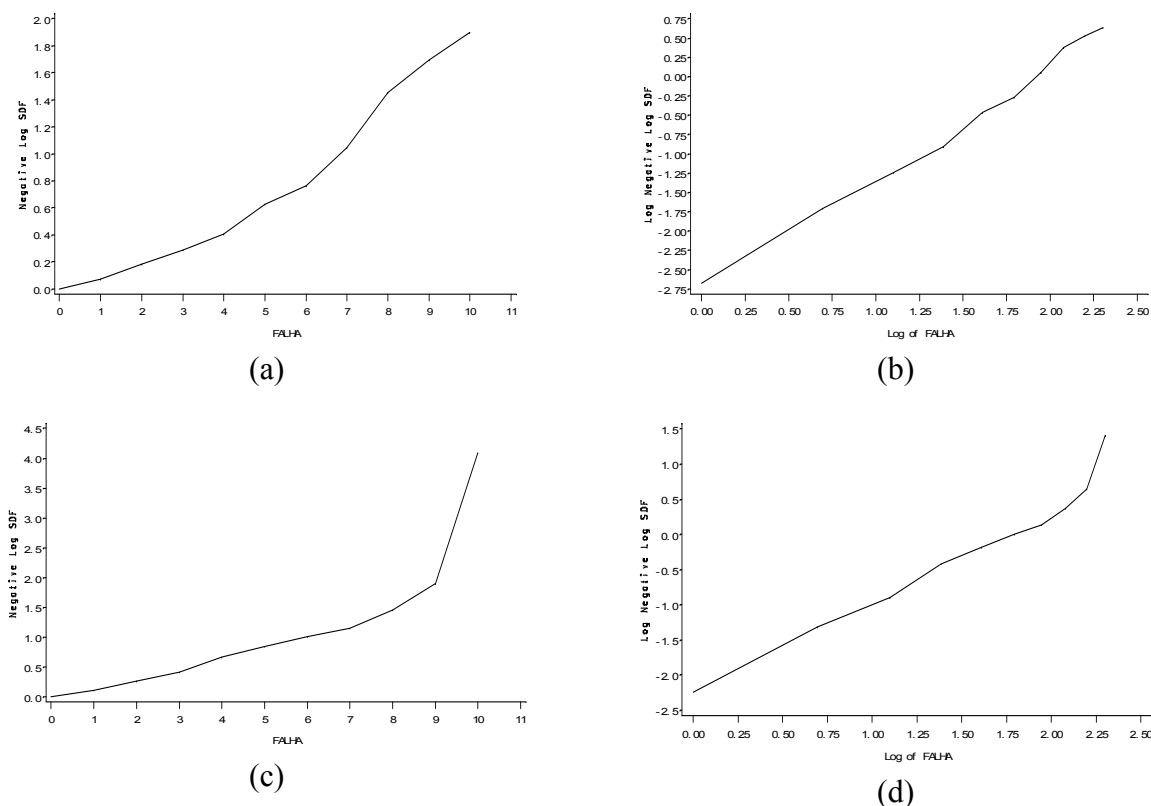


Gráfico 3 - Adequação de probabilidade normal(a) e *Weibull* (b) para Cultivar M.Soy, Gráficos de probabilidade normal(c) e weibull(d) para Cultivar Serra Negra
Fonte: Autor.

O Gráfico 3 mostra que, pelo método dos gráficos de linearização, foi possível perceber que o tempo, até a perda da viabilidade de uma subamostra dos dados experimentais, segue o modelo probabilístico *Weibull* ficando evidente que a cultivar M.Soy, com probabilidade de inviabilidade inicial 0,10, possui melhor ajuste que a cultivar Serra Negra com $p_{di} = 0,15$. Os dados utilizados para o ajuste foram: número de amostras (n_a)=60, tamanho da amostra (n)=2000, probabilidade de inviabilidade (p)=0,20, tamanho da subamostra (n_1)=10 e probabilidade de inviabilidade inicial (p_{di})=0,10 para a cultivar M.Soy e probabilidade de inviabilidade inicial (p_{di})=0,15 para a cultivar Serra Negra.

3.4 Predição Pontual e Intervalar para número futuro de sementes não viáveis utilizando dados reais

Abaixo, a Tabela 3 contendo valores de \hat{b} , $\hat{\eta}$, \hat{q} , x , predição pontual (\hat{Y}) e predição intervalar $[\underline{Y}; \tilde{Y}]$, variando: número de amostras(n_a), tamanho da amostra(n), probabilidade de inviabilidade inicial(p_{di}), probabilidade de inviabilidade(p) e tamanho da subamostra(n_1).

