

**COMPARAÇÃO DE GENÓTIPOS DE FEIJÃO-DE-VAGEM POR MEIO  
DE CURVAS DE CRESCIMENTO, UTILIZANDO IDENTIDADE DE  
MODELOS LINEARES DE REGRESSÃO**

**JOSÉ CARLOS DE OLIVEIRA JÚNIOR**

**UNIVERSIDADE ESTADUAL DO NORTE FLUMINENSE  
DARCY RIBEIRO**

**CAMPOS DOS GOYTACAZES – RJ  
ABRIL – 2014**

## **FICHA CATALOGRÁFICA**

Preparada pela Biblioteca do **CCTA / UENF** 061/2014

Oliveira Júnior, José Carlos de

Comparação de genótipos de feijão-de-vagem por meio de curvas de crescimento, utilizando identidade de modelos lineares de regressão / José Carlos de Oliveira Júnior. – 2014.

60 f.

Orientador: José Tarcísio Lima Thiébaud

Dissertação (Mestrado - Produção Vegetal) – Universidade Estadual do Norte Fluminense Darcy Ribeiro, Centro de Ciências e Tecnologias Agropecuárias. Campos dos Goytacazes, RJ, 2014.

Bibliografia: f. 43 – 49.

1. Regressão linear 2. Identidade de modelos 3. Variável Dummy  
4. Modelos lineares 5. Feijão-de-vagem I. Universidade Estadual do Norte Fluminense Darcy Ribeiro. Centro de Ciências e Tecnologias Agropecuárias. II. Título.

CDD – 635.651

**COMPARAÇÃO DE GENÓTIPOS DE FEIJÃO-DE-VAGEM POR MEIO  
DE CURVAS DE CRESCIMENTO, UTILIZANDO IDENTIDADE DE  
MODELOS LINEARES DE REGRESSÃO**

**JOSÉ CARLOS DE OLIVEIRA JÚNIOR**

Dissertação apresentada ao Centro de Ciências e Tecnologias Agropecuárias da Universidade Estadual do Norte Fluminense Darcy Ribeiro, como parte das exigências para obtenção do título de Mestre em Produção Vegetal

Orientador: Prof. José Tarcísio Lima Thiébaud

**CAMPOS DOS GOYTACAZES – RJ  
ABRIL – 2014**

**COMPARAÇÃO DE GENÓTIPOS DE FEIJÃO-DE-VAGEM POR MEIO  
DE CURVAS DE CRESCIMENTO, UTILIZANDO IDENTIDADE DE  
MODELOS LINEARES DE REGRESSÃO**

**JOSÉ CARLOS DE OLIVEIRA JÚNIOR**

Dissertação apresentada ao Centro de Ciências e Tecnologias Agropecuárias da Universidade Estadual do Norte Fluminense Darcy Ribeiro, como parte das exigências para obtenção do título de Mestre em Produção Vegetal

Aprovada em 29 de abril de 2014

Comissão Examinadora

---

Prof. Ernany Santos Costa (D.Sc., Produção Vegetal) – IFF – Câmpus Bom Jesus

---

Prof. Rogério Figueiredo Daher (D. Sc. em Produção Vegetal) – UENF

---

Prof. Geraldo de Amaral Gravina (D. Sc. em Fitotecnia) – UENF  
(Coorientador)

---

Prof. José Tarcísio Lima Thiébaud (D. Sc. em Ciência Animal) – UENF  
(Orientador)

Dedico este trabalho aos meus pais José Carlos (in memorian) e Ana Emilia por tudo que fizeram por mim, à minha esposa Rejane pelo incentivo, compreensão e paciência e aos meus filhos Fernanda e Gustavo.

## AGRADECIMENTOS

Ao Instituto Federal Fluminense Campus Bom Jesus do Itabapoana juntamente com os colegas professores Lanusse, Cavichini, Sebastião Ney, Antônio Alonso, Clinimar, Kleberon e Juarez, pelo apoio, amizade e ajuda para a realização deste trabalho;

Ao meu orientador e amigo, Professor José Tarcísio Lima Thiébaud, pelos ensinamentos, competência, paciência e incentivo durante todas as fases da realização dessa dissertação;

Aos Professores Rogério Figueiredo Daher e Ernany Santos Costa, por aceitarem fazer parte desta banca e pelas contribuições dadas a este trabalho;

Ao coorientador professor Geraldo de Amaral Gravina, pela compreensão, amizade e orientação para realização deste trabalho;

A FAPERJ pelo auxílio financeiro ao desenvolvimento do projeto;

Agradeço à Universidade Estadual do Norte Fluminense Darcy Ribeiro, ao Programa de Pós-Graduação em Produção Vegetal e ao Laboratório de Engenharia Agrícola, pela oportunidade de realização deste curso;

E a todos que de alguma maneira contribuíram para a realização deste trabalho.

## SUMÁRIO

RESUMO .....	vi
ABSTRACT .....	viii
1. INTRODUÇÃO .....	1
1.1. Objetivos Gerais .....	2
1.2. Objetivos específicos .....	2
2. REVISÃO DE LITERATURA .....	3
2.1. O feijoeiro .....	3
2.2. A regressão linear .....	7
2.2.1. Análise de variância da regressão .....	9
2.2.2. Coeficiente de determinação .....	11
2.2.3. Notação matricial .....	12
2.3. O teste de Scott-Knott .....	21
2.4. Identidade de modelos .....	23
2.4.1. O teste de Chow .....	23
2.4.2. Variável Dummy .....	24
2.5. Seleção da melhor equação .....	27
3. MATERIAL E MÉTODOS .....	30
3.1 Histórico dos experimentos no programa de melhoramento do feijão- vagem da UENF .....	30

3.2. Localização e caracterização da área experimental .....	32
3.3. Amostragem simples ao acaso .....	32
4. RESULTADOS E DISCUSSÃO .....	35
5. CONCLUSÕES .....	44
6. REFERÊNCIAS BIBLIOGRÁFICAS .....	46
7. APÊNDICE .....	53



## RESUMO

OLIVEIRA JÚNIOR, J. C. de, M.Sc., Universidade Estadual do Norte Fluminense Darcy Ribeiro, abril de 2014. Comparação de genótipos de feijão-de-vagem por meio de curvas de crescimento, utilizando identidade de modelos lineares de regressão. Orientador: Prof. José Tarcísio Lima Thiebaut, D. Sc. Coorientador: Prof. Geraldo de Amaral Gravina, D. Sc.

Tendo em vista a importância econômica e social do feijão-de-vagem nas regiões Norte e Noroeste Fluminense, o objetivo deste trabalho foi a avaliação de 17 genótipos de feijão-de-vagem por meio de curvas de crescimento utilizando modelos de regressão linear e proceder a comparação dos genótipos, empregando a identidade de modelos. Os dados foram coletados em um experimento de campo e foi conduzido no campus Bom Jesus do Itabapoana do Instituto Federal Fluminense, seguindo as recomendações para a cultura. Foram amostradas ao acaso 20 plantas de feijão-de-vagem de cada um dos 17 genótipos avaliados. Foi avaliada a altura de cada planta amostrada a partir de 12 dias após a semeadura, com intervalos de quatro dias entre as avaliações, num total de nove avaliações. Foram ajustados modelos de regressão linear simples para todas as plantas amostradas, sendo a altura a variável dependente e o tempo em dias após a primeira avaliação como variável independente. A seguir, foi realizada análise amostral para os estimadores dos parâmetros (intercepto e declividade) para todas as plantas amostradas. O genótipo 15 foi eliminado devido a valor muito alto do coeficiente de variação para a declividade

(CV=35,378%). Realizou-se o teste de agrupamento de Scott-Knott a nível de 5% de probabilidade para os estimadores dos parâmetros dos dezessete genótipos, obtendo-se 5 grupos para a declividade e 6 para o intercepto. Após, formaram-se 5 grupos que pertenciam simultaneamente aos grupos formados pela declividade e intercepto, e um sexto grupo formado pelos genótipos não pertencentes a nenhum dos grupos formados anteriormente. Foi realizada a análise de identidade de modelos com uso das variáveis Dummy, testando os estimadores dos parâmetros com o teste “t” de Student a nível de 5% de probabilidade. Para o grupo 3, formado pelos genótipos 03 e 12 e o grupo 5, pelos genótipos 06, 07 e 11, não existiram diferenças estatísticas significativas a nível de 5% de probabilidade entre as equações dos genótipos pertencentes a cada grupo. Para todos os outros grupos formados, não houve diferença significativa entre os interceptos dos genótipos pertencentes a cada grupo, porém, diferiram em relação à declividade a nível de 5% de probabilidade.

## ABSTRACT

OLIVEIRA JÚNIOR, J. C. de, M.Sc., Universidade Estadual do Norte Fluminense Darcy Ribeiro, april, 2014. Comparison of growth curves of snap beans, using identity linear regression models. Advisor: Prof. José Tarcísio Lima Thiebaut, D. Sc. Coadvisor: Prof. Geraldo de Amaral Gravina, D. Sc.

Due to the economic and social importance of snap bean in the North and Northwest of Rio de Janeiro, the objective of this study was the evaluation of 17 genotypes of snap bean by means of growth curves using linear regression models and carry out comparison of genotypes, using the model identity. Data were collected in a field experiment where was conducted at the Fluminense Federal Institute campus Bom Jesus do Itabapoana, following the recommendations for this culture. 20 plants pole beans each of the 17 genotypes were sampled at random. The height of each sampled plant was assessed from 12 days after sowing, at intervals of four days between assessments, a total of nine reviews. Simple linear regression models were adjusted for all plants sampled, the height the dependent variable and time in days after the first review as an independent variable. The following sample analysis was performed for the estimators of the parameters (intercept and slope) for all plants sampled. The genotype 15 was eliminated due to a very high coefficient of variation for the slope (CV= 35.378%). We carried out the grouping of the Scott-Knott 5% level of probability for the estimators of the parameters of the seventeen genotypes

testing, obtaining groups 5 and 6 for the slope to the intercept. After, 5 groups were formed belonging to both groups formed by the slope and intercept, and a sixth group formed by genotypes not belonging to any of the previously formed groups. Analysis model identity with the use of Dummy variables was performed by testing the estimators of the parameters with the "t" test at the 5% probability test. In group 3, formed by genotypes 3 and 12 and the group 5 by genotypes 06, 07 and 11 there were no significant at the 5% probability equations between the genotypes in each group statistical differences. For all other groups formed, there was no significant difference between the intercepts of the genotypes in each group, however, differed in the slope at the 5% probability.

## 1. INTRODUÇÃO

O feijão-vagem, feijão-de-vagem ou simplesmente vagem, é planta anual, muito difundida nas diversas regiões brasileiras. Trata-se de uma fabácea (leguminosa), hortaliça da mesma espécie botânica do feijoeiro-comum (*Phaseolus vulgaris* L.) e se caracteriza por ser colhida quando as vagens estão verdes e tenras (Filgueira, 2008).

As principais cultivares apresentam hábito de crescimento indeterminado, podendo alcançar mais de 2,50 m de altura, o que exige o tutoramento. As vagens apresentam seção circular (grupo macarrão) ou chatas (grupo manteiga), sendo este grupo o de maior volume de comercialização (Filgueira, 2008).

É produzido principalmente por pequenos agricultores (agricultores familiares) em todos os estados brasileiros, sendo quase a totalidade de sua produção destinada ao consumo “*in natura*” (Viggiano, 1990).

Santos (2013) destaca a necessidade da pesquisa com feijão-vagem, devido a sua importância econômica e social, pois, é cultivado principalmente por pequenos produtores.

A análise de regressão é um método que procura estabelecer uma relação funcional entre duas ou mais variáveis (uma variável resposta e uma ou mais variáveis explicativas), ou seja, determinar um modelo que descreva melhor da melhor maneira a resposta das variáveis a serem analisadas (Vanni, 1998; Draper e Smith, 1998).

Na análise estatística de experimentos em que os tratamentos são em níveis crescentes de um fator quantitativo, como níveis de adubação, tempo, altura, idade, temperatura ou concentrações de agrotóxicos, o efeito do tratamento deve ser avaliado, em geral, por meio de uma análise de regressão, pois, neste caso, o uso de procedimentos para comparação múltiplas (testes de médias) não é o mais indicado (Elían, 1988; Banzzato e Kronka 1995; Regazzi, 1999).

Com muita frequência, estuda-se a relação funcional entre as variáveis Y e X. Alguns problemas têm aplicações importantes, como determinar se um conjunto de curvas tem intercepto comum, se são paralelas, se são idênticas ou não (Regazzi, 1993).

### **1.1. Objetivos Gerais**

Avaliação de 17 genótipos de feijão-vagem por meio de curvas de crescimento, utilizando modelos de regressão linear e proceder a comparação dos genótipos empregando a identidade de modelos.

### **1.2. Objetivos específicos**

- Fazer a regressão linear de 1<sup>o</sup> grau para todas as amostras;
- Análise amostral para o intercepto e a declividade das equações de regressão linear para cada genótipo;
- Agrupar os genótipos em relação ao intercepto e a inclinação, utilizando o teste de Scott-Knott;
- Realizar análise de identidade de modelos para os genótipos agrupados pelo teste de Scott-knott, utilizando o método das variáveis Dummy.

## 2. REVISÃO DE LITERATURA

### 2.1. O feijoeiro

O feijão-de-vagem (*Phaseolus vulgaris* L.) é uma fabaceae que possui distribuição cosmopolita, incluindo cerca de 650 gêneros e, aproximadamente, 18000 espécies, representando uma das maiores famílias de Angiospermas e também uma das principais do ponto de vista econômico. No Brasil, ocorrem cerca de 175 gêneros e 1500 espécies (Souza e Lorenzi, 2008).

Cronquist (1988) classifica o feijão-de-vagem como pertencente à subclasse Rosidae, ordem Fabales e família Fabaceae. Suas espécies, especialmente o feijão comum (*Phaseolus vulgaris* L.), são amplamente distribuídas no mundo todo, cultivadas nos trópicos e nas regiões temperadas dos hemisférios Norte e Sul.

Segundo Peixoto et al. (2002), da mesma forma que o feijão comum, o feijão-de-vagem é uma cultura que tem uma boa adaptação a climas quentes e amenos, dentro de uma faixa térmica (18°C a 30°C). Em temperaturas superiores a 35°C, há queda de flores e deficiência de polinização, resultando em vagens deformadas e ocasionando queda da produtividade. Temperaturas abaixo de 15°C retardam o bom desenvolvimento das plantas, além de ser um fator limitante à cultura, por promover maior incidência de ferrugem.

Quanto ao efeito de fotoperiodismo, o feijão-de-vagem responde indiferente, podendo produzir em dias longos ou curtos (Filgueira, 2008).

A temperatura do solo é um dos fatores que pode afetar a germinação e, conseqüentemente, a emergência das plântulas. Pode ser o fator mais importante, tendo em vista que na maioria das vezes o produtor não tem total controle sobre ela. Temperaturas muito baixas ou muito altas podem alterar tanto a velocidade quanto a porcentagem final de germinação. Em geral, as temperaturas baixas reduzem a velocidade de germinação, enquanto as altas aumentam. Em condições extremas de temperatura, a germinação poderá não ocorrer, e, em alguns casos, a semente é levada à condição de dormência. A temperatura ótima para a germinação do feijão-de-vagem está entre 20 e 30°C, sendo 16°C o valor mínimo e 35°C o valor máximo (Nascimento, 2012).

