FINAL REPORT OF
UGC STARTUP RESEARCH PROJECT
[No.F. 30-438/2018 (BSR)]

Latent Class Analysis of Survey data on Cheating Behaviour

SUBMITTED TO

UNIVERSITY GRANT COMMISSION
BAHADUR SHAH ZAFAR MARG
NEW DELHI-110002

SUBMITTED BY

DR. SUNIL KUMAR
Assistant Professor
Department Of Statistics
University Of Jammu
Jammu-180006

Phone : (O) 0191-2457254

DEPARTMENT OF STATISTICS
NEW UNIVERSITY CAMPUS
UNIVERSITY OF JAMMU
JAMMU (J&K)–180 006 (INDIA)

Ref. No. Stat./ 21/0012 **Dated.** 05-04-2021

Dr Sunil Kumar, M.Sc., M.Phil., Ph.D.,
Assistant Professor
Department of Statistics, University of Jammu, Jammu

To
Deputy Secretary
Basic Scientific Research Unit
University Grants Commission
Ministry of Human Resource Development, Govt. of India
Bahadurshah Zafar Marg, New Delhi – 11002

Respected Sir/Madam,

Subject: Submission of Final Report for UGC Start Up Grant

Greetings, I express my heartfelt thanks for releasing and approving the UGC-BSR Start-Up grant for newly recruited faculty at Assistant Professor Level in Science departments of various Universities for the session 2018-2020. Grant released vide order no.: F.30-438/2018(BSR) dated 07/06/2018. **Due to the COVID 19 and sad demise of my father**, I am unable to submit the report along with the required documents (Statement of Expenditure, Utilization Certificate, etc).

I am here by submitting consolidated final report and will submit the other required documents as soon as possible. Kindly acknowledge the same.

Thanking you

Yours Sincerely,

Dr. Sunil Kumar
Principal Investigator
Department of Statistics
University of Jammu
Jammu

Phone : (O) 0191-2457254

DEPARTMENT OF STATISTICS
NEW UNIVERSITY CAMPUS
UNIVERSITY OF JAMMU
JAMMU (J&K)–180 006 (INDIA)

Ref. No. Stat./22/146 **Dated.**04-07-2022

To

Deputy Secretary
Basic Scientific Research Unit
University Grants Commission
Ministry of Human Resource Development, Govt. of India
Bahadurshah Zafar Marg, New Delhi – 11002

Subject: Submission of refund of unspent balance of **Rs. 4, 33, 165/-** to Secretary, UGC (New Delhi)

Respected Sir/Madam,

Greetings, I express my heartfelt thanks for releasing and approving the UGC Start – Up Grant to me under the scheme UGC, BSR Start – Up grant for the session 2018-20. Grant released vide order no. F.30-438/2018(BSR) dated 07/06/2018. I had **submitted the hard copy of the report with reference no Stat/21/0012 dated 05/04/2021**. Further, I had **submitted Utilization Certificate and Statement of Expenditure with reference no. Stat/283-284 dated 17/11/2021**. But, I have not received any receipt from your side about the same.

I am submitting that the refund of unspent balance plus bank interest of **Rs. 4,33,176 has been refunded to the Secretary, UGC** and credited to the account No. 8627101002122, IFSC Code: CNRB0008627 of UGC Branch, New Delhi vide UTR No. JAKAR52022060900070322 dated 09/06/2022.

In this regard, request you for **closing of the accounts of the project and acknowledge the same.**

Thanking you

Yours Sincerely,

04/07/2022
Dr Sunil Kumar
Principal Investigator
Department of Statistics
University of Jammu
Jammu

Enclosure: Letter from University to Dr Sunil Kumar about refund to Secretary UGC (New Delhi)

University of Jammu

No: Grants/2022-23/185
Dated :- 29/06/2022

Dr. Sunil Kumar
Principal Investigator
Department of Statistics
University of Jammu

Subject: Refund of unspent balance of Rs. 4,33,167/- to Secretary, UGC (New Delhi).

Sir

This is in reference to your letter No PGD/Stat/22/99 dated: 27/05/2022 regarding refund of the unspent balance for an amount of Rs. 4,33,167/- of the project sanctioned to you under the scheme **UGC-BSR Start-up grant newly recruited faculty at Assistant Professor level in science department** by the UGC. In this connection, the same has been refunded to the Secretary, UGC and credited to the account No. 8627101002122, IFSC Code : CNRB0008627 of UGC, Branch, New Delhi vide UTR No. JAKAR52022060900070322 dated 09/06/2022.

You are, therefore, requested to inform the UGC accordingly for closing of the accounts of the project.

Thanking You

Yours faithfully

Deputy Registrar (Grants)

28/6

Copy to:
1. P.A to Joint Registrar (Finance) for information

University of Jammu

No: Grants/2022-23/76
Dated: 9-6-22

Manager,
J&K Bank Ltd,
New University Campus,
Jammu.

Sub: Payment through NEFT/RTGS

Sir,

On the Authority of this Letter, Kindly credit the following Amount by debit to our NEW Virtual Saving Account No. 0345040160000078

ENCLOSURES: (Canara Bank)

S.No.	Particulars	Amount	Head	Department
76	Secretary, UGC (New Delhi)	Rs.4,33,167 /-	U-141 (Refund of the unspent balance)	Statistics

Rupees Four Lac Thirty Three Thousand One Hundred & Sixty Seven Only

Joint Registrar (Finance)
9/6/22

JAKARTA
520226090070322

Dr. Teacher
Cr.

Adva
Dr. Teach
Cr. Bank

Cr. Teacher Name
Dr. Bank

Dr. Bank
Cr. Teach

UNIVERSITY GRANTS COMMISSION
BAHADUR SHAH ZAFAR MARG
NEW DELHI - 110002

Statement of expenditure in respect of Start-Up-Grant 2018-2020

1. Name of Principal Investigator: **Dr Sunil Kumar**
2. Department of Principal Investigator: **Department of Statistics, University of Jammu**
3. UGC approval Letter no. & Date: **No. F.30-438/2018(BSR) dated 23/05/2018**
4. Title of the Research Project: **Latent Class Analysis of Survey Data on Cheating Behaviour.**
5. Effective date of starting the project: **07/06/2018**
6. Period of expenditure: **07/06/2018 to 31/03/2020**
7. Details of Expenditure:

S. No.	Grant Released by the UGC		Expenditure		Total Expenditure	Balance
	Budget Head	Amount	2018-2019	2019-2020		
1.	Minor Equipment	200000.00	82423.00	71186.00	153609.00	46391.00
2.	Consumables	50000.00	18519.00	18100.00	36619.00	13381.00
3.	Contingency	100000.00	44418.00	36900.00	81318.00	18682.00
4.	Field Work	100000.00	NIL	NIL	NIL	100000.00
5.	Travel	100000.00	17097.00	45785.00	62882.00	37118.00
6.	Books	150000.00	NIL	91146.00	91146.00	58854.00
7.	Software	100000.00	NIL	NIL	NIL	100000.00
	Total	**800000.00**	**162457.00**	**263117.00**	**425574.00**	**374426.00**

Signature of the Principal Investigator

Dr. Sunil Kumar
Principal Investigator
Department of Statistics
University of Jammu
Jammu

Registrar/Principal

(Seal)

Registrar
University of Jammu

15/9

15/09/2021

KRA & Co.

UNIVERSITY GRANTS COMMISSION
BAHADUR SHAH ZAFAR MARG
NEW DELHI - 110002

UTILIZATION CERTIFICATE

Certified that the grant of Rs 10,00,000/- (Ten lacs only) sanctioned to Dr Sunil Kumar, Assistant Professor, Department of Statistics, University of Jammu, by the University Grant Commission vide their letter No. **F.30-438/2018(BSR) dated 23/05/2018** towards UGC-BSR Research Project "Start-Up-Grant" entitled "Latent Class Analysis of Survey Data on Cheating Behaviour" for newly recruited faculty at Assistant Professors level in Science Department of various Universities. Out of the grant 8,00,000/- has been released by the UGC and Rs 4,25,574/- is utilized during the period 2018-2020 **for which it was sanctioned and in accordance with the terms and conditions as laid down by the commission**. The remaining unspent amount Rs **3,74,426/-** should be return back to UGC.

Dr Sunil Kumar
Principal Investigator
Department of Statistics
University of Jammu
Jammu

Signature of the
Principal Investigator

Registrar
University of Jammu

Registrar/Principal
(Seal)

15/9/2021

KRA & Co.
CA
FRN : 020256N

सत्यमेव जयते

विश्वविद्यालय अनुदान आयोग
University Grants Commission
मानव संसाधन विकास मंत्रालय, भारत सरकार
(Ministry of Human Resource Development, Govt. of India)
बहादुर शाह जफर मार्ग नई दिल्ली–110002
Bahadurshah Zafar Marg, New Delhi - 110002
Phone : 011-23604410, 011-23604425

ज्ञान-विज्ञान विमुक्तये

FD Diary No.1865
Dated : 23.05.2018

No.F.30-438/2018(BSR)

Dated: May, 2018

The Under Secretary FD-III Section,
University Grants Commission
Bahadur Shah Zafar Marg,
New Delhi – 110002.

- 7 JUN 2018

Subject:- **Approval-cum-Sanction letter for UGC-BSR Research Start-Up-Grant for newly recruited faculty at Assistant Professors level in Science Departments of various Universities – Release of the grant for the year 2018-2019 under Revenue.**

Sir,

The University Grants Commission convey its approval and allocate a sum of **₹10,00,000/- (Rupees Ten Lakhs Only) @ ₹10.00 Lakh to each Faculty** to the **Registrar, University of Jammu, 1st floor, New Administration Block, Jammu-180 006, (J&K)** being the UGC-BSR Research **Start-Up-Grant** for newly recruited faculty at Assistant Professors level of Science Departments.

Accordingly, I am further directed to convey the sanction of the University Grants Commission for payment of **₹8,00,000/- (Rupees Eight Lakhs only) (80% of the approved Grant ₹10.00 Lakhs to each Faculty)** to the **Registrar, University of Jammu, 1st floor, New Administration Block, Jammu-180 006, (J&K)** towards UGC-BSR Research Start-Up-Grant for newly recruited faculty at Assistant Professors level in Science Departments as per details given below, the expenditure to be incurred during 2018-2019.

Name of the Item	Head of Account	Name of Faculty / Professor	Name of Departments	Amount Approved (₹)	Amount being released (80%) (₹)
UGC-BSR Start-up grant for newly recruited faculty at Assistant Professor level in science department	3(A) 16(x) 31	Dr. Sunil Kumar	Stastics	10,00,000/-	8,00,000/-
			Total:	10,00,000/-	8,00,000/-

4. The sanctioned amount is debitable to the major Head **3(A) 16(x) 31** and is valid for payment during the financial year 2018-19 only.

3. The amount of the Grant shall be drawn by the Under Secretary (Drawing and Disbursing Officer) UGC on the Grants-in-aid bill and shall be disbursed to and credited to the **Registrar, University of Jammu, 1st floor, New Administration Block, Jammu-180 006, (J&K)** through Electronic mode as per the following details:

a.	**Details (Name & Address) of Account Holder**	:	**Registrar, University of Jammu, 1st floor, New Administration Block, Jammu-180 006, (J&K)**
b.	**Account No.**	:	**0345040160000001**
c.	**Name & Address of Bank Branch**	:	**Jammu & Kashmir Bank Ltd., New University Campus, Jammu-180 006, J&K**
d.	**MICR Code**	:	**180051018**
e.	**IFSC Code**	:	**JAKA0CANAAL**
f.	**Type of Account**	:	**Saving**

4. The Grant is Subject to the adjustment on the basis of Utilization Certificate in the prescribed proforma submitted by the University/Institution.

5. The University / Institution shall maintain proper accounts of the expenditure out of the Grants which shall be utilized only on the approved items of expenditure.

Latent Class Analysis of Survey data on Cheating Behaviour

Acknowledgements

First of all I would like to thank Almighty God Shri Mahakaleshwar Ji for showering his blessings on me and making my life more bountiful. May your name be exalted, honoured, and glorified!

The researcher gratefully acknowledges the following: his family, who through their love and example have been a source of strength and commitment; his teachers Prof J P S Joorel, Director, INFLIBNET, Gandhinagar, Gujarat, India, who motivates him for UGC start-up grant project and invaluable direction provided the inspiration and format for this study; Prof Diganta Mukherjee, Professor at ISI Kolkata, India, whose strict attention to detail and sound advice facilitated the completion of this work, and whose encouragement and support were unfailing and a source of motivation to succeed; Prof Pawan Kumar, Head, Department of Statistics, University of Jammu, Jammu for his immense help in planning and executing the works in time. On the other hand, profound knowledge and their valuable suggestions of Prof Rahul Gupta, Professor and Director HRDC, Dr Parmil Kumar, Associate Professor and Dr Vijay Shivgotra, Assistant Professor, Department of Statistics, University of Jammu,Jammu, came as a boon for me. The confidence and dynamism with which my elder brother Dr Sandeep Bhougal, Assistant Professor in SMVD University, Jammu supported the work requires no elaboration. I will be failing in my duty if I do not mention the staff and administrative staff of this department specially Mr Chaman Lal and the librarian Mr Jagdev Sharma for their timely help.

The researcher also acknowledges **UGC for providing funding** which enabled the researcher to make significant progress towards his research work.

I wish I would never forget the company I had from my research scholars of my department specially Miss Apurba Vishal Dabgotra and Miss Vishwantra Sharma for their help in every stage, it would have been impossible for me to finish this work.

My deepest gratitude goes to my family for their unflagging love and support throughout my life; this work is simply impossible without them. I would also like to gratefully acknowledge the support of a very special individual my lovely wife Monika, she is always available to chat with me and cheer me up. I appreciate my beautiful daughters Hitakshi and Avika for abiding my ignorance and the patience they showed during this work. Words would never say how grateful I am to all of you. I consider myself the luckiest in the world to have such a lovely and caring family, standing beside me with their love and unconditional support.

Sunil Kumar

Contents

List of Tables

Abstract

Over the years, academic cheating continues to be an endemic issue that has always been a threat to academic honesty and social values. Academic dishonesty among the students is a perplexing phenomenon, that exists in every stage of our education system, especially in colleges and universities. College and university administrators admit that academic dishonesty is an issue on campus but they often lack in preparing effective policies and procedures to monitor it and to deal with it. In addition, indecisive perceptions regarding academic dishonesty has adverse effects on paradoxical situation of education. The current study provides the details about the causes that motivate the students to cheat and describes the different forms of cheating practices performed by the students.

The purpose of this study is to determine the insight of the students of the University of Jammu on cheating, to determine the proportion of students who honestly admit that they cheats in their academics life, to determine the proportion of students who cheat but did not admit it and to determine the factors associated with cheating behaviour.

This study deals with cheating behaviour of students at college and university levels. Thus, data from the sample of 2317 students enrolled in second year of graduation, post graduates students of the University of Jammu and its affiliated colleges, were obtained from anonymous questionnaire comprises of 46 closed ended questions related to factors leading to students' cheating, types of cheating and other related questions.

From the current study, the main reasons due to which students go for cheating are not knowing/understanding the study material, performance pressure, inadequate exam preparations, etc. Also, around 42%, 6% and 52% of the respondents are found to be occasional, persistent and instantaneous cheaters, respectively. We suggest that students must understand that cheating is wrong not only for the society but also for their own knowledge because by indulging in cheating students prevent them from learning what they are studying and hence, deteriorate the intellectual human resource of the country.

Chapter 1

Introduction

1.1 Background

Education is a very important investment for the future. It allows one to attain a future which one cannot grab otherwise. So, it is necessary to give a child good education for him/her to survive and prosper. But, the real problem comes when we realize that the child is not capable to study well or more precisely we do not know in advance in which field he/she could execute well. Parents should not force the child to choose a career in which he/she has no interest and talent. If he/she is compelled to take a path in which he/she is not capable, he/she may be forced to resort to ways that dodge the problem, like resorting to cheating?

The phenomenon of cheating is the most prominent and serious practice which exits at every level of education system. Academic cheating can be of various types, suggested by Holleque (1982) and k12 academics, such as copying from another's person, stealing examination papers and lecture notes, using prohibited material like crib notes, deception, sabotaging, impersonation, forgery, plagiarism, fabrication, data manipulation, padding bibliographies, and many more. So, however and whenever it happens, it creates a serious problem, not only for faculty and institutions of higher education, but also for students.

Academic dishonesty is an escalating phenomenon that is plaguing educational institutions around the world. It has been and continues to be a major problem in schools, colleges and universities. Infact, it has become an inescapable activity, especially, in colleges and universities, where grades earned directly effect the academic careers of students for many years to come. The rising pressure to get the best grades in school, get into the best college, and land the best paying job is a cycle that has made academic dishonesty increase exponentially. The subject of academic cheating has attracted the attention of not only academics but also public communities. Recent studies have proven that the issues of academic cheating among undergraduates have increased along the years. For example, 76 percent of the students confessed to having involved in academic cheating (Jeergal et al., 2015).

Academic cheating is becoming so common these days that it sometimes becomes necessary for some persons to cheat in order to get better ranks. Hence, it is quite important to identify the

causes that motivate students to cheat. Various studies have been conducted to identify the causes of cheating and the ways for controlling it: see Rettinger & Kramer (2009), Munir et al. (2011), Griffin et. al (2015), Curtis & Vardanega (2016), and Hussein et al. (2018). There may be several reasons that might motivate students to get involved in cheating practices. The one major reason for students' cheating is that they may not have studied well or they may not understand the material well as a result of which they are unable to write an exam or to complete the assignment by their own, so they resort to cheating. Another factors encouraging students to cheat includes tough question papers, laziness, peer influence, parental pressure, performance pressure, lack of time management skills, etc. Whatever be the type of cheating it is serious threat to the academic integrity of our education system.

1.2 Content of This Document

This report contains three chapters and three appendices. The present chapter i.e., Chapter One gives an introduction to the report, and Chapter Two describes the basic concepts, definitions and methodology used for the analysis of the data used. Chapter 3 contains the summary of the main findings on the cheating behaviour of the students obtained from the survey data. Appendix A gives the detailed description of the sample size selected from considered sampling frame, appendix B provides a detailed questionnaire used for the data collection of the present study and appendix C provides a detailed description of variables used in analysis.

The current chapter provides a detailed background of the study along with the literature that have reviewed for this study. It also contains the objectives of the current study. In addition to the above this chapter provide detailed description of the sample design and the data collection.

Chapter two contains the basic concepts including the definitions of academics cheating , types of academics cheating & factors influencing students to cheat in their academics life and a detailed explanation of the methodology used for the analysis of the collected data. It also provide a brief description of the data collected for this study.

And, finally chapter three comprises of all the results and findings of the study along with interpretations and conclusions followed by the references.

1.3 Literature Review

Academic cheating is an immoral way of achieving a goal in the field of academic (Kalhori, 2014). In the context of this study, academic cheating is generally composed of two different forms, cheating in tests and cheating in assignments. Test cheating is defined as an act of deception using forbidden items and information during examinations in order to gain unfair advantage over others (Muchai, 2014). Meanwhile, assignment cheating is an act that deceives, misleads or fools the lecturer into thinking that the assignment submitted by the student was a student's own work (Davis et al., 2009).

Various theoretical perspectives have been applied for understanding cheating. Michaels and Miethe (1989) examined the applicability of several theories of divergence to cheating, including

deterrence theory, rational choice theory, social bond theory, etc. Thereafter, Beck and Ajzen (1991) proposed the integrated model of cheating, which shows that students' intentions to engage in disfunctional behaviours may be influenced by attitudes, subjective norms, perceived behavioural control and moral obligations. Previous studies conducted on academic cheating have looked into various perspectives. Some studies focused on the relationship between motivation and cheating behavior, some focused more on the relationship between socio-demographic factors and cheating behavior (e.g., Jung-In et al. (2015)). Others (e.g., Donse and Van De Goep (2013)) have carried out studies to determine what causes students to cheat.

The existence of academic cheating has always been a major concern for various researchers. Many studies have been conducted in order to identify various types of academics cheating among college students and this can be achieved by using anonymous questionnaire distributed or mailed to the students. The estimates of students that are involved in academic cheating during their college lives, ranges from 49% for marketing students (Tom & Borin, 1988) to 88% for premedical students (Sierles et al., 1980).

Dishonesty in an academic setting has been a consistent and paramount problem for many years at all educational levels (Harding et al.2004), and it is a serious educational issue(Orosz et al. 2016; Koul et al. 2009). Considerable progress has thus, been made in identifying factors that influence cheating behaviour. Alarape & Onkoya (2003) have found that age and self esteem are the two major factors that are positively correlated with the cheating behaviour of students. Khodaie et al. (2011) found that higher socio-economic status along with stress, depression and family crisis are the major factors that promote students to cheat in their academic lives.

Technological factors such as the use of internet for assignment completion has increased the possibility of academic misconduct (Hosny & Shameem, 2014). In fact, the growth of technology through social media has provided electronic storage opportunities as well as platforms for sharing and archiving exams or answers for anyone who searches for it (Smith et al., 2007). Another study discovered that students actively cheat on assignments rather than on examinations (King & Case, 2014). Additionally, the internet which has been used by students in their personal lives has helped them to get the information needed for assignments (Anitsal et al., 2009). Hutton (2006) also stated that the easy access of the internet has increased the opportunity for students to share work in unethical ways.

Various guidelines for controlling cheating have also been proposed like Houston (1976, 1983) & Aiken (1991) suggested that cheating can be controlled by making it as difficult as possible for students to cheat, Jenderk (1992) suggested that the cheating can be avoided by discussing the consequences of cheating with the students, etc. Despite these findings and recommendations, the prevalence of cheating is on the rise (Daniel et al., 1991; Davis et al., 1992). Marsh and Campion (2018) suggest that academic integrity should be strongly assumed as an institutional concern, instead of just students' responsibility. Moreover, through a collaboration approach and using workshops and open educational resources settled to address paraphrasing, summarizing and quotation, Marsh and Campion (2018) concluded that "better collaboration and co-operation among faculty staff, learning advisors and librarians is therefore essential" (Marsh and Campion 2018). Still there are many questions that remain to be answered concerning the nature, cause and type of academic cheating.

