

DETERMINAÇÃO DA SUPOSIÇÃO DA NORMALIDADE DOS PARÂMETROS PRODUTIVIDADE EM CUSTO DE PRODUÇÃO DE CAFÉS DO BRASIL

JOSÉ C. SOBRINHO¹, FERNANDO B. MENEGHETTI², DIOGO J. CARDOSO³, JADIR N. SILVA⁴.

¹ Engenheiro Agrícola, Professor Associado, UFSM – Campus Silveira Martins. Depto. Multidisciplinar, Silveira Martins – RS. Tel. 55 3224 4701, jcardosos@yahoo.com.br

² Bacharel em Sistema de Informação. Analista de TI da UFSM Campus Silveira Martins. Silveira Martins – RS.

³ Acadêmico do Curso de Ciência da Computação. Campus Central da UFSM – Santa Maria – RS

⁴ Matemático, Professor Titular – Departamento de Engenharia Agrícola – DEA/UFV – Viçosa – MG.

Apresentado no
XLIV Congresso Brasileiro de Engenharia Agrícola - CONBEA 2015
13 a 17 de setembro de 2015- São Pedro- SP, Brasil.

RESUMO: Objetivou-se determinar a suposição da normalidade das variáveis que compõem o custo de produção de cafés no Brasil, bem como da produtividade do grão nas 15 regiões produtoras de café no Brasil. As variáveis avaliadas foram produtividade do grão, custeio da lavoura, despesas com pós-colheita, despesas financeiras, depreciações, outros custos fixos e renda de fatores, sendo a produtividade do grão expressa em $\text{kg}\cdot\text{ha}^{-1}$ e as demais em $\text{R}\$.ha^{-1}$. Foram utilizados os parâmetros relativos ao café arábica e robusta (conilon) publicados pela CONAB em abril de 2014, referentes à safra de 2013/14. Para avaliar a normalidade das distribuições, utilizaram-se o teste de Shapiro-Wilk, correlação r_Q do Q-Q plot e procedimento gráfico Q-Q plot. A correlação r_Q do teste de Q-Q plot apresentou os mesmos resultados obtidos pelo teste de Shapiro-Wilk, seja ele apontando a distribuição como normal ou não, contudo, o procedimento gráfico Q-Q-plot não apresentou resultados conclusivos, uma vez que foram obtidos vários pontos fora da curva. As distribuições das variáveis: custeio de lavoura, despesas com pós-colheita e renda de fatores apresentaram distribuição normal. As distribuições das variáveis produtividade do grão, despesas financeiras, depreciações e outros custos fixos não apresentaram distribuição normal.

PALAVRAS-CHAVE: grãos, correlação, estatística.

DETERMINATION OF THE ASSUMPTION OF NORMALITY OF THE PARAMETERS PRODUCTIVITY AND COST OF PRODUCTION OF COFFEE IN BRAZIL.

ABSTRACT: This study aimed to determine the assumption of normality of the variables that make up the cost of production of coffee in Brazil, and the grain productivity in 15 coffee producing regions in Brazil. The variables evaluated were grain productivity, crop costing, post-harvest expenses, financial expenses, depreciation, other fixed costs and income factors, grain productivity is expressed in $\text{kg}\cdot\text{ha}^{-1}$ and the others in $\text{R}\$.ha^{-1}$. It was used the parameters related to arabica and robusta coffee (conilon) published by CONAB in April 2014, referring to the harvest of 2013/14. To assess the normality of the distributions, we used the Shapiro-Wilk test, Q-Q plot correlation and Q-Q plot graphic procedure. The r_Q correlation from the Q-Q plot test presented the same results obtained using the Shapiro-Wilk test, indicating whether or not as a normal distribution, however, the Q-Q plot graphic procedure did not present conclusive results, since several points were obtained outside of the curve. The distributions of the variables: crop cost, post-harvest expenses and income factors were normally distributed. The distributions of the variables: grain productivity, financial expenses, depreciation and other fixed costs were not normally distributed.

KEYWORDS: grains, correlation, statistical

INTRODUÇÃO: Devido às suas características o café teve rápida adaptação às condições climáticas e de solo do Brasil e isto fez do café, um produto de grande relevância no mercado. Hoje o café é um dos principais itens de exportação do País. No início, era restrito ao Pará e Maranhão, hoje a produção de café expandiu-se e destacam-se os estados de MG, ES, SP, BA, PR e RO.

