EUROPEAN ACADEMIC RESEARCH
Vol. III, Issue 6/ September 2015

ISSN 2286-4822
www.euacademic.org

Impact Factor: 3.4546 (UIF)
DRJI Value: 5.9 (B+)

Factors Identification of Drug Addiction in Youngsters in Dera Ismail Khan

MUHAMMAD AMEEQ
Department of Statistics
Bahauddin Zakariya University
Multan, Pakistan
MEHR ALI
Department of Statistics
Bahauddin Zakariya University
Multan, Pakistan
MUHAMMAD MUNEEB HASSAN
Department of Sociology
Bahauddin Zakariya University
Multan, Pakistan

Abstract:

The purpose of this study is to factors identification of drug addiction in youngsters (10-18) years for only male. This study was carried out in Dera Ismail Khan (DIK) KPK. Distract Dera Ismail Khan consists of Four tehsils such as Dera Ismail Khan, Paharpur, Kulachi and Prova, and its total population is almost 1018020 (2013) in which the total male are 587397 and total male having 10-18 years age are 124528. For this study the data was collected from every tehsil by simple random sampling and a questionnaire was also created which was consisting fifty six (56) questions. Tehsil one DIK: the total male (10-18) population of the tehsil DIK was 44083 (35.4%) and 431 samples were selected. Tehsil one Paharpur: the total male (10-18) population of the tehsil Paharpur was 28641 (23%) and 320 samples were selected. Tehsil one Kulachi: the total male (10-18) population of the tehsil Kulachi was 22415 (18%) and 280 samples were selected. Tehsil one Prova: the total male (10-18) population of the tehsil Prova was 28389 (23.6%) and 330 samples were selected. In methods of study the first step was using pie chart with descriptive statistics. For main

work the logistics regression was used because study model the response variable is in categorical form which tells that the variables significant. After logistic regression the stepwise regression technique was used this helped to identify significant variables after that again logistic regression was applied. For this study total 1361 samples were selected among 984 were drug addict and 377 were non-drug addict. This data was collected from different places like school, factories, bus stand, workshop, sport club and hotels etc. There are many factors of drug addiction but society and poverty are major factors which are causing increase in this addiction day by day moreover more than forty articles were studied for this purpose.

Key words: drug addiction, youngsters, Dera Ismail Khan, Pakistan

Introduction

Definition of Drug Addiction:

It is a condition in which there is an irresistible craving to continue taking a drug, chemical, medicine, to which one has become habitual through frequent consumption because it produces a specific effect, usually modification of mental condition. Nowadays Drug addiction is a big problems in society. It is quickly increasing day by day. We want to determine the factors identification of drug addiction in youngsters in DIK. Here youngsters age between 10 to 18 years.

Why youngsters take drug:

Nowadays life has become so busy and full of tensions as well. Few people think that these drugs can make their life better or temporarily remove their tensions. Few people take drugs to bring changes in their life style. There are multiple reasons for drug addiction, most important are:

- People think drugs will keep them fit and fine.
- People take drugs to forget painful memories.
- To escape from the depression and anxieties.

- To relieve tediousness.
- Just for experimental purpose or for a bet.
- To look grown up.

People think that these drugs are the real solution for every problem but they are actually putting themselves in much bigger problem. Many factors are taking drug addiction but Poverty and society are main factors of drug addictions in DIK. Many people take drug for relaxation himself due to stress. Some take drug as a fashion. Addicted youngsters create big problems in society.

How Drugs Work?

Drugs in minute quantities can act as stimulant (speeds you up). If the dose of the drug is increased it will act as sedative (slows you down). A very high dose may be toxic and even lead to death. (Foundation for Drug –free world)

Do Drugs Affect the Mind?

Human mind is really a true gift of nature. It screens out any information very quickly which is already stored in it. The use of abuse drugs causes the formation of black spots which actually blur the memory. Person mind fail to gather information from the stored memory which is under the stress of narcotics.

Illegal drugs are consumed by approximately 208 million people worldwide. The study was conducted in 2007, the results indicated that in America 19.9 million people or 8% of the total population aged 12 year or older take narcotics in the month before investigation. Alcohol is the most commonly used and abused beverage in America. In US the 2nd leading cause of teen death is accidents. Marijuana is most frequently used abuse drug. It is estimated by United Nations 2008 World Drug Report that about 3.9% of the world's population having age

range of 15 to 64 abuse marijuana. (National survey on Drug use and Health)

Drugs addiction is injurious to the addicted persons. It also has a very bad impact on the surrounding environment. Drug addiction imparts devastating changes in the complete physiology of the addict. The most worrying feature of dependence is that it is growing very fastly all around the globe. Addiction can occur from all kind of drugs like prescription and non-prescription drugs etc. Non-prescription drugs comprise of cocaine, morphine, heroin, crack, marijuana and meth, while prescription drugs include hose such as benzodiazepines and barbiturates. In the US, addiction to drugs is quite widespread. One study on addiction showed that 1:5 American aged between 16 and 59 had used at least one drug. About 46% of Americans, aged between 16 and 21, faced occasion hospital admission at different ages of their lives. It is estimated that 20 million individuals between the age of 15 and 45 years, have used banned or unlawful drugs in the America.

Drug Addiction in Pakistan:

Drug Addiction in Pakistan Pakistan is also facing the alarming percentage of drug addiction. In South Asia in comparison to other countries like India, Bhutan, Afghanistan, Maldives as well as Nepal Pakistan is at top most regarding this issue. Pakistan is considered as largest heroin consumption market. 50 Masses of opium is unlawfully smuggled to Pakistan for the production of heroine. Pakistan is considered as largest exporter of heroin. Pakistan, Afghanistan, Nepal, Bangladesh, Bhutan, Sri Lanka and even India are opposed with drug addiction, abuse issues where a high population ratio is addicted to various abuse drugs**.** Pakistan is disreputable for many things, but most common and of great concern is drug production and addiction in the last 20 years. Drug addiction is also a major issue like other human improvement problems, such as poverty, illiteracy unemployment but it is often

surpassed. But in fact drug abuse is growing day by bay like a cancer in Pakistan and South Asia.

National Survey on Drug Abuse 1993 revealed, In Pakistan there are more than four million drug addicts while in KPK the drug addicts are 0.60 million out of which 71% drug abusers are below 35 years of age. In these, the 60% are educated and amazingly a same percentage was on job. Considerable distortions of Social Economical & Political system of the Pakistan especially in KPK was imparted by Afghan turmoil in 20 years. Considerable increase in the number of heroin user's especially young generation was found after enormous migration of Afghans to KPK Pakistan. Addiction has soiled its root in every single province. (KPK; social welfare and women development)

Among four million drug addicts in Pakistan, Ahmad is one of them, who use Cannabis. United Nations Office on Drugs and Crime (UNODC) made a report in collaboration with the Ministry of Narcotics Control and the Pakistan Bureau of Statistics. Report stated that 11% of KPK population use drugs. 6.5% (Sindh), 5% (Baluchistan), and 4.8% (Punjab) population use drugs in different parts of the country. 18% of KPK population who use Cannabis lie in the age of 15 to 19 years which is then followed by use of Amphetamines. 11% of the population of KPK use which is much higher than any other province. There is continuous increase in drug abuse i.e. at the rate of 7% per annum as revealed by a survey conducted in the year 1980, 1982,1988,1993,2000 and 2006. It is estimated by National Assessment Report on Problem Drug Use in Pakistan 2006 that there are approximately 6 lac opiate consumers, out of these 77 percent (482000) are heroin consumers. Although there is massive production of heroin and opium in Afghanistan, the relative stability was found in the numbers of the opiate addict people from the year 2000 to 2000 is a prominent accomplishment. However, the count of injecting drug users (IDUs) in 2006 was found double (125,000) than the

figure found in the year 2000. HIV/AIDS transmission risk (UNODC. 2006) has given special attention to this rise. The usage frequency for abusive drugs range from 0.4 % in Punjab and Sindh to 0.7 % in the KPK and 1.1 % in Baluchistan. The KPK and Baluchistan have a very long shared border with Afghanistan. In Pakistan there is slight increase in the percentage of drug user every year. The Intravenous drug users are increased from 15 % in 2000 to 29% in year 2006 (UNODC).

