DOI: 10.17311/sciintl.2014.64.71

Consumer Behavioural Buying Patterns on the Demand for Detergents Using Hierarchically Multiple Polynomial Regression Model

[1]H.J. Zainodin, [1]Noraini Abdullah and [2]G. Khuneswari
[1]School of Science and Technology, Universiti Malaysia Sabah, Kota Kinabalu, Sabah, 88400, Malaysia
[2]University of Glasgow, School of Mathematics and Statistics, G12 8QQ Glasgow, United Kingdom

ABSTRACT

Background: Market studies on consumer preferences on product items had shown that consumer's behavioural patterns and intentions are sources of business profit level. In the advent wave of global businesses, the behavioural buying patterns of consumers have to be studied and analysed. Hence, this research illustrated the procedures in getting the best polynomial regression model of the consumer buying patterns on the demand for detergent that had included interaction variables. **Methods:** The hierarchically multiple polynomial regression models involved were up to the third-order polynomial and all the possible models were also considered. The possible models were reduced to several selected models using progressive removal of multicollinearity variables and elimination of insignificant variables. To enhance the understanding of the whole concept in this study, multiple polynomial regressions with eight selection criteria (8SC) had been explored and presented in the process of getting the best model from a set of selected models. **Results:** A numerical illustration on the demand of detergent had been included to get a clear picture of the process in getting the best polynomial order model. There were two single independent variables: the "price difference" between the price offered by the enterprise and the average industry price of competitors' similar detergents (in US$) and advertising expenditure (in US$). **Conclusion:** In conclusion, the best cubic model was obtained where the parameters involved in the model were estimated using ordinary least square method.

Key words: Hierarchically multiple polynomial regressions, ordinary least square, coefficient test, eight selection criteria, best cubic model

Science International 2 (3): 64-71, 2014

INTRODUCTION

Regression analysis allows the researcher to estimate the relative importance of independent variables in influencing a dependent variable. It also identifies a mathematical equation that describes the relationship between the independent and dependent variables. According to Sandy[1], simple linear regression is a technique that is used to describe the effects of changes in an independent variable on a dependent variable. However, the Multiple Regression is a technique that predicts the effect on the average level of the dependent variable of a unit change in any independent variable while the other independent variables are held constant. According to Gujarati[2], a model with two or more independent variables is known as a multiple regression model. The term "multiple" indicates that there are multiple independent variables involved in the analysis and multiple influences will affect the dependent variable. According to Pasha [3] using too few independent variables may give a biased prediction, while using too many independent variables, can cause a varied fluctuation prediction value.

Most cost and production functions are curves with changing slopes and not a straight line with constant slopes[4]. The polynomial model enables the estimation of these curves[5,6]. In a polynomial model, independent variables can be squared, cubed or raised to any power or exponential power. The studies of[7] is an example of a squared polynomial or quadratic model, while[8] fitted a third order polynomial model of the best R^2 correlation. The coefficients still appear in the regression in a linear fashion, allowing the regression to be estimated as usual. As in[2], for example, the polynomial model is given as:

$$Y = \beta_0 + \beta_1 X_1 + \beta_{11} X_1^2 + u$$

Since the same independent variables appear more than once in different forms, it will appear to be highly correlated. The effect of multicollinearity may not exist because the independent variables are not linearly correlated.

Corresponding Author: H.J. Zainodin, School of Science and Technology, Universiti Malaysia Sabah, Kota Kinabalu, Sabah, 88400, Malaysia

The main objective of this study is to obtain the best model with a polynomial order which would represent the whole structure of the collected data so that further analysis can be carried out. There is no unique statistical procedure for doing this and personal judgment will be a necessary part of any of the statistical methods discussed.

METHODOLOGY

A Hierarchical Multiple Regression Model relates a dependent variable Y, to several independent variables $W_1, W_2, ..., W_k$. The model is in the following form:

$$Y = S_0 + S_1W_1 + S_2W_2 + \ ... \ + S_{k-1}W_{k-1} + S_kW_k + u \quad (1)$$

Basic assumptions of Multiple Regression Models are made about the error terms, u_i (for $i = 1, 2, ..., n$) and the values of independent variables $W_1, W_2, …, W_k$ are shown in Table 1 with S_0 denotes the constant term of the model, S_j denotes the j-th coefficient of independent variable W_j (for $j = 1, 2, …, k$) and u denotes the random residual of the model. The k denotes the number of independent variables $(k+1)$ denotes the total number of parameters.

The general Multiple Regression Model with k independent variables is defined in Eq. 1 where Y denotes the dependent variable, W_j denotes the j-th independent variable (which can be single independent quantitative variable, or interaction variable (first-order interaction, second-order interaction, third-order interaction, ...), or generated variable (dummy or/and polynomial or/and categorical variables) or transformed variable (Ladder transformation and Box-Cox transformation). The S_0 denotes the constant term of the model, S_j denotes the j-th coefficient of independent variable W_j (for $j = 1, 2, …, k$) and u denotes the random residual of the model. The k denotes the number of independent variables, $(k+1)$ denotes the total number of parameters.

