

## Métodos estatísticos aplicados à experimentação para avaliação da produção de leite

### Statistical methods applied to experimentation to assess milk production

### Métodos estadísticos aplicados a la experimentación para evaluar la producción de leche

Recebido: 23/08/2021 | Revisado: 03/09/2021 | Aceito: 04/09/2021 | Publicado: 06/09/2021

#### **Ana Maria Duarte Cabral**

ORCID: <https://orcid.org/0000-0002-6101-1104>  
Universidade Federal Rural de Pernambuco, Brasil  
E-mail: amdcabral@gmail.com

#### **Girlene Cordeiro de Lima Santos**

ORCID: <https://orcid.org/0000-0002-6979-6321>  
Universidade Federal Rural de Pernambuco, Brasil  
E-mail: girlenecordeiros@gmail.com

#### **Francisco Fernando Ramos de Carvalho**

ORCID: <https://orcid.org/0000-0001-9211-0263>  
Universidade Federal Rural de Pernambuco, Brasil  
E-mail: ffrcarvalho@hotmail.com

#### **Maiara Guedes Lima**

ORCID: <https://orcid.org/0000-0002-5485-5616>  
Universidade Federal Rural de Pernambuco, Brasil  
E-mail: maiara.guedesac@gmail.com

#### **Andrezza Araújo de França**

ORCID: <https://orcid.org/0000-0001-5050-8812>  
Universidade da Integração Internacional da Lusofonia Afro-Brasileira, Brasil  
E-mail: andrezza.franca@unilab.edu.br

#### **Maria Josilaine Matos Santos da Silva**

ORCID: <https://orcid.org/0000-0003-1990-8257>  
Universidade Federal de Alagoas, Brasil  
E-mail: lainematos@gmail.com

#### **Lucas do Vale Duarte da Silva**

ORCID: <https://orcid.org/0000-0002-2480-8798>  
Universidade Federal Rural de Pernambuco, Brasil  
E-mail: lucasdovalduarten.1997@gmail.com

#### **Bruno Willian Araujo**

ORCID: <https://orcid.org/0000-0001-5424-1318>  
Universidade Federal Rural de Pernambuco, Brasil  
E-mail: bruno.willian2010@hotmail.com

#### **Resumo**

O objetivo desse estudo foi identificar e descrever os tipos de métodos estatísticos aplicados à experimentação para avaliação da produção de leite, a partir de um levantamento de estudos publicados em um período de 10 anos na Revista Brasileira de Zootecnia (RBZ). Realizou-se uma busca eletrônica de artigos publicados na RBZ, que avaliaram a produção de leite caprino (PPLC) ou bovino (PPLB) no período de 2010 a 2019. Em cada artigo, foi avaliado às análises estatísticas, testes de médias, níveis de significância estatística e softwares utilizados. Identificou-se ( $n = 107$  publicações) como principais métodos estatísticos, a análise de variância (58% = PPLB; 36,36% = PPLC) e análise de regressão (35,85% = PPLB; 54,55% = PPLC); testes de médias (teste Tukey = 32,08%; teste F = 22,64; teste  $t = 16,98\%$ ) para PPLB e (testes Tukey = 54,55%; teste F = 0,00%; teste  $t = 9,09\%$ ) para PPLC; os níveis de significância 5% (77,36%), 1% (5,66%) e 10% (0,00%), para a PPLB; e, 5% (90,91%), 1% (0,00%) e 10% (9,09%), para a PPLC; e, os softwares SAS (PPLB = 35,85%; PPLC = 18,18%) e SAEG (PPLB = 41,51%; PPLC = 63,64%). Os principais métodos estatísticos utilizados na experimentação para avaliar a produção de leite são, análise de variância e análise de regressão; o teste de média Tukey; o nível de significância estatísticas de 5% de probabilidade; e, para análise dos dados, os softwares SAS e SAEG são os mais empregados.

