Interpretation of Stepwise Multiple Linear Regression Analysis in Behavioral Sciences in APA Format

Abstract

In several methodological steps, the paper simplifies the procedure for interpreting the stepwise multiple linear regression. Nowadays the researchers in the field of behavioral sciences frequently use the statistical software's like SPSS, AMOS etc. for statistical calculations, and the important thing that these software's cannot do is the interpretation. The researcher finds it very difficult to draw meaningful conclusions while doing any statistic from the output tables of these software's. Therefore, in this paper one statistic technique (Stepwise Multiple Linear Regression Analysis) was applied with the help of SPSS on the data of one dependent variable and multiple independent variables with the purpose of drawing meaningful interpretations. The paper also highlights the important points that need to be take care of while interpreting the output while doing the Stepwise Multiple Linear Regression Analysis.

Mukhtar Ahmad Wani
Lecturer,
School Education Department,
J & K

Introduction

Regressions techniques are primarily used in order to create an equation which can be used to predict values of dependent variables for all members of the population. A secondary function of using regression is that it can be used as a means of explaining causal relationships between variables. Stepwise regression is a semi-automated process of building a model by successively adding or removing variables based solely on the *t*-statistics of their estimated coefficients. Properly used, the stepwise regression option in SPSS (or other stat packages) puts more power and information at your fingertips than does the ordinary multiple regression option, and it is especially useful for sifting through large numbers of potential independent variables and/or fine-tuning a model by poking variables in or out. Improperly used, it may converge on a poor model while giving you a false sense of security. It's like doing carpentry with a chain saw: you can get a lot of work done quickly, but it leaves rough edges and you may end up cutting off your own foot if you don't read the instructions, remain sober, engage your brain, and keep a firm grip on the controls. It is not a tool for beginners or a substitute for education and experience.

How It Works

Suppose you have some set of potential independent variables from which you wish to try to extract the best subset for use in your forecasting model. (These are the variables you will select on the initial input screen.) The stepwise option lets you either begin with *no* variables in the model and proceed *forward* (adding one variable at a time), or start with *all* potential variables in the model and proceed *backward* (removing one variable at a time). At each step, the program performs the following calculations: for each variable currently *in* the model, it computes the *t*-statistic for its estimated coefficient, *squares* it, and reports this as its "*F*-to-remove" statistic; for each variable *not* in the model, it computes the *t*-statistic that its coefficient *would* have *if it were the next variable added*, squares it, and reports this as its "*F*-to-enter" statistic. At the next step, the program automatically enters the variable with the highest *F*-to-enter statistic, or removes the variable with the lowest *F*-to-remove statistic, in accordance with certain control parameters you have specified. So the key relation to remember is: **F = t-squared**

In the multiple regression procedure in most statistical software packages, you can choose the stepwise variable selection option and then specify the method as "Forward" or "Backward," and also specify *threshold*

Remarking An Analisation

values for *F*-to-enter and *F*-to-remove. (You can also specify "None" for the method--which is the default setting--in which case it just performs a straight multiple regression using all the variables.) The program then proceeds automatically. Under the *forward* method, at each step, it *enters the variable with the largest F-to-enter statistic*, provided that this is *greater* than the threshold value for *F*-to-enter. When there are no variables left to enter whose *F*-to-enter statistics are above the threshold, it checks to see whether the *F*-to-remove statistics of any variables added previously have fallen *below* the *F*-to-remove threshold. If so, it removes the worst of them, and then tries to continue. It finally stops when no variables either in or out of the model have *F*-statistics on the wrong side of their respective thresholds. The *backward* method is similar in spirit, except it starts with all variables in the model and successively *removes the variable with the smallest F-to-remove statistic*, provided that this is *less* than the threshold value for *F*-to-remove.

Whenever a variable is entered, its new *F*-to-remove statistic is initially the same as its old *F*-to-enter statistic, but the *F*-to-enter and *F*-to-remove statistics of the *other* variables will generally all change. (Similarly, when a variable is removed, its new *F*-to-enter statistic is initially the same as its old *F*-to-remove statistic.) Until the *F*-to-enter and *F*-to-remove statistics of the other variables are recomputed, it is impossible to tell what the *next* variable to enter or remove will be. Hence, this process is *myopic*, looking only one step forward or backward at any point

Assumptions

Normality

All errors should to be normally distributed, which can be tested by looking at the skewness, kurtosis, and histogram plots. Technically normality is necessary only for the t-tests to be valid, estimation of the coefficients only requires that the errors be identically and independently distributed

Independence

The errors associated with one observation are not correlated with the errors of any other observation.