The Multiple Regression Model as defined in Eq. 1 is a hierarchically well-formulated models[9] and can be written as a system of n equations, as shown by the following:

$$\begin{aligned} \text{For } i = 1,\ Y_1 &= S_0 + S_1W_{11} + S_2W_{21} + \ ... \ + S_kW_{k1} + u_1 \\ i = 2,\ Y_2 &= S_0 + S_1W_{12} + S_2W_{22} + \ ... \ + S_kW_{k2} + u_2 \\ I = n,\ Y_n &= S_0 + S_1W_{1n} + S_2W_{2n} + \ ... \ + S_kW_{kn} + u_n \end{aligned} \quad (2)$$

This system of n equations can be expressed in matrix terms as the following:

$$Y = WS + u \quad (3)$$

Where:

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ . \\ . \\ . \\ Y_n \end{bmatrix}_{n\times 1} \quad W = \begin{bmatrix} 1 & W_{11} & ... & W_{k1} \\ 1 & W_{12} & ... & W_{k2} \\ . & . & & . \\ . & . & ... & . \\ . & . & & . \\ 1 & W_{1n} & ... & W_{kn} \end{bmatrix}_{n\times(k+1)} \quad \Omega = \begin{bmatrix} \Omega_0 \\ \Omega_1 \\ \Omega_2 \\ . \\ . \\ . \\ \Omega_k \end{bmatrix}_{(k+1)\times 1} \quad u = \begin{bmatrix} u_1 \\ u_2 \\ . \\ . \\ . \\ u_n \end{bmatrix}_{n\times 1}$$

The Y, S and u are in a form of vectors. The W is in a form of matrix. For example, the regression model for two single independent variables (X_1 and X_2) with possible interaction variable (X_{12} is the product of independent variables X_1 and X_2) can be expressed as follows:

$$Y = \$_0 + \$_1X_1 + \$_2X_2 + \$_{12}X_{12} + \$_{11}X_1^{\ 2} + \$_{22}X_2^{\ 2} + \$_{111}X_1^{\ 3} + \$_{222}X_2^{\ 3} + \$_{122}X_1X_2^{\ 2} + u \quad (4)$$

This Eq. 4 can be written in general form as in Eq. 1:

Table 1: Details of multiple regression assumptions

Name	Description	Explanation
Linearity of model	Linear	Linear in the unknown parameters, S_j for $j = 0, 1, ..., k$
Some of the observed data are different	Var(independent variable) > 0	The variance of each W_j is greater than zero for $j = 1, 2, ..., k$. At least one of the observations in each independent variable is different
Conditional mean of u, given W, is zero	u_i is a random error variable for $i = 1, 2, ..., n$	$E(u \mid W_1, W_2, …, W_k) = 0$. The resulting expected value of u, given independent variables
W are given and hence can be treated as non random	$Cov(W_s, u_i) = 0$	The covariance of each independent variable is uncorrelated with u
Homoscedasticity	$Var(u_i \mid W_1, W_2, …, W_k) = F^2$	Variance of u_i, given $W_1, W_2, …, W_k$ has a constant value
Serial independence	$Cov(u_s, u_i \mid W_1, W_2, …, W_k) = 0$	The covariance between u_s and u_i are zero, given $W_1, W_2, …, W_k$, for all s … t independently distributed
Sample size is greater than number of parameters	$n > k$	Number of observations (n) is greater than the number of estimated parameters (k), $n > k$
Normality of errors	Error term is normally distributed	u_i is normally distributed, given independent variables $W_1, W_2, …, W_k$

$$Y = S_0 + S_1W_1 + S_2W_2 + S_3W_3 + S_4W_4 + S_5W_5 + S_6W_6 + S_7W_7 + S_8W_8 + u$$

Where:

$S_0 = \$_0$ constant
$S_1 = \$_1$ and $W_1 = X_1$ single independent
$S_2 = \$_2$ and $W_2 = X_2$ single independent
$S_3 = \$_{12}$ and $W_3 = X_{12} = X_1X_2$ (First-order interaction variable)
$S_4 = \$_{11}$ and $W_4 = X_1^2 = X_1X_1$ (Quadratic variable)
$S_5 = \$_{22}$ and $W_5 = X_2^2 = X_2X_2$ (Quadratic variable)
$S_6 = \$_{111}$ and $W_6 = X_1^3 = X_1X_1X_1$ (Cubic variable)
$S_7 = \$_{222}$ and $W_7 = X_2^3 = X_2X_2X_2$ (Cubic variable)
$S_8 = \$_{122}$ and $W_8 = X_1X_2^2 = X_1X_2X_2$ (interaction variable)

where, $k = 8$ and $(k+1) = 9$.

