

**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
DEPARTAMENTO ACADÊMICO DE ELETROTÉCNICA
ESPECIALIZAÇÃO EM ENGENHARIA DA CONFIABILIDADE**

VAGNER DE ALMEIDA CONSELVAN

**ANÁLISE DE CONFIABILIDADE APLICADA NA OTIMIZAÇÃO DE
SISTEMAS DE PRODUÇÃO ANIMAL**

MONOGRAFIA DE ESPECIALIZAÇÃO

CURITIBA

2017

VAGNER DE ALMEIDA CONSELVAN

**ANÁLISE DE CONFIABILIDADE APLICADA NA OTIMIZAÇÃO DE
SISTEMAS DE PRODUÇÃO ANIMAL**

Monografia apresentada como requisito parcial à obtenção do título de Especialista em Engenharia da Confiabilidade, do Departamento Acadêmico de Eletrotécnica, da Universidade Tecnológica Federal do Paraná.

Orientador: Prof. Dr. Emerson Rigoni

CURITIBA

2017



TERMO DE APROVAÇÃO

ANÁLISE DE CONFIABILIDADE APLICADA NA OTIMIZAÇÃO DE SISTEMAS DE PRODUÇÃO ANIMAL

por

Vagner de Almeida Conselvan

Esta monografia foi apresentada em 03 de 11 de 2017, como requisito parcial para obtenção do título de Especialista em Engenharia da Confiabilidade, outorgado pela Universidade Tecnológica Federal do Paraná. O aluno foi arguido pela Banca Examinadora composta pelos professores abaixo assinados. Após deliberação, a Banca Examinadora considerou o trabalho aprovado.

Prof. Emerson Rigoni, Dr. Eng.
Professor Orientador - UTFPR

Prof. Carlos Henrique Mariano Dr.
Membro Titular da Banca - UTFPR

Prof. Marcelo Rodrigues Dr.
Membro Titular da Banca - UTFPR

O Termo de Aprovação assinado encontra-se na Coordenação do Curso.

Dedico este trabalho à minha família, por sua capacidade de acreditar em mim e investir em mim. Também à minha noiva que não mediu esforços para que eu chegasse até essa etapa de minha vida.

AGRADECIMENTOS

Certamente estes parágrafos não irão atender a todas as pessoas que fizeram parte dessa importante fase de minha vida. Portanto, desde já peço desculpas àquelas que não estão presentes entre essas palavras, mas elas podem estar certas que fazem parte do meu pensamento e de minha gratidão.

Agradeço ao meu orientador Prof. Dr. Emerson Rigoni, pela sabedoria com que me guiou nesta trajetória.

Aos meus colegas de sala.

A Secretaria do Curso, pela cooperação.

Gostaria de deixar registrado também, o meu reconhecimento à minha família e noiva, pois acredito que sem o apoio deles seria muito difícil vencer esse desafio.

Enfim, a todos os que por algum motivo contribuíram para a realização desta pesquisa.

“Embora isso possa parecer um paradoxo, toda ciência exata é dominada pela ideia de aproximação.”

Bertrand Russel
1872-1970

RESUMO

CONSELVAN, de Almeida. Análise de Confiabilidade Aplicada na Otimização de Sistemas de Produção Animal. 2017. 45 folhas. Monografia (Especialização em Engenharia da Confiabilidade) - Universidade Tecnológica Federal do Paraná. Curitiba, 2017.

Ainda que a Engenharia de Confiabilidade se concentre no estudo e discussões sobre a vida de objetos inanimados, como equipamentos, componentes, itens, subitens ou sistemas, esses mesmos conceitos podem ser aplicados em outras áreas, tal como análise de sobrevivência na otimização de sistemas de produção animal – objeto de estudo deste trabalho em que a variável resposta, decorrente de um experimento fatorial 2×3 , descreve o número de dias transcorridos entre o nascimento e a primeira postura, evento de interesse, de codornas das linhagens amarela, azul e vermelha e submetidas a dois níveis de dieta alimentar. Buscou-se selecionar o melhor modelo probabilístico que descreve de forma razoável o comportamento da idade, em dias, até a primeira postura, bem como estimar pontualmente e por intervalo, o tempo médio da primeira postura e, por fim, avaliar o efeito do tratamento dieta nas linhagens. Os procedimentos analíticos e gráficos, usados para avaliação do melhor ajuste, mostram que a distribuição Gama Generalizada, na maioria dos grupos, é a mais adequada para descrever a variável resposta de interesse. De posse de um modelo probabilístico capaz de descrever de forma eficaz o fenômeno de interesse, tem-se implicações significativas no processo decisório com relação ao manejo animal e conseqüentemente na otimização do sistema de produção. Nessa modelagem, a linhagem amarela apresentou menor tempo de postura em torno de 40 dias e a dieta 1, independente da linhagem, proporcionou menor tempo de postura para todos os tratamentos.

Palavras-chave: Confiabilidade. Distribuição de probabilidade. Verossimilhança. Seleção de modelos. Produção animal.

ABSTRACT

CONSELVAN, de Almeida. Análise de Confiabilidade Aplicada na Otimização de Sistemas de Produção Animal. 2017. 45 folhas. Monografia (Especialização em Engenharia da Confiabilidade) - Universidade Tecnológica Federal do Paraná. Curitiba, 2017.

Although Reliability Engineering focuses on the study and discussion of the life of inanimate objects, such as equipment, components, items, sub-items or systems, these same concepts can be applied in other areas, such as survival analysis in the optimization of animal production - object of study of this work in which the target variable, resulting from a 2 x 3 factorial experiment, describes the number of days elapsed between birth and the first posture, event of interest, from yellow, blue and red quails and submitted to two levels of diet. It was attempted to select the best probabilistic model that describes in a reasonable way the behavior of the age, in days, until the first posture, as well as to estimate punctually and by interval, the average time of the first posture and, finally, to evaluate the effect of the treatment diet in the bloodlines. The analytical and graphical procedures, used to evaluate the best fit, show that the Generalized Gama distribution in most groups is the most adequate to describe the target variable of interest. With a probabilistic model capable of describing effectively the phenomenon of interest, there are significant implications in the decision process regarding animal management and consequently in the optimization of the production system. In this modeling, the yellow bloodline presented a shorter posture time around 40 days and the diet 1, independent of the lineage, provided a shorter posture time for all treatments.

Palavras-chave: Reliability. Probability distributions. Likelihood function. Models selection. Animal production.

LISTA DE ILUSTRAÇÕES

Figura 1 - Forma típica da função taxa de falha para $\beta < 1$, $\beta = 1$ e $\beta > 1$, segundo modelo Weibull.....	20
Figura 2 - Representação da função taxa de falha conhecida por curva da banheira	21
Figura 3 - Frequência relativa das distribuições do número de codornas segundo dieta e linhagem (representatividade por dieta e linhagem).....	32
Figura 4 - Frequência absoluta das distribuições do número de codornas segundo dieta e linhagem	32
Figura 5 - Frequência absoluta das distribuições do número de codornas segundo dieta e linhagem	33
Figura 6 - Função de sobrevivência empírica e distribuição teórica Gama Generalizada ajustadas a linhagem Amarela segundo dietas 1 e 2.....	36
Figura 7 - Função de sobrevivência empírica e distribuição teórica Gama Generalizada e Weibull 3 parâmetros ajustadas a linhagem Vermelha segundo dietas 1 e 2.....	37
Figura 8 - Função de sobrevivência empírica e distribuição teórica Gama Generalizada e Weibull 3 parâmetros ajustadas a linhagem Azul segundo dietas 1 e 2	38
Figura 9 - Curvas da probabilidade da primeira postura segunda distribuição Gama Generalizada para os alinhamentos e dietas	39
Figura 10 - Estimativas pontuais e por intervalo (coeficiente de confiança de 95%) da idade média da primeira postura, em dias, para cada um dos tratamentos segundo a distribuição selecionada	40

LISTA DE TABELAS

Tabela 1 - Distribuição do número de codornas na combinação dieta consoante linhagem.....	31
Tabela 2 - Estatísticas descritivas dos tempos (dias) de postura das codornas, por linhagem e dieta	33
Tabela 3 - Análise de variância para os tempos até primeira postura segundo a combinação linhagem e dieta.....	34
Tabela 4 - Valores das estatísticas e testes de discriminação segundo linhagem amarela e dietas 1 e 2.....	35
Tabela 5 - Valores das estatísticas e testes de discriminação segundo linhagem vermelha e dietas 1 e 2.....	35
Tabela 6 - Valores das estatísticas e testes de discriminação segundo linhagem azul e dietas 1 e 2.....	35
Tabela 7 - Estimativas de máxima verossimilhança do tempo médio até a primeira postura e da probabilidade de ocorrência da primeira postura ao tempo médio	39

LISTA DE SIGLAS

IBGE	Instituto Brasileiro de Geografia e Estatística
ADV	Análise de Dados de Vida
FEI	Fazenda Experimental de Iguatemi
MF	Matriz de Informação de Fisher
MMV	Método da Máxima Verossimilhança
MTTF	Tempo Médio até Primeira Falha
UEM	Universidade Estadual de Maringá

SUMÁRIO

1 INTRODUÇÃO	13
1.1 PREMISSAS E PROBLEMA DE PESQUISA	14
1.2 OBJETIVOS	15
1.2.1 Objetivo Geral	15
1.2.2 Objetivos Específicos	15
1.3 JUSTIFICATIVA	16
1.4 PROCEDIMENTOS METODOLÓGICOS	16
1.5 ESTRUTURA DO TRABALHO	17
2 REFERENCIAL TEÓRICO.....	18
2.1 ANÁLISE DE DADOS DE VIDA	18
2.2 ESPECIFICANDO O TEMPO DE CONFIABILIDADE / SOBREVIVÊNCIA.....	19
2.2.1 Função de Sobrevivência	19
2.2.2 Função de Taxa de Falha.....	19
2.2.3 Tempo Médio de Vida	21
2.3 MODELOS PROBABILÍSTICOS - PARAMÉTRICOS	22
2.3.1 Distribuição Weibull	22
2.3.2 Distribuição Log-normal.....	23
2.3.3 Distribuição Log-logística	24
2.3.4 Distribuição Gama	25
2.3.5 Estimação dos Parâmetros dos Modelos	26
2.3.5.1 Método da Máxima Verossimilhança.....	26
2.4 TÉCNICA NÃO-PARAMÉTRICA: KAPLAN-MEIER	27
2.5 ESCOLHA DO MODELO	28
2.5.1 Critérios de Discriminação.....	28
2.5.1.1 Métodos Gráficos	29
2.5.1.2 Estatísticas de Qualidade do Ajuste	29
3 RESULTADOS E DISCUSSÃO	31
3.1 ANÁLISE EXPLORATÓRIA DOS DADOS	31
3.2 AJUSTE DOS MODELOS	34
3.3 DISTRIBUIÇÃO GAMA GENERALIZADA.....	38
4 CONCLUSÃO	41
4.1 TRABALHOS FUTUROS.....	41
REFERÊNCIAS.....	42

1 INTRODUÇÃO

No Brasil, o efetivo de codorna em 2015 foi de aproximadamente 22 milhões de aves (IBGE, 2015), apresentando um aumento de 8,2% com relação ao registrado em 2014. A Região Sudeste destaca-se como a maior produtora nacional de codornas, independentemente da finalidade, seja para produção de carne ou de ovos.

