UNIVERSIDADE DO VALE DO RIO DOS SINOS (UNISINOS)
UNIDADE ACADÊMICA DE GRADUAÇÃO
CURSO DE LICENCIATURA EM MATEMÁTICA

ALEXSANDRO SCHOFFEN

USANDO MODELOS DE REGRESSÃO LINEAR PARA ESTIMAR O PESO DE UMA PRODUÇÃO AVÍCULA

SÃO LEOPOLDO
2023

ALEXSANDRO SCHOFFEN

USANDO MODELOS DE REGRESSÃO LINEAR PARA ESTIMAR O PESO DE UMA PRODUÇÃO AVÍCULA

Trabalho de Conclusão de Curso apresentado como requisito parcial para obtenção do título de Licenciado em Matemática, pelo Curso de Licenciatura em Matemática da Universidade do Vale do Rio dos Sinos (UNISINOS).

Orientador(a): Prof.ª Ms. Ana Cristina Gerhard

SÃO LEOPOLDO
2023

Dedico este trabalho a minha professora orientadora que deu associação para as várias ideias que tinha em mente, mesmo estando confusas. Espero que se sinta orgulhosa. E dedico esse trabalho também a mim, que apesar de tudo, não desisti.

AGRADECIMENTOS

Para iniciar este momento, gostaria de agradecer à professora orientadora, Ms. Ana Cristina, que sempre esteve disposta a me ajudar e teve a sabedoria para organizar meus pensamentos.

Agradeço, também, aos meus familiares e amigos que compreenderam os momentos de ausência ou quando fui omisso em minhas relações. Espero que consigam entender que, às vezes, para pegarmos impulso e voar, precisamos nos afastar.

Aos inúmeros professores que encontrei ao longo da minha trajetória como estudante, meu muito obrigado, pois foram eles que me incentivaram a chegar até aqui.

E, por fim, não poderia deixar de a agradecer a mim, que, mesmo nos momentos mais difíceis, tive persistência para prosseguir com meu sonho. Também, a Deus, que sempre me deu forças e inspiração para chegar até aqui.

RESUMO

Com o presente trabalho, buscou-se verificar a possibilidade de estimar o desenvolvimento de uma produção avícola através dos modelos de regressão linear, considerando as variáveis: dias de alojamento e ração consumida pelas aves. Define-se como modelo de regressão linear uma relação linear simples entre uma variável dependente y e uma variável independente x. Para Chein (2019, p.11) a “variável dependente ou a variável endógena, y, é aquela cujo comportamento será explicado pela variável x, chamada de variável explicativa, regressor ou variável independente”. Obteve-se o resultado a partir dos dados coletados de uma produção avícola em dois aviários que têm uma quantidade total de 49900 aves alojadas. Para a coleta de dados, realizou-se uma pesagem de 1% do total alojado, nos dias em que os frangos estavam 7, 14, 21, 28, 35 e 42 dias na granja. Através destes resultados, concluiu-se que essa relação não se dá por um modelo de regressão linear, uma vez que não conseguimos explicar a variável “peso” através das variáveis independentes observadas.

Palavras-chave: regressão linear; produção avícola.

LISTA DE FIGURAS

LISTA DE TABELAS

SUMÁRIO

1. INTRODUÇÃO

Ao iniciar, peço licença aos legentes, pois com o intuito de demonstrar a pessoalidade dessa pesquisa, escrevo em primeira pessoa. A pesquisa surgiu com o intuito de ajudar a classe dos avicultores, ou seja, produtores do sistema de integração de frangos, a obterem melhores resultados com a estimativa do peso de sua produção. Uma vez que, ao não conseguir realizar essa estimativa de forma precisa, acabam por serem punidos, com um desconto no valor do resultado.

Assim sendo, juntamente com a minha orientadora, estipulamos o tema desse projeto que é a aplicação de conceitos da regressão linear na avicultura. Nosso objetivo principal com esse tema é analisar a possibilidade de estimar o peso de uma produção avícola através de uma reta de regressão, quando este peso depende de diversas variáveis. E o objetivo adicional é estimar o modelo de uma reta de regressão do tipo $y = a + bx$ que nos permita realizar o cálculo do peso médio.

Utilizamos os modelos de regressões lineares como um instrumento estatístico, para resumir informações e dados. Ou seja, podemos prever o comportamento de uma variável, com base no valor de uma outra.

Esse tipo de análise determina os coeficientes entre uma ou mais variáveis independentes, que preveem com uma maior precisão os resultados da variável dependente. No entanto, quando falamos em uma análise de regressão, sempre há preocupação com uma dependência estatística entre as variáveis, pois, na grande maioria dos casos, trabalha-se com variáveis aleatórias que possuem entre si uma distribuição de probabilidades.

Deste modo, acreditamos ser possível considerar uma reta de regressão para estimar o peso de uma produção avícola, uma vez que o peso médio é uma variável que depende de diversos fatores, como o consumo de ração em um determinado período e a quantidade de dias que se passaram. Para descobrir se, de fato, nosso objetivo foi concretizado, realizamos uma pesquisa bibliográfica em diversas fontes como artigos e livros, que contêm estudos de diferentes áreas; e fizemos uma pesquisa *in loco* coletando dados amostrais de uma produção de frangos.

2. FUNDAMENTAÇÃO TEÓRICA

Nesse capítulo, aprofundaremos o referencial teórico. Em um primeiro momento, falaremos sobre o Coeficiente de Correlação. Na sequência, trataremos sobre como se dá um modelo de regressão e a análise do mesmo, assim como seus parâmetros.

2.1 Coeficiente de correlação

O coeficiente de correlação, segundo Virgillito, (2010, p. 402), “trata de medir a intensidade” das relações entre as variáveis. Essas relações, sendo detectadas, poderão ocorrer de maneira casual, ou seja, a relação a ser estudada pode ter intensidade menor na população.

Para quantificarmos a relação entre duas variáveis quantitativas utilizamos o coeficiente de correlação de Pearson. Este coeficiente, entre as duas variáveis X e Y, é dada por:

$$r_{x,y} = \frac{\Sigma(XY) - \frac{(\Sigma X)(\sum Y)}{n}}{\sqrt{\left[\Sigma X^2 - \frac{(\sum X)^2}{n}\right] \cdot \left[\Sigma Y^2 - \frac{(\Sigma Y)^2}{n}\right]}}$$

onde

$n =$ número de variáveis observadas.

$\Sigma(XY) =$ somatório dos resultados das multiplicações dos valores da variável X, pela respectiva variável Y.

$(\Sigma X)(\sum Y) =$ soma dos valores da variável X multiplicado pela soma dos valores da variável Y.

$\Sigma X^2 =$ somatório dos quadrados dos valores de cada variável X.

$(\sum X)^2 =$ quadrado da somatória dos valores de X.

$\Sigma Y^2 =$ somatório dos quadrados dos valores de cada variável Y.

$(\sum Y)^2 =$ quadrado da somatória dos valores de Y.

Virgillito (2010, p. 405 e 406) afirma, também, que “independentemente das variáveis estudadas, o grau de correlação denominado como coeficiente de correlação, estará sempre contido entre dois valores”:

$$-1 \leq r_{x,y} \leq 1$$

Dessa forma, podemos afirmar que, existirá uma correlação perfeita negativa ou uma correlação perfeita positiva. Da mesma maneira que, poderá ter uma correlação perfeitamente nula ou uma correlação não existente. Concluímos, assim, que a partir do coeficiente $r_{x,y}$, conseguimos tirar conclusões sobre a direção e a intensidade da relação existente entre X e Y.