Em que x é o número de subamostra falhas em $[0, t_c]$ semanas e Y é o número de subamostras que falharão em $[t_c, t_w]$ semanas. Considerando neste caso do experimento real, $[0, t_c]=[0, 4]$ e $[t_c, t_w]=[4, 10]$, isto é, as 6 semanas seguintes (1 mês e meio).

Tabela 3 - Valores das predições pontual (\hat{Y}) e intervalar $[\underline{Y}; \tilde{Y}]$ das cultivares M.Soy e Serra Negra para $n_a = 60$, $n = 2000$, $p_{di} = 0.10$ e $p = 0.20$ e $n_a = 60$, $n = 2000$, $p_{di} = 0.15$ e $p = 0.20$ respectivamente.

	x	\hat{b}	$\hat{\eta}$	\hat{q}	\hat{Y}	$[\underline{Y}; \tilde{Y}]$
$n_1=10$ M.Soy	16	0.55	33.61	0.13	8	[4;13]
$n_1=10$ Serra Negra	21	0.56	18	0.16	10	[5;15]

Fonte: Autor.

Pode-se observar que as sementes se tornam não viáveis mais rapidamente na cultivar M Soy.

3.5 Determinação da Probabilidade de Cobertura do Intervalo de Predição Bilateral

Por meio de simulações utilizando um programa desenvolvido em linguagem C, foi possível obter os valores das probabilidades de cobertura condicionais para o intervalo de predição $[\underline{Y}; \tilde{Y}]$ do número de subamostras falhas no intervalo de tempo futuro $[4, 16]$ semanas, para um nível de confiança nominal de 95%. Com os resultados a probabilidade de cobertura não condicional será:

$$E_x\{PC[IP(95\%)] | x\} = 0.9548$$

Observou-se que a probabilidade de cobertura não condicional é muito próxima do valor nominal, isto é, o intervalo de predição $[\underline{Y}; \tilde{Y}]$ para a quantidade aleatória $Y =$ número de subamostras falhas no intervalo de tempo $[t_c, t_w]$, demonstra ser confiável.

4 CONCLUSÃO

É possível conhecer a distribuição dos tempos de vida quando se trabalha com ensaios destrutivos (caso das sementes). A avaliação de sementes inviáveis determinadas no teste de tetrazólio na predição do número futuro de subamostras de sementes falhas, permitirá a previsão do comportamento de sementes de soja durante o armazenamento em relação a qualidade fisiológica dos lotes. O delineamento amostral envolvendo subamostras falhas apresentou-se como uma alternativa viável quando não é possível realizar a medição da variável de estudo nas próprias unidades amostrais. A predição intervalar do número de subamostras falhas no intervalo de tempo $[t_c, t_w]$ mostrou-se confiável, conforme demonstra a análise da probabilidade de cobertura. No estudo via simulação, encontrou-se que a Distribuição *Weibull* pode ser uma boa

alternativa para ajustá-la à variável tempo, até a perda da viabilidade das sementes do lote. No estudo com dados reais, ao comparar os valores das estimativas pontual e intervalar do número de subamostras de sementes falhas, estas se mostraram aceitáveis.

REFERÊNCIAS

ELLIS, R. H. The meaning of viability. In: DICRIE, J. B; LININGTON, S. H; WILLIAM, J. T. (Ed.). **Seed management technique for genebanks**. Rome: International Boards for Plant Genetic Resources, 1984. p. 146-178.

ESCOBAR, L. A; MEEKER, W, Q. Statistical prediction based on censored life data. **Technometrics**, v. 41, n. 2, p. 113-124, May 1999.

FRANÇA NETO, J. B; KRZYZANOWSKI, F. C; COSTA, N. P. da. **O teste de tetrazólio em sementes de soja**. Londrina: EMBRAPA-CNPSO, 1998.

NELSON, W. Statistical methods for the ratio of two multinomial proportions. **The American Statistician**, v. 26, n. 3, p. 22-27, 1972.

_____. Weibull analysis of reliability data with few or no failures. **Journal of Quality Technology**, v. 17, p. 140-146, Jul. 1985.

_____. Weibull prediction of a future number of failures. **Quality and Reliability Engineering International**, v. 16, p. 23-26, Sept. 2000.

_____; HAHN, G. J. A survey of prediction intervals and their applications. **Journal of Quality Technology**, v. 5, p. 178-188, Oct. 1973.

NORDMAN, D. J; MEEKER, W. Q. Weibull prediction intervals for a future number of failures. **Technometrics**, v. 44, p. 15-23, 2002.

SAS[®] INSTITUTE. **SAS procedures guide for computer**. 4. ed. Cary, NC: SAS Institute Inc., 1990.

Recebido em: 17/06 /2015

Aprovado em: 18/08/2015

Publicado em: 15/01/2016

ANEXO A – Procedimento para obtenção dos dados via simulação

