International Journal of Agricultural Science Research Vol. 6(1), pp. 001-007, March 2017
Available online at http://academeresearchjournals.org/journal/ijasr
ISSN 2327-3321 ©2017 Academe Research Journals

Full Length Research Paper

ANOM Technique for Evaluating Practical Significance of Observed Difference Among Treatment Groups

Soner Yigit and Mehmet Mendes*

Çanakkale Onsekiz Mart University, Faculty of Agriculture, Department of Animal Science, Biometry and Genetics Unit, 17100/Çanakkale, Turkey

Accepted 23 February, 2017

It is quite important to take into account the practical significance along with the statistical significance since the presence of statistically significant differences between the group means don't show that these differences are also practically significant. Similarly, the statistically non-significant differences do not mean that they are not practically significant too. That is why; evaluating statistical and practical significances simultaneously, particularly in applied sciences based studies would be a great advantage for the researchers. In this study, it is focused on how benefit could be derived from Analysis of Mean (ANOM) Technique while evaluating both the statistical as well as the practical significances of the results related to animal science based studies along with their effect in size and measurements. Three data sets which were obtained from an experiment to investigate the effect of different feed restriction regimes were used as material. The obtained results of ANOM technique showed that any statistical significant difference was not found among the feed restriction groups in terms of dry matter, acidity, and fat percentage. However, both ANOM and Epsilon effect size measure showed that although differences in the groups in terms of fat percentage in chest meat were not found as statistically significant (P=0.058), these differences may be practically significant. As a result, observed differences among the groups in terms of fat percentage may be considered practically significant.

Keywords: ANOM, practical significance, statistical significance, effect size measure

Introduction

Results of the statistical analyses should be reported informative and understandable as much as possible. In practice, the researchers report commonly the p-value that shows only statistically significance of the observed difference among the group means. However, finding a statistically significant difference among the group means does not mean that this difference is also practically significant. Notwithstanding, most of the researchers believe that finding a smaller p-value shows that the observed difference among the group means is very significance (Nickerson, 2000). However, statistical significance is a function of sample size. Thus, tiny differences may be found as statistically significant when studying with large samples. Clear differences, on the other hand, may not be found statistically significant in case of working with small sample sizes (Fan, 2000; Mendeş, 2012; Yiğit and Mendeş, 2016). That is why, only the reporting of p-value is not enough for both evaluating statistical and practical significances of observed difference among the treatment groups. Therefore, especially in studies related to applied science, it will be very beneficial to evaluate both the statistical and the practical significance of the observed difference simultaneously. That way, it will be possible to get more detail information about the effect of the factor(s) in the study. That is why, most of the reputable journals are looking for such authors for reporting some effect size measures that would provide information regarding to practical significance along with P-value. Especially in recent years, it is remarkable that many researchers started reporting some effect size measures namely; Eta-Squared, Partial Eta-Squared, Omega-

*Corresponding author's email: mmendes@comu.edu.tr

Squared, and Epsilon-Squared (Cohen, 1973; Keselman, 1975; Carroll and Nordholm; 1975; Fritz et al., 2012; Skidmore and Thompson, 2012; Okada, 2013; Yiğit and Mendeş, 2016). Since the commonly used statistical package programs such as Minitab, SPSS, NCSS, Statistical etc. directly report Eta-squared or Partial eta-squared when reporting the results of the analysis of variance. However, the majority of the authors have some challenges in terms of understanding and interpreting the above mentioned measurements. That is why, it is suggested to make familiar about some effect size measurements and statistical methods that would be able to provide more accurate information regarding practical significance of the observed differences. Analysis of Means (ANOM) can easily be used for this purpose along with some effect size measures such as Epsilon-Squared and Omega-Squared. ANOM is a graphical method that helps for understanding and interpreting the results of any research work very easy especially, for non-statisticians. ANOM technique enables the researchers for assessment of the practical significance as well as the statistical significance of the differences which were observed among the compared group means, simultaneously (Ott, 1967; Nelson, 1983; Nelson et al., 2005; Balamurali, and Kalyanasundaram, 2011; Mendeş and Yiğit, 2013; Ashit et al., 2015; Chakraborty and Khurshid, 2015). In this study, due to this kind of advantageous of ANOM, it was used to evaluate practical significance of statistical results of animal science-based studies. For this aim, ANOM has been applied to different data sets related to experiments to investigate the effect of three different feed restriction programson acidity, dry matter and fat content of thigh and chest meats in broilers chickens. In this way, it has been informed about both of the practical significance as well as the statistical significance related to different feed restriction programs.