Our Objectives and data:

Our study's aim is to determine the factors identification of drug addiction in youngsters (10-18) years in DIK. Our research work is to endeavour to dig out factors identification of drug addiction in youngsters of age group (10-18) years especially in males. In order to accomplish this work, questionnaire was constructed and distributed among the sample of 1361 including two categories (addict drug is 984, not addict is 377) respondents belonging to Dera Ismail Khan (DIK), a Khyber Pakhtunkhwa (KPK) district. Our data collection tool is "questionnaire". The questionnaire carries different categories to be studied including education, drug addiction, smoking, occupation, behaviour, games, financial supports and environments of school, home and society. Different places i.e. School/college, Factories, Bus stand, Hotels, workshop and sports clubs were used to come across with the selected respondents. There are four Tehsils (subdivisions) in DIKHAN, and we have collected data from all these Tehsils (Subdivisions). First of all Tehsil 1 DIK is observed for data collection.

Tehsil 1 (DIK): We have collected data from schools, factories, workshops and sports clubs. The school names are Government higher secondary school muryalli, Government middle school Jhoke qurashian, Government high school Malana, Father School and Nadik public school. The total

population of this tehsil is 44083(35.6%) and 431 samples draw from this tehsil. The factories names are Huqqa Soap factory, Chashma Sugar Mills 1and the hotels names are Jan's hotel, Quraishi hotel and Waziristan hotel and also from Sports club DIK and Ali workshop. Tehsil 2 (Paharpur): The schools names are Government middle school Dhakki, Government High school paharpur and Al abbas public school Choora, The factories names are Al Moiz sugar mills, Sultan ghee mills and ZamZam Soda water factory. The hotels names are Sawan hotel and Kamran hotel. The total population of this tehsil is 28641(23%) and 320 samples draw from this tehsil. Tehsil 3 (Kulachi): The schools names are Government high school Kulachi, Government middle school Hathala, Oxford public school kulachi, the hotels names are Gandapur hotel Hathala and Khan Hotel kulachi. The total population of this tehsil is 22415(18%) and 280 samples draw from this tehsil. Tehsil 4 (Prova): The schools names are Government middle school Naiwela, Government high college Prova and Government High school Mahra.The factories names are Chashma sugar mills 2 and Meran Sugar mills. The hotels names areBaloch hotel parova, Niazi Hotel lunda and Shan hotel Naiwela. The total population of this tehsil is 29389(23.6%) and 330 samples draw from this tehsil.

Methodology first of all we are used the simple random sampling for collecting the data from youngsters in DIK. We are used the percentage, pie chart and Chi-square test. We are used the logistic regression technique for main purpose of our work. We also used the stepwise regression for selected the significant variables in our data.

The some variables have association in our data such as the Go to school/college/work (X_1), Took parents monthly progress (X_2) and Spend time after school/work (X_3), Take/addict niswar (X_{10}) and Take/addict niswar daily (X_{11}), Play the game (X_{48}) and Type of the game (X_{49}), Take breakfast (X_{52}) and Place the take breakfast (X_{54}).

The logistic regression and stepwise technique show the result that the six variables are significant such as go to school/college/work (X_1),Spend time after school/work (X_3), Incharge at their home (X_{26}), Behaviours of family member with them (X_{34}), Environment at their home (X_{37}) and Worries their friends (X_{42}).

Literature Review:

Sharma, N & Joshi, S (1956). The ratio of children living in streets is increasing at very alarming rate in developing countries and among these India has the highest proportion. This study aimed to investigate the strategies to prevent the use of drugs among these street children. This study indicated that street children in India constitute 10% of all the world's street children out of which 2/3rd were boys. It was revealed that due to poverty and peer pressure they were forced to live in streets, furthermore, large proportion of them were drug addicts i.e. nicotine and alcohol. Drug usage adversely affects the vital organs of the body and causes oral, heart, facial, respiratory, heart and digestive diseases.

Mohasoa, I, P (2010). Various theoretical situations were applied to identify the reasons, consequences and the strategies to deal with the problem of drug abuse. The sample was composed of 12 adult males ranging from 12 to 15 years selected from rural area of Zee-Rust and North-West bank of South Africa. Results indicated that people in selected sample was addict of heroin, alcohol, cannabis and nicotine. Primary factor was peer pressure while other factors include individual, family stress and environmental conditions.

Feidlerk et al (2002). This article reviewed all the literature over the past decade 1992-2002 on drug abuse among adults and further suggested treatment practices and policies for future. It was found that due to excessive use of alcohol and other drugs; Americans over age 55 had serious health

problems and an effective treatment system was needed to cope with this problem. This study was aimed to assess the reasons, patterns and the consequences of substance abuse. The researchers on the basis of results concluded that treatment providers, researchers and policy makers must be assisted to make effective decisions regarding aging problems of adults.

Foundation et al (1999). The purpose of the study was to identify the factors behind the drug abuse. Identified reasons include biological inclination, personality traits, past life, weak social relations, family problems and many others. Josser's behavior syndrome and Social stress model of Rhodes & Jason were used to understand the problem of drug abuse and it was concluded that to solve this complex problem, productive solutions were required.

Fox, A, M. (2008). In this article, 16 different factors were discussed regarding delinquency, drug addiction and other concerns in youth. Researchers collected data from 22 schools of five districts through surveys. The sample was composed of students enrolled in class 3 and 5. Response rate of the survey was recorded as 36%. It was named Trinidad and Tobago Youth Survey which measured factors in four different domains i.e. family, peers, school and community. The results revealed that improvement was required in the system.

Rammala, M, S & Dissertation, M. (2009). This study was intended to evaluate the factors that contribute in 12th standard students' poor performance. Two schools were selected as sample population and for this purpose both home and schools environments were assessed. Data was collected using multiple methods like personal interviews, field observations and secondary analysis of documents. Findings of the study indicated home environment was worst due to poverty, high unemployment, less education, child labour and emotional problems etc. Moreover, school environment was equally disturbing due to insufficient facilities, discipline, and workload

and so on. It was concluded that both environments needed to be improved to enhance the performance of students.

Dube, D, J. (2007). This research article was based on the objective to analyse the effects of social factors on adults' drug usage. This study was conducted in Attiring village Tshwane South. Researchers were interested to evaluate the relationship in drug addiction at primary school level and the social problems of South Africa. Interviews were conducted using purposive sampling method i.e. high school students (drug addicts) were interviewed. Researchers discovered that drug addiction negatively affected their health, social life and lead them towards crime.

Tresidder J. V. (2003). Less educated and unemployed adults are more likely to become drug addicts. This made the researchers to investigate the reasons and behaviours of school going children of 16-years old. Two-stage cluster sampling was used as data collection method in 1994 to recruit teenagers through Commonwealth Employment Service (CES). The findings showed that the sample group (both girls and boys) used alcohol and other drugs more than those who were still in school.

Hafisa, K. (2008). According to this article use of cannabis was rapidly increasing worldwide. Makerere University of Uganda conducted this study because use of cannabis in the country doubled (5-10%) in just 2-years i.e. 2008-2010 and users were approached and imprisoned. Makin dye Divisionissituated in the South-Eastern Kampala and has got 21 districts with 24% of the city's population whereas 70% of Makin dye's population was low income earners. Main objective of the study was to assess the effects of cannabis usage among residents of Makin dye division. Cross-sectional survey was used to collect data from 263 respondents using cluster and multi-stage cluster sampling. Usage ratio was 15.6%/34.6%. About 75.7% were male above 30-years and out of which 62% were aware of the consequences. Low education

level, family and peer pressure, and living environment were attributed as the main reasons of addiction.

Heim et al. (2004). This articles was aimed to explore the public perception regarding alcohol consumption. This study was conducted in Glasgow Scotland, UK while samples were drawn from Pakistan, India and Chinese composed of young generation ages ranging from 16-25 years. Purposive sampling was used to select sample while data was collected using interviewer controlled questionnaires. Researchers discovered that rate of alcohol consumption was less among targeted population than general one. And the reason behind was religious and ethnic affiliations.