**Palavras-chave:** Estatística experimental; Teste de médias; Pecuária leiteira.

#### **Abstract**

The objective of this study was to identify and describe the types of statistical methods applied to experimentation for the evaluation of the production of milk, based on a survey of studies published in a period of 10 years in the Brazilian Journal of Zootechnics (RBZ). An electronic search for articles published in the RBZ was carried out, which will assess the production of goat (PPLC) or bovine (PPLB) milk from 2010 to 2019. In each article, it was analyzed the statistical analyzes, media tests, Statistical significance and software used. Identified ( $n = 107$  publications) as the main statistical methods, analysis of variance (58% = PPLB; 36.36% = PPLC) and regression analysis (35.85% =

PPLB; 54.55% = PPLC) ; means tests (test Tukey = 32.08%; test F = 22.64; test t = 16.98%) for PPLB e (test Tukey = 54.55%; test F = 0.00%; test t = 9.09%) for PPLC; The significance levels are 5% (77.36%), 1% (5.66%) and 10% (0.00%), for PPLB; e, 5% (90.91%), 1% (0.00%) and 10% (9.09%), for a PPLC; e, SAS software (PPLB = 35.85%; PPLC = 18.18%) and SAEG (PPLB = 41.51%; PPLC = 63.64%). The main statistical methods used in experimentation to evaluate milk production they are, variance analysis and regression analysis; or Tukey media test; or statistical significance level of 5% probability; e, to analyze data, the SAS and SAEG software are the most common.

**Keywords:** Experimental statistic; Test of means; Dairy farming.

### Resumen

El objetivo de este estudio fue identificar y describir los tipos de métodos estadísticos aplicados a la experimentación para evaluar la producción de leche, a partir de una encuesta de estudios publicados durante un período de 10 años en la Revista Brasileira de Zootecnia (RBZ). Se realizó una búsqueda electrónica de artículos publicados en la RBZ, que evaluó la producción de leche de cabra (PPLC) o leche bovina (PPLB) en el período de 2010 a 2019. En cada artículo was evaluated, análisis estadísticos, pruebas de medias, niveles de significación estadística y software utilizado. Los principales métodos estadísticos (n = 107 publicaciones) fueron análisis de varianza (58% = PPLB; 36,36% = PPLC) y análisis de regresión (35,85% = PPLB; 54,55% = PPLC); pruebas de medios (prueba de Tukey = 32,08%; prueba F = 22,64; prueba t = 16,98%) para PPLB e (pruebas de Tukey = 54,55%; prueba F = 0,00%; prueba t = 9,09%) para PPLC; los niveles de significancia 5% (77,36%), 1% (5,66%) y 10% (0,00%) para PPLB; y 5% (90,91%), 1% (0,00%) y 10% (9,09%) para PPLC; y el software SAS (PPLB = 35,85%; PPLC = 18,18%) y SAEG (PPLB = 41,51%; PPLC = 63,64%). Los principales métodos estadísticos utilizados en la experimentación para evaluar la producción de leche son, análisis de varianza y análisis de regresión; o prueba de medios de Tukey; o nivel de significación estadística del 5% de probabilidad; e, para analizar datos, los software SAS y SAEG son los más comunes.

**Palabras clave:** Estadísticas experimentales; Prueba de medios; La ganadería lechera.

## 1. Introdução

A experimentação animal foi difundida como técnica sistemática de pesquisa para solucionar problemas corriqueiros nos distintos sistemas de produção observados na pecuária. O princípio dessa metodologia, inclui a utilização de um rigor científico e metodologias preestabelecidas, para formular hipóteses e testá-las por meio de um experimento. Esses experimentos animais permitem a obtenção de dados, que após submetidos a análises estatísticas respondem diferentes perguntas ou simulam o mais próximo possível uma situação real (Sampaio et al., 2008; Salman & Giachetto, 2014; Souza, 2015; Montgomery, 2017).

Logo, a estatística experimental se constitui a melhor forma para as avaliações de dados experimentais, uma vez que, possibilita acurácia e precisão dos resultados achados, e assim conclusões assertivas sobre o problema em questão.

O planejamento de um experimento, inclui, além da formulação de hipóteses, a escolha da unidade experimental, determinação dos fatores e/ou níveis e variáveis a serem testadas. Um experimento quando mal conduzido, compromete as análises e interpretação dos resultados. Por esse motivo, a fim de evitar confundimento e inconfiabilidade dos resultados, é de fundamental importância que o pesquisador tenha o conhecimento dos tipos de delineamentos comumente aplicadas em sua respectiva área de atuação, do mesmo modo, dos métodos estatísticos aplicados nas análises dos dados experimentais (Köche, 2015; Silva et al., 2016).

Uma maneira muito eficaz para o preenchimento dessa lacuna, é a condução de levantamentos documentais na literatura em periódicos da área que se deseja explorar. Esse tipo de pesquisa consiste na busca e sistematização de informações, feita com apoio da estatística descritiva, e tem como objetivo organizar dados e mapear diferentes temas científicos, organizando-os de modo que seja possível avaliá-los de forma tanto quantitativa como qualitativa (Sampaio, 2010; Donato, & Donato, 2019; Santos et al., 2020). De modo geral, nesse tipo de estudo busca-se determinar a frequência com que algo ocorre através de dados obtidos mediante um levantamento documental, ou seja, de registros compilados em diferentes publicações (Callegari-Jacques, 2009; Köche, 2015; Pereira et al., 2018).

