Is My Hypothesis Statistically Significant—Part II

Dr. (Mrs) Marie Farrell and Dr. (Miss) Aparna Bhaduri

"How do I find out by looking at the Chi-square value whether or not this is significant?"

You can look at the Chi-square table given in most statistical books. In order to read the table we need to understand "degrees of freedom" or "df". For finding df we minus 1 each from the total columns and total rows. For example, in a 2 x 2 table, df would be (2—1) (2—1)=1; in the case of Table 3 which is a 4 x 4 table, df=(4—1) (4—1)=9. The Chi-square table indicates that with 1 df a Chi square of 6.64 or more is significant at .01 level and 3.84 or more is significant at .05 level. Computed X^2 for Table 1=5.31, df=1.

"Can you state the analysis in statistical terms?'

For your easy reference let's recall, the null hypothesis: "There is no difference in women's preference for trained health workers for deliveries in villages A and B." The null hypothesis is rejected as the Chi-square is 5.31, which is significant at .05 level and the probability that our obtained difference is due to sampling errors is less than five per cent. If we had set .01 as the level of significance, we state that Chi-square is 5.31 which is not significant at .01 levels and thus we have failed to reject the null hypothesis.

"What are the basic rules that I need to remember for getting data where the Chi-square test is applied?"

Use of Chi-square

The essential point is to develop a sound method for the categorization of data. The most pertinent points to remember for categorization are as follows: (1) categories must be exhaustive, (2) categories must be mutually exclusive, (3) they should be independent and, (4) the categories are based on a single principle of classification. Let's examine each of these rules.[2]

The categories must be exhaustive means that the data related to each subject must find a place somewhere in the categories listed. If you take categories say, for occupation of the subjects as "teachers", "farmers", "technical worker", etc, you need to include "others" as another category for those who do not fall in categories of teachers farmers or technical workers. It is not possible to make 100 categories to include occupations. So we take the categorious where most of the workers would fall and keep "others" as a category where the remaining subjects are put together.

"How are mutually exclusive categories different from exhaustive categories?"

We are actually discussing the rules for categorization. It is not a question of one or the other, but we are to see that we take into consideration all relevant aspects of the variables under consideration.

Let us examine the meaning of mutually exclusive categories. In a study where the subjects are health supervisors, we want to know the previous experience of the subjects. The previous experience listed are: worked as "nurse", "health educator", "sanitary inspector", etc. It is possible that subjects have had the experience in more than one category listed above. This will obviously lead to percentages which would exceed 100 per cent as one respondent may respond to more than one category. Hence in each category each response has to be handled as a separate item; e.g. "has worked as nurse/has not worked as nurse", "has worked as health education/has not worked as health educators", etc.

"Does this mean each should be 'independent'?"

Yes, that's the third consideration to be made when we say that the categories should be independent. The allocation of a subject to one category must not affect the allocation of another subject to that or any other category. If the categories are "high income group", "middle income group", and "low income group" and we have to categories the responses of 100 participants, we may be faced with the consequences of a mismatch between the "variable assessed" and the "sample assessed."

The next rule is to see that the categories are based on a single principle of classification. The sample given in Table 2 includes assessment of nutritional status at the beginning and end of the

school year. The norms that govern the classification that a child's nutritional status falls into categories in A, B, C or D should be same for both measurements.

The t-test

"Now I know about a test of significance when the data are in ordinal level. What happens when data are in the form of continuous scores?"

Let us discuss the use of t-test here. The t-test, as we have mentioned earlier, is based on the assumption of normality.

"Does this mean that the variables I am working with are normally distributed in the population from which I have taken the sample?"

Yes, although there are chances that you may not find them normally distributed in the sample if the sample size is small. When the sampling technique is appropriate and the sample size is fairly large, the distribution tends to be normal if in the parent population the distribution is normal. However, there are statistical tests by which one can check whether or not the sample is normally distributed.

"When do I use a t-test?"