Existem vários fatores que têm contribuído para o aumento da criação de codornas no país, entre eles se destacam: o rápido crescimento, a precocidade na produção, a maturidade sexual (35 a 42 dias), a alta produtividade (média de 300 ovos/ano), pequenos espaços para grandes populações, a grande longevidade em alta produção (14 a 18 meses), o baixo investimento e, conseqüentemente, o rápido retorno financeiro (PINTO et al., 2002). Ainda, segundo Silva et al. (2009), a codorna é uma excelente alternativa para alimentação humana, pois pode ser utilizada tanto para a produção de ovos como para a produção de carne, que é aceita universalmente por ser um produto de excelente qualidade e rica em aminoácidos essenciais, apresentando, também, alto conteúdo proteico e baixa quantidade de gordura.

Consoante esse cenário, observa-se que o crescimento do setor coturnicultura no Brasil é significativo. No entanto, o material genético disponível para produção de codornas, tanto para a produção de ovos como para a produção de carne no Brasil, é de baixa qualidade, necessitando importar matrizes a fim de obter melhor peso e conformação de carcaça no caso de codornas de corte, ideais para o atendimento dos mercados internos e de exportação (GARCIA, 2002). Para se obter materiais genéticos de qualidade, são necessários programas de melhoramento genéticos bem fundamentados, embasados em parâmetros acurados e precisos (DIONELLO et al., 2008).

Logo, percebe-se que pouco se conhece sobre o potencial produtivo de codornas de corte e postura no Brasil (MÓRI et al., 2005). Para melhorar os índices produtivos e minimizar a desuniformidade dos lotes, um dos principais problemas da criação, é preciso melhorar e muito o manejo, a genética e a nutrição das codornas (GRIESER, 2012).

1.1 PREMISSAS E PROBLEMA DE PESQUISA

Análise de dados de vida - ADV - em engenharia de confiabilidade refere-se ao estudo e modelagem do ciclo de vida de produtos. Os dados de vida podem ser a vida útil dos produtos no mercado, tal como, tempo em que o produto funcionou com sucesso ou tempo em que o produto operou antes de falhar. Esses tempos podem ser medidos em horas, quilômetros, ciclos de falha, ciclos de estresse ou alguma outra métrica que mensure a vida ou a exposição do produto.

Ainda que a Engenharia de Confiabilidade se concentre no estudo e discussões sobre a vida de objetos inanimados, como equipamentos, componentes, itens, subitens ou sistemas, os mesmos conceitos aqui estudados podem ser aplicados em outras áreas, tal como análise de sobrevivência na otimização de sistemas de produção animal – objeto de estudo desse trabalho.

Apesar de o Brasil estar entre os três maiores produtores mundiais de codornas (MINVIELLE, 2004), são poucas as instituições que trabalham com essa espécie animal, e principalmente com avaliações genéticas. Segundo Martins (2002), os programas de melhoramento genético devem ser realizados da forma mais eficiente, utilizando metodologias modernas e adequadas, para que a seleção resulte em vantagens. No caso da produção de ovos, a eficiência do processo de seleção é em parte devida à escolha adequada da função que descreve o comportamento da postura.

Dessa forma, a utilização de modelos que reflitam de modo parcimonioso a realidade biológica, grande parte dos efeitos naturais, é de grande importância na produção animal. Na coturnicultura, tratando-se de codornas criadas para a produção de ovos, ferramentas intrínsecas as áreas de análise de sobrevivência/confiabilidade podem ser utilizadas na estimação de parâmetros, associados com o tempo até a primeira postura. Uma vez selecionada uma distribuição de probabilidade, medidas sumárias como idade média e a idade mais provável da primeira postura, importantes no processo de seleção de linhagens e/ou dietas mais efetivas, podem ser estimadas (LEE e WANG, 2003).

Uma inquirição importante e usual na análise de dados de vida, refere-se a escolha da distribuição de probabilidade mais apropriada, que apresente melhor ajuste aos dados. Nos últimos anos foram propostas uma infinidade de novas distribuições para análise de dados relacionados ao tempo até a ocorrência de

algum evento de interesse. Contudo, distribuições de probabilidade tradicionais, como Gama, Log-Normal, Log-Logística e Weibull ainda se mostram muito úteis (XIE et al., 2002).

Nesse contexto, torna-se evidente a importância de conduzir experimentos científicos a fim de fornecer informações aprofundadas sobre a postura das codornas de diferentes linhagens genéticas na presença de diferentes dietas para obter informações condizentes com as características genéticas atuais, contribuindo com os sistemas de produção animal.

1.2 OBJETIVOS

Nesse tópico apresentam-se os objetivos geral e específicos do presente trabalho.

1.2.1 Objetivo Geral

Tem-se como objetivo geral a seleção de uma distribuição de probabilidade que descreva de forma razoável o comportamento da idade em dias até a primeira postura – evento – de codornas das linhagens amarela, azul e vermelha submetidas a dois níveis de dieta alimentar constituindo um planejamento de experimento fatorial 2×3 que descreve o número de dias transcorridos entre o nascimento e a primeira postura.

1.2.2 Objetivos Específicos

Para atingir o principal objetivo desse trabalho, faz-se necessário percorrer as seguintes etapas:

- Estudar o banco de dados resultando do delineamento amostral do experimento;
- Revisão bibliográfica sobre modelos de sobrevivência/confiabilidade;
- Estudar as ferramentas computacionais utilizadas no presente trabalho;
- Realizar análise exploratória/descritiva dos dados no ambiente *R*;

- Avaliar os procedimentos analíticos e gráficos usados como ferramentas de discriminação para seleção de modelos;
- Selecionar a melhor distribuição via programa *Reliasoft Weibull++* e avaliar a qualidade do ajuste;
- Interpretar os parâmetros do modelo e obter medidas sumárias de interesse;
- Comparar a eficiência de postura entre as diferentes linhagens na presença de dois níveis de dieta alimentar.

1.3 JUSTIFICATIVA

Uma vez adotada uma distribuição de probabilidade, parâmetros de interesse como a idade média e a idade mais provável da primeira postura podem ser estimados. A partir dessas medidas e contrastes gráficos, pode-se comparar a eficiência de postura das codornas para os diferentes fatores do delineamento experimental auxiliando no processo de seleção de linhagens e/ou dietas mais efetivas. Desse modo, contribui-se com os programas de melhoramento genético aumentando os índices produtivos e minimizando a dissemelhança dos lotes.

1.4 PROCEDIMENTOS METODOLÓGICOS

Os dados utilizados nessa análise referem-se ao tempo em dias desde o nascimento até a primeira postura de codornas (*Coturnix coturnix japonica*) provenientes de três linhagens e submetidas a duas dietas alimentares. O experimento foi realizado no setor de coturnicultura da Fazenda Experimental de Iguatemi (FEI), pertencente à Universidade Estadual de Maringá (UEM), Paraná, Brasil, no período de agosto 2006 a janeiro 2007 com a utilização de 308, 374 e 378 codornas de postura das linhagens amarela, azul e vermelhas (cor das anilhas), respectivamente.

Inicialmente, avaliou-se no banco de dados as marcações de postura durante o período experimental (dia a dia) a fim de calcular o tempo, em dias, até a primeira postura, obtendo-se, assim a variável resposta em questão.

De posse dos dados e da marcação de performance da variável resposta, realiza-se uma análise exploratória dos dados no programa R (*software* livre, trata-se

de uma linguagem e ambiente para análise estatística) considerada de extrema importância, já que de forma simples, proporciona métodos eficazes para resumo e análise da informação. Também, auxilia encontrar problemas – inconsistências – nos dados e possibilita antever possíveis padrões/tendências a serem investigados numa análise mais robusta posterior.

Com o programa *Reliasoft Weibull++*, realiza-se a análise ADV, ajuste e diagnóstico de modelos, segundo uma gama de modelos e conjunto de critérios de seleção já implementados no *software*. Considerou-se a estimação via Método da Máxima Verossimilhança (MMV) para obtenção do vetor paramétrico e os intervalos de confiança para os mesmos foram obtidos considerando a normalidade assintótica dos estimadores de máxima verossimilhança, via matriz de informação de Fisher (MF).

Por fim, consoante modelo adotado, calcula-se as métricas de confiabilidade, interpreta-se os parâmetros e, também, avalia-se os contrastes entre os diferentes níveis de tratamento do experimento, indicando o nível mais eficiente.

1.5 ESTRUTURA DO TRABALHO

O presente trabalho está estruturado em 4 capítulos, sendo esse o primeiro capítulo em que se faz uma introdução ao estudo, apresenta-se o tema, objetivos gerais e específicos, justificativa do trabalho e, também, os procedimentos metodológicos utilizados ao longo da análise.

No capítulo 2 é apresentado a fundamentação teórica necessária para o desenvolvimento do trabalho no que diz respeito ao processo de modelagem do tempo até ocorrência do evento de interesse. Apresentação das possíveis distribuições de probabilidade adequadas ao estudo, método de estimação dos parâmetros da curva, bem como os critérios de seleção e ajuste do modelo.

No capítulo 3 tem-se a apresentação e discussão dos resultados da análise exploratória dos dados e da aplicação da metodologia de análise dos dados de vida.

Finalmente, no capítulo 4 apresenta-se as conclusões e proposições para trabalhos futuros.

2 REFERENCIAL TEÓRICO

Nesse trabalho, para descrever o tempo até a primeira postura de codornas num delineamento fatorial 2 x 3 utilizou-se das ferramentas práticas da Engenharia de Confiabilidade cujo embasamento teórico deriva da Análise de Sobrevivência.