The variable X_1X_2 denotes the product of variables X_1 and X_2. According to Zainodin and Khuneswari[10], the variable X_1X_2 can also be written as X_{12} and defined it as the first-order interaction variable. In general, the cross product between q single independent quantitative variables, the interaction variable is (q-1)-order interaction variable.

In the development of the mathematical model, there are four phases involved. These phases are possible models, selected models, best model and goodness-of-fit test[11]. In the beginning, all the possible models are listed. Once this has been done, the next step is to estimate the coefficients for the entire possible model and then carry out the tests to get selected models. All the possible models must be run one by one to obtain selected models. In the process of getting the selected models from possible models, multicollinearity test and elimination procedure (elimination of insignificant variables) should be carried out in order to obtain a selected model. These selected models should be free from the multicollinearity sources and insignificant variables.

Multicollinearity arises when two independent variables are closely linearly related. This is equivalent to saying that a coefficient of determination greater than 0.95 represents a strong linear relation. An absolute correlation coefficient greater than 0.95 (i.e., $|r| > 0.95$) defines a strong multicollinearity. The algorithm for the multicollinearity test procedures have also been described by Zainodin *et al.*[11]. The Global test is then carried out as shown by Zainodin and Khuneswari[10]. If there are no more multicollinearity source variables, then the next step is to carry out the elimination procedures on the insignificant variables. This is done by performing the coefficient test for all the coefficients in the model. According to Abdullah *et al.*[12] the coefficient test is carried out for each coefficient by testing the coefficient of the corresponding variable with the value of zero. The insignificant variables will be eliminated one at a time. Subsequent removals are carried out on all the models until all the insignificant variables are eliminated. The Wald test is then carried out on all the resultant selected models so as to justify the removal of the insignificant variables[10].

A best model will be selected from a set of selected models that have been obtained from previous phases in the model-building procedures. The best model will be selected with the help of the eight selection criteria (8SC)[13,11]. The best model is chosen when a particular model has the most of the least criteria value (preferably all the 8 criteria chose the same model).

Finally, the residual analysis should be carried out on the best model to verify whether the residuals are randomly and normally distributed. Bin Mohd *et al.*[14] stated that randomness test should be carried out to investigate the randomness of the residuals produced. One of the MR assumptions is that the residuals should follow a normal distribution. Besides that, Shapiro-Wilk test is used (for $n < 50$) to check the normality assumption of the residuals. Scatter plot, histogram and box-plot of the residuals are to a get a clear picture of distribution of the residual. These plots are used as supporting evidence in addition to the two quantitative tests.

RESULTS

This section will describe the process in getting the best polynomial regression model by including a numerical illustration. In this study, the data was collected from 30 sale regions of detergent 'FRESH' during 1989, as published in Bowerman and O'Connell[15].

In this study, the variables used are the demand (Y) for the detergent bottles to the factors affecting, such as the "price difference" (X_1) between the price offered by the enterprise and the average industry prices of competitors' similar detergents and advertising expenditure (X_2). The aim is to analyse what is the contribution of a specific attribute in determining the demand. Table 2 shows the descriptive statistics for each variable used. As stated by Crawley[16], the distribution of a variable is said to be normal if the value of skewness fall within [-0.4472, 0.4472] and kurtosis fall within [-0.8944, 0.8944].

There are two single quantitative independent variables involved in this data in order to determine the demand of detergent bottles. Figure 1 shows the scatter plots of all the variables involved and their possible trends.

Based on the plots in Fig. 1, it could be seen significantly that Y and X_2 have a polynomial shaped relationship. The fitted curves show a quadratic line relationship. The linear line between Y and X_1 shows a linear relationship and it is supported with the Table 3 that shows Y and X_1 is highly correlated. This shows that there exists possible non linear variable that contributes

Table 2: Descriptive statistics for all independent variables

	Variables							
Statistics	Y	X_1	X_2	X_1^2	X_2^2	X_{12}	X_1^3	X_2^3
Mean	8.3827	0.2133	6.4533	0.0955	41.9670	1.4723	0.0441	274.7323
Standard Error (SE)	0.1244	0.0415	0.1044	0.0213	1.3022	0.2854	0.0122	12.2681
Median	8.3900	0.2000	6.6250	0.0400	43.9030	1.3000	0.0080	291.0859
Mode	8.7500	0.2000	6.8000	0.0025	46.2400	1.3000	0.0080	314.4320
Standard deviation	0.6812	0.2274	0.5719	0.1166	7.1325	1.5632	0.0668	67.1953
Sample variance	0.4641	0.0517	0.3271	0.0136	50.8700	2.4435	0.0045	4515.2037
Kurtosis (SE = 0.8944)	-1.0858	-1.1797	-0.3880	-0.0363	-0.5797	-1.1817	1.2724	-0.7308
Skewness (SE = 0.4472)	-0.1328	0.2221	-0.8348	1.1305	-0.7132	0.3197	1.5291	-0.5921
Minimum	7.1000	-0.1500	5.2500	0.0000	27.5650	-0.7875	-0.0034	144.7031
Maximum	9.5200	0.6000	7.2500	0.3600	52.5650	4.3500	0.2160	381.0781

