Agric. Res. J. Kerala 2X (1990) : .16 - 39

DUMMY VARIABLE APPROACH FOR ANALYSING YIELD RESPONSE OF PLANT GENOTYPES TO NITROGEN LEVELS: CASE STUDY OF TOMATO

K.V. Subrahmanyam, S.K. Tikoo, B.S. Prabhakar, Vishnu Shukla and V.R.Srinivasan
Indian Institute of Horticultural Research, Bangalore 560 080, India

Abstract: For analysing the data from field experiments involving qualitative and quantitative variables, a new approach of using dummy variables for qualitative variables in the regression equation was tried in with the help of two years data from the experiments conducted on tomato crop involving four genotypes and four nitrogen levels. The results of analysis showed that using dummy variables for qualitative variable viz., genotypes in the regression analysis has the advantage of single equation taking care of differences between genotypes as well as their individual contribution to yield along with other quantifiable inputs like nitrogen. Besides having comparable predicting efficiency of yield with that of individual functions for each genotype based on ANOVA, dummy variable approach saves a lot of time and labour especially when a number of genotypes are involved in the experiment, as otherwise a separate equation has to be fitted for each genotype to catch their effect on yield.

INTRODUCTION

The plant genotype is an important qualitative variable which will effect the yield and also the use of other inputs. The data generated from field experiments involving genotypes and other inputs are usually analysed using analysis of variance technique developed for different designs of experiments. In this technique a separate regression is fitted for each genotype when the interaction between genotype and the other variable is significant or pooled one with the other quantifiable variable when the interaction is not significant. Hence, this technique fails to measure simultaneously the effect of all genotypes and the other inputs on the yield of crops, because genotype cannot be quantified. To overcome this, a single regression equation which can incorporate both qualitative and quantitative variable viz., dummy variable approach where qualitative variables can be incorporated using dummy variables was tried to study the effect of genotypes and nitrogen levels on yield of tomato crop.

MATERIALS AND METHODS

The data for this study were obtained from a field experiment conducted during winter season of 1986 and 1987 at the experimental farm of the Indian Institute of Horticultural Research, Hessarghatta. The treatments consisted of four nitrogen levels (0,60,120 and 180 kg N/ha) and four genotypes (IHR 707, IHR 726, Arka Vikas and Pusa Ruby) laid out in factorial RBD design. All plots received an uniform dose of 100 kg P_2O_5/ha and 50 kg K_2O/ha at the rime of planting.

The qualitative variable such as genotypes will often play an important role in determining the yield response to inputs such as nitrogen and must be incorporated. In the usual analysis, as genotype is not quantifiable, we will be fitting four nitrogen response functions representing the four genotypes. Therefore the yield response of different genotypes to nitrogen can be represented as

$$Y = \alpha + \beta N + \gamma N^2 + u$$

where Y = yield of the crops; α, β, γ are regression coefficients; N = nitrogen and u = error term.

As there were four genotypes in the experiment, four equations are possible. To avoid these dummy vatiables can be introduced to represent different genotypes in a single equation as shown below

$$Y = \alpha_1 + \alpha_2 G_2 + \alpha_3 G_3 + \alpha_4 G_4 + \gamma_1 N + \gamma_2 N^2 + \theta_1 (G_2 N) + \theta_2 (G_3 N) + \theta_3 (G_4 N) + u$$

Three dumy variables are introduced; (G) i = 2,3,4 and three interaction terms to represent four genotypes and their interaction with nitrogen in the model to avoid the dummy variable trap while fitting the equation with the computer (Johnston, 1972). But as the preliminary analysis of the data has shown that there is no significant interaction between the genotypes and the nitrogen levels, the interaction variables were dropped. As there are two years data which represent two different climatic situations one dummy variable (T) is introduced for the years also and the final fitted equation was of the following nature

$$Y = \alpha_1 + \alpha_2 G_2 + \alpha_3 G_3 + \alpha_4 G_4 + \beta T + \gamma_1 N + \gamma_2 N_2 + u$$

where G = 1 if the observation relates to genotypes, i = 2, 3, 4
i = 0 otherwise
T = 1 if the observation belongs to year 1986-1987, 0 if the observation belongs to year 1985-1986 and N = nitrogen applied (kg/ha).

The interpretation of the coefficients of dummy variables is slightly different from the usual regression coefficient of the variables. The α_1 represents the contribution of genotypes 1 in period I. and α_2, α_3 and α_4 represent differential genotype effects on yields as compared to genotype 1. The absolute contribution of different genotypes to yield is derived from the above equation as follows.

For example the contribution of genotype 2 is given in period I by $\alpha_1 + \alpha_2 + \beta$ and in period II by $\alpha_1 + \alpha_2$ as T will have zero value in period II. The average in period I and II gives the contribution for pooled data. Similarly for genotype 3, it will be $\alpha_1 + \alpha_3 + \beta$ in period I and $\alpha_1 + \alpha_3$ in period II and for genotype 4, $\alpha_1 + \alpha_4 + ft$ and $\alpha_1 + \alpha_4$ in period I and II respectively.