2.1 ANÁLISE DE DADOS DE VIDA

Confiabilidade em análise de dados de vida refere-se ao estudo e modelagem da vida de produtos. Dados de vida podem ser, por exemplo, o tempo de vida de um determinado produto no mercado, isto é, tempo até ocorrência do evento de interesse, no caso de um produto, tempo até a sua falha. Assim os tempos de falha de um produto podem ser abrangidos ao termo dados de vida, ou mais especificamente, dados de vida do produto. A análise e predição subsequente desses dados são descritas como Análise de Dados de Vida.

Na Engenharia de Confiabilidade os estudos, as discussões e aplicações restringem-se a tempos de vida de objetos inanimados, logo se faz uso da teoria de Análise de Sobrevivência estendendo-a à equipamentos, componentes e sistemas. Dessa forma todo o instrumental analítico desenvolvido pela engenharia de confiabilidade para aplicação industrial também se pode aplicar na análise de tempos de vida de seres animados, havendo, portanto, uma contribuição mútua entre as áreas.

Ao se realizar uma análise de dados de vida, o objetivo é fazer predições sobre os tempos de vida de uma determinada população por meio do ajuste de uma distribuição estatística – modelo – aos dados de vida de uma amostra representativa da população de interesse. Em geral, o processo de análise de dados de vida pode ser dividido em cinco etapas:

1. Obter os dados de vida do evento de interesse;
2. Selecionar as potenciais distribuições de probabilidade para o ajuste do modelo;
3. Estimar os parâmetros que irão ajustar as distribuições aos dados de vida;
4. Selecionar a distribuição – modelo – mais aderente aos dados;
5. Gerar gráficos e medidas de interesse, como taxa de falha e tempo médio de vida.

2.2 ESPECIFICANDO O TEMPO DE CONFIABILIDADE / SOBREVIVÊNCIA

A variável aleatória não-negativa T , usualmente contínua, que representa o tempo de falha, é geralmente especificada em análise de sobrevivência pela sua função de sobrevivência ou pela função de taxa de falha. Essas duas funções, e funções relacionadas, que são extensivamente usadas na análise de dados de vida são apresentadas a seguir (COLOSIMO; GIOLO, 2006).

2.2.1 Função de Sobrevivência

Principal função probabilística usada para descrever estudos de análise de dados de vida. Essa função é definida como a probabilidade de uma observação não falhar até um certo tempo t , ou seja, a probabilidade de uma observação sobreviver ao tempo t . Em termos probabilísticos escreve-se como:

$$S(t) = P(T \geq t).$$

Em consequência, a função de distribuição acumulada é definida como a probabilidade de uma observação não sobreviver ao tempo t , isto é:

$$F(t) = 1 - S(t).$$

2.2.2 Função de Taxa de Falha

Na literatura de análise de sobrevivência a função de risco também é conhecida por função de taxa de falha, definida como sendo a taxa de ocorrência de um evento por unidade de tempo (Lawless, 2003). Ela descreve a forma em que a taxa instantânea de falha muda com o tempo. Essa função em T é definida como:

$$\lambda(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t}$$

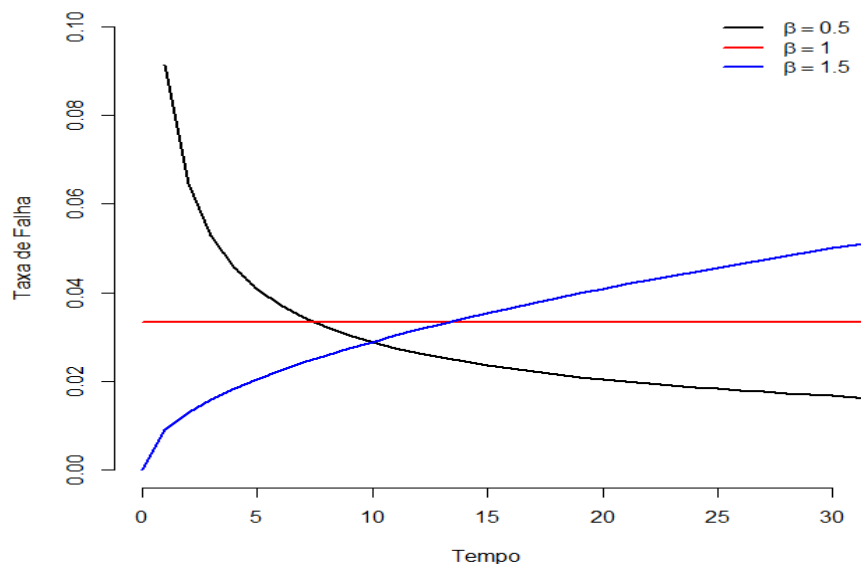
Em uma distribuição com taxa de falha unimodal o número de eventos por unidade de tempo cresce até atingir um ponto de máximo e após esse ponto de máximo ser atingido essa taxa tende assintoticamente para zero ou algum outro valor à medida que o tempo tende ao infinito. Na prática, muitas vezes, tem-se o

interesse em estimar esse ponto de máximo, também conhecido como ponto de mudança. A Figura 1 mostra três funções de taxa de falha segundo o modelo *Weibull* com a seguinte parametrização:

$$f(t) = \frac{\beta}{\eta^\beta} t^{\beta-1} e^{-\frac{t^\beta}{\eta}}, \quad t \geq 0$$

em que β , representa o parâmetro de forma e η o parâmetro de escala, ambos positivos.

Figura 1 - Forma típica da função taxa de falha para $\beta < 1$, $\beta = 1$ e $\beta > 1$, segundo modelo *Weibull*



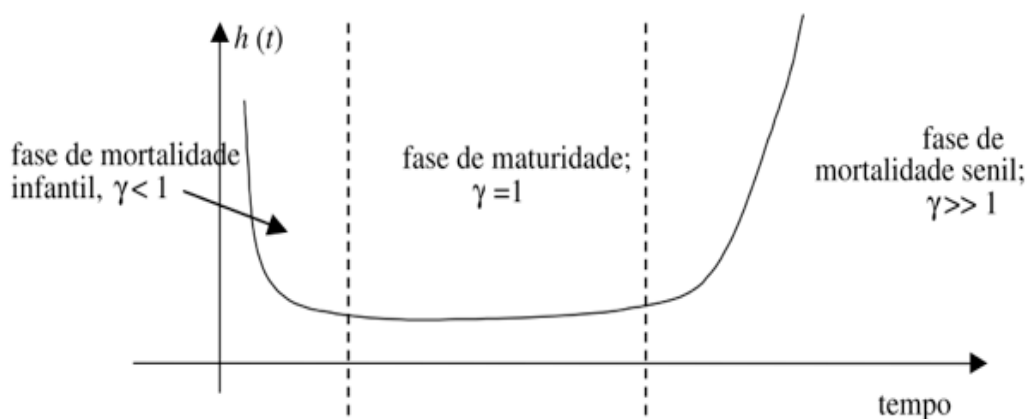
Fonte: autor (2017).

Para essa distribuição, tem-se que a função taxa de falha é estritamente crescente para $\beta > 1$, estritamente decrescente para $\beta < 1$ e constante para $\beta = 1$. Essas diferentes formas podem revelar indícios sobre o mecanismo de falha.

- Se $\beta < 1$, há indícios de falhas prematuras que podem ser ocasionadas por problemas no controle de qualidade e montagem;
- Se $\beta = 1$, tem-se falhas que ocorrem de modo ocasional, decorrentes de solicitações normais de uso, diferentes combinações de uso, acidentes causados por uso incorreto e manutenção inadequada;
- $\beta > 1$, representa falhas devido ao processo natural de envelhecimento ou desgaste do produto.

Sabe-se, ainda, que a taxa de falha para o tempo de vida de seres humanos é uma combinação das curvas apresentadas na Figura 1 em diferentes períodos de tempo. Ela é conhecida como *curva da banheira* e tem uma taxa de falha decrescente no período inicial, representando a mortalidade infantil, constante na faixa intermediária e crescente na porção final (COLOSIMO; GIOLO, 2006). Uma representação dessa curva é mostrada na Figura 2.

Figura 2 - Representação da função taxa de falha conhecida por curva da banheira



Fonte: autor (2017).

É importante observar que a função taxa de falha é mais informativa do que a função de sobrevivência. Diferentes funções de sobrevivência podem ter formas semelhantes, enquanto as respectivas funções taxas de falhas podem diferir drasticamente. Desta forma, a modelagem da função de taxa de falha é um importante método para dados de sobrevivência.

2.2.3 Tempo Médio de Vida

Outra quantidade de interesse na análise de dados de vida é: o tempo média de vida, obtido pela área sob a função de sobrevivência ou confiabilidade. Isto é:

$$E(T) = t_m = \int_0^{\infty} S(t) dt.$$

Pela propriedade de invariância dos estimadores de máxima verossimilhança estima-se $E(T)$, que descreve, nesse estudo, a idade média (em dias) da primeira postura.

2.3 MODELOS PROBABILÍSTICOS - PARAMÉTRICOS

A função de densidade de probabilidade (pdf) e a função distribuição acumulada (cdf) são as duas funções estatísticas mais importantes em confiabilidade e estão estreitamente relacionadas. Quando uma dessas funções é conhecida, quase todas as outras medidas de confiabilidade podem ser obtidas ou derivadas a partir da mesma. A função de confiabilidade, função de taxa de falha, função do tempo médio, função do tempo mediano podem todas serem obtidas diretamente da *pdf* ou $f(t)$.

Nesse trabalho para descrever o tempo até a primeira postura de codornas, avaliou-se as potenciais distribuições de probabilidade: Weibull, Log-Normal, Log-Logística e Gama. Tais distribuições, denominadas modelos probabilísticos ou paramétricos, têm se mostrado bastante adequadas para descrever, em particular, os tempos de vida de produtos industriais além da aplicação em seres animados.

2.3.1 Distribuição Weibull

A ampla aplicabilidade dessa distribuição se deve ao fato dela apresentar uma grande variedade de formas e uma função de taxa de falhas monótona; ou seja, uma função que é sempre crescente, decrescente ou constante. Além disso, ela também descreve adequadamente o tempo de vida de equipamentos que são formados por várias partes e onde a falha ocorre quando a primeira dessas partes falhar (FREITAS; COLOSIMO, 1997). Para uma variável aleatória T com distribuição Weibull, tem-se a seguinte função de densidade de probabilidade:

$$f(t) = \frac{\beta}{\eta^\beta} t^{\beta-1} e^{-\frac{t^\beta}{\eta}}, \quad t \geq 0$$

em que β , representa o parâmetro de forma e η o parâmetro de escala, ambos positivos.