Portanto, o objetivo desse estudo foi identificar e descrever os tipos de métodos estatísticos aplicados à experimentação para avaliação da produção de leite, a partir de um levantamento de estudos publicados em um período de 10

anos na Revista Brasileira de Zootecnia (RBZ).

## 2. Metodologia

O estudo foi realizado na Universidade Federal Rural de Pernambuco (UFRPE), Unidade Acadêmica de Serra Talhada (UAST), entre o período de março de 2018 a abril de 2019. Durante este período realizou-se uma revisão, por meio de uma busca eletrônica de artigos publicados na Revista Brasileira de Zootecnia (RBZ), que avaliaram a produção de leite caprino ou bovino no período de 2010 a 2019 (Ludke & André, 1986; Köche, 2016; Estrela, 2018; Pereira et al., 2018).

A escolha do recorte temporal de 10 anos e a RBZ como base de consulta única, é baseado na estatística inferencial, que consiste na utilização uma amostra não aleatória de uma determinada população para se obter estimações sobre um determinado assunto a partir de dados observados (Ferreira, 2005; Frei, 2020). Além disso, a RBZ é um periódico oficial mantido pela sociedade brasileira de Zootecnia (SBZ), que publica artigos inéditos em vários segmentos voltados a produção animal. Até 2017 publicações da RBZ eram realizadas mensalmente, e a partir 2018 passaram a ser realizadas a cada quatro meses. Nesse periódico os artigos são publicados nas seguintes áreas da produção animal: Forragicultura, Produção e Nutrição de Ruminantes, Genética, Sistemas de Produção, Aquicultura, Produção de não ruminantes, Biometeorologia em bem-estar animal e Nota técnica. Sendo que, nesse estudo, como critério de inclusão dos dados, utilizou-se apenas resultados de pesquisas publicadas nas áreas de Forragicultura, Produção e Nutrição de Ruminantes, Genética e Sistemas de Produção.

Inicialmente, os artigos encontrados foram divididos em dois grupos: pesquisa com produção de leite bovino (PPLB) e pesquisa com produção de leite caprino (PPLC). Em cada artigo, foi avaliado os tipos de análises estatísticas conduzidas, testes de médias, níveis de significância estatística e softwares utilizados para realizar as análises dos dados experimentais do estudo.

Com o objetivo de extrair o máximo de informações de interesse nesse estudo, a seleção dos artigos foi conduzida de maneira criteriosa, avaliando o resumo ou abstract, material e métodos e conclusão dos trabalhos.

Após identificados os dados foram compilados e sistematizados com o auxílio de planilha eletrônica Microsoft Excel®. Posteriormente, os dados foram organizados mediante estatística descritiva e submetidos a testes de correlação.

Todos os delineamento e métodos estatísticos identificadas nos artigos foram minuciosamente detalhados nesse estudo a partir de informações disponíveis na literatura.

## 3. Resultados e Discussão

Em 10 anos de publicações na RBZ, constam um total de 107 artigos que avaliaram a produção de leite, sendo 86 artigos relativos a PPLB e 21 a PPLC.

Para as análises estatísticas, 50% das pesquisas referentes a PPLB utilizaram análise de variância, enquanto que nos estudos onde avaliou-se a PPLC, 50% aplicaram a análise de regressão (Tabela 1). O teste de Tukey, ao nível de 5% de significância, e o uso do SAEG são destaques na PPLC, enquanto que na PPLB os testes de médias utilizados são mais variados, e tanto o SAS como SAEG foram comumente utilizados (35 e 41%), respectivamente. Salienta-se que em ambas as pesquisas houve falta de algum tipo de informação, não esclarecendo totalmente a metodologia utilizada. As médias de ambos os estudos são avaliadas ao nível de significância de 5%, resultados que expressam valores máximos adequados com grau de 95% de confiabilidade nos experimentos com PPLB e PPLC.