Table 3: Details of models and their derivatives (interaction, quadratic and cubic)

Model	All possible models
Basic all possible models	
M1	$Y = \beta_0 + \beta_1X_1 + u$
M2	$Y = \beta_0 + \beta_2X_2 + u$
M3	$Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + u$
M4	$Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_{12}X_{12} + u$
Quadratic variables added to models M1 and M2	
M5	$Y = \beta_0 + \beta_1X_1 + \beta_{11}X^2_1 + u$
M6	$Y = \beta_0 + \beta_2X_2 + \beta_{22}X^2_2 + u$
M7	$Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_{11}X_1^2 + \beta_{22}X^2_2 + u$
M8	$Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_{12}X_{12} + \beta_{11}X^2_1 + \beta_{22}X^2_2 + \beta_{111}X^3_1 + \beta_{112}X_1^2X_2 + \beta_{122}X_1X^2_2 + \beta_{222}X_2^3 + u$
Cubic variables added to models M5 and M6	
M9	$Y = \beta_0 + \beta_1X_1 + \beta_{11}X^2_1 + \beta_{111}X^3_1 + u$
M10	$Y = \beta_0 + \beta_2X_2 + \beta_{22}X^2_2 + \beta_{222}X^3_2 + u$
M11	$Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_{11}X_1^2 + \beta_{22}X_2^2 + \beta_{111}X_1^3 + \beta_{2222}X_2^3 + u$
M12	$Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_{12}X_{12} + \beta_{11}X_1^2 + \beta_{22}X_2^2 + \beta_{1112}X_1^3X_2 + \beta_{1222}X_1X_2^3$ $+ \beta_{11112}X_1^4X_2 + \beta_{12222}X_1X_2^4 + \beta_{111}X_1^3 + \beta_{222}X_2^3 + u$

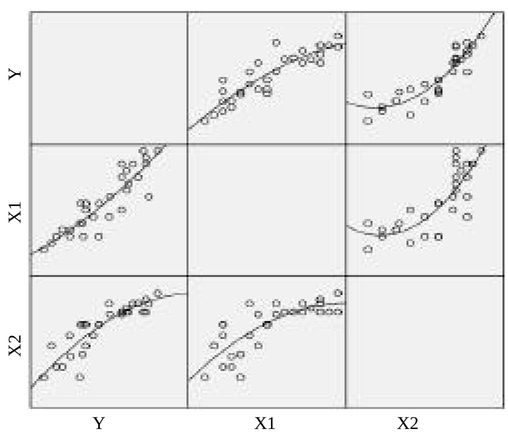

Fig. 1: Scatter plots of all the variables involved and possible trends

to Y. All the quadratic and cubic terms introduced in the models are thus shown in Table 3.

Table 3 shows all the possible models with interaction and polynomial variables. There are four all possible models (M1-M4) for two single independent variable (X_1 and X_2). The polynomial and interaction with polynomial variables are added in model M1 to model M4. Therefore, there are 12 possible models which include quadratic and cubic variables.

Table 4 shows the Pearson correlations matrix for example, the possible model M8.0. It can be seen that there exists multicollinearity among 9 pairs of independent variables (shaded values). The multicollinearity source variables should firstly be removed according to the Zainodin-Noraini Multicollinearity Remedial Techniques[11].

Since there is a tie, the variable X_1^3 (correlation with Y is 0.732) is removed because it has the smallest correlation coefficient with dependent variable (this is a case B of the Zainodin-Noraini multicollinearity remedial technique). After the removal of variable X_1^3, model then becomes model M8.1. The correlation coefficient is recalculated as shown in Table 5. Here, it could be seen that there exists multicollinearity among the 7 pairs of independent variables (shaded values). There exist a tie and the variable X_2 (correlation with Y is 0.876) is then removed because it has the smallest correlation coefficient with the dependent variable (this is again a case B).

After removing variable X_2, the new model is model M8.2. The correlation coefficient is recalculated as shown in Table 6.