A suposição da normalidade dos dados amostrais ou experimentais é uma condição exigida para a realização de muitas inferências válidas a respeito de parâmetros populacionais. O uso da distribuição normal é muito importante na estatística seja na teórica ou na aplicada. Aplicações em distribuições de dados coletados da natureza, por exemplo, variáveis físicas, biológicas ou psicológicas comportam-se de modo aproximadamente simétrico, podendo ser bem representadas por essa distribuição. Quando se verifica a suposição da normalidade em um conjunto de dados visa-se avaliar o comportamento dos mesmos, ou seja, a forma que a distribuição assume, a qual se espera que seja gaussiana em forma de sino (SANTOS & FERREIRA, 2003). Na literatura encontram-se vários testes para testar a normalidade univariada dos dados de uma distribuição dentre eles usam-se (SHAPIRO & WILK, 1965), (JOHNSON & WICHERN, 1998) e (THODE JR, 2002), sendo o primeiro um teste baseado em estatísticas de ordem da distribuição normal e de seus respectivos valores esperados ou médios; o segundo é baseado na correlação entre os quantis observados e esperados da distribuição normal padrão e o terceiro avalia a distribuição quanto à sua normalidade ou não pelo procedimento gráfico.

Dessa forma, neste trabalho, avaliou-se a suposição da normalidade das variáveis que compõem o custo de produção de cafés no Brasil e da produtividade do grão nas 15 regiões produtoras de café no Brasil.

MATERIAL E MÉTODOS: Determinou-se a suposição da normalidade das variáveis que compõem o custo de produção de cafés no Brasil, bem como da produtividade do grão nas 15 regiões produtoras de café no Brasil. As variáveis avaliadas foram produtividade do grão kg ha^{-1} (*PROD*), despesas com custeio de lavoura (*DCL*), despesas com pós-colheita (*DPC*), despesas financeiras (*DF*), depreciações (*D*), outros custos fixos (*OCF*) e renda de fatores (*RF*), todas em $\text{R\$ ha}^{-1}$. O Cálculo da normalidade das distribuições foi feito utilizando o teste de Shapiro-Wilk (1965), coeficiente de correlação r_Q de Johnson e Wichern (1998) e pelo método gráfico Q-Q plot que consistiu da representação em um plano cartesiano os quantis amostrais observados versus os quantis esperados da distribuição normal. Para o teste de Shapiro-Wilk (*W*) as hipóteses foram:

H_0 : A amostra provém de uma população Normal;

H_1 : A amostra não provém de uma população Normal;

Ocorrerá a rejeição de H_0 ao nível de significância 5% com tamanho de amostra 15 se W_{calc} for menor que $W_{5\%}$ tabelado com 15 graus de liberdade. Para o teste de normalidade usando o coeficiente r_Q , a hipótese de normalidade deve ser rejeitada se o valor de r_Q calculado for menor que o valor crítico correspondente obtido por Johnson e Wichern (1998) a α igual a 5% de probabilidade com 15 graus de liberdade. O método gráfico Q-Q plot consistiu em representar em um plano cartesiano os quantis amostrais observados versus os quantis esperados da distribuição normal.

RESULTADOS E DISCUSSÃO: Os parâmetros que apresentaram normalidade em suas distribuições foram despesas com custeio de lavoura (*DCL*), despesas com pós-colheita (*DPC*) e renda de fatores (*RF*). A normalidade dos fatores deu-se tanto para o estatística *W* de Shapiro-Wilk (1965) quanto para o teste de r_Q . O teste gráfico apresentado nas Figuras 1 e 2 mostra que para estas variáveis não há pontos fora da curva, o que indica que as variáveis apresentam distribuição normal.

Tabela 1 – Valores calculados para a estatística *W* de Shapiro-Wilk (1965) e de r_Q para os parâmetros produtividade, custeio de lavoura, despesas com pós-colheita, despesas financeiras, depreciações, outros custos fixos e renda de fatores para cafés produzidos em 15 regiões do Brasil.

Teste	<i>PROD</i>	<i>DCL</i>	<i>DPC</i>	<i>DF</i>	<i>D</i>	<i>OCF</i>	<i>RF</i>
<i>W</i>	0,768 ns	0,910*	0,945*	0,646 ns	0,630 ns	0,588 ns	0,953*
r_Q	0,876 NS	0,953**	0,968**	0,788 NS	0,782 NS	0,751 NS	0,978**

ns – não significativo ao nível de 5% de probabilidade pelo teste de Shapiro Wilk (1965).

* – significativo ao nível de 5% de probabilidade pelo teste de Shapiro Wilk (1965).

NS – valores de r_Q calculado menor que r_Q tabelado a 5% de probabilidade, não significativo.

** – valores de r_Q calculado maior que r_Q tabelado a 5% de probabilidade, significativo.

Na Tabela 2 apresentam-se os parâmetros produtividade em kg ha⁻¹ e custo de produção em R\$ ha⁻¹; cujos valores estão em forma de tabela primitiva.