Para essa distribuição, as funções de confiabilidade ou sobrevivência e taxa de falha são, respectivamente:

$$S(t) = \exp \left\{ - \left(\frac{t}{\eta} \right)^\beta \right\}$$

e

$$\lambda(t) = \frac{\beta}{\eta^\beta} t^{\beta-1}$$

para $t \geq 0$, β e $\eta > 0$.

É importante observar que para $\beta = 1$, tem-se a distribuição exponencial e, sendo assim, a distribuição exponencial é um caso particular da distribuição Weibull. As expressões para média e variância da Weibull incluem o uso da função gama, isto é,

$$\mu = E[T] = \alpha \Gamma[1 + (1/\alpha)]$$

$$Var[T] = \alpha^2 [\Gamma[1 + (2/\alpha)] - \Gamma[1 + (1/\alpha)]^2]$$

2.3.2 Distribuição Log-normal

A distribuição log-normal, assim como a de Weibull, é muito utilizada para descrever o tempo de vida de produtos e indivíduos. Essa distribuição é a que melhor descreve tempos de vida de mecanismos que envolvem reações químicas como processos de corrosão, acúmulo superficial de cargas elétricas, degradação de contatos. Também é bastante utilizada para descrever situações clínicas, como o tempo de vida de pacientes com leucemia (FREITAS; COLOSIMO, 1997). A função densidade de probabilidade de uma variável aleatória com distribuição log-normal é dada por:

$$f(t) = \frac{1}{\sqrt{2\pi}t\sigma} \exp \left\{ -\frac{1}{2} \left(\frac{\log(t) - \mu}{\sigma} \right)^2 \right\}, \quad t > 0$$

em que μ é média do logaritmo do tempo de falha, assim como σ é o desvio-padrão.

Como o próprio nome sugere, o logaritmo de uma variável aleatória com distribuição log-normal de parâmetros μ e σ tem distribuição normal com média μ e desvio-padrão σ . Essa relação significa que dados de uma distribuição log-normal podem ser analisados segundo uma distribuição normal, desde de que se conheça o

logaritmo dos dados em vez dos valores originais. As funções de confiabilidade e taxa de falha dessa distribuição são dados, respectivamente, por:

$$R(t) = \Phi \left(\frac{-\log(t) + \mu}{\sigma} \right)$$

e

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

em que $\Phi(\cdot)$ é a função de distribuição acumulada de uma normal padrão. A média e a variância de uma variável com distribuição log-normal são dadas, respectivamente, por:

$$\mu = \exp \left\{ \mu + \frac{\sigma^2}{2} \right\}$$

$$Var[T] = \exp\{2\mu + \sigma^2\} (\exp\{\sigma^2\} - 1).$$

2.3.3 Distribuição Log-logística

Uma distribuição que, em muitas situações práticas, tem se apresentado como uma alternativa à Weibull e à Log-normal. Para uma variável aleatória T com essa distribuição, a função de densidade é expressa por:

$$f(x) = \frac{\beta}{\eta^\beta} t^{\beta-1} \left(1 + \left(\frac{t}{\eta} \right)^\beta \right)^{-2}, \quad t > 0$$

sendo $\eta > 0$ o parâmetro de escala e $\beta > 0$ o de forma. As funções de confiabilidade ou sobrevivência e de taxa de falha são expressas, respectivamente por:

$$S(t) = \frac{1}{1 + \left(\frac{t}{\eta} \right)^\beta}$$

e

$$\lambda(t) = \frac{\beta \left(\frac{t}{\eta} \right)^{\beta-1}}{\eta \left(1 + \left(\frac{t}{\eta} \right)^\beta \right)}.$$

As expressões para a esperança e variância da log-logística são, respectivamente

$$E(T) = \frac{\pi \eta \operatorname{Csc}\left(\frac{\pi}{\beta}\right)}{\beta}$$

e

$$\operatorname{Var}(T) = \frac{2\pi\eta^2 \operatorname{Csc}\left(\frac{2\pi}{\beta}\right)}{\beta} - E(T)^2$$

em que Csc = cossecante.

2.3.4 Distribuição Gama

Esta distribuição tem sido usada em problemas de confiabilidade, pois a mesma se ajusta adequadamente a uma variedade de fenômenos nessa área, principalmente quando se envolve efeitos aleatórios, como é o caso dos modelos de fragilidade em que se assume essa distribuição para modelar tais componentes.

A função densidade da distribuição gama é caracterizada por dois parâmetros, κ e η , em que $\kappa > 0$ é chamado parâmetro de forma e $\eta > 0$ de escala, é expressa por:

$$f(x) = \frac{1}{\Gamma(\kappa)\eta^\kappa} t^{\kappa-1} \exp\left(-\left(\frac{t}{\eta}\right)\right), \quad t > 0$$

Com $\Gamma(\kappa)$ sendo a função gama. Para $\kappa > 1$, esta função de densidade apresenta um único pico em $t = (\kappa - 1)/\eta$. A respectiva função de confiabilidade dessa distribuição é dada por:

$$S(t) = \int_t^\infty \frac{1}{\Gamma(\kappa)\eta^\kappa} u^{\kappa-1} \exp\left\{-\left(\frac{u}{\eta}\right)\right\} du .$$

A função de taxa de falha é obtida da relação: $\lambda(t) = \frac{f(t)}{S(t)}$, apresenta um padrão

crescente ou decrescente convergindo, no entanto, para um valor constante quando t cresce de 0 a infinito. A média e a variância são dadas respectivamente por $\kappa\eta$ e $\kappa\eta^2$. A distribuição gama com o parâmetro κ restrito a valores inteiros é conhecida como distribuição de Erlang (LEE e WANG, 2003).

2.3.5 Estimação dos Parâmetros dos Modelos

O termo estimação de parâmetros refere-se ao processo do uso de dados amostrais (em engenharia da confiabilidade, dados de tempos até falha) para estimar os parâmetros das distribuições selecionadas como potenciais candidatas ao ajuste do melhor modelo (RELIAWIKI, 2017). Há diversos métodos disponíveis na literatura para estimar parâmetros, talvez o mais conhecido seja o método de mínimos quadrados, no entanto esse método é inapropriado para estudos de tempo de vida, pois impossibilita a incorporação de censuras no seu processo de estimação (COLOSIMO e GIOLO, 2006).

Nesse estudo, para estimação dos parâmetros das distribuições candidatas, utilizou-se o método da máxima verossimilhança, considerada uma das técnicas mais robusta para estimação de parâmetros, já que além de incorporar as censuras, é relativamente simples de ser entendido e possui propriedades ótimas para grandes amostras. A ideia básica do método consiste em obter o valor mais provável para o parâmetro, para uma dada distribuição, que melhor irá descrever os dados.

2.3.5.1 Método da Máxima Verossimilhança

O princípio de máxima verossimilhança é um dos procedimentos mais usados para se obter estimadores. Esse método é formulado matematicamente como se segue:

Se x é uma variável aleatória contínua, como tempo até ocorrência do evento de interesse, com uma determinada pdf, $f(x; \theta_1, \theta_2, \theta_3, \dots, \theta_k)$, onde $\theta_1, \theta_2, \theta_3, \dots, \theta_k$ são parâmetros desconhecidos a serem estimados, com N observações independentes, $x_1, x_2, x_3, \dots, x_N$, que nesse caso correspondem aos dados de vida dos tempos até falha ou dados dos tempos até a primeira postura de codornas. A função de verossimilhança é dada por:

$$f(\theta_1, \theta_2, \theta_3, \dots, \theta_k | x_1, x_2, x_3, \dots, x_N) = L = \prod_{i=1}^N f(x; \theta_1, \theta_2, \theta_3, \dots, \theta_k)$$

$i = 1, 2, 3, \dots, N$ e o logaritmo verossimilhança é dado por:

$$\Lambda = \ln L = \sum_{i=1}^N \ln f(x; \theta_1, \theta_2, \theta_3, \dots, \theta_k)$$

assim os estimadores de verossimilhança ou os valores dos parâmetros de $\theta_1, \theta_2, \theta_3, \dots, \theta_k$, são obtidos maximizando Λ . Logo, os valores dos parâmetros são a solução simultânea do sistema de K equações segundo as derivadas parciais para cada parâmetro e igualando a zero, tal como:

$$\frac{\partial \Lambda}{\partial \theta_j} = 0, \quad j = 1, 2, 3, \dots, k.$$

2.4 TÉCNICA NÃO-PARAMÉTRICA: KAPLAN-MEIER

Em confiabilidade ou sobrevivência, o principal componente da análise descritiva envolvendo dados de tempo de vida é a função de confiabilidade ou sobrevivência. Nesta situação, o procedimento inicial é encontrar uma estimativa para a função de confiabilidade/sobrevivência e, então, a partir dela, estimar as estatísticas de interesse que usualmente são o tempo médio ou mediano, alguns percentis ou certas frações de falhas em tempo fixos de acompanhamento (COLOSIMO e GIOLO, 2006).

Na análise não-paramétrica não se assume um mecanismo gerador dos dados de falha ou tempos de vida, isto é, não se assume uma distribuição de probabilidade para os dados. É claro, há vantagens e desvantagens em se utilizar apenas a técnica não-paramétrica. Ao se analisar os dados sem assumir uma pdf para os mesmos, evita-se o potencial erro devido a suposição de uma distribuição errônea para os dados. Por outro lado, os intervalos de confiança associados a análise não-paramétrica são, usualmente, bem maiores quando comparados com àqueles calculados via técnicas paramétricas (RELIAWIKI, 2017).

O estimador não-paramétrico de Kaplan-Meier, proposto por Kaplan e Meier (1958) para estimar a função de sobrevivência, é também chamado de estimador limite-produto. Ele é adaptação da função de sobrevivência empírica que, na ausência de censuras, é definida como:

$$\hat{S}(t) = \frac{\textit{n}^{\textit{o}} \textit{ de observações que não falharam até o tempo } t}{\textit{n}^{\textit{o}} \textit{ total de observações no estudo}}$$

$\hat{S}(t)$ é uma função escada com degraus nos tempos observados de falha de tamanho $(1/n)$, em que n é o tamanho da amostra. Se existirem empates em um certo tempo t , o tamanho do degrau fica multiplicado pelo número de empates. Naturalmente, o estimador de Kaplan-Meier se reduz à função de sobrevivência empírica se não existirem censuras na análise.