Linearity

The relationship between the IVs and DV should be linear.

Homoscedasticity

The variances of the residuals across all levels of the IVs should be consistent, which can be tested by plotting the residuals.

Objective of the Study

To make the learners familiar about the interpretation of stepwise multiple regression in APA Format in behavioral sciences.

Review of Literature

The term "Regression" was first used by Francis Galton with reference to the inheritance of stature. Galton found that children of tall parents tend to be less tall, and children of short parents less short, than their parents. In other words, the heights of the offspring tend to 'move back' toward the mean height of the general population. This tendency toward maintaining the "mean height" Galton called the

principle of regression. Similarly, regression analysis presently comes to mind since we are attempting to establish (predict) the value of one variable based on the knowledge of others. Behavioral sciences dealing with dynamic aspects of behavior frequently use regression analysis with the help of computer statistics. Many researchers use regression analysis technique for calculation like Aejaz Masih (1992), Josse Antony (2001), Akram (2010), Subnana Rais (2011), Hera Singh (2012), Nayyer Jabeen (2013), Mehraj Bhat (2015), Bilal Ahmad (2017) and predict one variable from other variables. What is the difference in these studies is the type of interpretation they have followed. The studies which have used manual calculations like Masih (1992) have interpreted in one of the best ways while presenting the results. But today the researchers have to follow different formats like MLA, APA etc while reporting the results. At the present researchers do not use manual calculations and doing calculations with the help of computer software, it becomes very difficult to report the results. Therefore, this paper is based on the theme that how to interpret the regression analysis output while reporting the results in APA format.

As an Empirical Study

Empiricalresearch is thestudy using empirical evidence. An empirical research reports a study that used quantitative research methods, which generate numerical data and seek to establish causal relationships between two or more variables. Hence this study also falls under the empirical research by employing the Correlational (Prediction) design of research. The participants for this study were selected through the multi-stage sampling technique from different Govt. Higher Secondary Schools of two districts (Anantnag & Kulgam) of South Kashmir (J&K). Total sample of 705 students participated in the study.

Duration of the Study

To make the researchers understand the interpretation process after conducting regression analysis with help of SPSS, the investigator first collect the data from the sample participants which took about six months. Then entering, tabulation, analysis and interpretation of data took almost about next six months. The paper becomes ready after one year of work done on it.

Analysis through SPSS

Let us take the example to understand the stepwise multiple linear regression analysis and its interpretation. I am having the data on achievement motivation and different dimensions of school environment. There are eight dimensions of school environment that act as predictors or independent variables and achievement motivation as criterion variable or dependent variable. *Creative Stimulation (A), Cognitive Encouragement (B), Permissiveness (C), Acceptance (D), Rejection (E), Control (F), Physical aspects of the School (G) and Student Teacher Interaction (H) were entered as independent variables and Achievement motivation (AM) as dependent variable in SPSS 21. Table 1, 2 & 3 presents the output.*

Table 1: Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.486[a]	.237	.236	15.47887	.237	217.866	1	703	.000
2	.602[b]	.363	.361	14.15417	.126	138.747	1	702	.000
3	.632[c]	.399	.396	13.75531	.036	42.301	1	701	.000
4	.648[d]	.420	.417	13.51562	.022	26.084	1	700	.000
5	.664[e]	.441	.437	13.28710	.020	25.285	1	699	.000
6	.672[f]	.452	.447	13.16478	.011	14.050	1	698	.000

a. Predictors: (Constant), C
b. Predictors: (Constant), C, B
c. Predictors: (Constant), C, B, A
d. Predictors: (Constant), C, B, A, F
e. Predictors: (Constant), C, B, A, F, H
f. Predictors: (Constant), C, B, A, F, H, D