Segundo Filgueira (2008), em regiões de baixa altitude, com inverno ameno é possível a semeadura ao longo do ano possibilitando a comercialização na entressafra. Em regiões de maior altitude, as semeaduras devem ser evitadas no inverno, sendo uma das alternativas o plantio em casa de vegetação em rotação com outras culturas.

As maiores produtividades de vagens ocorrem em solos de textura média, não compactados, profundos, férteis, ricos em matéria orgânica, com boa drenagem e disponibilidade de água em todo o seu desenvolvimento, sendo que aqueles excessivamente argilosos são menos indicados.

A quase totalidade das fabáceas, não tolera alta acidez no solo, produzindo melhor na faixa de pH 5,6 a 6,8. Em solos mais ácidos, a calagem é benéfica, elevando-se a saturação por base para 70%, procurando-se atingir pH 6,5, fornecendo cálcio um macronutriente altamente exigido pela cultura (Filgueira, 2008).

De acordo com Filgueira (2008), experimentos conduzidos no Brasil evidenciam resultados acentuados à aplicação de fósforo e cálcio, não havendo resposta à aplicação de potássio em termos de produtividade. O nitrogênio é o segundo nutriente mais exigido pelas hortaliças. Seu fornecimento via adubação na cultura tutorada é fracionada em três aplicações, sendo 20 kg/ha por vez, a primeira por ocasião do desbaste das plântulas; no início da floração e durante o período produtivo, favorecendo colheita contínua e produtividade.

A recomendação de adubação do feijão-vagem, de acordo com a disponibilidade dos nutrientes e a textura do solo, pode ser vista no quadro 1 (Carrijo et al, 1999b).



Quadro 1 - Recomendação de adubação mineral NPK/ano para o feijão-vagem

Disponibilidade de P ou de K	Textura do Solo			Dose Total	
	Argilosa	Média	Arenosa	K <sub>2</sub> O	N
	-----Dose P <sub>2</sub> O <sub>5</sub> -----				
-----kg.ha <sup>-1</sup> -----					
Baixa	280	230	180	120	150
Média	230	180	130	90	150
Boa	180	130	80	60	150
Muito boa	130	80	50	30	150

Segundo Carrijo et al. (1999a), o parcelamento da adubação recomendada no quadro 1 é feito da seguinte maneira: plantio – aplicar 30% do nitrogênio, 50% do potássio e todo o fósforo; cobertura – parcelar em duas aplicações o restante do nitrogênio (70%) e do potássio (50%), aos 30 e aos 60 dias da emergência das plântulas. A adubação orgânica deverá ser feita utilizando 10 t/ha de esterco de curral bem curtido ou se em rotação com o tomate ou pepino, não há necessidade.

As classes de interpretação para a disponibilidade de fósforo e potássio podem ser avaliadas de acordo com o quadro 2 (Alvarez V., V. H. et al,1999).

O cultivo de feijão-de-vagem é realizado por semeadura direta em sulcos ou cova feita manualmente ou com semeadora de tração mecânica ou manual, sendo intolerante ao transplântio (Filgueira, 2008).

De acordo com Barbosa Filho e Silva (2001), semeiam-se duas a três sementes em cada cova ou no sulco de plantio, no espaçamento de 1,0 a 1,2 m entre linhas e 0,2 a 0,5 m entre plantas, para as cultivares de crescimento indeterminado. Para as cultivares de crescimento determinado, recomenda-se espaçamento menor, variando de 0,5 a 0,9 m entre linhas e 0,15 a 0,5 m entre plantas. A profundidade de semeadura oscila entre 4 e 7 cm, conforme a textura do terreno, variando de argiloso a arenoso, respectivamente.

Quadro 2 - Classes de interpretação da disponibilidade para o fósforo de acordo com o teor de argila do solo e para o potássio

Característica	Classificação				
	Muito baixo	Baixo	Médio	Bom	Muito bom
	-----( $\text{mg.dm}^{-3}$ )-----				
Argila (%)	Fósforo disponível (P)				
60 – 100	$\leq 2,7$	2,8 – 5,4	5,5 – 8,0	8,1 – 12,0	$> 12,0$
35 – 60	$\leq 4,0$	4,1 – 8,0	8,1 – 12,0	12,1 – 18,0	$> 18,0$
15 – 35	$\leq 6,6$	6,7 – 12,0	12,1 – 20,0	20,1 – 30,0	$> 30,0$
0 – 15	$\leq 10,0$	10,1 – 20,0	20,1 – 30,0	30,1 – 45,0	$> 45,0$
	Potássio disponível (K)				
	$\leq 15$	16 – 40	41 – 70	71 – 120	$> 120$

A cultura tutorada é mais exigente em tratamentos culturais e, conseqüentemente, em mão-de-obra, tendo custo mais elevado. O primeiro trato cultural é o desbaste manual, deixando-se apenas de duas a três plantas selecionadas, no espaçamento preestabelecido, nas cultivares de porte alto; para as anãs, apenas uma planta é recomendada. Na cultura rasteira deixam-se plantas isoladas, corretamente espaçadas. O tutoramento é um trato necessário à planta de crescimento indeterminado. Pode-se tutorar o feijão-de-vagem com bambu, mourões, barbante, arame e ráfia. O mais utilizado é em cerca cruzada, o mesmo usado pelos tomaticultores, porém com bambu e mourões mais finos. Não são necessários os amarrios, pois o caule tipicamente volúvel da planta cresce contornando o suporte. As capinas são realizadas manualmente, ou por meio de cultivadores, conforme as necessidades (Filgueira, 2008).

Essa é uma hortaliça menos exigente em água, sendo que na fase inicial, procura-se manter 80% de água útil no solo junto às raízes durante o outono-inverno. A maior exigência hídrica é do início da floração até o término da colheita, obtendo-se maior produtividade e vagens de melhor qualidade com fornecimento de água adequado (Filgueira, 2008).

Nas cultivares de crescimento indeterminado as colheitas iniciam-se aos 60-70 dias após a semeadura e são frequentes (até três vezes por semana) prolongando-se por 30 dias ou mais, dependendo do estado nutricional e

fitossanitário. As colheitas são sempre do tipo manual, muito trabalhosa e de baixo rendimento, entretanto, colheitas frequentes elevam a produtividade, por não deixar vagens “passadas” nas plantas.

Crescimento é um termo quantitativo relacionado principalmente a mudança de tamanho que ocorre nos seres vivos, sendo consequência da distensão e divisão celular (Taiz e Zeiger, 2004).

O crescimento pode ser influenciado por fatores endógenos (internos) como a genética ou exógenos (externos) como a disponibilidade e qualidade da água, intensidade de luz, fotoperíodo, disponibilidade de nutrientes, temperatura, dentre outros (Taiz e Zeiger, 2004).

## 2.2. A regressão Linear

Sempre é interessante conhecer o efeito que algumas variáveis exercem ou que parecem exercer sobre outras. Mesmo quando não existe uma relação causal entre elas. Quando existem duas ou mais variáveis, podemos relacioná-las por meio de uma expressão matemática, denominada modelo matemático (Charnet et al, 2008).

Tal relação pode ser representada por

$$y = f(x_1, x_2, x_3, \dots, x_n) \quad (01)$$

Em que  $y$  representa a variável dependente e os  $x_i$  representa as variáveis independentes.

Considerando duas variáveis  $x$  e  $y$ , relacionadas por uma função matemática  $y = f(x)$ . Se colocarmos os pontos do conjunto de valores  $x_i$  ( $i = 1, 2, \dots, n$ ) e os correspondentes valores de  $y_i$  ( $i=1, 2, \dots, n$ ) em um gráfico, verificamos que eles pertencem à curva que representa o modelo matemático, figura 1 (Hoffmann e Vieira, 1977).

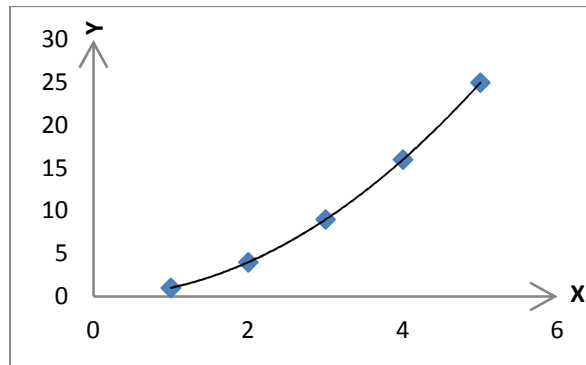


Figura 1 - Curva representando o modelo matemático.

No modelo estatístico, a relação pode ser representada por

$$y = f(x) + \varepsilon_i \quad (02)$$

Em que  $\varepsilon_i$  é o erro aleatório, que pode ser considerado como o efeito resultante de várias características que não estão explícitas no modelo

Numa relação estatística, os pontos não estão necessariamente sobre a curva da relação. (figura 2)

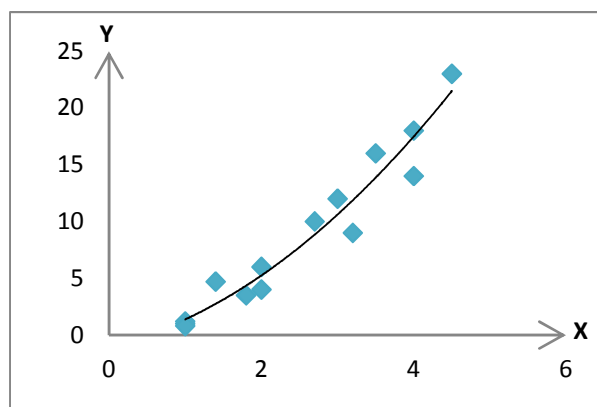


Figura 2 - Curva representando o modelo estatístico.

Quando se tem uma única variável independente o modelo estatístico é da forma (Graybill, 1976):

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad (03)$$

Em que:

$y_i$ : variável resposta aleatória (dependente)

$x_i$ : constante (variável independente)

$\beta_0$  e  $\beta_1$ : parâmetros

$\varepsilon_i$ : erro aleatório

O modelo de regressão linear simples é dito linear nos parâmetros e na variável independente. É linear nos parâmetros porque esses não aparecem como expoente ou multiplicando ou dividido por outro parâmetro. É linear na variável independente porque essa variável é de primeira ordem (Azevedo, 1997; Gujarati, 2011)

Sendo o valor esperado  $y$  conhecido como função de regressão, sendo:

$$E(y_i) = E(\beta_0 + \beta_1 x_i + \varepsilon_i) = \beta_0 + \beta_1 x_i \quad (04)$$

Pois,  $E(\varepsilon_i) = 0$ , e o estimador de  $y_i$  igual a  $\hat{y}_i$ , com os respectivos estimadores dos parâmetros.

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i \quad (05)$$

### 2.2.1. A análise de variância da regressão

Para se realizar a análise de variância e de regressão alguns pressupostos devem ser atendidos, são eles:

a) Aditividade – é um dos pressupostos mais importante. Sua violação leva a inadequação do modelo aos dados ou dos dados ao modelo. A aditividade pode ser testada pelo método de Tukey, o qual desdobra a soma de quadrados do erro em um grau de liberdade para a não aditividade e  $n-1$  graus de liberdade para o resíduo da não aditividade, onde  $n$  são os graus de liberdade do resíduo. O quadrado médio da não aditividade é testado contra o resíduo da não aditividade,

pelo teste F. A não aditividade pode revelar a necessidade de transformação de dados, sendo indicada a transformação logarítmica caso os modelos sejam multiplicativos (Resende, 2007)

b) Normalidade – A não normalidade dos dados e dos erros é restritiva apenas em situações muito drásticas. O teste F de Fisher-Snedecor é robusto a pequenos desvios de normalidade, apresentando resultados satisfatórios quando se verificam pequenos desvios na normalidade (Resende, 2007).

Existem diversos testes para verificar a normalidade dos dados ou erros. Podemos destacar os testes de Kolmogorov-Smirnov, Shapiro-Wilk, Lilliefors e de aderência do qui-quadrado, ou análise gráfica dos resíduos (Bussab, 1988; Barbin, 2003; Ferreira, 2009).

Os resultados desses testes são aproximados quando se utiliza os resíduos padronizados, pois esses não representam uma amostra aleatória (Souza, 1998)

c) Independência - Apesar de existir um teste formal (teste de Durbin-Watson) para se avaliar a ocorrência de correlação entre os erros, é importante salientar que tal correlação resulta da não aleatoriedade na obtenção da amostra de dados.

A independência dos erros é até certo ponto, garantida pelo princípio da casualização e, pode ser verificada pela análise gráfica da dispersão dos resíduos.

d) Homocedasticidade - A violação de qualquer das outras pressuposições pode conduzir a heterogeneidade de variância dos erros.

A homogeneidade pode ser verificada com base em análises de resíduos, gráficos (verificar a correlação entre a média e a variância dos tratamentos) e testes (Barbin, 1993).

O balanceamento do experimento, ou seja, a adoção de números iguais de repetição por tratamento é um grande passo na proteção contra os efeitos da heterogeneidade de variâncias.

Existem vários testes para inferência sobre a existência ou não de homogeneidade de variâncias, como o de Bartlett, o de Hartley, o de Cochran e o de Levene (Barbin, 2003; Vieira, 1999).

No caso em que a heterogeneidade da variância é comprovada, têm-se as opções de padronização e transformação dos dados (Resende, 2007).

Antes de realizar a análise de variância da regressão, devemos realizar a análise de variância para tratamento, somente se os tratamentos forem significativos, ou seja, existir diferenças significativas entre os tratamentos.

Definindo:

$\sum_{i=1}^n (y_i - \bar{y}_i)^2$  é soma de quadrados de desvios Total – SQTotal, representa

a variação de y em torno da média;

$\sum_{i=1}^n (y_i - \hat{y}_i)^2$  é soma de quadrados do Resíduo – SQRes, representa a

variação de y em torno da reta;

$\sum_{i=1}^n (\hat{y}_i - \bar{y}_i)^2$  é soma de quadrados de regressão – SQReg, representa a

variação das esperanças de y em torno de sua média.

### 2.2.2. Coeficiente de determinação – $R^2$

O coeficiente de determinação é uma medida descritiva da qualidade do ajustamento obtido. Representa a proporção da variação em y que é explicada pela variável x na equação de regressão.

Sendo definido por:

$$R^2 = \frac{SQReg}{SQTotal} \quad (06)$$

O coeficiente de determinação calculado anteriormente é muito influenciado pelo tamanho da amostra, amostras pequenas tendem a apresentar

$R^2$  alto. Neste sentido, torna-se necessário a correção do  $R^2$  para graus de liberdade, e definido como (Kutner et al, 2005):

$$\bar{R}^2 = R^2 - \frac{1}{n-2}(1-R^2) \quad (07)$$

$$\rho(\widehat{X}, \widehat{Y}) = \widehat{\beta}_1 \left[ \frac{S_{XX}}{S_{XY}} \right]^{\frac{1}{2}} \quad (08)$$

$$\rho(\widehat{X}, \widehat{Y})^2 = \widehat{\beta}_1^2 \left[ \frac{S_{XX}}{S_{XY}} \right] = \frac{SQReg}{SQTotal} \quad (09)$$

### 2.2.3. A notação matricial

Os modelos de regressão linear ou modelos lineares podem ser escritos da seguinte forma:

$$E(Y) = X\beta \quad (10)$$

ou,

$$Y = X\beta + \varepsilon \quad (11)$$

Onde:

$Y$  = vetor de observações ( $n \times 1$ );

$$Y_{n \times 1} = \begin{bmatrix} Y_1 \\ Y_2 \\ \dots \\ Y_n \end{bmatrix} \quad (12)$$



$X$  = matriz de delineamento ou matriz de constantes conhecidas ( $n \times p + 1$ );

$\beta$  = vetor de parâmetros desconhecidos ( $p + 1 \times 1$ ); e

$\varepsilon$  = vetor de erros ( $n \times 1$ )

O modelo linear geral pressupõe que  $\varepsilon \sim NID(\Phi; I\sigma^2)$  (erro é normalmente distribuído com média  $\Phi$ , e variância  $I\sigma^2$ ), denominado modelo linear de Gauss-Markov normal (Searle, 1971).