Material and Methods

In order to show how to evaluate practical significance of the observed difference among the groups along with statistical significance (p-value) by using ANOM technique, three data sets which were obtained from an experiment to investigate the effect of different feed restriction regimes (ad libitum (AD) group, the group with the 20% feed restriction based on ad libitum groups (RF), and the group that was not fed between 9 am and 3 pm (NF)) on thigh and chest fat percentage, acidity (lactic acid percent) and dry matter of 30 Ross 308 line male chickens were used as a material for this study (Mendeş et al., 2007). ANOM Technique was used to evaluate practical significance as well as the statistical significance related to different feed restriction programs. Practical significance of the observed difference among the groups was also evaluated by Epsilon-Squared effect size measure.

Analysis of Means (ANOM) Technique

The Analysis of Means (ANOM), introduced by Ott (1967), is a useful graphical alternative to the Analysis of Variance (ANOVA) for comparing independent group means. ANOM graph displays each group mean, the overall mean, and the upper and lower decision lines. Nelson et al (2005) reported that when studying main effects, ANOM has two advantages over ANOVA: a) if any of the group or treatment is statistically different, ANOM indicates exactly which ones is different easily and b) ANOM can be presented in a graphical form, which allows one to easily evaluate both the statistical and practical significance of the differences.

The null hypothesis for ANOM and ANOVA are the same but, the alternative hypothesis of ANOM is basically different. Because the alternative hypothesis for ANOM states that the group mean is different from the overall mean; while the alternative hypothesis for ANOVA states that at least one means is different from the others. Therefore, although similar results are obtained at the end of both methods in general, there are some cases where the two methods might lead to different results. For example, if mean of one group is above the overall mean and another group of means are below the overall mean, the ANOVA might reject the null hypothesis while the ANOM might accept.

ANOM is performed based on confidence interval. The steps of constructing decision lines are given below:

1. Calculate the mean and the variance of each group
2. Calculate the grand mean
3. Calculate the Mean Square Error
4. Determine the ANOM critical value *h*, based on the level of significance, the number of groups k, and the total sample size *N*
5. Upper (UDL) and Lower (LDL) decision lines are computed.

Upper Decision (UDL) and Lower Decision Lines (LDL) are computed as below (Nelson et al., 2005; Mendeş and Yiğit, 2013):

If sample sizes are equal:

$$UDL = \overline{Y}_{..} + h\,(\alpha, k, N-k)\sqrt{MSE}\,\sqrt{\frac{(k-1)}{N}}$$

$$LDL = \overline{Y}_{..} - h\,(\alpha, k, N-k)\sqrt{MSE}\,\sqrt{\frac{(k-1)}{N}}$$

If sample sizes are not equal:

$$UDL = \overline{Y}_{..} + h\,(\alpha, k, N-k)\sqrt{MSE}\,\sqrt{\frac{(N-n_i)}{Nn_i}}$$

Table 1. Results of ANOVA (P-values) and Epsilon-Squared values

	Characteristics	Statistical Significance (P-^{value})	Practical Significance (Epsilon-Squared)
Thigh	Acidity	0.214	18.59
	Dry Matter	0.879	1.74
	Fat (%)	0.578	14.45
Chest	Acidity	0.457	9.91
	Dry Matter	0.581	6.98
	Fat (%)	**0.058**	31.59

$$LDL = \overline{Y}_{..} - h(\alpha,\ k,\ N-k)\sqrt{MSE}\sqrt{\frac{(N-n_i)}{Nn_i}}$$

Where k is the number of treatment groups, N is the total number of observation, n_i is the sample size for the i th group, MSE is Mean Square Error and h(n , k ,N − k) is the critical values based on significance level (α), number of means being compared (k) and degrees of freedom for means square error (N − k) .

