**Greetings everyone,**

**Problem:** Need suggestions for Methodology/Statistical Section.

**Note:** Dear members, no help is required for conducting or writing analyses. Only need help with finalizing the appropriate design/vetting the validity of the design.

I am conducting a study (for publishing in a Journal) to compare the effects of gender and locus of control in relation to stress levels, in a sample of late adolescents in the state of Odisha, India. I currently live in Odisha, India.

The data has already been collected and is as follows-

Males- 74

Females- 67

Total Participants- 74+67=141

Age- 17 to 19 years.

**Variables in the Study**

**IV 1: Gender**- 2 levels ( Males, Females):

**IV 2: Locus of Control**- 2 levels of LOC ( Internal LOC vs External LOC); : Measurement using Rotter's LOC.

[Locus of control is of two types-internal LOC and External LOC. People with intrinsic LOC believe they control their successes and failures. Contrary to the beliefs of internals, externals believe that successes and failures are beyond their control, and that fate and destiny dictate their lives. ]

**DV** : Stress levels of Adolescents; Measurement using Sheldon Cohen Perceived Stress Scale (PSS).

**Instruments**

**For LOC: Rotter's Scale of Locus of Control:** This scale is a 23-item questionnaire, and it measures generalized expectancies for internal versus external control of reinforcement.

Scores range from 0 to 23.

Norms of Scales- A low score indicates an internal control, while a high score indicates external control. No predefined absolute cut-off point for discriminating between Internals vs externals for this scale has been mentioned in the norms.

**For Stress Levels: Sheldon Cohen Perceived Stress Scale (PSS):** The Perceived Stress Scale (PSS), published in 1983 by Sheldon Cohen, consists of 10 items that measure the degree to which situations in one’s life are appraised as stressful.

Range of Scores- 0 to 40

Norms of Scales-

· Scores ranging from 0-13 would be considered low stress.

· Scores ranging from 14-26 would be considered moderate stress.

· Scores ranging from 27-40 would be considered high perceived stress.

**Choice of Statistical Test**

There are two independent variables: Gender and Locus of Control. The independent variable of gender has two levels: male and female. The independent variable of locus of control has two levels: internal locus of control and external locus of control. An investigation will be carried out for the effect of these independent variables on stress scores. Stress has been chosen as the dependent variable. Since there are two categorical independent variables and one continuous dependent variable; a two factor ANOVA/Regression can be applied to examine whether gender and locus of control influences the stress of subjects or not.

**Problems Encountered-** I have two options for conducting two way ANOVA.

**Option 1**: Take unequal sample sizes ( Males 74 and Females 67). See how many of them are Internals or Externals as per norms of the LOC scale. Then conduct ANOVA for unequal sample sizes. **Hurdle:** If I take all the samples, then there will be unequal number of internal and external subjects per cell. In other words, the LOC variable will partially become dependent on gender. If this is so, will there be any problem when considering LOC as an IV? How to deal with this ?

**Possible Option:** Use ANOVA with Type III SS. Will it work?

**Option 2**: Use Median split to decide a cut-off point. The second scale LOC is not a categorical scale, as it does not defines absolute point for differentiating between Internals and Externals. But conducting a median split for a non-categorical scale will confound the categorization for those subjects whose score will lie around the median.

**Possible Option:** After median split, arrange the participants in ascending order and from this list, only select subjects with extreme scores, from both top and bottom. Equal number of subjects will be taken from top and bottom and then be divided into groups with equal sample sizes, say each group with n participants. In other words: out of N= 141 a total of X extreme case samples selected from top and bottom; this X extreme samples will be divided as X/4 = n, such that there will be n Male adolescents with an Internal locus of control, n Male adolescents with an External locus of control, n Female adolescents with an Internal locus of control and n Female adolescents with External LOC;

where n + n + n + n = X; X ⊂ N.

**Summary of my Questions/Problems:**

A. Is the Option 2 a valid method? What are the flaws and how to improve it?

B. Which option I should choose- Option 1 or Option 2? Or any other suggestion?

C. If you were the Reviewer/External for this paper, which portions you would highlight as weak/confounding/flawed areas in the context of statistics for-

1. **Sample Size Selection (Total Participants- 74 males +67 females =141)**
2. **Instruments Selection**
3. **Choice of Statistical Test (Two Way ANOVA vs Regression)**
4. **Categorization of Subjects (using Median Split or Using Unequal Samples)**
5. **Statistical Analyses**

I request you to kindly suggest methods to deal with the above problems.

Thank you for your feedback.

Anmol