



SDI Review Form 1.6

Journal Name:	British Journal of Mathematics & Computer Science
Manuscript Number:	2015_BJMCS_18616
Title of the Manuscript:	CLUSTER ANALYSIS AS A STRATEGY OF GROUPING TO CONSTRUCT GOODNESS-OF-FIT TESTS WHEN THE CONTINUOUS COVARIATES PRESENT IN THE LOGISTIC REGRESSION MODEL
Type of the Article	Original Research Article

General guideline for Peer Review process:

This journal's peer review policy states that **NO** manuscript should be rejected only on the basis of '**lack of Novelty**', provided the manuscript is scientifically robust and technically sound.

To know the complete guideline for Peer Review process, reviewers are requested to visit this link:

(<http://www.sciencedomain.org/page.php?id=sdi-general-editorial-policy#Peer-Review-Guideline>)



SDI Review Form 1.6

PART 1: Review Comments

	Reviewer's comment	Author's comment (if agreed with reviewer, correct the manuscript and highlight that part in the manuscript. It is mandatory that authors should write his/her feedback here)
<u>Compulsory</u> REVISION comments	Use of classical Goodness-of-fit-tests (GOF) , i.e., Pearson or deviance chi-square, in logistic regression are often not practical in majority of recent applications- specially when continuous predictors are present or when number of profiles are less than 5. The proposed approach is practical and also address the main issue of lack-of-fit and how to improve the fit in logistic regression. The manuscript provide a comprehensive review of the literature of how the different GOF statistics were evolved and summarizes other approaches proposed in the literature. Finally, the authors propose a new approach based on cluster analysis and provide simulation results for their GOF statistics. This is a good manuscript that can be improved. These are the main comments: 1. The authors need to provide in their discussion the exact differences between their approach and that of Xie et al[15]. If possible, show how the proposed method works better than Xie et al .[15]. 2. HL GOF and other methods are not stable with respect to number of groupings (e.g, 10 or 15 can affect the HL and its p-value). The proposed method in this manuscript use a cluster analysis as a strategy , but it does not address the number of groupings. Please add a paragraph in the	Thank you for all comments 1- Please refer to subsection 4-2-1 where we change the simulation study to compare the performance of the proposed GOF tests with Xie et al. [15] GOF tests. 2- Please refer to subsection 2.1 and subsection 4.1, where we addressed some statistical methods on how to choose the number of clusters and we suggested to use a multiple runs



SDI Review Form 1.6

	discussion to address recent studies on how to choose the number of groupings and future studies.	in the range of 5-15 and chose the appropriate number of clusters.
<u>Minor</u> REVISION comments	Line 29: Revise this sentence “because it is measuring how far is the model from the data; more than how much is the model good” Line 59: Revise: “A different approach is proposed by Tsiatis [13], instead of grouping observations by their predicted outcomes; he partitions the multidimensional...” Line 101: Revise : All the reviewed works above “ARE” based on the ...” Line 103: “Ward’s method “ : Provide reference. Line 190: revise: This problem is how many clusters are present in a given data set?	We revised it. We revised it. We revised it. Xie, <i>et al.</i> [15]. We revised it.
<u>Optional/General</u> comments	-	