Animal Feed Science and Technology 159 (2010) 131–137

Contents lists available at ScienceDirect

Animal Feed Science and Technology

journal homepage: www.elsevier.com/locate/anifeedsci

The distribution of crude protein and amino acid content in maize grain and soybean meal

N. Sriperm [a], G.M. Pesti [a,*], P.B. Tillman [b]

[a] Department of Poultry Science, University of Georgia, Athens, GA 30602, United States
[b] Ajinomoto Heartland LLC, Chicago, IL 60631, United States

ARTICLE INFO

Article history:
Received 8 September 2009
Received in revised form 25 May 2010
Accepted 28 May 2010

Keywords:
Protein content
Amino acid content
Maize grain
Soybean meal
Normal distribution

ABSTRACT

This study examines the assumptions of normal distributions for crude protein (CP) and amino acid (AA) contents in feedstuffs. Data for maize grain and soybean meal (SBM) were collected from the Ajinomoto Heartland LLC laboratory analysis database between 2002 and 2008. Tests of normality for CP and selected AA were performed for both feedstuffs by using graphical methods (histogram and normal quantile–quantile plot) and numerical methods (skewness and Shapiro–Wilk procedure (*W*)). Relationships between CP and AA were also computed using linear and quadratic regression and *W* were used to test for normality of the internally Studentized residuals of the regression model. Results indicated that methionine (Met) and arginine (Arg) were not normally distributed in maize grain ($P<0.05$). In addition, CP, lysine (Lys), threonine (Thr), Met, isoleucine (Ile) and tryptophan (Trp) were not normally distributed in SBM ($P<0.05$). There were linear relationships between CP and most of the AA in maize grain and SBM, except for the relationship between CP and Thr, and CP and Ile in maize grain and CP and total sulfur amino acids (TSAA), and CP and Arg in SBM which were found to be non-linear (significant quadratic terms at $P<0.05$). The results indicate the need for normality testing of AA levels in feed ingredients prior to generating prediction equations for AA levels from CP. Assuming a normal distribution of CP and AA in critical feed ingredients may lead to an over or under estimated nutrient content in feed formulation.

Even though the regression residuals are normally distributed in maize grain and SBM, other models beside linear and quadratic regression could be applied in order to accurately predict AA contents based on CP.

1. Introduction

Maize grain and soybean meal (SBM) are the major feed components of poultry feed in the United States, often representing greater than 85% of a diet. Habitually, most nutritionists formulate diets for poultry based on average reference values of maize grain and SBM from the National Research Council (1994) or some other ingredient composition tables; nevertheless, variability in all ingredients has been observed. Wet-chemistry laboratory analysis is the most accurate method to determine the nutritional content of ingredients; but the analysis cost and promptness of results are issues of concern. Therefore, many

Abbreviations: AA, amino acids; CP, crude protein; Lys, lysine; Thr, threonine; Met, methionine; Cys, cystine; TSAA, total sulfur amino acids; Val, valine; Ile, isoleucine; Arg, arginine; Trp, tryptophan; *W*, Shapiro–Wilk test; SBM, soybean meal.

* Corresponding author. Tel.: +1 706 542 1351.
E-mail address: gpesti@uga.edu (G.M. Pesti).

0377-8401/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.anifeedsci.2010.05.009

Table 1
Descriptive statistics and Shapiro–Wilk test of crude protein and amino acids in maize grain samples.

Component	*n*	MIN[a]	MAX[b]	Mean	Median	Diff[c]	SD	CV	Skewness	*W*[d]	*P*[e]	NRC (1994)[f]
Crude protein	336	6.379	9.531	7.905	7.935	0.030	0.621	7.861	−0.196	0.992	0.057	8.50
Lysine	340	0.206	0.301	0.252	0.253	0.001	0.017	6.642	0.030	0.994	0.183	0.26
Threonine	305	0.237	0.325	0.281	0.281	0.000	0.017	6.178	0.176	0.993	0.139	0.29
Methionine	330	0.127	0.200	0.167	0.167	0.000	0.016	9.349	−0.163	0.989	0.012	0.18
TSAA	323	0.291	0.397	0.345	0.345	0.000	0.021	6.002	0.200	0.991	0.050	0.36
Valine	335	0.304	0.439	0.370	0.369	0.001	0.025	6.710	0.231	0.992	0.065	0.40
Isoleucine	334	0.222	0.321	0.267	0.269	0.002	0.018	6.734	−0.032	0.993	0.101	0.29
Arginine	327	0.313	0.473	0.393	0.390	0.003	0.031	7.866	0.293	0.989	0.016	0.38
Tryptophan	290	0.047	0.074	0.060	0.060	0.000	0.005	8.880	−0.087	0.991	0.063	0.06