Como nesse delineamento experimental não há presença de censuras nos dados de tempos até a primeira postura, então o estimador não-paramétrico, aqui, reduz-se a função empírica de sobrevivência enunciada acima.

2.5 ESCOLHA DO MODELO

A escolha do modelo a ser utilizado para predições de interesse é um tópico extremamente importante no processo de modelagem em ADV. O método MMV somente pode ser aplicado após ter sido definido um modelo probabilístico adequado para os dados. Algumas vezes existem evidências provenientes de testes realizados no passado de que um certo modelo se ajusta bem aos dados. No entanto, em muitas situações, este tipo de informação não se encontra disponível. Logo, a solução para esses casos é basicamente empírica. Daí a importância de se conhecer técnicas não-paramétricas baseadas na distribuição empírica dos dados.

Sabe-se que as distribuições apresentadas na Subseção 2.3 são potenciais candidatas para os dados de tempos até a primeira postura de codornas. Assim, a proposta empírica consiste em ajustar os modelos probabilísticos apresentados e, com base na comparação entre valores estimados e observados, decidir qual deles melhor explica os dados amostrais.

2.5.1 Critérios de Discriminação

Nessa aplicação, para discriminar o melhor ajuste e obter a distribuição mais apropriada, baseou-se em análises gráficas e, também, nos valores de estatísticas de qualidade do ajuste apresentadas a seguir.

2.5.1.1 Métodos Gráficos

A forma mais simples e eficiente de selecionar o modelo mais adequado a ser usado para o conjunto de dados é por meio de técnicas gráficas. Contudo, testes de hipóteses com modelos encaixados (COX e HINKLEY, 1974) também podem ser utilizados para essa finalidade.

O primeiro método gráfico considerado nesse planejamento experimental, consiste na comparação da função de confiabilidade/sobrevivência dos modelos propostos com o estimador empírico de Kaplan-Meier. Nesse procedimento ajustam-se os modelos propostos ao conjunto de dados e, a partir das estimativas dos parâmetros de cada modelo, estimam-se suas respectivas funções de sobrevivência. Também, para o conjunto de dados, obtêm-se a estimativa de Kaplan-Meier para função de sobrevivência.

Finalmente, comparam-se graficamente as funções de sobrevivência estimadas para cada modelo proposto. O modelo mais aderente é aquele em que a curva de sobrevivência mais se aproximar daquela do estimador de Kaplan-Meier. Assim o melhor modelo é aquele cujos pontos da função de sobrevivência estimada estiverem mais próximos dos valores obtidos pelo estimador de Kaplan Meier.

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, pois violações de linearidade podem ser rapidamente verificadas visualmente.

O gráfico utilizado é o de uma transformação que lineariza a função de sobrevivência do modelo proposto. Isto produz, como resultado final, uma reta, se o modelo proposto for adequado. A ideia é, novamente, comparar o estimador de Kaplan-Meier com o ajuste do modelo proposto.

2.5.1.2 Estatísticas de Qualidade do Ajuste

Como já mencionado anteriormente, as técnicas gráficas são extremamente úteis no processo de seleção de modelos. Entretanto, as conclusões a partir delas podem diferir para diferentes analistas, ou seja, existem nas técnicas gráficas o componente subjetivo inerente a sua interpretação. Dessa forma, outra maneira de

discriminar modelos é por meio de testes de hipótese. Nesse caso, a conclusão é direta e não envolve, portanto, qualquer componente subjetivo na sua interpretação, uma vez que se faz analogia de forma quantitativa e não qualitativa.

Para avaliar o poder de discriminação dos modelos propostos, utilizou-se das estatísticas de qualidade de ajuste já implementadas no software *Weibull++* da *Reliasoft*. Faz-se a classificação das distribuições segundo três testes estatísticos de qualidade do ajuste:

1. Teste de Kolmogorov-Smirnov (KS) – avalia a máxima diferente absoluta entre a função distribuição acumulada proposta para os dados e a função distribuição empírica dos dados. Como critério, compara-se essa diferença com um valor crítico, para um dado nível de significância;
2. Teste do Coeficiente de Correlação – quantifica o quão estão os pontos plotados próximo a uma linha reta;
3. Critério da função log-verossimilhança – calcula o valor do log da função de verossimilhança dado os parâmetros da distribuição: $-2\log(L)$ em que $L = L(\hat{\theta}|x)$ indica a função de verossimilhança localmente em θ .

Em todos os critérios apresentados, decide-se em favor do modelo que apresentar o menor valor da estatística tendendo à zero.

Para o caso de discordância entre algum critério no momento da classificação, pode-se atribuir pesos a cada um dos testes consoante o método de estimação dos parâmetros, assinalando, assim uma maior importância aquele critério que melhor se relaciona com método adotado.

3 RESULTADOS E DISCUSSÃO

Neste capítulo apresenta-se a análise exploratória dos dados, as discussões e os resultados obtidos no estudo.

3.1 ANÁLISE EXPLORATÓRIA DOS DADOS

Os dados utilizados referem-se ao tempo, em dias, do nascimento até a primeira postura de codornas (*Coturnix coturnix japonica*) provenientes de três linhagens genéticas distintas: Amarela, Azul e Vermelha (cor das anilhas), que foram submetidas a duas dietas alimentares. Os dados foram obtidos num experimento conduzido entre agosto de 2006 e janeiro de 2007 no Setor de Coturnicultura da Fazenda Experimental da Universidade Estadual de Maringá. Aos 28 dias de idade as codornas passaram pelo processo de sexagem e as fêmeas foram divididas em dois grupos e alojadas em gaiolas individuais. Cada grupo passou a receber um de dois tipos de dietas: a primeira com nível preconizado de energia metabolizável de 2.900 kcal/dia (dieta 1) e a segunda com nível de energia de 2.500 kcal/dia (dieta 2).

Para toda a análise descritiva dos dados utilizou-se o programa R (R DEVELOPMENT CORE TEAM, 2017).

O número de codornas por linhagem e dieta estão distribuídos segundo os valores apresentados na Tabela 1.

Tabela 1 - Distribuição do número de codornas na combinação dieta consoante linhagem

Dieta	Linhagem		
	Amarela	Azul	Vermelha
1	163	186	201
2	145	188	177

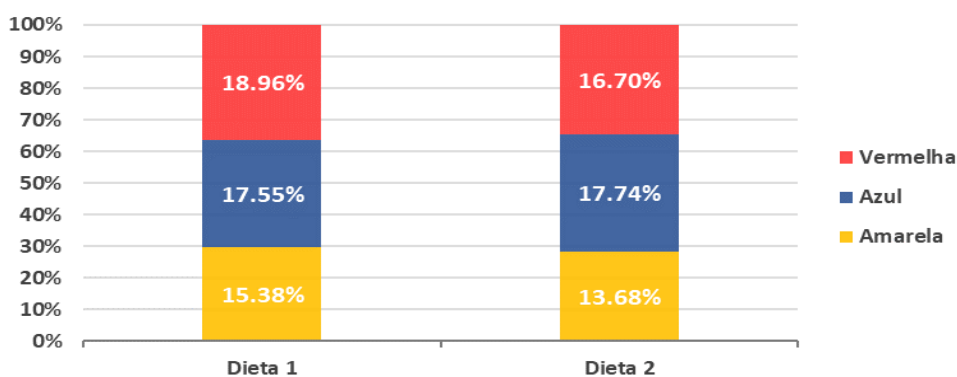
Fonte: o autor (2017).

Ainda que na análise de sobrevivência seja comum a presença de observações censuradas, o evento definido como postura foi observado em todas as unidades amostrais ao longo do experimento.

As Figuras 3 e 4 mostram a distribuição do número de codornas segundo o planejamento experimental. Observa-se que as frequências relativas e absolutas do

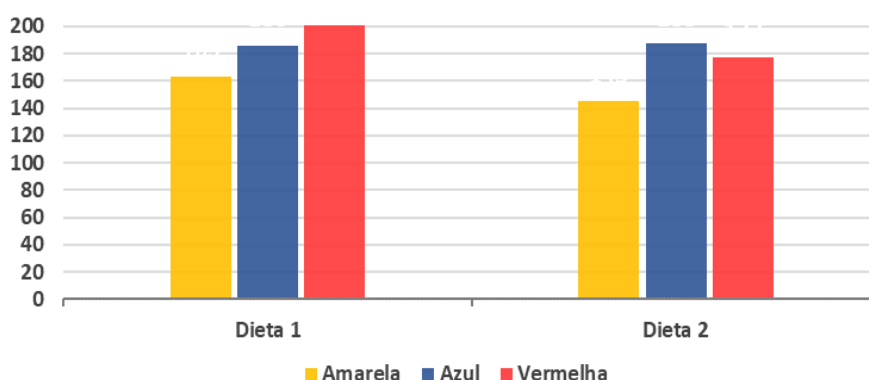
número de codornas conforme linhagens e dietas são equilibradas e proporcionais, garantindo representatividade para cada um dos fatores de tratamento no delineamento amostral.

Figura 3 - Frequência relativa das distribuições do número de codornas segundo dieta e linhagem (representatividade por dieta e linhagem)



Fonte: o autor (2017).

Figura 4 - Frequência absoluta das distribuições do número de codornas segundo dieta e linhagem



Fonte: o autor (2017).

A Tabela 2 apresenta, para cada dieta e linhagem, os tempos: mínimo, máximo, médio e desvio-padrão. Verifica-se que independente da linhagem a dieta 1 (mais calórica) parece ser mais eficiente com um tempo médio de postura menor em relação a dieta 2. Também, nota-se que independente da dieta, a linhagem amarela aparenta ser mais eficiente em relação as demais, também exibindo um tempo médio de postura menor.

