Abstract

Proportional data is a common occurrence in postharvest data where various treatments or treatments with varying concentrations are applied to fruits to increase the shelf life. The experimental unit is for example set of twenty (20) fruits on which the treatment may be applied but the observational unit on which measurement is taken is the individual fruits. The response variable for each individual fruit is the either yes or no response depending if it is still marketable or not marketable. The data generated may contain the observed proportion of fruits that are marketable for various treatments applied. Such data generally exhibit what is called a *binomial distribution* rather than a normal distribution. One of the characteristics of such a distribution is that the variances are related to the means and tend to be small at the two ends of the range of values (close to zero and 100%) but larger in the middle. Such data is generally transformed using *arcsine* transformation (taking the inverse sine of the square root of each observation, expressed as a proportion of a total) and then subjecting the transformed data to standard analysis of variance procedure (GLM -ANOYA). The means are later back transformed to original scale. Series of steps involved in the transformation and then back-transformation are quite rigorous and confusing to many that eventually end up using standard-ANOYA. Using the standard ANOYA on the original data may give significant results with higher R^2 and F value higher than the arcsine transformed data but the analysis may not be valid as one of the assumptions of ANOYA has not been satisfied. This article therefore presents a simpler and direct method of analyzing proportion data without subjecting to rigorous transformation. The procedure that could be used is Generalized Linear Model (GLMs) Procedure or Generalized Liner Mixed Model (GLMM) procedure depending on block being fixed or random. Arbitrary data was used in this article to demonstrate how the two models could be used to analyze proportion data using GENSTAT statistical software.