Decision Rule: If all means fall between the decision lines (UDL and LDL), then accept the hypothesis of k equal means while any of the group mean falls outside the decision lines, then the hypothesis of equality of k means is rejected.

Effect Size Measure

In order to estimate population effect size, many effect size measures namely Eta-squared, Partial eta-squared, Omega-squared, and Epsilon-squared have been developed. In this study, Epsilon-Squared which is one of the most popular effect size measures, was taken into considered.

$$\hat{\varepsilon}^2 = \frac{SS_{Effect} - df_{Effect} MS_{Error}}{SS_{Total}}$$

Where SS_{Total}: Total sum of squares, SS_{Effect}: Sum of squares of effect, SS_{Error}: Error sum of squares, MS_{Error}: Mean square error and df_{Effect}: Degree of freedom of effect (Kelly, 1935; Hays, 1963; Glass and Hakstian, 1969; Keselman, 1975; Skidmore and Thompson, 2012; Okada, 2013; Yiğit and Mendeş, 2016).

Results

The results regarding to P-values and Epsilon-Squared are presented in Table 1. ANOM graphs that were used to evaluate the practical significance of the observed differences among the group means along with statistical significance have been given in Figure 1-6. When P-values which provide only statistical significance of the observed differences among the groups are considered, it is seen that there are not any statistically significant differences among the groups in terms of acidity (P=0.457), dry matter (P=0.587) and fat percentage (P=0.058) in the chest meat of the broiler chickens (Table 1). Differences among the groups in terms of these characteristics are also not statistically significant in thigh meat of chickens as well (P=0.214; P=0.879; P=0.587). Therefore, when the authors report only P-values, the readers or other authors will understand that the feed restriction programs have not significant impact on acidity, dry matter and fat percentage of both chest and thigh meat of chickens. However, when ANOM graphs and Epsilon-Squared results showed that the observed differences among the groups, especially in terms of fat percentage of the chest meat, may be practically significant for the researchers. Thus, it will be useful to evaluate these differences as a practically along with statistically.

Practical significance of the observed differences among the groups is evaluated by using two different approaches namely Epsilon-Squared values (Table 1) and ANOM graphs (figure 1-6). When practical significance is evaluated by using Epsilon-Squared (due to Epsilon-Squared is an unbiased estimator of population, our evaluations are based upon Epsilon – Squared (Carroll and Nordholm,1975; Okada, 2013; Yiğit and Mendeş, 2016), it is possible to conclude that especially the differences among the groups in terms of fat percentage in chest meat may be practically significant. Epsilon-Squared value for fat percentage is found as 31.59% that means 31.59% of the variation in the fat percentage can be explained by the feed restriction regimes. The Epsilon-Squared values for acidity and dry matter are found as 9.91% and 6.98% respectively. Therefore, the 9.91% of the variation in acidity and 6.98% of variation in dry matter can be explained by the groups. In this case, it is possible to conclude that the differences among the groups in terms of fat percentage may be practically significant although it is not found as statistically significant while the differences among the groups in terms of acidity and dry matter most probably will not practically significant.