Tabela 2. Parâmetros relativos à produtividade e custo de produção de acordo com a localidade em que o grão foi produzido.

Local	PROD	Custos de Produção – R\$ ha ⁻¹						
		DCL	DPC	DF	D	OCF	RF	CTL
1	3.000,00	6.503,92	536,00	341,80	1.638,61	105,11	1.040,27	10.165,71
2	3.600,00	10.338,52	374,43	859,26	825,77	90,50	606,91	13.095,39
3	1.200,00	6.553,29	209,22	297,92	494,48	75,11	483,03	8.113,05
4	3.300,00	8.755,88	891,76	217,39	3.198,04	431,32	1.231,16	14.725,55
5	1.800,00	9.442,08	338,58	324,15	812,98	78,98	394,32	11.391,09
6	1.800,00	7.249,38	337,50	278,81	725,95	95,17	409,29	9.096,10
7	1.800,00	9.270,69	475,20	252,96	752,00	102,96	419,70	11.273,51
8	1.440,00	6.754,85	288,48	136,60	765,17	35,67	448,21	8.428,98
9	1.800,00	7.502,86	537,66	253,66	509,61	16,78	652,58	9.473,15
10	1.800,00	7.411,09	579,00	287,18	937,53	62,33	859,83	10.136,96
11	1.500,00	8.614,42	464,54	293,01	908,25	70,27	889,11	11.239,60
12	1.800,00	8.738,96	268,36	350,49	1.072,57	77,82	322,38	10.830,58
13	1.800,00	7.585,37	543,57	255,47	541,12	38,83	783,77	9.748,13
14	1.200,00	3.186,13	109,52	154,20	463,68	9,85	158,43	4.081,81
15	1.500,00	3.603,81	167,26	189,65	581,78	6,52	195,43	4.744,45

1 – Luiz Eduardo Magalhães – BA; 2 – Pinheiro ES; 3 – Venda Nova do Emigrante – ES; 4 – Cristalina – GO; 5 – Guaxupé – MG – lavoura de serra; 6 – Guaxupé – MG – lavoura mecanizada – 7 – Guaxupé – MG – lavoura semi-mecanizada; 8 – Manhuaçu – MG – café semi-adensado; 9 – Patrocínio – MG – lavoura mecanizada; 10 – São Sebastião do Paraíso - 70% mecanizado; 11 – São Sebastião do Paraíso - 30% mecanizado; 12 – Londrina – PR – cultivo adensado; 13 – Franca – SP – Lavoura Mecanizada; 14 – Ji-Paraná – RO – café robusta; 15 – Rolim de Moura – RO – café robusta.

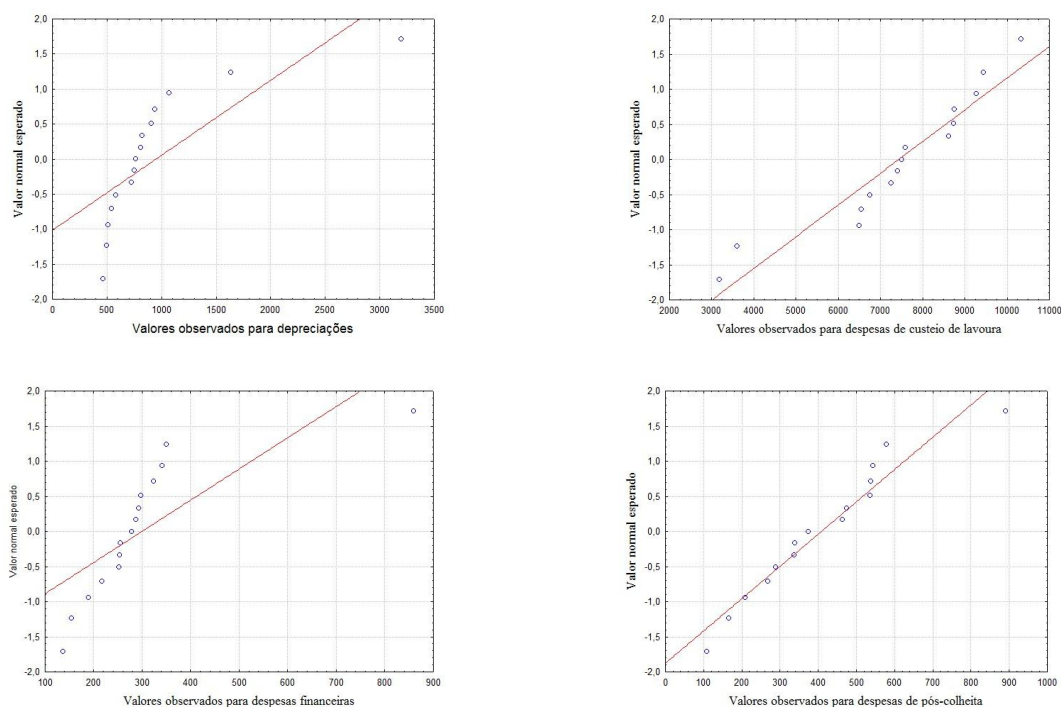


Figura 1 – Gráficos de dispersão entre valores observados para depreciações, custeio de lavoura, despesas financeiras e despesas de pós-colheita nas 15 localidades brasileira.