Para o modelo linear com uma variável independente, temos:

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_{1i} \quad (13)$$

Supondo  $n$  observações, obtemos:

$$\begin{aligned} Y_1 &= \beta_0 + \beta_1 X_{11} + \varepsilon_1 \\ Y_2 &= \beta_0 + \beta_1 X_{12} + \varepsilon_2 \\ &\vdots \\ Y_n &= \beta_0 + \beta_1 X_{1n} + \varepsilon_n \end{aligned} \quad (14)$$

Que na forma matricial pode ser representado como (Seber e Lee, 2003):

$$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \dots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} \\ 1 & X_{12} \\ 1 & X_{13} \\ \vdots & \vdots \\ 1 & X_{1n} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \quad (15)$$

Para o modelo polinomial de 2º grau com 1 variável independente, temos:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i}^2 + \varepsilon_i \quad (16)$$

Obtendo as seguintes observações:

$$\begin{aligned}
Y_1 &= \beta_0 + \beta_1 X_{11} + \beta_2 X_{21}^2 + \varepsilon_1 \\
Y_2 &= \beta_0 + \beta_1 X_{12} + \beta_2 X_{22}^2 + \varepsilon_2 \\
&\vdots \\
Y_n &= \beta_0 + \beta_1 X_{1n} + \beta_2 X_{2n}^2 + \varepsilon_n
\end{aligned} \tag{17}$$

Representado na forma matricial como (Cecon et al, 2012):

$$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \dots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{21}^2 \\ 1 & X_{12} & X_{22}^2 \\ 1 & X_{13} & X_{23}^2 \\ \vdots & \vdots & \vdots \\ 1 & X_{1n} & X_{2n}^2 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \tag{18}$$

Considerando o modelo de regressão com  $p$  variáveis independentes ou regressoras (Weisberg, 2005):

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_p X_{pi} + \varepsilon_i \tag{19}$$

Supondo  $n$  observações, em que  $n > p + 1$ , obtêm-se as seguintes observações:

$$\begin{aligned}
Y_1 &= \beta_0 + \beta_1 X_{11} + \beta_2 X_{21} + \dots + \beta_p X_{p1} + \varepsilon_1 \\
Y_2 &= \beta_0 + \beta_1 X_{12} + \beta_2 X_{22} + \dots + \beta_p X_{p2} + \varepsilon_2 \\
&\vdots \\
Y_n &= \beta_0 + \beta_1 X_{1n} + \beta_2 X_{2n} + \dots + \beta_p X_{pn} + \varepsilon_n
\end{aligned} \tag{20}$$

Pode-se representar tais equações na forma matricial como (Rawlings et al, 1998):

$$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \dots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{21} & \dots & X_{p1} \\ 1 & X_{12} & X_{22} & \dots & X_{p2} \\ 1 & X_{13} & X_{23} & \dots & X_{p2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{1n} & X_{2n} & \dots & X_{p2} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \tag{21}$$

O qual é obtido segundo o modelo a seguir, pela estimação do vetor  $\beta$ , obtenção de  $\hat{\beta}$ .

$$\hat{Y} = X \hat{\beta} \quad (22)$$

Minimizando os erros, procura-se a solução que minimiza a norma do vetor  $\varepsilon$ .

$$Y = X\beta + \varepsilon \Rightarrow \varepsilon = Y - X\beta \quad (23)$$

Assim:

$$\begin{aligned} Z &= \|\varepsilon\|^2 = \|Y - X\beta\|^2 \\ Z &= (Y - X\beta)'(Y - X\beta) \\ Z &= Y'Y - Y'X\beta - \beta'X'Y + \beta'X'X\beta \end{aligned} \quad (24)$$

Como as matrizes  $Y'X\beta$  e  $\beta'X'Y$  de dimensão 1x1 (escalar) e sendo a transposta do escalar o próprio escalar, pode-se reescrever:

$$Z = Y'Y - \beta'X'Y - \beta'X'Y + \beta'X'X\beta \quad (25)$$

$$Z = Y'Y - 2\beta'X'Y + \beta'X'X\beta \quad (26)$$

Como  $Z$  é uma função quadrática com ponto de mínimo, derivando  $Z$  em função de  $\beta$ .

$$\frac{dZ}{d\beta} = -2(d\beta')X'Y + (d\beta')X'X\beta + \beta'X'X(d\beta) \quad (27)$$

Como  $(d\beta')X'X\beta$  e  $\beta'X'X(d\beta)$  são escalares e transpostas entre si:

$$\frac{dZ}{d\beta} = -2(d\beta')X'Y + 2(d\beta')X'X\beta \quad (28)$$

$$\frac{dZ}{d\beta} = -2(d\beta')(X'X\beta - X'Y) \quad (29)$$

Fazendo  $\frac{dZ}{d\beta} = 0$  e sendo  $d\beta'$  arbitrário, tem-se:

$$X'X\beta - X'Y = 0$$

$$X'X\beta = X'Y \rightarrow \text{Sistema de equações normais} \quad (30)$$

Para o caso de modelo de regressão de posto completo, então  $X'X$  é não-singular, portanto, apresenta inversa. Logo:

$$\begin{aligned} X'X\hat{\beta} &= X'Y \\ (X'X)^{-1}X'X\hat{\beta} &= (X'X)^{-1}X'Y \\ I\hat{\beta} &= (X'X)^{-1}X'Y \end{aligned} \quad (31)$$

Pois,  $(X'X)^{-1}X'X = I$

$\hat{\beta} = (X'X)^{-1}X'Y$ , estimador de mínimos quadrados.

Como:

$$Y_i - \bar{Y} = (Y_i - \hat{Y}_i) + (\hat{Y}_i - \bar{Y}) \quad (32)$$

Manipulando algebricamente chega-se a:

$$\sum_{i=1}^n (Y_i - \bar{Y})^2 = \sum_{i=1}^n [(Y_i - \hat{Y}_i) + (\hat{Y}_i - \bar{Y})]^2 \quad (33)$$

Como os desvios totais são independentes, isto é, o duplo produto é igual a zero, logo, pode-se reescrever:

$$\sum_{i=1}^n (Y_i - \bar{Y})^2 = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 + \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2 \quad (34)$$

$$SQ_{Total} = SQ_{resíduo} + SQ_{Regressão}$$

Como a soma dos quadrados total é igual:

$$SQ_{Total} = \sum_{i=1}^n (Y_i - \bar{Y})^2 = \sum_{i=1}^n Y_i^2 - \frac{\left(\sum_{i=1}^n Y_i\right)^2}{n} \quad (35)$$

Em que:

$n$  = número total de observações.

Matricialmente, pode-se escrever:

$$SQ_{Total} = Y' \left[ I - \frac{u \cdot u'}{n} \right] Y$$

Em que:

$I$  = matriz identidade de ordem  $n$ ;

$u$  = vetor de 1's de dimensão  $n \times 1$ , e;

$C$  = correção, dada por

$$C = \frac{\left(\sum_{i=1}^n Y_i\right)^2}{n} \quad (36)$$

ou

$$C = \frac{1}{n} Y' u u' Y \quad (37)$$

A soma de quadrados da regressão pode escrita como:

$$SQ_{Reg} = \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2 = \sum_{i=1}^n \hat{Y}_i^2 - \left(\sum_{i=1}^n Y_i\right)^2 \quad (38)$$

Matricialmente, pode-se escrever:

$$SQRe g = \hat{Y}'\hat{Y} - \frac{1}{n}Y'u.u'Y \quad (39)$$

E, sendo:

$$X'X\hat{\beta} = X'Y \quad (40)$$

$$SQRe g = \hat{\beta}'X'X\hat{\beta} - \frac{1}{n}Y'u.u'Y \quad (41)$$

$$SQRe g = \hat{\beta}'X'Y - C \quad (42)$$

A soma de quadrados de resíduo é igual a:

$$SQRes = (Y - \hat{Y})'(Y - \hat{Y}) \quad (43)$$

$$SQRes = (Y - X\hat{\beta})'(Y - X\hat{\beta}) \quad (44)$$

$$SQRes = (Y - \hat{\beta}'X')(Y - X\hat{\beta}) = Y'Y - Y'X\hat{\beta} - \hat{\beta}'X'Y + \hat{\beta}'X'X\hat{\beta} \quad (45)$$

Sendo as matrizes  $Y'X\hat{\beta}$  e  $\hat{\beta}'X'Y$  de dimensão 1 x 1 (escalar) e sendo uma a transposta da outra, logo,  $Y'X\hat{\beta} = \hat{\beta}'X'Y$ . Assim,

$$SQRes = Y'Y - 2\hat{\beta}'X'Y + \hat{\beta}'X'X\hat{\beta} \quad (46)$$

Como,

$$X'X\hat{\beta} = X'Y \quad (47)$$

Então,

$$SQRes = Y'Y - 2\hat{\beta}'X'Y + \hat{\beta}'X'Y \quad (48)$$

$$SQRes = Y'Y - \hat{\beta}'X'Y \quad (49)$$

Manipulando-se algebricamente, cheg-se a:

$$SQRes = Y'Y - C - \hat{\beta}' X'Y + C \quad (50)$$

$$SQRes = (Y'Y - C) - (\hat{\beta}' X'Y - C) \quad (51)$$

$$SQRes = SQTotal - SQReg \quad (51)$$

Tabela 01- Tabela de Análise de Variância da regressão em notação matricial

Fonte de variação	Graus de liberdade	Soma dos quadrados	Quadrados médios	F
Regressão	$p$	$\hat{\beta}' X'Y - C$	$QMReg$	$QMReg / QMRes$
Resíduo	$n - p - 1$	$Y'Y - \hat{\beta}' X'Y$	$QMRes$	
Total	$n - 1$	$Y'Y - C$		

Em que  $p$  = número de coeficientes de regressão ou número de parâmetros menos 1.

Para a determinação das estimativas das variâncias das estimativas dos parâmetros, partimos da definição da covariância do estimador de  $\beta$  (Rawlings, 1998).

$$\text{cov}(\beta) = E[(\hat{\beta} - \beta)(\hat{\beta} - \beta)'] \quad (52)$$

Sabe-se que:

$$\hat{\beta} = (X'X)^{-1} X'Y = (X'X)^{-1} X'(X\beta + \varepsilon) = (X'X)^{-1} X'X\beta + (X'X)^{-1} X'\varepsilon \quad (53)$$

$$\hat{\beta} = I\beta + (X'X)^{-1} X'\varepsilon = \beta + (X'X)^{-1} X'\varepsilon \quad (54)$$

Logo,

$$\hat{\beta} - \beta = (X'X)^{-1} X'\varepsilon \quad (55)$$

Assim,

$$\text{cov}(\hat{\beta}) = E\{[(X'X)^{-1}X'\varepsilon][(X'X)^{-1}X'\varepsilon]'\} \quad (56)$$

$$\text{cov}(\hat{\beta}) = E[(X'X)^{-1}X'\varepsilon\varepsilon'X(X'X)^{-1}], \text{ pois, } (X'X)^{-1} \text{ é simétrica} \quad (57)$$

$$\text{cov}(\hat{\beta}) = (X'X)^{-1}X'E(\varepsilon\varepsilon')X(X'X)^{-1} \quad (58)$$

Mas,

$$E(\varepsilon\varepsilon') = V(\varepsilon) = I\sigma^2 \quad (59)$$

Portanto,

$$\text{cov}(\hat{\beta}) = (X'X)^{-1}X'I\sigma^2X(X'X)^{-1} \quad (60)$$

$$\text{cov}(\hat{\beta}) = (X'X)^{-1}X'X(X'X)^{-1}I\sigma^2 \quad (61)$$

$$\text{cov}(\hat{\beta}) = (X'X)^{-1}\sigma^2 \quad (62)$$

Obtendo-se o estimador:

$$\widehat{\text{cov}}(\hat{\beta}) = (X'X)^{-1}s^2 = (X'X)^{-1}QM \text{ Re } s \quad (63)$$

Tendo como resultado a matriz (Steel et al, 1997):

$$\widehat{\text{cov}}(\hat{\beta}) = \begin{bmatrix} s^2(\hat{\beta}_0) & \widehat{\text{cov}}(\hat{\beta}_0, \hat{\beta}_1) & \widehat{\text{cov}}(\hat{\beta}_0, \hat{\beta}_2) & \cdots & \widehat{\text{cov}}(\hat{\beta}_0, \hat{\beta}_p) \\ \widehat{\text{cov}}(\hat{\beta}_1, \hat{\beta}_0) & s^2(\hat{\beta}_1) & \widehat{\text{cov}}(\hat{\beta}_1, \hat{\beta}_2) & \cdots & \widehat{\text{cov}}(\hat{\beta}_1, \hat{\beta}_p) \\ \widehat{\text{cov}}(\hat{\beta}_2, \hat{\beta}_0) & \widehat{\text{cov}}(\hat{\beta}_2, \hat{\beta}_1) & s^2(\hat{\beta}_2) & \cdots & \widehat{\text{cov}}(\hat{\beta}_2, \hat{\beta}_p) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \widehat{\text{cov}}(\hat{\beta}_p, \hat{\beta}_0) & \widehat{\text{cov}}(\hat{\beta}_p, \hat{\beta}_1) & \widehat{\text{cov}}(\hat{\beta}_p, \hat{\beta}_2) & \cdots & s^2(\hat{\beta}_p) \end{bmatrix} \quad (64)$$

Onde os elementos da diagonal são as estimativas das variâncias dos estimadores dos parâmetros.



Essas estimativas são importantes para a determinação dos intervalos de confiança para os parâmetros e, para realização dos testes de hipóteses sobre os parâmetros utilizando o teste “t” de Student.

O teste “t” de Student pode ser estruturado como a seguir (Chatterjee e Hadi, 2006):

$$H_0 : \beta_i = 0$$

$$H_a : \beta_i \neq 0$$

$$t_{calc} = \frac{\hat{\beta}_i}{s(\beta_i)}$$

$$t_{tab} = t_{\frac{\alpha}{2}}(n - p - 1)$$

Se  $|t_{calc}| > |t_{tab}|$  rejeita-se  $H_0$  pelo teste “t” de Student ao nível  $\alpha$  de probabilidade. Em que  $\alpha$  é o nível de probabilidade de ocorrer o erro tipo 1.

O teste “t” de Student é equivalente ao teste F, pois ambos levam ao mesmo nível de significância para o erro tipo 1 e, conseqüentemente a mesma conclusão.

As conclusões somente serão as mesmas para o teste “t” de Student bilateral e para o modelo de regressão linear simples, ou quando se testa cada parâmetro isoladamente na regressão linear múltipla.