A análise estatística dos dados experimentais é uma fase muito importante do experimento, pois nela pode-se avaliar se a diferença entre dois ou mais tratamentos, por meio da análise de variância ou da análise de regressão, sendo esta última empregada para verificar a dependência entre variáveis. Desta forma, de acordo com a forma de distribuição dos tratamentos, faz-se uso de uma ou das duas análises (Fisher, 1970; Ferreira, 2000; Devore, 2010). A seguir são apresentadas características

e conceitos necessários para realização da análise de variância (ANOVA) e análise de regressão, testes de médias, níveis de significância e softwares estatísticos de acordo com os resultados obtidos nesse estudo.

**Tabela 1.** Percentual de artigos publicados entre 2010 a 2019 na Revista Brasileira de Zootecnia (RBZ) sobre a produção de leite bovino (PPLB) e caprino e (PPLC), classificados segundo a análise estatística utilizada, os testes de comparação de médias, níveis de significância e softwares utilizados para análise estatística.

Itens de Classificação	Percentual (%)	
	PPLB <sup>1</sup>	PPLC <sup>2</sup>
<b>Análise Estatística</b>		
Análise de Variância	58,49	36,36
Análise de Regressão	35,85	54,55
<b>Teste de Médias</b>		
Testes Tukey	32,08	54,55
Teste F	22,64	0,00
Teste t	16,98	9,09
Não relatado	0,00	36,36
<b>Software Estatístico</b>		
SAS	35,85	18,18
SAEG	41,51	63,64
Não relatado	11,32	0,00
<b>Nível de Significância</b>		
5%	77,36	90,91
1%	5,66	0,00
10%	0,00	9,09

<sup>1</sup>Publicações sobre a produção de leite bovino.

<sup>2</sup>Publicações sobre a produção de leite caprino.

Fonte: Revista Brasileira de Zootecnia (2010 - 2019).

### Análise de Variância

A análise de variância foi utilizada em 58% dos estudos relacionados a PPLB e em 36,36% nos da PPLC (Tabela 1).

A Análise de Variância é uma técnica que permite decompor a variação total existente entre todas as observações, nas variações devidas a causas conhecidas independentes e causas desconhecidas de natureza aleatória. A variação total entre os dados é medida em termos de variância (valores das observações), e é calculada pela soma de quadrados total (SQT) dessas observações, assim como o número de graus de liberdade (GL), os quais representam o numerador e o denominador da equação de variância, respectivamente. E por intermédio do desdobramento da soma de quadrados total de duas ou mais amostras de dados, obtém-se as suas respectivas somas de quadrados entre e dentro de amostragens (Perecin & Cargnelutti Filho, 2008; Callegari-Jacques, 2009).

O quadrado médio (QM), corresponde à estimativa da variância do erro experimental ( $s^2$ ), cujo valor é utilizado nos testes de hipóteses. Nesse sentido, a partir do QM, é possível verificar se existe ou não diferença significativa entre os

tratamentos avaliados. As somas de quadrados divididas pelos seus respectivos graus de liberdade fornecem os quadrados médios (variâncias) entre e dentro de amostragens, respectivamente, os quais são confrontados através de um teste de hipóteses (por exemplo, o teste F) para verificar se as amostras avaliadas diferem significativamente ou não, com relação a alguma variável (Resende, 2007; Costa, 2003; Callegari-Jacques, 2009; Montgomery, 2017).

A análise de variância associada ao teste F para avaliação da igualdade entre tratamentos raramente fornece respostas que são suficientemente específicas, pois, um resultado significativo implica meramente que nem todos os tratamentos apresentam o mesmo efeito, sendo assim, as diferenças entre as médias, podem ser determinadas em um nível de significância pré-estabelecido, através de diversos tipos de testes de médias (Monteiro, 2002; Callegari-Jacques, 2009).

Os dados obtidos relativos aos SQ, GL e QM irão constituir uma tabela (Tabela 2), conhecida como “quadro da Análise de Variância”, ou mais resumidamente como ANOVA. Todavia, para cada tipo de experimento irão ser considerados, fórmulas matemáticas e processo de análises de variância específicas. A análise de variância constitui a ferramenta estatística mais empregada em experimentos com animais, entretanto, para ser aplicada se faz necessário respeitar algumas exigências para se obter diagnósticos seguros e exatos. Dentre essas exigências estão: os efeitos estudados devem ser aditivos; os erros experimentais devem ser independentes, ou seja, o erro em uma observação não deve ser influenciado pelo(s) erro(s) de outra observação; os erros experimentais devem ter variância homogênea, ou seja, o erro não pode ser influenciado ou ser proporcional à dimensão da observação, ou das médias dos tratamentos; e os erros experimentais devem seguir a distribuição normal de probabilidade. Caso uma destas exigências não seja atendida, poderão ocorrer alterações nas propriedades da técnica de análise de variância, acarretando em prejuízo com relação à aplicação dos testes de hipóteses, resultando em valores aproximados ou não-exatos, implicando desta forma a probabilidade de ocorrer erros estatísticos (Costa, 2003; Percin & Cargnelutti Filho, 2008).