No entanto, sabemos que não existe uma classificação padrão para os coeficientes de correlação. Optamos, então, por seguir a classificação considerada por Santos (2007), que está apresentada abaixo na Tabela 1.

Tabela 1 - Interpretação do coeficiente de Correlação de Pearson Santos (2007)

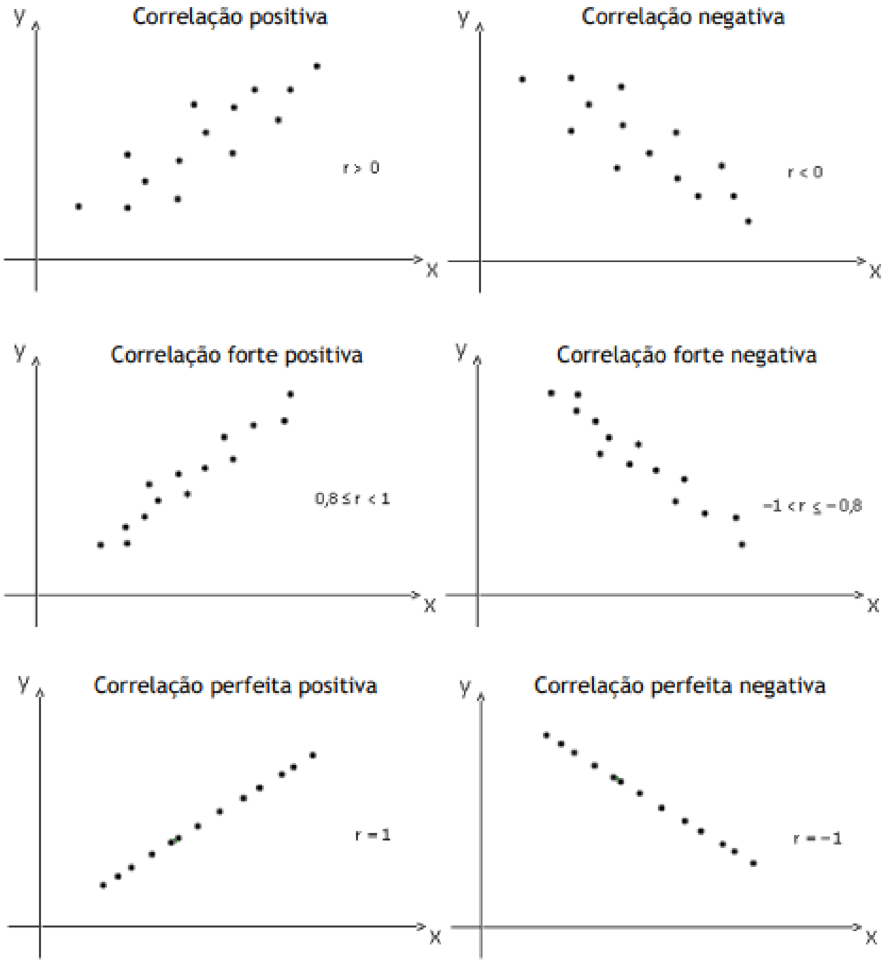
Coeficiente de Correlação	**Correlação**
$R_{x,y} = 1$	Perfeita positiva
$0{,}8 \leq R_{x,y} < 1$	Forte positiva
$0{,}5 \leq R_{x,y} < 0{,}8$	Moderada positiva
$0{,}1 \leq R_{x,y} < 0{,}5$	Fraca positiva
$0 \leq R_{x,y} < 0{,}1$	Ínfima positiva
$R_{x,y} = 0$	Nula
$-0{,}1 \leq R_{x,y} < 0$	Ínfima negativa
$-0{,}5 \leq R_{x,y} < -0{,}1$	Fraca negativa
$-0{,}8 \leq R_{x,y} < -0{,}5$	Moderada negativa
$-1 \leq R_{x,y} < -0{,}8$	Forte negativa
$R_{x,y} = -1$	Perfeita Negativa

Fonte: Santos (2007)

Para realizarmos a investigação da relação entre as variáveis X e Y, podemos organizar os pontos em um gráfico de dispersão. Rodrigues (2012, p. 2) afirma que, “existe uma relação linear quando os dados se aproximam de uma linha reta”.

Assim, conseguiremos verificar através da observação se a correlação entre as duas variáveis é mais ou menos forte, de acordo com a aproximação dos pontos de uma reta. Na figura 1, podemos observar alguns exemplos de gráficos de dispersão e a sua correlação.

Figura 1 – Classificação da correlação através do diagrama de dispersão.



Fonte: Santos (2007)

Como vimos, o coeficiente de correlação nos dá a intensidade da relação linear entre as variáveis. Agora, definiremos o coeficiente de determinação que é o quadrado do coeficiente de correlação.

O coeficiente de determinação é indicado para medir a explicação da reta de regressão. Temos, assim, que, quanto mais o valor se aproximar de 1, maior será a porcentagem de variação de Y, explicada pela reta e, como consequência, teremos uma qualidade maior do ajuste.

Deste modo, o coeficiente de determinação é dado por:

$$r^2_{x,y} = \left(\frac{\Sigma(XY) - \frac{(\Sigma X)(\sum Y)}{n}}{\sqrt{\left[\Sigma X^2 - \frac{(\sum X)^2}{n}\right] \cdot \left[\Sigma Y^2 - \frac{(\Sigma Y)^2}{n}\right]}} \right)^2$$

Logo, o coeficiente de determinação assumirá valores entre o *zero* e o *um*. E a qualidade do ajuste se torna maior quando o ${r^2}_{x,y}$ se aproximar de 1.

Assim, podemos averiguar a presença ou a ausência de uma relação entre as variáveis a partir de dois pontos, citados por Rodrigues (2012):

a) utilizando a análise da correlação;
b) fazendo o uso da análise de regressão.

2.2 Análise de regressão simples

“O termo ‘regressão’ foi proposto pela primeira vez por Francis Galton em 1885 num estudo que demostrou que a altura dos filhos não tende a refletir a altura dos pais, mas tende a regredir a média da população” (Marôco,2003). Estatístico, antropólogo, geógrafo, meteorologista, psicólogo e inventor, Francisc Galton (1822 - 1911), primo de Charles Darwin e neto de Erasmos Darwin, foi considerado uma das mentes mais criativas do século XIX.

Atualmente, quando utilizamos o termo “Análise de Regressão”, referimo-nos a um amplo conjunto estatístico que permite modelar relações entre duas ou mais variáveis para prognosticar o comportamento de uma variável dependente.

Definimos o modelo de regressão linear como uma relação linear simples entre uma variável dependente (y) e uma variável independente (x). Chein (2019, p.11) afirma que a “variável dependente ou a variável endógena, y, é aquela cujo comportamento será explicado pela variável x, chamada de variável explicativa, regressor ou variável independente”.