### 2.3. O teste de Scott-Knott

O procedimento de Scott e Knott (1974) utiliza a razão de verossimilhança para testar a significância de que os  $n$  tratamentos podem ser divididos em dois grupos que maximizem a soma de quadrados entre grupos e seja mínima dentro do grupo (Ramalho et al., 2000). Sejam, por exemplo, 3 tratamentos, A, B e C. O processo consiste em determinar uma partição, em dois grupos, que maximize a soma de quadrado entre grupos. Verifica-se que nesse caso são possíveis  $(2^{n-1} - 1)$  grupos, isto é, A vs B e C, B vs A e C e C vs A e B. Com um número pequeno de tratamentos como o do exemplo, é fácil obter todos os grupos,

contudo, quando o número ( $n$ ) de tratamentos é grande, o número de grupos cresce exponencialmente, dificultando a aplicação do teste.

Para atenuar esse problema, basta ordenar as médias dos tratamentos. Fisher (1958) citado por Scott-Knott (1974). Nessa situação, o número de partições possíveis passa a ser obtido por  $n - 1$ , obtidas pela ordenação das médias e sua divisão entre duas partições sucessivas. Após encontrar a melhor separação entre dois grupos, repete-se o processo em cada subgrupo. A partir daí, prossegue-se com a subdivisão até que os grupos resultantes sejam considerados estatisticamente iguais pelo teste  $\chi^2$ . Ordenadas as médias, procede-se do seguinte modo, fazendo inicialmente o número de tratamentos envolvidos no grupo de médias considerado ( $g$ ) igual ao o número total de tratamentos ( $n$ ).

i) Determinar a partição entre dois grupos que maximiza a soma de quadrados (SQ) entre grupos. Seja  $T_1$  e  $T_2$  os totais dos dois grupos com  $k_1$  e  $k_2$  tratamentos em cada um, a soma de quadrados  $B_0$  é estimada por:

$$B_0 = \frac{T_1^2}{k_1} + \frac{T_2^2}{k_2} - \frac{(T_1 + T_2)^2}{k_1 + k_2} \quad (65)$$

$$T_1 = \sum_{i=1}^{k_1} \bar{y}_{(i)} \quad (66)$$

e,

$$T_2 = \sum_{i=k_1+1}^g \bar{y}_{(i)} \quad (67)$$

em que  $\bar{Y}_{(i)}$  é a média do tratamento da posição ordenada  $i$ .

Os dois grupos deverão ser identificados por meio da inspeção das somas de quadrados das  $g - 1$  partições possíveis, sendo  $g$  o número de tratamentos envolvidos no grupo de médias considerado.

ii) Determinar o valor da estatística  $\lambda$ :

$$\lambda = \frac{\pi}{2(\pi - 2)} \times \frac{B_0}{\hat{\sigma}_o^2} \quad (68)$$

em que  $\hat{\sigma}_o^2$  é o estimador de máxima verossimilhança de  $\hat{\sigma}_{\bar{y}}^2$  dado por:

$$\hat{\sigma}_o^2 = \frac{1}{g + v} \left[ \sum_{i=1}^g (\bar{y}_{(i)} - \bar{y})^2 + v s_{\bar{y}}^2 \right] \quad (69)$$

em que,  $v$  é o grau de liberdade do erro associado a este estimador,  $\bar{y}$  é a média das médias dos  $g$  tratamentos e  $s_{\bar{y}}^2 = \frac{QME}{r}$  o estimador não viesado de  $\sigma_{\bar{y}}^2$ , sendo  $QME$  o quadrado médio do erro e  $r$  o número de repetições.

iii) Se  $\lambda \geq \chi_{(\alpha; g/(\pi-2))}^2$  rejeita-se a hipótese de que os dois grupos são idênticos em favor da hipótese alternativa de que os dois grupos diferem.  $\chi_{(\alpha; g/(\pi-2))}^2$  é um valor tabelado obtido na tabela de Qui-quadrado, correspondente a combinação entre o nível de significância estabelecido ( $\alpha$ ) e o valor dado por  $g/(\pi - 2)$ .

iv) No caso de rejeitar esta hipótese, os dois subgrupos formados serão independentemente submetidos aos passos i) a iii), fazendo respectivamente  $g = k_1$  e  $g = k_2$ . O processo, em cada subgrupo, encerra-se ao se aceitar  $H_0$  no passo iii) ou se cada subgrupo contiver apenas uma média.

## 2. 4. Identidade de modelos

### 2.4.1. O teste de Chow

Chow (1960) sugeriu um teste para comparar duas equações lineares, cujo algoritmo segue os seguintes passos:

1. Dadas as seguintes relações lineares: referentes a dois conjuntos de observações.

$$\begin{aligned} y_{1i} &= a_1 + b_1 x_{1i} + \varepsilon_{1i} & i &= 1, 2, \dots, n_1 \\ y_{2i} &= a_2 + b_2 x_{2i} + \varepsilon_{2i} & i &= 1, 2, \dots, n_2 \end{aligned} \quad (70)$$

2. Combinam-se todas as  $n_1 + n_2$  observações e calcula-se a estimativa de quadrados mínimos de  $a$  e  $b$  na regressão combinada  $y = a + bx + \varepsilon$ . Desta equação obtém-se a soma de quadrados de resíduo ( $SQRes_1$ ) com grau de liberdade igual a  $n_1 + n_2 - p$ , em que  $p$  é o número de parâmetros a ser estimado. Neste caso,  $p = 2$ .

3. Obtém-se a soma de quadrados de resíduo para as duas equações, ou seja,  $SQRes_2$  e  $SQRes_3$ , com os graus de liberdade  $n_1 - p$  e  $n_2 - p$ , respectivamente. Somam-se estas duas somas de quadrados de resíduo, isto é,  $SQRes_4 = SQRes_2 + SQRes_3$  e seus graus de liberdade  $n_1 + n_2 - 2p$ .

4. Obtém-se  $SQRes_5 = SQRes_1 + SQRes_4$ .

5. Calcula-se a estatística  $F$  como:

$$F_{calc} = \frac{\frac{SQRes_5}{p}}{\frac{SQRes_4}{n_1 + n_2 - 2p}} \quad (71)$$

com  $p$  e  $n_1 + n_2 - 2p$  graus de liberdade.

Se  $F_{calc} > F_{tabelado}$ , para um determinado nível de significância  $\alpha$ , rejeita-se a hipótese de que os parâmetros  $a$ 's e  $b$ 's são os mesmos para os dois conjuntos de observações.

Segundo Gujarati (1970b), o teste de Chow (1960) permite uma avaliação geral da equação, assegurando apenas se duas regressões lineares são iguais ou diferentes. Caso sejam diferentes, não especificam se a diferença é devida a interceptos ou inclinações.

#### 2.4.2. A variável Dummy

A especificação de uma equação de regressão com variáveis Dummy permite saber qual a origem da variabilidade, isto é, em quais parâmetros ocorre essa variabilidade, além de estimar em uma única equação de regressão, todas

as observações, constituindo assim, uma forma mais eficiente de testar a estabilidade dos parâmetros, aumentando os graus de liberdade e, conseqüentemente, melhorando a precisão relativa da estimação dos parâmetros (Valle, 2002).

Gujarati (1970a) apresentou uma metodologia para verificar a identidade de modelos de regressão linear utilizando as variáveis Dummy, conhecidas também como variáveis binárias.

A metodologia pode ser ilustrada a seguir

Dado o modelo linear:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad (72)$$

Considere-se quatro conjuntos de observações com seus respectivos modelos:

$$\begin{aligned} y_{1i} &= \beta_1 + \beta_2 x_{1i} + \varepsilon_i \\ y_{2i} &= \beta_3 + \beta_4 x_{2i} + \varepsilon_i \\ y_{3i} &= \beta_5 + \beta_6 x_{3i} + \varepsilon_i \\ y_{4i} &= \beta_7 + \beta_8 x_{4i} + \varepsilon_i \end{aligned} \quad (73)$$

Precisa-se saber se as quatro equações anteriores são iguais, ou seja, se os interceptos  $\beta_1 = \beta_3 = \beta_5 = \beta_7$  e os coeficientes angulares  $\beta_2 = \beta_4 = \beta_6 = \beta_8$ .

A seguir, a equação que representa simultaneamente as quatro equações

$$\begin{aligned} y_j &= a_0 + a_1 D_1 + a_2 D_2 + a_3 D_3 + a_4 x_j + a_5 D_1 x_j + a_6 D_2 x_j + a_7 D_3 x_j + \varepsilon_j \\ j &= 1, 2, 3, \dots, (n_1 + n_2 + n_3 + n_4) \end{aligned} \quad (74)$$

Sendo  $D_1, D_2, D_3$  as variáveis Dummy, que podem assumir o valor 1 ou 0.

Quando:

$D_1, D_2,$  e  $D_3 = 0$  as observações encontram-se no primeiro grupo;

$D_1 = 1$  e,  $D_2$  e  $D_3 = 0$  as observações encontram-se no segundo grupo;

$D_2 = 1$  e,  $D_1$  e  $D_3 = 0$  as observações encontram-se no terceiro grupo;

$D_3 = 1$  e,  $D_1$  e  $D_2 = 0$  as observações encontram-se no quarto grupo.

Os vários  $a_i$ 's são interpretados como:

$a_0$  = intercepto do grupo 1;

$a_1$  = diferencial para o intercepto do grupo 2;

$a_2$  = diferencial para o intercepto do grupo 3;

$a_3$  = diferencial para o intercepto do grupo 4;

$a_4$  = coeficiente angular do grupo 1;

$a_5$  = diferencial para o coeficiente angular do grupo 2;

$a_6$  = diferencial para o coeficiente angular do grupo 3;

$a_7$  = diferencial para o coeficiente angular do grupo 4.

A partir do diferencial para o intercepto e o coeficiente angular, pode-se determinar os valores para o intercepto e o coeficiente angular para cada grupo, como descrito a seguir:

Grupo 1:  $y = \alpha_0 + \alpha_4 x$

Grupo 2:  $y = (\alpha_0 + \alpha_1) + (\alpha_4 + \alpha_5)x$

Grupo 3:  $y = (\alpha_0 + \alpha_2) + (\alpha_4 + \alpha_6)x$

Grupo 4:  $y = (\alpha_0 + \alpha_3) + (\alpha_4 + \alpha_7)x$

Para determinar as equações para os grupos 1, 2, 3 e 4 descrito acima, precisa-se estimar pelo método dos mínimos quadrados, os coeficientes angulares e interceptos a partir da equação 74, desde que os pressupostos usuais sejam atendidos, ou seja,  $E(\varepsilon_j) = 0$ ,  $E(\varepsilon_j, x_{ij}) = 0$  e  $E(\varepsilon_j \varepsilon_{j+k}) = \sigma_u^2$  para  $k = 0$ , e zero se  $k \neq 0$ .

Dependendo da significância estatística dos diferenciais estimados, pode-se descobrir se os grupos de equações são estatisticamente iguais, se apresentam o mesmo intercepto, se possuem a mesma declividade ou se são estatisticamente diferentes.

Figueiredo (2005), utilizou o método das variáveis Dummy para estimar o volume individual de árvores, para quatro grupos de interesse (volume comercial com casca, volume comercial sem casca, volume total sem casca e volume total com casca), em função diâmetro na altura do peito.

Magalhaes et al (2004), avaliaram os métodos de identidade de modelos (Chow), das variáveis Dummy e da análise de variância para a comparação de modelos lineares simples, sinalizando resultados bem semelhantes em relação a ocorrência de erro tipo I e erro tipo II, sendo que o método das variáveis Dummy foi o mais eficiente, apresentando os menores percentuais de erro tipo I e erro tipo II.

Estudando modelos de regressão linear simples, na área médica, Magalhães et al (2009a) e Magalhães et al (2009b), verificaram que a aplicação do método da identidade de modelos foi equivalente ao método das variáveis Dummy.

#### 2.4.3. Seleção da melhor equação

A análise de regressão linear múltipla é muito utilizada por pesquisadores de vários ramos da pesquisa científica, entretanto, pode ser perigosa quando mal utilizada e interpretada. Na obtenção da equação de regressão devemos observar dois critérios, geralmente antagônicos, são eles (Draper e Smith, 1998):

- a) Equação com o maior número de variáveis independentes para que a equação forneça resultados mais precisos;
- b) Equação com o menor número de variáveis independentes a fim de simplificar os cálculos.

A seleção da melhor equação de regressão leva em consideração a conciliação entre estes dois critérios (Draper e Smith, 1998).

Os principais procedimentos para a seleção da melhor equação de regressão são (Snedecor e Cochran, 1967; Draper e Smith, 1998):

- a) Todas as equações possíveis;
- b) Eliminação indireta ou passo atrás (“*backward*”);
- c) Seleção por etapas ou passo a passo (“*stepwise*”);
- d) Passo a frente (“*forward*”).

O procedimento que leva em consideração todas as equações possíveis requer o ajustamento de todas as equações. O número de equações é igual  $2^p - 1$ , sendo  $p$  o número de variáveis independentes. O procedimento tem os seguintes passos:

- i) Ajustar todas as equações possíveis;
- ii) Agrupar as equações em conjuntos de 1, 2, ... ,  $p$ , variáveis independentes;
- iii) Ordenar as equações dentro de cada conjunto em função de coeficiente de determinação;
- iv) Examinar as melhores equações e escolher a mais indicada por alguns critérios adicionais, como o quadrado médio do resíduo ou estatística  $C_p$  de Mallows (Draper e Smith, 1998).

Em geral, a análise de todas as regressões possíveis não é muito útil, principalmente em equações com muitas variáveis.

A eliminação indireta ou passo a trás é um aperfeiçoamento do método de todas as equações possíveis, consistindo basicamente em ajustar um modelo de regressão utilizando todas variáveis independentes possíveis e determinar a contribuição de cada variável para a soma dos quadrados da regressão. De acordo com a contribuição, pode existir uma redução do número de variáveis independentes na equação, até a obtenção da melhor equação (Draper e Smith, 1998).

A seguir, os passos que devem ser seguidos na operação:

- i) Obter a equação de regressão com todas as variáveis independentes;
- ii) Calcular o valor parcial de  $F$  para cada variável, como se fosse a última a entrar no modelo;
- iii) Selecionar a variável de menor valor de  $F$  e comparar com ao valor de  $F$  tabelado, a um determinado nível de significância pré-estabelecido;
- iv) Se  $F$  calculado  $<$   $F$  tabelado, elimina-se a variável e determina-se uma nova equação de regressão com as variáveis restantes;
- v) Se  $F$  calculado  $>$   $F$  tabelado, escolhemos a equação de regressão conforme calculada e finaliza-se o processo.

É um processo mais rápido quando deseja-se ver todas as variáveis no modelo, porém, pode aparecer multicolinearidade.



Como o procedimento utilizando o teste “F” é trabalhoso, sugere-se o teste “t” de *Student* neste processo. Quando se tem um grau de liberdade para a hipótese, o teste “t” leva ao mesmo resultado do teste “F” (Myers, 1989).

### 3. MATERIAL E MÉTODOS

#### **3.1. Histórico dos experimentos no programa de melhoramento do feijão de vagem da UENF**

Na Universidade Estadual do Norte Fluminense Darcy Ribeiro (UENF), iniciou-se um programa de melhoramento genético com feijão-de-vagem de hábito indeterminado, com o objetivo de selecionar genótipos produtivos e de qualidade comercial para o Norte e Noroeste Fluminense. O programa teve início com a caracterização e estudo da diversidade genética de 25 acessos do Banco de Germoplasma da UENF, de hábito indeterminado. A partir de então foram realizados os cruzamentos entre cinco acessos divergentes e com características desejáveis, obtendo dez híbridos dialélicos. Foram realizadas seleções nas populações  $F_2$ , em campo; avançando as gerações  $F_3$ ,  $F_4$  e  $F_5$  pelo método SSD ("single seed descent" - descendente de uma única semente por planta), em casa de vegetação, abrindo e selecionando linhagens em  $F_6$ . A partir daí foram selecionadas 27 linhagens promissoras desta geração  $F_6$  com a qual se realizou um trabalho em três estações experimentais: Campos dos Goytacazes (RJ); Itaocara (RJ) e Bom Jesus do Itabapoana (RJ) obtendo a geração  $F_7$ . Em sequência buscou-se a geração  $F_8$  realizando um novo experimento utilizando 27 linhagens selecionadas de feijão-de-vagem da geração  $F_7$  e mais três testemunhas (duas variedades comerciais, FELTRIN, TOP SEED Blue Line e um dos progenitores, 19 UENF-1445), de hábito de crescimento indeterminado, do

Programa de Melhoramento da Universidade Estadual do Norte Fluminense. O experimento foi conduzido no período de maio a setembro de 2010 (Araujo, 2011).