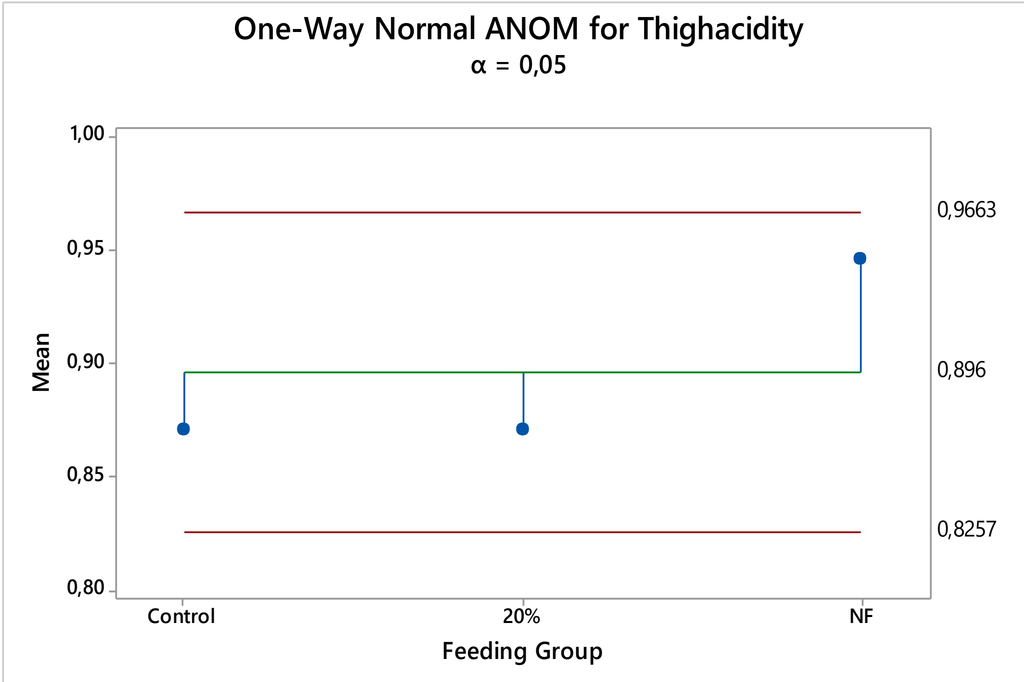

Figure 1. ANOM graph for comparing group means in terms of acidity in thigh meat

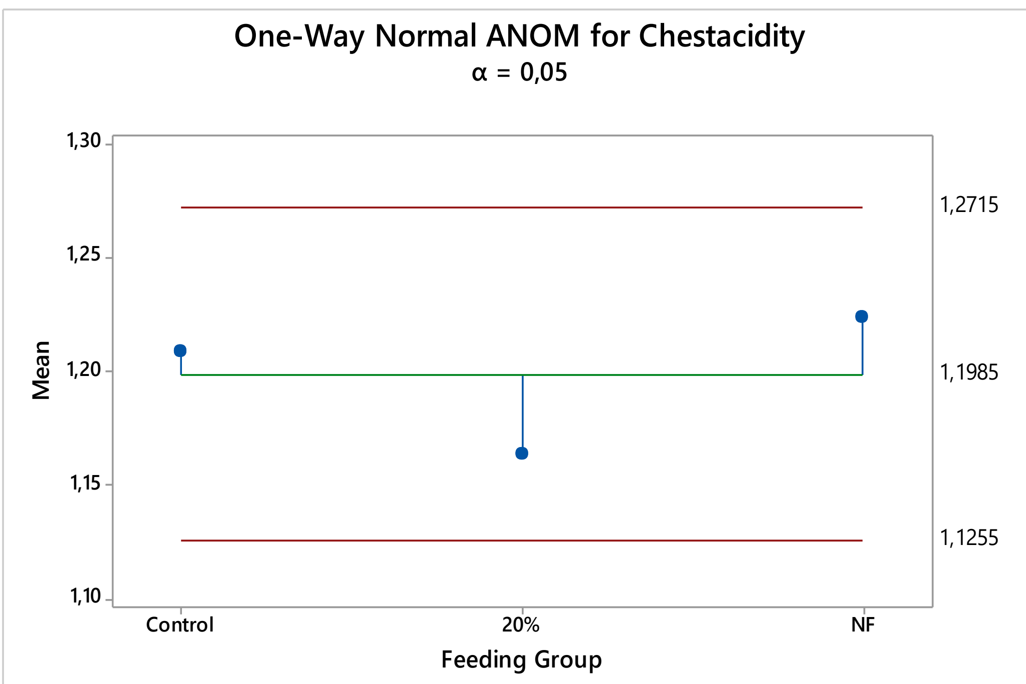

Figure 2. ANOM graph for comparing group means in terms of acidity in chest meat

Discussion

Analysis of Variance (ANOVA) is commonly used for comparing the differences among independent group means (Zar, 2010; Mendeş and Yiğit, 2013). However, at