ABSTRACT

Objective: Substance abuse research data containing categorical and continuous variables often violate assumptions of parametric statistical methods. Clustering of individuals, lack of repeated measurements, missing data, non-representative samples, etc. aggravates the problem. Quality of life (QoL) measures suffer from meaningful application of statistical methods. The paper describes statistical approaches which fit well with structure of drug addiction data and measures of QoL, facilitating better analysis and interpretations of results.

Method: Converting ordinal item scores to normally distributed continuous scores in the range [1-100], irrespective of number of response-category in items. Such transformations fit well with structure of drug addiction data and measures of QoL and facilitate better analysis and interpretations of results.

Results: Proposed method enables parametric statistical analysis leading to meaningful comparisons and inferences, finding equivalent scores, computation of responsiveness of the scale i.e., ability to assess changes across time and psychometric qualities like reliability, as per definition, Factorial validity reflecting the main factor for which the test was developed.

Conclusions: Considering theoretical advantages, the proposed method generating normally distributed scores is recommended. Future studies with longitudinal data suggested finding sensitivity with emphasis on progression of disease and to different therapeutic interventions.

Keywords: Contingency Table; Factorial Validity; Logistic regression; Normal distribution; Parallel tests; Quality of life; Substance abuse; Theoretical Reliability

Introduction

Drugs of different types are being used increasingly to relieve pain and enjoy “Feel-good” effects, alleviate stress, cope with mental and emotional pain, alter or avoid reality. Some drugs act brain and alter mood in ways which are not approved socially like feeling of euphoria, restlessness, nightmare, hallucinations and delusions [1].

Drug addiction (substance use disorder) is a global disease associated with significant morbidity, mortality and is characterized by compulsive and uncontrollable drug cravings. Drugs bring changes in behaviors by affecting brain and central nervous system which in turn produce physical, psychological, behavioral changes and harm the social fabric of people’s lives and social relationships [2]. However, different types of drugs affect differently on one’s brain and Quality of Life (QoL) of substance users. Need for specific questionnaires covering substance use, misuse and dependency and their effects on QoL was felt [3].

Using logistic regression [4] identified five factors for addiction viz. religion, employment status, previous family history, accessibility of drug and peer pressure. Unsanctioned psychoactive drugs including cannabis accounts for an estimated 83 disability-adjusted life years lost per 100 000 population [5].

Attempts made to estimate prevalence and addiction-related harms, number of drug addicts at global and national levels. However, due to poor quality of data, comparisons of countries
and regions across time and space are rather indicative. Major sources of primary data are websites of the World Health Organization (WHO), the United Nations Office on Drugs and Crime (UNODC), Institute for Health Metrics and Evaluation, Alberta Gambling Research Institute, etc. Need for standardized and rigorous methods for data collection, and reporting have been highlighted for better assessment in substance use and addiction burden [6]. Substance abuse research data containing categorical and continuous variables, clustering of individuals in selected areas, significant missing data and outliers, non-representative samples, lack repeated measurements. Even the independent observations assumption may be violated when data are collected from participants who are clustered within larger units, say schools, clinics, communities. Such cluster-wise dependency may not result in independent observations [7].

Thus, statistical analysis of addiction data and the relevant factors need to be innovative for answering questions how, why, and for whom prevention programs are effective. For example, error scores may not follow normal distribution when the dependent variable is binary. Here, logistic regression could be useful where the regression coefficients are the logarithm of the odds ratio.

The paper aims at describing statistical approaches which fit well with structure of drug addiction data and measures of QoL and facilitate better analysis and interpretations of results.

**Analysis of Drug Addiction Data**

**Frequency based approaches**

Statistical approaches based on frequencies includes selection of dependent variable (Y) drug addiction (like Peer pressure (yes – no), accessibility of drug and a set of independent variables which may cause drug abuses and conducting Chi-square test of independence by

\[
X^2 = \sum_{i=1}^{r} \left( \frac{\text{observed value} - \text{expected value}}{\text{expected value}} \right)^2
\]

where expected value of (i-j) cell = \( \frac{i-th \rowtotal \times j-th \columntotal}{\text{Grand Total}} \)

To investigate gender effect on drug use, consider the following 2×2 contingency table with sample size \( N \):

<table>
<thead>
<tr>
<th>Drug addiction</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive (Success)</td>
<td>Negative (Failure)</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>( f_{11} )</td>
</tr>
<tr>
<td>Female</td>
<td>( f_{21} )</td>
</tr>
<tr>
<td>Total</td>
<td>Column total (( C_1 ))</td>
</tr>
</tbody>
</table>

Table 1: Contingency Table – Gender vs. Drug addiction.