Moal, M, L &Koob G, F. (2006). According to scientific definition, medically, drug infatuation is a chronic degenerating disease. Due to its complexity there is strong need of productive investigations to completely understand the situation. Keeping in mind the real life cases, stress factor, other reasons and theoretical concerns, this study suggested treatments for different stages of disease starting from control disorder to driven disorders. Furthermore, relations among different variables were also identified.

Kilpatrick et al. (2000). This article was proposed to collect data about drug usage, experiences of victims and their reactions are recorded for "Diagnostic & Statistical manual of Mental Disorders" regarding drug abuse. Using age and ethnicity as variables, data was collected from 4023 participants, age group ranging from 12-17years but ethnicity data was available for 3907 participants only. The main findings were: adults having violent experiences in past were drug addicts, stress after traumas caused increased use of marijuana etc., when other variables were controlled ethnicity affected adversely e.g. African Americans were more likely to be addicts than Hispanics or Native Americans. US department of Justice and other organizations of Justice supported this study.

Mullen et al (1993). To evaluate the relationship between the sexual abuse in childhood and adulthood mental health was the purpose of conducting this study. This study was conducted by randomly selecting the women of the community and they concluded that there was positive relationship between the abuse in earlier life and psychopathology of later life. Drug addiction and suicidal behavior were more commonly recorded among women who were sexually abused in childhood. Researchers reported that there was higher risk of adult psychopathology among those women who were physically abused in their early lives by using logistic regression analysis. Furthermore, the overlap between the effects of sexual abuse and the disadvantages invoked doubts about it being an independent casual element. On the other hand, long-term effects were not reported and it was decided that the effects were due to the context not the sexual abuse.

Methodology

3.1 Data Description:-

Our research work is to endeavour to dig out factors identification of drug addiction in youngsters of age group (10-18) years especially in males. In order to accomplish this work, questionnaire was constructed and distributed among the sample of 1361 including two categories (addict drug is 984, not addict is 377) respondents belonging to Dera Ismail Khan (DIK), a Khyber Pakhtunkhwa (KPK) district. Our data collection tool is "questionnaire". The questionnaire carries different categories to be studied including education, drug addiction, smoking, occupation, behavior, games, financial supports and environments of school, home and society. Different places i.e. School/college, Factories, Bus stand, Hotels, workshop and sports clubs were used to come across with the selected respondents.

There are four Tehsils (subdivisions) in DIKHAN, and we have collected data from all these Tehsils (Subdivisions). First of all Tehsil 1 DIK is observed for data collection.

Tehsil 1 (DIK): We have collected data from schools, factories, workshops and sports clubs. The school names are Government higher secondary school Muriyali, Government middle school Jhoke Qurashi, Government high school Malana, Father School and Nadik public school. The total population of this tehsil is 44083(35.6%) and 431 samples draw from this tehsil.
The factories names are Huqqa Soap factory, Chashma Sugar Mills 1and the hotels names are Jan's hotel, Qureshi hotel and Waziristan hotel and also from Sports club DIK and Ali workshop.
Tehsil 2 (Paharpur): The schools names are Government middle school Dhakki, Government High school paharpur and Al abbas public school Choora, The factories names are Al Moiz sugar mills, Sultan ghee mills and ZamZam Soda water factory. The hotels names are Sawan hotel and Kamran hotel. The total population of this tehsil is 28641(23%) and 320 samples draw from this tehsil.
Tehsil 3 (Kulachi): The schools names are Government high school Kulachi, Government middle school Hathala, Oxford public school kulachi, the hotels names are Gandapur hotel Hathala and Khan Hotel kulachi. The total population of this tehsil is 22415(18%) and 280 samples draw from this tehsil.
Tehsil 4 (Prova): The schools names are Government middle school Naiwela, Government high college Prova and Government High school Mahra.The factories names are Chashma sugar mills 2 and Meran Sugar mills.The hotels names areBaloch hotel prova, Niazi Hotel lunda and Shan hotel Naiwela.The total population of this tehsil is 29389(23.6%) and 330 samples draw from this tehsil.

The detailed lists of the variable names with notations are presented in the bellow table;

Table3.1 Names and Notations of selected Variables

Sr. No	Variable Names	Notation
1	Go to school/college/work	X_1
2	Took parents monthly progress	X_2
3	Spend time after school/work	X_3
4	Vacation per week	X_4
5	Smoke cigarettes daily	X_5
6	Get started smoking	X_6
7	Any person smoke in their family	X_7
8	Tried to stop smoking	X_8
9	Thinking that increase smoking day by day	X_9
10	Take/addict niswar	X_{10}
11	Take/addict niswar daily	X_{11}
12	Drug bad impression on their health	X_{12}
13	Type of bad impression on their health	X_{13}
14	Type of work they do	X_{14}
15	Type of transportation mode use of going to school/college/work	X_{15}
16	Role of parents for drug	X_{16}
17	They role in their family	X_{17}
18	Type of drug for they addict	X_{18}
19	Please for taken drug	X_{19}
20	Reason for taken drug	X_{20}
21	Their parents react for they addict drug	X_{21}
22	Anyone addict drug in their family except they.	X_{22}
23	Feel during	X_{23}
24	Need help to stop using drug	X_{24}
25	Introduced to drug	X_{25}
26	Incharge at their home	X_{26}
27	Spend time at their home	X_{27}
28	Spend time at their homework	X_{28}
29	Their family back ground	X_{29}
30	Own household	X_{30}
31	People during taken drug	X_{31}
32	Behavior of people with them	X_{32}
33	Behavior of friends with them	X_{33}
34	Behavior of family member with them	X_{34}
35	Behavior of parents with them	X_{35}
36	The person he contacts first whenever faces problems	X_{36}
37	Environment at their home	X_{37}
38	Environment at their school	X_{38}
39	Places where they spend most of their time	X_{39}
40	They face troubles	X_{40}
41	Worries their family member	X_{41}
42	Worries their friends	X_{42}
43	Close relation drug user	X_{43}

44	Topic discuss mostly with their friends	X_{44}
45	Safe in their area	X_{45}
46	Facilities drug in their area	X_{46}
47	Society avoid the drug	X_{47}
48	Play the game	X_{48}
49	Type of the game	X_{49}
50	Attend the sports club	X_{50}
51	Use computer	X_{51}
52	Take breakfast	X_{52}
53	Reason not going to school/college	X_{53}
54	Place the take breakfast	X_{54}
55	Member of any organization	X_{55}
56	Financially support the drug	X_{56}

Analysis Tools:

In this section we will use the different technique of the statistics such as the descriptive statistics tool are the percentage and pie chart. We will use the Pearson Chi- square test for check the association between variables. The data are in categorical form and we will use the binary logistic regression for the response variable. We also use the stepwise regression for the model selection. So the details about all these techniques are below.

Descriptive statistics:

A collection of information when described quantitatively regarding main characteristics is known as descriptive statistics. Simple summaries regarding the sample and observations are obtained by descriptive statistics .these summaries can be quantitative, e.g. summary statistics, or illustration, e.g., simple to understand graphs. These summaries can be form the basis of the primary description of the data like a part of a more widespread statistical analysis, or they may be enough in specific investigations.

Percentage:

Expressing a proportion by some specific way is known as percentage. A percentage is equivalent to the proportion times 100.

Pie Chart:

Definition of Pie Chart:

It is a circle shape chart which is consists of sectors in which the area possessed by every sector represents the size of the data. .Theses charts represent data in proportion. All the sectors collectively form a complete diskette.

Chi-Square test:

For testing independence and goodness of fit a chi square test is used. The independence test tells whether about observations across two populations are dependent on each other or not while. Goodness of fit test tells about an experiential frequency distribution matches a theoretical frequency distribution or not. Following is the general procedure for testing Pearson chi-square:

I). **Null hypothesis**:H_0 : The k classifications are independent of each other.

Alternative hypothesis:

H_1 : The k classifications are dependent of each other

II) **Level of the significant**: significance level α is a level for decision, which is the rejection probability of null hypothesis in case it is true the general value of the α are 0.01, 0.05 and 0.1.