Table 4: Pearson correlation for model M8.0

	Y	X_1	X_2	X_{12}	X_1^2	X_2^2	X_1^3	$X_1^2X_2$	$X_1X_2^2$	X_2^3
Y	1	0.890	0.876	0.891	0.779	0.886	0.732	0.783	0.890	0.895
X_1	0.890	1	0.760	0.999	0.924	0.770	0.876	0.923	0.995	0.779
X_2	0.876	0.760	1	0.762	0.635	0.999	0.576	0.640	0.763	0.997
X_{12}	0.891	0.999	0.762	1	0.936	0.774	0.889	0.936	0.999	0.784
X_1^2	0.779	0.924	0.635	0.936	1	0.650	0.987	0.999	0.945	0.663
X_2^2	0.886	0.770	0.999	0.774	0.650	1	0.590	0.655	0.775	0.999
X_1^3	0.732	0.876	0.576	0.889	0.987	0.590	1	0.988	0.899	0.604
$X_1^2X_2$	0.783	0.923	0.640	0.936	0.999	0.655	0.988	1	0.945	0.669
$X_1X_2^2$	0.890	0.995	0.763	0.999	0.945	0.775	0.899	0.945	1	0.787
X_2^3	0.895	0.779	0.997	0.784	0.663	0.999	0.604	0.669	0.787	1

Correlation is significant at the 0.01 level (2-tailed)

Table 5: Pearson correlation for model M8.1

	Y	X_1	X_2	X_{12}	X_1^2	X_2^2	$X_1^2X_2$	$X_1X_2^2$	X_2^3
Y	1	0.890	0.876	0.891	0.779	0.886	0.783	0.890	0.895
X_1	0.890	1	0.760	0.999	0.924	0.770	0.923	0.995	0.779
X_2	0.876	0.760	1	0.762	0.635	0.999	0.640	0.763	0.997
X_{12}	0.891	0.999	0.762	1	0.936	0.774	0.936	0.999	0.784
X_1^2	0.779	0.924	0.635	0.936	1	0.650	0.999	0.945	0.663
X_2^2	0.886	0.770	0.999	0.774	0.650	1	0.655	0.775	0.999
$X_1^2X_2$	0.783	0.923	0.640	0.936	0.999	0.655	1	0.945	0.669
$X_1X_2^2$	0.890	0.995	0.763	0.999	0.945	0.775	0.945	1	0.787
X_2^3	0.895	0.779	0.997	0.784	0.663	0.999	0.669	0.787	1

Table 6: Pearson correlation coefficient for model M8.2

	Y	X_1	X_{12}	X_1^2	X_2^2	$X_1^2X_2$	$X_1X_2^2$	X_2^3
Y	1	0.890	0.891	0.779	0.886	0.783	0.890	0.895
X_1	0.890	1	0.999	0.924	0.770	0.923	0.995	0.779
X_{12}	0.891	0.999	1	0.936	0.774	0.936	0.999	0.784
X_1^2	0.779	0.924	0.936	1	0.650	0.999	0.945	0.663
X_2^2	0.886	0.770	0.774	0.650	1	0.655	0.775	0.999
$X_1^2X_2$	0.783	0.923	0.936	0.999	0.655	1	0.945	0.669
$X_1X_2^2$	0.890	0.995	0.999	0.945	0.775	0.945	1	0.787
X_2^3	0.895	0.779	0.784	0.663	0.999	0.669	0.787	1

Table 7: Pearson correlation coefficient for model M8.3

	Y	X_1	X_{12}	X_1^2	X_2^2	$X_1^2X_2$	X_2^3
Y	1	0.890	0.891	0.779	0.886	0.783	0.895
X_1	0.890	1	0.999	0.924	0.770	0.923	0.779
X_{12}	0.891	0.999	1	0.936	0.774	0.936	0.784
X_1^2	0.779	0.924	0.936	1	0.650	0.999	0.663
X_2^2	0.886	0.770	0.774	0.650	1	0.655	0.999
$X_1^2X_2$	0.783	0.923	0.936	0.999	0.655	1	0.669
X_2^3	0.895	0.779	0.784	0.663	0.999	0.669	1

Here, it could be seen that there exists multicollinearity among the 5 pairs of independent variables. The most commonly appearing variables are X_1, X_{12} and $X_1X_2^2$. Since there is a tie, the variable $X_1X_2^2$ (correlation with Y is 0.890) is removed because it has the smallest correlation coefficient with dependent variable (this is again a case B). The new model after removing variable $X_1X_2^2$ is model M8.3. Then the correlation coefficient was recalculated as shown in Table 7. From Table 7, it can be seen that there exists multicollinearity among the 3 pairs of independent variables.

There are no common variables among the multicollinearity source variables. Therefore, the variable with the smallest correlation coefficient with dependent variable in each pair should be removed. The three variables X_1, X_1^2 and X_2^2 are removed from the model M8.3 (this is a case C where frequency of all the variables is one).

Table 8: Pearson correlation coefficient for model M8.6

Varaibles	Y	X_{12}	$X_1^2X_2$	X_2^3
Y	1	0.891	0.783	0.895
X_{12}	0.891	1	0.936	0.784
$X_1^2X_2$	0.783	0.936	1	0.669
X_2^3	0.895	0.784	0.669	1

The new model after removing the variables X_1, X_1^2 and X_2^2 is then model M8.6. The correlation coefficient is thus recalculated as shown in Table 8. Based on the highlighted triangle in Table 8, it could be seen that there is no more multicollinearity left in the model.