Table 2: ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	52199.758	1	52199.758	217.866	.000[b]
	Residual	168435.641	703	239.596		
	Total	220635.399	704			
2	Regression	79996.422	2	39998.211	199.651	.000[c]
	Residual	140638.977	702	200.340		
	Total	220635.399	704			
3	Regression	88000.194	3	29333.398	155.032	.000[d]
	Residual	132635.205	701	189.209		
	Total	220635.399	704			
4	Regression	92765.025	4	23191.256	126.956	.000[e]
	Residual	127870.374	700	182.672		
	Total	220635.399	704			
5	Regression	97229.037	5	19445.807	110.145	.000[f]
	Residual	123406.362	699	176.547		
	Total	220635.399	704			
6	Regression	99664.026	6	16610.671	95.843	.000[g]
	Residual	120971.373	698	173.311		
	Total	220635.399	704			

a. Dependent Variable: AM
b. Predictors: (Constant), C
c. Predictors: (Constant), C, B
d. Predictors: (Constant), C, B, A
e. Predictors: (Constant), C, B, A, F
f. Predictors: (Constant), C, B, A, F, H
g. Predictors: (Constant), C, B, A, F, H, D

Table 3: Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	77.421	3.489		22.187	.000					
	C	1.940	.131	.486	14.760	.000	.486	.486	.486	1.000	1.000
2	(Constant)	45.507	4.186		10.871	.000					
	C	1.656	.123	.415	13.511	.000	.486	.454	.407	.961	1.040
	B	1.564	.133	.362	11.779	.000	.444	.406	.355	.961	1.040
3	(Constant)	34.601	4.400		7.864	.000					
	C	1.513	.121	.379	12.491	.000	.486	.427	.366	.930	1.076
	B	1.425	.131	.330	10.896	.000	.444	.381	.319	.936	1.069
	A	.397	.061	.198	6.504	.000	.344	.239	.190	.929	1.076

146

Model	Variable	B	Std. Error	Beta	t	Sig.				Tolerance	VIF
4	(Constant)	51.051	5.391		9.469	.000					
	C	1.432	.120	.359	11.920	.000	.486	.411	.343	.913	1.095
	B	1.385	.129	.321	10.755	.000	.444	.377	.309	.932	1.073
	A	.401	.060	.200	6.690	.000	.344	.245	.193	.929	1.077
	F	-.620	.121	-.149	-5.107	.000	-.235	-.190	-.147	.975	1.026
5	(Constant)	38.900	5.825		6.678	.000					
	C	1.381	.119	.346	11.653	.000	.486	.403	.330	.907	1.103
	B	1.182	.133	.274	8.898	.000	.444	.319	.252	.846	1.182
	A	.378	.059	.188	6.390	.000	.344	.235	.181	.923	1.083
	F	-.617	.119	-.148	-5.166	.000	-.235	-.192	-.146	.975	1.026
	H	.747	.149	.152	5.028	.000	.337	.187	.142	.870	1.149
6	(Constant)	32.635	6.008		5.431	.000					
	C	1.326	.118	.332	11.205	.000	.486	.390	.314	.893	1.120
	B	1.088	.134	.252	8.116	.000	.444	.294	.227	.816	1.225
	A	.355	.059	.177	6.026	.000	.344	.222	.169	.913	1.095
	F	-.590	.119	-.141	-4.972	.000	-.235	-.185	-.139	.971	1.030
	H	.707	.148	.144	4.788	.000	.337	.178	.134	.865	1.155
	D	.455	.121	.112	3.748	.000	.320	.140	.105	.884	1.131

a. Dependent Variable: AM

Being freshers in the field of research in behavioral sciences, a student finds it difficult to draw meaningful conclusion from these above tables. Therefore, for the convenience of students in behavioral science I tried to simplify the output tables 1, 2 & 3 by converting them in tables 4, 5 & 6 followed by the meaningful interpretation of the calculations.

Table 4: Model Summary: Showing the different variables of school environment as predictors of achievement motivation

Predictors Variables Entered	R	R Square	Adjusted R Square	Change Statistics		
				R Square Change (ΔR^2)	F Change	Sig. F Change
C (Permissiveness)	.486	.237	.236	.237	217.866	.000
B (Cognitive Encouragement)	.602	.363	.361	.126	138.747	.000
A(Creative Stimulation)	.632	.399	.396	.036	42.301	.000
F (Control)	.648	.420	.417	.022	26.084	.000
H (Student teacher interaction)	.664	.441	.437	.020	25.285	.000
D (Acceptance)	.672	.452	.447	.011	14.050	.000

Table 5: ANOVA Table for Final Model

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	99664.026	6	16610.671		
Residual	120971.373	698	173.311	95.843	.000
Total	220635.399	704			