Geralmente, essas exigências não são bem aplicadas devido à presença de observações discrepantes (outliers), causadas, muitas vezes, por erros eventuais de mensuração, anotação ou digitação, por comportamento anormal de alguns tratamentos ou parte do experimento (pode ser oriundo de erros na instalação experimental), presença de interações não-consideradas, e por alguns outros problemas associados à instalação experimental (Callegari-Jacques, 2009).

A variância residual deverá ser avaliada de forma bastante criteriosa, caso não, o teste F e demais testes não serão confiáveis. A variação que contribui para o erro experimental pode ser de dois tipos: o erro referente a própria variação natural do material a ser estudado ou o erro proveniente do ambiente experimental. Um exemplo prático desta situação é que um dado experimento busca-se avaliar a influência de uma dieta X na produção de leite de vacas das raças Y. Neste caso, problemas fisiológicos ou sanitários de alguns dos animais podem ser classificados como elementos que favoreçam o erro do material e ser estudado, já que estas variações são referentes aos próprios animais e que podem influenciar nos resultados finais a serem observados. Nesta mesma situação, pode haver variação de temperatura nas baias dos animais, tal situação, pode ser classificada como um erro provocado pelo ambiente experimental, já que esta heterogeneidade pode provocar variações na produção de leite dos animais que não estarão relacionados com a dieta ofertada e sim, com a baixa uniformidade no ambiente em que os animais estão mantidos (Resende, 2007; Callegari-Jacques, 2009; Montgomery, 2017).

Outra verificação de erros, agora inerentes, à condução do experimento é a determinação do coeficiente de variação (CV) entre as médias, que representa o desvio padrão residual expresso como porcentagem da média geral, dado pela seguinte expressão:

$$CV = s / \bar{X} * 100$$

em que:  $\bar{X}$  = média amostral; e,  $s$  = desvio padrão.

O CV é uma medida de dispersão adimensional, que permite a comparação de resultados experimentais por diferentes tratamentos, sendo ainda útil na comparação da variabilidade entre populações distintas, sendo válido, diante de qualquer delineamento utilizado na experimentação. Assim é necessário que o pesquisador esteja atento aos valores razoáveis de CV para as variáveis avaliadas em sua área, que no caso da avaliação da produção de leite, de acordo com Judice (2003) encontram-se classificados da seguinte maneira: CV Baixo =  $\leq 13,22$ ; CV Médio =  $13,22 < CV \leq 36,83$ ; CV Alto =  $36,83 < CV \leq 46,20$ ; e, CV Muito alto =  $> 46,20$ .

### **Análise de Correlação e Regressão**

Na Tabela 1, é possível observar que, para a análise de regressão, 35,85% compreendeu aos estudos com a PPLB e 54,55 % a PPLC.

Aplica-se a análise de correlação e regressão quando em um experimento, se considera a existência de uma variável quantitativa X, que apresenta alguma relação com outra variável quantitativa Y. Na experimentação aplicada a produção de leite, por exemplo, pode-se avaliar as relações entre o consumo de determinado concentrado e teor de gordura no leite; ordem de partos e quantidade de leite produzido; temperatura ambiental e consumo de matéria seca, dentre outros. Nestas situações é pertinente a construção de um gráfico de dispersão dos valores de X versus os valores de Y, que se constitui numa ferramenta estatística simples, porém muito útil, para investigar a existência de uma possível relação entre essas duas variáveis. A correlação determina o grau de relação entre as variáveis e a análise de regressão consiste na estimativa de uma variável dependente a partir de outra variável independente (Paternelli, 2017; Montgomery, 2017 Silva et al., 2019).

O Coeficiente de correlação é calculado tomando-se como base a seguinte equação:

$$r = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{N}}{\sqrt{\left[ \sum X^2 - \frac{(\sum X)^2}{N} \right] \left[ \sum Y^2 - \frac{(\sum Y)^2}{N} \right]}}$$

Esse coeficiente resulta sempre em um valor entre -1 e 1 e sua interpretação depende do seu valor numérico e do seu sinal. Quanto mais próximo de -1 e 1, mais forte é o grau de relação linear existente entre X e Y e, quanto mais próximo de 0, mais fraca é esta relação, e o valor 0 significa que não existe correlação. No estudo de regressão, o pesquisador, geralmente, escolhe os valores da variável independente e depois estabelece a relação existente entre os valores das duas variáveis, tal relação é expressa por uma função matemática (equação de regressão), onde se diz que a variável dependente (Y) é uma função da variável independente (X) (Montgomery, 2017).