Measures of association considering cell frequencies of contingency table are:

- Chi-square measure of association
  \[
  X^2 = \sum \left( \frac{f_{\text{observed}} - f_{\text{expected}}}{f_{\text{observed}}} \right)^2
  \]
  \( r \times s \) degrees of freedom (df) for \( r \times s \) contingency table

- Pearson’s Contingency Coefficient \( \phi = \sqrt{\frac{X^2}{N + X^2}} \) to measure relative (strength) of an association between two variables where \( 0 \leq \phi \leq 1 \). The measure can be used for incidence and prevalence studies

- Cramer’s V-Coefficient (\( V \)) \( X^2 \) may get increased by large sample size, even if the variables may not have any substantive relationship. Cramer’s V-Coefficient improves association by
  \[
  v = \sqrt{\frac{X^2}{N(q-1)}}
  \]
  \( q = \min (r, c) \) and \( 0 \leq v \leq 1 \)

Hypothesis frequently used in measures of association is

\( H_0 \): The two attributes are independent (no significant association between the two attributes) is tested against the alternate hypothesis

\( H_1 \): The two attributes are dependent (there is significant association between the two attributes)

\( X^2 \) measure of association, C-coefficient (\( C \)), V-coefficient, etc. can be applied to find association with clinical findings related to the drug addiction along with testing of significance of association.

Here, proportion of success are

\[
z = \frac{p - P_0}{\sqrt{p(1-p)\left(\frac{1}{R_1} + \frac{1}{R_2}\right)}}
\]

Let population proportions are \( \Pi_1 = \Pi_2 \) for large/moderate sample size is

\[
z = \frac{p - P_0}{\sqrt{p(1-p)\left(\frac{1}{R_1} + \frac{1}{R_2}\right)}} \sim N(0,1)\text{ where } p = \frac{f_{11} + f_{21}}{N}
\]

For small sample, Fisher’s exact test is preferred to test equality of population proportions

Wald Confidence interval (CI) for \( (\Pi_1 - \Pi_2) \) at 95% level is

\[
w = \left( f_1 - P_2 \right) \pm 1.96 \sqrt{\frac{P_1(1-P_1)}{R_1 + 2} + \frac{P_2(1-P_2)}{R_2 + 2}}
\]

For small samples, Wald CI for \( (\Pi_1 - \Pi_2) \) is improved by considering

\[
\hat{P}_1 = \frac{f_{11} + 1}{R_1 + 2} \quad \hat{P}_2 = \frac{f_{21} + 1}{R_2 + 2} \quad \hat{P}_1(1-P_1) + \hat{P}_2(1-P_2)
\]

in each of the two samples and taking Agresti-Caffo-CI for \( (\Pi_1 - \Pi_2) \) at 95% level as

\[
\hat{P}_1(1-P_1) + \hat{P}_2(1-P_2)
\]

\[
\pm 1.96 \sqrt{\frac{\hat{P}_1(1-P_1)}{R_1 + 2} + \frac{\hat{P}_2(1-P_2)}{R_2 + 2}}
\]

The Agresti-Caffo CI is closer to the nominal level than the Wald-CI. Thus, Agrestic-Caffo-CI is preferred. 2x2 table can also help to find prevalence of exposure by

\[
\text{Odd Ratio (OR)} = \frac{f_{11} \times f_{22}}{f_{12} \times f_{21}} \quad \text{OR Relative Risk} \quad \frac{\Pi_1}{\Pi_2}
\]

which is estimated by \( \frac{P_1}{P_2} \). OR compares categorical outcome data where

\[
\text{OR} = \frac{\Pi_1}{\Pi_2} \quad \text{OR} = \frac{1-\Pi_2}{1-\Pi_1}
\]

of exposure between two groups formed by outcome and \( RR \) compares incidence of disease between two groups formed by outcome. Both fail if the assumption of independence is violated. CI of OR and RR can be formed to reflect range of uncertainty. OR compares categorical outcome
Correlations