the end of ANOVA, it is only obtained information about statistical significance of the observed differences. ANOVA test does not give any information about its practical significance (Glass and Hakstian, 1969). In other words, it does not show how different the compared

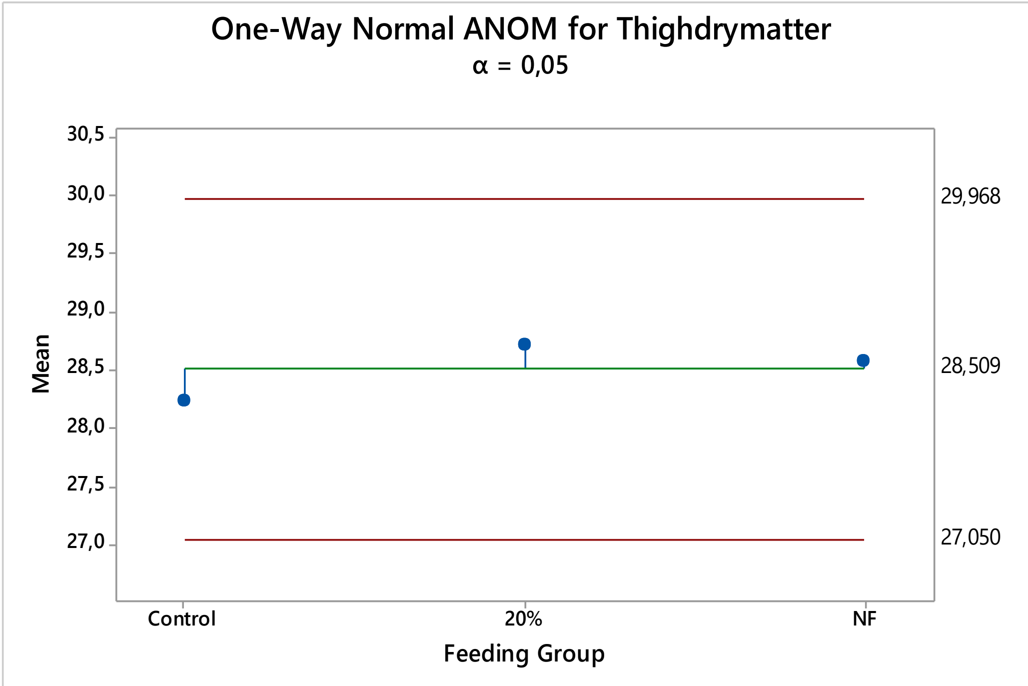

Figure 3. ANOM graph for comparing group means in terms of dry matter in thigh meat