2.2.1 Modelo teórico

A equação que representa os modelos de regressão linear simples, segundo Berenson, Krehbiel, Levine e Stephan (2012), é dada por:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \qquad (2.1)$$

em que:

$\beta_0 =$ intercepto de y para a população

$\beta_1 =$ inclinação da população

$\varepsilon_i =$ erro aleatório em Y para a observação i

$Y_i =$ variável dependente para a observação i

$X_i =$ variável independente para a observação i

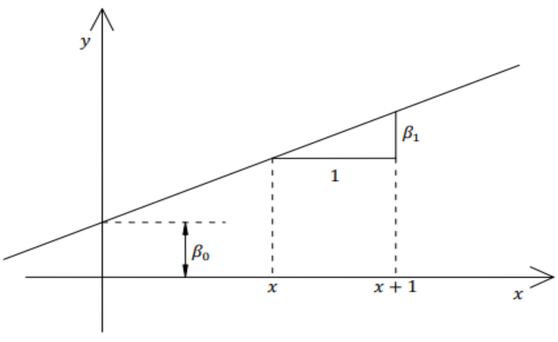
A parcela $Y_i = \beta_0 + \beta_1 X_i$, expressa na equação (2.1), para o modelo de regressão simples, corresponde a uma função afim. Berenson, Krehbiel, Levine e Stephan (2012, p.483), afirmam que

> “a inclinação da linha, β_1, representa a variância esperada em Y por uma unidade de variação em X. Representa a média aritmética da quantidade em que Y varia para uma unidade em alteração em X. O intercepto de Y, β_0, representa a média aritmética do valor de Y quando X é igual a 0.” Berense, Krehbiel, Levine e Stephan (2012, p.483)

Já ε_i, corresponde ao erro aleatório em Y para cada uma das observações. Em outras palavras, é a distância vertical, acima ou abaixo, do valor previsto em Y.

Na figura 2, conseguimos observar e interpretar os parâmetros β_0 e β_1 geométricamente.

Figura 2 - Representação geométrica dos parâmetros β_0 e β_1.



Fonte: Rodrigues (2012, p. 6)

2.2.2 Pressuposto do modelo

Rodrigues (2012, p.6-7) afirma que, ao definirmos um modelo da equação linear como 2.1, pressupomos que:

a) exista uma relação linear entre as variáveis X e Y;

b) os erros são independentes e com média nula:
   Assim, temos que $E(\varepsilon_i) = 0$, logo:

$$\begin{aligned} E(y_i) &= E(\beta_0 + \beta_1 X_i + \varepsilon_i) \\ &= \beta_0 + \beta_1 X_i + E(\varepsilon_i) \\ &= \beta_0 + \beta_1 X_i \end{aligned} \tag{2.2}$$

em que E representa o erro da amostra na variável Y.

Rodrigues (2012, p. 6) diz que "podemos afirmar que o erro de uma observação é independente do erro de outra observação, o que significa que"

$$cov(e_{i'}e_j) = E(e_i e_j) - E(e_i).E(e_j) = E(e_i e_j) = 0, \quad para\, i \neq j \quad i, j = 1, \dots, n.$$

onde, cov é a covariância do erro. Virgillito (2010, p. 413) nos traz que, quando estamos trabalhando com duas populações, obtemos a covariância através da fórmula:

$$COV_{x,y} = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X}) \cdot (Y_i - \bar{Y})$$

No entanto, quando tratamos de duas séries contidas por amostras de variáveis, deveremos utilizar a equação:

$$COV_{x,y} = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y}).$$

c) É constante a variância do erro, isto implica que $var(e_i) = \sigma^2, \ i = 1, \dots, n.$

Temos, assim:

$$var(e_i) = E(e_i^2) - [E(e_i)]^2 = E(e_i^2) = \sigma^2 \,,$$
$$[E(e_i)]^2 = 0$$

consequentemente:

$$var(y_i) = var(\beta_0 + \beta_1 X_i + \varepsilon_i) = var(\beta_0 + \beta_1 X_i) + var(e_i) = \sigma^2$$
$$(\beta_0 + \beta_1 X_i)\ constante$$
$$var(\beta_0 + \beta_1 X_i) = 0$$
$$var(e_i) = \sigma^2$$

na qual, var é a variância. Sabemos que a variância, assim como o desvio padrão, são medidas de dispersão. A variância equivale ao quadrado do desvio padrão (σ^2). Temos então que o desvio padrão é obtido através de:

$$\sigma = \sqrt{\frac{\sum(x_i - \bar{x})^2 \cdot f_i}{\Sigma f_i}}$$

Como resultado, obtemos a variância a partir de:

$$\sigma^2 = \frac{\sum(x_i - \bar{x})^2 \cdot f_i}{\Sigma f_i}$$

em que, f_i representa a frequência relativa de cada uma das variáveis.

d) Os erros de $e_i,\ i = 1, \dots, n$ estão normalmente distribuídos.

Quando falamos em uma distribuição normal, Virgillito (2010), assim como Berenson, Krehbiel, Levine e Stephan (2012), nos ensinam que a distribuição de uma variável aleatória contínua em X poderá ser dita como normal, quando ela apresentar as seguintes características:

1. O ponto máximo de $f(X)$ é o ponto $\bar{x} = \mu(média)$
2. Os pontos afastamento da função são $X = \mu - \sigma$ e $X = \mu + \sigma.$
3. A curva da função é simétrica em relação a μ, obtendo, assim, o formato de um sino.
4. E a variável X terá uma distribuição normal das suas probabilidades se a função for definida por:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{(x-\mu)^2}{2\sigma^2}}.$$

Assim, das proposições b) e c) podemos concluir que:

$$e_i \sim N(0, \sigma^2)\ , \qquad i = 1, \dots, n$$

portanto,

$$y_i \sim N(\beta_0 + \beta_1 X_i, \sigma^2), \qquad i = 1, \dots, n$$

2.3 Estimação dos parâmetros do modelo

Supõe-se que exista uma relação entre as variáveis $X\ e\ Y$, colocamos, então, em indagação como estimar os parâmetros $\beta_0\ e\ \beta_1$.

Karl Gauss, entre 1877 e 1885, propôs estimar os parâmetros $\beta_0 \ e \ \beta_1$, visando minimizar a soma dos quadrados dos desvios $e_i, i = 1, \ldots, n$, chamando esse processo de método nos mínimos quadrados. (Maroco, 2003).

2.3.1 Método dos mínimos quadrados

Sabemos que o método dos mínimos quadrados consiste em medir e comparar o quadrado das Variações não Explicadas de cada equação em ajuste. Segundo Virgillito (2012, p.422), “o melhor ajuste, ou seja, a equação que melhor explica o comportamento dos dados através da Correlação de suas variáveis é aquele que produz o menor das variações não explicadas, o que obviamente produzirá o maior coeficiente de variação”.

O método que nos produzirá a menor das variações é aquele em que a função (equação) melhor nos descreve a trajetória e o comportamento das variáveis. Essa função é denominada linha de previsão. Com essa linha, conseguimos notar os valores previstos para Y.

Berenson, Krehbiel, Levine e Stephan (2012) nos ensinam que a equação da linha de previsão é dada por:

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i \qquad (2.3)$$

onde, o valor previsto de Y é igual ao intercepto de Y somado a inclinação vezes o valor de X.

Sendo assim, a equação (2.3) requer a determinação de dois coeficientes da regressão β_0 (o intercepto de Y na amostra) e β_1 (inclinação da amostra). Para determinarmos estes coeficientes, a abordagem mais comum é o método dos mínimos quadrados. Segundo Berenson, Krehbiel, Levine e Stephan (2012, p.485), “esse método minimiza a soma das diferenças dos quadrados entre os valores verdadeiros de y_i e os valores previstos de $\hat{y}_i$, utilizando a equação da regressão linear simples.”