1.4 Objectives

This study was conducted to determine the intuition of the students attaining at least second year of graduation, post graduates of the University of Jammu, Jammu, J&K, India, on cheating. Following are the major objectives of the study:

1. To define students, what is meant by cheating;
2. To find out the main reasons, why student opt cheating;
3. To identify whether cheating is a product of laziness or some other circumstances;
4. To determine the latest ways used by the students for cheating;
5. Do the extreme response patterns significantly alter the information on cheating behavior? In other words is the response representative?

1.5 Sample Design

Our target population comprised of all the undergraduates students enrolled in 3rd semester and above, BEd students, masters students of University of Jammu and its affiliated Colleges. So, the total population size in our study was 66091. We had followed the NSSO methodology to select the sample with 95% confidence level, 2% margin of error and 50% of population share, using formulae:

$$n = \frac{\frac{z^2 \times p(1-p)}{e^2}}{1 + \frac{z^2 \times p(1-p)}{e^2 N}}$$

$$\begin{aligned} where, n &= sample\ size \\ z &= z\ score \\ p &= population\ proportion \\ e &= margin\ of\ error \\ N &= size\ of\ populationsize \end{aligned}$$

Thus, a sample of 2317 had selected for the collection of data. And we had selected sample based on the multi-stage stratified random sampling design. The first stage units (FSU) comprises of the main campus, off campuses and different colleges affiliated to the Jammu University. The ultimate stage units (USU) comprises of different courses offered by these institutes. Then, we apply proportional allocation for selecting sufficient number of sample units from each of the USU.

Table 1.1 provide details of sample selected from the different institutes. Detailed description of the sample selected from each of the sampling unit is given in Appendix A.

Table 1.1: Sample size selected from each of the institute

	Institute	Sample size
Off Campuses	Kathua Campus	58
	Reasi Campus	1
	Udhampur Campus	3
	Poonch Campus	1
	Kistwar Campus	2
	Ramnagar Campus	1
	Bhaderwah Campus	4
Affiliated Colleges	Govt. Degree College, Doda	117
	Govt. Degree College, Rajouri	250
	Govt. Degree College, Poonch	105
	Govt. Degree College, Kishtwar	41
	Govt. Degree College, Kathua	271
	Govt. Degree College,Samba	98
	Govt. Degree College, Ramban	39
	Govt. Degree College, Udhampur	237
	Govt. Degree College, Reasi	42
	Govt. Degree College, RS Pura	34
	Govt. Degree College, Akhnoor	32
	Govt. Degree College, Bishnah	10
	Govt. Degree College, Paloura	31
	Govt. College of Engineering	25
	Model Institute of Engineering and Technology	35
	MBS College of Engineering	27
	Private law colleges	18
	Private BEd colleges	50
On Campus	University of Jammu	785
Total		2317

1.5.1 Sampling Frame

For the sampling frame, we had collected information from the department of Statistical Planning and Research Unit of University of Jammu. But we had considered all the 7 Off campuses of University of Jammu, 9 government colleges from each district other than Jammu district, 4 government colleges from the Jammu district, 2 private law colleges of Jammu district, 16 private BEd colleges of Jammu district, 1 government engineering college, 2 private engineering colleges of Jammu district and finally, the campus of University of Jammu was considered as the sampling frame.

Stratification

Stratum had formed at district level. Within each district of Jammu division, nine basic strata were formed (one for each district). However, within the Jammu district, different government degree colleges, government engineering college, private engineering colleges, private BEd colleges and

private law colleges formed a separate basic stratum and the remaining off campuses of University of Jammu i.e., Ramnagar campus and Bhaderwah campus and campus of University of Jammu, itself was considered as another basic stratum.

Sub-Stratification

Different sub-stratification were done for the courses offered by each of the selected campuses and colleges. Seven (7) different streams were considered as the courses offered by these institute. They were: Science, Technology, Commerce, Arts, Management, Law and BEd.

Allocation used in Sub Strata

Within each stream(course) offered by University of Jammu, its Affiliated colleges and its different off campuses, the respective sample size was allocated to the different strata in proportion to the number of students enrolled in that stream.

The above discussed technique has employed on each of the district except for Jammu District. For Jammu District, 75% of the remaining sample, after selected from the other districts and off campuses, had selected from the campus of the University of Jammu, using proportional allocation. And the rest of sample had selected from the remaining colleges of Jammu district using proportional allocation

Selection of USUs

For the different districts, from each sub stratum, required number of sample were selected by Simple Random Sampling Without Replacement (SRSWOR) procedure.

1.6 Data Collection

A descriptive design is used to study cheating behaviour of the students. So, a questionnaire was designed to analyze the cheating behaviour of the students. The survey period of this study was from March to September 2019. The required information was collected from a selected number of students by visiting their respective departments, colleges or institutes, personally and by asking them to fill in an anonymous questionnaire based on different cheating behaviour. But due the disturbances in the state during the survey period, it was not possible to collect the data from the highly disturbed areas such as Doda, Kishtwar and Rajouri. So, the data was collected from only 1906 students which was the relevant sample size for our study.

Broadly, the following information was collected in this study from each selected student in order to understand their general perception about the cheating and to analyze their cheating behaviour:

1. the demographics characteristics of the students i.e., age, gender, stream, status etc.,
2. general questions on students' cheating,
3. factors influencing students to cheat including general as well as personal factors,
4. types of cheating students are involved in.

Visited Bahderwah Campus, University of Jammu to collect survey data

Visited Kathua Campus, University of Jammu to collect Survey Data

Chapter 2

Concepts And Methodology

2.1 Basic Concepts

2.1.1 Academics Cheating

The phenomenon of cheating is the most prominent and serious practice which exits at every level of education system. Rajendran defined cheating as an activity which is performed to complete a work in an unethical way by a person when he does not know how to do that work in a legal way. However, this is not the only way of defining cheating; different people have their different perception about cheating like Holleque (1982) defined cheating as changing one or more answers when correcting own examination and/or not marking two or more in correct responses, Mccabe (1999) defined cheating as an immoral activity in the academic environment, Romney & Steinbart (2003) defined cheating as an act of using any means of unfair and unjust privileges that include: lying, concealing the truth, deceive, deceit and violation of trust to achieve something. But academic cheating/academic dishonesty may be defined as the students' behaviour of contradicting the fundamental values of their academic lives (Holleque, 1982).

2.1.2 Types of Academics Cheating

Following is the list of all possible types of academics cheating in which students are involved intentionally or unintentionally:

1. **Bribery :** It is an act of giving money or gift that alters the behaviour of recipient in the favour of provider.
2. Using or having material like crib notes, gadgets etc. in an exam that is not specifically approved by the instructor.
3. **Deception :** It is an act of providing false information to a teacher concerning a formal academics exercise.
4. **Fabrication :** It is an act of falsification of data, information or citation.
5. **Impersonation :** It is a form of cheating whereby a different person other than the students, complete an exam or assignment assigned to him.

6. **Plagiarism :** It is an act of using the language and thoughts of another author and representation of them as one's own.
7. **Sabotage :** It is a form of cheating in which student prevents other student from completing their work.
8. **Copying :** It is an of act of writing an examination/ test/ assignment from someone else. It includes looking over someone's shoulder, forbidden sharing of information between students etc.
9. Attempting to obtain or accepting assistance from any other person (student or professor) during exam.
10. Unauthorized collaboration i.e., collaborating on an assignment unless specifically allowed by the professor.
11. Obtaining a copy of an examination or test prior to it.
12. Submitting any work for academic credit as their own when one is not the sole author or creator.
13. Submitting any work that has been previously accepted for academic credit.
14. **Stealing :** It is an act of illicitly obtaining materials needed to complete assignment or test.
15. Misrepresenting a family or personal situation to get an extension in the course.
16. **Forgery :** It is act of forging a faculty/ family member's signature on a permission form or add/ drop form or any other form.

2.1.3 Factors Influencing Academics Cheating

Following is the list of factors/ circumstances that can possible influence the students' decision to cheat in the exams or assignments:

- Student has outside job
- Type of exam/ level of test difficulty
- Instructor friendliness
- Effect of course grades on long-term goals
- Instructor view on cheating
- Family view on cheating
- Financial support for good grades
- Spacing of students in the exam room
- course workload/ Academic overload

- Sitting by the side of good students/ friend in the exam hall
- Type of course
- Friends' view on cheating
- Family view on grades
- Value of the course material
- Instructor vigilance
- Excelling in exam/ To have good grades/ Performance pressure
- Fear of failure
- Ranks required for further admissions
- Other classmates also cheat/ Peer influence
- Stiff competition/ Excessive competition
- Lack of facilities
- Anxiety (exam phobia)/ Stress
- Ineffective time management skills or overload
- Exams focus on memory rather than comprehension
- Lack of self-confidence that they can pass
- Inadequate preparations
- Like cheating and believe everyone cheats
- Unfairness of instructor/ Strict instructor
- Laziness
- No realization that they are cheating and consider it as little help.

2.2 Reliability and Validity Analysis

Generally, the questionnaires are designed with the objective of obtaining relevant information in most valid and reliable manner, for any study. Thus, it becomes necessary to test the consistency and accuracy of a questionnaire, by using different reliability and validity tests. Reliability may be defined as the degree to which an assessment tool produces stable and consistent results (Carmines and Zeller, 1979). Reliability testing is important as it indicates the consistency across the different parts of a research instrument i.e. questionnaire (Huck, 2007). Whereas, validity may be refers to how accurate an instrument is at measuring, what is intended to be measured (Field, 2005). Although reliability is necessary, it is not sufficient alone. In other words, for a research instrument to be reliable, it also needs to be valid (Wilson, 2010).

The thumb rules of reliability and validity are as follows:

- Internal consistency reliability can be tested by using Cronbach Alpha coefficient. Minimum value of 0.70 is recommended for internal consistency reliability (Whitley, 2002, Robinson, 2009).
- Face validity can be tested using Cohen's Kappa Index (CKI). Minimally acceptable value Kappa index is 0.60.
- Construct validity (discriminant and convergent validity) can be determined using factor analysis (Koh and Nam, 2005, Wee and Quazi, 2005). The loading of at least 0.40 & no cross loading of items above 0.40 are required for establishing discriminant validity and eigen values of 1 & loading of at least 0.40 is essential for convergent validity (Straub et al., 2004)
- Criterion validity is tested using correlation analysis. If significant value < 0.05, then the instrument is declared to be valid.

Reliability, construct validity and criterion validity tests are performed using SPSS software whereas face validity tests are carried out 'irr' package (Gamer et al., 2012) of R software.

2.3 Latent Class Analysis

There are various phenomenon that often cannot be directly observed or to analyze certain phenomenon, not all variables of interest may exist in practice nor are able to be directly measured. So, latent variable modeling, in which the value of the latent variable (unknown variable) cannot be directly observed, rather its value is deduced from observed variables, can be used in those cases. Latent variables may be defined as an unobserved random variables which are hidden from us (SKrondal, 2004) and are unknown to us in any particular study, whereas, manifest variables may defined as variables which are the observable and are designed specially to measure the unknown latent variable. Moreover, the manifest variables are known to us in that particular study. Out of all possible manifest variables only few are capable to be indicators of the latent variable. Indicator variables are the manifest variables which can measure the unknown latent variable. The value of unknown latent variable can be estimated on the basis of the responses made by the individuals to the different indicator variables. We also have a third variable, referred to as a grouping variable such as gender (G), which is used to identify an individual's membership in two or more population subgroups. however, there is no statistically rigorous way to test whether a particular choice of a grouping variable satisfies the assumptions of LC models. Instead, the grouping variable is based largely on subjective criteria, by considering whether the assumptions are plausible for a it or not.

Latent Class Analysis (LCA) is appropriate when the latent variable and all the indicator variables are discrete in nature. It establishes a relationship between a set of observed discrete variables (manifest variables) and a set of unknown discrete variables (latent variables). LCA is a methodology that allows us to identify hidden population subgroups/classes. A class of any latent variable is specified by pattern of responses made to the different manifest variables by the respondents, in terms of conditional probabilities. These probabilities show the possibility that the latent variables can take any particular values. LCA is a valuable tool for assessing measurement error, identifying flawed items of the questionnaire, assessing mode extreme responses effects and adjusting

non response bias. LCA involves both absolute fit of a particular model and relative fit of two or more competing models. It gives different types of model statistics measures like AIC (Akaike information criterion), BIC (Bayesian information criterion), chi- square, classification error, and many more, which makes us able to find right number of classes and we can try different options before choosing the optimum one. Moreover, the Expectation- Maximization (EM) algorithm used in latent class analysis for finding the latent classes is highly robust.

The basic analytical approach used in this study is the LCA for dichotomous manifest variables. In this approach we envisage a relationship between discrete indicator variables and discrete latent variables across different number of groups. We have used poLCA (Linzer and Lewis, 2011) package of R software for performing the Latent class analysis (LCA).

2.3.1 Latent Class Models

LCA models comprises of two types of probabilities which include

- the probability indicating the likelihood of a response by respondents in each of the classes and
- the probability representing the latent class size or the proportion of individuals who are members of a particular latent class.

Former one represents the probability of a particular responses to a manifest variable, conditioned on latent class membership and can be interpreted as factor loading for Factor Analysis in which both the observed or latent variables are continuous. LCA provides a clustering of individuals in a population, based on the response patterns of individuals to the different observed variables.

- Sample can be treated as if it were a simple random sample without replacement from an infinite population i.e. data is sampled without replacement from a large population units using SRS.
- The indicators are locally independent within a latent class, means all the indicator variables have nothing in common except latent variable, i.e., after accounting for latent variable X , there is no association between indicator variables.
- The response probabilities are homogeneous, i.e., the probability of selection of any two units or individuals from the population are same.
- The indicator variables are univocal, i.e., the indicator variables can measure one and only one latent variable.

Following the notation used by Linzer and Lewis (2011), suppose we have J polytomous categorical manifest variables (the observed variable) each of which contain K_j possible outcomes, for individuals $i = 1, 2, 3, ..., N$. Let Y_{ijk} be the observed values of the J manifest variables such that

$$\left\{ \begin{array}{ll} Y_{ijk} = 1: & \text{if } i^{th} \text{ respondent give the } k^{th} \text{ response to the } j^{th} \text{ variable} \\ Y_{ijk} = 0: & \text{otherwise} \end{array} \right\}$$

where j=1,2,...,J and k=1,2,...,K_j.

The LC models approximates the observed joint distribution of the manifest variables as the weighted sum of a finite number, R, of constituent cross-classification tables. Let π_{jrk} denote the cross-conditional probability that an observation in class r=1,2,...,R produces the k^{th} outcome on the j^{th} variable with

$$\sum_{k=1}^{K_j} \pi_{jrk} = 1$$

. Let p_r be the prior probabilities of latent class membership, as they represent the unconditional probability that an individual will belong to each class before taking into account the responses Y_{ijk} provided on the manifest variables. The probability that an individual i in class r produces a particular set of J outcomes on the manifest variables, assuming conditional independence of the outcomes Y given class membership, is the product

$$f(Y_i; \pi_r) = \prod_{j=1}^{J} \prod_{k=1}^{K_j} (\pi_{jrk})^{Y_{ijk}}, \tag{2.1}$$

The probability density function across all classes is the weighted sum

$$f(Y_i|\pi, p) = \sum_{r=1}^{R} f(Y_i; \pi_r) = \sum_{r=1}^{R} P_r \prod_{j=1}^{J} \prod_{k=1}^{K_j} (pi_{jrk})^{Y_{ijk}}, \tag{2.2}$$

The parameters P_r and π_{jrk} are estimated by the latent class model.

Given estimates $\hat{P}_r$ and $\hat{\pi}_{jrk}$ of P_r and π_{jrk} respectively, the posterior probability that each individual belongs to each class, conditional on the observed values of the manifest variables, are calculated by

$$\hat{P}(r_i|Y_i) = \frac{\hat{p}_r f(Y_i; \hat{\pi}_r)}{\sum_{q=1}^{R} \hat{p}_q f(Y_i; \hat{\pi}_q)'} \tag{2.3}$$

where $r_i \in (1, 2, \ldots, R)$.

It is important that the condition $R\sum_j (K_j - 1) + (R - 1) \leq n$ on the number of parameters should hold. Also, $R\sum_j (K_j - 1) + (R - 1) \leq (3^{10} - 1)$ i.e. one fewer than the total number of cells in the cross-classification table of the manifest variables, as then the latent class model will be unidentified. Under the assumptions of multinomial distribution, the log likelihood function can be given as:

$$lnL = \sum_{i=1}^{n} ln \sum_{r=1}^{R} p_r \prod_{j=1}^{J} \prod_{k=1}^{K_j} (\pi_{jrk})^{Y_{ijk}}, \tag{2.4}$$

LCA not only builds a classification model but it also explain a relation of the class membership to explanatory variables by including covariates (Vermunt,2010) in the model. Grouping variables

can be used in LC models in order to model the unexplained heterogeneity in the data. In that case latent class membership probabilities are predicted by covariates through a logistic link.

2.3.2 Parameter Estimation

In LCA, observed data corresponds to the observed responses of individuals on each item which are considered as categorical observed variables and the missing data are the unobserved scores in a latent categorical variable whose categories are called latent classes. Dempster et al.(1977) provided maximum likelihood estimation in case of observed and missing data involved in the analysis.

The unknown parameters of the LC models can be estimated by maximizing (2.4) with respect to p_r and π_{jrk},using the expectation-maximization (EM) algorithm (Dempster et al.(1977), McLachlan and Peel (2000) and Linzer and Lewis (2011)). The EM algorithm, begin with arbitrary initial values of $\hat{p}_r$ and $\hat{\pi}_{jrk}$, and denote them $\hat{p}_r^0$ and $\hat{\pi}_{jrk}^0$. The expectation step, calculate the missing class membership probabilities using equation (8), substituting $\hat{p}_r^0$ and $\hat{\pi}_{jrk}^0$ in place of $\hat{p}_r$ and $\hat{\pi}_{jrk}$. The maximization step, update the estimates of the parameters by maximizing the log likelihood function given these posterior $\hat{P}(r_i|Y_i)$, with

$$p_r^{new} = \frac{1}{N}\sum_{i=1}^{N} P(r_i|Y_i)$$

as the new prior probability and

$$\hat{\pi}_{jr}^{new} = \frac{\sum_{i=1}^{N} Y_{ij} P(r_i|Y_i)}{\sum_{i=1}^{N} P(r_i|Y_i)}$$

as the new class conditional outcome probabilities; $\hat{\pi}_{jr}^{new}$ is the vector of length K_j of class-r conditional outcome probabilities for the j^{th} manifest variable; and Y_{ij} is the $N \times K_j$ matrix of observed outcome Y_{ijk} on that variable. The algorithm repeats these steps several times until the overall log-likelihood reaches a local maximum and further increments are less than some arbitrarily small value.

2.3.3 Model Selection

Different LCA models have different number of latent classes. Usually, models with more parameters(i.e, more latent classes) provide a better fit, and more parsimonious models tend to have a somewhat poorer fit. So, there is always very close agreement between goodness of fit and parsimony of the latent class models. We can test the goodness of fit of an estimated LCA models by the Pearson Chi-square(χ^2) or the Likelihood Ration Chi-square(L^2) . However, the likelihood ratio Chi-square test, although extensively used in statistical literature, has a number of important limitations.These limitations can be controlled by making use of several information criteria, such as the Akaike information criterion (AIC) (Akaike (1973)) and Bayesian information criterion (BIC) (Schwartz (1978)), each of which is designed to penalize models with larger numbers of parameters. AIC and BIC on the number of parameters in the model:

$$AIC = L^2 - 2 \times d.f.$$

and

$$BIC = L^2 - d.f. \times ln(n),$$

where n is the sample size.

These information criterion are commonly used for selecting the optimal number of latent classes in a model . By comparing models with different number of latent classes, a model with lower AIC and BIC is selected.

2.4 Data Description

This study had used survey method to obtain the needed data and to determine the perception of students regarding the cheating problem. A questionnaire was prepared which comprises of 46 closed ended questions related to factors leading to students' cheating, types of cheating and other related questions. The whole questionnaire is divided into following 5 major sections.

Respondent's Details : This section tracks the record of the demographic characteristics of the students which includes gender, age, religion, status and stream.

General Questions : This section consist of general questions about cheating in order to get the general perception of students on academics cheating and to get idea from the students about the main reason why students cheat during the exams?

Situational Aspects : This section consist of questions about the situational factors that can influence the students to cheat during the exams.

Personal Aspects : This section consist of the questions about the personal factors that can influence the students to cheat during the exams.

Types Of Cheating : This section comprises of the questions about the types of cheating students are involved in. This section was prepared with the objective of making students aware of the types of academics cheating and to know about their own cheating behaviour.

A detailed Questionnaire is in Appendix B.

2.5 Statistical Hypothesis

The following null hypothesis were considered and tested, at the 5% level of significance, in the current study :

$H_{\circ 1}$ **:** Cheating is independent of gender.

$H_{\circ 2}$ **:** Cheating is independent of current Status of the student.

$H_{\circ 3}$ **:** Cheating is independent of stream of the student.

$H_{\circ 4}$ **:** Cheating is independent of students' perception of academic cheating.

$H_{\circ 5}$ **:** Cheating is independent of students' understanding of material.

H_{o6} **:** Cheating is independent of type of exam given by the teacher to the students.

H_{o7} **:** Cheating is independent of performance pressure on the students.

H_{o8} **:** Cheating is independent of the students' perception about humiliation due to failure.

H_{o9} **:** Cheating is independent of the students' confidence lacking.

H_{o10} **:** Cheating is independent of the students' parental pressure.

H_{o11} **:** Cheating is independent of the students' attitude towards grades.

H_{o12} **:** Cheating is independent of sitting plan of the students in the exam.

H_{o13} **:** Cheating is independent of punishment severity in the educational atmosphere.

H_{o14} **:** Cheating is independent of students' inadequate exam preparations.

H_{o15} **:** Cheating is independent of students' ineffective time management skills.

H_{o16} **:** Cheating is independent of students' habit of laziness.

H_{o17} **:** Cheating is independent of students' perception about instructor vigilance.

H_{o18} **:** Cheating is negatively correlated to the students' subject liking.

H_{o19} **:** Cheating is negatively correlated to the students' interest in the subject.

H_{o20} **:** Cheating is negatively correlated to the students' instructor liking.

H_{o21} **:** Cheating is independent of the student's perception that cheating inevitable.