A princípio, qualquer relação funcional entre um conjunto de variáveis regressoras e um conjunto de variáveis dependentes, representada por  $y = f(X_1, X_2, \dots, X_k)$ , pode ser chamada de modelo de regressão, sendo tal modelo fixo para determinado conjunto de dados. Um modelo de regressão linear é dito simples quando envolve somente uma variável regressora X. Quando os dados quantitativos de duas variáveis dependentes e independentes não são lineares, seu comportamento é avaliado a partir de regressão polinomial, a qual poderá ter resultados quadráticos e/ou cúbicos. No caso em que há a avaliação de uma variável dependente em função de diversas variáveis independentes, utiliza-se a regressão multivariada para avaliar suas relações. Entretanto, estes dois últimos tipos de regressão, não serão explicitados nesta revisão (Costa, 2003; Montgomery, 2017).

Sejam  $n$  pares de dados de duas variáveis  $(X_i, Y_i)$  com  $i = 1, 2, \dots, n$ . Admitindo que a relação entre  $Y_i$  e  $X_i$  é uma reta, tem-se o modelo de regressão linear simples:

$$y_i = a + b_1 X_i + e_i$$



em que:  $e_i$  é o erro experimental associado a observação  $y_i$ ;  $a$  e  $b$  são parâmetros correspondentes ao coeficiente linear ou termo constante (intercepto da reta) e coeficiente angular ou de regressão, respectivamente;  $e$ ,  $X$  = variável independente (Costa, 2003; Montgomery, 2017). Estes parâmetros podem ser calculados através de quadrados mínimos, e em casos de resultados oriundos de experimentos pelas seguintes fórmulas:

$$\hat{a} = \bar{Y} - b\bar{X} ; \hat{b} = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{N}}{\sum X^2 - \frac{(\sum X)^2}{N}}$$

A fim de comprovar se o modelo se ajusta satisfatoriamente aos dados, ou seja, para verificar a validade da equação da reta, faz uso do coeficiente de determinação  $R^2$ , que nos fornece a porcentagem da variação total de  $Y$  explicada pelo modelo, ou seja, o percentual da variabilidade da variável dependente  $Y$  explicada pela variável independente  $X$  (Costa, 2003).

O Coeficiente de Determinação pode ser calculado elevando-se ao quadrado o coeficiente de correlação  $r$ , ou ainda através da análise de variância da regressão (Tabela 2). Quando os dados apresentam repetições, o coeficiente de determinação poderá ser estimado pela fórmula:

$$R^2 = \frac{SQ \text{ Regressão} \times 100}{SQ \text{ Tratamentos}}$$

em que:  $SQ$  Regressão = soma das  $SQ$  Regressão até a regressão de maior grau que apresentou significância estatística pelo teste  $F$  na análise de variância; e,  $SQ$  Tratamentos = soma de quadrados de tratamentos da análise de variância. Sem repetições, será estimado substituindo o  $SQ$  Tratamentos, por  $SQ$  Total, indicado pela soma de quadrados total da análise de variância (Costa, 2003; Montgomery, 2017).

**Tabela 2.** Quadro da ANOVA para análise de regressão.

Causa de Variação	GL	SQ	QM	F
Regressão	1	SQ Regressão	$\frac{SQ \text{ Regressão}}{1}$	$\frac{QM \text{ Regressão}}{QM \text{ Resíduo}}$
Desvios da Regressão	$n - 1$	SQ Desvios	$\frac{SQ \text{ Desvios}}{(n-2)}$	$\frac{QM \text{ Desvios da Regressão}}{QM \text{ Resíduo}}$
Total	$n - 1$	SQ Total		

GL = número de graus de liberdade; SQ = soma de quadrados; QM = quadrado médio; F = valor calculado do teste F.  
 Fonte: Ferreira (2000).

### Testes de médias e nível de significância

O percentual com os diferentes tipos de testes de médias (teste Tukey = 32,08; teste  $F$  = 22,64; teste  $t$  = 16,98) para os estudos da PPLB e PPLC (testes Tukey = 54,55; teste  $F$  = 0,00; teste  $t$  = 9,09), em 10 anos de publicações ( $n = 107$ ) na RBZ são apresentados na Tabela 1. Observe que, 36,36% dos estudos da PPLB não relataram se utilizaram teste de médias em seus resultados experimentais.