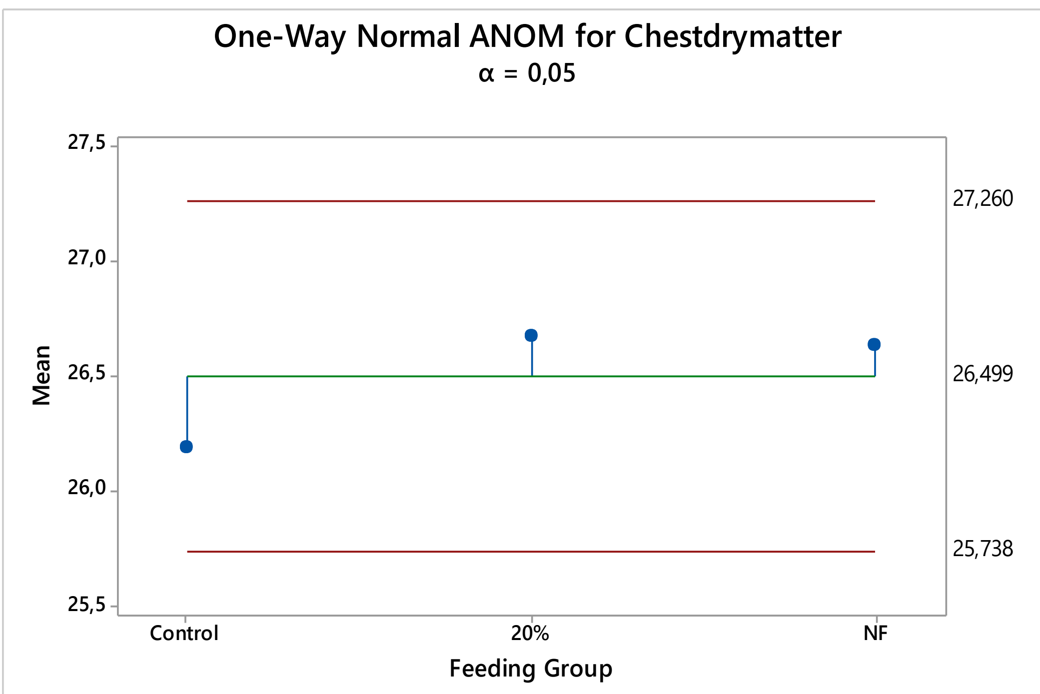

Figure 4. ANOM graph for comparing group means in terms of dry matter in chest meat

group means are from each other or how much of the difference occurred in the dependent variable results from the groups. Therefore, evaluating practical significance of the observed differences among the treatments along with statistical significance is an important issue, particularly for applied sciences. It would be a great advantage while assessing the statistical significance along with the practical significance at the