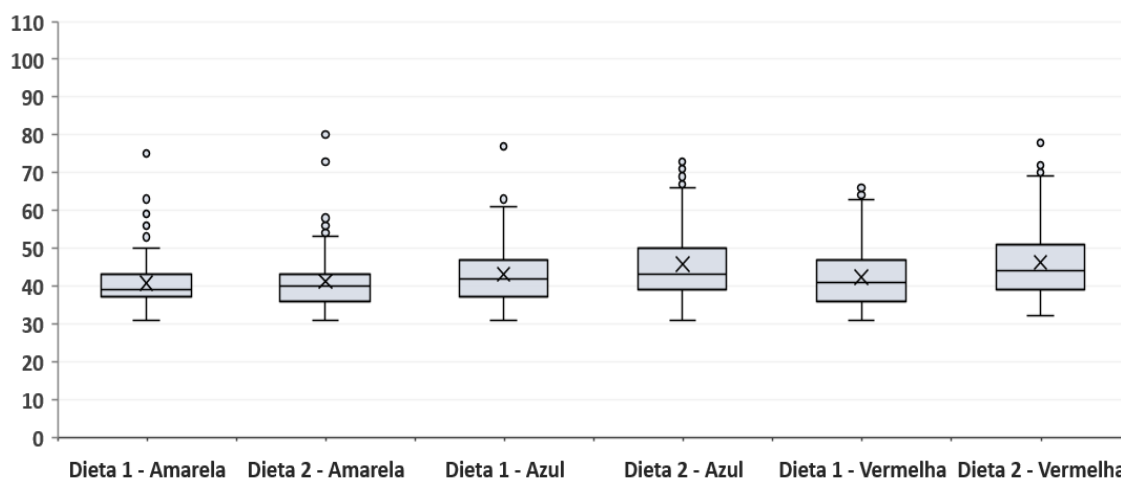
Tabela 2 - Estatísticas descritivas dos tempos (dias) de postura das codornas, por linhagem e dieta

Tratamento		Tempos			
Linhagem	Dieta	Mínimo	Máximo	Médio	Desvio-Padrão
Amarela	1	31	76	40,730	6,791
	2	31	80	41,152	7,322
Azul	1	31	77	43,124	7,395
	2	31	73	45,697	9,192
Vermelha	1	31	66	42,343	7,251
	2	32	105	46,650	10,461

Fonte: o autor (2017).

Na Figura 5 tem-se o resultado da análise dos tempos de postura em função dos fatores de tratamento que aponta para uma inexistência de uma relação melhor definida entre as variáveis. Contudo, assim como na análise da Tabela 2, sugere-se um leve favorecimento para a performance da linhagem amarela e dieta 1. Logo para corroborar o resultado intuitivo e visual e, também, quantificar essa relação faz-se necessário a realização de uma análise de variância (ANOVA) para avaliar se há efeito de tratamento segundo os fatores e interação.

Figura 5 - Frequência absoluta das distribuições do número de codornas segundo dieta e linhagem



Fonte: o autor (2017).

Segundo os resultados obtidos na análise de variância (Tabela 3), admitindo pelo menos 95% de confiança, mostra-se que a interação entre linhagem e dieta é significativa. Além disso, como $F_{2,1054} = 3,008$ e $F_{1,1054} = 3,851$, tem-se que os efeitos

principais tipos da linhagem e tipos da dieta também são significativos. A Tabela 3 também mostra os *valores p* para os testes estatísticos.

Tabela 3 - Análise de variância para os tempos até primeira postura segundo a combinação linhagem e dieta

Fonte de variação	SQ	gl	MQ	Valor F	Valor p
Linhagem	2.616	2	1.308,000	19,450	5,07E-08
Dieta	1.740	1	1.740,200	25,876	4,31E-07
Linhagem : Dieta	638	2	319,000	4,743	8,90E-03
Resíduo	70.884	1.054	67,300		
Total	75.878	1.059			

Fonte: o autor (2017).

Como avaliado na Tabela 3, tanto a interação entre os fatores como os efeitos principais são significativos, logo há de se ajustar todas as potenciais distribuições de probabilidade candidatas para todas as combinações do delineamento fatorial 2 x 3 (dieta versus linhagem).

3.2 AJUSTE DOS MODELOS

Para ajuste e diagnóstico dos modelos candidatos utilizou-se o programa da *Reliasoft Weibull++*, uma vez que há uma gama de modelos e critérios de seleção já implementados no *software*. Considerou-se a estimação via MMV para obtenção do vetor paramétrico e os intervalos de confiança para os mesmos foram calculados considerando a normalidade assintótica dos estimadores de máxima verossimilhança via matriz de informação de Fisher.

As Tabelas 4, 5 e 6 apresentam, para todos os modelos e grupos (linhagem versus dieta), os valores das estatísticas usadas como critérios de seleção de uma dentre todas as distribuições consideradas. Observa-se que na maioria dos tratamentos a distribuição Gama Generalizada apresenta, segundos os critérios de discriminação, o melhor ajuste em relação às outras distribuições, seguida da distribuição Weibull de três parâmetros.

Tabela 4 - Valores das estatísticas e testes de discriminação segundo linhagem amarela e dietas 1 e 2

Linhagem Amarela				
Modelo	Dieta	TA	GA	RV
Lognormal	1	97,23	4,30	-522,68
	2	99,78	5,66	-474,26
Weibull-3P	1	93,97	4,45	-518,55
	2	98,03	5,01	-466,34
Gama-G	1	51,35	2,08	-508,17
	2	63,60	2,72	-456,37
Loglogística	1	62,93	2,27	-514,45
	2	89,37	3,39	-466,75

Fonte: o autor (2017).

Tabela 5 - Valores das estatísticas e testes de discriminação segundo linhagem vermelha e dietas 1 e 2

Linhagem Vermelha				
Modelo	Dieta	TA	GA	RV
Lognormal	1	95,28	3,16	-670,55
	2	96,53	3,73	-631,80
Weibull-3P	1	68,11	1,94	-661,33
	2	69,30	2,27	-622,33
Gama-G	1	82,74	2,27	-663,30
	2	26,66	1,66	-623,72
Loglogística	1	89,09	2,93	-673,59
	2	84,54	3,09	-634,32

Fonte: o autor (2017).

Tabela 6 - Valores das estatísticas e testes de discriminação segundo linhagem azul e dietas 1 e 2

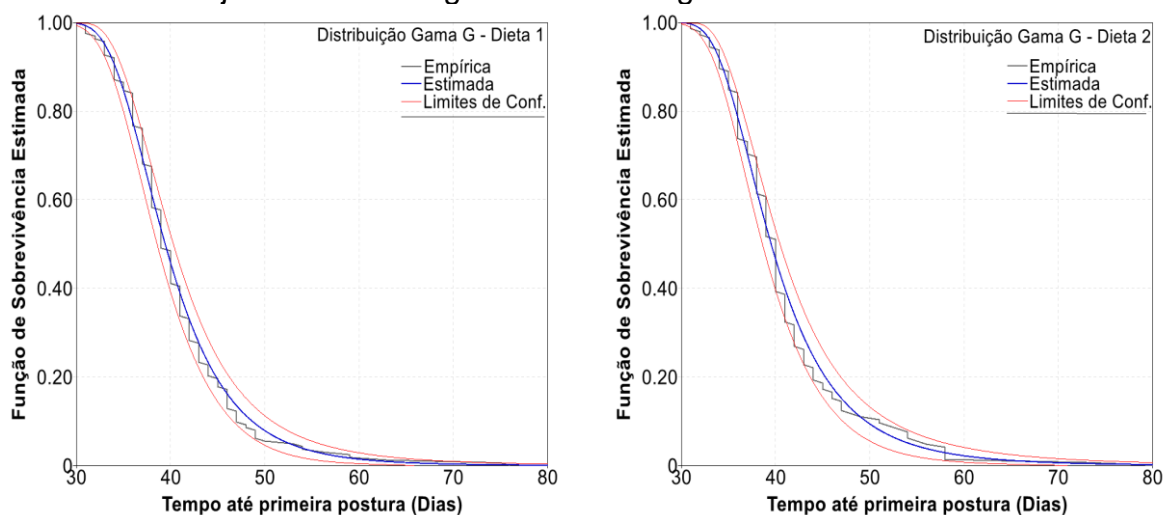
Linhagem Azul				
Modelo	Dieta	TA	GA	RV
Lognormal	1	61,90	1,93	-624,71
	2	98,10	4,60	-667,59
Weibull-3P	1	43,88	1,62	-619,64
	2	88,33	3,53	-659,89
Gama-G	1	65,33	1,65	-621,18
	2	39,96	1,72	-656,34
Loglogística	1	68,73	1,83	-626,33
	2	87,92	3,51	-669,47

Fonte: o autor (2017).

Nas Figuras 6, 7 e 8 tem-se as curvas de sobrevivência empíricas e as baseadas nas estimativas de máxima verossimilhança. Tanto as curvas empíricas como as provenientes dos ajustes dos modelos podem ser nomeadas de curvas de

posturas e estimam a probabilidade do tempo até a primeira postura ser superior a um número de dias quaisquer.

Figura 6 - Função de sobrevivência empírica e distribuição teórica Gama Generalizada ajustadas a linhagem Amarela segundo dietas 1 e 2