[a] MIN, minimum value (unit: g per 88 g dry matter).
[b] MAX, minimum value (unit: g per 88 g dry matter).
[c] Diff, Mean–Median value (unit: g per 88 g dry matter).
[d] *W*, the Shapiro–Wilk test for normally distributed data.
[e] *P*, the probability distribution of the Shapiro–Wilk test at 0.05 significance level.
[f] CP and AA contents of the ingredient listed for poultry by National Research Council, 1994 (unit: g per 100 g).

nutritionists are attempting to predict the essential amino acid (AA) contents of their ingredients, as accurately as possible, using least squares regression or near-infrared spectroscopy. When least squares regression is used to estimate AA contents, coefficients in prediction equations from historical laboratory samples relating AA composition to the CP (g per 100 g of nitrogen × 6.25) content of an ingredient, are usually based on linear models (National Research Council, 1994). One of the model assumptions is normality in the distribution of the error terms (or residuals). If this assumption is incorrect, then the estimation by a linear model may be inefficient or misleading. Applying an appropriate transformation to the response variable may improve the model accuracy (Box and Cox, 1964). Cravener and Roush (1999) suggested that linear regression of maize grain and soybean meal has less accurate predictions among their studied models. It provided the lowest r^2 compared with the other models and a low r^2 value indicates a low level of precise prediction.

The starting point for detecting non-normality in datasets is to view the data graphically using either a histogram, stem and leaf plot, normal probability plot or normal quantile–quantile plot (Q–Q plot) (Mendenhall and Sincich, 2003). Secondarily, one could use numerical calculations to evaluate summary statistics such as skewness, kurtosis or use one of several statistical tests for normality (D'Agostino and Stephans, 1986; Park, 2008). A statistical analysis program such as that from the SAS® Institute (SAS, 2004) offers four statistical tests for determining if a dataset or a models' error terms are normally distributed; Shapiro–Wilk, Kolmogorov–Smirnov, Cramer–von Mises, and Anderson–Darling. In this study, we primarily considered the Shapiro–Wilk test (*W*) because Shapiro and Wilk (1965), Royston (1982, 1983) and Park (2008) agree that this is the most reliable test for non-normality. Moreover, Royston (1982) extended the test for the sample size range between 7 and 2000 from the original *W* statistic (Shapiro and Wilk, 1965) which had a small sample size range between 3 and 50. Therefore, the *W* statistic is appropriate for this dataset as the sample size of this study was less than 2000. The *W* statistic, which is positive and always less than or equal to one, was determined by the ratio between the best estimator of the variance and the usual corrected sum of squares estimator of the variance (SAS, 2004). The null hypothesis for this test is that the probability distribution of the response, *y*, is normal. Hence a computed value of *W* less than W_α suggests the hypothesis of normality at a specified α level, should be rejected (Shapiro, 1986).

The purpose of this work were (1) to determine if using the mean value of CP and AA (Lys, Thr, Met, TSAA, Val, Ile, Arg, and Trp) in maize grain and SBM is adequate for feed formulation and (2) to test the assumption of normality of the relationships between CP and AA in those two major feed ingredients. This paper will also provide a guideline for testing whether an individual AA or CP distribution in a selected feed ingredient is normally distributed.

2. Materials and methods

Amino acid and CP contents, for maize grain and SBM samples submitted between 2002 and 2008, were determined at Ajinomoto Heartland LLC's AA laboratory. The ingredient samples tested, while having an unidentified regional origin, were collected from across the United States. The AA and CP contents were adjusted to a g per 88 g of dry matter based upon dry matter determination in a forced air oven at 135 °C for 2 h (AOAC, 2000; 930.15). The stem and leaf procedure was used to analyze data for outliers by identifying extreme values. The digits of the individual values were ordered in a histogram and the extreme values, listed individually outside the scale of the histogram, were removed from the data (Statistix9, 2008). Once the determined outliers were removed, the number of analyzed samples of each component (CP and each AA) was between 290 to 340 for maize grain and 489 to 571 for SBM (Tables 1 and 2, respectively). The maize grain and SBM samples were analyzed for CP (Assoc. of Official Analytical Chemists (Method 990.03, nitrogen divides by 0.16, AOAC, 2000)) and AA (AOAC, 2000; 982.30 E [a, b]) composition. The AA used HPLC (HITACHI 8900, 2010) with ninhydrin in a post-column derivatization. Regarding AA recovery, internal reference samples are run with each batch of hydrolysates in order to monitor the degree of hydrolysis. In addition, nor-leucine is used as an internal standard. A known sample is used to monitor hydrolysis efficiency and HPLC conditions. Sample results are as-is, and represent the average of two analyses.