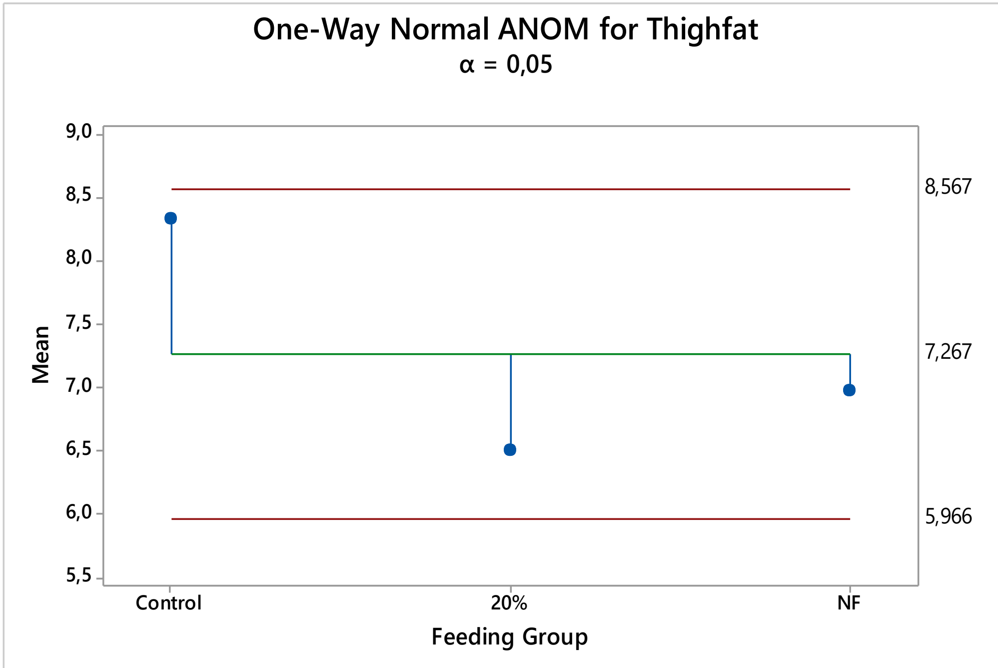

Figure 5. ANOM graph for comparing group means in terms of fat % matter in thigh meat

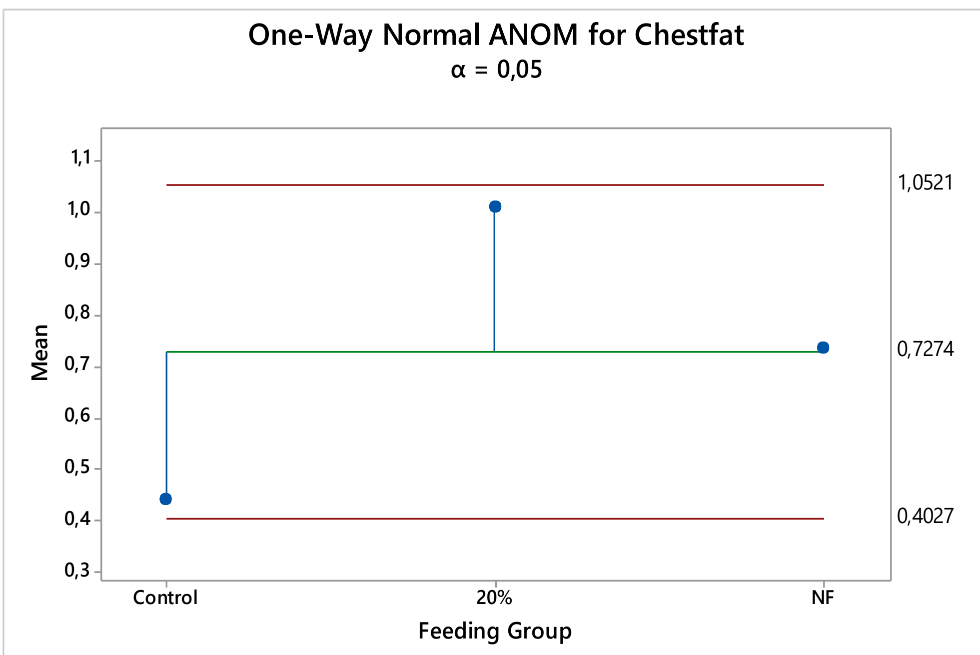

Figure 6. ANOM graph for comparing group means in terms of fat % matter in chest meat

same time. In this study, two different approaches namely ANOM and an effect size measure, Epsilon-squared, have been considered in assessing practical significance of the observed difference (Cumming and Finch, 2001; Nelson et al., 2005; Skidmore and Thompson, 2012; Okada, 2013). That way, it will be possible to get more detail information about the effect of feed restrictions on chemical properties of broiler chickens.

Results of this study showed that although similar results are obtained when practical significance of the differences among the groups is evaluated by using ANOM technique (figure 1-6), evaluation of practical significance of observed differences among the groups by using ANOM technique is easier than evaluation that differences by using effect size measures. It is because ANOM technique is a graphical method and enables the researchers to figure out the differences among the groups visually. And it also enables us to evaluate the observed differences both statistically and practically simultaneously. That is why; using the ANOM to assess the practical significance is easier than that of effect size measures especially for non-statisticians. When ANOM graphs are examined it is clearly seen that the fat percentage mean of the control group is obviously smaller than the overall mean while the mean of NF group is higher than that of the overall mean.