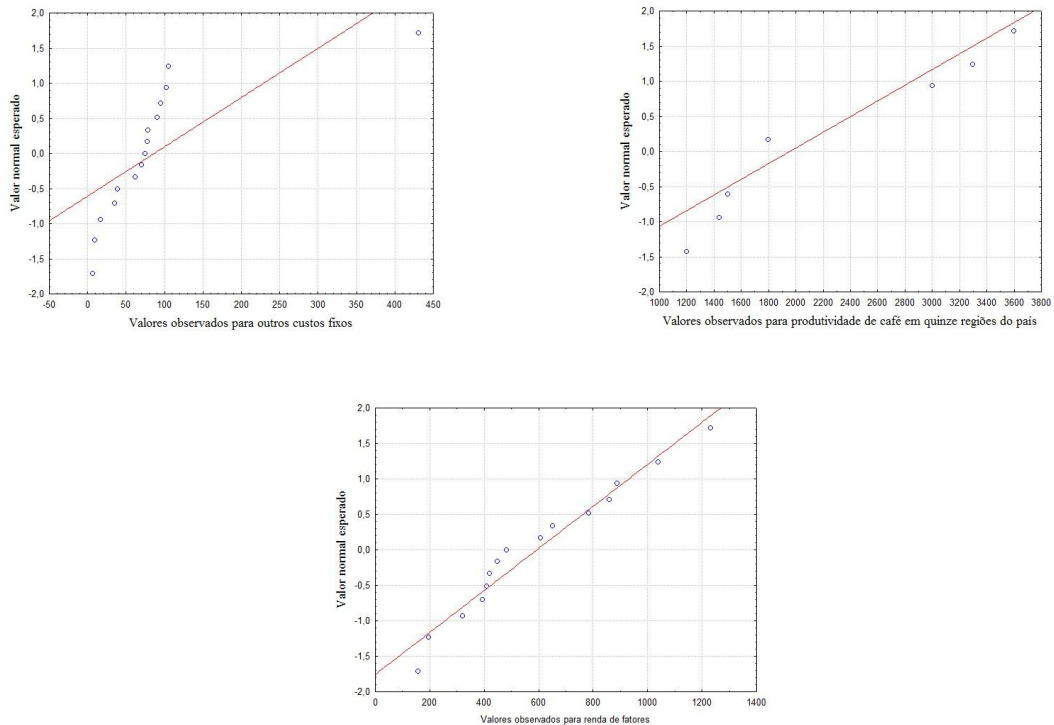


Figura 2 – Gráficos de dispersão entre valores observados para outros custos fixos, produtividade e rendas de fatores em cafés em 15 localidades brasileiras.

Os gráficos nas Figuras 1 e 2 apresentaram vários pontos fora da curva, contudo, as variáveis que apresentaram distribuição normal utilizando a estatística W ou rQ tiveram seus dados mais próximos da linha média na dispersão dos dados.

CONCLUSÕES: As distribuições das variáveis: despesas com custeio de lavoura, despesas com pós-colheita e renda de fatores apresentaram distribuição normal, tanto no teste de Shapiro Wilk (1965) quanto naquele realizado quando se calculou o valor de rQ. As distribuições das variáveis produtividade do grão, despesas financeiras, depreciações e outros custos fixos não apresentaram distribuição normal.

REFERÊNCIAS

- JOHNSON, R.A., WICHERN, D.W. **Applied multivariate statistical analysis**. New Jersey - USA. Prentice Hall. 1998. 816p.
- SANTOS, A. C.; FERREIRA, D. F. Definição do tamanho amostral usando simulação Monte Carlo para o teste de normalidade baseado em assimetria e curtose: I. Abordagem univariada. *Ciência e agrotecnologia*. Lavras – MG. v.27, n.2, p. 432-437, 2003.
- SHAPIRO, S. S.; WILK, M.B. An Analysis of Variance Test for Normality (complete samples), *Biometrika*, London, v.52. n. 3-4, p. 591-611, Dec: 1965.
- THODE JR., H.C. **Testing for Normality**. New York. Marcel Dekker, 2002. 463p.