Temos, assim,

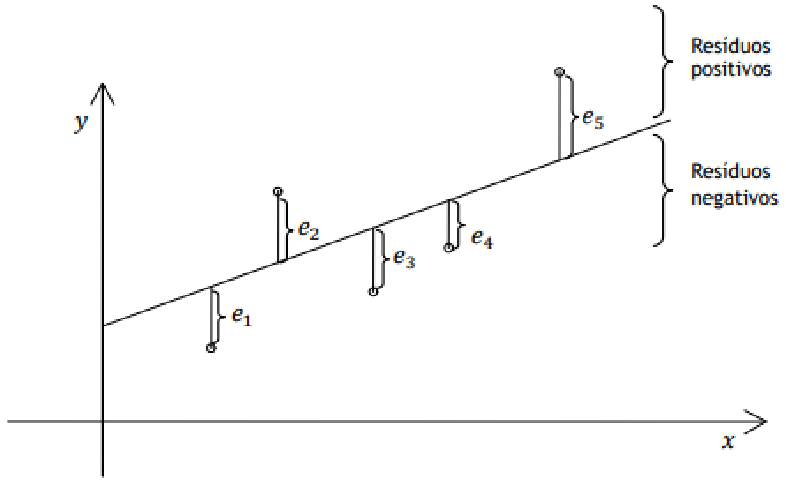
$$\sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

uma vez que $\hat{y}_i = \beta_0 + \beta_1 X_i$,

$$\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} [y_i - (\beta_0 + \beta_1 X_i)]^2$$

Em termos gráficos, podemos observar que estes resíduos representam as distâncias verticais entre os valores ajustados e os valores observados, como mostra a figura 3.

Figura 3 - Resíduos representados graficamente.



Fonte: Rodrigues (2012, p. 8)

Carvalho e Dachs (1984), assim como Rodrigues (2012 p. 8-9), nos ensinam que, inicialmente, precisamos realizar o cálculo das derivadas parciais das somas dos quadrados dos resíduos (SQE), em ordem a β_0 e β_1, assim obtemos:

$$\begin{cases} \dfrac{\partial SQE}{\partial \beta_0} = -2\displaystyle\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 X_i) \\ \dfrac{\partial SQE}{\partial \beta_1} = -2\displaystyle\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 X_i)X_i \end{cases}$$

Na sequência, igualamos as derivadas a zero e realizamos a troca dos parâmetros β_0 e β_1, por $\hat{\beta}_0$ e $\hat{\beta}_1$, para, então, indicarmos valores concretos para estes parâmetros, obtemos, assim:

$$\begin{cases} \dfrac{\partial SQE}{\partial \beta_0} = -2\displaystyle\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0 \\ \dfrac{\partial SQE}{\partial \beta_1} = -2\displaystyle\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)x_i = 0 \end{cases}$$

$$\begin{cases} n\hat{\beta}_0 = \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} \hat{\beta}_1 x_i \\ \hat{\beta}_0 \sum_{i=1}^{n} x_i + \hat{\beta}_1 \sum_{i=1}^{n} X_i^2 - \sum_{i=1}^{n} x_i \, y_i = 0 \end{cases} \quad (2.4)$$

$$\left\{ \hat{\beta}_0 = \frac{\sum_{i=1}^{n} y_i}{n} - \frac{\hat{\beta}_1 \sum_{i=1}^{n} x_i}{n} \right.$$

$$\left\{ \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \right. \quad (2.5)$$

onde, $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ e $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ corresponde as médias de Y e X, respectivamente.

Agora, pegaremos a segunda equação em 2.4 e, consequentemente, tentaremos chegar à equação do parâmetro $\hat{\beta}_1$.

$$\hat{\beta}_1 \sum_{i=1}^{n} {x_i}^2 = \sum_{i=1}^{n} x_i \, y_i - \hat{\beta}_0 \sum_{i=1}^{n} x_i$$

$$\hat{\beta}_1 \sum_{i=1}^{n} {x_i}^2 = \sum_{i=1}^{n} x_i \, y_i - (\bar{y} - \hat{\beta}_1 \bar{x}) \sum_{i=1}^{n} x_i$$

Como

$$\sum_{i=1}^{n} x_i \, y_i - \bar{y} \sum_{i=1}^{n} x_i + \hat{\beta}_1 \bar{x} \sum_{i=1}^{n} x_i$$

$$= \sum_{i=1}^{n} x_i \, y_i \, n\bar{x}\bar{y} + n\hat{\beta}_1 \bar{x}^2 \, ,$$

vem

$$\hat{\beta}_1 \sum_{i=1}^{n} X_i^2 = \sum_{i=1}^{n} x_i \, y_i - n\bar{x}\bar{y} + n\hat{\beta}_1 \bar{x}^2$$

$$\hat{\beta}_1 \left(\sum_{i=0}^{n} x_i - n\bar{x}^2 \right) = \sum_{i=0}^{n} x_i \, y_i - n\bar{x}\bar{y}$$

$$\hat{\beta}_1 = \frac{\sum_{i=0}^{n} x_i y_i - n\bar{x}\bar{y}}{\sum_{i=0}^{n} x_i^2 - n\,\bar{x}_i{}^2} \qquad (2.6)$$

os parâmetros $\hat{\beta}_0$ e $\hat{\beta}_1$ determinados em (2.5) e (2.6) são determinados como os Estimadores dos Mínimos Quadrados de β_0 e β_1.

Em seguida, apresentaremos propriedades dos mínimos quadrados.

- Como apresentado por Berenson Krehbiel, Levine e Stephan (2012, p.485), os resíduos correspondem diferenças dos quadrados entre os valores verdadeiros de y_i e os valores previstos de $\hat{y}_i$, temos, assim, que:

$$e_i = y_i - \hat{y}_i$$

$$= y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i)$$

com:
$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i \qquad i = 1, \dots, n.$$

a) $\sum_{i=0}^{n} e_i = 0$, isso significa que a soma dos resíduos sempre será igual a 0.
b) $\sum_{i=0}^{n} e_i{}^2$ é mínima.
c) $\sum_{i=0}^{n} y_i = \sum_{i=0}^{n} \hat{y}_i$, temos que a soma dos valores observados para y_i é sempre igual aos valores observados para $\hat{y}_i$.
d) A reta que obtemos pelo método dos mínimos quadrados passa pelo ponto determinado por $(\bar{x}, \bar{y})$.

Demonstração:
$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

$$= \beta_0 + \beta_1 x_i + \varepsilon_i + \beta_1\bar{x} - \beta_1\bar{x}$$

$$= \beta_0 + \beta_1(x_i - \bar{x}) + \beta_1\bar{x} + \varepsilon_i$$

$$= (\beta_0 + \beta_1\bar{x}) + \beta_1(x_i - \bar{x}) + \varepsilon_i$$

$$= \beta_0^0 + \beta_1(x_i - \bar{x}) + \varepsilon_i$$

com

$$\beta_0^0 = \beta_0 + \beta_1\bar{x}.$$

Logo, os valores estimados serão dados por:

$$\hat{y}_i = \beta_0 + \beta_1\bar{x} + \beta_1(x_i - \bar{x}_i)$$

$$= \left(\bar{y} + \hat{\beta}_1\bar{x}\right) + \beta_1\bar{x} + \beta_1(x_i - \bar{x}_i)$$

$$= \bar{y} + \beta_1(x_i - \bar{x}_i).$$

Vimos que,

$$\hat{\beta}_0 = \bar{y}_i + \hat{\beta}_1\bar{x}$$

Assim, temos que no ponto da abcissa $\bar{x}$:

$$\hat{y}_i = \bar{y} + \hat{\beta}_1(\bar{x} - \bar{x}) = \bar{y}$$

2.4 Testes e intervalos de confiança parâmetros

Nesse capítulo, realizaremos a construção de hipóteses para os parâmetros β_0 e β_1. Junto com a elaboração das hipóteses, faremos, também, a elaboração de intervalos de confiança para os parâmetros.