Point bi-sarial correlation is suitable to find correlation between categorical variable (Y) (say drug addict or not) and a continuous variable (X) (say age, monthly income, etc.) and is given by

\[ r_{pb} = \frac{M_1 - M_0}{\sqrt{\left(\frac{n_1 n_2}{n}\right)}} \]

\[ M_1 : \text{the group with } Y = 1 \text{ (sample size}_1) \]

\[ M_0 : \text{the group with } Y = 0 \text{ (sample size}_2) \text{ } n = n_1 + n_2 \]

\[ S_n = \sqrt{\frac{1}{n} \sum (X_i - \bar{X})^2} \text{ and } -1 \leq r_{pbs} \leq 1 \]

For \( H_0: r_{pbs} = 0 \), the test statistics is

\[ t = \frac{r_{pbs} - 1}{\sqrt{\frac{1 - r_{pbs}^2}{n}}} \]

\[ (n_1 - n_2 - 2) \text{ df.} \]

Logistic Regression (LR)

LR is used when the dependent variable is qualitative or categorical. For categorical dependent variable, several assumptions of the ordinary least square are violated. Qualitative response variable is either binary or have multiple categories. The linear logistic regression model is given by:

\[ \ln \frac{\pi}{1 - \pi} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k \text{ where } 0 \leq \pi \leq 1 \]

is the probability of success and \( (1 - \Pi) \): the probability of failure

\[ \beta_0 : \text{Constant term} \]

\[ \beta_i : \text{Coefficients of } i-th \text{ independent variables, } i = 1, 2, 3, k \]

\[ X_i : i-th \text{ independent variable } \forall i = 1, 2, 3, k \]

Here, probabiity of sucess

\[ \text{probability of failure} = \frac{\Pi}{1 - \Pi} \]

is odd ratio of success and (ii) \( \text{Exp}(\beta) \) where \( j = 1, 2, \ldots k \) is a factor by which the odds of occurrence of success change by a unit increase in the \( j \)-th independent variable.

Test of individual LR regression coefficient i.e., \( H_0: \beta = 0 \) against \( H_1: \beta \neq 0 \) is done by the Wald test by

\[ \frac{\hat{\theta}_m}{\text{standard error of } \hat{\theta}_m} \]

where standard error (SE) of \( \hat{\theta}_m \)

\[ \hat{\theta}_m = \sqrt{\text{var}(\beta)} \text{ and } N(0,1) \text{ asymptotically.} \]

Equivalently,

\[ W = \left( \frac{\hat{\theta}_m}{\sqrt{\text{var}(\hat{\theta}_m)}} \right)^2 \]

can be used where \( W \sim \chi^2 \)

approximately \( H_0 \). Rejection of \( H_0 \Rightarrow X \) adds something to the model.

One can use CI for the odds ratio to determine whether odds ratio \( = 1 \). If the CI does not contain one, then we conclude that the odds ratio is statistically significant. For small samples, reliability of Wald statistics is questionable [4]. However, LR method is sensitive to dependent variables and the researcher need to choose them correctly.

Multilevel analysis

Intra-class correlation is used to evaluate cluster-based dependency data [8]. However, both positive and negative intra-class correlation may influence Type I error (unjust rejection of). Multilevel analysis enabling simultaneous analysis of several levels of data can better analyze clustered data for cluster data, multilevel model includes individuals at Level 1 and schools (say) at Level 2. Relationships between \( i-th \) individual and \( j-th \) school are given by the following three equations:

Individial (Level 1): \[ Y_{ij} = \beta_0 + \beta_1 X_{ij} + \epsilon_{ij} \] (1)

School (Level 2) \[ \beta_{oj} = \gamma_{00} + \gamma_{0j} W_j + u_{oj} \] (2)

\[ \beta_{ij} = \gamma_{10} + u_{ij} \] (3)

The three equations show relationship between individual-level predictors \( X \)-matrix (gender, age, other relevant covariates) and school-level predictors in the \( W \)-matrix (like assignment to program or control groups or other school characteristics). Estimation of error terms \( \varepsilon \), \( \eta \), \( u \) in \((1,2,3) \) respectively allow non-zero intra-class correlations which are incorporated in the analysis. In addition, multilevel models can also be used to study effects at the cross-levels (say, school effect on individuals) or third order cross-levels. 