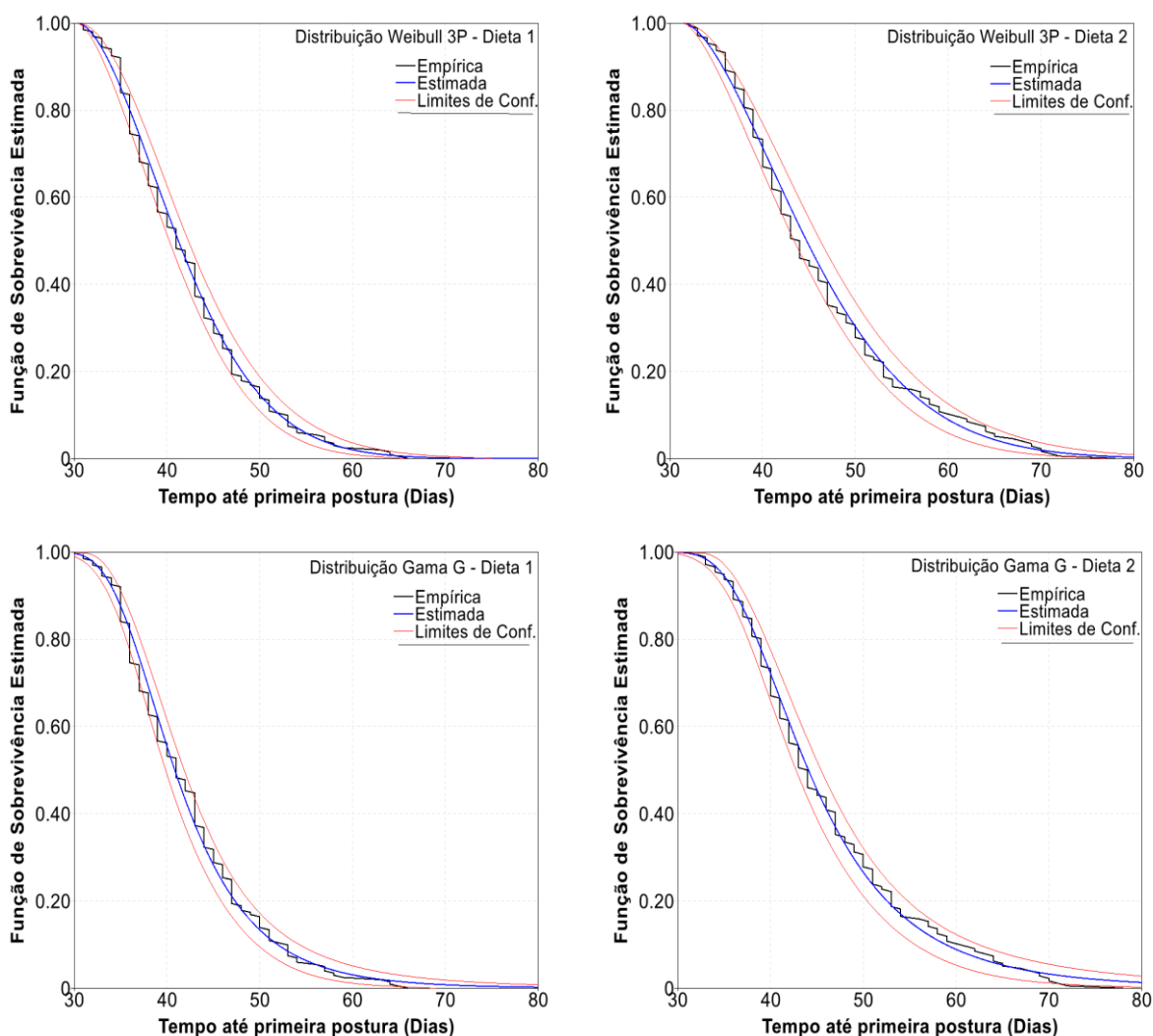
Fonte: o autor (2017).

A partir da Figura 6 tem-se o ajuste da distribuição Gama Generalizada, melhor ranqueada para linhagem amarela independente da dieta, comparada com a distribuição empírica estimada via método não-paramétrico de Kaplan-Meier. Observa-se, a partir de uma inspeção visual, um ajuste aparente de qualidade segundo a boa sobreposição das curvas empíricas e estimadas dentro dos limites de confiança de 95%, corroborando as estatísticas apresentadas na Tabela 4.

A Figura 7 apresenta o ajuste dos modelos teóricos Weibull 3P e Gama Generalizado para linhagem vermelha segundo as dietas 1 e 2, bem como as curvas empíricas. Ao analisar visualmente o ajuste das distribuições teóricas para a dieta 1, verifica-se que aparentemente não há diferença significativa entre os dois modelos, pois parecem estar bem aderentes a curva empírica, ratificando o resultado da Tabela 5 que apesar de mostrar o modelo Weibull 3P como melhor ajuste para dieta 1, o mesmo apresenta valores próximos as estatísticas de discriminação do modelo Gama Generalizado caracterizando um ajuste similar. Avaliando os ajustes para a dieta 2, percebe-se visualmente que há diferença entre os modelos, uma vez que a distribuição Gama Generalizada parece estar mais aderente ao ajuste empírico do que a distribuição Weibull 3P, mais uma vez confirmando os resultados da Tabela 5 que para a dieta 2 apresenta valores mais distantes entre as estatísticas de

discriminação dos modelos teóricos, destacando o modelo Gama Generalizado como melhor ajuste.

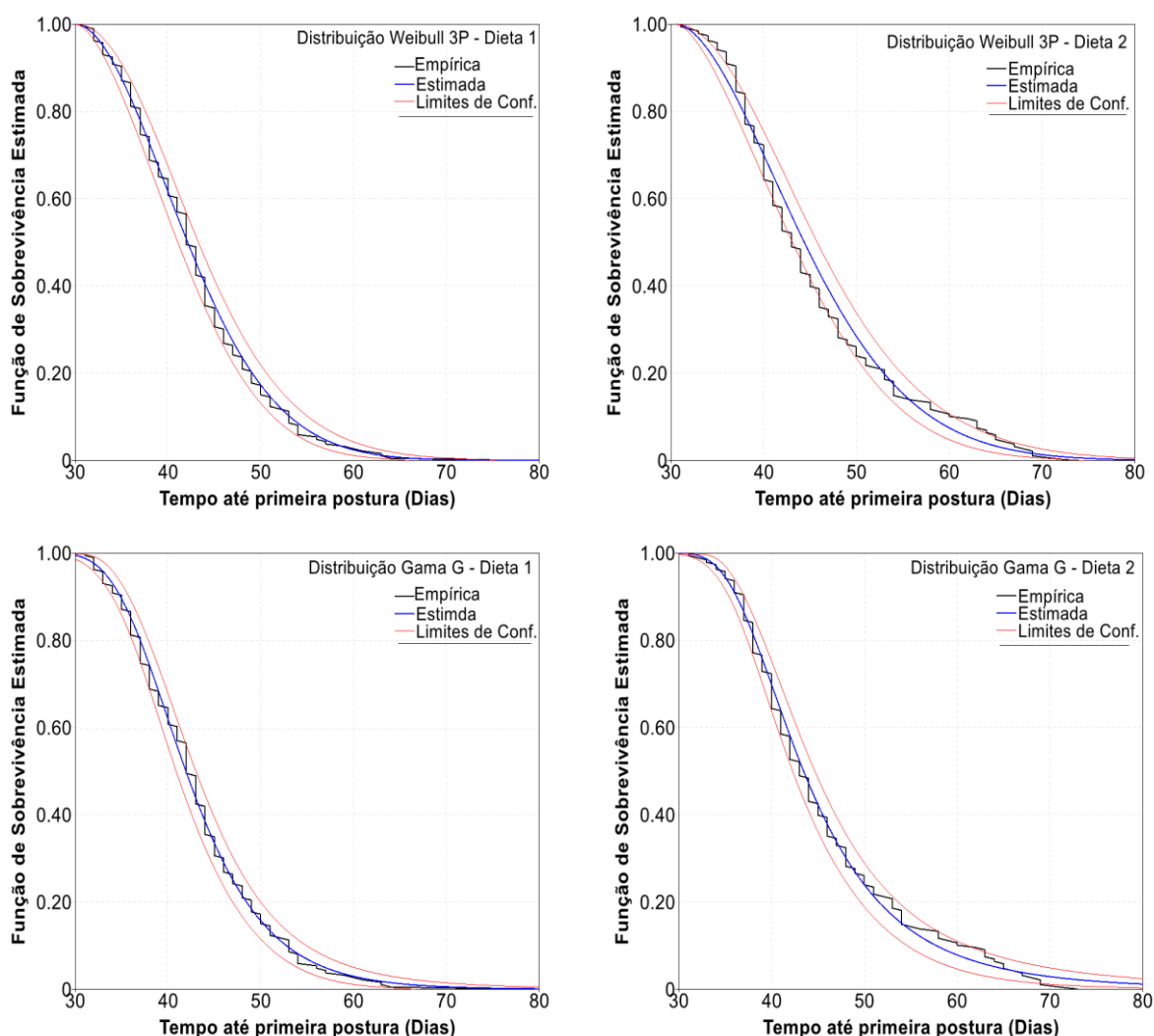
Figura 7 - Função de sobrevivência empírica e distribuição teórica Gama Generalizada e Weibull 3 parâmetros ajustadas a linhagem Vermelha segundo dietas 1 e 2



Fonte: o autor (2017).

Para Figura 8 e Tabela 6 tem-se uma análise similar a anterior, ou seja, para dieta 1 verifica-se um ajuste muito similar entre os dois modelos que aparentemente estão bem sobrepostos à curva empírica e também apresentam valores de aderência próximos. Já para a dieta 2, destaca-se o modelo Gama Generalizado que visualmente está mais aderente a curva e também apresenta melhores valores para os testes de aderência conforme apresentado na Tabela 6.

Figura 8 - Função de sobrevivência empírica e distribuição teórica Gama Generalizada e Weibull 3 parâmetros ajustadas a linhagem Azul segundo dietas 1 e 2



Fonte: o autor (2017).

3.3 DISTRIBUIÇÃO GAMA GENERALIZADA

Após análise das estatísticas de qualidade do ajuste e análise gráfica conclui-se pelo modelo Gama G que apresentou ajuste superior as demais distribuições na maioria dos tratamentos e, nos casos que não foi classificado como o melhor, apresentou valores de aderência muito próximos àqueles do modelo probabilístico melhor classificado, bem como não houve diferença aparente na análise gráfica ao se comparar com a distribuição empírica dos dados.

A Tabela 7 apresenta as estimativas de máxima verossimilhança, pontual e por intervalo segundo modelo probabilístico Gama Generalizado, do tempo médio

até a primeira postura (MTTF), em dias, e da probabilidade de ocorrência desse evento ao tempo médio estimado. Observa-se que a primeira postura ocorre mais rapidamente na linhagem amarela em torno de 40 dias e, também, independente da linhagem a dieta 1 proporciona menor tempo de postura para todas as linhagens.