H_{o22} **:** Cheating is independent of the student's perception of peer influence.

Chapter 3

Summary of Findings

3.1 Introduction

This study had conducted in order to make students aware of exactly what is meant by academics cheating and to find out the factors or circumstances under which students are more likely to cheat. This study also categorizes the students according to their response patterns to different manifest variables using LCA.

The current chapter summarizes all the results of the survey and discusses the main findings of the survey about the cheating behaviour of the students. The discussion in this chapter begins with the reliability and validity analysis which forms initial stage of the analysis procedure. Then, it explains the way of treating missing data values followed by the descriptive statistics.

This chapter also, explains the hypothesis testing and LCA results which includes LC models, model selection, selection of optimum number of latent classes along with its interpretation. Finally, it is concluded with the conclusions of the study.

3.2 Reliability and Validity Analysis

We had conducted a pilot survey on 9 departments of the University of Jammu, in order to test the reliability and validity of our questionnaire, before conducting the full survey. The Value of Cronbach Alpha coefficient in this case comes out to be 0.735, which indicates a high reliability according to Hinton et al. (2004).

Cohen's Kappa index, in this case was 0.61, which shows that face validity holds for our questionnaire. Although, face validity holds in this study but it is the weakest form of validity and many would suggest that it is not a form of validity in the right sense of the word. So, we had employed other validity tests also, as it is always advisable to test different forms of validity in order to get valid data. We had performed factor analysis utilizing principle component analysis with varimax rotation method for testing construct validity and correlation analysis for testing the criterion validity. Both the validities (criterion and construct) holds in this study.

3.3 Dealing With Missing Data

Our data set consists of 1906 observations, but out of that 7 were deleted due to non response and a total of 1899 responses were considered for the further analysis. These responses were, then tested for the identification of missing data values. The term missing data is, commonly, refers to the absence of one or more values within a study variable(s) contained in the dataset. In general, a greater number of missing values within a dataset reflects a greater challenge to the data analyst. So, it become quite essential to assess and treat missing values. The keys to effectively assessing and treating missing data values within a dataset involve specifying how missing data will be defined in a study, assessing the amount of missing data, identifying the pattern of the missing data and selecting the best way to treat the missing data values.

We had used multiple imputation method to deal with missing data values. With singular imputation methods, mean, median, or some other statistics is used to impute the missing values. However, using single values carries with it a level of uncertainty about which values to impute. Multiple imputation narrows uncertainty about missing values by calculating several different options. We had used SPSS software for performing multiple imputations on over dataset.

3.4 Descriptive Statistics

The collected data were organized into frequency and contingency table. Frequency distribution tables were used to indicate the proportion of those students who admitted that they had cheated in their academic lives through the general questions and the direct questions. Contingency table was utilized in comparing cheating behaviors of the students with their general perception about cheating.

Table 3.1 provides the frequency table of the respondents demographic characteristics. From the table 3.1 it is found that the proportion of female in the sample is 0.60 and male is 0.40; whereas in population this proportion was 0.66 and 0.34.

Table 3.1: Frequency table for demographic characteristics of the respondents

Variables	**Categories**								**Total**
	1	2	3	4	5	6	7	8	
Gender	768	1131	-	-	-	-	-	-	1899
Age	744	653	402	100	-	-	-	-	1899
Religion	1533	231	402	100	-	-	-	-	1899
Current Status	1158	122	619		-	-	-	-	1899
Stream	451	-	146	57	862	57	102	224	1899

Also, the proportion of undergraduates, graduates, postgraduates in the sample is 0.61, 0.06 and 0.33, respectively. It is also found that the proportion of the students belonging to science, Technology, Commerce, Arts, Management, Law and other streams is 0.24, 0.08, 0.03, 0.45, 0.03, 0.05 and 0.12, respectively.

Table 3.2: Frequency table for the questionnaire

S.No.	Question No.	Sub Questions	Categories						Total
			1	2	3	4	5	6	
1	Q1		1723	176	-	-	-	-	1899
2	Q2		596	717	481	105	-	-	1899
3	Q3		781	601	432	85	-	-	1899
4	Q4		386	1069	64	223	157	-	1899
5	Q5		300	220	369	129	758	123	1899
6	Q6		1576	323	-	-	-	-	1899
		a	154	130	205	1087	-	-	1576
7	Q7		201	140	856	146	556	-	1899
8	Q8		1513	386	-	-	-	-	1899
		a	1249	264	-	-	-	-	1513
		b	1071	442	-	-	-	-	1513
		c	1116	397	-	-	-	-	1513
		d	1163	350	-	-	-	-	1513
9	Q9		1227	672	-	-	-	-	1899
10	Q10		1017	882	-	-	-	-	1899
11	Q11		1562	335	-	-	-	-	1899
12	Q12		1359	540		-	-	-	1899
13	Q13		1410	489	-	-	-	-	1899
14	Q14		1177	722	-	-	-	-	1899
15	a		1653	246	-	-	-	-	1899
16	b		1655	244	-	-	-	-	1899
17	c		1008	891	-	-	-	-	1899
18	d		1083	816	-	-	-	-	1899
19	e		1059	840	-	-	-	-	1899
20	Q15		1393	506	-	-	-	-	1899
		a	357	1036	-	-	-	-	1393
		b	1113	280	-	-	-	-	1393
		c	733	660	-	-	-	-	1393
		d	539	854	-	-	-	-	1393
		e	269	1124	-	-	-	-	1393
		f	407	986	-	-	-	-	1393
		g	373	1020	-	-	-	-	1393
21	Q16		1097	802	-	-	-	-	1899
		a	827	270	-	-	-	-	1097
		c	642	455	-	-	-	-	1097
		d	386	711	-	-	-	-	1097
		e	417	680	-	-	-	-	1097
22	Q17		590	1309	-	-	-	-	1899
23	Q18		673	1226	-	-	-	-	1899

S.No.	Question No.	Sub Questions	Categories						Total
			1	2	3	4	5	6	
24	Q19		367	260	821	345	106	-	1899

Table 3.2 provides the frequency distribution for the questionnaire. The following are the main findings from the table 3.2:

- 91% of the students believed that students cheat during their academic lives.
- Only 73% of the students have admitted that they have actually cheated in the tests and exams.
- Around 58% of the students have cheated in assignments or projects.
- On average 33% of the students are involved in other forms of cheating.
- Around 83% of students believed that inadequate preparation is the major factor that enforces a student to go for cheating.

Table 3.3: Contingency table

		Q15		Total
		1	**2**	
Q1	**1**	1280	443	1723
	2	113	63	176
Total		1393	506	1899

Table 3.3 provides the contingency table which compares students cheating perception in general with their own cheating behaviour. Table 3.3 shows that 23% of the respondents believe that students cheat during their academic lives but have denied that they have cheated in the exams or tests; whereas around 6% of the students have actually cheated in exams or tests but they don't believe, in general, that students cheat in their academic lives

3.5 Testing of Hypothesis

Table 3.4 provides the summary of the hypothesis testing. From table 3.4 it is clear that the cheating behaviour of the students is independent of the their current-status, type of exam given by the teacher, perception about humiliation due to failure, confidence lacking, parental pressure, attitude towards grades, punishment severity in the educational atmosphere, ineffective time management skills, habit of laziness and perception about instructor vigilance.

Table 3.4: Hypothesis test summary

S.No.	Hypothesis	χ^2	Significant value	Decision
1	H_{o1}	17.574	0.000	Reject the null hypothesis
2	H_{o2}	5.504	0.138	Accept the null hypothesis
3	H_{o3}	17.795	0.013	Reject the null hypothesis
4	H_{o4}	8.309	0.004	Reject the null hypothesis
5	H_{o5}	20.702	0.000	Reject the null hypothesis
6	H_{o6}	6.919	0.140	Accept the null hypothesis
7	H_{o7}	38.973	0.000	Reject the null hypothesis
8	H_{o8}	1.735	0.188	Accept the null hypothesis
9	H_{o9}	0.208	0.648	Accept the null hypothesis
10	H_{o10}	2.023	0.155	Accept the null hypothesis
11	H_{o11}	2.763	0.096	Accept the null hypothesis
12	H_{o12}	9.199	0.002	Reject the null hypothesis
13	H_{o13}	1.098	0.295	Accept the null hypothesis
14	H_{o14}	15.933	0.000	Reject the null hypothesis
15	H_{o15}	2.709	0.100	Accept the null hypothesis
16	H_{o16}	0.314	0.575	Accept the null hypothesis
17	H_{o17}	0.119	0.731	Accept the null hypothesis
18	H_{o18}	6.906	0.009	Reject the null hypothesis
19	H_{o19}	23.675	0.000	Reject the null hypothesis
20	H_{o20}	14.010	0.000	Reject the null hypothesis
21	H_{o21}	6.298	0.012	Reject the null hypothesis
22	H_{o22}	11.629	0.001	Reject the null hypothesis

Also, Cheating behaviour of students depends on their gender and stream. It also depends on their perception about academic cheating, understanding of material, performance pressure, sitting plan in the exams, inadequate exam preparations, belief that cheating is inevitable and peer influence. Cheating behaviour of the students is found to be positively correlated to the their subject liking or disliking, interest in the subject and instructor liking.

3.6 Latent Class Analysis

3.6.1 Latent Class Models/ Path Models

The path model diagram is the graphical method of displaying the causal relationships among variables in a LCA. Latent variables are represented by circles and manifest variables (indicators as well as grouping variables), by rectangles. The arrows leading from the latent variable to each indicator shows the direction of causal influence. Thus, each arrow represents a conditional probability or interaction of the corresponding indicator with respect to latent variable . Likewise, the absence of a line between two variables represents the conditional independence of the variables. The order in which manifest variables appear in the figure can be used to indicate the order in which the indicator variables were observed.

A detailed description of the variable used in the current study for performing LCA is in Appendix C. Given the relatively large number of observed variables measuring the latent variable and the number of response categories per variable, the number of parameters is fairly high. For this reason larger models were not considered. But out of those models, only 4 models provide the efficient results, consequently we proposed following 4 models for estimating the cheating behaviour of the students using LCA.

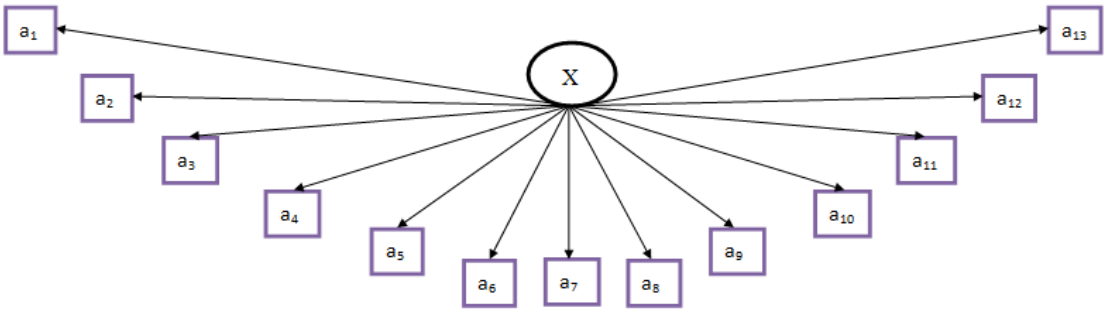

Figure 3.1: Model 0

Model 0 : This model is the simple LC model without any grouping variable, preserving the assumption of local independence. It will estimates the Cheating behaviour (X) on the basis of the individuals' response pattern to the different indicator variables.

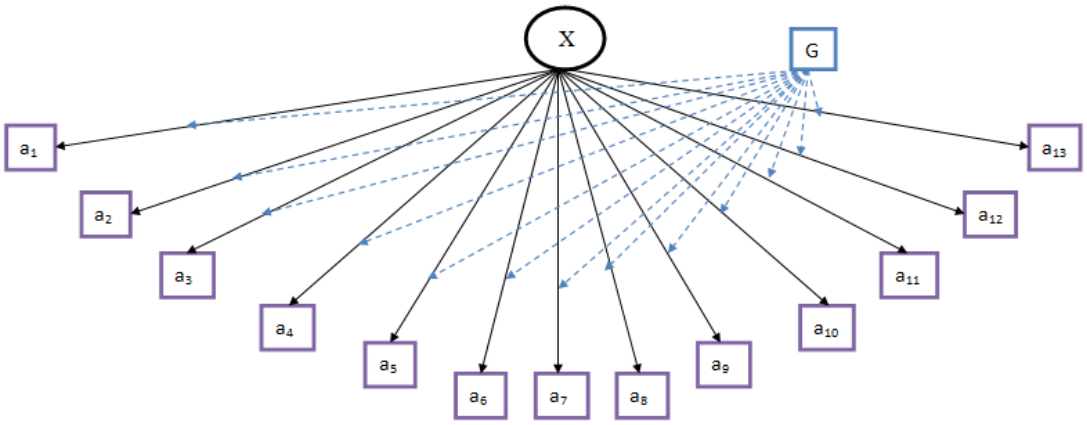

Figure 3.2: Model 1

Model 1 : This model represents the variation in the indicator variables with the inclusion of grouping variable, gender (G) and estimates the Cheating behaviour on the basis of the individuals' response pattern to the different indicator variables through grouping variable.

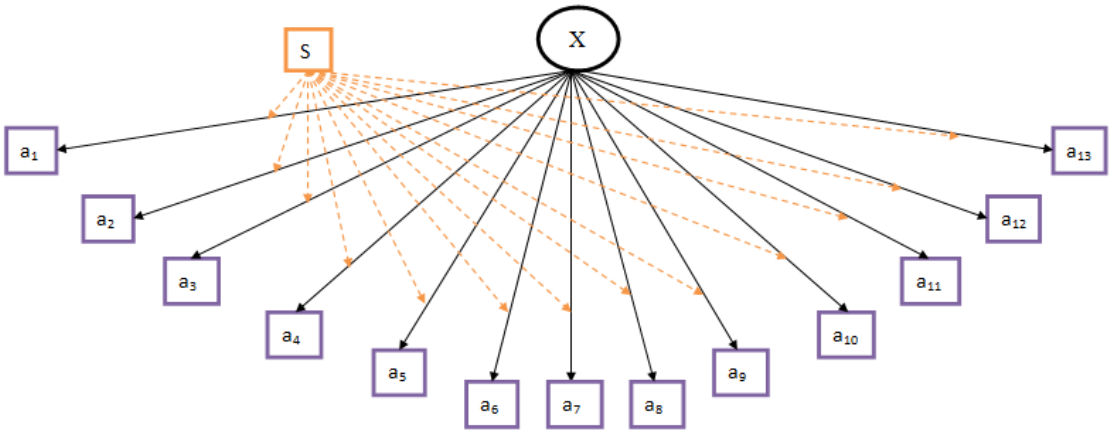

Figure 3.3: Model 2

Model 2 : This model represents the variation in the indicator variables with the inclusion of grouping variable, stream (S) and will test the influence of stream on the response pattern of the individual to estimate the cheating behaviour.

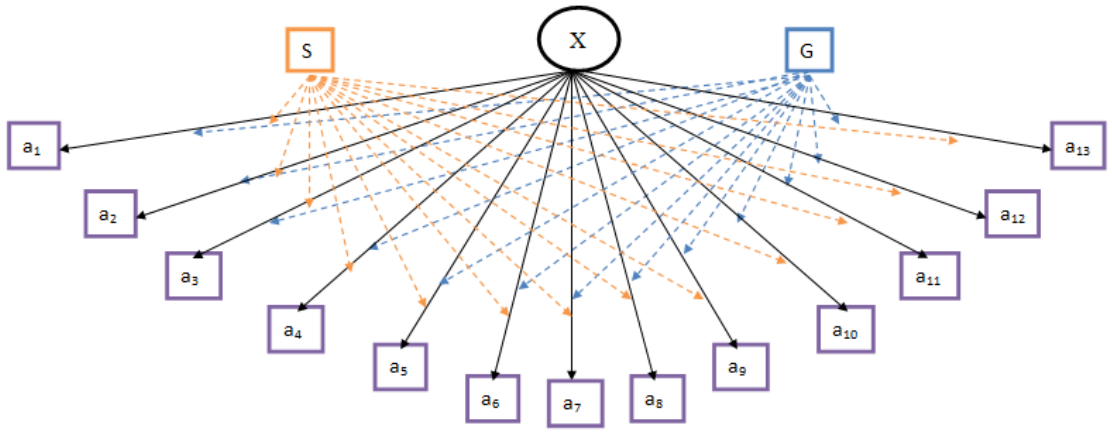

Figure 3.4: Model 3

Model 3 : This model is a complex representation of the variation in the indicator variables on the inclusion of two grouping variables i.e., gender (G) and stream (S). It will estimate the cheating behaviour from the response patterns of the respondents to the different indicator variables under the influence of the gender and stream of the respondents.

3.6.2 Selection of best fitted model

Our data consist of 13 indicator variables which was used for predicting the cheating behaviour of the students. We had incorporated 2 grouping variables with these variables in order to estimate the LC models. As a result, we had to select one best model from the list of the proposed models, which provides the better fit and optimum number of latent classes.

Table 3.5 provides necessary model statistics for different competing models, using poLCA package of R software. Best fitted model is selected on the basis of BIC value (Lin and Dayton, 1997). The model with lowest BIC value is always preferred since it provides the best balance between the two factors namely, model fit and model parsimony. From Table 3.5, it can be seen that BIC value of model 3 is lowest i.e. 14954.7 . Also, it's AIC and LL value (14725.37 and -7315.686, respectively) is also satisfactory. Hence, model 3 is used for further analysis.

Table 3.5: Model diagnostics

Model	**d.f.**[1]	**No. of parameters**	**LL**[2] **value**	**AIC**[3]	**BIC**[4]
Model 0	931	41	-7344.114	14770.23	14970.28
Model 1	929	43	-7330.693	14747.38	14957.2
Model 2	929	43	-7330.405	14746.81	14956.62
Model 3	925	47	-7315.686	14725.37	14954.7

[1] $d.f. : Degrees of freedom of error term$
[2] $LL : Log likelihood value$
[3] $AIC : Akaike information criterion$
[4] $BIC : Bayesian information criterion$

3.6.3 Selection of optimum number of classes

A latent class or class of any latent variable is specified by pattern of responses made to the different manifest variables by the respondents, in terms of conditional probabilities. These probabilities show the possibility that the latent variables can take any particular values depending on the responses of the respondents. It also forms the underlying subgroups of respondents based on the observed attributes. In the present study, classes specify the number of categories into which the responses about the personal cheating behaviour falls. In order to identify optimum number of latent classes, we had performed LCA on the selected model 3 to optimize the number of latent classes. This would help us to find a parsimonious model which provides better fit.

Table 3.6 provides the goodness of fit statistics of model 3 for different number of latent classes. From table 3.6 it is clear that the data-set was best fitted for Model 3 with 3 latent classes as the corresponding BIC as well as AIC values of that model were lowest.

Table 3.6: Goodness of fit statistics of model.

Number of classes (n)	n=2	n=3	n=4
Estimated n-class population shares	0.6489 0.3511	0.4181 0.0659 0.516	0.3632 0.3087 0.2686 0.0595
Predicted n-class memberships	0.6605 0.3395	0.3981 0.0689 0.5329	0.356 0.356 0.2263 0.0617
No. of observations	972	972	972
No. of parameters	30	47	64
Residual degrees of freedom	942	925	908
Maximum log likelihood	-7417.268	-7315.686	-7371.238
AIC	14894.54	14725.37	14770.48
BIC	15040.92	14954.7	14982.76
χ^2	15781.95	11435.73	10963.64

Therefore, the underlying latent classes can be identified as "Occasional Cheaters" (latent class 1) which represents the group of students who are frequent cheaters and cheat commonly in their academic lives , "Persistent Cheater" (latent class 2) which represents the students who are occasional cheaters and cheat rarely in tests, exams or assignments and "Instantaneous Cheaters" (latent class 3) represents the group of students who are instant cheaters and cheats in their academic lives whenever got chance to do so.

Table 3.7 provides the estimated conditional item response membership probabilities for each of the indicator variables. 1^{st} sub row of table 3.7 provides the results for the students who actually admitted that they are involved in the cheating activities and 2^{nd} sub row provides the results for students who have denied for being involved in the cheating.

Table 3.7: Estimated conditional item response probabilities

Indicator variables	Categories of indicators	Latent class 1	Latent class 2	Latent class 3
a_1	1	0.5031	0.9208	0.0470
	2	0.4969	0.0792	0.9530
a_2	1	0.6765	0.9530	0.8734
	2	0.3235	0.0470	0.1266
a_3	1	0.6524	1.0000	0.4183
	2	0.3476	0.0000	0.5817
a_4	1	0.5634	0.8724	0.2874
	2	0.4366	0.1276	0.7126
a_5	1	0.3954	1.0000	0.0304
	2	0.6046	0.0000	0.9696

Indicator variables	Categories of indicators	Latent class 1	Latent class 2	Latent class 3
a_6	1	0.4293	0.9870	0.1661
	2	0.5707	0.0130	0.8339
a_7	1	0.4746	0.9751	0.1029
	2	0.5254	0.0249	0.8971
a_8	1	0.6907	0.9871	0.7658
	2	0.3093	0.0129	0.2342
a_9	1	0.6466	0.9797	0.5114
	2	0.3534	0.0203	0.4886
a_{10}	1	0.5077	0.8925	0.1525
	2	0.4923	0.1075	0.8475
a_{11}	1	0.5674	0.9468	0.1590
	2	0.4326	0.0532	0.8410
a_{12}	1	0.5117	0.7281	0.1882
	2	0.4883	0.2719	0.8118
a_{13}	1	0.5288	0.7153	0.3156
	2	0.4712	0.2847	0.6844

Figure 3.5 provides the graphical representation of the class membership probabilities for estimation of the 3 class lc model. Each group of red bars represents the conditional probabilities of the indicator variables given the latent variable.