School Individual Classroom. However, applications of higher order cross-levels are rare for drug prevention studies [9]. One limitation of multilevel analysis is weakened causal inference without randomization. However, there could be situations where randomization is either not possible or may be unethical [10-11] studied Fighting Back program focusing on the use of multiple comparison sites for each treatment community where only respondents were selected randomly and not the cities or communities or sites. In addition, coding for each variable is important, since intercepts and slopes at each level are explained at the next highest level. If intercepts or slopes have no useful interpretation, then it is not meaningful to explain them at higher levels. Multilevel analysis has been used to substance abuse prevention and cross-site evaluation of CSAP’s high-risk youth programs CSAP’s community partnerships [7].

Covariance Structure Analysis (CSA)

To know the extent to which predictors differ across subgroups (gender, income, racial groups, etc.), CSA are used to test the equality or invariance of effects across groups, using test comparing the model with and without parameters freed across groups. Disadvantages of CSA are needed to check assumption of multivariate normal distribution and computation of fourth order moments.

Asymmetric data

Drug addiction data are often skewed and non-normal. Computer-intensive methods like bootstrap, randomization tests, etc. use observed data and finds the significance of an effect without making assumptions about underlying distributions and thus, help in analysis of substance abuse research data. However, such computer-intensive approaches have not been widely used in substance abuse data.
Quality of Life

People with substance use generally have lower QoL-scores than those without such problems [12]. Substance use does not have enduring benefits for subjective feelings of wellbeing [13]. Like drug addiction data, QoL measures suffer from problems relating to meaningful application of statistical analysis. In the area of substance abuse, QoL has been evaluated covering among others functioning, well-being and life satisfaction [14–; 15] Studies to find relationship between QoL, treatment and substance use have mostly considered habits of consuming alcohol with few studies reporting on drug dependent samples. While abstinence alone may have little effect on QoL [16], gain on QoL may accrue gradually with increasing length of abstinence exceeding the initial six months [17] Investigation of QoL in the addiction field is in its infancy and further works are suggested [18].

QoL for substance users have been evaluated using generic instruments like 36-Item Short Form Health Survey questionnaire (SF-36), World Health Organization Quality of Life Assessment-BREF (WHOQOL-BREF), specific tool like Drug Users Quality of Life Scale (DUQOL) [19-20].

Psychometric properties of SF-36 and WHOQOL-BREF emerging from various populations may not be applicable to substance users. Moreover, scales for assessing QoL differ with respect to number of items, item formats, scoring methods and are not comparable [21]. felt need of clarifications of WHOQOL-BREF.

In SF 36, original item scores are rescaled to range between 0 to100. A high score indicates higher favorable health-state. However, total SF-36 scores are not defined due to several dimensions being measured by the scale (http://www. webcitation.org/6cfeefPkf). Descriptive statistics, reliability, validity, etc., are obtained separately for each sub-scale. Mean, SD and shape of distribution, reliability, validity, and discriminating power are different for Yes-No type, K-point scales for K= 3, 5, 6-point [22]. Studies to investigate factor structure of SF-36 using factor analysis [23], structural equation model [24]. confirmed multidimensional structure of the SF-36 [25]. found large number of articles reporting calculation of total SF-36 scores by different ways including arithmetic average of eight sub-classes.