Table 6: Coefficients for The Final Model

Model Variables Entered	Standardized Coefficients	t	Sig.	r	Collinearity
	Beta				Tolerance
(Constant)		5.43	.000		
C (Permissiveness)	.332	11.20	.000	.39	.893
B (Cognitive Encouragement)	.252	8.11	.000	.29	.816
A (Creative Stimulation)	.177	6.02	.000	.22	.913
F (Control)	-.141	-4.97	.000	-.18	.971
H (Student teacher interaction)	.144	4.78	.000	.17	.865
D (Acceptance)	.112	3.74	.000	.14	.884

Interpretation in APA Format

A stepwise multiple regression was conducted to evaluate whether different dimensions of School Environment (A, B, C, D, E, F, G, & H) predict Achievement Motivation. It is observed in tables 4, 5 and 6 that different dimensions of School Environment (A, B, C, D, E, F, G, & H) that appeared as significant predictors of Achievement Motivation (criterion variable), only six i.e. C, B, A, F, H and D were upheld as significant predictors that in combination contributed significantly to Achievement Motivation **(F (6, 698) = 95.84, p < 0.01)** of higher secondary school students. As observed from coefficients table, all tolerance values much above the zero. Tolerance values are a measure of the correlation between the predictor variables and can carry a value between 0 to

1. The closer to 0 the tolerance is for a variable, the stronger relationship between this and the other predictor variables. We should worry about variables that have a very low tolerance. Thus, on account of tolerance values, it is concluded that a low level of multicollinearity is present (tolerance = .89, .81, .91, .97 .86 and .88) for C (Permissiveness), B (Cognitive Encouragement), A (Creative Stimulation), F (Control), H (Student teacher interaction) and D (Acceptance) respectively. Multiple correlation is found to be R= .67 for C, B, A, F, H, and D which accounted for 45.2% variance achievement motivation scores. As apparent from the standardized coefficient (β), bearing t value that is significant at 0.01 level, Permissiveness (C) has the predictive power (β = .33; p < 0.01) for achievement motivation of higher secondary school students and contributed about 23.7% in achievement motivation of higher secondary school students (ΔR^2 = .237, F (1, 703) = 217.86; p < 0.01). This indicates that as the teachers provide a school climate in which students express their views freely, their achievement oriented level also increases (r = .39).

Cognitive Encouragement (B)

Cognitive Encouragement (B) emerged as the next significant potential predictor (β = -.25; p < 0.01) of Achievement motivation of higher secondary school students which contributed approximately 12.6% (ΔR^2 = .126, F (1, 702) = 138.74; p < 0.01) of variance in Achievement motivation of higher secondary school students. It means that when the teachers behavior stimulates the cognitive development of students by encouraging their actions, they increase the level of achievement motivation of their students (r = .29).

Creative Stimulation (A)

In the same way Creative Stimulation (A) further increases the variance by 3.6% making the prediction to improve further in a significant manner (ΔR^2 = .036, F (1, 701) = 42.30; p < 0.01) and exerts a positive influence (β = .17; p < 0.01) on achievement motivation of higher secondary school students (r = .22). This means that increase in creative stimulation by the teachers in schools leads to the increase in the achievement motivation scores.

Control (F)

Control (F) also increases the variance by 2.2% making the prediction to improve further in an inverse significant manner (ΔR^2 = .022, F (1, 700) = 426.08; p < 0.01) and exerts a significant negative influence (β = -.14; p < 0.01) on achievement motivation of higher secondary school students (r = -.18). It means decrease in autocratic atmosphere in the school leads to the increase in achievement motivation scores and vice-versa.

Student teacher interaction (H)

Student teacher interaction (H) dimension also increases the variance by 2.0% making the prediction to improve further in a significant manner (ΔR^2 = .020, F (1, 699) = 25.28; p < 0.01) and exerts a significant positive influence (β = .14; p < 0.01) on achievement motivation of higher secondary school students (r = .17), indicating that as the teachers provide an empathetic and joyful atmosphere for learning environment in the school, it will consequently improve the achievement oriented level of their students.

Acceptance (D)

Lastly Acceptance (D) increases the variance by 1.1% making the prediction to improve further in a significant manner (ΔR^2 = .011, F (1, 698) = 14.05; p < 0.01) and exerts a significant positive influence (β = .11; p < 0.01) on achievement motivation of higher secondary school students (r = .14), indicating that as the teachers accept the feelings of their students in a non-threatening manner, it will improve the achievement oriented level of their students.