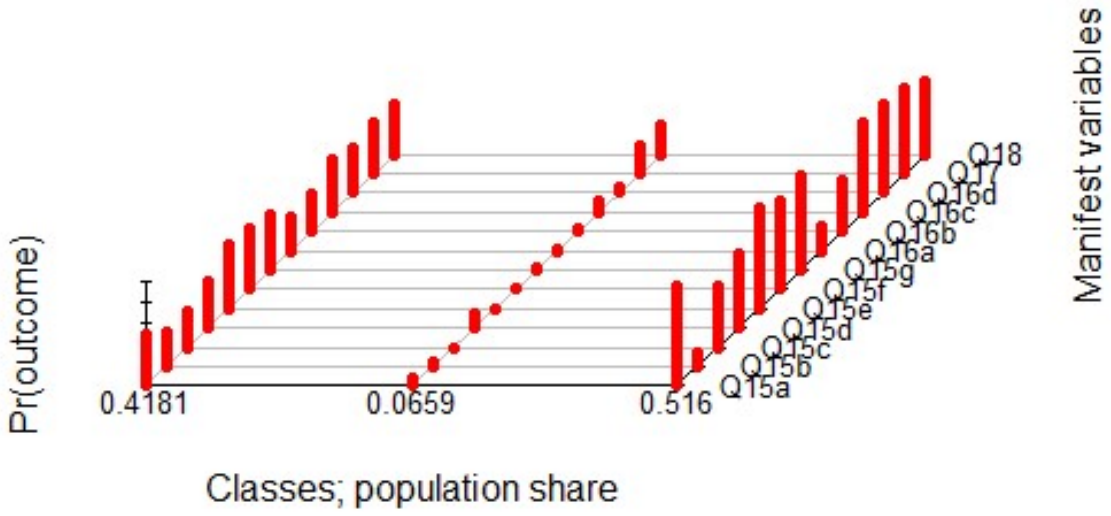

Figure 3.5: Graphical representation of class membership probabilities

3.7 Conclusion

Based on the objectives, defined hypothesis and the analysis of collected data, following conclusions are drawn:

- The current study defines students, by means of questionnaire, exactly what is meant by academics cheating and make them aware of different types of cheating in which they are involved knowingly or unknowingly.
- Also, it is observed that the main reasons because of which student opt for cheating are not knowing or understanding the study material as it is quite difficult and boring; performance pressure especially because of fear of humiliation due to failure ; inadequate exam preparations; ineffective time management skills; Laziness; disliking of the subject/course and lack of interest in the topic.
- On the basis of identified factors and circumstances in which students choose cheating it can be said that the cheating in the universities or any other higher educational institutes is the product of not only students' laziness but also students' not knowing or understanding the study material as they don't like the subject/ they are not interested in the topic/ they find it boring and difficult, performance pressure on them and finally, their inadequate exam preparations.
- From the current study, the latest way of cheating used by students include gadgets, internet, bluetooth devices and writing on stationery material carried by them.
- Also from the table 3.7 it is clear that the extreme responses do not alter the information on cheating behaviour. Hence, we can say that the responses in our study are representative.
- The latent classes, in present study, have been identified as "occasional cheaters", "persistent cheaters" and "instantaneous cheaters", consequently around 42%, 6% and 52% of our respondents are concluded to be occasional, persistent and instantaneous cheaters, respectively.

Academic cheating is a 'disorder' that should be taken seriously to restrain the behaviour of academic dishonesty. It can be challenging to overcome the behaviour of academic dishonesty but an ongoing effort must be taken to lessen its occurrence. The institutions of higher learning should organize programs to promote academic integrity and inculcating an ethical behaviour amongst tertiary students. Students should also be made aware on the negative implications they will receive if they are found to be involved in the academic cheating. Institutions of higher learning should also implement a clear and strict policy on the act of academic dishonesty.

Bibliography

[1] Aiken, L. S. & West, S. G. (1991), *Multiple regression: Testing and interpreting interactions.* Newbury Park, CA: Sage Publications.

[2] Akaike, H. (1973), *Information Theory and an Extension of the Maximum Likelihood Principle.* In B Petrov, F Csake (eds), Second International Symposium on Information Theory, pp. 267-281 Akademiai Kiado, Budapest, Hungary.

[3] Alarape, A. I. & Onakoya, A. Y. (2003), *Correlates of examination cheating behavior among university students.* IFE Psychologia: An International Journal, 11: 71-79.

[4] Anitsal, I. Anitsal, M. & Elmore, R. (2009), *Academic dishonesty and intention to cheat : A model on active versus passive academic dishonesty as perceived by.* 13(2): 17–27.

[5] Beck, L. & Ajzen, I. (1991), *Predicting dishonest actions using the theory of planned behavior.* Journal of Research in Personality, 25: 285-301.

[6] Carmines, E. G. & Zeller, R. A. (1979), *Reliability and Validity Assessment.* Newbury Park, CA, SAGE.

[7] Curtis, G. & Vardanega, L. (2016), *Is Plagiarism changing over time? A 10-year time-lag study with three points of measurement.* Higher Education Reserach & Development, 35:1-13.

[8] Davis, S. F., Grover, C. A., Becker, A. H. & Mcgregor, L. N.(1992), *Academic Dishonesty: Prevalence, Determinants, Techniques, and Punishments.* Teaching of Psychology 19(1): 16-20.

[9] Davis, S, F., Drinan, P. F. & Gallant, T. B. (2009) , *Cheating in school: What we know and what we can do.* United Kingdom: Wiley-Blackwell.

[10] Daniel, L. G., Blount, K. D. & Ferrell, C. M. (1991), *Academic misconduct among teacher education students: a descriptive–correlational study.* Research in Higher Education, 32(6): 703–724.

[11] Dempster A. P., Laird N. M. & Rubin D. B. (1977), *Maximum likelihood from incomplete data via the EM algorithm.* Journal of the Royal Statistical Society Series B (Methodological), 39(1):1-38.

[12] Donse, L. & Van De. Goep, I. H. (2013), *Academic dishonesty among college students.* Predictors and interventions, Social Cosmos, 10(1): 40–50.

[13] Field, A. P. (2005), *Discovering Statistics Using SPSS*. Sage Publications Inc.

[14] Gamer, M., Lemon, J., Fellows, I. & Singh, P. (2012), *Package irr: Various Coefficients of Inter rater Reliability and Agreement*. `https://cran.r-project.org/package=irr`

[15] Griffin, D. J., Bolkan, S. & Goodboy, A. K. (2015), *Academic dishonesty beyond cheating and plaigiarism: Students' interpersonal deception in the college classroom*. Qualitative Research Reports in Communication, 16(2):9-19.

[16] Harding, T., Carpenter, D. , Finelli, C. & Passow, H. (2004), *Does academic dishonesty relate to unethical behavior in professional practics? An exploratory study*. Conference paper in 3rd conference on Ethics and Social Responsibility in Engineering, 10: 311-324.

[17] Hinton, P. R., Brownlow, C., Mcmurray, I. & Cozens, B. (2004), *SPSS explained*. East Sussex, England, Routledge Inc.

[18] Holleque, K. L. (1982), *Cheating behavior of college students. Doctoral dissertation, Montana State University*. Dissertation Abstracts International 43: 88. University Microfilms No. 82-14,649.

[19] Hosny, M. & Shameem, F. (2014), *Attitude of Students Towards Cheating and Plagiarism*. University Case Study

[20] Houston, J. P. (1976), *The assessment and Prevention of Answer Copying on Undergraduate Multiple-Choice Examinations*. Research in Higher Education, 5: 301-311.

[21] Houston, J. P. (1983), *Alternate test forms as a means of reducing multi-ple-choice answer copying in the classroom*. Journal of Educational Psy-chology, 75: 572–575.

[22] Huck, S. W. (2007), *Reading Statistics and Research*. United States of America, Allyn & Bacon.

[23] Hussein, N., Rahman, N. A. A., Rusdi, S. D., Omar, M. K., & Aziz, Z. Z. A. (2018), *Factors that Influence Self-Perceived Academic Cheating: An Empirical Evidence of Business Students*. International Journal of Academic Research in Business and Social Sciences, 8(11): 758-767.

[24] Hutton, P. (2006), *Understanding student cheating and what educators can do about it*. College Teaching, 54(1): 171-172.

[25] Jendrek, M. P. (1989), *Faculty reactions to academic dishonesty*. Journal of College Student Development, 30: 401–406.

[26] Jung In, S. B., Jun, M. K. & Shane, C. B. (2015), *Students' Unethical Academic Behaviors: A Self Determination Theory Approachm*. Journal of Tourism and Hospitality Management, 3(2): 60–72.

[27] Jeergal, P. A., Surekha, R., Sharma, P., Anila, K., Jeergal, V. A. & Rani, T. (2015), *Prevalence, perception and attitude of dental students towards academic dishonesty and ways to overcome cheating behaviors*. Journal of Advanced Clinical & Research Insights, 2: 2–6.

[28] Kalhori, Z. (2014), *The Relationship between Teacher-student Rapport and Students willingnes to cheat*. Procedia-Social and Behavioral Sciences, 136: 153-158.

[29] Khodaie, E., Moghadamzadeh, A. & Salehi, K. (2011), *Factors affecting the probability of academic cheating school students in Tehran.* Procedia-Social and Behavioral Sciences, 29: 1587-1597.

[30] King, D. L. & Case, C. J. (2014), *E Cheating: Incidence and Trends Among College Students.* Issues in Information Systems, 15(1): 20–27.

[31] Koh, C. E. & Nam, K. T. (2005), *Business use of the internet: a longitudinal study from a value chain perspective.* Industrial Management & Data Systems, 105: 85-95.

[32] Koul, R., Clariana, R. B., Jitgarun, K., Songsriwittaya, A. (2009), *The influence of achievement goal orientation on plagiarism.* Learning and Individual Differences, 19: 506-512.

[33] K12 Academics., *Types of Cheating.*
`https://www.k12academics.com/Academic%20Dishonesty/types-cheating`

[34] Linzer , D. A. and Lewis J. (2011), *poLCA: Polytomous Variable Latent Class Analysis.* Annals of Applied Statistics, 11(4): 2085-2115.
`http://CRAN.R-project.org/package=poLCA`

[35] Marsh, J. D., & Campion, J. (2018), *Academic integrity and referencing: whose responsibility is it?.* Journal of Academic Language and Learning, 12(1), A213–A226.

[36] Mccabe, D. (1999), *Toward a Culture of Academic Integrity.* Journal of higher education 46(8): B7.

[37] McLachlan, G. & Peel, D. (2000), *Finite Mixture Models.* John Wiley & Sons, New York.

[38] Michaels, J. W. and Miethe, T. D. (1989), *Academic effort and college grades* Social Forces, 68(1): 309–319.

[39] Muchai, J. (2014) *An investigation into factors that contribute to cheating in examinations in technical institutions in Central Province, Kenya.* University oF South Africa, Pretoria.

[40] Munir, M., Ahmed, Z. & Shahzadi, E. (2011), *A Study on Academic Dishonesty of University Students.* Conference: 8th international conference on Recent Advances in Statistics, volume: Lahore, Pakistan. DOI:10.13140/2.1.1196.8002.

[41] Orosz, G., Dombi, E., Toth-kiraly, I., Bothe, B. Jagodics, B. & Zimbardo, P. G. (2016), *Academic Cheating and time perspective: Cheaters live in the present instead of the future.* Learning and Individual Differences, 52: 39-45.

[42] Rajendran, A. (unpublished), *Research Paper on Cheating.*

[43] Rettinger, D. & Kramer, Y. (2009), *Situational and Personal Causes of Student Cheating.* Research in Higher Education, 50: 293-313.

[44] Robinson, J. (2009), *Triandis theory of interpersonal behaviour in understanding software privace behaviour in the South African context.* Masters degree, University of the Witwatersrand.

[45] Romney, M. & Steinbart, P.(2003), *Accounting Information Systems.* 9th Edition, Prentice Hall, Upper Saddle River, NJ.

[46] Schwartz, G. (1978), *Estimating the Dimension of a Model*. The Annals of Statistics, 6: 461-464.

[47] Sierles, F., Hendrickx, I. & Circle, S. (1980), *Cheating in medical school*. Journal of Medical Education, 55: 124-125.

[48] Skrondal, A., Rabe-Hesketh, S. (2004), *Generalized latent variable modeling: multilevel, longitudinal, and structural equation models*. Interdisciplinary statistics 1st ed., Chapman & Hall; Boca Raton, Fla.

[49] Smith, M., Ghazali, N. & Noor Minhad, S. F. (2007), *Attitudes towards plagiarism among undergraduate accounting students: Malaysian evidence*. Asian Review of Accounting, 15(2): 122–146.

[50] Straub, D., Boudreau, M. C. & Gefen, D. (2004), *Validation guidelines for IS positivist research*. Communications of the Association for Information Systems, 13, 380-427.

[51] Tom, G. & Borin, N. (1988), *Cheating in academic*. Journal of Education for Business, 63: 153-157.

[52] Vermunt, K. & Jeroen k. (2010), *Latent class modelling with covariates: Two improved three-step approaches*. Political Analysis, 18: 450-469.

[53] Wee, Y. S. & Quazi, H. A. (2005), *Development and validation of critical factors of environmental management*. Industrial Management & Data Systems, 105, 96-114.

[54] Whitley, B. E. (2002), *Principals of Research and Behavioural Science*. Boston, McGraw-Hill.

[55] Wilson, J. (2010), *Essentials of business research: a guide to doing your research project*. SAGE Publication.

Appendices

Appendix A

Detailed Description of district wise, sample size selected from considered sampling frame

S.No	Institutions	Course	Sample Size
Off Campuses			
1	Kathua Campus	MBA	1
		MCA	1
		LAW	4
		BCA	4
		BBA	3
		BSC	107
		BED	47
		BCOM	10
		BA	152
2	Reasi Campus	Sociology	1
		BCA	0
		BSC	9
		BED	3
		BCOM	0
		BA	30
3	Udhampur Campus	M.Com	2
		BCA	2
		BBA	3
		BSC	64
		BED	28
		M.A Economics	2
		BCOM	12
		BA	127

S.No	Institutions	Course	Sample Size
4	Poonch Campus	Sericulture	1
		BCA	0
		BSC	29
		BCOM	0
		BA	75
5	Kishtwar Campus	Kashmiri	1
		MSC IT	0
		MSC GEOLOGY	0
		BSC	19
		BCOM	0
		BA	23
6	Ramnagar Campus	Sociology	1
7	Bhaderwah Campus	MCA	1
		English	1
		Geography	1
		MBA	0
		B.com	1
8	Doda	BSC	33
		BCA	1
		BCOM	0
		BA	83
9	Rajouri	BSC	73
		BCA	0
		BED	18
		BCom	1
		BA	158
10	Samba	BSC	32
		BED	13
		BA	49
		BCom	4
11	Ramban	BSC	6
		BA	32
Affiliated Colleges (Jammu)			
12	Dogra Law College	LAW	9
	K.C Law College		9
13	Govt. College of Engineering	ENGG	25
	Model Inst. Of Engg.&Tech.		35
	MBS College of Engg.&Tech.		27
14	MC Khalsa College of Education	BEd	6
	Chenab College of Education		2
	Trikuta College of Education		1
	National College of Education		4
	Calliope College of Education		7
	Harvard College of Education		4
	K.C College of Education		4

S.No	Institutions	Course	Sample Size
	Guru Nanak College of Education		1
	B.N College of Education		2
	Sai Shyam College of Education		1
	J.K College of Education		2
	Sant Mela Singh College of Education		3
	K.C Gurukul College of Education		4
	Ranjit College of Education		4
	Sacred Heart College of Education		4
	Galaxy College of Education		1
15	Govt. Degree College, R.S Pura	BSC	10
		BA	24
	Govt.Degree college Akhnoor		32
	Govt Degree College Bishnah		10
	Govt.Degree college Paloura		31
Main Campus			
16		Buddhist Studies	11
17		Dogri	16
18		English	29
19		Hindi	29
20		Punjabi	9
21		Sanskrit	10
22		Urdu	28
23		History	29
24		Political Science	30
25		Economics	31
26		B.Lib	8
27		M.Lib	6
28		Sociology	17
29		Psychology	14
30		Home Science	12
31		Chemistry	22
32		Geology	16
33		Geography	16
34		Physics	22
35		Electronics	15
36		Remote Sensing	3
37		Bio Tech	8
38		Bio-Chem.	5
39		Micro-Biology	4
40		Botany	15
41		Environment science	17
42		Zoology	16
43		Human Genetics	3
44		Mathematics	26

S.No	Institutions	Course	Sample Size
45		Statistics	14
46		Computer science	27
47		Business school	36
48		Commerce	28
49		Education	29
50		Life-long learning	17
51		Law	84
52		BBA	14
53		M.Ed	25
54		Physical Education(BPED)	27
55		Physical Education(MPED)	16
Total			2317

Appendix B

Questionnaire Used For the Survey

Respondents' details

A	**Sex**	Male		Female	
B	**Age**	18-20	20-22	22-24	24&above
C	**Religion**	Hinduism	Islam	Sikhism	Christianity	Any other
D	**Status**	Under-graduates	Graduates	Post-graduates	
E	**Stream**	Science	Medical Science	Technology	Commerce
		Arts	Management	Law	Any other, please specify

PART 1: General Questions

1. Do you think students cheat during their academic life in schools, colleges or university?
   a) Yes **b)** No
2. What percentage of students cheats **regularly** during exams or tests or assignments?
   a) 20-40% **b)** 40-60% **c)** 60-80% **d)** 80-100%
3. What percentage of students cheats **occasionally** during exams or tests or assignments?
   a) 20-40% **b)** 40-60% **c)** 60-80% **d)** 80-100%
4. Students are more likely to cheat during:
   a) Group projects **b)** In-class assignments **c)** On-line courses
   d) Take-home exams **e)** Classes that are unimportant
5. Which is the one main reason why students cheat during exam?
   a) Didn't have enough time to study **b)** Course was too hard
   c) Bunk class too often **d)** Professor gave very hard tests
   e) Was lazy during semester **f)** any other ,please specify

PART 2: a) Situational aspects

6. Students are most likely to cheat if they don't know or understand the material

a) Agree b) Disagree

If Agree, then in which of the situation content of course is responsible for that?

When course is:

a) Easy and interesting b) Easy and boring

c) Difficult and interesting d) Difficult and boring

7. Type of assignment or exam given by instructor will influence students' decision of cheating, if they have

a) Essay type question b) Short answers c) Multiple choice questions

d) Computational/Mathematical questions e) Doesn't matter

8. Performance pressure(pressure of obtaining good grades) will influence students' decision of cheating

a) Agree b) Disagree

If Agree, then do you think that following circumstances are going to influence their decision of cheating

	Circumstances	I	II
a	Humiliation due to failure	Agree	Disagree
b	Lack of self confidence (students lack confidence that they can pass)	Agree	Disagree
c	Parental pressure for better performance	Agree	Disagree
d	Grades affecting the ability to get into a program or stay in a program	Agree	Disagree

9. Sitting plan in exam will influence their decision of cheating

a) Agree b) Disagree

10. Punishment for cheating will influence their decision of cheating

a) Agree b) Disagree

11. Inadequate preparations will influence their decision of cheating

a) Agree b) Disagree

12. Ineffective time management skills will influence their decision of cheating

a) Agree b) Disagree

13. Laziness will influence their decision of cheating

a) Agree b) Disagree

14. Instructor vigilance will influence their decision of cheating

a) Agree b) Disagree

b) Personal interest aspects

	Circumstances	I	II
a	When student likes the subject/course they are less likely to cheat	Agree	Disagree
b	When students are interested in the topic, they are less likely to cheat	Agree	Disagree
c	When students like the instructor, they are less likely to cheat	Agree	Disagree
d	When students believe that cheating is inevitable and everybody cheats, they are more likely cheat	Agree	Disagree
e	Peer influence (other classmates also cheat) will influence their decision of cheating	Agree	Disagree

PART 3: Types of cheating

15. Have you ever cheated in tests or exams in school, college or university?

a) Yes **b)** No

If yes, then whether you engaged in following types of cheating activities during exams or tests?

		I	II
a	Used any sort of prohibited material (like crib sheets, written material, any gadget) in the exams	Yes	No
b	Deliberately looked at another student's test sheet or made someone else to look at your test sheet	Yes	No
c	Passed answers to another person during a test or take answers from them (forbidden sharing of information during exams between students)	Yes	No
d	Planned with another student how would you cheat prior to an exam	Yes	No
e	Obtain a copy of an exam paper or test paper before exam	Yes	No
f	Ever made attempt to obtain or accept assistance from any other person (student, staff or professor) during exam	Yes	No
g	Lied to your instructor for conducting an exam or test again / for not appearing in exams or tests	Yes	No

16. Have you ever cheated in assignments or projects given at school, college or university?

a) Yes **b)** No

If yes, then whether you are engaged in following types of cheating activities during exams or tests?