(Table 2) describes major features of QoL scales used for substance users

<table>
<thead>
<tr>
<th>Scale</th>
<th>Sub-scales (Dimensions)</th>
<th>No. of levels (K)</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>SF 36 (36 items)</td>
<td>Physical functioning</td>
<td>K=3</td>
<td>1 is recorded as 0 and 3 is recorded as 100</td>
</tr>
<tr>
<td></td>
<td>Role limitations due to physical health</td>
<td>K =2: Binary</td>
<td>1 is recorded as 0 and 2 is recorded as 100</td>
</tr>
<tr>
<td></td>
<td>Role limitations due to emotional problems</td>
<td>K =2: Binary</td>
<td>1 is recorded as 0 and 2 is recorded as 100</td>
</tr>
<tr>
<td></td>
<td>Energy/ Fatigue</td>
<td>K=6</td>
<td>For item 23, 27 (1:100 and 6:0), requiring reverse scoring to ensure higher score ⟹ better QoL. For item 29, 31(6:100 and 1:0)</td>
</tr>
<tr>
<td></td>
<td>Emotional well-being</td>
<td>K =6</td>
<td>Item 24, 25, 26, 28, 30 (6:100 and 1:0) Item 26 (1:100 and 6:0) – requires reverse scoring</td>
</tr>
<tr>
<td></td>
<td>Social functioning</td>
<td>K =5</td>
<td>Item 20 (1:100 and 5:0) requires reverse scoring and Item 32 (1:0 and 5:100)</td>
</tr>
<tr>
<td></td>
<td>Pain</td>
<td>K=6 for Item 21 &amp;</td>
<td>21 (1:100 and 6:0) and 22 (1:100 and 5:0)</td>
</tr>
<tr>
<td></td>
<td>General health</td>
<td>K=5</td>
<td>Item 1, 34, 36 (1:100 and 5:0). For item 33, 35 (1:0 and 5:100)</td>
</tr>
<tr>
<td>WHOQOL-BREF (26-items)</td>
<td>Physical health</td>
<td>7 items, K =5</td>
<td>Computation of Domain scores (1) reversing scoring two items of physical domain and one item of psychological domain, (2) using the mean item-score of each domain and (3) multiplying the mean scores by 4 to get domain scores, Domain scores are transformed linearly to a 0-100 scale.</td>
</tr>
<tr>
<td></td>
<td>Psychological health</td>
<td>6 items, K =5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Social relationships</td>
<td>3 items, K =5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Environmental health</td>
<td>8 items, K =5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>General health</td>
<td>2 items, K =5</td>
<td></td>
</tr>
<tr>
<td>DUQOL (22-items)</td>
<td></td>
<td>K=7 for each item</td>
<td>Item score ranges from 1 (very unsatisfied) to 7 (very satisfied). Here higher score indicates higher satisfaction with quality of life.</td>
</tr>
</tbody>
</table>

Table 2: Major features of illustrative QoL scales.

Observations

Generated data are ordinal and discrete.

Assume constant distance between two successive levels of an item (Equidistance property). For a 7-point item, equidistant property requires constant value of distance between j-th and (j+I)-th levels $\forall j = 1, 2, 3, 4, 5, 6$

Arithmetic averages requiring equidistant scores are not meaningful for ordinal item scores [26].

Summative scores for domain and test giving equal importance to items and domains are un-justified due to different values of inter-item correlations, item-total correlations and factor loadings of the items and domains.

Mean, SD of test tends to increase with increase in number of levels and may influence mean more than the underlying variable [27].

Distribution of scores of items, domain and test are different and skewed. If two variables X and Y follow two different distributions, $X + Y = Z$ is most meaningful if
\((p = Z = z) = p(X = Y = Z = X)\) for discrete case and
\(p(z = z) = P(X + Y \leq Z) = \int_0^z \left( \int f_{X,Y}(x,t-x)dx \right) dt\) for continuous case. Thus, it is necessary to know probability density function (pdf) of \(X\) and \(Y\) and their convolution.

PCA, FA, \(t\)-test, paired \(t\)-test, \(F\)-test, etc., assume normal distribution of the variables under study. Results may go wrong if assumptions of the techniques used are violated. Even, high \(r_w\) may not imply linearity between \(X\) and \(Y\). If \(r_1, 2, 3, 4, 5 > 0.9\) and \(r_{x, x'} > 0.9\) despite each of \(X^2 X^3\), is non-linear function of \(X\). Linearity between \(X\) and \(Y\) can be tested by checking normality of error score by testing \(H_i: S_i^2 = 0\) where \(S_i^2 = \frac{1}{n} \sum (Y_i - \bar{Y}_i)^2\) for a sample size \(n\) denotes variance of error scores.