Dos experimentos realizados, foram selecionadas 14 linhagens de feijão-de-vagem, da geração F<sub>8</sub>. Dando segmento ao programa de melhoramento genético da UENF, foi realizado um trabalho com as linhagens selecionadas e três cultivares de feijão-de-vagem (duas variedades comerciais e um dos progenitores da UENF), em Bom Jesus do Itabapoana, RJ, para obtenção da geração F<sub>9</sub> (Araujo, 2011; Almeida, 2011).

Tabela 02: Relação dos 17 Genótipos de Feijão-de-vagem utilizados nos estudos, sendo 14 linhagens e três testemunhas

Genótipos	Identificação
01	Progenitor UENF 19 (UENF 1445)
02	Feltrin
03	Top Seed Blue Line
04	UENF 7-3-1
05	UENF 7-4-1
06	UENF 7-5-1
07	UENF 7-6-1
08	UENF 7-9-1
09	UENF 7-10-1
10	UENF 7-12-1
11	UENF 7-14-1
12	UENF 7-20-1
13	UENF 9-24-2
14	UENF 14-3-3
15	UENF 14-4-3
16	UENF 14-6-3
17	UENF 15-23-4

Os genótipos de 1 a 3 são testemunhas e de 4 a 17 linhagens.

### 3.2. Localização e caracterização da área experimental

O experimento foi instalado no Instituto Federal Fluminense, Campus Bom Jesus do Itabapoana, Coordenadas UTM 224021 m E e 7660399 m S, zona 24K, altitude 84 m, com temperatura média anual de 23°C, numa área de relevo plano e precipitação média anual em torno de 1480 mm (IBGE, 2012).

Análise do solo efetuada na área de plantio revelou as seguintes características: pH = 5,7; fósforo = 14 mg/dm<sup>3</sup>; potássio = 211 mg/dm<sup>3</sup>; enxofre = 4,0 mg/dm<sup>3</sup>; Ca = 4,3 cmol/dm<sup>3</sup>; Mg = 1,5 cmol/dm<sup>3</sup>; Al = 0,0 cmol/dm<sup>3</sup>; H+Al = 3,7 g/kg; C = 11,5 g/kg e M.O = 19,8 g/kg.

Com base nesta análise do solo e levando em consideração as exigências da cultura de feijão-de-vagem, foi realizada a seguinte adubação de acordo com Carrijo et al (1999b):

- Nitrogênio = 150 kg ha<sup>-1</sup> de N - sendo aplicado 30% deste na semeadura e o restante em cobertura (30 e 60 dias após o plantio), usando ureia;
- Fósforo = 180 kg ha<sup>-1</sup> de P<sub>2</sub>O<sub>5</sub> (superfosfato simples) - sendo aplicado 100% na semeadura;
- Potássio = 30 kg ha<sup>-1</sup> de K<sub>2</sub>O (cloreto de potássio) - sendo aplicados 50% no plantio e o restante juntamente com as demais aplicações de nitrogênio.

A semeadura foi realizada em doze de junho de 2012, colocando-se duas sementes por cova, a uma profundidade de 2,5 cm. A emergência iniciou, em média, dez a quinze dias após a semeadura. Foi realizado o desbaste deixando-se apenas uma planta por cova. Cerca de quinze dias após a emergência, as plantas foram tutoradas com bambu e arame. Durante a condução do experimento, foram realizados os tratos culturais e fitossanitários recomendados para a cultura, segundo Filgueira (2008), bem como irrigação por aspersão.

### 3.3. Amostragem simples ao acaso

No plantio dos dezessete genótipos descritos, foram selecionadas ao acaso 20 plantas de cada um dos genótipos.

As alturas das plantas foram medidas com régua de madeira com dois metros, com intervalos de quatro dias, iniciando as avaliações 12 dias após o plantio, até que as plantas atingissem a altura máxima do espaldeamento, o que ocorreu 32 dias após a primeira avaliação para os genótipos de maiores taxas de crescimento.

A análise amostral foi realizada para cada genótipo, em relação à média, desvio padrão, erro padrão, coeficiente de variação, amplitude, valor de “t”, intervalo de confiança a 5% e amostra ideal ao nível de 10% de significância.

Inicialmente foi realizada a análise de variância a 5% de probabilidade para verificar se os genótipos apresentavam efeitos significativos, ou seja, se houve diferença significativa entre os intervalos de tempo. Confirmada a diferença, foi realizada a análise de regressão para o modelo linear simples para cada planta individualmente.

Após, foi realizado o teste de Scott-Knott para agrupamento dos genótipos a nível de 5% de probabilidade, em relação ao intercepto ( $\beta_0$ ) e ao coeficiente angular ( $\beta_1$ ) ao nível de 5% de probabilidade. Novos grupos foram formados por aqueles genótipos que pertenciam, simultaneamente, ao grupo do intercepto ( $\beta_0$ ) e da declividade ( $\beta_1$ ) e, um grupo para os genótipos que não pertenciam a nenhum dos grupos formados anteriormente.

Para identificar as diferenças entre as curvas dos genótipos que compunham cada grupo, foi realizado o teste de identidade de modelos, utilizando a técnica de variáveis Dummy, para os grupos formados pelo teste de agrupamento de Scott-Knott, e para o grupo formado pelos genótipos que não pertenciam a nenhum grupo, conforme descrito por Gujarati (1970a).

Os parâmetros da regressão linear, foram testados pelo teste “t” de Student a nível de 5% de probabilidade, eliminando os parâmetros que apresentavam maior probabilidade de erro tipo 1, desde que estes apresentassem probabilidade de erro maior que 5% e, procedendo-se nova análise sem o parâmetro eliminado, conforme processo descrito por Draper e Smith (1998). Continuando o processo até que não existissem nenhum parâmetro com probabilidade de erro maior que 5%. Obtendo-se assim, a equação que representa o grupo de genótipos.

As análises estatísticas foram realizadas com o auxílio do programa estatístico SAEG (2007).

#### 4. RESULTADOS E DISCUSSÃO

O coeficiente de variação experimental pode ser classificado em baixo quando menor que 10%; médio, entre 10% e 20%; alto, entre 20% e 30%; e muito alto, quando maior que 30% (Pimentel-Gomes, 2000).

Avaliando a variação experimental devido aos fatores não controláveis no experimento e de acordo com a classificação proposta por Pimentel-Gomes (2000), em relação aos coeficientes de variação (CV) estimados no experimento de campo, para o intercepto ( $\beta_0$ ), tabela 3, os genótipos 2, 3, 7, 9, 11, 12 e 14 foram considerados médios, enquanto que os genótipos 1, 4, 5, 6, 8, 10, 13 e 17 considerados altos e o genótipo 15 e 16, considerados muito altos.

Em relação a declividade ( $\beta_1$ ), tabela 4, os genótipos 1, 2, 3, 4, 5, 6, 7, 9, 11, 12, 13, 14 e 17 foram considerados de CV médio, os genótipos 8, 10 e 16, altos, enquanto que o genótipo 15, considerado muito alto.

Observa-se que o genótipo 15 teve coeficiente de variação (CV) muito alto para o intercepto ( $\beta_0$ ) e a declividade ( $\beta_1$ ), enquanto que os genótipos 8 e 10 apresentaram CV alto, e os genótipos 2, 3, 7, 9, 11, 12 e 14 apresentaram CV médio (tabelas 3 e 4).

O genótipo 15 em relação à amostra ideal para a declividade ( $\beta_1$ ), foi superior a 53 e para o intercepto ( $\beta_0$ ) superior a 52, valores muito superiores ao número de observações realizados, indicando possível incompleta endogamia.

Tabela 3 – Estatística descritiva para a amostra do intercepto ( $\beta_0$ ) dos dezessete genótipos analisados

Genótipos	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Número de Observações	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20
Média Geral	-12,344	-38,426	-29,296	-21,247	-24,201	-28,194	-30,939	-19,081	-22,628	-21,242	-26,932	-28,217	-14,603	-35,402	-18,578	-15,842	-17,517
Desvio Padrão	3,483	5,837	4,881	4,542	5,089	5,705	5,567	5,221	3,487	5,243	4,406	5,043	3,271	6,626	6,450	5,507	3,646
Erro Padrão	0,779	1,305	1,091	1,016	1,138	1,276	1,245	1,167	0,780	1,172	0,985	1,128	0,731	1,482	1,442	1,232	0,815
Coeficiente de Variação	28,2%	15,2%	16,7%	21,4%	21,0%	20,2%	18,0%	27,4%	15,4%	24,7%	16,4%	17,9%	22,4%	18,7%	34,7%	34,8%	20,8%
Amplitude	12,511	26,533	14,378	15,789	19,278	22,456	20,367	20,456	13,833	22,089	20,033	19,122	15,067	23,856	25,433	25,544	11,678
Teste de t	-15,850	-29,440	-26,840	-20,921	-21,265	-22,101	-24,857	-16,344	-29,022	-18,120	-27,335	-25,025	-19,967	-23,892	-12,880	-12,864	-21,486
Intervalo de Confiança P(0.05)	1,628	2,728	2,281	2,123	2,378	2,666	2,601	2,440	1,630	2,450	2,059	2,357	1,529	3,097	3,015	2,574	1,704
Amostra Ideal (10%)	34,776	10,079	12,127	19,959	19,318	17,885	14,140	32,704	10,372	26,607	11,692	13,950	21,913	15,304	52,660	52,795	18,924

Tabela 4 – Estatística descritiva para a amostra da declividade ( $\beta_1$ ) dos dezessete genótipos analisados

Genótipos	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Número de Observações	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20
Média Geral	3,190	7,350	6,674	4,309	5,793	5,947	5,812	4,151	4,746	4,125	5,754	6,557	3,754	7,653	3,288	4,006	3,696
Desvio Padrão	0,479	1,253	0,993	0,662	1,158	0,897	1,016	0,914	0,661	0,923	0,769	0,694	0,555	0,886	1,163	0,923	0,602
Erro Padrão	0,107	0,280	0,222	0,148	0,259	0,201	0,227	0,204	0,148	0,206	0,172	0,155	0,124	0,198	0,260	0,206	0,135
Coeficiente de Variação	15,017	17,046	14,883	15,351	19,989	15,080	17,481	22,021	13,930	22,380	13,366	10,577	14,776	11,571	35,378	23,053	16,297
Amplitude	2,140	5,002	3,485	2,577	3,523	3,794	4,250	3,358	2,808	3,627	2,929	2,525	2,052	3,867	4,431	4,210	1,952
Teste de t	29,780	26,236	30,049	29,133	22,373	29,657	25,582	20,308	32,103	19,983	33,459	42,282	30,266	38,649	12,641	19,399	27,442
Intervalo de Confiança P(0.05)	0,224	0,586	0,464	0,309	0,541	0,419	0,475	0,427	0,309	0,431	0,359	0,324	0,259	0,414	0,544	0,432	0,281
Amostra Ideal (10%)	9,851	12,692	9,675	10,294	17,453	9,933	13,349	21,182	8,477	21,879	7,804	4,887	9,537	5,849	54,673	23,215	11,601

As médias de intercepto ( $\beta_0$ ), declividade ( $\beta_1$ ) e altura máxima dos 17 genótipos com suas respectivas comparações (agrupamentos), pelo critério de Scott-Knott a nível de 5% de probabilidade são apresentadas na tabela 5.

Em relação à média das alturas máximas dos genótipos avaliados, no momento da última avaliação, de acordo com o critério de agrupamento de Scott-Knott, a nível de 5% de probabilidade, os genótipos 2 e 14 foram os que apresentaram a maior média de altura para a última avaliação, enquanto que as menores foram dos genótipos 1, 13, 15, 16 e 17 (tabela 05).

Araujo (2011) em trabalho com os mesmo genótipos, verificou que a maior altura alcançada foi do genótipo 1 e a menor do genótipo 22, apesar de não existir diferença significativa entre as médias das alturas, pelo critério de agrupamento de Scott-Knott a nível de 5% de probabilidade.

Tabela 05 – Médias das estimativas dos estimadores de  $\beta_0$  e  $\beta_1$  e da última avaliação das alturas dos genótipos avaliados

Genótipos	Identificação	Média ( $\beta_0$ )	Média ( $\beta_1$ )	Altura
1	Progenitor UENF 19 (UENF 1445)	-12,3444A	3,1901E	109E
2	Feltrin	-38,4256F	7,3501A	238A
3	Top Seed Blue Line	-29,2961E	6,6741B	209B
4	UENF 7-3-1	-21,2472D	4,3094D	152D
5	UENF 7-4-1	-24,2006D	5,7929C	188C
6	UENF 7-5-1	-28,1939E	5,9472C	198B
7	UENF 7-6-1	-30,9394E	5,8122C	189C
8	UENF 7-9-1	-19,0811C	4,1507D	139D
9	UENF 7-10-1	-22,6283D	4,7461D	159D
10	UENF 7-12-1	-21,2417D	4,1250D	144D
11	UENF 7-14-1	-26,9317E	5,7541C	187C
12	UENF 7-20-1	-28,2172E	6,5570B	204B
13	UENF 9-24-2	-14,6033A	3,7539E	123E
14	UENF 14-3-3	-35,4017F	7,6527A	242A
15	UENF 14-4-3	-18,5778C	3,2880E	118E
16	UENF 14-6-3	-15,8417B	4,0057D	129E
17	UENF 15-23-4	-17,5172B	3,6955E	131E

Médias seguidas de uma mesma letra, na coluna, pertencem a um mesmo grupo, pelo critério de agrupamento de Scott-Knott, a 5% de probabilidade. A última avaliação ocorreu 37 dias após a semeadura.



Os genótipos 14 e 2 foram os que apresentaram as maiores declividades ( $\beta_1$ ) e 1, 13, 15 e 17 as menores, pelo critério de agrupamento de Scott-Knott a nível de 5% de probabilidade (tabela 5).

Em relação ao intercepto, os menores ou mais distantes da origem foram os genótipos 2 e 14 e os maiores os genótipos 1 e 13 (tabela 5).

Após a realização do teste de Scott-Knott, para as médias do intercepto ( $\beta_0$ ) e da declividade ( $\beta_1$ ), a nível de 5% de probabilidade, foram formados cinco grupo para a inclinação ( $\beta_1$ ) e seis para o intercepto ( $\beta_0$ ). A seguir foram formados cinco grupos onde o intercepto ( $\beta_0$ ) e a declividade ( $\beta_1$ ) pertenciam simultaneamente aos mesmos grupos.

Grupo 1: Genótipos 01 e 13;

Grupo 2: Genótipos 02 e 14;

Grupo 3: Genótipos 03 e 12;

Grupo 4: Genótipos 04,09, e 10;

Grupo 5: Genótipos 06, 07 e 11.