The estimation of yield at different N levels from the above equation for different genotypes will be as follows:

For genotype 1 =

$$Y_1 = \alpha_1 + \beta T + \gamma_1 N + \gamma_2 N^2$$

For genotype 2 =

$$Y_2 = \alpha_1 + \alpha_2 + \beta T + \gamma_1 N + \gamma_2 N^2$$

For genotype 3 =

$$Y_3 = \alpha_1 + \alpha_3 + \beta T + \gamma_1 N + \gamma_2 N^2$$

For genotype 4 =

$$Y_4 = \alpha_1 + \alpha_4 + \beta T + \gamma_1 N + \gamma_2 N^2$$

Where T takes Q or 1 values for period I and II respectively and N = nitrogen (kg/ha).

RESULTS AND DISCUSSION

The yield response function fitted with dummy variables to genotypes and nitrogen for the above data was

$$Y = 367.8429 - 52.50\ G2 + 54.25\ G3 - 63.125\ G4 - 323.5625\ T + 6.0221\ N - 0.0261\ N$$

$R^2 = 0.69$, F = 33.02

where Y = yield (g/ha), G and T are

Table 1. Contribution of genotypes to yield of tomato based on dummy variable equation at 0 level of nitrogen

Genotype	Quantity* (q/ha)
IHR 707	206
IHR 726	154
Arka Vikas	260
Pusa Ruby	143

* Average of period I and period II

dummy variables representing the three genotypes and year respectively and N = nitrogen (kg/ha).

The fitted equation could explain 69 per cent of the variation in yield, which shows a lot of improvement over the variation explained by nitrogen variable alone fitted for pooled data for all genotypes based on ANOVA results which was explaining hardly 25% of the variation ($Y = 190.7179 + 6.0221 N - 0.0261 N^2$).

The modified significant tests conducted on the regression coefficients of dummy variables are significant at 5% level and the time dummy variable was not significant.

Contribution of genotypes: The contribution of genotypes to yield was estimated from the above equation by using the porocedure explained earlier and the results are presented in Table 1. From the results it could be observed that Arka Vikas was found to be superior compared to all other genotypes followed by IHR 707 and IHR 726 and Pusa Ruby. This shows that all other genotypes are superior to Pusa Ruby which is now a commercially cultivated variety.

Comparison of yield estimates using dummy variable equation and individual genotype equation: Four yield response function equations with nitrogen for each genotype were fitted (Table 2) for comparing the yield estimated from the dummy variable equaitions at different nitrognen levels for different genotypes with individual genotype equah'ons.

The comparison of the yields predicted based on the above two functions, viz. dummy variable function and individual genotype functions, at the three levels of nitrogen viz. 60, 120 and 180 kg N/ha is presented in Table 3. It could be observed that there was not much difference between the yield estimated by the two types of functions.

From the above discussion it is clear that the dummy variable approach offers some advantages over the ANOVA approach. The dummy variable provides a method for quantifying qualitative variables, like genotype, in field experiments and also the seasonal effects when the experiments are conducted for two or more seasons. Besides, a single equation will take care of the differences between genotypes as well as their individual contribution to yield along with other quantifiable variables. The dummy variable approach saves a lot of time and labour especially when a number of genotypes are involved in the experiment as otherwise a separate function has to be fitted for each genotype to catch the genotype effect. In predicting efficiency also the dummy variable function gave comparable results with that of individual function approach based on ANOVA.

Table 2. Nitrogen response equation for different genotypes of tomato

Genotype	Response equation
IHR 707	$Y = 173.3695 + 7.2096 N - 0.0320 N^2$
IHR 726	$Y = 169.9418 + 5.1376 N - 0.0211 N^2$
Arka Vikas	$Y = 251.4072 + 5.7954 N - 0.0238 N^2$
Pusa Ruby	$Y = 168.0419 + 5.9482 N - 0.0276 N^2$

y = yield (q/ha) N = Nitrogen (kg/ha)

Table 3. Comparison of estimated yields (q/ha) based on dummy variable function and individual genotype function at different levels of nitrogen*

Genotype	60		120		180	
	DVF	IGF	DVF	IGF	DVF	IGF
IHR 707	473	491	553	578	444	434
IHR 726	421	402	500	483	392	411
Arka Vikas	528	513	607	604	499	523
Pusa Ruby	410	426	490	484	381	344

DVF = dummy variable function, IGF = individual genotype function
* Average of period I and II

ACKNOWLEDGEMENT

The authors wish to thank Dr. R.M. Pandey, Director, Indian Institute of Horticultural Research, Bangalore for providing the necessary facilities for carrying out this works.

REFERENCE

Johnston, J. 1972. *Econometric Mehthods*. Mc Graw-H Kogakusha Ltd., Tokyo, p. 176-186

D □