Proposed method

[28] proposed method of converting ordinal scores of \(i\)-th item to continuous scores satisfying equidistant property (scores) in ratio scale and further transforming -scores linearly to proposed scores (scores) in ratio scale and further transforming -scores linearly to proposed scores (scores) where \(1 \leq P \leq 100\) and normal distribution, irrespective of number of levels (\(K\)) of an item, \(K = 2, 3, 4, 5\) and so on. Equidistant scores are obtained by assigning different weights to different levels of different items. Let be the weight to the \(j\)-th level of the \(i\)-th item (say 5-point) so that \(w_{i1}, 2w_{i2}, 3w_{i3}, 4w_{i4}, 5w_{is}\) forms and arithmetic progression. Clearly \(j, w_y = (j - 1)w_{ij-1}\) is constant, value of which will be different for different items. A positive value of constant will make E-scores monotonic. Here \(w_y, s\) are based on empirical probabilities obtained from the basic item score matrix. If \(f_{ij} = 0\) for a particular \(j\)-th level of an item, the method fails and can be taken as zero value for scoring Likert items as weighted sum. Different weights to different levels of different items may break ties of individual total scores which are common in usual summative scores and thus, improve discriminating power of the test or sub-test.

\(E_i\)-scores are normalized by \(z_i = \frac{E_i - \overline{E}}{SD(E)} \sim N(0,1)\) To avoid negative values, \(Z\)-scores are transformed to proposed scores \((p - scores)\) by \(p = \frac{99*(Z_i - Minz_i)}{Maxz_i - Minz_i} + 1\) \(1 \leq p \leq 100\) follows \(P_i\) normal. Domain score is the sum of \(\sum\) of the domain and test score is sum of all domain scores, each following normal distribution, parameters of which can be estimated from the data.

Advantages of proposed scores

Better admissibility of arithmetic average where scores are normally distributed.
Measurement of total test score for each individual like \(SF_{total}\).
Dimension score \(D_i\) and proposed scale scores are continuous, monotonic, normal and facilitate undertaking parametric analysis including estimation of population mean \((\mu)\) population variance \((\sigma^2)\) confidence interval of testing statistical hypothesis like \(H_0: \mu_1 = \mu_2\) or \(H_0: \sigma_1^2 = \sigma_2^2\). either for longitudinal data or snap-shot data.

\(P\)-scores reduce significantly number of tied scores. Thus, most of the individuals can be given unique ranks.

Progress path of a patient or a group of patients across time i.e., trajectories over time is analogous to hazard function and can be used to compare response to treatments from the start. This is especially important for substance users who are undergoing treatments. Such trajectories can help to identify high-risk groups.

For two QoL scales \(X\) with normal pdf and \(Y\) with normal pdf one can find regression equation of the form \(y = \alpha + \beta X\). Predict \(X\) with knowledge of \(Y\). However, the two regression lines differ and thus, relationship between \(X\) and \(Y\) will not be unique. Better is to find equivalent score combinations \((X_i, Y_i)\) of the two QoL scales by solving the equation \(\int_{-\infty}^{\infty} P(x)dx = \int_{-\infty}^{\infty} g(y)dy\) for a given value say i.e. area of the curve up to \(=\) area of the curve up to [29]. This avoids the problems of linear equating or percentile equating. The equation can be solved using standard normal table. The method of finding equivalent score-combinations is possible even if the scales have different number of items or dimensions.