		I	II
a	Copied another person's assignment/research/thoughts through online/offline mode and passed it off as yours own	Yes	No
b	Complete the work which is assigned to someone else or made any other person complete the work assigned to you	Yes	No
c	Illicitly obtain material or steal material needed to complete an assignment?	Yes	No
d	Misrepresent a family or personal situation (made excuses) to get an extension in assignment?	Yes	No

17. Have you ever prevented other students from completing their work?

a) Yes **b)** No

18. Have you ever forged (copy) a faculty/family/friend's signature on permission form or add/drop form?

a) Yes **b)** No

19. What do you think is the latest or most frequently used way of cheating?

a) Gadgets **b)** Material **c)** Copying

d) Writing on desks **e)** Any other, please specify

Name of the Investigator	Signature of the Investigator	Date of the interview

Appendix C

Detailed description of variables used in LCA

S.No.	Variables	Descriptions
1	X	Cheating behaviour of the students
2	a_1	Used any sort of prohibited material in the exam
3	a_2	Deliberately looked at another student's test sheet or made someone else to look at your test sheets
4	a_3	Passed answers to another person during a test or take answers from them
5	a_4	Planned with another student how to cheat prior to exam
6	a_5	Obtain a copy of an exam paper or test paper before exam
7	a_6	Ever made attempt to obtain or accept assistance from any other person during exam
8	a_7	Lied to an instructor for conducting an exam or test again/ for not appearing in exams or tests
9	a_8	Copied another person's assignment/ research/ thoughts through online/ offline mode and passed it off as your own
10	a_9	Complete the work which is assigned to someone else or made any other person complete the work assigned to you
11	a_{10}	Illicitly obtain material or steal material needed to complete assignment
12	a_{11}	Misrepresenting a family or personal situation (made excuses) to get an extension in assignment
13	a_{12}	Ever prevented other students from from completing their work
14	a_{13}	Ever forged (copy) a faculty/ family/ friend's signature on permission form or add/ drop form
15	G	Gender of the student
16	S	Stream of the student

The Journal of Mathematical Sociology

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/gmas20

Latent class analysis of multigroup heterogeneity in propensity for academic dishonesty

Sunil Kumar, Apurba Dabgotra & Diganta Mukherjee

To cite this article: Sunil Kumar, Apurba Dabgotra & Diganta Mukherjee (2024) Latent class analysis of multigroup heterogeneity in propensity for academic dishonesty, The Journal of Mathematical Sociology, 48:1, 81-99, DOI: 10.1080/0022250X.2023.2179999

To link to this article: https://doi.org/10.1080/0022250X.2023.2179999

THE JOURNAL OF MATHEMATICAL SOCIOLOGY
2024, VOL. 48, NO. 1, 81–99
https://doi.org/10.1080/0022250X.2023.2179999

Routledge
Taylor & Francis Group

Latent class analysis of multigroup heterogeneity in propensity for academic dishonesty

Sunil Kumar [a], Apurba Dabgotra [a], and Diganta Mukherjee [b]

[a]Department of Statistics, University of Jammu, Jammu, J&K, India; [b]Sampling and Official Statistical Unit, Indian Statistical Institute, Kolkata, India

ABSTRACT

Latent class analysis (LCA) is a cross-sectional latent variable mixture modeling (LVMM) approach. Like all LVMM approaches, LCA aims to find heterogeneity within the population by identifying homogenous subgroups of individuals, with each subgroup (called latent class) possessing a unique set of characteristics that differentiate it from other subgroups. LCA can be carried out with categorical latent and indicator variables. But, LCA is unable to examine the association between respective items and the latent variable among categories of individuals. Multiple-group LCA, in particular, is a useful extension of LCA which enables the testing of homogeneity of the class patterns between groups of the individual through a series of constraints. In this paper, we have performed a multi-group latent class analysis for measuring self reported academic dishonesty among the students of University of Jammu. From the analysis, three general behaviors of academic cheaters are identified as rare, frequent, and instant cheaters. Further, from the multi-group LCA, it is envisaged that female students of University of Jammu are more instantaneous cheaters than male students. Students who are self-reported cheaters from sciences and humanities of the University of Jammu are persistent in cheating whereas from professional courses they are more occasional.

ARTICLE HISTORY
Received 30 December 2021
Revised 3 November 2022
Accepted 8 December 2022

KEYWORDS
AIC; BIC; grouping variables; Latent Class Analysis (LCA); Multigroup Latent Class Analysis (MLCA)

1. Introduction

Latent class analysis (LCA) is a popular technique that is used to measure certain variables (called latent variables) which cannot be directly observed either because of its hypersensitive nature or due to the unavailability/difficulty in the measurement of a variable required to analyze them. Latent variables are not directly observed rather their values are deduced from observed variables. It also identifies the relationship between categorical latent and observed variables in the form of membership among the subjects using observed variables. LCA enumerates the latent classes in which respondents with similar response patterns in terms of observed variables are in the same class. In traditional LCA, two sets of parameters are estimated i.e., class

CONTACT Apurba Dabgotra apurvavdabgotra@gmail.com Department of Statistics, University of Jammu, Jammu, Jammu & Kashmir 180006, India

membership probabilities and item-response probabilities conditioned on class membership.

LCA is one of the finite mixture modeling approaches that can also capture within a group as well as between group heterogeneity. In every data set, there is a mixture distribution that is heterogeneous across the sample but homogeneous within sub-samples. So, LCA captures the unobserved heterogeneity in the population, which is depicted by latent variable and has a direct physical interpretation.

LCA was first introduced by Lazarsfeld (1950) for performing the clustering for dichotomous items based on categorical observed variables. Subsequently, this technique was widely studied by various other statisticians. Goodman, (1974) extended it to nominal variables by developing an algorithm for obtaining maximum likelihood estimates of the parameters of Latent Class (LC) models, so that it could be practically applicable. In the same period of time Haberman (1979) put forward a relationship between LC models and log-linear models for frequency tables containing unknown cell counts. Hagenaars, (1990) also proposed a general framework for categorical data analysis with discrete latent variables. Various other important studies have also been carried out since then, such as the development of LC models containing covariates, several latent variables, multigroup latent class analysis, etc.

The multigroup extension of the standard LC model was first proposed by Clogg & Goodman, (1985). It was originally developed for the analysis of latent structures of categorical latent variables across the different groups (Kankaras & Moors, 2009). It now serves as a useful tool for segmentation, which enables the testing of homogeneity of the class patterns between groups of the individual through a series of constraints. Multiple group LCA is especially useful when there are existing subgroups in the data representing different populations and when the differences across these groups need to be compared within the latent class model (A. McCutcheon, 2002; Collins & Lanza, 2010). In other words, multiple group LCA examines differences between populations where the underlying latent group can be measured directly and this latent membership can be accordingly examined with the differences of population groups.

Three types of categorical variables are involved in multigroup LC models which include a set of indicator variables; one or more latent variables that account for the relationships between the indicator variables; and a grouping variable, which is a known categorical variable that can be related to both indicator variables and latent variable(s). The major assumption of this model is that the indicator variables are mutually independent and conditioned on the grouping variable and the latent variable(s). This is usually referred to as the assumption of local independence in multigroup LCA (Lazarsfeld & Henry, 1968). In multiple-group LCA, both the measurement part and

structural part of the model can vary across groups. Multigroup LCA offers a flexible alternative to the more commonly used multigroup confirmatory factor analysis (MCFA) and multigroup item response theory (IRT) approaches, which both rely on stronger distributional assumptions than LC analysis.

The purpose of this paper is to perform multi-group LCA using R software to take into consideration the heterogeneity among individuals which otherwise is not possible to extricate through LCA with covariates.

2. Methodology

Following the terminology of Kankaraš et al. (2010))we assume an LC model with five observed polytomous variables A, B, C, D, and E having I (i = 1, 2, . . ., I), J (j = 1, 2, . . ., I), K (k = 1, 2, ., K), L (l = 1, 2, . . ., L) and M (m = 1, 2, . . ., M) categories, respectively; one latent polytomous variable X with T classes (t = 1, 2, . . ., T), and one grouping variable G with S groups indexed by s = 1, 2, . . ., S. The variables A, B, C, D, and E are observed in each of these S groups. Thus, we have a set of S $\times$ five-way ($I \times J \times K \times L \times M$) observable contingency tables, or one $\times$ Six-way table ($I \times J \times K \times L \times M \times S$). Then the multigroup LC model takes the following form:

$$\pi_{ijklmts}^{ABCDEX|G} = \pi_{ts}^{X|G}\pi_{its}^{A|XG}\pi_{jts}^{B|XG}\pi_{kts}^{C|XG}\pi_{lts}^{D|XG}\pi_{mts}^{E|XG} \tag{1}$$

Here, $\pi_{ijklmts}^{ABCDEX|G}$ denotes the conditional probability that an individual who belongs to the s^{th} group will be at level (i, j, k, l, m, t) with respect to variables A, B, C, D, E and X. The conditional probability of 'X' taking level 't' given the s^{th} group of grouping variable 'G' is denoted with $\pi_{ts}^{X|G}$, which determines the LC proportion for the s^{th} group. $\pi_{its}^{A|XG}$ is the conditional probability of an individual taking level 'i' of variable A, for a given level 't' of the latent variable 'X' and a given group membership 's' of the grouping variable 'G.' Similarly, the parameters $\pi_{jts}^{B|XG}$, $\pi_{kts}^{C|XG}$, $\pi_{lts}^{D|XG}$ *and* $\pi_{mts}^{E|XG}$ are defined accordingly.

It should be noted that Eq. (1) implies that indicator variables A, B, C, D, and E are independent of each other, given the value of the latent variable X. This is usually referred to as the assumption of local independence (Lazarsfeld & Henry, 1968). The latent class and conditional response probabilities are constrained to a sum of 1 i.e., $\Sigma_i \pi_{ts}^{X|G} = 1$ and $\Sigma_i \pi_{its}^{A|XG} = 1$, and so on.

The model presented in Eq. (1) is heterogeneous model since all model parameters differ across groups (Clogg & Goodman, 1985). showed standard LC models as a special case of the more general multigroup LC model, eg. for $S = 1$ in Eq. (1), we have

$$\pi_{ijklmt}^{ABCDEX} = \pi_t^X \pi_{it}^{A|X} \pi_{jt}^{B|X} \pi_{kt}^{C|X} \pi_{lt}^{D|X} \pi_{mt}^{E|X} \tag{2}$$

The probabilistic LC model presented in Eq. (1) can also be parameterized using log-linear terms (Goodman, 1974; Haberman, 1979; Hagenaars & McCutcheon, 2002). The conditional response probabilities from the probabilistic parameterization can be obtained from log-linear terms as follows:

$$\pi_{its}^{A|XG} = \frac{exp(\lambda_i^A + \lambda_{it}^{AX} + \lambda_{is}^{AG} + \lambda_{its}^{AXG})}{\Sigma_i exp(\lambda_i^A + \lambda_{it}^{AX} + \lambda_{is}^{AG} + \lambda_{its}^{AXG})}, \quad etc. \tag{3}$$

where λ_i^A and λ_{it}^{AX} represent the parameters of the single-group standard LC model, while λ_{is}^{AG} and λ_{its}^{AXG} are the log-linear parameters. Parameters λ_{its}^{AXG} are the interaction effects of the latent variable with grouping variables on indicator variables. In a similar manner, λ_{is}^{AG} refers to be a direct effect of the grouping variable 'G' on the indicator variable 'A.'

The class membership probabilities $\pi_{ts}^{X|G}$ can be defined in terms of log-linear parameters; that is,

$$\pi_{ts}^{X|G} = \frac{exp(\gamma_t^X + \gamma_{ts}^{XG})}{\Sigma_{i=1}^T exp(\gamma_t^X + \gamma_{ts}^{XG})} \tag{4}$$

where the symbol γ denotes a log-linear parameter of the marginal distribution of the latent variable X (Vermunt & Magidson, 2005).

The multigroup LC model by using a logistic regression-type of the equation for the item response probabilities, i.e., the model for indicator variable 'A' is

$$\pi_{its}^{A|XG} = \frac{exp(\alpha_{is}^{A|G} + \beta_{its}^{AX|G})}{\Sigma_i exp(\alpha_{is}^{A|G} + \beta_{its}^{AX|G})} \tag{5}$$

where $\alpha_{is}^{A|G}$ represents the group-specific intercepts and $\beta_{its}^{AX|G}$ the group-specific slope parameters. The slope parameter $\beta_{its}^{AX|G}$ indicates the strength of the relationships between the latent and indicator variable across the categories of grouping variables. It can be interpreted as a factor loading expressed in log-linear terms (Vermunt & Magidson, 2005). Note that there is a straightforward relation between the log-linear and the logistic formulations of the multigroup LC model presented in Eq. (3) and (5) as

$$\alpha_{is}^{A|G} = \lambda_i^A + \lambda_{is}^{AG} \tag{6}$$

and

$$\beta_{its}^{AX|G} = \lambda_{it}^{AX} + \lambda_{its}^{AXG} \tag{7}$$

In their unrestricted form, the three parameterizations of the multigroup LC model are essentially equivalent, estimating the same number of parameters and producing identical conditional probabilities. However, they allow for slightly different types of model restrictions which have important implications for the procedures to test measurement equivalence. First, in the probabilistic parameterization, equivalence is studied by restricting probabilities to be group invariant, in the log-linear parameterization by eliminating interaction and direct effects, and in logistic formulation by restricting intercepts and slopes to be invariant. Second, the latter two parameterizations are needed to formulate models in which indicator or latent variables are treated as discrete-ordinal.

LC models are usually estimated using maximum-likelihood (ML) under the assumption of a multinomial distribution for the indicator variables in the model. In this study, we have used the poLCA package of R software for estimating the LC models. Model selection is based on the Akaike information criterion (AIC), Bayesian information criterion (BIC), modified AIC (AIC3), and consistent AIC (CAIC), each of which is designed to penalize models with larger numbers of parameters. Since more parameters in a model increase its likelihood, the information criteria reduce that likelihood by a certain amount which is a function of the increased number of estimated parameters. They differ in the specific function with which they calculate the penalizing value for each additional parameter in a model.

Usually, the poLCA package is unable to quantify the effect of different groups of individuals on their conditional item response probabilities as well as on the membership probabilities. So, to compute the conditional item response probabilities and membership probabilities given the different groups of individuals, we perform ordinary LCA for different groups of respondents separately. By using this technique we can not only evaluate different conditional probabilities given a particular group of respondents but also categorize that particular group of individuals according to the considered latent variable.

3. Empirical example

In this section, we present an example showing the use of the proposed technique of multigroup LCA for measuring academic dishonesty among the students of Universty of Jammu. This example involves a modified multigroup LCA as both latent and indicator variables are ordinal. In this example, we used data that we have collected from the students (undergraduates students, BEd students, masters students) of the University of Jammu and its affiliated colleges for studying the academic dishonesty of the students. We have used this particular example because there exists subgroups (males/females, sciences/professionals/humanities) in the data representing different populations and compare the differences across these groups within the latent class

model. We have broadly followed the National Sample Survey Office (NSSO) methodology to select our sample and calculated the sample size which comes out to be 1906. The method used for the calculation of the sample size of the study is provided in Appendix A. We have then employed a multistage stratified random sampling design for selecting the sample, where the first stage unit (FSU) comprises the University main campus, off campuses, and different affiliated colleges and the ultimate stage unit (USU) comprises different courses offered by these institutes. After that, proportional allocation is used for selecting a sufficient number of sample units from each USU. A detailed description of the sampling frame is provided in Appendix B.

We have used a descriptive design to study the cheating behavior of the students, so, a questionnaire is prepared with 26 closed-ended questions related to various factors leading to cheating, types of cheating, and some other related questions. The questionnaire was designed in such a way that it will cover all possible types of academic cheating (Shaughnessy, 1988; (Genereux & McLeod, 1995)) in which students are involved knowingly or unknowingly along with all the factors that encourage those cheating practices. Due to the sensitive nature of study, the face to face paper and pencil survey was conducted and data was collected from the students by personally visiting their respective departments, colleges, or institute. Thus, we have collected data from 1906 students but due to unit non-response, 7 observations were eliminated, and the rest of 1899 were used for the analysis. The questionnaire is divided into the following 5 major sections.

Respondent's Details: This section records the demographic characteristics of the students which include gender, age, religion, status, and stream.

General Questions: This section consists of general questions about cheating to have an idea about the general perception of students on academic cheating and to get an idea from the students about the main reason why students cheat during exams.

Situational Aspects: This section consists of questions related to the situational factors that can influence students to cheat during exams.

Personal Aspects: This section consists of questions about the personal factors that can influence students to cheat during exams.

Types Of Cheating: This section comprises questions about the types of cheating students are involved in. This section was prepared with the objective of making students aware of the types of academic cheating and to know about their cheating behavior.

Table 1 provides the frequency table of the demographic characteristics of the students which include gender (male, female), age (18–20, 20–22, 22–24, 24 & above), religion (Hinduism, Islam, Sikhism, Christianity), status (undergraduates, graduates, postgraduates) and stream (sciences,[1] professional

[1]Sciences here includes life sciences, mathematical sciences, social sciences, Physical sciences, library science, environmental science, home science, etc.

Table 1. Frequency table for demographic characteristics of the respondents.

Variables	Categories				Total
	1	2	3	4	
Gender	768	1131	-	-	1899
Age	744	653	402	100	1899
Religion	1533	231	402	100	1899
Current Status	1158	122	619		1899
Stream	452	473	974		1899

courses,[2] humanities[3]). The proportion of females in the sample is 0.60 and males is 0.40 whereas in the population it was 0.66 and 0.34. Also, the proportion of science students, professional courses students, and humanities students in the sample is 0.24, 0.25, and 0.51, respectively.

In this example, LCA with covariates is performed on the responses of those respondents who have accepted that they have cheated in their academic lives, to identify the latent classes irrespective of the respondent's heterogeneity. The results of LCA with covarites Gender and Stream are shown in the first row of Table 3. It is clear from the values of AIC and BIC that the model with three latent classes is the best fit. We performed LCA with covariates to find the grouping variables that can affect the membership and conditional

Table 2. Detailed description of variables used in LCA.

S.No.	Variables	Descriptions
1	Latent Variable	Cheating behaviour of the students
2	a_1	Used any sort of prohibited material in the exam
3	a_2	Deliberately looked at another student's test sheet or made someone else to look at your test sheets
4	a_3	Passed answers to another person during a test or take answers from them
5	a_4	Planned with another student how to cheat prior to exam
6	a_5	Obtain a copy of an exam paper or test paper before exam
7	a_6	Ever made attempt to obtain or accept assistance from any other person during exam
8	a_7	Lied to an instructor for conducting an exam or test again for not appearing in exams or tests
9	a_8	Copied another person's assignment/research/thoughts through online/offline mode and passed it off as your own
10	a_9	Complete the work which is assigned to someone else or made any other person complete the work assigned to you
11	a_{10}	Illicitly obtain material or steal material needed to complete assignment
12	a_{11}	Misrepresenting a family or personal situation (made excuses) to get an extension in assignment
13	a_{12}	Ever prevented other students from completing their work
14	a_{13}	Ever forged (copy) a faculty/family/friend's signature on permission form or add/drop form
15	Gender	Gender of the student
16	Stream	Stream of the student

[2]Professional courses includes engineering, computer application, bio-technology, commerce, business administration, BEd, MEd, physical education, etc.
[3]Humanities here includes law, all language courses offered by univerty of Jammu, history, sociology, etc.

probabilities of the latent variable. But we were not able to spot the difference across the groups of grouping variable with in LC models. So, we decided to perform multigroup LCA for different groups of respondents separately. The analysis follows the procedure outlined in Section 2, by first fitting different models and then choosing the optimum number of classes for the best-fitted model. We use the BIC statistic as the main model selection fit criterion but the AIC statistic also provides control mechanisms for the sample size of about 2000 cases or less.

In this analysis, we have considered only 13 indicator variables that are related to different types of cheating for predicting the cheating behavior of the students. Two grouping variables are incorporated i.e., gender (Male\Female) and stream (Sciences\Professional\Humanities) with the manifest variables to ensure that the considered LC model is identifiable and it improves the fit of the model. Table 3 provides a detailed description of the variable used in the

Table 3. Goodness of Fit Statistics of Models.

[1]GV		Classes	[2] LL value	[3]AIC	[4]BIC	G^2	χ^2
	LCA with covariates	2	−7417.268	14894.54	15040.95	-	15781.95
		3	−7315.686	14725.37	14954.7	-	11435.73
		4	−7371.238	14770.48	14982.76	-	12963.64
Gender	Males	2	−3378.509	6749.876	6963.651	1898.674	11498.80
		3	−3297.489	6676.977	6843.401	1875.258	11345.68
		4	−3428.784	6790.786	6983.346	1910.854	11529.25
	Females	2	−4085.987	8189.974	8299.681	2015.543	9797.652
		3	−4011.229	8104.459	8280.716	1996.724	9748.828
		4	−4120.295	8203.058	8315.542	2045.657	9826.801
Stream	Sciences	2	−1698.471	3487.267	3609.783	1035.874	11172.30
		3	−1681.041	3444.082	3587.128	992.9755	11105.83
		4	−1723.963	3506.820	3645.543	1062.597	11201.02
	Professional	2	−1217.602	2484.108	2592.840	880.624	12714.514
		3	−1178.722	2439.443	2565.781	835.2384	12660.14
		4	−1269.283	2509.362	2612.071	903.814	12740.634
	Humanities	2	−3799.940	7638.481	7808.761	2077.829	8993.156
		3	−3758.417	7598.834	7770.637	2019.148	8956.801
		4	−3803.751	7670.393	7856.560	2109.107	9023.076
Gender-Stream	Males-Sciences	2	−858.618	1769.076	1879.678	765.464	7650.403
		3	−736.857	1555.714	1663.338	584.9922	7095.459
		4	−883.536	1865.132	1937.263	801.145	7948.910
	Males-Professional	2	−872.227	2048.310	1927.738	838.264	9434.671
		3	−793.8941	1669.788	1779.756	627.9064	9110.694
		4	−943.862	2169.033	2016.641	985.174	9740.843
	Males-Humanities	2	−1597.490	3684.124	3378.681	1279.219	9583.715
		3	−1467.805	3017.611	3150.739	1064.469	9235.73
		4	−1949.491	3850.134	3645.768	1389.763	9759.168
	Females-Sciences	2	−1004.274	1995.567	2113.653	674.598	29537.2
		3	−900.8617	1883.723	2004.331	577.7109	28373.8
		4	−1117.385	2103.421	2147.674	752.679	29759.28
	Females-Professional	2	−617.872	1374.182	1352.438	582.786	9234.824
		3	−549.0374	1180.075	1276.171	461.7203	8768.129
		4	−638.743	1518.243	1498.761	603.045	9508.843
	Females-Humanities	2	−2379.954	4768.810	4980.271	1502.029	9913.156
		3	−2318.207	4718.413	4870.811	1439.194	9804.279
		4	−2434.275	4867.139	5126.215	1589.183	9987.476

1, GV: Grouping variable, 2, LL: Log-likelihood value 3, AIC: Akaike information criterion, 4, BIC: Bayesian information criterion

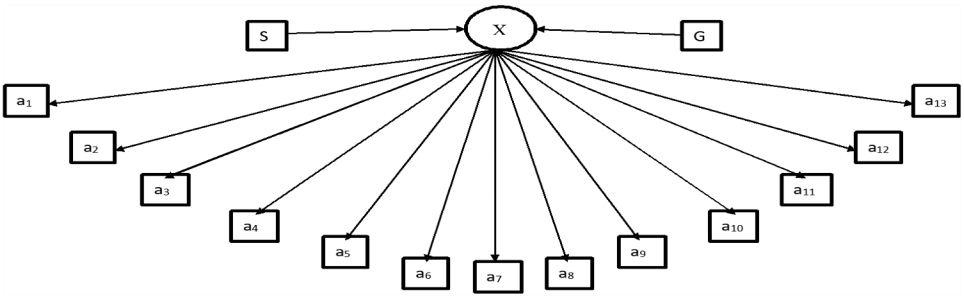

Figure 1. Path Model for Multigroup LCA.

current study for performing LCA and Figure 1 provides the path model for the multigroup LCA.