2.4.1 Teste de Hipóteses

É comum em nossos dias realizarmos algumas afirmações. Estas, que na grande maioria dos casos, vêm acompanhadas de bases que nos fazem acreditar nessas alegações.

Sendo assim, os Testes de Hipóteses ajudam a verificar a veracidade da afirmação feita a respeito dos parâmetros populacionais.

Virgillito (2010, p. 280) nos ensina que os Testes de Hipóteses

> [...] podem ser divididos em dois grupos: os testes paramétricos, assim chamados por se referirem as hipóteses sobre os parâmetros populacionais e aqueles chamados não paramétricos, pois na verdade se referem exclusivamente a forma de distribuição, ou em outras palavras, dada uma distribuição de frequências, determina-se a que tipo de distribuição de probabilidades esta 'adere' melhor. Virgillito (2010, p. 280)

Através do Teste de Hipótese testamos uma hipótese inicial, H_0. Essa, é inicialmente tida como verdadeira. A partir de uma amostra válida da população, tentaremos provar se a nossa hipótese inicial é realmente verdadeira; caso não seja, realizaremos a substituição por uma hipótese alternativa, H_1.

2.4.2 Teste e Intervalo de Confiança $\boldsymbol{\hat{\beta}_1}$

Na equação 2.6, vimos que o estimador do parâmetro $\hat{\beta}_1$ é dado por:

$$\hat{\beta}_1 = \frac{\sum_{i=0}^{n} x_i y_i - n\bar{x}\bar{y}}{\sum_{i=0}^{n} x_i^2 - n {x_i}^2}$$

Em seu estudo, Rodrigues (2012 p. 20) nos mostra que ao testarmos as hipóteses,

$$\begin{cases} H_0: \beta_1 = 0 \\ H_1: \beta_1 \neq 0 \end{cases}$$

concluímos que o intervalo de confiança de β_1 para $(1-\alpha) \times 100\%$ é dado por:

$$\left[\hat{\beta}_1 - t\left(1 - \frac{\alpha}{2}; n-2\right) S_{\hat{\beta}_1}; \widehat{\beta}_1 + t\left(1 - \frac{\alpha}{2}; n-2\right) S_{\hat{\beta}_1}\right]$$

onde, α indica o nível de significância.

2.4.3 Teste e Intervalo de Confiança $\boldsymbol{\hat{\beta}_0}$

Já na equação 2.5, vimos que o estimador do parâmetro $\boldsymbol{\hat{\beta}_0}$ é dado por:

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

Rodrigues (2012 p. 21) também mostrou que, ao realizarmos o teste de hipótese,

$$\begin{cases} H_0: \beta_0 = 0 \\ H_1: \beta_0 \neq 0 \end{cases}$$

obtemos que o intervalo de confiança, assim como em β_1, $(1-\alpha) \times 100\%$, será dado por:

$$\left[\hat{\beta}_0 - t\left(1-\frac{\alpha}{2};n-2\right)S_{\widehat{\beta}_0} ; \widehat{\beta}_0 + t\left(1-\frac{\alpha}{2};n-2\right)S_{\widehat{\beta}_0}\right]$$

3. METODOLOGIA

Nesse segmento, trabalharemos com a metodologia utilizada para a realização deste trabalho e a sequência utilizada na coleta de informações.

Nesta pesquisa, apuramos resultados de um lote de produção de frangos com data de alojamento em 03 de fevereiro e data de abate em 20 de março de 2023. Os dados foram coletados em dois aviários de uma granja avícola, situada no interior da cidade de Tupandi, município pertencente à região do Vale do Rio Caí. Um aviário possuía 22900 aves alojadas e, o outro, uma quantidade alojada de 27000 aves. Para uma melhor organização, denominamos de aviário A e aviário B, respectivamente.

Define-se como data de alojamento o dia em que os frangos são transportados da incubadora até à granja, e, data do abate, o dia em que os frangos são carregados em caminhões específicos na granja e vão para o frigorífico. O abate das aves ocorre, em média, 45 dias após a data do alojamento.

A coleta foi feita seguindo as normas da empresa de integração, ou seja, devemos realizar a pesagem dos frangos nos dias em que se completa o ciclo de 7, 14, 21, 28, 35 e 42 dias dentro da granja, sendo necessário ser pesado uma quantidade correspondente até, no mínimo, 1 % da quantidade inicial alojada. Os dados foram organizados em forma de planilhas, que foram elaboradas no Excel e se encontram anexas neste trabalho.

Para estimarmos o peso médio semanal, foi calculado o quociente entre o somatório dos pesos e da quantidade de aves em cada amostra. Como o consumo de ração não pode ser controlado de forma separada, pois ambos os aviários consomem ração vindo do mesmo reservatório, para realizarmos a modelagem da situação, calculamos a proporção consumida por cada frango. Ou seja, pegamos o quociente que resulta da razão entre o total consumido com o total de frangos na granja.

4 ANÁLISE DOS RESULTADOS

Neste segmento, faremos a análise de dados para chegarmos, assim, a uma conclusão sobre nosso problema estabelecido inicialmente. A organização ocorre por subtítulos.

4.1 Aviário A: relação entre peso e dias de alojamento.

Neste aviário, realizamos uma pesagem de 229 aves, ou seja, 1% da quantidade alojada. Na tabela abaixo, organizamos os pesos semanais a partir dos resultados obtidos. Na mesma, encontram-se os valores que serão necessários para calcularmos o coeficiente de correlação.

Além do mais, como a variável do peso depende da quantidade de dias que se passaram, temos que dias (x) é a nossa variável independente e o peso (y) como variável dependente.

Tabela 2 - Aviário A: relação entre peso e dias de alojamento

Dias (X)	Peso (Y)	x.y	x^2	y^2
7	180	1260	49	32400
14	420	5880	196	176400
21	862	18102	441	743044
28	1474	41272	784	2172676
35	2074	72590	1225	4301476
42	2660	111720	1764	7075600
ΣX 147	$(\sum Y)$ 7670	$\Sigma(XY)$ 250824	ΣX^2 4459	ΣY^2 14501596

Fonte: elaborada pelo autor.(2023)

A partir dessa tabela, calculamos o coeficiente de correlação, trazido por Virgillito (2010, p. 405). O coeficiente de correlação foi calculado através da equação:

$$r_{x,y} = \frac{\Sigma(XY) - \frac{(\Sigma X)(\sum Y)}{n}}{\sqrt{\left[\Sigma X^2 - \frac{(\sum X)^2}{n}\right] \cdot \left[\Sigma Y^2 - \frac{(\Sigma Y)^2}{n}\right]}}$$

onde,

$n = 6$, pois é o número de elementos da nossa amostra.