The next step is to carry out the coefficient test of the elimination procedures on the insignificant variables and the Wald test on model M8.6.0. The Wald test is carried out to justify the elimination process. There is only one variable that is omitted from the model M8.6 as shown in Table 9.

Table 10 shows the Wald test for model M8.6.1. From Table 10, F_{cal} is 0.7103 and F_{table} = F(1, 26, 0.05) = 4.23 from the F-distribution table.

Table 9: Illustration of elimination procedure in getting selected model M8.6.1

	Models	
Coefficients	M8.6	M8.6.1
Constant	6.6986	6.6462
$\$_{12}$	0.2877 (2.9457)	0.2138 (4.8555)
$\$_{222}$	0.0049 (4.5542)	0.0052 (5.0523)
$\$_{112}$	-0.1328 (-0.8489)	-
SSE	1.3926	1.4310

*$t_{critical}$ is 2.052 and value in parentheses is the t_{cal}

Table 10: Wald test for model M8.6.1

Source of variations (R-U)	Sum of squares	df	Mean sum of squares	F_{cal}
Differences	0.0380	1	0.0380	0.7103
Unrestricted (U)	1.3926	26	0.0535	
Restricted (R)	1.4310	27		

Table 11: Corresponding selection criteria values for the selected models

Selected models	(k+1)	AIC	FPE	GCV	HQ	RICE	SCH WARZ	SGMA SQ	SHI BATA
M1.0.0	2	0.1069	0.1069	0.1074	0.1101	0.1079	0.1173	0.1002	0.1060
M2.0.0	2	0.1193	0.1193	0.1198	0.1229	0.1205	0.1310	0.1119	0.1183
M3.0.0	3	0.0624	0.0625	0.0631	0.0653	0.0639	0.0718	0.0568	0.0613
M4.1.0	3	0.0625	0.0625	0.0631	0.0653	0.0639	0.0719	0.0568	0.0614
M7.1.1	3	0.0598	0.0598	0.0604	0.0625	0.0612	0.0688	0.0544	0.0587
M8.6.1	5	0.0648	0.0650	0.0668	0.0698	0.0696	0.0818	0.0557	0.0619
M10.2.0	2	0.1021	0.1021	0.1026	0.1052	0.1031	0.1121	0.0958	0.1013
M11.3.1	3	0.0574	0.0575	0.0581	0.0601	0.0588	0.0661	0.0523	0.0564

Since F_{cal} is less than F_{table}, H_0 is accepted. The removal of insignificant variables in the coefficient test is therefore justified. The selected model M8.6.1 is:

$$\hat{Y} = 6.6460 + 0.2140X_{12} + 0.0050X_2^3$$

Similar procedures and tests (Global test, Coefficient test and Wald test) are carried out on the remaining selected models. There are 8 selected models obtained in this phase. For each selected model, the eight selection criterion (8SC) values are obtained and the corresponding values are shown in Table 11.

It can be seen from Table 11 that all the criteria (8SC) indicate that model M11.3.1 as the best model. Thus, the best model is M11.3.1 is given by:

$$\hat{Y} = 6.6398 + 1.4673X_1 + 0.0052X_2^3$$

Referring to Table 4, the Pearson correlation shows the correlation where the "price difference" (X_1) (i.e., r = 0.890) between the price offered by the enterprise and the average industry price of competitors' similar detergents, are highly correlated with the demand for detergent bottles (Y). The components of model M11.3.1 further indicate the presence of polynomial-order (cubic) of the significant independent variables. The variable X_2^3 (i.e., r = 0.8949) are highly correlated with the demand for detergent bottles (Y). The interpretations of the coefficients for best model M11.3.1 are depicted in Table 12.

Table 12: Interpretation of coefficients for best model M11.3.1

Variables	Coefficients	Comments
X_1	1.4673	"Price difference" between the price offered by the enterprise and the average industry price of Competitors' Similar detergents Main factor One unit increase in X_1 will increase the number of detergent demanded (Y) by 1.4673 bottles
X_2^3	0.0052	Advertising expenditure Cubic factor One unit increase in X_2^3 will increase the number of detergent demanded (Y) by 0.0052 bottles

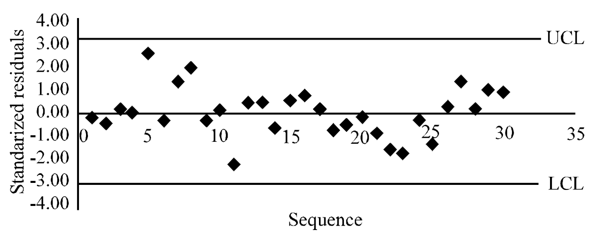

Fig. 2: Scatter plot of standardized residuals for best model M11.3.1

Based on the best model, the residuals analyses are obtained. Several tests on the model's goodness-of-fit are carried out. It is found that all the basic assumptions are satisfied and the residuals plots are shown in Fig. 2-4. Using the residuals obtained, randomness test and the normality test are carried out. Both randomness test and residuals scatter plot, as shown in Fig. 2 and 3, indicate