Discussion

Under the model summary shown in table 4, SPSS tells us what the dependent variable was and what the predictors were in each of the six models. In the column labeled R are the values of the multiple correlation coefficients between the predictors and the outcome (DV). The next column gives us a value of R^2 which is the measure of how much of the variability in the dependent variable is accounted for by the predictors as explained in the interpretation section. The Adjusted R^2 gives us idea of how well our model generalizes and ideally we would like its value to be the same as, or the value of R^2. In our example the difference for the final model is small (.452 - .447 = .005 or 0.5%). This shrinkage means that if the model were derived from the population rather than a sample it would account for approximately 0.5% less variance in the outcome. This indicates that the cross validity of this model is very good. The change statistics are provided only if requested in SPSS and these tell us whether the change in R^2 is significant and how much variance one independent variable leads to change in dependent variable separately. Similarly the F ratio for the final model represents the ratio of the improvement in prediction through its Sig Value. Through ANOVA we can directly manipulate the factors and measure the resulting change in the dependent variable. In multiple regression we simply measure the naturally occurring scores on a no. of predictor variables and try to establish which set of the observed variables gives rise to the best prediction of the criterion variable.

The Beta (β) values (Standard Regression Coefficients) is a measure of how strongly each predictor variable influences the criterion variable. E.g. a beta of 2.5 indicates that a change of one standard deviation in the predictor variable will result in a change of 2.5 standard deviations in the criterion variable. Thus the higher the beta value the greater the impact of the predictor variable on the criterion variable. Making interpretations in behavioral sciences we mainly focuss on standardized beta weights and in economics and finance we focus on unstandardized beta weights. Partial correlation in regression analysis lists the partial correlation for each predictor as it was evaluated for its weighing in the model (the correlation between predictor and the dependent variable) when other predictors are treated as covariates.

Remarking An Analisation

Conclusion

In stepwise regressions decision about the order in which predictors are entered into the model are based on a purely mathematical criterion. In this method, an initial model is defined that contains only the constant (I.V.). The computer then searches for the predictor (out of the ones available) that best predicts the outcome variable. It does this by selecting the predictor that has the highest simple correlation with the outcome- then this predictor is retained in the model and the SPSS searches for a second predictor. The precondition for doing and interpreting regression modeling is that our independent variable must not be too strongly correlated with one another. If they are, this will cause serious problems in estimating the relationship between the dependent and independent variables, as it becomes hard to calculate the individual contribution of each variable.

Regression analysis allows us to do a number of things. Firstly, we can look at how well all our predictor variables together predict the outcome variable. R^2 will give us a statistic (Between 0 and 1) that will tell us that. Secondly, we can look at the relationship between each of our predictors separately and the outcome variable. For each predictor, we can calculate a relationship that takes into account the effect of all the other independent variables. The coefficient Beta (β) gives us the amount the dependent variable changes by if our predictor goes up by 1. Beta (β) standardized version that allows us to determine which of our predictors has the strongest relationship with the outcome variable. The

significance of the relationship is also calculated. In multiple linear regression, the independent variables can be continuous, ordinal or nominal (if we use dummy variable). The dependent variable must be continuous.

References

1. Anastasi, A. (1987). Psychological Testing. New York: Macmillan Co.
2. Cunningham, J. B., & Aldrich, J. O. (2012). Using SPSS: An Interactive Hands-On Approach. New Delhi: Sage Publications India Pvt. Ltd.
3. Ferguson, G. A., & Yashio, T. (1989). Statistical Analysis in Psychology and Education. New York: McGraw Hill Book Co.
4. Field, A. (2014). Discovering Statistics using IBM SPSS Statistics (4th ed.). New Delhi: Sage Publications India Pvt. Ltd.
5. Garret, H. E. (1973). Statistics in Psychology and Education. Bombay Vakils: Feffer & Simons Pvt Ltd.
6. Guilford, J. P. (1987). Psychometric Methods. New York: McGraw Hill.
7. Guilford, J. P., & Beryamin, F. (1946). Fundamental Statistics in Psychology and Education. Singapore: McGraw Hill Book Co.
8. Kerlinger, F. N. (1983). Foundations of Behavioural Research (2nd ed.). Delhi: Surjeet Pub. Molt, Rinchart & Winston, Inc.
9. Mujis, D. (2011). Doing Quantitative Research in Education with SPSS (2nd ed.). New Delhi: Sage Publications India Pvt. Ltd.