III) **Test Statistics**:

Chi-square statistic is

$$\chi Chi - Square = \sum_{i=1}^{n} \frac{(O_i - E_i)}{E_i}$$

Where, i= 1, 2, 3… n.

Where O_i represents the observed frequency while E_i the expected frequency.chi-square test results, beside with the degrees of freedom , for finding p-value a formerly calculated table of chi-square distributions is used. This p-value then used to describe the worth of the test.

IV) **Computation**: The value of the test-statistic is computed from the sample data for taking decision whether the null hypothesisH_0.rejected/accepted.

V) **Critical Region**: It describes the rejection region so that the probability of rejecting the null hypothesis H_0, if it is true is equivalent α which is signigicance level. The rejection region location depends upon the type of H_1. The alpha (a) separates the acceptance region from the rejection region.

VI) **Conclusion:** Reject the null hypothesis if the computed value of the test-statistic falls in the critical region and conclude that alternative hypothesis is true and otherwise null hypothesis is true.

Logistic regression:

The connection between a dependent variable and one or more independent variables is known as logistic regression which are typically continuous, predicted values are used as probability scores of the dependent variable. As it treats the similar set of problems as the probit regression using same methods; the previous supposed logistic function, and the last supposed as standard normal distribution function.

Let us assume that there is a one expletory variable X, a quantitative variable. For a dual response variable Y, review

that P(X) represents the success probability at value X and this is the parameter for the binomial distribution. This model has linear form for the logit of this probability,

$$logit[P(X)] = log\left(\frac{P(X)}{1-P(X)}\right) = \alpha + \boldsymbol{\gamma} X$$

The formula suggests that $P(X)$ rises or falls as an S-shaped function of X.

This formula suggests the following formula for the probability $P(X)$ by the help of exponential function.

$$\exp(\alpha + \boldsymbol{\gamma} X) = e^{(\alpha+\gamma X)}$$

$$P(X) = \frac{\exp(\alpha + \boldsymbol{\gamma} X)}{1 + \exp(\alpha + \boldsymbol{\gamma} X)}$$

It can be seen that Logistic regression is exceptional case of generalized linear model and therefore analogous to linear regression. The logistic regression model, though, is grounded on quite different norms from those of linear regression. (Agrasti 2007).

Logistic function

The details of logistic regression start with a detail of the logistic function, its values between zero and one (Stanley, L. (2000))

$$\boldsymbol{F(X) = \frac{e^{xi}}{1+e^{xi}} = \frac{1}{1+e^{-xi}}}$$

A linear function of an explanatory variables, the logistic function can be expressed as:

$$\boldsymbol{F(X) = \frac{1}{1+e^{-(\gamma o+\gamma 1x1+\gamma 2x2+\ldots,+\gamma pxp)}}}$$

i= 1, 2, 3. . . p, j= 1, 2, 3. . . p

This is understood as the probability of the dependent variable equating success or "case" rather than a failure or non-case. It's obvious to see that the response variables Y_i are not equally dispersed:

$$P(Yi \mid X1, X2, X3, \ldots, Xp)$$

Varies from one data point X_i to another; however they are independent given design matrix X and shared with parameters β_j.

Binary Logistic Regression:

In statistics the activity of modelling the association between regressed and repressors variables is a basic activity. Usually simple linear regression is used to find the relationship between single explanatory and single response variable. For several variables the regression used is multiple, but most often the regressed is not a numerical value.

The response is simply the designation of two possible results. For instance head, tail good, bad or success failure. Even though the responses provide the number of successes and failures but the binary characteristic of the response remains similar. The data obtained from the relationship of explanatory and binary responses can be abounding in engineering to natural sciences as well as in medicine and education.

The effect of more variables that variables having binary outcome can be analysed by logistic regression. The outcome having binary nature is given the code zero and one. Often one is known for success and zero with failure. The following conditions should be fulfilled for logistic regression.

1) A variable with two possible outcomes preferably in the form of zero and one.
2) The probability (p) of the observed value of the outcome variable is needed to be estimate.
3) The logistic function should be used for relating the outcome variable to the repressor's variables.

4) The coefficients of the regression equation should be estimated by away. We need to find that way.
5) Goodness of fit of the regression models need to be tested along with confidence interval estimate of the coefficients.

Bring to mind that in multiple linear regressions the fundamental step is to draw the least square line about which the values of y are spread.

On the other hand the probability that an individual will fall into the outcome group or the other to be estimated.

The probability helps to understand the coefficients in the logistic regression model, same as we give meaning to the coefficients in linear regression model.

For measuring probability one asks about the odds of an event happening. If "P" denotes the probability an event, then (1 − P) is the probability of the event not happenings. The odds of the event would be as:

$$\text{Odds} = {}^{P}/_{(1-P)}$$

Its means that (Probability of event happening) ÷ (Probability of event not happening).

OR

$${}^{P}/_{(1-P\ )}$$

The relative amount of the probability of success, also called the odds of success. So

$$P = \frac{odd\ ratio}{1 + odd\ ratio}$$

If there are many reasons involved in deciding the ultimate outcome we calculate the odds ratio for each one one by one. The joint effect of all the independent variables may be spoken precisely as:

$$\mathbf{Odds} = \boldsymbol{P}/(\mathbf{1}-\boldsymbol{P}) = \boldsymbol{e}^{\alpha+\gamma_1 X_1+\gamma_2 X_2+\ldots+\gamma_p X_p}$$

Here $\boldsymbol{e}$ is a mathematical term and numerical value of e is to 2.71828.

The independent involvements of several factors i.e X_1, X_2 … X_p mentioned above in the expression to the total odds ratio can be determined by applying log equality in the equation.

$$ln\{Probability\ of\ success(P)/probability\ of\ failure\ \ (1-P)\} = Log\boldsymbol{e}^{\alpha+\gamma_1 X_1+\gamma_2 X_2+\ldots+\gamma_p X_p}$$

Where ln=$\log_e$

$$ln\{Probability\ of\ success(P)/probability\ of\ failure\ \ (1-P)\}$$

is considered as a logistic transformation regarding probability P denoted as logit (P), which means logistic unit. The changing of calculated proportion P means getting free of the disadvantage of probability which differs from zero to one while logit differs from negative infinity (-∞) to positive infinity (+∞).

$$ln\{Probability\ of\ success(P)/probability\ of\ failure\ \ (1-P)\} = \text{logit of probability (P)} = \alpha + \boldsymbol{\gamma}_1 X_1 + \boldsymbol{\gamma}_2 X_2 + \ldots + \boldsymbol{\gamma}_p X_p$$

It is observed that right side of equation is similar as in multiple linear regressions which are known to us. As the parameters are estimated by the help of least square method. So by this way the coefficients nominated are such that the squared distance of observed and predicted values are smallest. The estimated done by logistic regression in such a way that coefficients make our results more likely. This method is known as maximum likelihood. Let us assume that the possible values in the above mentioned equation of a variables are existence or not existence of illness with (existence of drug addiction is equal to one and not existence of drug addiction is equal to zero).

The total factors of drug addiction danger is represented by a where γ_1 is a portion by which the danger gets bigger or smaller. When a unit change occur in X_1, and γ_2 is a portion by which the drug addiction factors are transformed when unit change in X_2 occurs and so on. The independent variables may to measurable or non-measurable. For instance by using

(dummy) variables. However it should be kept in mind that logit of a proportion P is log of the relative odds.

If γ is the co-efficient of X variable then a unit rise in X rises the log of odds by a quantity equal to γ. Which implies that the odds are themselves raising by a factor of e^{γ}. Therefore if γ=1.8 than the odds $e^{1.8}=$

There is a rule for changing the log of overall odds back to probability. i.e

Probability is equal to the ratio between exponential of variables and one minus exponential of variables and the mathematical form:

Let probability is denoted by P and variable is denoted by X

$$P = \frac{e^{X}}{1-e^{X}}$$

Likelihood Function for Logistic Regression

Logistic regression gives only probabilities, not only classes, it can be fit by applying likelihood. There is a vector of features for every training point that is Xi. Also an observed class, Yi. The probability attained by that class was might p, now if Yi =1 or 1−p, and if Yi =0. The likelihood is then

L (γ0, γj) =$\prod_{i=1}^{n} p(Xi)^{Yi}\,(1-p(Xi))^{1-Yi}$

1

Where i =1, 2, 3. . . n and j=1, 2 3 . . . p.