Ao concluir que existe diferença significativa entre tratamentos, através da ANOVA e, portanto, do teste  $F$ , é interessante avaliar a magnitude destas diferenças, a fim de inferir, por exemplo, sobre qual tratamento deve ser utilizado em detrimento de outro. Isto pode ser feito através de teste de comparações múltiplas. São vários os testes disponíveis: teste  $t$ , DMS, Dunnett, Duncan, Scheffe e Tukey, sendo este último o mais utilizado na experimentação animal, por sua maior rigidez

para comparação das médias, em relação aos outros testes, além de ser fácil de ser aplicado (Costa, 2003; Figueiró, 2014; Montgomery, 2017).

Na experimentação animal, o teste de Tukey permite testar qualquer contraste entre duas médias e baseia-se na diferença mínima significativa (DMS) entre elas. Representado pela expressão:

$$\Delta = q \frac{s}{\sqrt{r}}$$

em que: q = valor da amplitude total estudentizada; s = estimativa do desvio padrão residual; r = número de repetições (Costa, 2003; Montgomery, 2017).

Para o nível de significância, os resultados (Tabela 1), compreenderam: 5% (77,36%); 1% (5,66%) e 10% (0,00%), para a PPLB; e, 5% (90,91%), 1% (0,00%) e 10% (9,09%), para a PPLC. O nível de significância com o qual o teste é utilizado indica a probabilidade destes dados serem verdadeiros, representando o grau da confiabilidade da relação observada entre as variáveis como um representativo da população estudada (Costa, 2003; Montgomery, 2017).

### Softwares estatísticos

Observou-se que, entre os principais softwares estatísticos utilizados foram, SAS (PPLB = 35,85; PPLC = 18,18) e SAEG (PPLB = 41,51; PPLC = 63,64), sendo que, 11,32% dos estudos da PPLB, não relataram o tipo de software utilizado nas análises estatísticas.

No entanto, para realização dos processos matemáticos da análise estatística diversos programas computacionais (Statistical Analysis System - SAS, Sistema de Análise de Variância - SISVAR, Sistemas para análises estatísticas e genéticas - SAEG, Minitab, Software for Experimental Statistics in Genetics - GENES) podem ser utilizados. Do mesmo modo, de acordo com a complexidade dos dados, calculadoras especiais gráficas (50, T1 - 83) e programas operacionais como Excel, podem ser empregados. Desta forma, é fundamental que o pesquisador tenha o conhecimento de quais as etapas e ferramentas são necessários na avaliação e interpretação de seus dados experimentais.

## 4. Conclusão

Os principais métodos estatísticos utilizados na experimentação para avaliar a produção de leite são análises de variância e análise de regressão. A Análises de variância é utilizada com maior frequência na avaliação da produção de leite bovinos em detrimento a produção de leite caprino. A Análise de variância é fundamentalmente aplicada para mensurar diferenças entre tratamentos, computando possíveis erros dentro de um experimento. Independentemente do tipo de delineamento, quando se avalia uma variável em função de outra, utiliza-se a análise de regressão, sendo possível ainda, efetuar correlações entre as variáveis.

De modo geral, o teste de Tukey, quando comparado com teste F e o teste t, é o teste de médias mais usado em experimentos para avaliar a produção de leite caprino e bovino. Os testes de média são aplicados para identificar a significância das diferenças das médias observadas afim de se conhecer a confiabilidade dos dados. Sendo, o nível de significância estatísticas “5% de probabilidade” o mais empregado, quando comparado a 1% e 10%. Todas as análises estatísticas na experimentação animal são conduzidas predominantemente nos softwares SAS e SAEG.

Sugere-se pesquisas, que avaliem diferentes métodos estatísticos a partir de uma abordagem mais detalhadas das análises e validação de dados obtidos em experimentos animais.



## Agradecimentos

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES - Código Financeiro 001) pela bolsa de estudos.