$(\Sigma X) = 147$

$(\sum Y) = 7670$

$\Sigma(XY) = 250824$

$\Sigma X^2 = 4459$

$\Sigma Y^2 = 4501496$

A partir disso temos:

$$r_{x,y} = \frac{250824 - \frac{(147)(7670)}{6}}{\sqrt{\left[4459 - \frac{(147)^2}{6}\right] \cdot \left[14501496 - \frac{(7670)^2}{6}\right]}}$$

$$r_{x,y} = \frac{250824 - \frac{1127490}{2}}{\sqrt{\left[4459 - \frac{21609}{6}\right] \cdot \left[14501496 - \frac{58828900}{6}\right]}}$$

$$r_{x,y} = \frac{250824 - 563745}{\sqrt{[4459 - 3601{,}5] \cdot [14501496 - 98038150]}}$$

$$r_{x,y} = \frac{-312921}{\sqrt{857{,}5} \cdot [-83536654]}$$

$$r_{x,y} = \frac{-312921}{29{,}2831 \ \cdot [-83536654]}$$

$$r_{x,y} = \frac{-312921}{-2446212{,}7474}$$

$$r_{x,y} = 0{,}1279$$

Logo, podemos afirmar que existe relação entre as variáveis. A partir do coeficiente de correlação, determinamos o coeficiente de determinação, que é o quadrado do coeficiente de correlação. Assim obtemos:

$$(r_{x,y})^2 = (0{,}1279)^2$$

$$(r_{x,y})^2 = 0{,}01623076$$

Com o coeficiente de determinação, chegamos à conclusão de que existe uma relação entre as variáveis, porém, como o valor está distante do 1, essa relação se dá de forma fraca. A partir disso, estimamos os parâmetros do modelo de acordo com as equações encontrada em 2.5 e 2.6.

Inicialmente, pegamos a equação 2.6, pois, sabendo o valor do parâmetro $\hat{\beta}_1$, conseguimos calcular com uma maior facilidade o parâmetro $\hat{\beta}_0$. Temos assim:

$$\hat{\beta}_1 = \frac{\sum_{i=0}^{n} x_i\, y_i - n\bar{x}\bar{y}}{\sum_{i=0}^{n} x_i^2 - n\, \bar{x}_i^{\,2}}$$

onde:

$\hat{\beta}_1$= é o valor do parâmetro

$\sum_{i=0}^{n} x_i\, y_i$ = somatório do produto entre as variáveis x e y;

$n\bar{x}\bar{y}$= é a multiplicação entre o número de elementos da amostra e a média das variáveis x e y;

$\sum_{i=0}^{n} x_i^2$= é o somatório do quadrado da variável x.

$n\bar{x}_i^{\,2}$= é o quadrado do produto entre o número de elementos e a média da variável x.

$$\hat{\beta}_1 = \frac{250824 - 6(24{,}5)(1278{,}3333)}{4459 - 6(24{,}5)}$$

$$\hat{\beta}_1 = \frac{250824 - 187914{,}9951}{4459 - 147}$$

$$\hat{\beta}_1 = \frac{62909{,}0049}{4312}$$

$$\hat{\beta}_1 = 14{,}589286$$

Sabendo o valor do parâmetro $\hat{\beta}_1$, calculamos na equação 2.5 o valor do parâmetro$\hat{\beta}_0$. Temos assim:

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x}$$

$$\hat{\beta}_0 = 1278{,}3333 - (14{,}589286)(24{,}5)$$

$$\hat{\beta}_0 = 1278{,}3333 - 357{,}437507$$

$$\hat{\beta}_0 = 920{,}895793$$

Agora, como temos os coeficientes $\hat{\beta}_1$e$\hat{\beta}_0$, substituímos os valores na equação $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$, e, a partir disso, calculamos o peso estimado de cada semana. Deste modo, dispomos da equação $\hat{y}_i = 920{,}895793 + 14{,}589286x_i$ para estimar os valores das amostras.

$$\hat{y}_1 = 920{,}895793 + 14{,}589286.7 \quad = \quad 920{,}895793 + 102{,}125002 = 1023{,}020795$$

$$\hat{y}_2 = 920{,}895793 + 14{,}589286.14 \quad = \quad 920{,}895793 + 204{,}250004 = 1125{,}145797$$

$$\hat{y}_3 = 920{,}895793 + 14{,}589286.21 \quad = \quad 920{,}895793 + 306{,}375006 = 1227{,}270799$$

$$\hat{y}_4 = 920{,}895793 + 14{,}589286.28 \quad = \quad 920{,}895793 + 408{,}500008 = 1329{,}395801$$

$$\hat{y}_5 = 920{,}895793 + 14{,}589286.35 \quad = \quad 920{,}895793 + 510{,}620803 = 1023{,}020795$$

$$\hat{y}_6 = 920{,}895793 + 14{,}589286.42 \quad = \quad 920{,}895793 + 612{,}750012 = 1533{,}645805$$

Com os valores previstos calculados, demanda, na sequência, o cálculo para verificar o erro em cada amostra. Como apresentado por Berenson, Krehbiel, Levine e Stephan (2012, p. 498), obtemos o valor a partir de $e_i = y_i - \hat{y}_i$. Ou seja, o erro corresponde a diferenças dos quadrados entre os valores verdadeiros de y_i e os valores previstos de $\hat{y}_i$. Verificamos, assim, o tamanho dos nossos erros.

$$e_1 = y_i - \hat{y}_i = 180 - (1023{,}020795) = -843{,}020795$$

$$e_2 = y_i - \hat{y}_i = 420 - (1125{,}145797) = -705{,}145797$$

$$e_3 = y_i - \hat{y}_i = 862 - (1227{,}270799) = -265{,}270799$$

$$e_4 = y_i - \hat{y}_i = 1474 - (1329{,}395801) = 144{,}604199$$

$$e_5 = y_i - \hat{y}_i = 2074 - (1431{,}520803) = 642{,}479197$$

$$e_6 = y_i - \hat{y}_i = 2660 - (1533{,}645805) = 1126{,}354195$$

Realizando o $\sum_{i=0}^{n} e_i$, obtemos 100,0002. A partir disso, podemos concluir que, com a equação encontrada, não conseguimos estimar o peso de uma produção avícola em relação aos dias que se passaram, pois, como nos mostram Berenson, Krehbiel, Levine e Stephan (2012) ao explicar as propriedades dos quadrados, é necessário que $\sum_{i=0}^{n} e_i = 0$, o que de fato não acontece na situação descrita acima.

4.2 Aviário B: relação entre peso e dias de alojamento.

Neste subcapítulo, aplicaremos a mesma sequência aplicada em 4.1, porém, agora, temos uma coleta de dados com 270 aves, que correspondem aos mesmos 10% do segmento anterior. Na tabela abaixo, estão organizados os dados que serão necessários para aplicarmos a mesma sequência que foi aplicada no segmente 4.1.