The log-likelihood turns products in to sums:

$$log(\boldsymbol{\gamma}0,\boldsymbol{\gamma}j) = \sum_{i=1}^{n} yi\ logP(xi) + (1-yi)log1 - P(xi)$$

$$log(\boldsymbol{\gamma}0,\boldsymbol{\gamma}j) = \sum_{i=1}^{n} log1 - P(xi) + \sum_{i=1}^{n} yi \log P(xi)/(1-P(xi)$$

$$= \sum_{i=1}^{n} log1 - P(xi) + \sum_{i=1}^{n} yi \log(\gamma 0 + \mathrm{xi}\gamma \mathrm{j})$$

$$= \sum_{i=1}^{n} -log1 + e^{\gamma 0,+\mathrm{xi}\gamma \mathrm{j}} \sum_{i=1}^{n} yi \log(\gamma 0 + \mathrm{xi}\gamma \mathrm{j})$$

characteristically, log likelihood is differentiated regarding parameters to find the maximum likelihood estimates with respect to the parameters, the derivatives are set equal to zero, and solution is to be done. Take the derivative with respect to additional than one component of γ, say γ_j

$$\frac{\delta l}{\delta \boldsymbol{\gamma} j} = \sum_{i=1}^{n} \frac{1}{1 + e^{\boldsymbol{\gamma} 0 + xi\boldsymbol{\gamma} j}} e^{\boldsymbol{\gamma} 0 + xi\boldsymbol{\gamma} j} \; xij \sum_{i=1}^{n} yi \; xij$$

$$= \sum_{i=1}^{n} (yi - p(\, xi;\, \boldsymbol{\gamma} o, \boldsymbol{\gamma} j))\; xij$$

We aren't able to set this as zero and solve precisely.

Stepwise Regression:

The selection method consists of forward and backward section is known as stepwise regression but the selection procedure known as multivariate variables selection is usually better.

Stepwise regression is a modification form of forward selection in such way that entire candidate variables are checked regression significance whether their significance is reduced below the level or not. The non-significant variable is removed from the model.

There two significance levels required in stepwise regression. These levels are used for adding and removing variables cut-off. The probability for adding variables must be smaller than the cut-off probability for removing variables. In this process must not get into an infinite loop.

Software:

We have used SPSS 19 and Minitab 16 for analysis of the data.

Data Analysis

In this chapter we analyze our data. We use the addict or not addict status as response variable and all the remaining variables are explanatory variables such as already discuss in chapter 3. We use the different statistical tool for analysis our data. Firstly we use the descriptive statistics such as percentage and pie charts. We use the Pearson Chi-square foe association between the variables. We use the binary logistic regression for model building and after that we select a best model by using the stepwise regression technique. The analysis is explaining below;

Descriptive Statistics

We use the percentage for those variables have two category i.e. (yes, no or dependent, independent or one, more than one).

These variables X_1, X_2, X_8, X_9, X_{10}, X_{12}, X_{17}, X_{21}, X_{22}, X_{24}, X_{27}, X_{30}, X_{31}, X_{42}, X_{45}, X_{46}, X_{47}, X_{48}, X_{50}, X_{52}, X_{55}have two categories.

A total of 1361 youngsters were used in this analysis in which the 984 youngsters have addict and 377 youngsters have not addict in drug.

Now the results of the percentage are shown below one by one which those variables have two categories.

This table show that the percentage of those variables have two categories:

Variables	Valid value	Count	Percentage
X_1	0	780	53.3
	1	581	42.7
X_2	0	339	24.9
	1	1022	75.1
X_8	0	211	21.1
	1	789	78.9
X_9	0	133	13.3
	1	867	86.7
X_{10}	0	694	51.0
	1	667	49.0

X_{12}	0	29	2.9
	1	971	97.1
X_{17}	0	695	51.1
	1	666	48.9
X_{21}	0	188	18.8
	1	795	79.5
	Missing	17	1.7
X_{22}	0	100	10.0
	1	833	88.3
	Missing	17	1.7
X_{24}	0	302	30.2
	1	684	68.4
	Missing	14	1.4
X_{27}	0	537	39.5
	1	824	60.5
X_{30}	0	373	27.4
	1	988	72.6
X_{31}	0	290	29.0
	1	693	69.3
	Missing	17	1.7
X_{42}	0	46	3.4
	1	1315	96.6
X_{45}	0	146	10.7
	1	1215	89.3
X_{46}	0	430	31.6
	1	931	68.4
X_{47}	0	760	55.8
	1	601	44.2
X_{48}	0	97	7.1
	1	1264	92.9
X_{50}	0	712	52.3
	1	649	47.7
X_{52}	0	38	2.8
	1	1323	97.2
X_{55}	0	903	66.3
	1	458	33.7

Interpretation:

For X_1:

The results indicate that 42.7% of the total youngsters have attended school/college/work while remaining 57.3% have not.

For X_2:

The results indicate that 75.1% of the total youngsters have attended performance report from school/college/work while remaining 24.9% have not.

For X_8:

The results indicate that 78.9% of the total youngsters have attended tried to stop smoking while remaining 21.1% have not.

For X_9:

The results indicate that 86.7% of the total youngsters have attended trend to smoking increased day by day while remaining 13.3% have not.

For X_{10}:

The results indicate that 49.0% of the total youngsters have taken Niswar while the remaining 51.0% have not.

For X_{12}:

The results indicate that 97.1% of the total youngsters have feel bad impression on their health of addiction drugs while remaining 2.9% have not.

For X_{17}:

The results indicate that 48.9% of the total youngsters have independent in their family while the remaining 51.1% have not.

For X_{21}:

The results indicate that 79.5% of the total youngsters have attended their parents react to their involvement in drug while remaining 18.8% have not and 1.7% are missing values.

For X_{22}:

The results indicate that 88.3% of the total youngsters have attended their family member addict drug except them while remaining 10.0% have not and 1.7% are missing values.

For X_{24}:

The results indicate that 68.4% of the total youngsters have attended need help to stop using drug while remaining 30.2% have not and 1.4% are missing values.

For X_{27}:

The results indicate that 39.5% of the total youngsters have spent twelve hour time at your home while the remaining 60.5% have not.

For X_{30}:

The results indicate that 72.6% of the total youngsters have owner household while the remaining 27.4% have not.

For X_{31}:

The results indicate that 69.3% of the total youngsters have attended how many peoples with their taking drug while remaining 29.0% have not and 1.7% are missing values.

For X_{42}:

The results indicate that 96.6% of the total youngsters have worried their friends while the remaining 3.4% have not.

For X_{45}:

The results indicate that 89.3% of the total youngsters have attended safe in their areas while the remaining 10.7% have not.

For X_{46}:

The results indicate that 68.4% of the total youngsters have attended avail facilities of drug in their area while the remaining 31.6% have not.

For X_{47}:

The results indicate that 44.2% of the total youngsters have attended avoid the drug in their society while the remaining 55.8% have not.

For X_{48}:

The results indicate that 92.9% of the total youngsters have attended played the game while the remaining 7.1% have not.

For X_{50}:

The results indicate that 47.7% of the total youngsters have attended the sports club while remaining 52.3% have not.

For X_{52}:

The results indicate that 97.2% of the total youngsters have taken breakfast while remaining

2.8% have not.

For X_{55}:

The results indicate that 33.7% of the total youngsters have active member of organization while remaining 66.3% have not.

Check the association:

Association between X_1 and X_2:

X1 * X2 Cross tabulation

Count

		X2		
		.00	1.00	Total
X1	.00	212	568	780
	1.00	127	454	581
Total		339	1022	1361

Pearson Chi-square value = 5.040, df = 1, P-value = 0.025

The test shows that there is an association between mode of school/college/work concern and parents look after in taking their progress reports.