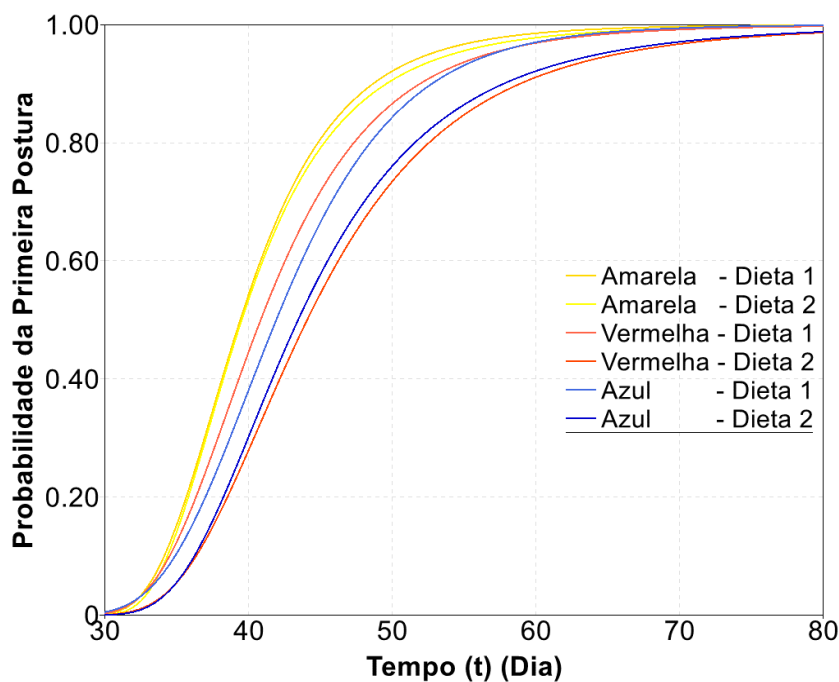
Tabela 7 - Estimativas de máxima verossimilhança do tempo médio até a primeira postura e da probabilidade de ocorrência da primeira postura ao tempo médio

Tratamento		T. Médio (Dia)	Limites de Conf. 95%		P(T _{médio})	Limites de Conf. 95%	
Linhagem	Dieta		L.I.	L.S.		L.I.	L.S.
Amarela	1	40.7003	39.7595	41.6633	0.5903	0.5257	0.6519
	2	41.1247	40.0693	42.2078	0.6079	0.5442	0.6716
Vermelha	1	42.3822	41.3027	43.4899	0.5904	0.5286	0.6495
	2	46.4082	44.9103	47.9560	0.6065	0.5449	0.6681
Azul	1	43.1316	42.0226	44.2700	0.5685	0.5055	0.6295
	2	45.7836	44.4243	47.1845	0.6118	0.5538	0.6699

Fonte: o autor (2017).

Também, verifica-se na Tabela 7 e Figura 9 que a probabilidade de ocorrência do evento postura ao tempo médio estimado se concentra em torno de 60% de chance para todos os tratamentos.

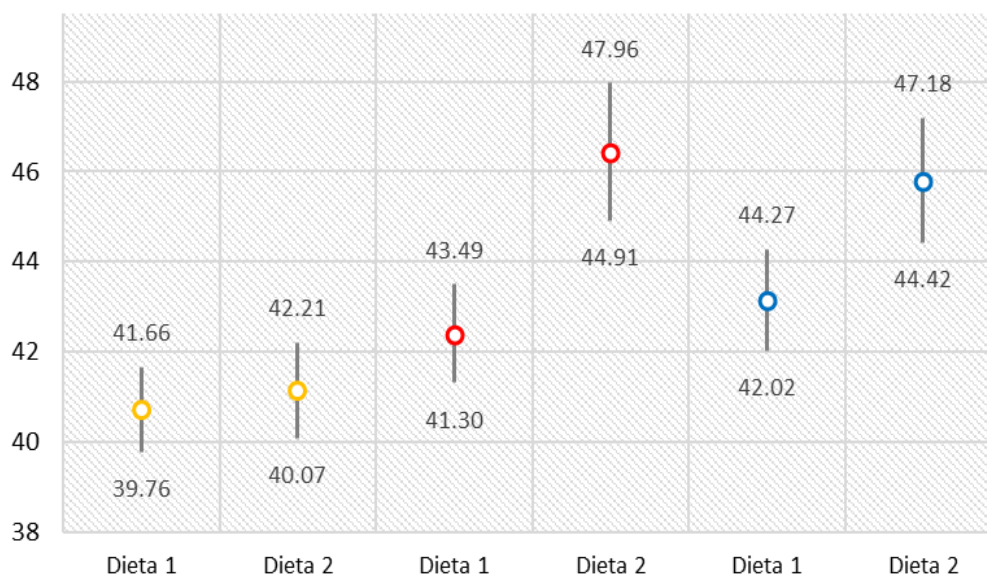
Figura 9 - Curvas da probabilidade da primeira postura segunda distribuição Gama Generalizada para os alinhamentos e dietas



Fonte: o autor (2017).

Além disso, observa-se a maior probabilidade de ocorrência da primeira postura a favor da linhagem amarela independente da dieta sobre as demais linhagens (Figura 9).

Figura 10 - Estimativas pontuais e por intervalo (coeficiente de confiança de 95%) da idade média da primeira postura, em dias, para cada um dos tratamentos segundo a distribuição selecionada



Fonte: o autor (2017).

A Figura 10 apresenta as estimativas, segundo a distribuição selecionada, idade média da primeira postura para todas as linhagens e dietas.

De forma subjetiva, verifica-se na linhagem amarela ausência de efeito da dieta em função da proximidade dos valores estimados, ou seja, parece não haver diferença significativa entre as dietas, corroborando os resultados da análise exploratória dos dados realizada na seção anterior. Já para as linhagens vermelha e azul parece haver diferença entre as dietas, isto é, subjetivamente existe efeito de tratamento, também já observado na análise descritiva dos dados que se confirma com a estatística inferencial.

Toda análise preditiva foi feita segundo o modelo probabilístico Gama Generalizado, uma vez que essa distribuição melhor se ajustou aos dados consoantes os critérios avaliados na seção anterior. Com base nas estimativas obtidas, tem-se que a linhagem amarela destaca-se com o menor tempo de postura independente do nível de dieta alimentar e a dieta 1 mostrou-se mais efetiva para todas as linhagens.

4 CONCLUSÃO

Nesse trabalho buscou-se encontrar um modelo probabilístico para explicar o comportamento da idade até a primeira postura de codornas num planejamento de experimento fatorial 2 x 3 delineado a partir de três linhagens genéticas de codornas e duas dietas distintas. Também, buscou-se avaliar os efeitos da dieta para as diferentes linhagens, cujo maior objetivo é obter melhor linhagem e dieta a fim de otimizar sistema de produção de codornas.

Segundo vários critérios de discriminação, incluindo análise gráfica, chegou-se ao modelo probabilístico Gama Generalizado que apresentou ajuste superior em relação as demais distribuições candidatas rivais de um total de 6 tratamentos – linhagens versus dietas. A partir do modelo selecionado, sob o ponto de vista prático, conclui-se pela ausência de efeito da dieta na linhagem amarela. Já nas linhagens vermelha e azul, o tempo médio de postura na dieta 1 ocorre 4 e 2 dias antes do que na dieta 2 respectivamente. Contudo, independente da dieta a linhagem amarela apresenta um tempo de médio de postura inferior as demais linhagens.

Portanto, consoante os resultados obtidos, tem-se uma diferença de 2 a 4 dias de precocidade postural entre os diferentes grupos, o que já é suficiente para se decidir entre uma dieta ou outra dieta, ou até mesmo por alguma linhagem segundo a necessidade do negócio.

4.1 TRABALHOS FUTUROS

Tal análise também poderia ser aplicada a outros sistemas de produção animal fazendo uso do ferramental da engenharia de confiabilidade e expandindo a fronteira de suas aplicações indo além da falha de seres inanimados.

REFERÊNCIAS

- COLOSIMO, E. A.; GIOLO, S. R. **Análise de Sobrevivência Aplicada**. São Paulo: Edgard Blucher, 2006.
- COX, D. R.; HINKLEY, D. V. 1974. **Theoretical Statistics**. Chapman and Hall, London.
- DIONELLO, N. J. L.; CORREA, G. S. S.; SILVA, M. A. et al. Estimativas da trajetória genética do crescimento de codornas de corte utilizando modelos de regressão aleatória. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v. 60, n. 2, p. 454-460, 2008.
- FREITAS, M. F.; COLOSIMO, E. A. **Confiabilidade: Análise de Tempo de Falha e Testes de Vida Acelerados**. Belo Horizonte: Fundação Christiano Ottoni, Escola de Engenharia da UFMG, 1997.
- GARCIA, E. A. Codornas para produção de carne. In: SIMPÓSIO INTERNACIONAL DE COTURNICULTURA – Novos conceitos aplicados à produção de codornas, 1., 2002, Lavras. **Anais...** Lavras: UFLA/NECTA, 2002, p. 97-108.
- GRIESER, D. O. **Estudo do crescimento e composição corporal de linhagens de codornas de corte e postura**. 2012. 87 f. Dissertação (Mestrado em Zootecnia) - Universidade Estadual de Maringá. Maringá, 2012.
- INSTITUTO BRASILEIRO DE GEOGRAFIA ESTATÍSTICA (IBGE). Produção da Pecuária municipal 2015. Disponível em: <https://ww2.ibge.gov.br/home/estatistica/economia/ppm/2016>. Acesso em: 22 ago. 2017.
- LEE, E. T.; WANG, J. W. **Statistical methods for survival data analysis**. Third. [S.l.]: Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, 2003. (Wiley Series in Probability and Statistics).
- MINVIELLE, F. The future of Japanese quail for research and production. **World's Poultry Science Journal**, v. 60, p. 50-507, 2004.
- MÓRI, C.; GARCIA, E. A.; PAVAN, A. C. et al. Desempenho e rendimento de carcaça de quatro grupos genéticos de codornas para produção de carne. **Revista Brasileira de Zootecnia**, v. 34, n. 3, p. 870-876, 2005.
- PINTO, R.; FERREIRA, A. S.; ALBINO, L. F. T.; GOMES, P. C.; VARGAS, J. G. J. Níveis de Proteína e Energia para Codornas Japonesas em Postura. **Revista Brasileira de Zootecnia**, v. 31, n. 4, p. 1761-1770, 2002.
- R DEVELOPMENT CORE TEAM. **R: a language and environment for statistical computing**. R Foundation for Statistical Computing: Vienna, Austria, 2017. Disponível em: <http://www.R-project.org>. Acesso em: 28 ago 2017.

RELIAWIKI, **Life Distributions**, Disponível em: http://www.reliawiki.org/index.php/Life_Distributions. Acesso em: 02 ago 2017.

SILVA, R. M.; FURLAN, A. C.; TON, A. P. S.; MARTINS, E. N.; SCHERER, C.; MURAKAMI, A. E. Exigências nutricionais de cálcio e fósforo de codornas de corte em crescimento. **Revista Brasileira de Zootecnia**, v. 38, n. 8, p. 1509-1517, 2009.

XIE, M.; TANG, Y.; GOH, T. N. A modified Weibull extension with bathtub-shaped failure rate function. **Reliability Engineering & System Safety**, v. 76, n. 3, p. 279-285, 2002.