## Referências

- Callegari-Jacques, S. M. (2009). *Bioestatística: princípios e aplicações*. Artmed Editora.
- Costa, J. R. (2003). *Técnicas experimentais aplicadas às ciências agrárias*. Embrapa Agrobiologia-Documents (INFOTECA-E).
- Devore, J. L. (2010). *Probabilidade e estatística para engenharia e ciências*. Cengage Learning Edições Ltda.
- Donato, H., & donato, M. (2019). Etapas na Condução de uma Revisão Sistemática. *Acta Médica Portuguesa*, 32(3).
- Estrela, C. (2018). *Metodologia científica: ciência, ensino, pesquisa*. Artes Médicas.
- Ferreira, P. V. (2000). *Estatística Experimental Aplicada à Agronomia*. (3a ed.), EDUFAL, 422p.
- Ferreira, P. L. (2005). *Estatística descritiva e inferencial: breves notas*.
- Figueiró, R. (2014). *Manual prático de Bioestatística Computacional*. Centro Universitário de Volta Redonda - UniFOA Campus Três Poços.
- Fisher, R. A. (1970). *Statistical methods for research workers*. Oliver and Boyd.
- Frei, F. (2020). Aprendizagem baseada em problemas (abp) aplicada ao ensino de estatística inferencial não paramétrica no Ensino Superior. *Revista de Ensino de Ciências e Matemática*, 11(1), 13-26.
- Judice, M. G., Muniz, J. A., & Aquino, L. (2003). *Metodologias para classificação do coeficiente de variação na experimentação com animais*. In: Simpósio De Estatística Aplicada À Experimentação Agrônômica (SEAGRO), 10.
- Köche, J. C. (2015). *Fundamentos de metodologia científica: teoria da ciência e iniciação à pesquisa*. 122-123. Vozes, 2015, 185p.
- Köche, J. C. (2016). *Fundamentos de metodologia científica*. Editora Vozes.
- Ludke, M., & André, M. (1986). *Pesquisa em educação: abordagens qualitativas*. Em Aberto, 5(31).
- Monteiro, L. S. (2002). A experimentação e a experiência. *Revista portuguesa de ciências veterinárias*, 117, 1-16.
- Montgomery, D. C. (2017). *Design and analysis of experiments*. John wiley & sons.
- Perecin, D., & Cargnelutti Filho, A. (2008). Efeitos por comparações e por experimento em interações de experimentos fatoriais. *Ciência e Agrotecnologia*, 32, 68-72.
- Pereira, A. S., Shitsuka, D. M., Parreira, F. J., & Shitsuka, R. (2018). *Metodologia da pesquisa científica*. UFSM.
- Peternelli, L. A. (2017). *Correlação Amostral, Regressão linear Simples e Regressão Linear Múltipla*. <http://www.dpi.ufv.br/~peternelli/inf162.www.16032004/materiais/CAPITULO9.pdf>.
- Resende, M. D. V. (2007). *Matemática e estatística na análise de experimentos e no melhoramento genético*, Embrapa Florestas, Colombo. Forestry Embrapa, Colombo, PR, Brazil.
- Salman, A. K. D., & Giachetto, P. F. (2014). Conceitos estatísticos aplicados à experimentação zootécnica. *PUBVET*, 8, 1416-1550.
- Sampaio, A. A. S., de Azevedo, F. H. B., Cardoso, L. R. D., de Lima, C., Pereira, M. B. R., & Andery, M. A. P. A. (2008). Uma introdução aos delineamentos experimentais de sujeito único. *Interação em Psicologia*, 12(1).
- Sampaio, I. B. M. *Estatística Aplicada Á Experimentação Animal*. (3a ed.), FEP MVZ Editora, 2010. 264p.
- Santos, G. C. L., Gonzaga Neto, S., Bezerra, L. R., & de Medeiros, A. N. *Análise exploratória, conceitual e metodológica do uso de meta-análise aplicada às ciências animais*. In: Carlos Alexandre Oelke. (Org.). Zootecnia Nutrição e Produção Animal. Científica Digital, 2020, 11-24.
- Silva, G. M., da Silva, F. F., de Oliveira Rodrigue, E. S., de Araújo Rufino, C., Pimentel, L. R., de Souza, J., & Santos, M. (2016). Planejamento de experimento a pasto: Revisão. *PUBVET*, 10(5), 399-405.
- Silva, T. G. F., Santos, G. C. L., Duarte, A. M. C., Turco, S. H. N., da Cruz Neto, J. F., Jardim, A. M. D. R. F., & Dos Santos, T. S. (2019). Black globe temperature from meteorological data and a bioclimatic analysis of the Brazilian Northeast for Saanen goats. *Journal of thermal biology*, 85, 102408.
- Souza, F. A. D. (2015). *Experimentos Fatoriais Fracionados Assimétricos para Avaliação de Modelos para Previsão de Chuva no Nordeste do Brasil* (Master's thesis, Brasil).