Association between X_1 and X_3 :

X_1 * X_3 Cross tabulation

Count

		X3			
		.00	1.00	2.00	Total
X1	.00	142	146	492	780
	1.00	58	165	358	581
Total		200	311	850	1361

Pearson Chi-square = 29.090, df= 2, P-value= 0.000

The relationship between modes of school/college/work concern is shown here by this and spend most of the time after come to school/college/work.

Association between X_2 and X_3 :

X_2 * X_3 Cross tabulation

Count

		X3			
		.00	1.00	2.00	Total
X2	.00	118	85	136	339
	1.00	82	226	714	1022
Total		200	311	850	1361

Pearson Chi-square = 161.318, df= 2, P-value= 0.000

The association is shown by the this test in between parents looks after in taking their progress reports concern and parents look after in taking their progress reports.

Association between X_{10} and X_{11}:

X_{10} * X_{11} Cross tabulation

Count

		x11								
		.00	1.00	2.00	3.00	4.00	5.00	6.00	7.00	Total
X10	.00	694	0	0	0	0	0	0	0	694
	1.00	1	95	235	172	90	34	39	1	667
Total		695	95	235	172	90	34	39	1	1361

Pearson Chi-square = 1357.004, df= 7, P-value= 0.000

The test shows that there is an association between taken/addict, Niswar concern and how many times per day.

Association between X_{48} and X_{49} :

X_{48} * X_{49} Cross tabulation

Count

		X49				
		.00	1.00	2.00	3.00	Total
X48	.00	10	2	74	1	87
	1.00	797	304	163	0	1264
Total		807	306	237	1	1351

Pearson Chi-square = 309.378, df= 3, P-value= 0.000

The test shows that there is an association between played any game concerned and type of the game

Association between X_{52} and X_{54}:

X_{52} * X_{54} Cross tabulation

Count

		X54				
		.00	1.00	2.00	3.00	Total
X52	.00	13	13	12	0	38
	1.00	65	1162	95	1	1323
Total		78	1175	107	1	1361

Pearson Chi-square = 95.623, df= 3, P-value= 0.000

The test shows that there is an association between taken breakfast concern and taken breakfast place.

Binary Logistic Regression:

Link Function: Logit

Response Information

Variable	Value	Count
Y	1	973 event
	0	358
	Total	1331
	Missing	30

Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio	95% CI Lower	95% CI Upper
Constant	-2.77721	1.34266	-2.07	0.039			
X_1	0.386586	0.205502	1.88	0.060	1.47	0.98	2.20
X_2	-0.06336	0.211653	-0.30	0.765	0.94	0.62	1.42
X_3	**-0.23387**	**0.120225**	**-1.95**	**0.052**	**0.79**	**0.63**	**1.00**
X_4	0.0753774	0.342841	0.22	0.826	1.08	0.55	2.11
X_{10}	21.7433	2613.40	0.01	0.993	2.77342E+09	0.00	*
X_{11}	0.0197840	853.141	0.00	1.000	1.02	0.00	*
X_{17}	0.0333471	0.191120	0.17	0.861	1.03	0.71	1.50
X_{26}	**-0.407970**	**0.146285**	**-2.79**	**0.005**	**0.66**	**0.50**	**0.89**
X_{27}	0.185421	0.195234	0.95	0.342	1.20	0.82	1.76
X_{28}	0.0762548	0.132880	0.57	0.566	1.08	0.83	1.40
X_{29}	-0.108483	0.107980	-1.00	0.315	0.90	0.73	1.11
X_{30}	-0.220210	0.199039	-1.11	0.269	0.80	0.54	1.19
X_{32}	0.118793	0.128776	0.92	0.356	1.13	0.87	1.4
X_{33}	-0.124891	0.108510	-1.15	0.250	0.88	0.71	1.09
X_{34}	**0.316524**	**0.128846**	**2.46**	**0.014**	**1.37**	**1.07**	**1.77**
X_{35}	-0.054995	0.134536	-0.41	0.683	0.95	0.73	1.23
X_{36}	-0.039306	0.0934440	-0.42	0.674	0.96	0.80	1.15

X_{37}	**0.356968**	**0.137443**	**2.60**	**0.009**	**1.43**	**1.09**	**1.87**
X_{38}	-0.123849	0.129702	-0.95	0.340	0.88	0.69	1.14
X_{39}	-0.048752	0.139154	-0.35	0.726	0.95	0.73	1.25
X_{40}	-0.059919	0.160656	-0.37	0.709	0.94	0.69	1.29
X_{41}	-0.044334	0.104937	-0.42	0.673	0.96	0.78	1.18
X_{42}	**1.90871**	**0.621001**	**3.07**	**0.002**	**6.74**	**2.00**	**22.78**
X_{44}	-0.115559	0.0767277	-1.51	0.132	0.89	0.77	1.04
X_{45}	-0.038122	0.263796	-0.14	0.885	0.96	0.57	1.61
X_{46}	-0.079066	0.206096	0.701	-0.38	0.92	0.62	1.38
X_{47}	-0.090638	0.179009	-0.51	0.613	0.91	0.64	1.30
X_{48}	-0.524903	0.382526	-1.37	0.170	0.59	0.28	1.25
X_{49}	-0.092391	0.130500	-0.71	0.479	0.91	0.71	1.18
X_{50}	0.135094	0.181608	0.74	0.457	1.14	0.80	1.63
X_{51}	0.111915	0.0949582	1.18	0.239	1.12	0.93	1.35
X_{52}	1.40190	0.792967	1.77	0.077	4.06	0.86	19.22
X_{54}	0.131445	0.250765	0.52	0.600	1.14	0.70	1.86

Log-Likelihood = -434.598

Test that all slopes are zero	G	DF	P-Value
	680.706	33	0.000

Goodness-of-Fit Tests:

Method	**Chi-Square**	**DF**	**P**
Pearson	420.609	795	1.000
Deviance	550.325	795	1.000
Hosmer-Lemeshow	4.530	8	0.806

Table of Observed and Expected Frequencies:

For one

Value	1	2	3	4	5	6	7	8	9	10	Total
Obs	29	58	65	73	87	130	133	133	133	132	973
Exp	33.2	53.5	63.2	72.7	87.1	132.3	133	133	133	132	

For zero

Value	1	2	3	4	5	6	7	8	9	10	Total
Obs	104	75	68	60	47	4	0	0	0	0	358
Exp	99.8	79.5	69.8	60.3	46.9	1.7	0.0	0.0	0.0	0.0	

Total	133	133	133	133	134	134	133	133	133	132	1331

Measures of Association:

Pairs	Number	Percent	Summary Measures
Concordant	309677	88.9	Somers' D 0.78
Discordant	37885	10.9	Goodman-Kruskal Gamma 0.78
Ties	772	0.2	Kendall's Tau-a 0.31
Total	348334	100.0	

Interpreting the results

Response Information:

There are three column in this table first is variable second is value of the variable and third is count of the value of variable. In variables column, we have only one variable which known as response variable and is represented by Y. In the second column there are two value of the variable such as one and zero. The third column show that the value one occur up to how many time as well as the value zero. The table also determines the missing values.

In our data value one occurred 973 times, the value zero 358 times and missing values are 30. Hence total value is 1361.

Logistic Regression Table:

There are eight columns in this table. The first column is predictors' column in which there is one constant and thirty three variables. In the second column there is description of co-efficient of all variables separately. In third column the standard error of co-efficient is described. In fourth column the Z-value of each variable is determined. In fifth column the P-value is described. The P-value tells us about the significant and non-significant of the variables.

In our data five variables are significant i.e. X_3, X_{26}, X_{34}, X_{37} and X_{42}. In the sixth column the odd ratio of each variable is described, in the seventh and eight columns the lower and upper value is described regarding 95% confidence interval respectively.

Log-Likelihood.

In log-likelihood table we get log-likelihood results as well as the results of G statistics, degree of freedom and also p-value. We observed in this table that all slopes of test are zero. . where G = 680.706, DF=33 with a p-value of 0.000, indicating that there is sufficient evidence that at least one of the coefficients is different from zero, given that your accepted α-level is greater than 0.000.

Goodness of Fit Tests:

Here three tests are shows Pearson, deviance, and Hosmer-Lemeshow goodness-of-fit tests. The goodness-of-fit tests, with p-values 1.0, 1.0 and 0.806 respectively. The P-value of all test is insignificant so we cannot reject the null hypothesis.