Next, we perform the multigroup LCA to see the effects of different streams as well as of gender of the students on latent classes obtained earlier. Table 3 provides the goodness of statistics of various models for different numbers of latent classes. The best fitted model in each case is selected based on the BIC value (Lin and Dayton, 1997). The model with the lowest BIC value is always preferred since it provides the best balance between the two factors namely, model fit and model parsimony. From Table 2, it is clear that the best fit is for 3 latent classes in the LC model and multigroup LC models, as the corresponding BIC, as well as AIC values, are minimum. In addition to it, maximum likelihood values, chi-square (χ^2) values and likelihood ratio chi-square (G^2) are also favorable for the models with 3 latent classes.

Since the optimum number of latent classes obtained in each model is 3 and we have considered only those responses who have accepted they have cheated in their academic lives, therefore, the underlying latent classes can be identified as "Occasional Cheaters" (latent class 1) which represents the students who are occasional cheaters and cheat rarely in tests, exams or assignments, "Persistent Cheaters" (latent class 2) which represents the group of students who are frequent cheaters and cheat commonly in their academic lives and "Instantaneous Cheaters" (latent class 3) represents the group of students who are instant cheaters and cheats in their academic lives whenever they get a chance to do so. The first two classes i.e., occasional cheaters & persistent cheaters come under the category of planned cheaters where students plan to cheat before appearing in exams or doing their work either sometimes or often and the last class i.e., instantaneous cheaters come under the category of unplanned cheaters where students beforehand don't have any intention to cheat but when they got an opportunity they immediately go for cheating.

Table 4 provides the estimated membership probabilities in each of the considered cases. From Table 4 it is clear that around 0.47 of the total male students opt for cheating whenever they got a chance to do so whereas when it comes to the female students this proportion increases to 0.57. Also, 0.64 of the

Table 4. Membership probabilities for different models.

	Occasional cheaters	Persistent cheaters	Instantaneous cheaters
LCA with covariates	0.4181	0.0659	0.516
Males	0.4461	0.0866	0.4673
Females	0.3798	0.0533	0.57
Sciences	0.1891	0.6409	0.17
Professional	0.5681	0.2005	0.2314
Humanities	0.0598	0.5313	0.409
Males-Sciences	0.2106	0.6552	0.1342
Males-Professional	0.1408	0.456	0.4032
Males-Humanities	0.1202	0.7266	0.1532
Females-Sciences	0.4395	0.1007	0.4598
Females-Professional	0.2174	0.2116	0.5711
Females-Humanities	0.3343	0.2924	0.3733

total students from the science stream are persistent cheaters whereas this proportion decreases to around 0.53 for students from humanities and around 0.57 of the total students from professional courses are occasional cheaters. When we consider two categories simultaneously then we can observe that the Males in all the streams are persistent cheaters whereas females in all the streams are instantaneous cheaters.

In Figure 2, we have considered graphs of the estimated conditional item response probabilities ($P[a_i = 1|X = j] = \pi_{ij}^{A|XG}; i = 1, 2, \ldots, 13; j = 1, 2, 3$ for LCA with covariates, then for multigroup LCA, we have $P[a_i = 1|X = j, g = G, S] = \pi_{ijg}^{A|XG}; i = 1, 2, \ldots, 13; j = 1, 2, 3;$ $G = males, females; S = sciences, professionals, humanities$) of responding 'yes' to each of the indicator variables for all models i.e., it provides the results for the students who actually admitted that they are involved in the different cheating activities. The table of estimated conditional item response probabilities is given in Appendix C.

Table 5 provides the estimated conditional item response probabilities in gender – stream (combined) LC models of responding "Yes" to the manifest items for each of the latent variables i.e., it provides the results for the students in combined categories who have accepted that they are involved in cheating.

From Table 5 it can be seen that about 78% of the males from sciences sometimes carry prohibited material in their exams whereas as this percentage increases to about 93% for males from professional courses and reduces to about 65%, and 62% for females from professional courses and humanities, respectively. Also, around 35% of females in sciences often carry prohibited material in their exams and almost all the males from humanities admit that they instantly decide to carry prohibited material in their exams.

Around 96% of males from humanities sometimes look at another person's answer sheet or made another person look at their answers sheets. About 88%, 87%, and 92% of males from sciences, females from Sciences and professional courses often looked at another person's answer sheet or made someone else look at their answer sheets during their exams. And almost all the males from

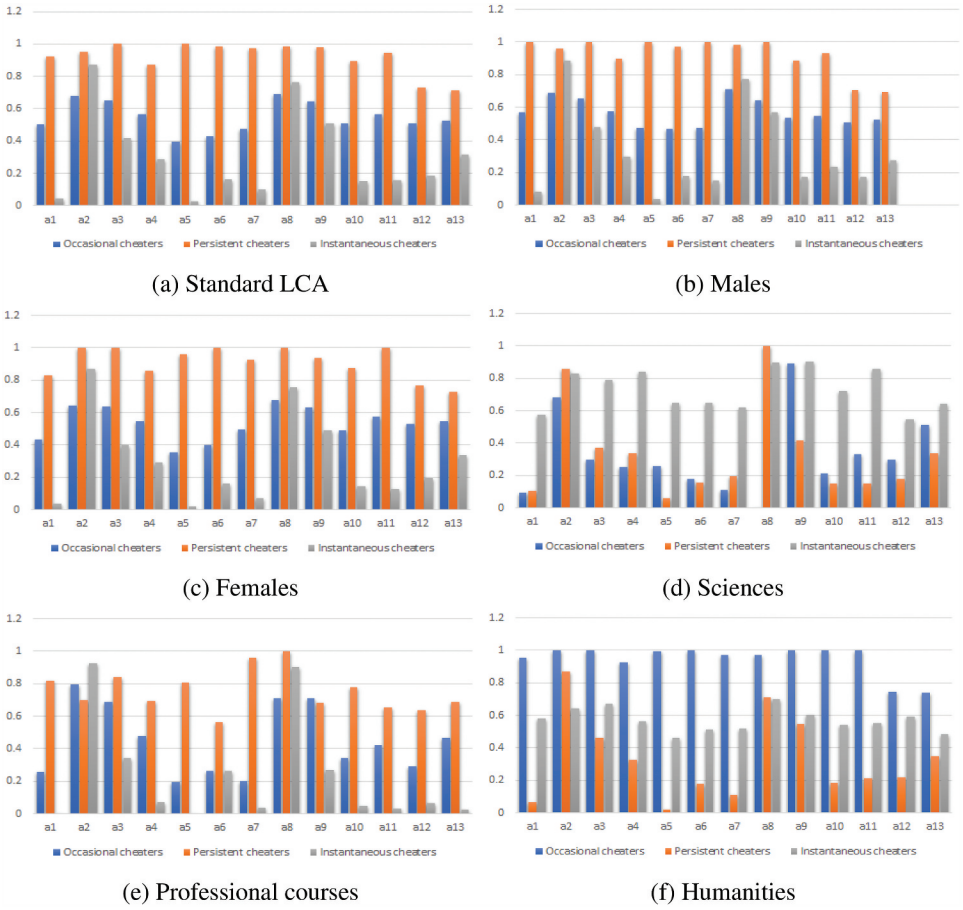

Figure 2. Estimated conditional probabilities for each the considered model.

professional courses as well as females from humanities instantly decide to do the same.

Around 81% of males from sciences sometimes pass/receive answers to/from another person in exams whereas this percentage reduces to about 80%, 78% and 69% for males from professional courses, females from professional courses and humanities, respectively. About 79% of females from sciences frequently pass/receive answers to/from other person and around 86% of males from humanities at sudden decide to pass/receive answers to/from other persons in the exams whenever they have a chance to do so.

About 92% of males from sciences sometimes planned to cheat with another person before entering exam halls whereas this percentage increases to 98% for the males from humanities and reduces to about 64% & 67% for the females from professional courses & humanities. About 78% of females from sciences often plan to cheat in their exams with another student before their exams.

Table 5. Estimated conditional item response probabilities in gender – stream (combined) LC models of responding "Yes" to the manifest items for each of the latent variables.

		Occasional cheaters			Persistent cheaters			Instantaneous cheaters		
		Sciences	Professionals	Humanities	Sciences	Professionals	Humanities	Sciences	Professionals	Humanities
a_1	Male	**0.7790**	**0.9314**	0.8715	0.1562	0.0821	0.2462	0.1331	0.5004	**1.0000**
	Female	0.0936	**0.6492**	**0.6242**	**0.3483**	0.0000	0.1432	0.0673	0.2305	0.0250
a_2	Male	0.8214	0.0000	**0.9564**	**0.8804**	0.8824	0.8537	0.4031	**1.0000**	0.3880
	Female	0.7957	0.8097	0.7191	**0.8708**	**0.9184**	0.5561	0.8663	0.6933	**1.0000**
a_3	Male	**0.8144**	**0.8049**	0.8404	0.4863	0.4936	0.5630	0.0000	0.7682	**0.8616**
	Female	0.3284	**0.7811**	**0.6851**	**0.7908**	0.2109	0.6763	0.3518	0.7616	0.3019
a_4	Male	**0.9204**	0.3285	**0.9834**	0.3316	0.0887	0.4471	0.5162	**0.7266**	0.4059
	Female	0.4476	**0.6366**	**0.6706**	**0.7764**	0.1295	0.3306	0.1633	0.4829	0.3125
a_5	Male	0.6338	**0.7875**	**1.0000**	0.0496	0.0209	0.1190	**0.6628**	0.5740	0.8151
	Female	0.0763	**0.5782**	**0.5824**	**0.7231**	0.0566	0.0654	0.1105	0.0546	0.0000
a_6	Male	**0.7402**	0.3262	**1.0000**	0.1674	0.1850	0.3014	0.1396	**0.5726**	0.6029
	Female	0.2069	**0.4683**	**0.6059**	**0.6679**	0.2675	0.2812	0.1220	0.2230	0.1094
a_7	Male	**0.4944**	**0.6573**	**0.9559**	0.2217	0.0453	0.2287	0.3334	**0.6608**	0.7417
	Female	0.2084	**0.8268**	**0.5848**	**0.9079**	0.0638	0.2599	0.0835	0.1392	0.0393
a_8	Male	**0.9029**	0.7258	**0.8743**	0.8867	0.7438	0.6930	0.0978	0.9492	**0.8711**
	Female	**1.0000**	0.9364	0.7728	0.9352	**1.0000**	0.5301	0.5792	0.6149	**0.7784**
a_9	Male	**0.9133**	0.4761	**1.0000**	0.5744	0.5218	0.6293	0.8760	**0.7133**	0.5595
	Female	0.0000	**0.7982**	**0.6665**	**1.0000**	0.1122	0.5121	0.9150	0.7689	0.5089
a_{10}	Male	**0.6851**	0.4626	**0.9624**	0.2312	0.1527	0.3058	0.4990	**0.5614**	0.6802
	Female	0.1421	**0.7067**	**0.6194**	**0.7710**	0.1088	0.2873	0.0992	0.3502	0.1367
a_{11}	Male	**0.9470**	0.4570	**1.0000**	0.2435	0.1828	0.3707	0.6122	**0.5293**	0.3427
	Female	0.1391	**1.0000**	**0.6377**	**0.7684**	0.1791	0.3918	0.1493	0.2580	0.1177
a_{12}	Male	0.4518	**0.5239**	**0.8301**	0.1880	0.0980	0.3324	**0.4917**	0.4641	0.5547
	Female	0.1992	**0.7119**	**0.6667**	**0.6839**	0.0912	0.2346	0.2035	0.3320	0.2678
a_{13}	Male	0.6214	**0.5915**	**0.7371**	0.2404	0.2066	0.3799	**0.9197**	0.5473	0.4025
	Female	0.4712	**0.9121**	**0.5480**	**0.5837**	0.0000	0.4280	0.3539	0.4033	0.3372

Also, around 73% of males from professional courses suddenly plan to cheat after reading question papers right before starting their exams.

About 79% of males from professional courses seldomly obtain a copy of an exam paper before an exam whereas this percentage increases to almost 100% for the males from humanities and reduces to about 58% for the females from professional courses as well as humanities. Around 72% of females from sciences frequently obtain a copy of an exam paper before an exam and about 66% of males from sciences instantly accept a copy of an exam paper if they have a chance to do so.

Around 74% of males from sciences sometimes try to obtain or accept assistance from any staff or other person during exam whereas this percentage increases to almost 100% for males from humanities and reduces to around 47% & 61% for females from professional courses and humanities, respectively. Also, about 68% of females from sciences often try to obtain/accept assistance from any staff or any other person during exams and about 57% of males from professional courses at sudden accept assistance from any staff member or other person during exams whenever they got a chance to do so.

Around 49% of males from sciences had sometimes lied to their instructor for not appearing in exam/for conducting exam again whereas this percentage increases to around 96%, 83% and 58% for males from humanities, females from professional courses and humanities, respectively. Also, about 91% of females from sciences often lied to their instructor for not appearing in exams/for conducting exams again, and about 66% of males from professional courses sometimes either planned to/at sudden lied to their instructor for not appearing in exam/for conducting exam again.

About 90% of males from sciences had sometimes copied a person's work or thoughts and passed it on as their own whereas this percentage increases to almost 100% for the females from sciences. Also, almost all the females in professional courses often copy a person's work or thoughts and passed it on as their own. Around 95% & 78% of males from professional courses & females from humanities suddenly decide to copy person's work or thoughts and passed it on as their own during their deadlines and about 87% of males from humanities sometimes either plan to/suddenly decide to copy a person's work or thoughts and passed it on as their own during their deadlines.

About 91% of males from sciences sometimes made any other person complete the work assigned to them/complete the work which is given to someone else whereas this percentage increases to almost 100% for the males from humanities and reduces to about 80% & 67% for the females from professional courses & humanities. Almost all the females from sciences more frequently either complete the work assigned to any other person/made others complete the work given to them. Also, around 71% of males from professional courses suddenly made any other person complete the work assigned to them/complete the work which is given to someone else.

About 68% of males from sciences admitted that sometimes they unethically obtain material to complete their work whereas this percentage increases to almost 96% & 71% for the males from humanities & females from professional courses and reduces to about 62% for the females from humanities. Almost 77% of females in sciences always tried to obtain material for their work through unethical ways. Also, around 56% of males from professional courses admitted that they didn't plan but suddenly obtain material to complete their work through illicit ways whenever they have to.

About 95% of males from sciences admitted that they occasionally make excuses to get an extension in their work whereas this percentage increases to almost 100% for the males from humanities & females from professional courses and reduces to about 64% for the females from humanities. Almost 77% of females from sciences always misrepresent their personal/family situation to get an extension for their work. Also, around 53% of males from professional courses admitted that they don't plan but instantly make excuses to get an extension for their work.

Around 52% of males from professional courses had seldom prevented other students from completing their work whereas this percentage increases to about 83%, 71% & 66% for the males from humanities, females from professional courses & females from humanities, respectively. Also, about 68% of females in sciences often prevent other students from completing their work and around 49% of the males in sciences suddenly prevent others from completing their work whenever they have a chance to do so.

Around 59% of males from professional courses had rarely forged a faculty/family/

Friend's signatures on permission form or add/drop form whereas this percentage increases to about 72% & 91% for the males from humanities & females from professional courses, respectively and reduces to about 55% for females from humanities. Also, about 58% of females from sciences always copy another friend/family/faculty's signatures and around 92% of the males from sciences instantly forged a faculty/family/friend's signature whenever they have to.

4. Conclusion

This paper has considered the problem of heterogeneity across individuals in LC models which results in the dependency of latent class probabilities on the grouping variable and provides a method of performing multi-group LCA using the poLCA. The poLCA package is not able to describe the effect of different groups of respondents on the results of LCA, so, as a remedy for this problem, we have performed LCA on different groups of individuals separately. By this technique, we can obtain the conditional probabilities and membership probabilities given different groups of individuals.

From the results of LCA, it can be observed that there are three classes of the latent variable which have been identified as "occasional cheaters(42%)," "persistent cheaters (6%)," and "instantaneous cheaters(52%)." Also, from the multi-group LCA, it is found that among the self reported cheaters around 47% of the males are instantaneous cheaters whereas this percentage increases to 57% for the females. It can also be observed that around 64% of the students from professional courses are instantaneous cheaters whereas this percentage reduces to 53% for the students from humanities and around 57% of the students from professional courses are occasional cheaters.

From multigroup LCA with gender-stream, it is observed that around 66% of males from sciences are persistent cheaters whereas this percentage reduces to around 46% for the males from professional courses and increases to about 73% for the males from humanities. Also, around 46% of females in sciences are instantaneous cheaters whereas this percentage increases to about 57% for the females from professional courses and reduces to around 37% for the females from humanities.

Next, we have computed conditional item response probabilities for each multi-group LC model corresponding to each of the manifest variables and it is noted that the males from sciences and humanities are mostly occasional cheaters, and males from professional courses are either occasional or instantaneous cheaters depending upon the types of cheating. Also, females from sciences are persistent cheaters whereas females from professional and humanities are mostly occasional cheaters.

We recommend the use of poLCA for analyzing the heterogeneity of individuals among the latent classes, which in this case is supported by the case of academic cheaters.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Sunil Kumar http://orcid.org/0000-0003-0249-8415
Apurba Dabgotra http://orcid.org/0000-0002-8056-7239
Diganta Mukherjee http://orcid.org/0000-0002-0842-8314

References

Clogg, C. C., & Goodman, L. A. (1985). Simultaneous latent structural analysis in several groups. In N. B. Tuma (Ed.), *Sociological Methodology* (pp. 81–110). Jossey-Bass Publishers.

Collins, L. M., & Lanza, S. T. (2010). *Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral and Health Sciences*. Wiley.

Genereux, R. L., & McLeod, B. A. (1995). Circumstances Surrounding Cheating: A Questionnaire Study of College Students. *Research in Higher Education*, *36*(6), 687–704.

Goodman, L. A. (1974). Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika*, *61*(2), 215–231.

Haberman, S. J. (1979). *Analysis of Qualitative Data* (1979, Vol. 2). Academic PressNew Developments.

Hagenaars, J. A. (1990). *Categorical Longitudinal Data – loglinear Analysis of Panel, Trend and Cohort Data*. Sage.

Hagenaars, J. A., & McCutcheon, A. L. (2002). *Applied Latent Class Analysis*. Cambridge University Press.

KankaraŠ, M., & Moors, G. (2009). Measurement equivalence in solidarity attitudes in Europe. Insights from a multiple group latent class factor approach. *International Sociology*, *24*(4), 1–23.

Kankaraš, M., Moors, G., & Vermunt, J. K. (2010). Testing for Measurement Invariance with Latent Class Analysis. In E. Davidov, P. Schmidt, & J. Billiet (Eds.), *Cross-Cultural Analysis: Methods and Applications* (pp. 359–384). Routledge.

Lazarsfeld, P. F. (1950). The logical and mathematical foundation of latent structure analysis and the interpretation and mathematical foundation of latent structure analysis. In S. A. S. er (Ed.), *Measurement and Prediction* (pp. 362–472). Princeton University Press.

Lazarsfeld, P. F., & Henry, N. W. (1968). *Latent Structure Analysis*. Houghton Mifflin.

Lin, T. H., & Dayton, C. M. (1997). Model Selection Information Criteria for Non-Nested Latent Class Models. *Journal of Educational and Behavioral Statistics*, *22*(3), 249–264. https://doi.org/10.2307/1165284

McCutcheon, A. (2002). Asic concepts and procedures in single- and multiple-group latent class analysis. In J. Hagenaars & A. McCutcheon (Eds.), *Applied latent class analysis* (pp. 56–88). Cambridge University Press.

Shaughnessy, M. F. (1988), *The Psychology of Cheating Behavior*. https://files.eric.ed.gov/fulltext/ED303708.pdf

Vermunt, J. K., & Magidson, J. (2005). Factor analysis with categorical indicators: A comparison between traditional and latent class approaches. In A. Van der Ark, M. A. Croon, & K. Sijtsma (Eds.), *New developments in categorical data analysis for the social and behavioral sciences* (pp. 41–62). Erlbaum.

Appendices

A Sample size calculation

The total population size in our study is 66,091. We have followed the same methodology of selecting sample as NSSO does in its survey, so, we selected sample with 95% confidence level, 2.21% margin of error and 50% of population share, using:

$$n = \frac{\frac{z^2 \times p(1-p)}{e^2}}{1 + \frac{z^2 \times p(1-p)}{e^2 N}}$$

where, n = sample size, z = z-score, p = population proportion, e = margin of error, N = population size.

B Detailed description of the sampling frame

For this study, we have formed stratum at district level. Within each district of Jammu division, nine main strata are formed (one for each district). However, within the Jammu district, different government degree colleges, government engineering college, private engineering colleges, private BEd colleges and private law colleges formed a separate basic main stratum and the remaining off campuses of University of Jammu i.e., Ramnagar campus and Bhaderwah campus and campus of University of Jammu, itself was considered as another basic stratum. Then, different sub-stratification are done for the courses offered in each of the selected campuses and colleges. Also, only three (3) different streams are considered as the courses offered by these institute. They were: Sciences, Professional courses and Humanities. Within each stream (course) offered by University of Jammu, its affiliated colleges and its different off campuses, the respective sample size was allocated to the different strata in proportion to the number of students enrolled in that stream.

The above discussed technique has employed on each of the district except for Jammu District. For Jammu District, 75% of the remaining sample, after selected from the other districts and off campuses, had selected from the campus of the University of Jammu, using proportional allocation. And the rest of sample had selected from the remaining colleges of Jammu district using proportional allocation. Finally, For the different districts (from each sub stratum) required number of sample were selected by simple random sampling without replacement (SRSWOR) procedure.