Tabela 3 - Aviário B: relação entre peso e dias de alojamento

Dias (X)	Peso (Y)	x.y	x^2	y^2
7	180	1260	49	32400
14	424	5936	196	179776
21	829	17409	441	687241
28	1452	40656	784	2108304
35	2104	73640	1225	4426816
42	2723	114366	1764	7414729
ΣX 147	$(\sum Y)$ 7712	$\Sigma(XY)$ 253267	ΣX^2 4459	ΣY^2 14849266

Fonte: elaborada pelo autor. (2023)

A partir das informações da tabela, calculamos o coeficiente de correlação de Pearson.

$$r_{x,y} = \frac{253267 - \frac{147.7712}{6}}{\sqrt{\left[4459 - \frac{(147)^2}{6}\right] \cdot \left[14849266 - \frac{(7712)^2}{6}\right]}}$$

$$r_{x,y} = \frac{253267 - \frac{1133664}{6}}{\sqrt{\left[4459 - \frac{21609}{6}\right] \cdot \left[14849266 - \frac{59474944}{6}\right]}}$$

$$r_{x,y} = \frac{253267 - 188944}{\sqrt{[4459 - 3601{,}5] \cdot [14849266 - 9912460{,}666]}}$$

$$r_{x,y} = \frac{64323}{\sqrt{857{,}5} \cdot 4936805{,}34}$$

$$r_{x,y} = \frac{64323}{29{,}283 \cdot 4936805{,}34}$$

$$r_{x,y} = \frac{64323}{144564470{,}77122}$$

$$r_{x,y} = \frac{64323}{144564470{,}77122}$$

$$r_{x,y} = 0{,}0004449434$$

Elevando o coeficiente de correlação ao quadrado, obtemos $0{,}000000197974629$ como coeficiente de determinação. Como o coeficiente de determinação é praticamente nulo, podemos concluir que a qualidade do nosso ajuste será praticamente nula. Ou seja, a explicação da variável y, através da variável x, por meio de um modelo de regressão linear, torna-se praticamente inviável.

4.3 Aviário A: relação entre peso e ração consumida.

A tabela abaixo, elaborada no Excel, traz os dados necessários para estudarmos a relação entre as variáveis: peso médio e consumo de ração. Nesta, temos como variável independente (x) o consumo de ração e na variável dependente (y) o peso médio. Uma vez que pressupomos o peso como dependente do consumo de ração.

Tabela 4 - Aviário A: relação entre peso e ração consumida

Consumo de Ração (X)	Peso (Y)	x.y	x²	y²
2775,48	180	499586,4	7703289,23	32400
11936,45	420	5013309	142478838,6	176400
14240,22	862	12275069,64	202783865,6	743044
17879,44	1474	26354294,56	319674374,7	2172676
27282,66	2074	56584236,84	744343536,7	4301476
26713,61	2660	71058202,6	713616959,2	7075600
ΣX 100827,86	$(\sum Y)$ 7670	$\Sigma(XY)$ 171784699	ΣX^2 2130600864	ΣY^2 14501596

Fonte: elaborada pelo autor. (2023)

A partir dos dados da tabela, calculamos o coeficiente de correlação e de determinação.

$$r_{x,y} = \frac{171784699 - \frac{(100827,86).(7670)}{6}}{\sqrt{\left[2130600864 - \frac{(100827,86)^2}{6}\right] \cdot \left[14501596 - \frac{(7670)^2}{6}\right]}}$$

$$r_{x,y} = \frac{171784699 - \frac{773349686,2}{6}}{\sqrt{\left[2130600864 - \frac{10166257352,1796}{6}\right] \cdot \left[14501596 - \frac{58828900}{6}\right]}}$$

$$r_{x,y} = \frac{171784699 - 128891614,4}{\sqrt{[2130600864 - 1694376225,363267]} \cdot [14501596 - 9804816,667]}$$

$$r_{x,y} = \frac{42893084,6}{\sqrt{[436224638,636733]} \cdot [4696779,333]}$$

$$r_{x,y} = \frac{42893084,6}{20885,99144 \cdot [4696779,333]}$$

$$r_{x,y} = \frac{42893084,6}{98096892944,60691}$$

$$r_{x,y} = 0,000437252239$$

Quando elevamos o coeficiente de correlação ao quadrado, obtemos $0{,}0000001911895073$ como coeficiente de determinação. Como é um valor que se inclina ao zero, podemos afirmar que a variável y não tende a ser explicada por x.

4.4 Aviário B: relação entre peso e ração consumida.

Para realizarmos os cálculos neste segmento, aplicaremos a mesma sequência dos itens anteriores. Os dados necessários encontram se na tabela abaixo:

Tabela 5 - Aviário B: relação entre peso e ração consumida

Consumo de Ração (X)	Peso (Y)	x.y	x²	y²
3274,52	180	589413,6	10722481,23	32400
14073,4	424	5967121,6	198060587,6	179776
16789,77	829	13918719,33	281896376,7	687241
21080,56	1452	30608973,12	444390009,9	2108304
32167,34	2104	67680083,36	1034737763	4426816
31496,39	2723	85764669,97	992022583	7414729
ΣX 118881,98	$(\sum Y)$ 7712	$\Sigma(XY)$ 204528981	ΣX^2 2961829801	ΣY^2 14849266

Fonte: elaborada pelo autor. (2023)

$$r_{x,y} = \frac{\Sigma(XY) - \frac{(\Sigma X)(\sum Y)}{n}}{\sqrt{\left[\Sigma X^2 - \frac{(\sum X)^2}{n}\right] \cdot \left[\Sigma Y^2 - \frac{(\Sigma Y)^2}{n}\right]}}$$

$$r_{x,y} = \frac{204528981 - \frac{(118881{,}98)(7712)}{6}}{\sqrt{\left[2961829801 - \frac{(118881{,}98)^2}{6}\right] \cdot \left[14849266 - \frac{(7712)^2}{6}\right]}}$$

$$r_{x,y} = \frac{204528981 - \frac{916817829{,}76}{6}}{\sqrt{\left[2961829801 - \frac{14132925168{,}7204}{6}\right] \cdot \left[14849266 - \frac{59474944}{6}\right]}}$$

$$r_{x,y} = \frac{204528981 - 152802971{,}6266667}{\sqrt{[2961829801 - 2355487528{,}120067]} \cdot [14849266 - 9912490{,}666666667]}$$

$$r_{x,y} = \frac{51726009{,}3733333}{\sqrt{606342272{,}879933} \cdot [4936775{,}333333333]}$$

$$r_{x,y} = \frac{51726009{,}3733333}{24624{,}01821149 \cdot [4936775{,}333333333]}$$

$$r_{x,y} = \frac{51726009{,}3733333}{121563245714{,}0346}$$

$$r_{x,y} = \frac{51726009{,}3733333}{121563245714{,}0346}$$

$$r_{x,y} = 0{,}0004255069$$

Desse modo, obtivemos como coeficiente de determinação $0{,}00000018105618$ que, assim como no caso anterior, tende ao zero. Logo, isso implica que explicar a variável y através de x se torna inviável.

5 CONSIDERAÇÕES FINAIS

A modelagem de situações a partir do uso da regressão linear é um instrumento estatístico, utilizado em diversos ramos para resumir informações e dados. Ou seja, quando usamos este modelo, desejamos prever o comportamento de uma variável endógena através de uma variável explicativa.

Pensando neste contexto, estabelecemos o objetivo principal do nosso estudo, que era observar as variáveis: peso de uma produção avícola com relação aos dias que se passaram e a quantidade de ração consumida em um determinado período de dias. E, assim, analisar se é possível estimar o peso de uma produção avícola através da linearidade estudada pelos modelos de regressão.

Por meio do nosso objetivo principal, chegamos ao nosso problema, que era observar se, de fato, é concebível traçar e modelar a função de uma reta de regressão, realizando a análise gráfica do peso de uma produção avícola, quando ele depende de variáveis como a quantidade de dias e/ou o consumo de ração na produção de frangos de corte. Como hipótese inicial, acreditávamos que seria possível traçar essa reta, já que o peso de uma produção depende de variáveis, como as que seriam estudadas na ocasião.