Psychometric qualities

Validity

Validity of a QoL instrument is often reported as correlation between the QoL score (\(X\)) in question and another scale (\(Y\)) which may be influenced by mismatch of dimensions covered, different score ranges, different distributions of scores of the two scales, etc. Moreover, reflects validity of \(X\) and \(Y\) also, \(r_{xy}\) high, need of two different scales may be questioned. Normality satisfies the assumptions of PCA enabling computation of eigenvalue and consider factorial validity as \(\lambda_i\), \( \sum \lambda_i \) is the highest eigenvalue associated with the first principal component. Factorial validity reflects the main factor for which the test was developed and accounts for \(\lambda_i \sum \lambda_i \times 100\) percent of overall variability. Such factorial validity from single administration of a test avoids the problems of construct validity and is independent of criterion scale [30].

Reliability

Avoiding uni-dimensionality assumption of Cronbach alpha, [31] proposed to dichotomize a test in two parallel subtests (\(g\)-th and \(h\)-th) and to compute test reliability as per theoretical definition \(r_{tt(\text{Theoretical})} = \frac{\text{True score variance (}S_g^2\text{)}}{\text{Observed score variance (}S_g^2\text{)}}\)

where \(Error\) variance \(S_g^2\) for Sample Size \(N\) is

\[S_g^2 = \frac{1}{N} \sum_{i=1}^{N} (||X_{g}\||^2 + ||X_{h}\||^2 - 2||X_{g}\|| ||X_{h}\|| \cos \theta_{gh})\]

where \(||X_{g}\||\) denotes length of the \(g\)-th vector \(||X_{h}\||\) is defined similarly and \(\theta_{gh}\) is the angle between the \(g\)-th and \(h\)-th vectors given by

\[\cos \theta_{gh} = \frac{\sum_{i=1}^{N} X_{g\text{-th}} X_{h\text{-th}}}{||X_{g}\|| \cdot ||X_{h}\||}\]
From (4),

\[ S_i = S_{i-1} - S'_{i-2} \]

and

\[ n(\text{Theoretical}) = \frac{S_i^2}{S_X} = 1 - \frac{S_X^2}{S_Y} = 1 - \frac{\text{[\{X_i^2 + |X_i|^2 - 2|X_i||X_i|\cos \varphi \}]}{\text{NS}_{i-1}^2} \]

[5]

Normally distributed P-scores help to know whether two subtests g-th and h-th are parallel if and only if \( H_{i}: \frac{P_g}{P_h} = \frac{P_{i-1}}{P_{i-2}} \). \( H_{i}: \sigma^2_{P_g} = \sigma^2_{P_{i-1}} = \sigma^2_{P_{i-2}} \) are accepted.

Equation (5), in terms of P-scores helps to test \( H_{0: \gamma(\text{Theoretical}) = 1} \) which boils down to test \( H_{i: \sigma^2_P = \sigma^2} \) by F-test [32].

Discussion and Conclusion

Assumptions of standard parametric statistical methods are usually not satisfied by drug addiction data. Innovative use of statistical methods helps to better analyze and interpret substance abuse data, including better assessment of resultant QoL of substance users. Frequency based approaches [33] are simple to comprehend and calculate and involve no assumptions. \( 2 \times 2 \) Contingency table can be extended to \( r \times S \) contingency table for \( r \)-rows and \( s \)-columns to accommodate levels of drug addictions say mild, moderate, severe, etc. and various level of income or age-group. However, very low frequencies in off-diagonal cells can distort the result.

For only two outcomes [34] and a set of continuous independent variables, point bi-serial correlation is suitable and logistic regression model helps to assess and test significant logistic regression coefficient. Multilevel models are [35] encouraging to investigate effects at the cross-levels [36] (say, school effect on individual) or third order cross-levels like School×Individual × Classroom etc. However, causal inference without randomization of all levels and cross-levels is a weak point since intercepts/slopes at higher levels depends on same at each lower level. But there are situations where randomization is not possible or may be unethical.

Proposed method of transforming ordinal raw scores of QoL items to follow normal distribution enables statistical analysis under parametric set up leading to meaningful comparisons and inferences, finding equivalent scores of [37]. two or more scales along with better computation of psychometric qualities like reliability, validity, responsiveness, etc. Considering theoretical advantages, the proposed method of transforming raw scores of QoL items to follow normal distribution is recommended. Future studies with longitudinal data can be undertaken to find sensitivity of the proposed score over time with emphasis on progression of disease and to different therapeutic interventions, etc.

Reference