Table of Observed and Expected Frequencies

This table shows that the two sub tables in this table, one for addict and other for non addicts regarding drugs. We are observing observed and expected frequencies. The total observed and expected frequencies are 973 for one and 358 for zero value of variables.

Measures of Relation

There are four columns in this table such as pairs, number, percent and summary measures. There are three types of pair functions such as concordant, discordant and tied. This table is showing number with percent of these pair as well as the summary measures. The total numbers are obtained by multiplying total observed and expected frequencies.

In this table the concordats number are 309677 with 88.9 percentages. The discordant numbers are 37885 with 10.9 percentages. The ties numbers are 772 with 0.2 percentages.

There are three summary measures also in this table. The concordant (Somers D= 0.78), the discordant (Goodman-

Kruskai gamma=0.78) and ties (Kendall's Tav-a=0.31) are summary measures.

Stepwise Regression:

Alpha-to-Enter: 0.15 Alpha-to-Remove: 0.15
Response is y on 33 predictors, with N = 1331
N (cases with missing observations) = 30,N (all cases) = 1361

Step	1	2	3	4	5	6
Constant	0.4688	0.2191	0.2546	0.2928	0.2678	0.2657
X10	0.531	0.524	0.521	0.523	0.523	0.522
T-Value	27.26	26.97	26.84	26.96	26.98	26.95
P-Value	0.000	0.000	0.000	0.000	0.000	0.000
X42		0.261	0.257	0.261	0.272	0.243
T-Value		4.60	4.53	4.61	4.78	4.13
P-Value		0.000	0.000	0.000	0.000	0.000
X26			-0.041	-0.043	-0.044	-0.043
T-Value			-2.65	-2.81	-2.88	-2.79
P-Value			0.008	0.005	0.004	0.005
X3				-0.028	-0.030	-0.034
T-Value				-2.14	-2.27	-2.55
P-Value				0.032	0.023	0.011
X1					0.043	0.047
T-Value					2.22	2.39
P-Value					0.027	0.017
X34						0.024
T-Value						1.87
P-Value						0.062
S	0.355	0.353	0.352	0.351	0.351	0.351
R-Sq	35.87	36.87	37.20	37.42	37.65	37.82
R-Sq (adj)	35.82	36.78	37.06	37.23	37.42	37.53
Mallows Cp	37.9	18.5	13.4	10.8	7.9	6.4

Interpretation:

Stepwise regression selects those variables which are significant. In our data six variables are significant according to stepwise regression such as X_1, X_{10}, X_{26}, X_{34}, X_{37} and X_{42}. In this table there is P-value and t-value is given for every variables. We can get standard deviation S, R-square (R-sq) and adjusted R-square (R-sq adj) as well as mallow C_p results.

Logistic Regression after stepwise technique:

Binary Logistic Regression:

Link Function: Logit

Response Information

Variable	Value	Count
Y	1	986 event
	0	374
	Total	1360
	Missing	1

Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio	95% CI Lower	95% CI Upper
constant	-2.32438	0.577348	-4.03	0.000			
X_1	**0.401262**	**0.162675**	**2.47**	**0.014**	**1.49**	**1.09**	**2.05**
X_{10}	21.6781	1070.77	0.02	0.984	2.5984E+09	0.00	
X_{26}	**-0.392387**	**0.132016**	**-2.97**	**0.003**	**0.68**	**0.52**	**0.87**
X_{34}	0.169399	0.101376	1.67	0.095	1.18	0.97	1.44
X_{37}	0.219026	0.116555	1.88	0.060	1.24	0.99	1.56
X_{42}	**1.93074**	**0.560181**	**3.45**	**0.001**	**6.89**	**2.30**	**20.67**

Log-Likelihood = -460.041

Test that all slopes are zero	G	DF	P-Value
	679.738	6	0.000

Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	49.8912	100	1.000
Deviance	59.9117	100	0.999
Hosmer-Lemeshow	0.9635	7	0.995

Table of Observed and Expected Frequencies

For one

Value	1	2	3	4	5	6	7	8	9	Total
Obs	42	64	70	86	101	189	146	144	144	986
Exp	43.6	59.9	73.7	84.9	100.9	189.0	146	144	144	

For zero

Value	1	2	3	4	5	6	7	8	9	Total
Obs	103	76	84	74	37	0	0	0	0	374
Exp	145	140	154	160	138	189	146	144	144	

Total	145	140	154	160	138	189	146	144	144	1360

Measures of Association:

Pairs	**Number**	**Percent**	**Summary Measures**
Concordant	320118	86.8	Somers' D 0.75
Discordant	42539	11.5	Goodman-Kruskal Gamma 0.77
Ties	6107	1.7	Kendall's Tau-a 0.30
Total	368764	100.0	

Interpreting the results:

Response Information:

There are three column in this table first is variable second is value of the variable and third is count of the value of variable. In variables column, we have only one variable which known as response variable and is represented by Y. In the second column there are two value of the variable such as one and zero. The third column show that the value one occur up to how many time as well as the value zero. The table also determines the missing values. In our data value one occurred 986 times, the value zero 374 times and missing values are 1. Hence total value is 1361.

Logistic Regression Table:

There are eight columns in this table. The first column is predictors' column in which there is one constant and thirty three variables. In the second column there is description of co-efficient of all variables separately. In third column the standard error of co-efficient is described. In fourth column the Z-value of each variable is determined. In fifth column the P-value is described. The P-value tells us about the significant and non-significant of the variables.

In our data five variables are significant i.e. X_1, X_{26} and X_{42}. In the sixth column the odd ratio of each variable is described, in the seventh and eight columns the lower and upper value is described regarding 95% confidence interval respectively.

Log-Likelihood:

In log-likelihood table we get log-likelihood results as well as the results of G statistics, degree of freedom and also p-value. We observed in this table that all slopes of test are zero. . where G = 679.738, DF=6with a p-value of 0.000, indicating that there is sufficient evidence that at least one of the coefficients is different from zero, given that your accepted α-level is greater than 0.000.

Goodness of Fit Tests:

Here three tests are shows Pearson, deviance, and Hosmer-Lemeshow goodness-of-fit tests. The goodness-of-fit tests, with p-values 1.0, 0.999 and 0.995 respectively. The P-value of all test is insignificant so we cannot reject the null hypothesis.

Table of Observed and Expected Frequencies:

This table shows that the two sub tables in this table, one for addict and other for non addicts regarding drugs. We are observing observed and expected frequencies.

Measures of Relation:

There are four columns in this table such as pairs, number, percent and summary measures. There are three types of pair functions such as concordant, discordant and tied. This table is showing number with percent of these pair as well as the summary measures. The total numbers are obtained by multiplying total observed and expected frequencies.

In this table the concordats number are 320118 with 86.6 percentages. The discordant numbers are 42539 with 11.5 percentages. The ties numbers are 6107 with 1.7 percentages.

There are three summary measures also in this table. The concordant (Somers D= 0.75), the discordant (Goodman-Kruskai gamma=0.77) and ties (Kendall's Tav-a=0.30) are summary measures.

Conclusions and Recommendations

The present study is conducted to assess the factors identification of drug addiction in youngster (10-18) in Dera Ismail Khan (DIK), in province Khyber Pakhtunkhwa (KPK). In order to do accomplish this work, questionnaires were constructed and these questionnaires carried different categories to be studied including/like education, drug addiction, smoking, occupations, behaviour, games, financial supports and environments of school, home and society. Our objective is to assess the factors identification of drug addiction in youngster (10-18) in Dera Ismail Khan (DIK) for this purpose we select 56questions; in each questionnaire and related variables as described in Chapter 3. We selected different categories to be studied including/like education, drug addiction, smoking, occupation, behaviour, games, financial supports and environments of school, home and society.

A total of 1361 youngsters were used/ in analysis. In which 984 youngsters were addict/s and/while 377 were /not addict in drug. Our data divided into two main parts first completed observation and other is missing observation. The complete data consist of 40 variables and uncompleted data is consist /on 13 variables and three variables are open-end. We were introduced the dummy variable as a response variable in our data. The dummy variable has two outcomes such as addict and not addict.