C Estimated conditional item response probabilities of responding 'Yes' to the indicator items for each of the latent variables

LCA with covariates			
Indicators	Occasional cheaters	Persistent cheaters	Instantaneous cheaters
a_1	0.5031	0.9208	0.0470
a_2	0.6765	0.9530	0.8734
a_3	0.6524	1.0000	0.4183
a_4	0.5634	0.8724	0.2874
a_5	0.3954	1.0000	0.0304
a_6	0.4293	0.9870	0.1661
a_7	0.4746	0.9751	0.1029
a_8	0.6907	0.9871	0.7658
a_9	0.6466	0.9797	0.5114
a_{10}	0.5077	0.8925	0.1525
a_{11}	0.5674	0.9468	0.1590
a_{12}	0.5117	0.7281	0.1882
a_{13}	0.5288	0.7153	0.3156
Males			
Indicators	**Occasional cheaters**	**Persistent cheaters**	**Instantaneous cheaters**
a_1	0.5703	1.0000	0.0852
a_2	0.6858	0.9604	0.8873
a_3	0.6524	1.0000	0.4768
a_4	0.5778	0.8979	0.2970
a_5	0.4756	1.0000	0.0355
a_6	0.4677	0.9736	0.1781
a_7	0.4729	1.0000	0.1481
a_8	0.7095	0.9809	0.7709
a_9	0.6437	1.0000	0.5701
a_{10}	0.5343	0.8874	0.1754
a_{11}	0.5467	0.9296	0.2357
a_{12}	0.5086	0.7043	0.1739
a_{13}	0.5247	0.6961	0.2750
Females			
Indicators	**Occasional cheaters**	**Persistent cheaters**	**Instantaneous cheaters**
a_1	0.4339	0.8287	0.0378
a_2	0.6406	1.0000	0.8718
a_3	0.6365	1.0000	0.3972
a_4	0.5442	0.8602	0.2905
a_5	0.3536	0.9628	0.0182
a_6	0.3968	1.0000	0.1635
a_7	0.4942	0.9264	0.0728
a_8	0.6743	1.0000	0.7591
a_9	0.6298	0.9362	0.4912
a_{10}	0.4900	0.8771	0.1445
a_{11}	0.5757	1.0000	0.1250
a_{12}	0.5312	0.7674	0.1944
a_{13}	0.5462	0.7280	0.3395
Sciences			
Indicators	**Occasional cheaters**	**Persistent cheaters**	**Instantaneous cheaters**
a_1	0.0930	0.1033	0.5767
a_2	0.6852	0.8605	0.8308
a_3	0.2971	0.3740	0.7876
a_4	0.2500	0.3349	0.8416
a_5	0.2578	0.0616	0.6477
a_6	0.1781	0.1563	0.6469
a_7	0.1101	0.1963	0.6202
a_8	0.0000	1.0000	0.8969
a_9	0.8932	0.4188	0.9035
a_{10}	0.2122	0.1523	0.7211
a_{11}	0.3328	0.1513	0.8582

(*Continued*)

LCA with covariates			
Indicators	Occasional cheaters	Persistent cheaters	Instantaneous cheaters
a_{12}	0.2966	0.1803	0.5460
a_{13}	0.5110	0.3359	0.6446
Professionals			
Indicators	**Occasional cheaters**	**Persistent cheaters**	**Instantaneous cheaters**
a_1	0.2574	0.8196	0.0000
a_2	0.7969	0.6979	0.9282
a_3	0.6902	0.8400	0.3425
a_4	0.4803	0.6931	0.0724
a_5	0.1952	0.8098	0.0000
a_6	0.2635	0.5616	0.2623
a_7	0.2034	0.9584	0.0390
a_8	0.7114	1.0000	0.9032
a_9	0.7091	0.6834	0.2706
a_{10}	0.3407	0.7766	0.0475
a_{11}	0.4233	0.6554	0.0303
a_{12}	0.2945	0.6357	0.0682
a_{13}	0.4686	0.6884	0.0247
Humanities			
Indicators	**Occasional cheaters**	**Persistent cheaters**	**Instantaneous cheaters**
a_1	0.9562	0.0649	0.5826
a_2	1.0000	0.8688	0.6443
a_3	1.0000	0.4633	0.6700
a_4	0.9282	0.3259	0.5634
a_5	0.9938	0.0219	0.4627
a_6	1.0000	0.1792	0.5130
a_7	0.9743	0.1088	0.5180
a_8	0.9726	0.7121	0.7016
a_9	1.0000	0.5474	0.6059
a_{10}	1.0000	0.1840	0.5419
a_{11}	1.0000	0.2142	0.5527
a_{12}	0.7431	0.2209	0.5915
a_{13}	0.7390	0.3499	0.4847

Exploring Academic Dishonesty Among University Students: University of Jammu, India

Sunil Kumar, Apurba Vishal Dabgotra

Department of Statistics, University of Jammu, Jammu, J&K, India

Corresponding author: Sunil Kumar, Email: sunilbhougal06@gmail.com

Over the years, academic cheating continues to be an endemic issue that has always been a threat to academic honesty and social values. Academic dishonesty among the students is a perplexing phenomenon, that exists in every stage of our education system, especially in colleges and universities. College and university administrators admit that academic dishonesty is an issue on campus but they often lack in preparing effective policies and procedures to monitor it and to deal with it. In addition, indecisive perceptions regarding academic dishonesty has adverse effects on paradoxical situation of education. The current study provides the details about the causes that motivate the students to cheat and describes the different forms of cheating practices performed by the students. The purpose of this study is to define the students of University of Jammu what is meant by academic cheating, to determine the factors associated with cheating behaviour and to classify of the students according to their cheating behaviour. From the current study, the main reasons due to which students go for cheating are not knowing/understanding the study material, performance pressure, inadequate exam preparations, etc. Also, around 42%, 6% and 52% of the respondents are found to be occasional, persistent and instantaneous cheaters, respectively. We suggest that students must understand that cheating is wrong not only for the society but also for their own knowledge because by indulging in cheating students prevent them from learning what they are studying and hence, deteriorate the intellectual human resource of the country.

Keywords:Academic cheating, Latent Class Analysis (LCA), AIC, BIC.

2023. In Saroj Hiranwal & Garima Mathur (eds.), *Artificial Intelligence and Communication Technologies*, 629–647. Computing & Intelligent Systems, SCRS, India. https://doi.org/10.52458/978-81-955020-5-9-60

1 Introduction

Academic dishonesty is an escalating phenomenon that is plaguing educational institutions around the world. Infact, it has become an inescapable activity, especially, in colleges and universities, where grades earned, directly effect the academic careers of students for many years to come. The rising pressure to get the best grades in school, get into the best college, and land the best paying job is a cycle that has made academic dishonesty increase exponentially. The subject of academic cheating has attracted the attention of not only academics but also public communities. Recent studies have proven that the issues of academic cheating among undergraduates have increased along the years. For example, 76 percent of the students confessed to having involved in academic cheating [16].

Rajendran defined cheating as an activity which is performed to complete a work in an unethical way by a person when he does not know how to do that work in a legal way ([27]). However, this is not the only way of defining cheating; different people have their different perception about cheating like [11] defined cheating as changing one or more answers when correcting own examination and/or not marking two or more in correct responses, [23] defined cheating as an immoral activity in the academic environment, [28] defined cheating as an act of using any means of unfair and unjust privileges that include: lying, concealing the truth, deceive, deceit and violation of trust to achieve something. But academic cheating/academic dishonesty may be defined as the students' behaviour of contradicting the fundamental values of their academic lives [11]. Academic cheating can be of various types, suggested by [11] and [19], such as copying from another's person, stealing examination papers and lecture notes, using prohibited material like crib notes, deception, sabotaging, impersonation, forgery, plagiarism, fabrication, data manipulation, padding bibliographies, and many more.

Various theoretical perspectives have been applied for understanding cheating. [25] examined the applicability of several theories of divergence to cheating, including deterrence theory, rational choice theory, social bond theory, etc. Thereafter, [4] proposed the integrated model of cheating, which shows that students' intentions to engage in disfunctional behaviours may be influenced by attitudes, subjective norms, perceived behavioural control and moral obligations. Some studies focused on the relationship between motivation and cheating behavior, some focused more on the relationship between socio-demographic factors and cheating behavior (e.g., [15]). Others (e.g., [8]) have carried out studies to determine what causes students to cheat. The existence of academic cheating has always been a major concern for various researchers. Many studies have

been conducted in order to identify various types of academics cheating among college students and this can be achieved by using anonymous questionnaire distributed or mailed to the students. The estimates of students that are involved in academic cheating during their college lives, ranges from 49% for marketing students [32] to 88% for premedical students [30]. Dishonesty in an academic setting has been a consistent and paramount problem for many years at all educational levels [9], and it is a serious educational issue([26]; [18]). Considerable progress has thus, been made in identifying factors that influence cheating behaviour. [3] have found that age and self esteem are the two major factors that are positively correlated with the cheating behaviour of students. [17] found that higher socio-economic status along with stress, depression and family crisis are the major factors that promote students to cheat in their academic lives.

Various guidelines for controlling cheating have also been proposed like [12], [13] and [1] suggested that cheating can be controlled by making it as difficult as possible for students to cheat, [14] suggested that the cheating can be avoided by discussing the consequences of cheating with the students, etc. Despite these findings and recommendations, the prevalence of cheating is on the rise ([6]; [5]). [22] suggest that academic integrity should be strongly assumed as an institutional concern, instead of just students' responsibility. Moreover, through a collaboration approach and using workshops and open educational resources settled to address paraphrasing, summarizing and quotation, [22] concluded that "better collaboration and co-operation among faculty staff, learning advisors and librarians is therefore essential" [22]. Still there are many questions that remain to be answered concerning the nature, cause and type of academic cheating.

The purpose of current study is to define the students what is meant by academic cheating, to identify the factors/circumstances associated with cheating behaviour of students and to classify of the students according to their cheating behaviour. The paper is structured as follows: The next section (i.e. section 2) explains the sample design in detail followed by data description in section 3. Section 4 explains the framework of LCA methodology along with the statistical hypothesis and results of the analysis in the section 5. Finally, we discuss conclusions of the study in Section 6.

2 Sample Design

Our target population comprised of all the undergraduates students enrolled in 3rd semester and above, BEd students, masters students of University of Jammu and its affiliated Colleges. So, the total population size in our study was 66091. We

had followed the NSSO methodology to select the sample with 95% confidence level, 2% margin of error and 50% of population share, using formulae:

$$n = \frac{\frac{z^2 \times p(1-p)}{e^2}}{1 + \frac{z^2 \times p(1-p)}{e^2 N}} \tag{60.1}$$

where, n= sample size, z= z score, p= population proportion, e= margin of error, and N= size of population size.

Thus, a sample of 2317 had selected for the collection of data. And we had selected sample based on the multi-stage stratified random sampling design. The first stage units (FSU) comprises of the main campus, off campuses and different colleges affiliated to the Jammu University. The ultimate stage units (USU) comprises of different courses offered by these institutes. Then, we apply proportional allocation for selecting sufficient number of sample units from each of the USU. For the sampling frame, we had collected information from the department of Statistical Planning and Research Unit of University of Jammu.

Stratum had formed at district level. Within each district of Jammu division, nine basic strata were formed (one for each district). However, within the Jammu district, different government degree colleges, government engineering college, private engineering colleges, private BEd colleges and private law colleges formed a separate basic stratum and the remaining off campuses of University of Jammu i.e., Ramnagar campus and Bhaderwah campus and campus of University of Jammu, itself was considered as another basic stratum. Different sub-stratification were done for the courses offered by each of the selected campuses and colleges. Seven (7) different streams were considered as the courses offered by these institute. They were: Science, Technology, Commerce, Arts, Management, Law and BEd. Within each stream(course) offered by University of Jammu, its Affiliated colleges and its different off campuses, the respective sample size was allocated to the different strata in proportion to the number of students enrolled in that stream.

The above discussed technique has employed on each of the district except for Jammu District. For Jammu District, 75% of the remaining sample, after selected from the other districts and off campuses, had selected from the campus of the University of Jammu, using proportional allocation. And the rest of sample had selected from the remaining colleges of Jammu district using proportional allocation. Finally, For the different districts (from each sub stratum) required number of sample were selected by simple random sampling without replacement (SRSWOR) procedure.

3 Data Description

A descriptive design is used to study cheating behaviour of the students. So, a questionnaire, comprises of 46 closed ended questions related to factors leading to students' cheating, types of cheating and other related questions, was designed to analyze the cheating behaviour of the students. The survey period of this study was from March to September 2019. The required information was collected from a selected number of students by visiting their respective departments, colleges or institutes, personally and by asking them to fill in an anonymous questionnaire based on different cheating behaviour. But due the disturbances in the state during the survey period, it was not possible to collect the data from the highly disturbed areas such as Doda, Kishtwar and Rajouri. So, the data was collected from only 1906 students which was the relevant sample size for our study. The whole questionnaire is divided into following 5 major sections.

Respondent's Details : This section tracks the record of the demographic characteristics of the students which includes gender, age, religion, status and stream.

General Questions : This section consist of general questions about cheating in order to get the general perception of students on academics cheating and to get idea from the students about the main reason why students cheat during the exams?

Situational Aspects : This section consist of questions about the situational factors that can influence the students to cheat during the exams.

Personal Aspects : This section consist of the questions about the personal factors that can influence the students to cheat during the exams.

Types Of Cheating : This section comprises of the questions about the types of cheating students are involved in. This section was prepared with the objective of making students aware of the types of academics cheating and to know about their own cheating behaviour.

4 Methodology

4.1 Latent Class Analysis

We have certain phenomenon that often cannot be directly observed or to analyze certain phenomenon not all variables can be measured directly. So, latent variable modeling, in which the value of the latent variable (unknown variable) cannot be directly measured, rather its value is deduced from observed (manifest) variables, can be used in those cases. Latent variables may be defined as an unobserved random variables which are hidden from us [31] and are unknown to us in any particular study, whereas, manifest variables may defined as variables which are the observable and are designed specially to measure the unknown

latent variable. Indicator variables are the manifest variables which can measure the unknown latent variable. The value of unknown latent variable can be estimated on the basis of the responses made by the individuals to the different indicator variables. We also have a third variable, referred to as a grouping variable such as gender (G), which is used to identify an individual's membership in two or more population subgroups. The grouping variable is based largely on subjective criteria, by considering whether the assumptions are plausible for a it or not.

Latent Class Analysis (LCA) is appropriate when the latent variable and all the indicator variables are discrete in nature. It establishes a relationship between a set of observed discrete variables (manifest variables) and a set of unknown discrete variables (latent variables). LCA is a methodology that allows us to identify hidden population subgroups/classes. A class of any latent variable is specified by pattern of responses made to the different manifest variables by the respondents, in terms of conditional probabilities. These probabilities show the possibility that the latent variables can take any particular values. In this approach we envisage a relationship between discrete indicator variables and discrete latent variables across different number of groups. We have used poLCA [21] package of R software for performing the Latent class analysis (LCA).

4.1.1 Latent Class Models

LCA models comprises of two types of probabilities which include the probability indicating the likelihood of a response by respondents in each of the classes and the probability representing the latent class size or the proportion of individuals who are members of a particular latent class.

Following the notation used by [21], suppose we have J polytomous categorical manifest variables (the observed variable) each of which contain K_j possible outcomes, for individuals $i = 1, 2, 3, ..., N$. Let Y_{ijk} be the observed values of the J manifest variables such that

$$\left\{ \begin{array}{ll} Y_{ijk} = 1: & \text{if } i^{th} \text{ respondent give the } k^{th} \text{ response to the } j^{th} \text{ variable} \\ Y_{ijk} = 0: & \text{otherwise} \end{array} \right\}$$

where j=1,2,...,J and k=1,2,...,K_j.

The LC models approximates the observed joint distribution of the manifest variables as the weighted sum of a finite number, R, of constituent cross-classification tables. Let π_{jrk} denote the cross-conditional probability that an observation in

class r=1,2,...,R produces the k^{th} outcome on the j^{th} variable with $\sum_{k=1}^{K_j} \pi_{jrk} = 1$.

Let p_r be the prior probabilities of latent class membership, as they represent the unconditional probability that an individual will belong to each class before taking into account the responses Y_{ijk} provided on the manifest variables. The probability that an individual i in class r produces a particular set of J outcomes on the manifest variables, assuming conditional independence of the outcomes Y given class membership, is the product

$$f(Y_i; \pi_r) = \prod_{j=1}^{J} \prod_{k=1}^{K_j} (\pi_{jrk})^{Y_{ijk}}, \tag{60.2}$$

The probability density function across all classes is the weighted sum

$$f(Y_i|\pi, p) = \sum_{r=1}^{R} f(Y_i; \pi_r) = \sum_{r=1}^{R} P_r \prod_{j=1}^{J} \prod_{k=1}^{K_j} (pi_{jrk})^{Y_{ijk}}, \tag{60.3}$$

The parameters P_r and π_{jrk} are estimated by the latent class model.

Given estimates $\hat{P}_r$ and $\hat{\pi}_{jrk}$ of P_r and π_{jrk} respectively, the posterior probability that each individual belongs to each class, conditional on the observed values of the manifest variables, are calculated by

$$\hat{P}(r_i|Y_i) = \frac{\hat{p}_r f(Y_i; \hat{\pi}_r)}{\sum_{q=1}^{R} \hat{p}_q f(Y_i; \hat{\pi}_q)'} \tag{60.4}$$

where $r_i \in (1, 2, ..., R)$. It is important that the condition $R\sum_j (K_j - 1) + (R-1) \leq n$ on the number of parameters should hold. Also, $R\sum_j (K_j - 1) + (R-1) \leq (3^{10} - 1)$ i.e. one fewer than the total number of cells in the cross-classification table of the manifest variables, as then the latent class model will be unidentified. Under the assumptions of multinomial distribution, the log likelihood function can be given as:

$$lnL = \sum_{i=1}^{n} ln \sum_{r=1}^{R} p_r \prod_{j=1}^{J} \prod_{k=1}^{K_j} (\pi_{jrk})^{Y_{ijk}}, \tag{60.5}$$

LCA not only builds a classification model but it also explain a relation of the class membership to explanatory variables by including covariates [33] in the model. Grouping variables can be used in LC models in order to model the unexplained heterogeneity in the data. In that case latent class membership probabilities are predicted by covariates through a logistic link.

4.1.2 Parameter Estimation and Model Selection

The unknown parameters of the LC models can be estimated by maximizing (60.5) with respect to p_r and π_{jrk},using the expectation-maximization (EM) algorithm ([7], [24] and [21]). The EM algorithm, begin with arbitrary initial values of $\hat{p}_r$ and $\hat{\pi}_{jrk}$, and denote them $\hat{p}_r^0$ and $\hat{\pi}_{jrk}^0$. The expectation step, calculate the missing class membership probabilities using equation (60.4), substituting $\hat{p}_r^0$ and $\hat{\pi}_{jrk}^0$ in place of $\hat{p}_r$ and $\hat{\pi}_{jrk}$. The maximization step, update the estimates of the parameters by maximizing the log likelihood function given these posterior $\hat{P}(r_i|Y_i)$, with

$$p_r^{new} = \frac{1}{N}\sum_{i=1}^{N} P(r_i|Y_i) \quad and \quad \hat{\pi}_{jr}^{new} = \frac{\sum_{i=1}^{N} Y_{ij} P(r_i|Y_i)}{\sum_{i=1}^{N} P(r_i|Y_i)} \tag{60.6}$$

as the new prior and class conditional outcome probabilities, respectively; $\hat{\pi}_{jr}^{new}$ is the vector of length K_j of class-r conditional outcome probabilities for the j^{th} manifest variable; and Y_{ij} is the $N \times K_j$ matrix of observed outcome Y_{ijk} on that variable. The algorithm repeats these steps several times until the overall log-likelihood reaches a local maximum and further increments are less than some arbitrarily small value.

Different LCA models have different number of latent classes. Usually, models with more parameters (i.e, more latent classes) provide a better fit, and more parsimonious models tend to have a somewhat poorer fit. So, there is always very close agreement between goodness of fit and parsimony of the latent class models. We can test the goodness of fit of an estimated LCA models by the Pearson Chi-square(χ^2) or the Likelihood Ration Chi-square(L^2) . However, the likelihood ratio Chi-square test, although extensively used in statistical literature, has a number of important limitations. These limitations can be controlled by making use of several information criteria, such as the Akaike information criterion (AIC) [2] and Bayesian information criterion (BIC) [29], each of which is designed to penalize models with larger numbers of parameters. LC models with different number of latent classes are compared and a model with lower AIC and BIC is selected.

4.2 Statistical Hypothesis

The following null hypothesis were considered and tested, at the 5% level of significance, in the current study :

$H_{\circ 1}$: Cheating is independent of gender.

$H_{\circ 2}$: Cheating is independent of current Status of the student.

$H_{\circ 3}$: Cheating is independent of stream of the student.

$H_{\circ 4}$: Cheating is independent of students' perception of academic cheating.

$H_{\circ 5}$: Cheating is independent of students' understanding of material.

$H_{\circ 6}$: Cheating is independent of type of exam given by the teacher to the students.

$H_{\circ 7}$: Cheating is independent of performance pressure on the students.

$H_{\circ 8}$: Cheating is independent of the students' perception about humiliation due to failure.

$H_{\circ 9}$: Cheating is independent of the students' confidence lacking.

$H_{\circ 10}$: Cheating is independent of the students' parental pressure.

$H_{\circ 11}$: Cheating is independent of the students' attitude towards grades.

$H_{\circ 12}$: Cheating is independent of sitting plan of the students in the exam.

$H_{\circ 13}$: Cheating is independent of punishment severity in the educational atmosphere.

$H_{\circ 14}$: Cheating is independent of students' inadequate exam preparations.

$H_{\circ 15}$: Cheating is independent of students' ineffective time management skills.

$H_{\circ 16}$: Cheating is independent of students' habit of laziness.

$H_{\circ 17}$: Cheating is independent of students' perception about instructor vigilance.

$H_{\circ 18}$: Cheating is negatively correlated to the students' subject liking.

$H_{\circ 19}$: Cheating is negatively correlated to the students' interest in the subject.

$H_{\circ 20}$: Cheating is negatively correlated to the students' instructor liking.

$H_{\circ 21}$: Cheating is independent of the student's perception that cheating inevitable.

$H_{\circ 22}$: Cheating is independent of the student's perception of peer influence.

5 Results

We had conducted a pilot survey on 9 departments of the University of Jammu, in order to test the reliability and validity of our questionnaire, before conducting the full survey. The Value of Cronbach Alpha coefficient in this case comes out to be 0.735, which indicates a high reliability according to [10]. Cohen's Kappa index, in this case was 0.61, which shows that face validity holds for our questionnaire. We had also performed factor analysis utilizing principle component analysis with varimax rotation method for testing construct validity and correlation analysis for testing the criterion validity. Both the validities (criterion and construct) holds in this study.