O genótipo 15 não foi incluído em nenhum grupo devido ao coeficiente de variação muito alto, principalmente em relação à inclinação ( $\beta_1$ ), que representa a taxa de crescimento.

Com os genótipos que não pertenciam a nenhum dos grupos anteriores foi formado o grupo 6 com os genótipos 05, 08, 16 e 17.

A análise de variância da regressão para todos os grupos formados, foi significativa ao nível de 1% de probabilidade.

No modelo geral utilizando as variáveis Dummy, os parâmetros foram testados utilizando o teste “t” de Student a nível de 5% de probabilidade.

Para o grupo 1, formados pelos genótipos 01 e 13, obteve-se a seguinte equação de regressão (equação 75):

$$\hat{y} = -13,4739 + 0,4641D_1Dia + 3,2399Dia \quad (75)$$

Quando  $D_1$  igual a zero, a equação representa o genótipo 01, se  $D_1$  igual a um, representa o genótipo 13, obtendo-se assim, as equações 76 e 77 para os genótipos 01 e 13 respectivamente

$$\hat{y} = -13,4739 + 3,2399Dia \quad (\text{Genótipo 01}) \quad (76)$$

$$\hat{y} = -13,4739 + 3,7040Dia \quad (\text{Genótipo 13}) \quad (77)$$

Em relação ao grupo 01, observa-se que não houve diferença ao nível de 5% de probabilidade em relação ao intercepto ( $\beta_0$ ), sendo -13,4739 o intercepto comum para os genótipos 01 e 13. E, 3,2399cm/dia o valor da declividade ( $\beta_1$ ) para o genótipo 01 e 0,4641cm/dia o diferencial de declividade do genótipo 13 em relação ao genótipo 01. A declividade é a taxa média de crescimento para o período de 32 dias entre a primeira e a última avaliação.

O coeficiente de determinação  $R^2$  corrigido obtido para a equação de regressão (equação 75) foi igual a 0,838013, significando que 83,8013% da variação na altura do feijão-de-vagem é explicada pela equação de regressão.

A equação de regressão (equação 78) representa o grupo 2, formado pelos genótipos 02 e 14.

$$\hat{y} = -36,9136 + 0,43226D_1Dia + 7,28715Dia \quad (78)$$

Quando  $D_1$  igual a zero, a equação representa o genótipo 02, se  $D_1$  igual a um, representa o genótipo 14, obtendo-se assim, as equações 79 e 80 para os genótipos 02 e 14 respectivamente.

$$\hat{y} = -36,9136 + 7,28715Dia \quad (\text{Genótipo 02}) \quad (79)$$

$$\hat{y} = -36,9136 + 7,71941Dia \quad (\text{Genótipo 14}) \quad (80)$$

Em relação ao intercepto ( $\beta_0$ ), não houve diferença significativa ao nível de 5% probabilidade entre os genótipos 02 e 14, ou seja, estatisticamente os dois genótipos tem a declividade comum igual a -36,9136cm/dia. A taxa de crescimento para o genótipo 14 para o período de 32 dias entre a primeira e a última avaliação foi de 7,28715cm/dia e o diferencial da taxa de crescimento do genótipo 14 em relação ao genótipo 02 foi de 0,43226cm/dia.

O coeficiente de determinação  $R^2$  corrigido obtido para a equação de regressão (equação 78) foi igual a 0,840243, o que significa que 84,0243% da variação na altura do feijão-de-vagem é explicada pela equação de regressão.

A equação de regressão (equação 81) representa o grupo 3, formado pelos genótipos 03 e 12:

$$\hat{y} = -28,7567 + 6,61552Dia \quad (\text{genótipos 03 e 12}) \quad (81)$$

Em relação à taxa de crescimento (inclinação da reta) e ao intercepto, não houve diferença significativa ao nível de 5% de probabilidade entre os genótipos 03 e 12. Pode-se concluir que em relação à taxa de crescimento e ao intercepto os genótipos, comportam-se como se fossem apenas um. O coeficiente de determinação  $R^2$  para o grupo foi de 0,864412 ou 86,4412% da variação na altura foi explicado pela equação de regressão.

O grupo 4, formado pelos genótipos 04, 09 e 10, foi representado pela equação 82.

$$\hat{y} = -21,7935 + 0,46356D_1Dia + 4,24565Dia \quad (82)$$

Quando  $D_1$  for igual a 1, a equação de regressão (equação 82) representa o genótipo 09 e, quando igual a zero, os genótipos 04 e 10, representados respectivamente, pelas equações 83 e 84:

$$\hat{y} = -21,7935 + 4,71921Dia \quad (\text{Genótipo 09}) \quad (83)$$

$$\hat{y} = -21,7935 + 4,24565Dia \quad (\text{Genótipos 04 e 10}) \quad (84)$$

Com  $0,46356cm/Dia$  representando o diferencial da taxa de crescimento do genótipo 09 em relação aos genótipos 04 e 10, não houve diferença entre os interceptos ao nível de 5% de probabilidade. Com coeficiente de determinação corrigido  $R^2$  igual a 0,784196 ou, 78,4196% da variação na altura das plantas de feijão-de-vagem foi explicado pela equação de regressão.

Para o grupo 5, formado pelos genótipos 06, 07 e 11, não houve diferença significativa ao nível de 5% de probabilidade em relação ao intercepto e a declividade (taxa de crescimento), com a equação 85 representando todos os genótipos citados anteriormente. O grupo 5, representado pelos genótipos 06, 07

e 11, em relação à taxa de crescimento, pode ser considerado como um único genótipo:

$$\hat{y} = -28,6883 + 5,8378Dia \quad (\text{Genótipos 06, 07 e 11}) \quad (85)$$

O coeficiente de determinação corrigido  $R^2$  foi de 0,827688, ou seja, 82,7688% da variação na altura das plantas de feijão vagem é explicado pela equação de regressão.

Para o grupo 6, formado pelos genótipos 05, 08, 16 e 17, não houve diferença significativa ao nível de 5% de probabilidade entre os intercepto, porém, existiu diferença significativa ao nível de 5% de probabilidade entre todos os genótipos, em relação à taxa de crescimento (declividade). Com a equação de regressão (equação 86) representando o grupo.

$$\hat{y} = -19,16014 + 5,557055Dia - 1,41633D_1Dia - 1,41841D_2Dia - 1,80254D_3Dia \quad (86)$$

Tabela 6 – Valores das variáveis Dummy para obtenção das equações de regressão dos genótipos

Variáveis Dummy			Genótipos
D <sub>1</sub>	D <sub>2</sub>	D <sub>3</sub>	
0	0	0	05
1	0	0	08
0	1	0	16
0	0	1	17

De acordo com a tabela 6, as representações das equações de regressão para os genótipos 05, 08, 16 e 17, são respectivamente:

$$\hat{y} = -19,16014 + 5,557055Dia \quad (\text{Genótipo 05}) \quad (87)$$

$$\hat{y} = -19,16014 + 4,140725Dia \quad (\text{Genótipo 08}) \quad (88)$$

$$\hat{y} = -19,16014 + 4,138645Dia \quad (\text{Genótipo 16}) \quad (89)$$

$$\hat{y} = -19,16014 + 3,754515Dia \quad (\text{Genótipo 17}) \quad (90)$$

Com os valores  $-1,41633$ ;  $-1,41841$  e  $-1,80254$ , representa o diferencial da taxa de crescimento das plantas de feijão-vagem dos genótipos 08, 16 e 17 em relação ao genótipo 05, respectivamente.

E, com coeficiente de determinação para o grupo 6 igual a  $0,814235$ , ou seja,  $81,4235\%$  da variação das alturas das plantas de feijão-de-vagem é explicada pela equação de regressão 86.

Vasconcelos et al (2010), utilizaram com sucesso o método de identidade de modelos para agrupar equações de regressão linear simples para estudo de adaptabilidade e estabilidade de genótipos de milho.

Figueiredo (2005), estudando o volume de madeira com e sem casca e volume total e comercial, concluiu que o emprego do método das variáveis Dummy possibilita a obtenção de diversas especificações de volumes.

A seguir tem-se o quadro com as equações para cada genótipo.

Tabela 7 – Genótipos e suas respectivas equações.

Genótipo	Equação
01	$\hat{y} = -13,4739 + 3,2399Dia$
02	$\hat{y} = -36,9136 + 7,28715Dia$
03	$\hat{y} = -28,7567 + 6,61552Dia$
04	$\hat{y} = -21,7935 + 4,24565Dia$
05	$\hat{y} = -19,16014 + 5,557055Dia$
06	$\hat{y} = -28,6883 + 5,8378Dia$
07	$\hat{y} = -28,6883 + 5,8378Dia$
08	$\hat{y} = -19,16014 + 4,140725Dia$
09	$\hat{y} = -21,7935 + 4,71921Dia$
10	$\hat{y} = -21,7935 + 4,24565Dia$
11	$\hat{y} = -28,6883 + 5,8378Dia$
12	$\hat{y} = -28,7567 + 6,61552Dia$
13	$\hat{y} = -13,4739 + 3,7040Dia$
14	$\hat{y} = -36,9136 + 7,71941Dia$
16	$\hat{y} = -19,16014 + 4,138645Dia$
17	$\hat{y} = -19,16014 + 3,754515Dia$

## 5. CONCLUSÕES

O uso das variáveis Dummy permitiu ajustar simultaneamente equações lineares de 1º grau, identificando a existência de diferenças nas respostas de crescimento em relação a altura para o feijão-de-vagem em função de diferentes períodos em Bom Jesus do Itabapoana, RJ.

## 6. REFERÊNCIAS BIBLIOGRÁFICAS

- ALMEIDA, S. (2011) Avaliação das características morfológicas e agronômicas de linhagens de feijão-de-vagem em Bom Jesus do Itabapoana, RJ, com potencial de recomendação. Dissertação (Mestrado em Produção vegetal) – Campos Goytacazes – RJ, Universidade Estadual do Norte Fluminense Darcy Ribeiro – UENF, 46p.
- ALVAREZ V., V. H.; NOVAES, R. F. de; BARROS, N. F. de; LOPES, A. S. (1999) Interpretação dos resultados das análises de solos. In: RIBEIRO, A. C.; GUIMARÃES, P. C. G.; ALVAREZ V. V. H (Ed.). Recomendações para o uso de corretivos e fertilizantes em Minas Gerais: 5ª aproximação. Viçosa, MG: Comissão de Fertilidade de Solo do Estado de Minas Gerais. 359 p.
- ARAUJO, L. C. (2011) Avaliação de linhagens melhoradas de feijão-de-vagem, em Bom Jesus do Itabapoana, RJ. Dissertação (Mestrado em Produção vegetal) – Campos Goytacazes – RJ, Universidade estadual do Norte Fluminense Darcy Ribeiro – UENF, 46p.
- AZEVEDO, P. R. M. de (1997) Modelos de Regressão Linear. Natal - RN: Editora da UFRN. 212p.

- BARBIN, D. (1993) Componentes de variância: teoria e aplicação. 2. ed. rev. e amp. Piracicaba: FEALQ. 120p.
- BARBIN, D. (2003) Planejamento e análise estatística de experimentos agronômicos. Arapongas: Midas. 208p.
- BARBOSA FILHO, M. P. B.; SILVA, O. F. (2001) Adubação de cobertura do feijoeiro irrigado com uréia fertilizante em plantio direto: um ótimo negócio. Piracicaba: POTAFOS. 20p.
- BANZATTO, D. A.; KRONKA, S. do N. (1995) Experimentação Agrícola. 3. ed. Jaboticabal: Funep. 247p.
- BUSSAB, W. O. (1988) Análise de Variância e Regressão: uma introdução. 2. ed. São Paulo: Editora Atual. 147p.
- CARRIJO, O. A.; SILVA, W. L. C.; MAROUELLI, W. A.; SILVA, H. R. (1999a) Tendências e desafios da fertirrigação no Brasil. In: FOLEGATTI, M. V. Fertirrigação: citrus, flores e hortaliças. Guaíba: Agropecuária, 1999, p.155-169
- CARRIJO, I, V.; FILGUEIRA, F. A. R.; TRANI, P. E. (1999b) Sugestões de adubação para hortaliças - Feijão-vagem (trepador). In: RIBEIRO, A. C.; GUIMARÃES, P. C. G.; ALVAREZ V. V. H (Ed.). Recomendações para o uso de corretivos e fertilizantes em Minas Gerais: 5ª aproximação. Viçosa, MG: Comissão de Fertilidade de Solo do Estado de Minas Gerais. 359 p.
- CECON, P. R.; SILVA, A. R. da; NASCIMENTO, M; FERREIRA, A. (2012) Métodos Estatísticos. Viçosa, Editora UFV. 229p. (Série Didática)
- CHARNET, R.; FREIRE, C. A. de L.; CHARNET, E. M. R.; BONVINO, H. (2008) Análise de Modelos de Regressão Linear com Aplicações. 2. ed. Campinas, SP: Editora Unicamp. 356p.



- CHATTERJEE, S.; HADI, A. S. (2006) Regression Analysis by Example. 4. ed. New Jersey: John Wiley & Sons, inc. 375p.
- CHOW, G. C. (1960) Tests of equality between sets of coefficients in two linear regressions. *Econometrica*, v.28, n.3, 591-605
- CRONQUIST, A. (1988). Devolution and classification of flowering plants. New York: Botanical Garden, 555 p.
- DRAPER, N. R.; SMITH, H. (1998) Applied Regression Analysis. 3rd ed. New York: John Wiley & Sons, inc. 706p.
- ELIAN, S. N. (1988) Análise de Regressão. São Paulo: Editora da USP. 232
- FERREIRA, D. F. (2009) Estatística Básica. 2. ed. rev. Lavras: Ed. UFLA. 664p.
- FIGUEIREDO, E. O. (2005) Método da variável fictícia para ajuste de modelos volumétricos estáveis e compatíveis em povoamentos florestais. Rio Branco: Embrapa Acre. 28p. (Embrapa Acre. Documentos, 95)
- FILGUEIRA, F. A. R. (2008) Novo Manual de olericultura: agrotecnologia moderna na produção e comercialização de hortaliças. 3. ed. rev. Viçosa: Universidade Federal de Viçosa, 421p.
- GRAYBILL, F. A. (1976) Theory and application of the linear model. Belmont: Duxbury. 704p.
- GUJARATI, D, N.; PORTER, D. C. (2011) Econometria básica. 5a. ed. Porto Alegre: Bookman. 920p.
- GUJARATI, D. (1970a) Use of Dummy variables in testing for equality between sets of coefficients in two linear regressions: a note. *The American Statistician*. v.24, n.1. 50-52.