We find the percentages of all those variables having two categories for completed observation and 14 variables have completed observation for finding the percentage. We were/ find the percentage of all those variables having two categories for

uncompleted observation and seven variables have uncompleted observation for finding percentage. We were/ constructed the pie charts of all those variables having three or more categories for completed observation and 19 variables have completed observation for pie chart. We were/ constructed the pie charts of all those variables having three or more categories for uncompleted observation and 13 variables have uncompleted observation for pie chart. The 11 variables have association in which the five variables have completed observation and two variables have uncompleted observation. We were/find the chi-square of those variables having association between each other. In chi square tables showed/ many results such as likelihood ratio and linear-by-linear association but we were just interpreted chi square Pearson result.

We were/ find checked that association between variables X_1(mode of school/college/work) between X_2(Took parents monthly progress) and their result is shown in table is Pearson-chi square (P-value) 0.025 and its means significant. Similarly checked the association between X1 (mode of school/college/work) and X3(Spend time after school/work) and their result is shown in table is Pearson-chi square (P-value) 0.00 and its means significant. Similarly checked the association between X_2(Took parents monthly progress) and X3(Spend time after school/work) and their result is shown in table is Pearson-chi square (P-value) 0.00 and its means significant. Similarly checked the association between X_{10}(Take/addict niswar) and X11(Take/addict niswar daily) and their result is shown in table is Pearson-chi square (P-value) 0.000 and its means significant. Similarly checked the association between X_{48}(Play the game) and X_{49}(Type of the game) and their result is shown in table is Pearson-chi square (P-value) 0.000 and its means significant. Similarly checked the association between X_{52}(Take breakfast) and X_{54}(Place the take

breakfast) and their result is shown in table is Pearson-chi square (P-value) 0.000 and its means significant.

We were find the checked that association between X_{12}(Drug bad impression on their health) and X_{13}(Type of bad impression on their health) their result is shown in table is Pearson-chi square (P-value) 0.642 and its means non-significant. We were interpreted based on the chi squared Pearson results if P-value has less than 0.05, it means that this is significant if P-value is more than 0.05 means that results is non-significant.

We were/ used the technique /of/ the binary logistic regression. Here, the logistic regression model was applied to choose the significant variables that are perceived to cause drug addiction in youngsters that are impressionable. We were /make the regression model. Our date/ have the categorical response variable (addict and non-addict) and the categorical predictor's variables. Binary logistics technique apply/ for 36 variables because 36 variables were completed and other were uncompleted variables. The binary logistic regression results show that information about 36 variables in which five variables are significant (P-value is less than .05) and other variables were non-significant variables. The significant variables X_3 (Spend time after school/work) have P-value 0.052, X_{26} (Incharge at their home) P-value 0.005, X_{34} (behaviour of family member with them) P-value 0.014, X_{37} (Environment at their home) P-value 0.009 and X_{42} (Worries their friends) P-value 0.002.All of other variables are non-significant.

The logistic regression tables show many results such as log-likelihood, goodness of fit and measure of association but we were interpreted that Pearson chi square value in goodness of fit table and the Pearson chi square value is 1.000 and its means non-significant.

The selection method consists of forward and backward section is known as stepwise regression but the selection

procedure known as multivariate variables selection is usually better.

The stepwise relapse shows the significant variables. The significant variables are X_{10} (Take/addict niswar) with the P-value is 0.000 at all six step and its means significant, X_{42} (Worries their friends)with the P-value is 0.000 at five step and its means significant, X_{26}(Incharge at their home) with the P-value is less than 0.05 at four step and its means significant, X_3 (Spend time after school/work) with the P-value is less than 0.05 at three step and its means significant,X_1(Go to school/college/work)with the P-value is less than 0.05 at two step and its means significant and X_{34} (Behavior of family member with them) with the P-value is less than 0.05 at one step and its means significant.

We make comparison of our study with some previous study to determine the factor identification of drug addiction.

Suggestion:

The importance of these two factors is reasons to keep youngsters away from drug are emphasized when their lack is mentioned and censured by drug users.

REFERENCES:

Sharma, N and Joshi, S (2013). "Preventing substance abuse among street children in India. Health" Science Journal Vol, **7** (2).

Mohasoa, I, P (2010)."Substance Abuse among Male Adolescents. University of South Africa."

Feidler, K and Associates, C(2002)."Substance Abuse Among Aging Adults. NEDS, J:\SARE\170\171150\Aging Lit Review\aging_lit.doc.

Foundation, N and Australia (1999). Causes and correlates of adolescent drug abuse and implications for treatment. Drug and Alcohol Review (1999) 18, 453-475.

Andrew M. Fox (2008).An Examination of Risk and Protective Factors Drug use and Delinquency in Trinidad and Tobago.

Mokgaetsi, Mini-Dissertation (2009).Factors Contributing Towards Poor Performance of Grade 12 Learners at Manoshi and Mokwatedi High Schools.

Dikeledi Josephine Dube (2007).Socail Factors InfluencingAdoleScent Drug Abuse in High Schools in Atteridgeville. Department of Educational Studies.

Julia Victoria Tresidder (2003).Binge drinking and related risk factors amongst out of school 16 year-olds in NSW. School of Public Health Faculty of Medicine,University of Sydney.

Michel Le Moal , George F. Koob(2006). Drug addiction: Pathways to the disease and pathophysiological perspectives. European Neuropsychopharmacology (2007) 17, 377–393, Received 16 May 2006; received in revised form 28 August 2006; accepted 24 October 2006.

Ron Acierno, Paula P. Schnurr(2000) . Risk Factors for Adolescent Substance Abuse and Dependence Data From a National Sample Dean G. Kilpatrick. Journal of Consulting and Clinical Psychology, February 2000 Vol. 68, No. 1, 19-30.

Sharma, N., & Joshi, S. (2014). Preventing-substance abuse among street children in India

P E Mullen, G P Herbison (1993).Childhood Sexual Abuse and Mental Health in Adult Life, British Journal of Psychiatry (1993), 163, page 721-732.

Mohasoa, I. P. (2010). Substance abuse among male adolescents.

Mattias Gunnarsson(2012). Psychological factors associated with substance use in adolescents, Mattias Gunnarsson,

Printed in Sweden, Department of Psychology. University of Gothenburg 2012, ISSN 1101-718X SBN 978-91-628-8493-2 , ISRN GU/PSYK/AVH--261—SE.

Zila van der Meer Sanchez ,SolangeAparecidaNappo (2005). Main reasons for non-use of illicit drugs by young population exposed to risk situations, Rev saudepublica 2005; 39(A) www.fsp.usp.br/rsp.

Judith S. Brook, Tine Pahl, B.A.(2006). "Predictors of Drug Use Among South African Adolescents" ,Published in final edited form as; J Adolesc Health. 2006 January ; 38(1): 26–34.

Mesfin Kassaye, "Ghimja Fissehaye (1999). Drug" use among high school students in Addis Ababa and Butajira" ,[Ethiop. J. Health Dev. 1999;13(2):101-106].

MCA Van Hout (2009). "Youth alcohol and drug use in rural Ireland -parents' views" , A licence to publish this material has been given to ARHEN http://www.rrh.org.au.

Michel Le Moal, George F. Koob (2006); Drug addiction: Pathways to the disease and pathophysiological perspectives, European Neuropsychopharmacology (2007) 17, 377–393.

All Rights Reserved. The Foundation logo is a trademark owned by the Foundation for a Drug-Free World. 2006–2014 Foundation for a Drug-Free World. Foundation for Drug –free world.find out the truth about drugs.

Tabachnick, Barbara and Linda Fidell.1996. Using Multivariate Statistics, Third edition. Harper Collins.

Trexler, J.C., and J. Travis.1993. Nontraditional Regression Analyses.Ecology 74:1629-1637.

Connor, E.F., Adams-Manson, R.H., Carr, T.G., and M.W. Beck.1994. The effects of host plant phenology on the demography and population dynamics of the leaf-mining moth, Cameraria hamadryadella (Lepidoptera: Gracillariidae). Ecological Entomology 19:111-120.