When ANOM graphs are examined it is clearly seen that none of the means fall outside the decision lines. That means, there is not any statistically significant difference among the groups in terms of fat percentage, acidity and dry matter. However, it is clearly seen that the fat percentage mean of the control group is obviously smaller than the overall mean while the mean of NF group is higher than that of the overall mean. When decision lines for control and NF groups are examined, it can be seen that although both lines are not exceeded the decision lines they are very close to the upper and lower lines (Figure 6). This situation may be accepted as an indicator of a practical significance among the groups rather than a statistical significance. The finding of high Epsilon-Squared value (31.59%) also supports this finding as well. And this finding may be important for the researchers.

References

Balamurali, S, and Kalyanasundaram, M (2011). An Investigation of the Effects of Misclassification Errors on the Analysis of Means. Tamsui Oxford J. Inform. Math. Sci., 27(2): 117-136.

Carroll, RM, Nordholm, LA (1975). Sampling Characteristics of Kelley's and Hays' . Educ. Psychol. Meas., 35:541–554.

Chakraborty, AB, Khurshid, A (2015). Measurement Error Effect on the Power of ANOM Control Chart for Doubly Truncated Normal Distribution under Standardization Procedure. Ann. Manage. Sci., 4(2):87-98.

Cohen, J (1973). Eta-Squared and Partial Eta-Squared in Fixed Factor Anova Designs. Educ. Psychol. Meas., 33:107–112.

Fan, X (2001). Statistical Significance and Effect Size in Education Research: Two Sides of a Coin. J. Educ. Res., 94:275–282.

Fritz, CO, Morris, PE, Richler, JJ (2012). Effect size estimates: current use, calculations, and interpretation. J. Exp. Psychol. Gen., 41:2–18.

Glass, GV, Hakstian, AR (1969). Measures of Association in Comparative Experiments: Their Development and Interpretation. Am. Educ. Res. J. 6:403–414.

Keselman, HJ (1975). A Monte Carlo investigation of three estimates of treatment magnitude: Epsilon squared, eta squared, and omega squared. Can. Psychol. Rev., 16:44–48.

Mendeş, M, Dinçer, E, Arslan, E (2007). Profile analysis and growth curve for body mass index of broiler chickens reared under different feed restrictions in early age. Arch. Tierz., Dummerstorf, 50:403-411.

Mendeş, M (2012). Uygulamalı Bilimler İçin İstatistik ve Araştırma Yöntemleri Çanakkale: Kriter

Mendeş, M, Yiğit, S (2013). Comparison of ANOVA-F and ANOM tests with regard to type I error rate and test power. J. Stat. Comput. Sim. 83:2093–2104.

Nelson, PP (1983). The analysis of means for balanced experimental designs. Journal of Quality Technology, 15: 45-54.

Nelson, PP, Wludyka PS, Copeland, KAF (2005). The analysis of means: A graphical method for comparing means, rates and proportions. SIAM, Philadelphia, USA.

Nickerson, RS (2000). Null hypothesis significance testing: a review of an old and continuing controversy. Psychol. Methods 5:241–301.

Okada, K (2013). Is omega squared less biased? A comparison of three major effect size indices in one-way ANOVA. Behaviormetrika 40:129–147.

Ott, ER (1967). Analysis of means- A graphical procedure. Industrial Quality Control, 24: 101-109.

Skidmore, ST, Thompson, B (2012). Bias and precision of some classical ANOVA effect sizes when assumptions are violated. Beh. Res. Methods 45:2 536–546.

Yiğit, S, Mendeş, M (2016). Which Effect Size Measure is Appropriate for One-Way and

Two-Way Anova Models? A Monte Carlo Simulation. Revs., (in pressing)

Zar, JH (2010). Biostatistical Analysis, Pearson, New Jersey.