- GUJARATI, D. (1970b) Use of Dummy variables in testing for equality between sets of coefficients in two linear regressions: a generalisation. *The American Statistician*. v.24, n.5. 18-21.
- HOFFMANN, R; VIEIRA, S. (1977) *Análise de Regressão: uma Introdução à Econometria*. 2. ed. São Paulo: Hucitec. 379p.
- IBGE – Instituto Brasileiro de Geografia e Estatística – IBGE – Cidades. Disponível em: <http://www.ibge.gov.br/cidadesat/topwindow.htm>. Acessado em 15/03/2012.
- KUTNER, M. H.; NACHTSHEIM, C. J.; NETER, J.; LI, W. (2005) *Applied Linear Statistical Models*. New York: McGraw-Hill/Irwin. 1396p.
- MAGALHÃES, S. R. S.; VEIGA, R. D.; SÁFADI, T.; CIRILLO, M. A.; MAGINI, M. (2004) Avaliação de Métodos para Comparação de Modelos de Regressão por Simulação de Dados. *Semina: Ciências Exatas e Tecnológicas, Londrina*, v. 25, n.2, p.117-122.
- MAGALHÃES, S. R.; ANDRADE, E. A. de (2009a) Teste para Verificar a Igualdade de Modelos de Regressão e uma Aplicação na Área Médica. *Exacta*, Belo Horizonte, v.2, n.1, p.34-41.
- MAGALHÃES, S. R., ANDRADE, E. A. de (2009b) Teste à Igualdade dos parâmetros de um Modelo de Regressão: uma Aplicação Especial das Variáveis Dummy. *Exacta*, Belo Horizonte, v.2, n.3, p.20 -25.
- MYERS, R. H. (1989) *Classical and Modern Regression with Applications*. 2. ed. Boston: Pws-Kent Publishing Company. 488p.
- NASCIMENTO, W. M. Temperatura x Germinação. Disponível em: <http://www.cnph.embrapa.br/public/textos/texto3.html> - Acesso em: 20/03/2012.

- PEIXOTO, N. (2002) Adaptabilidade e estabilidade em feijão-vagem de crescimento indeterminado. Horticultura Brasileira, Brasília, v.20, n.4, p.616-618.
- PIMENTEL-GOMES, F. (2000) Curso de Estatística Experimental. 14. ed. Piracicaba. 477p.
- RAMALHO, M. A. P.; FERREIRA, D. F.; OLIVEIRA, A. C. de (2000) Experimentação em genética e melhoramento de plantas. Lavras: UFLA, 303p.
- RAWLINGS, J. O.; PANTULA, S. G.; DICKEY, D. A. (1998) Applied Regression Analyses: a research tool. 2. ed. New York: Spring-Verlag, 658p.
- REGAZI, A. J. (1993) Teste para Verificar a Identidade de Modelos de Regressão e a igualdade de alguns parâmetros num modelo polinomial ortogonal. Revista Ceres, Viçosa, v.40, n.228, p.176-195
- REGAZI, A. J. (1999) Teste para Verificar a Identidade de Modelos de Regressão e a Igualdade de Parâmetros no Caso de Dados de Delineamentos Experimentais. Revista Ceres, Viçosa, Vol.46, n.266, p.383-409.
- RESENDE, M. D. V de. (2007) Matemática e estatística na análise de experimentos e no melhoramento genético. Colombo: Embrapa Florestas. 362p.
- SAEG (2007) Sistema para Análises Estatísticas, versão 9.1: Fundação Arthur Bernades – UFV.
- SANTOS, M. H. dos (2013) Melhoramento do Feijão-de-vagem: herança da resistência a *Xanthomonas axonopodis* pv. *plaseoli* e potencial agrônômico de linhas parcialmente endogâmicas obtidas pelo método SSD. Tese (doutorado em Produção Vegetal) – Campos dos Goytacazes – RJ, Universidade Estadual do Norte Fluminense Darcy Ribeiro) – UENF, 117p.

- SCOTT, A. J.; KNOTT, M. A. (1974) A cluster analysis method for grouping means in the analysis of variance. *Bometrics*, v.30, n.3, 507-512
- SEARLE, S. R. (1971) *Linear models*. New York: John Wiley & Sons, Inc. 532p.
- SEBER, G. A. F.; LEE, A. J. (2003) *Linear Regression Analysis*. 2. ed. New Jersey: John Wiley & Sons, inc. 557p.
- SNEDECOR, G. W.; COCHRAN, W. G. (1967) *Statistical Methods*. 6. ed. Iowa: The Iowa State University Press. 593p.
- SOUZA, G. S. (1998) *Introdução aos Modelos de Regressão Linear e Não-linear*. Brasília: Embrapa – Serviço de Produção de Informação. 633p.
- SOUZA, V. C.; LORENZI, H. (2008) *Botânica sistemática: guia ilustrado para identificação das famílias de Fanerógamas nativas e exóticas no Brasil, baseado em APG II*. 2. ed. Nova Odessa, SP: Instituto Plantarum. 704p.
- STEEL, R. G. D.; TORRIE, J. H.; DICKEY, D. A. (1997) *Principles and procedures of Statistics: a biometrical approach*. 3. ed. New York: McGraw-Hill. 666p.
- TAIZ, L.; ZEIGER, E. (2004) *Fisiologia vegetal*. Tradução de SANTARÉM, E. R.; MARIATH, J. E de A.; ASTARITA, L. V.; DILLENGURG, L. R.; ROSA, L. M. G.; OLIVEIRA, P. L. de. 3.ed. Porto Alegre: Artmed. 719p.
- VANI, S. M. (1998) *Modelos de Regressão: Estatística Aplicada*. São Paulo: Legnar Informática & editora. 177p.
- VALLE, P. O. (2002) Testes à estabilidade dos parâmetros de um modelo de regressão: uma aplicação especial dos regressores dummy. *Revista de Estatística*, III, p.17-40

VASCONCELOS, E. S. de; CRUZ, C. D.; REGAZZI, A. J.; BHERING, L. L.; ROSADO, T. B.; VASCONCELOS, F. S. de (2010) Agrupamento de modelos de regressão da análise de adaptabilidade e estabilidade de genótipos. *Pesq. Agrop. Bras.*, Brasília, v.45, n.12, p.1357-1362.

VIEIRA, S. (1999) *Estatística experimental*. 2. ed. São Paulo: Atlas. 185p.

VIGGIANO, J. (1990) Produção de semente de feijão-vagem, In: CASTELLANE, P. D.; NICOLOSI, W. R.; HASEGAWA, H. *Produção de semente de hortaliças*. Jaboticabal: FCAV/FINEP. p. 127-140

WEISBERG, S. (2005) *Applied Linear Regression*. 3. ed. New Jersey: John Wiley & Sons, inc. 310p.

.

## 7. APÊNDICE

Modelos inicialmente ajustados para o grupo 1

$$\hat{y} = -12,3441 - 2,25889D_1 + 0,56375D_1Dia + 3,1901Dia$$

Modelos inicialmente ajustados para o grupo 2

$$\hat{y} = -38,4256 + 3,02389D_1 + 0,29886D_1Dia + 7,35385Dia$$

Modelos inicialmente ajustados para o grupo 3

$$\hat{y} = -29,2961 + 1,07889D_1 - 0,11708D_1Dia + 6,67405Dia$$

Modelos inicialmente ajustados para o grupo 4

$$\hat{y} = -21,2472 - 1,38111D_1 - 0,257778D_2 + 0,43667D_1Dia - 0,16427D_2Dia + 4,30937Dia$$

Modelos inicialmente ajustados para o grupo 5

$$\hat{y} = -28,1939 - 2,74556D_1 + 1,26222D_2 - 0,13500D_1Dia - 0,19312D_2Dia + 5,94717Dia$$

Modelos inicialmente ajustados para o grupo 6

$$\hat{y} = -24,2006 + 5,11944D_1 + 8,35889D_2 + 6,68333D_3 - 1,64219D_1Dia - 1,78719D_2Dia - 2,09740D_3Dia + 5,79292Dia$$

## Grupo 1. Genótipos 01 e 13

Tabela 1A – Análise dos parâmetros de regressão referente ao grupo 01

Nome	Coefficiente	Valor t	Probabilidade
Constante	13,4793		
Dia	0,4641	35,10465	0,0001
D <sub>1</sub> Dia	3,2399	5,27388	0,0001

Tabela 2A -Análise de variância da regressão linear referente ao grupo 01

Fonte de variação	G L	Quadrado médio	F	Significância
Devido a regressão	2	234.964,1000	929,61	0,0001
Independ. da regressão	357	252,7552		

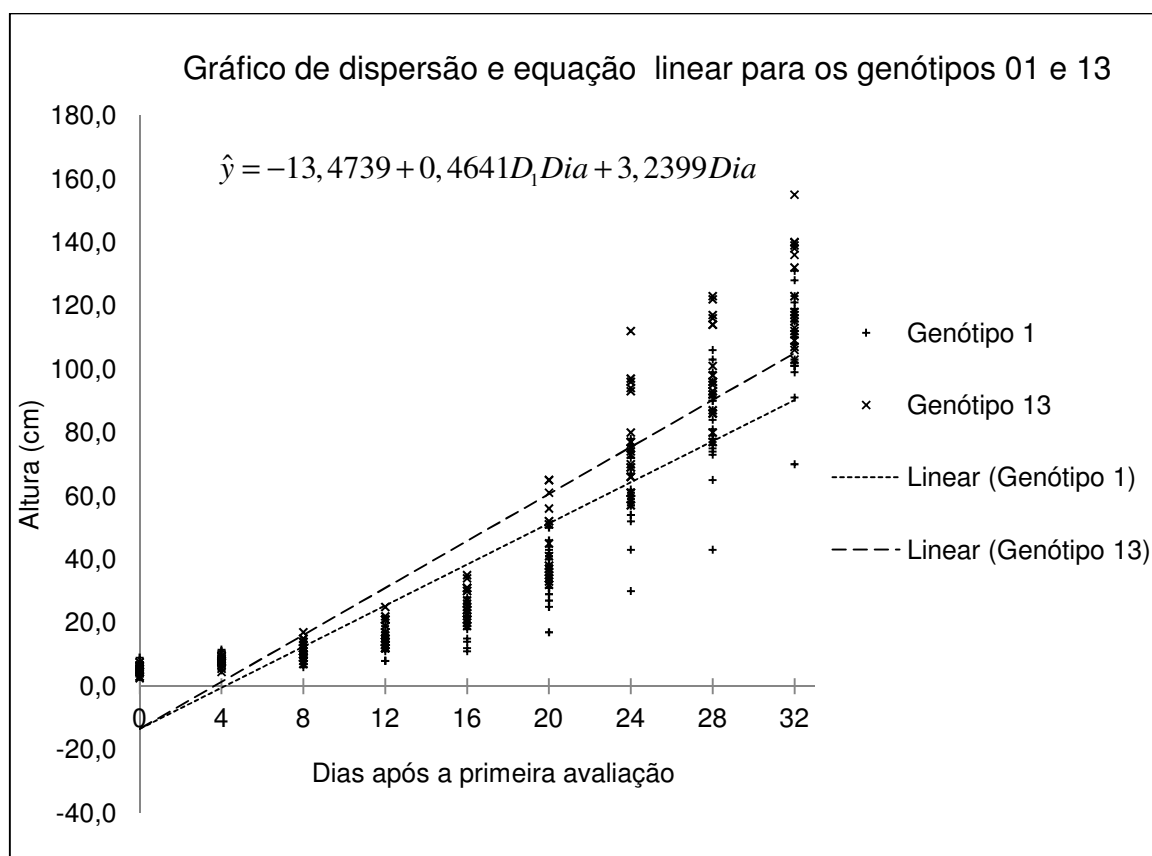


Figura 1A – Gráfico de dispersão e equação linear ajustada para o grupo 1.

## Grupo 2. Genótipos 02 e 14

Tabela 3A – Análise dos parâmetros de regressão referente ao grupo 02

Nome	Coefficiente	Valor t	Probabilidade
Constante	-36,9136		
Dia	7,28715	37,064399	0,0001
D <sub>1</sub> Dia	0,43226	2,305906	0,0001

Tabela 4A - Análise de variância da regressão linear referente ao grupo 02

Fonte de variação	G L	Quadrado médio	F	Significância
Devido a regressão	2	1.083.995,000	929,61	0,0001
Independ. da regressão	357	1.146,988		

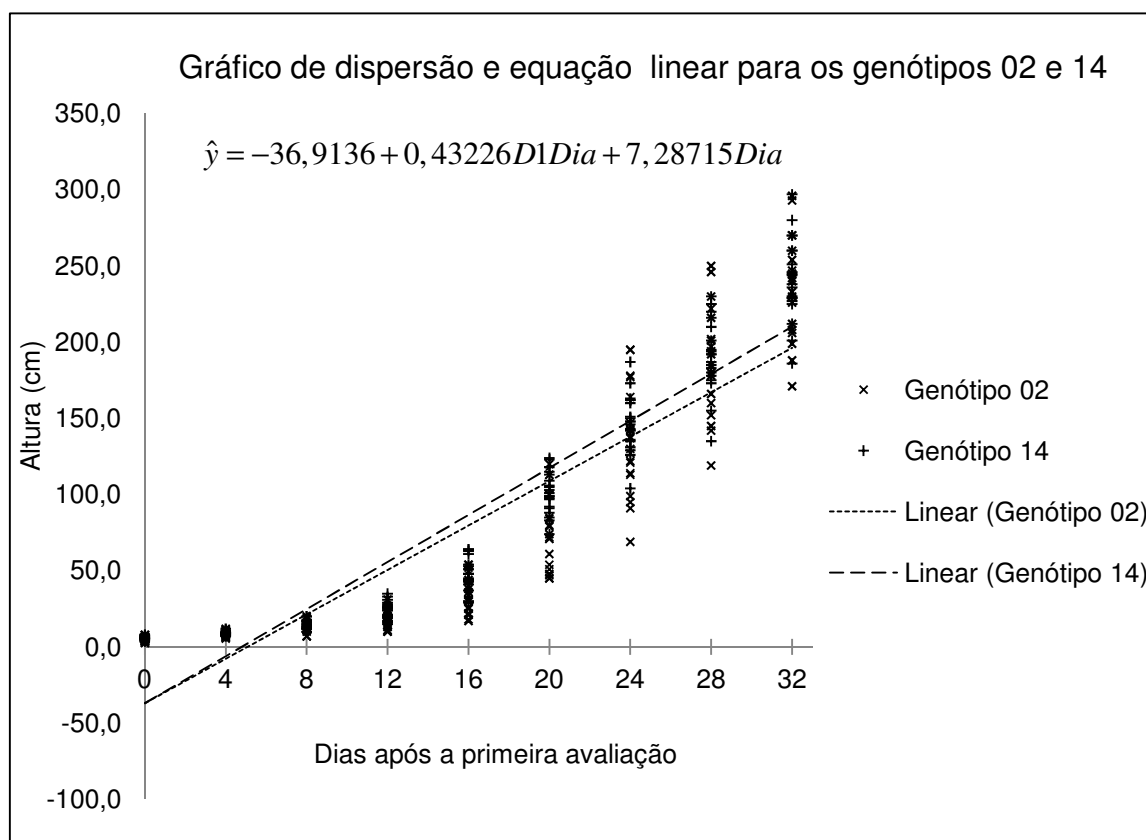


Figura 1A – Gráfico de dispersão e equação linear ajustada para o grupo 2.