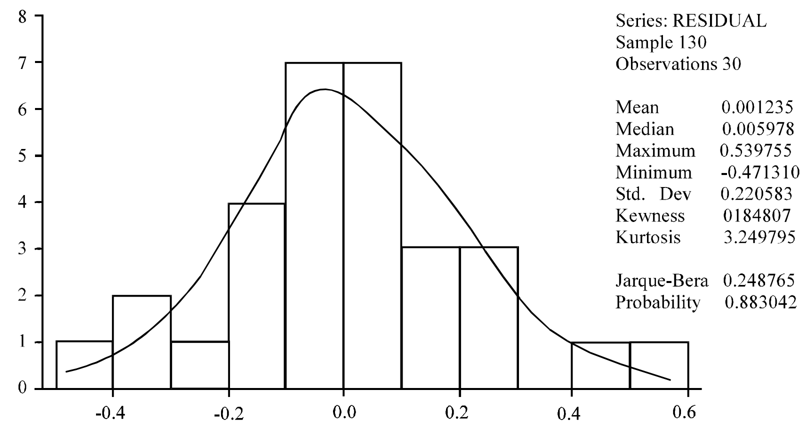

Fig. 3: Histogram of standardized residuals for best model M11.3.1

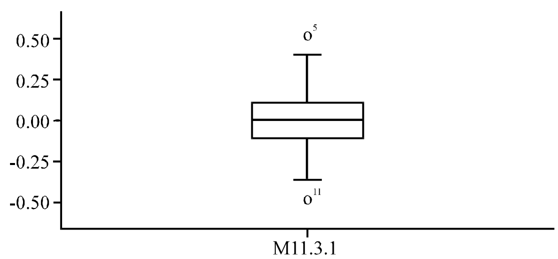

Fig. 4: Box-plot of residuals for best model M11.3.1

that the residuals are random, independent and normal. The total sum of residual of the best model, M11.3.1 is 0.0370 while the sum of square error is 1.4110. The randomness test carried out on residuals shows that resulting error term of best model M11.3.1 is random and independent. This strengthens the belief which is reflected in the residuals plot of Fig. 3 which confirms that no obvious pattern exists. This shows that the best model M11.3.1 is an appropriate model in determining the demand of detergent.

Besides that by taking ± 3 standard deviation (i.e., 99.73%) for Upper Control Limit (UCL) and Lower Control Limit (LCL) as in Fig. 2, the residuals are distributed between the ± 3 standard deviation lines which indicate that there are no outliers. The Shapiro-Wilk statistics of the normality plot in Fig. 3 shows that the residuals of model M11.3.1 are distributed normally (i.e., statistics $= 0.9865$, df $= 30$ and p-value $= 0.9603$). Figure 4 with its median positioned at the centre of the box further shows that model M11.3.1 is a well represented model to describe the demand of detergents. Thus, the model is ready to be used for further analysis. Now the demand model is ready for use in forecasting or estimating to make a logical decision in determining the appropriate demand of detergent.

DISCUSSIONS

Other goodness-of-fit tests can be used such as the root-mean square error (RMSE), Modelling Efficiency (EF) and many others[17,18]. However, in this research, the residuals analysis is based on the randomness and normality tests. The numerical illustration of this study shows that polynomial regression model M11.3.1 is the best model to describe the demand of detergent. The "price difference" ($= X_1$) between the price offered by the enterprise and the average industry price of competitors' with similar detergents is the main factor in determining the demand of detergent and the “advertising expenditure” ($= X_2$) as the cubic factor. The best model M11.3.1 (i.e., $\hat{Y} = 6.6398 + 1.4673X_1 + 0.0052X_2^{\ 3}$) in Table 12 shows that constant demand without any contributions of other factors is approximately 7 bottles (since the constant value is 6.6398). Every one unit increase in “price difference” will directly increase the number of detergent by 1.4673 bottles and every one unit increase in advertising expenditure will positively increase the number of detergent by 0.0052 bottles, respectively. This means that if “price difference” and expenditure are increased by one unit, then the number of detergent demanded will be 8 bottles (approximate value).

Similar to Chu[19], the cubic polynomial approach is used to forecast the demand of detergent. [20]had also attempted the use of price in the consumer purchase decision using price awareness with three distinct factors, namely, price knowledge, price search in store and between stores and the Dickson-Sawyer method[21]. The effect of advertising tools on the behaviour of consumers of detergents had also been surveyed and compared by Salavati *et al*.[22] in developing countries. Moreover, advance researches on promoting hygiene, such as hand hygiene, with criteria like fast and timely, easy and skin protection, are part of the campaigns involved in advertising expenditure[23].