Our data set consists of 1906 observations, but out of that 7 were deleted due to non response and a total of 1899 responses were considered for the further analysis. These responses were, then tested for the identification of missing data values. We had used multiple imputation method to deal with missing data values. With singular imputation methods, mean, median, or some other statistics is used to impute the missing values. However, using single values carries with it a level of uncertainty about which values to impute. Multiple imputation narrows uncertainty about missing values by calculating several different options. We had used SPSS software for performing multiple imputations on over dataset. The imputed dataset is thus, used for the statistical analysis.

5.1 Testing of Hypothesis

Table 1 provides the summary of the hypothesis testing. From table 1 it is clear that the cheating behaviour of the students is independent of the their current-status, type of exam given by the teacher, perception about humiliation due to failure, confidence lacking, parental pressure, attitude towards grades, punishment severity in the educational atmosphere, ineffective time management skills, habit of laziness and perception about instructor vigilance.

Also, Cheating behaviour of students depends on their gender and stream. It also depends on their perception about academic cheating, understanding of material, performance pressure, sitting plan in the exams, inadequate exam preparations, belief that cheating is inevitable and peer influence. Cheating behaviour of the students is found to be positively correlated to the their subject liking or disliking, interest in the subject and instructor liking.

Table 1: Hypothesis test summary

S. No.	Hypothesis	χ^2	Significant value	Decision
1	$H_{\circ 1}$	17.574	0.000	Reject the null hypothesis
2	$H_{\circ 2}$	5.504	0.138	Accept the null hypothesis
3	$H_{\circ 3}$	17.795	0.013	Reject the null hypothesis
4	$H_{\circ 4}$	8.309	0.004	Reject the null hypothesis
5	$H_{\circ 5}$	20.702	0.000	Reject the null hypothesis
6	$H_{\circ 6}$	6.919	0.140	Accept the null hypothesis
7	$H_{\circ 7}$	38.973	0.000	Reject the null hypothesis
8	$H_{\circ 8}$	1.735	0.188	Accept the null hypothesis
9	$H_{\circ 9}$	0.208	0.648	Accept the null hypothesis
10	$H_{\circ 10}$	2.023	0.155	Accept the null hypothesis
11	$H_{\circ 11}$	2.763	0.096	Accept the null hypothesis
12	$H_{\circ 12}$	9.199	0.002	Reject the null hypothesis
13	$H_{\circ 13}$	1.098	0.295	Accept the null hypothesis
14	$H_{\circ 14}$	15.933	0.000	Reject the null hypothesis
15	$H_{\circ 15}$	2.709	0.100	Accept the null hypothesis
16	$H_{\circ 16}$	0.314	0.575	Accept the null hypothesis
17	$H_{\circ 17}$	0.119	0.731	Accept the null hypothesis
18	$H_{\circ 18}$	6.906	0.009	Reject the null hypothesis
19	$H_{\circ 19}$	23.675	0.000	Reject the null hypothesis
20	$H_{\circ 20}$	14.010	0.000	Reject the null hypothesis
21	$H_{\circ 21}$	6.298	0.012	Reject the null hypothesis
22	$H_{\circ 22}$	11.629	0.001	Reject the null hypothesis

5.2 Latent Class Analysis

For the current study, the proportion of female in the sample is 0.60 and male is 0.40; whereas in population this proportion was 0.66 and 0.34. Also, the proportion of the students belonging to science, Technology, Commerce, Arts, Management, Law and other streams is 0.24, 0.08, 0.03, 0.45, 0.03, 0.05 and 0.12, respectively.

5.2.1 Latent Class Models/ Path Models

The path model diagram is the graphical method of displaying the causal relationships among variables in a LCA. A detailed description of the variable used

in the current study for performing LCA is in Appendix A. Given the relatively large number of observed variables measuring the latent variable and the number of response categories per variable, the number of parameters is fairly high. For this reason larger models were not considered. But out of those models, only 4 models provide the efficient results, consequently we proposed following 4 models for estimating the cheating behaviour of the students using LCA.

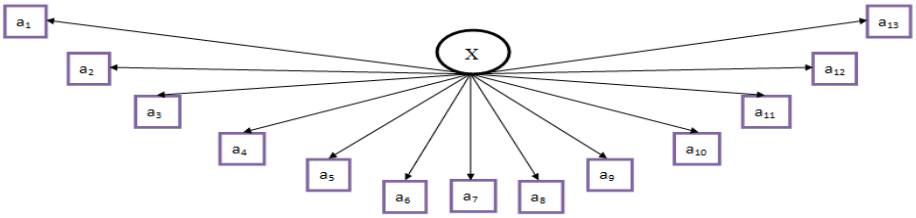

Figure 1: Model 0

Model 0 : Figure 1 is the path model for model 0. This model is the simple LC model without any grouping variable, preserving the assumption of local independence. It will estimates the Cheating behaviour (X) on the basis of the individuals' response pattern to the different indicator variables.

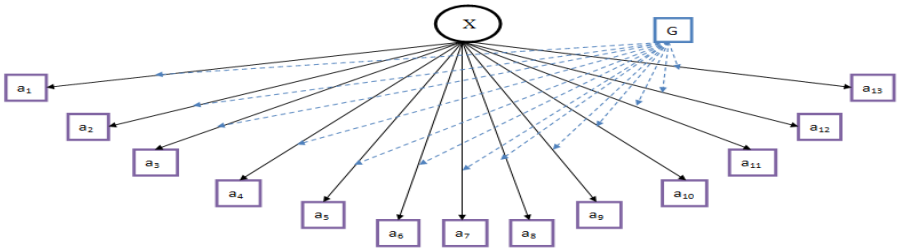

Figure 2: Model 1

Model 1 : Figure 2 is the path model for model 1. This model represents the variation in the indicator variables with the inclusion of grouping variable, gender (G) and estimates the Cheating behaviour on the basis of the individuals' response pattern to the different indicator variables through grouping variable.

Model 2 : Figure 3 is the path model for model 2. This model represents the variation in the indicator variables with the inclusion of grouping variable, stream (S) and will test the influence of stream on the response pattern of the individual to estimate the cheating behaviour.

Model 3 : Figure 4 is the path model for model 3. This model is a complex representation of the variation in the indicator variables on the inclusion of two grouping variables i.e., gender (G) and stream (S). It will estimate the cheating behaviour from the response patterns of the respondents to the different indicator variables under the influence of the gender and stream of the respondents.

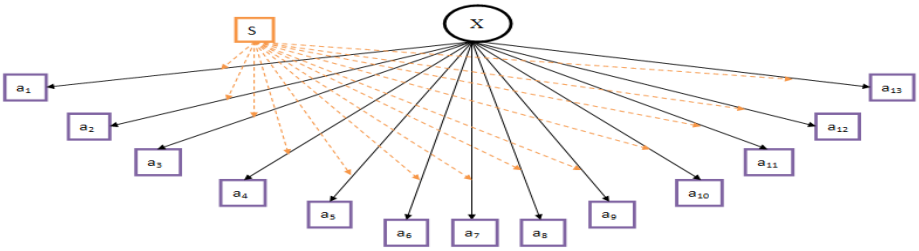

Figure 3: Model 2

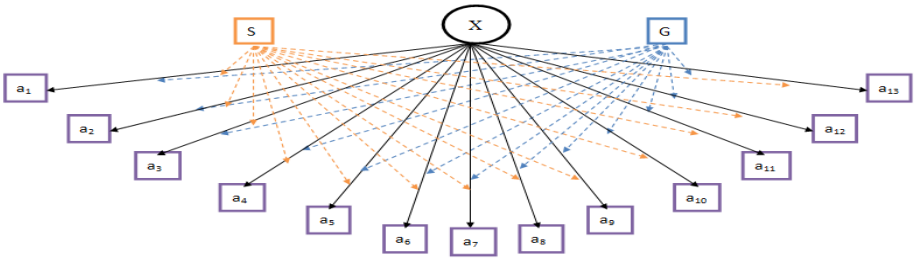

Figure 4: Model 3

5.2.2 Selection of best fitted model

Our data consist of 13 indicator variables which was used for predicting the cheating behaviour of the students. We had incorporated 2 grouping variables with these variables in order to estimate the LC models. As a result, we had to select one best model from the list of the proposed models, which provides the better fit and optimum number of latent classes. Table 2 provides necessary model statistics for different competing models, using poLCA package of R software. Best fitted model is selected on the basis of BIC value [20]. The model with lowest BIC value is always preferred since it provides the best balance between the two factors namely, model fit and model parsimony. From table 2, it can be seen that BIC value of model 3 is lowest i.e. 14954.7 . Also, it's AIC and LL value (14725.37 and -7315.686, respectively) is also satisfactory. Hence, model 3 is used for further analysis.

Table 2: Model diagnostics

Model	d.f.[1]	No. of parameters	LL[2] value	AIC[3]	BIC[4]
Model 0	931	41	-7344.114	14770.23	14970.28
Model 1	929	43	-7330.693	14747.38	14957.2
Model 2	929	43	-7330.405	14746.81	14956.62
Model 3	925	47	-7315.686	14725.37	14954.7

[1] $d.f.$: *Degrees of freedom of error term*
[2] LL : *Log likelihood value*
[3] AIC : *Akaike information criterion*
[4] BIC : *Bayesian information criterion*

5.2.3 Selection of optimum number of classes

A latent class or class of any latent variable is specified by pattern of responses made to the different manifest variables by the respondents, in terms of conditional probabilities. These probabilities show the possibility that the latent variables can take any particular values depending on the responses of the respondents. It also forms the underlying subgroups of respondents based on the observed attributes. In the present study, classes specify the number of categories into which the responses about the personal cheating behaviour falls. In order to identify optimum number of latent classes, we had performed LCA on the selected model 3 to optimize the number of latent classes. This would help us to find a parsimonious model which provides better fit.

Table 3 provides the goodness of fit statistics of model 3 for different number of latent classes. From table 3 it is clear that the data-set was best fitted for Model 3 with 3 latent classes as the corresponding BIC as well as AIC values of that model were lowest.

Therefore, the underlying latent classes can be identified as "Occasional Cheaters" (latent class 1) which represents the group of students who are occasional cheaters and cheat rarely in their academic lives , "Persistent Cheater" (latent class 2) which represents the students who are frequent cheaters and cheat commonly in tests, exams or assignments and "Instantaneous Cheaters" (latent class 3) represents the group of students who are instant cheaters and cheats in their academic lives whenever got chance to do so.

Table 4 provides the Estimated conditional item response probabilities for each of the indicator variables. 1^{st} sub row of table 4 provides the results for the students who actually admitted that they are involved in the cheating activities and 2^{nd} sub row provides the results for students who have denied for being involved in the cheating.

Figure 5 provides the graphical representation of the class membership probabilities for estimation of the 3 class lc model. Each group of red bars represents the conditional probabilities of the indicator variables given the latent variable.

6 Conclusion and Discussion

Based on the objectives, defined hypothesis and the analysis of collected data, following conclusions are drawn:

- The current study defines students (by means of questionnaire) exactly what is meant by academics cheating and make them aware of different types of cheating in which they are involved knowingly or unknowingly.

- Also, it is observed that the main reasons because of which student opt for cheating are not knowing or understanding the study material as it is quite difficult and boring; performance pressure espe-

Table 3: Goodness of fit statistics of model.

Number of classes (n)	n=2	n=3	n=4
Estimated n-class population shares	0.6489 0.3511	0.4181 0.0659 0.516	0.3632 0.3087 0.2686 0.0595
Predicted n-class memberships	0.6605 0.3395	0.3981 0.0689 0.5329	0.356 0.356 0.2263 0.0617
No. of observations	972	972	972
No. of parameters	30	47	64
Residual degrees of freedom	942	925	908
Maximum log likelihood	-7417.268	-7315.686	-7371.238
AIC	14894.54	14725.37	14770.48
BIC	15040.92	14954.7	14982.76
χ^2	15781.95	11435.73	10963.64

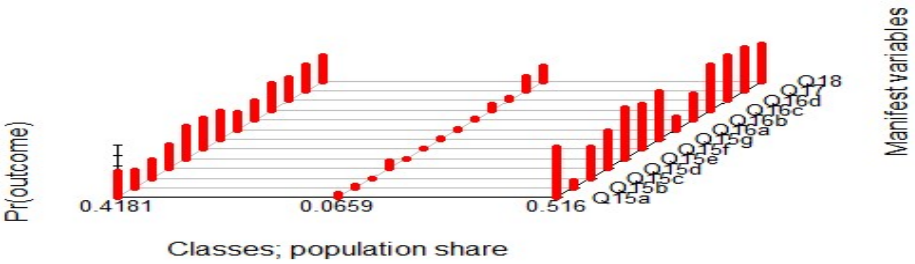

Figure 5: Graphical representation of class membership probabilities

cially because of fear of humiliation due to failure ; inadequate exam preparations; ineffective time management skills; Laziness; disliking of the subject/course and lack of interest in the topic.

- The latent classes, in present study, have been identified as "occasional cheaters", "persistent cheaters" and "instantaneous cheaters", consequently around 42%, 6% and 52% of our respondents are concluded to be occasional, persistent and instantaneous cheaters, respectively.

Academic cheating is a 'disorder' that should be taken seriously to restrain the behaviour of academic dishonesty. It can be challenging to overcome the behaviour of academic dishonesty but an ongoing effort must be taken to lessen its occurrence. The institutions of higher learning should organize programs to promote academic integrity and inculcating an ethical behaviour amongst tertiary students. Students should also be made aware on the negative implications they will receive if they are found to be involved in the academic cheating. Institutions of higher learning should also implement a clear and strict policy on the act of academic dishonesty.

Table 4: Estimated conditional item response probabilities

Indicator variables	Categories of indicators	Latent class 1	Latent class 2	Latent class 3
[t]2*a_1	1	0.5031	0.9208	0.0470
	2	0.4969	0.0792	0.9530
[t]2*a_2	1	0.6765	0.9530	0.8734
	2	0.3235	0.0470	0.1266
[t]2*a_3	1	0.6524	1.0000	0.4183
	2	0.3476	0.0000	0.5817
[t]2*a_4	1	0.5634	0.8724	0.2874
	2	0.4366	0.1276	0.7126
[t]2*a_5	1	0.3954	1.0000	0.0304
	2	0.6046	0.0000	0.9696
[t]2*a_6	1	0.4293	0.9870	0.1661
	2	0.5707	0.0130	0.8339
[t]2*a_7	1	0.4746	0.9751	0.1029
	2	0.5254	0.0249	0.8971
[t]2*a_8	1	0.6907	0.9871	0.7658
	2	0.3093	0.0129	0.2342
[t]2*a_9	1	0.6466	0.9797	0.5114
	2	0.3534	0.0203	0.4886
[t]2*a_{10}	1	0.5077	0.8925	0.1525
	2	0.4923	0.1075	0.8475
[t]2*a_{11}	1	0.5674	0.9468	0.1590
	2	0.4326	0.0532	0.8410
[t]2*a_{12}	1	0.5117	0.7281	0.1882
	2	0.4883	0.2719	0.8118
[t]2*a_{13}	1	0.5288	0.7153	0.3156
	2	0.4712	0.2847	0.6844

References

[1] Aiken, L. S. and West, S. G. (1991). Multiple regression: Testing and interpreting interactions. *Newbury Park*, CA: Sage Publications.

[2] Akaike, H. (1973). Information Theory and an Extension of the Maximum Likelihood Principle. In B Petrov, F Csake (eds), *Second International Symposium on Information Theory*, pp. 267-281 Akademiai Kiado, Budapest, Hungary.

[3] Alarape, A. I. and Onakoya, A. Y. (2003). Correlates of examination cheating behavior among university students. *IFE Psychologia: An International Journal*, 11: 71-79.

[4] Beck, L. and Ajzen, I. (1991). Predicting dishonest actions using the theory of planned behavior. *Journal of Research in Personality*, 25: 285-301.

[5] Davis, S. F., Grover, C. A., Becker, A. H. and Mcgregor, L. N.(1992). Academic Dishonesty: Prevalence, Determinants, Techniques, and Punishments. *Teaching of Psychology* 19(1): 16-20.

[6] Daniel, L. G., Blount, K. D. and Ferrell, C. M. (1991), Academic misconduct among teacher education students: a descriptive–correlational study. Research in Higher Education, 32(6): 703–724.

[7] Dempster, A. P., Laird, N. M. and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society Series B (Methodological)*, 39(1):1-38.

[8] Donse, L. and Van De. Goep, I. H. (2013). Academic dishonesty among college students. Predictors and interventions, Social Cosmos, 10(1): 40–50.

[9] Harding, T., Carpenter, D. , Finelli, C. and Passow, H. (2004). Does academic dishonesty relate to unethical behavior in professional practics? An exploratory study. Conference paper in *3rd conference on Ethics and Social Responsibility in Engineering*, 10: 311-324.

[10] Hinton, P. R., Brownlow, C., Mcmurray, I. and Cozens, B. (2004). SPSS explained. East Sussex, England, Routledge Inc.

[11] Holleque, K. L. (1982). Cheating behavior of college students. Doctoral dissertation, Montana State University. *Dissertation Abstracts International* 43: 88. University Microfilms No. 82-14,649.

[12] Houston, J. P. (1976). The assessment and Prevention of Answer Copying on Undergraduate Multiple-Choice Examinations. *Research in Higher Education*, 5: 301-311.

[13] Houston, J. P. (1983). Alternate test forms as a means of reducing multi-ple-choice answer copying in the classroom. *Journal of Educational Psy-chology*, 75: 572–575.

[14] Jendrek, M. P. (1989). Faculty reactions to academic dishonesty. *Journal of College Student Development*, 30: 401–406.

[15] Jung In, S. B., Jun, M. K. and Shane, C. B. (2015), Students' Unethical Academic Behaviors: A Self Determination Theory Approachm. *Journal of Tourism and Hospitality Management*, 3(2): 60–72.

2023. In Saroj Hiranwal & Garima Mathur (eds.), *Artificial Intelligence and Communication Technologies*, 629–647. Computing & Intelligent Systems, SCRS, India. https://doi.org/10.52458/978-81-955020-5-9-60

[16] Jeergal, P. A., Surekha, R., Sharma, P., Anila, K., Jeergal, V. A. and Rani, T. (2015). Prevalence, perception and attitude of dental students towards academic dishonesty and ways to overcome cheating behaviors. *Journal of Advanced Clinical & Research Insights*, 2: 2–6.

[17] Khodaie, E., Moghadamzadeh, A. and Salehi, K. (2011). Factors affecting the probability of academic cheating school students in Tehran. *Procedia-Social and Behavioral Sciences*, 29: 1587-1597.

[18] Koul, R., Clariana, R. B., Jitgarun, K. and Songsriwittaya, A. (2009). The influence of achievement goal orientation on plagiarism. *Learning and Individual Differences*, 19: 506-512.

[19] K12 Academics. Types of Cheating.
`https://www.k12academics.com/Academic%20Dishonesty/types-cheating`

[20] Lin , T. S. , and Dayton , C. M. (1997). Model - selection information criteria for non - nested latent class models. *Journal of Educational and Behavioral Statistics*, 22 , 249 – 264.

[21] Linzer , D. A. and Lewis J. (2011). poLCA: Polytomous Variable Latent Class Analysis. *Annals of Applied Statistics*, 11(4): 2085-2115.
`http://CRAN.R-project.org/package=poLCA`

[22] Marsh, J. D., and Campion, J. (2018). Academic integrity and referencing: whose responsibility is it?. *Journal of Academic Language and Learning*, 12(1), A213–A226.

[23] Mccabe, D. (1999). Toward a Culture of Academic Integrity. *Journal of higher education*, 46(8): B7.

[24] McLachlan, G. and Peel, D. (2000). Finite Mixture Models. *John Wiley & Sons*, New York.

[25] Michaels, J. W. and Miethe, T. D. (1989). Academic effort and college grades. *Social Forces*, 68(1): 309–319.

[26] Orosz, G., Dombi, E., Toth-kiraly, I., Bothe, B., Jagodics, B. and Zimbardo, P. G. (2016). Academic Cheating and time perspective: Cheaters live in the present instead of the future. *Learning and Individual Differences*, 52: 39-45.

[27] Rajendran, A. (unpublished), Research Paper on Cheating.

[28] Romney, M. and Steinbart, P.(2003). Accounting Information Systems. 9th Edition, Prentice Hall, Upper Saddle River, NJ.

[29] Schwartz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, 6: 461-464.

[30] Sierles, F., Hendrickx, I. and Circle, S. (1980). Cheating in medical school. *Journal of Medical Education*, 55: 124-125.

[31] Skrondal, A., Rabe-Hesketh, S. (2004). Generalized latent variable modeling: multilevel, longitudinal, and structural equation models. *Interdisciplinary statistics* 1st ed., Chapman & Hall; Boca Raton, Fla.

[32] Tom, G. & Borin, N. (1988). Cheating in academic. *Journal of Education for Business*, 63: 153-157.

[33] Vermunt, K. and Jeroen k. (2010). Latent class modelling with covariates: Two improved three-step approaches. *Political Analysis*, 18: 450-469.

7 Detailed description of variables used in LCA

S.No.	Variables	Descriptions
1	X	Cheating behaviour of the students
2	a_1	Used any sort of prohibited material in the exam
3	a_2	Deliberately looked at another student's test sheet or made someone else to look at your test sheets
4	a_3	Passed answers to another person during a test or take answers from them
5	a_4	Planned with another student how to cheat prior to exam
6	a_5	Obtain a copy of an exam paper or test paper before exam
7	a_6	Ever made attempt to obtain or accept assistance from any other person during exam
8	a_7	Lied to an instructor for conducting an exam or test again/ for not appearing in exams or tests
9	a_8	Copied another person's assignment/ research/ thoughts through online/ offline mode and passed it off as your own
10	a_9	Complete the work which is assigned to someone else or made any other person complete the work assigned to you
11	a_{10}	Illicitly obtain material or steal material needed to complete assignment
12	a_{11}	Misrepresenting a family or personal situation (made excuses) to get an extension in assignment
13	a_{12}	Ever prevented other students from from completing their work
14	a_{13}	Ever forged (copy) a faculty/ family/ friend's signature on permission form or add/ drop form
15	G	Gender of the student
16	S	Stream of the student