## Grupo 3. Genótipos 03 e 12

Tabela 5A – Análise dos parâmetros de regressão referente ao grupo 03

Nome	Coeficiente	Valor t	Probabilidade
Constante	-28,7567		
Dia	6,61552	47,85119i2	0,0001

Tabela 6A - Análise de variância da regressão linear referente ao grupo 03

Fonte de variação	G L	Quadrado médio	F	Significância
Devido a regressão	1	1.680.580	2.289,74	0,0000
Independ. da regressão	358	733,9624		

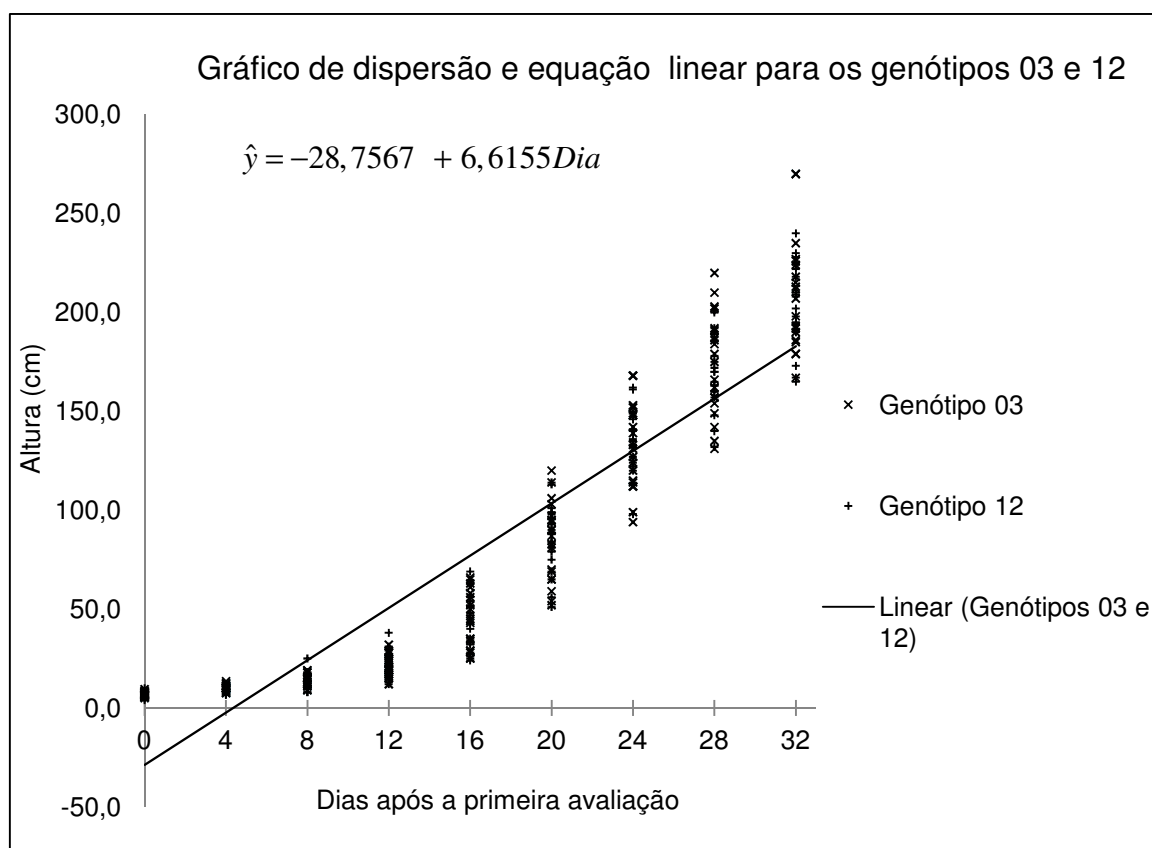


Figura 3A – Gráfico de dispersão e equação linear ajustada para o grupo 3

## Grupo 4. Genótipos 04, 09 e 10

Tabela 7A – Análise dos parâmetros de regressão referente ao grupo 04

Nome	Coeficiente	Valor t	Probabilidade
Constante	-21,7935		
Dia	4,24565	39,725307	0,0001
D <sub>1</sub> Dia	0,463557	4,037865	0,0001

Tabela 8A - Análise de variância da regressão linear referente ao grupo 04

Fonte de variação	G L	Quadrado médio	F	Significância
Devido a regressão	2	562.287,9	980,32	0,0000
Independ. da regressão	537	573,5768		

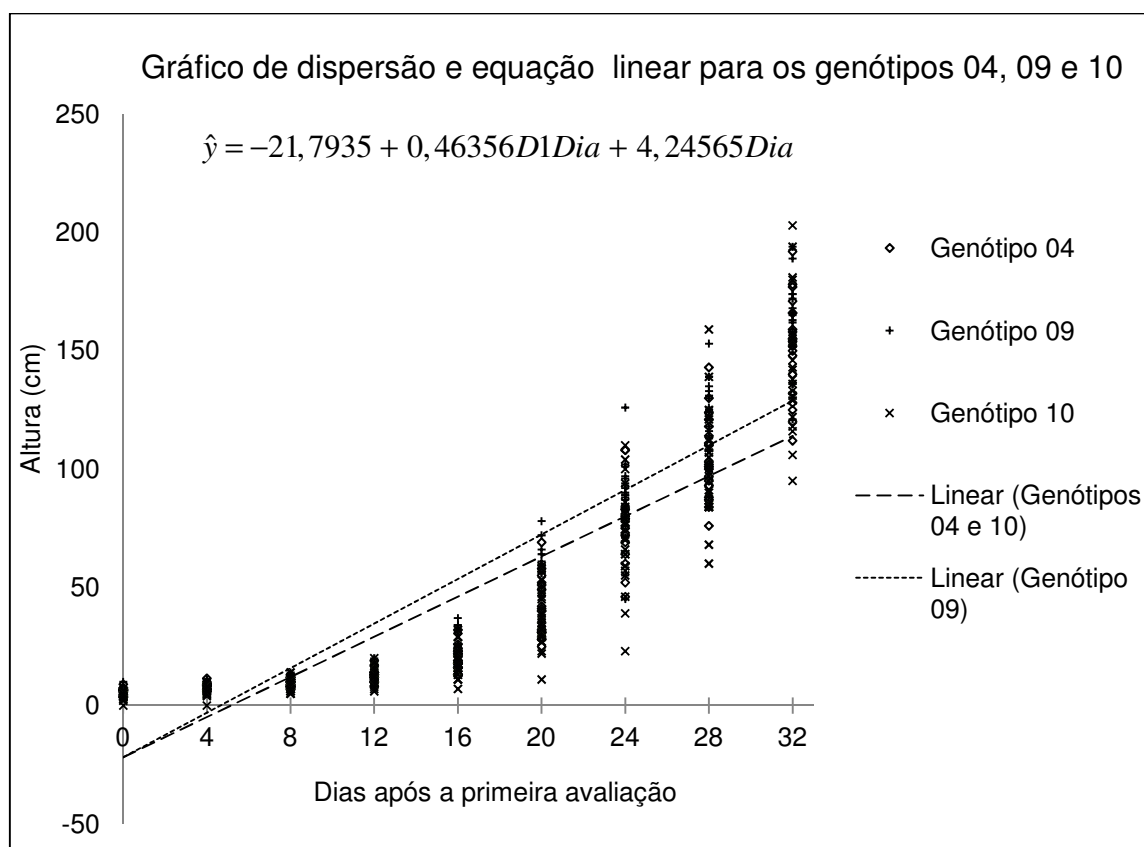


Figura 4A – Gráfico de dispersão e equação linear ajustada para o grupo 4

## Grupo 5. Genótipos 06, 07 e 11

Tabela 9A – Análise dos parâmetros de regressão referente ao grupo 05

Nome	Coefficiente	Valor t	Probabilidade
Constante	- 28,6883		
Dia	5,8378	50,892518	0,0001

Tabela 10A - Análise de variância da regressão linear referente ao grupo 05

Fonte de variação	G L	Quadrado médio	F	Significância
Devido a regressão	1	1.963.011	2.590,05	0,0000
Independ. da regressão	538	757,9052		

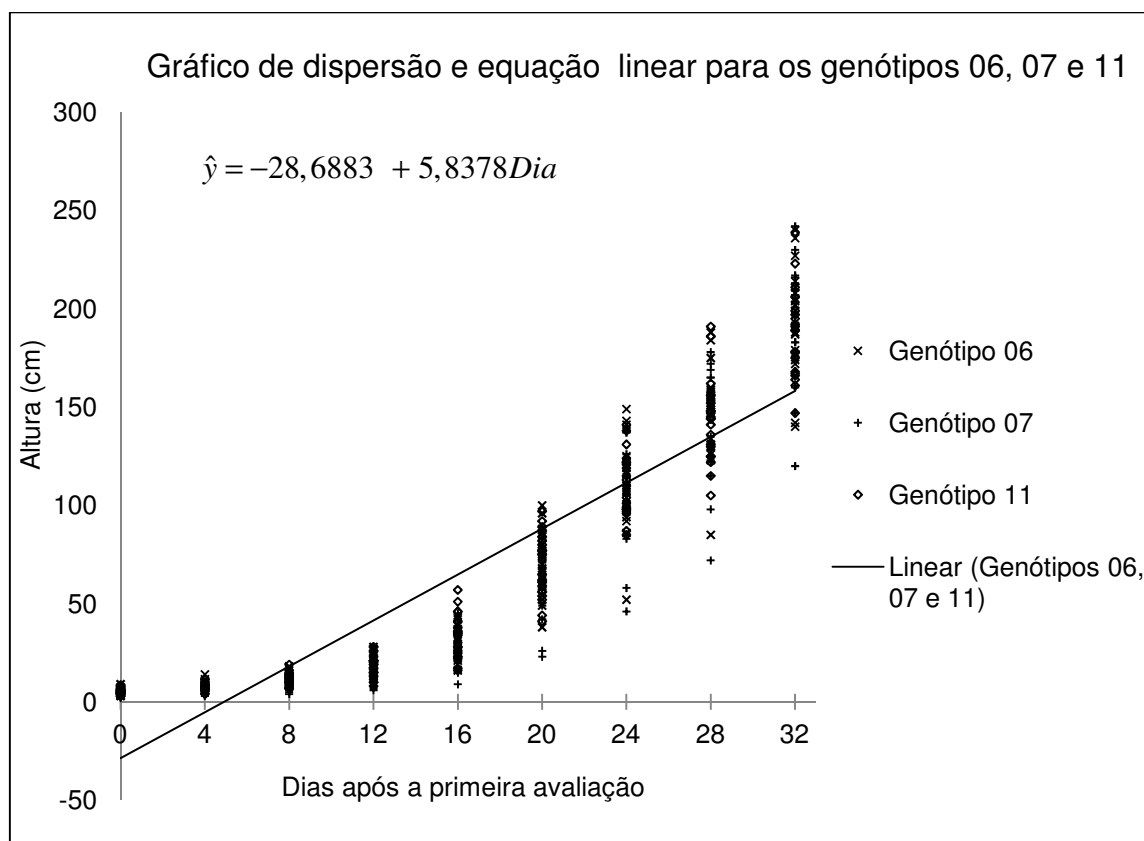


Figura 5A – Gráfico de dispersão e equação linear ajustada para o grupo 5

## Grupo 6. Genótipos 05, 08,16 e 17

Tabela 11A – Análise dos parâmetros de regressão referente ao grupo 066

Nome	Coeficiente	Valor t	Probabilidade
Constante	- 19,16014		
Dia	5,557055	49,6921	0,0000
D <sub>1</sub> Dia	- 1,41633	- 11,3005	0,0000
D <sub>2</sub> Dia	- 1,41841	- 11,3172	0,0000
D <sub>3</sub> Dia	- 1,80254	- 14,3820	0,0000

Tabela 12A - Análise de variância da regressão linear referente ao grupo 06

Fonte de variação	G L	Quadrado médio	F	Significância
Devido a regressão	4	404.471,0	788,8718	0,0000
Independ. da regressão	715	512,7208		

Tabela13A – Coeficientes de Determinação dos grupos

Grupo	R <sup>2</sup>	R <sup>2</sup> <sub>ajustado</sub>
1	0,838915	0,838013
2	0,841133	0,840243
3	0,864790	0,864412
4	0,784997	0,784196
5	0,828008	0,827688
6	0,815269	0,814235

Tabela 14A – Modelo da planilha para a determinação dos parâmetros de regressão, de cada planta, de cada genótipo e dos genótipos

Genótipo	Planta	Dia	Altura
1	1	0	6,5
1	1	4	11,0
1	1	8	14,0
1	1	12	15,0
1	1	16	26,0
1	1	20	50,0
1	1	24	78,0
1	1	28	84,0
1	1	32	108,0
1	2	0	7,5
1	2	∴	∴
1	2	32	99,0
∴	∴	∴	∴
1	20	32	91,0
∴	∴	∴	
17	19	0	4,5
17	∴	∴	∴
17	19	32	94,0
17	20	0	5,5
17	20	4	9,5
17	20	8	11,0
17	20	12	15,0
17	20	16	17,0
17	20	20	30,0
17	20	24	58,0
17	20	28	85,0
17	20	32	112,0

Tabela 15A – Modelo de planilha para o teste de Scott-Knott para os genótipos, em função dos betas ( $\beta_0$  e  $\beta_1$ )

Genótipos	Repetição	( $\beta_0$ )	( $\beta_1$ )
1	1	-8,9222	3,2833
1	2	-5,6778	2,7646
1	3	-7,0889	1,8042
1	4	-13,7111	3,1958
1	5	-10,7667	2,9708
1	:	:	:
1	18	-15,0333	3,6208
1	19	-18,1889	3,9354
1	20	,-11,9333	2,5583
2	1	-37,3222	6,8187
2	:	:	:
2	20	-43,7111	6,9958
:	:	:	:
17	1	-13,2556	3,1583
17	2	-19,0111	3,4104
17	3	-21,5444	4,0792
17	4	-12,0889	2,7833
17	5	-17,3000	3,4771
17	6	-20,4000	4,3479
17	:	:	:
17	17	-18,5444	6,6729
17	18	-19,1444	3,9917
17	19	-10,5000	2,7083
17	20	-12,6556	3,1729

Tabela 16A – Modelo da planilha para a para a análise de regressão e de variância utilizando as variáveis Dummy: Caso de dois genótipos

Genótipo	Planta	Dia	Altura	Dummy
1	1	0	6,5	0
1	1	4	11,0	0
1	1	8	14,0	0
1	1	12	15,0	0
1	1	16	26,0	0
1	1	20	50,0	0
1	1	24	78,0	0
1	1	28	84,0	0
1	1	32	108,0	0
1	2	0	7,5	0
1	2	⋮	⋮	0
1	2	32	99,0	0
1	⋮	⋮	⋮	0
1	20	32	91,0	0
13	1	0	4,5	1
13	⋮	⋮	⋮	1
13	1	32	118,0	1
13	⋮	⋮	⋮	1
13	20	0	5,5	1
13	20	4	7,5	1
13	20	8	12,0	1
13	20	12	21,0	1
13	20	16	31,0	1
13	20	20	51,0	1
13	20	24	77,0	1
13	20	28	98,0	1
13	20	32	117,0	1

Tabela 17A – Modelo da planilha para a para a análise de regressão e de variância utilizando as variáveis Dummy: Caso de três genótipos

Genótipo	Planta	Dia	Altura	Dummy 1	Dummy 2
4	1	0	7,0	0	0
4	1	4	10,0	0	0
4	1	8	10,0	0	0
4	1	12	13,0	0	0
4	1	16	22,0	0	0
4	1	20	42,0	0	0
4	1	24	79,0	0	0
4	1	28	93,0	0	0
4	1	32	120,0	0	0
4	2	0	4,5	0	0
4	2	∴	∴	0	0
4	2	32	150,0	0	0
4	∴	∴	∴	0	0
4	20	32	140,0	0	0
9	1	0	6,5	1	0
9	∴	∴	∴	1	0
9	1	32	154,0	1	0
9	∴	∴	∴	1	0
9	20	0	4,0	1	0
9	∴	∴	∴	1	0
9	20	32	136,0	1	0
10	1	0	5,0	0	1
10	∴	∴	∴	0	1
10	1	32	136,0	0	1
10	20	0	5,0	0	1
10	20	4	6,0	0	1
10	∴	∴	∴	0	1
10	20	28	109,0	0	1
10	20	32	155,0	0	1