CONCLUSIONS

The consumer buying behaviour will affect the demand for detergents. It is imperative in today's diverse global market that a firm can identify consumer behavioural attributes and needs, lifestyles and purchase processes and the influencing factors which are responsible for the consumer decisions. Hence, devising good marketing plans is necessary for a firm. While serving its targeted markets, it minimises dissatisfaction and still stay ahead of other competitors. It was seen from the best model that there is conclusive increase in production due to price difference and expenditure. In other words, mathematically, a detergent manufacturing company has to project its marketing and production plans to incorporate the impacts of pricing and expenditure on the consumers. Indirectly, the firm succeeded in meeting the demands of the consumers in a win-win relationship.

REFERENCES

1. Sandy, R., 1990. Statistics for Business and Economics. McGraw-Hill Inc., New York, USA.
2. Gujarati, D.N., 2006. Essentials of Econometrics. 3rd Edn., McGraw-Hill Inc., New York, USA.
3. Pasha, G.R., 2002. Selection of variables in multiple regression using stepwise regression. J. Res. Bahauddin Zakariya Univ. Multan Pak., 13: 119-127.
4. Smith, R.T. and R.B. Minton, 2007. Calculus: Early Transcendental Functions. 3rd Edn., McGraw-Hill, Inc., New York, USA., ISBN:139780071107518.
5. Beach, E.F., 1950. The use of polynomials to represent cost functions. Rev. Econ. Stud., 16: 158-169.
6. Wackerly, D.D., W. Mendenhall and R.L. Scheaffer, 2002. Mathematical Statistics with Applications. 6th Edn., Thomson Learning, South Western Ohio, USA.
7. Al-Hazza, M.H.F., E.Y.T. Adesta, A.M. Ali, D. Agusman and M.Y. Suprianto, 2011. Energy cost modeling for high speed hard turning. J. Applied Sci., 11: 2578-2584.
8. Fonseca, G.G., P.S. da Silva Porto and L.A. de Almeida Pinto, 2006. Reduction of drying and ripening times during the Italian type salami production. Trends Applied Sci. Res., 1: 504-510.
9. Jaccard, J., 2001. Interaction Effects in Logistics Regression. SAGE Publication, London, UK., ISBN-13: 9780761922070.
10. Zainodin, H.J. and G. Khuneswari, 2009. A case study on determination of house selling price model using multiple regression. Malaysian J. Math. Sci., 3: 27-44.
11. Zainodin, H.J., A. Noraini and S.J. Yap, 2011. An alternative multicollinearity approach in solving multiple regression problem. Trends Applied Sci. Res., 6: 1241-1255.
12. Abdullah, N., Z.H.J. Jubok and J.B.N. Jonney, 2008. Multiple regression models of the volumetric stem biomass. WSEAS Trans. Math., 7: 492-502.
13. Ramanathan, R., 2002. Introductory Econometrics with Application. 5th Edn., Harcourt College Publishers, South Western, ISBN: 9780030341861, Pages: 688.
14. Bin Mohd, I., S.C. Ningsih and Y. Dasril, 2007. Unimodality test for global optimization of single variable functions using statistical methods. Malaysian J. Math. Sci., 1: 205-215.
15. Bowerman, B.L. and R.T. O?Connell, 1990. Linear Statistical Models and Applied Approach. 2nd Edn., PWS-Kent Publishing Company, Boston .
16. Crawley, M.J., 2005. Statistics: An Introduction Using R. John Wiley and Sons Ltd., England, pp: 155-185.
17. Mohammadi, A., S. Rafiee, A. Keyhani and Z. Emam-Djomeh, 2009. Moisture content modeling of sliced kiwifruit (cv. *Hayward*) during drying. Pak. J. Nutr., 8: 78-82.
18. Meisami-Asl, E., S. Rafiee, A. Keyhani and A. Tabatabaeefar, 2009. Mathematical modeling of moisture content of apple slices (Var. Golab) during drying. Pak. J. Nutr., 8: 804-809.
19. Chu, F.L., 2004. Forecasting tourism demand: A cubic polynomial approach. Tourism Manage., 25: 209-218.
20. Kenesei, Z. and S. Todd, 2003. The use of price in the purchase decision. J. Empirical Generalisations Market. Sci., 8: 1-21.
21. Dickson, P.R. and A.G. Sawyer, 1990. The price knowledge and search of supermarket shoppers. J. Market., 54: 42-53.
22. Salavati, A., R. Shafei and M.Y. Mazhari, 2010. Surveying and comparing the effect of advertising tools on the behavior of consumers of detergents (case study in developing countries). Eur. J. Soc. Sci., 17: 96-108.
23. Pittet, D., 2001. Improving adherence to hand hygiene practice: A multidisciplinary approach. Emerging Inf. Dis., 7: 234-240.