Realizamos a nossa coleta de dados e, a partir disso, fomos para a análise deles. Nessa análise, o primeiro passo realizado foi calcular o coeficiente de correlação e de determinação, uma vez que esses nos dão a intensidade da relação entre as variáveis. A contar dos coeficientes, conseguimos notar que não existe uma relação forte entre as variáveis, e, como consequência, não teríamos uma qualidade de ajuste nos modelos da regressão. Assim, podemos afirmar que o nosso problema inicial não tem solução com base no estudo dos modelos de regressão linear simples, da mesma forma que a nossa hipótese inicial também foi refutada.

Ademais, o nosso estudo mostrou que não existe uma explicação linear para o crescimento do peso de uma produção avícola. Mas, baseado nesta conclusão, podemos ter uma relação de outra natureza. Isso seria uma motivação para estudos futuros, ou seja, relacionar a variável do peso com um outro modelo de estudo. Uma vez que não obtivemos êxito quando estudamos esse com relação a linearidade das variáveis.

REFERÊNCIAS

BERENSON. L. M; KREHBIEL, C. T; LEVINE, M. D; STEPHAN, F. D. Estatística Teoria e Aplicação. Rio de Janeiro: LTC, 2012.

CHEIN, F. Introdução aos Modelos de Regressão Linear. Brasília-DF: ENAP, 2019.

DACHS, W. N. J; CARVALHO, F. J. Diagnóstico em Regressão. Rio de Janeiro: UFRJ,1984.

MAROCO, J.; Análise Estatística – Com utilização do SPSS; Edições Sílabo; 2003.

RODRIGUES, S. Modelo de Regressão Linear e suas Aplicações. Portugal: UBI, 2012.

SANTOS, C. M. A.; Estatística Descritiva – Manual de auto-aprendizagem; Edições Sílabo; 2007.

VIRGILLITO, B.S. Estatística Aplicada.São Paulo: Edicon,2010.

ANEXOS

Coleta de dados

- Semana 1 – 7 dias

	7 dias Aviário A
Peso	Tamanho da amostra
3620	20
3780	20
3650	20
3620	20
3810	20
3600	20
3530	20
3630	20
3570	20
4080	24
4420	25
41310	**229**

	7 dias Aviário B
Peso	Tamanho da amostra
3650	20
3600	20
3665	20
3540	20
3430	20
3600	20
3330	20
3500	20
3760	20
3800	20
3680	20
3720	20
4260	24
1190	6
48725	**270**

- Semana 2 – 14 dias

	14 Dias Aviário A
Peso (em gramas)	Tamanho da amostra
8830	20
8650	20
8765	20
8495	20
3350	9
8215	20
7445	20
7670	20
8460	20
8355	20
9165	20
8805	20
96205	**229**

	14 Dias Aviário B
Peso (em gramas)	Tamanho da amostra
8825	20
8440	20
8755	20
8455	20
8690	20
7780	20
8325	20
7940	20
8400	20
8595	20
8215	20
8225	20
9060	20
4850	10
114555	**270**

- Semana 3 – 21 dias

	Peso 21 Dias Aviário A
Peso (em gramas)	Tamanho da Amostra
8530	10
8530	10
8000	10
8335	10
8795	10
7935	10
8840	10
8585	10
9285	10
8800	10
8695	10
8645	10
8855	10
9505	10
8725	10
8960	10
8380	10
9265	10
9040	10
7745	10
7745	10
8000	9
8380	10
197575	**229**

	Peso 21 dias Aviário B
Peso (em gramas)	Tamanho da amostra
8000	10
8445	10
7360	10
7535	10
8415	10
7645	10
7925	10
7910	10
8000	10
8710	10
8305	10
8430	10
9232	10
7725	10
8250	10
8865	10
8520	10
8790	10
8345	10
8550	10
8050	10
8455	10
8385	10
8280	10
9000	10
8500	10
8270	10
223897	**270**

- Semana 4 – 28 dias

	Peso 28 dias Aviário A
Peso (em gramas)	Tamanho da amostra
15660	10
15520	10
14210	10
14670	10
14680	10
14870	10
15090	10
15520	10
14270	10
16150	10
14150	10
13680	10
14380	10
13930	10
14440	10
14630	10
12430	9
15720	10
14930	10
14760	10
15500	10
14390	10
14130	10
337710	**229**

	Peso 28 dias Aviário B
Peso (em gramas)	Tamanho da Amostra
14820	10
15000	10
15170	10
15050	10
14130	10
14450	10
14170	10
14170	10
14950	10
14280	10
14120	10
14780	10
14790	10
15100	10
13850	10
14740	10
12840	10
14810	10
14410	10
14310	10
15470	10
14910	10
14060	10
14790	10
13730	10
14380	10
14780	10
392060	**270**

- Semana 5 – 35 dias

	Peso 35 dias Aviário A
Peso (em gramas)	Tamanho da amostra
11660	5
10050	5
10500	5
9780	5
9780	5
11270	5
10930	5
11110	5
9910	5
11090	5
10290	5
10310	5
10670	5
9925	5
11020	5
9960	5
10420	5
11180	5
10850	5
10440	5
10080	5
10300	5
9900	5
10890	5
10060	5
9000	5
9445	5
9830	5
8460	5
10420	5
10520	5
10520	5
10320	5
10020	5
11230	5
10520	5
10000	5
10840	5
11000	5
10920	5
10580	5
10160	5
10300	5
10160	5
10300	5
8050	4
474970	**229**

	Peso 35 dias Aviário B
Peso (em gramas)	Tamanho da amostra
9500	5
10280	5
10400	5
10040	5
10680	5
10610	5
9500	5
10330	5
10500	5
10000	5
10500	5
10600	5
10130	5
11760	5
10920	5
10700	5
10350	5
9965	5
10230	5
10650	5
11000	5
11040	5
10460	5
10330	5
10610	5
10270	5
9070	5
10740	5
9735	5
11370	5
9610	5
10190	5
10500	5
10610	5
11400	5
9860	5
10760	5
10600	5
9910	5
11230	5
10760	5
10860	5
11620	5
10670	5
11000	5
9870	5
10810	5
9500	5
11020	5
10710	5
11020	5
11270	5
11550	5
10500	5
568100	**270**

- Semana 6 – 42 dias

	Peso 42 dias Aviário A
Peso (em gramas)	Tamanho da Amostra
12570	5
12820	5
13730	5
13500	5
13050	5
11870	5
13560	5
12590	5
11940	5
13390	5
13540	5
13890	5
12630	5
12890	5
12900	5
13730	5
13560	5
14290	5
12340	5
11900	5
14300	5
14040	5
12070	5
14020	5
14210	5
12820	5
14370	5
13560	5
13250	5
13370	5
13330	5
14370	5
12910	5
13320	5
13590	5
12260	5
13360	5
14930	5
12980	5
12760	5
13760	5
14130	5
13830	5
14610	5
9845	4
12680	5
609365	**229**

	Peso 42 dias Aviário B
Peso (em gramas)	Tamanho da Amostra
13500	5
15270	5
13080	5
13790	5
12210	5
12950	5
13920	5
13500	5
14630	5
14450	5
12610	5
14010	5
14330	5
11790	5
13500	5
14430	5
13830	5
13030	5
13930	5
11720	5
12180	5
14740	5
13420	5
14120	5
13600	5
13420	5
13840	5
12440	5
13980	5
14000	5
13150	5
13750	5
13390	5
14160	5
13310	5
13790	5
14140	5
13280	5
14440	5
13780	5
13500	5
12610	5
13530	5
12940	5
14600	5
13730	5
13560	5
14410	5
13950	5
14500	5
13230	5
13520	5
13850	5
13940	5
735280	**270**