The t-test is used if the researcher is concerned with significance of the differences between the means of two sets of data, either taken from the same sample or taken from two sets of random samples belonging to a single population. The following example will clarify this point. To measure the effectiveness of a unique workshop format for teaching research, an experimental design was carried out on a sample of nurses prepared at the graduate level. There were two groups taken from the same population. The workshop format of teaching was given to the experimental group and pre and post-tests for research knowledge were administered to them. The control group was administered pre and post-tests for research knowledge. Two of the hypotheses formulated for the study are given below:

H_1 Nurses who have participated in a two week formalized workshop in nursing research will demonstrate significantly higher post-test than pre-test scores on a measure of research knowledge.

H_2 Nurses who have participated in a two week formalized workshop in nursing research will score significantly higher on a measure of knowledge than nurses who have not been exposed to such a course.[3]

Table 6 shows the data related to pre and post-test scores of knowledge, in the experimental group.

Table 6: Mean, Standard Deviation and t-Value of Pre and Post-test Scores of knowledge and Experimental Group

Variables	X	S.D.	t
Knowledge			
Pretest	9.30	4.37	5.38
Posttest	17.08	2.84	

Looking at the two means we find there is a difference of (17.08 — 9.30) 7.78 points. Now the question is whether the difference of 7.78 points is sufficiently large enough to be significant or can we say that the workshop method of teaching significantly increased the knowledge of research methods in the experimental group. The t value was found to be 5.38. Now let's refer to the t-table in the back of a standard statistical text.

"Should't I find out the degrees of freedom first as I did in case of Chi-square?"

Right you are! In the case of t-test, df is calculated from n (or sample size). Since we have two n (Pretest N_1 and post-test N_2) the df is equal to N_1+N_2 —2. There were 13 subjects in the experimental group; hence df=13+13 —2=24. Are you now ready to see the table for t distribution?

"Yes, I find that a t-value of 2.80 or more at df=24 is significant at .01 level of significance."

You are getting the idea! Now you can say a t-value of 5.38 at df=24 is significant at .01 level, and thus H_1 is rejected. This means that the nurses who participated in the research workshop showed significantly higher post-test scores than pre-test scores on a measure of research knowledge. "In the above example you have explained how to find the significant difference between the means of two sets of data obtained on the same person. Can t-test be applied to test the differences between the post-test scores of experimental and control groups."

Yes. We have mentioned this earlier. The t-test is also used for testing the hypothesis of zero difference (null hypothesis) between two means using independent samples. Let us examine H_2 stated earlier. See Table 7.

Table 7: Mean, Standard Deviation and t-Value of Post-test Scores of Knowledge of Experimental and Control Groups

Groups	Post-test Knowledge X	SD	t
Experimental	17.08	2.84	5.90*
Control	7.20	3.40	

There are two independent samples and the experimental group and the control group; we have means of post-test knowledge scores for both samples. The difference of the mean as shown in Table 7 is 9.88 (17.08 —7.20) points.

You could ask "Is this difference significant at the .01 level". Now look at a t-table. We find that a t-value of 5.90 at df=27 (13+16 —2) between Post-test means for the experimental and control group for knowledge is significant at .01 level. The null hypothesis is therefore rejected and thus the nurses who participated in the nursing research workshop scored significantly higher on a measure of knowledge than nurses who had not been exposed to such a course.

Use of t-test

"What is the reasoning behind t-test?"

When we test hypotheses with t-test, we use the same principle as with the normal curve distribution.

In order to study the relationships between two parameters we begin with an assumption that the relationship is zero. That is, if the two means are X_1 and X_2, then $X_1=X_2$ or in other words the two means are identical. A word of caution in this—we always find a relationship in terms of relative value and not in absolute value, as hypothesis testing is done at certain probability levels particularly in the behavioural sciences. In behavioural science the study of living organisms makes the research more complex as measures collected are often affected by many variables.

In the example of teaching research methods through a workshop, the researchers intended to evaluate the use of the format. In order to obtain a more meaningful interpretation of the effects of the format, a comparison with a group not exposed to the teaching would be most desirable. Thus, there are two things you can do—(1) You can compare the post-test mean with that of pre-test mean of the same group, or (2) by use of the mean from a control group (equated with the experimental group) which does not experience the experimental treatment i.e. the workshop method of teaching research.

"It seems in H_1 you have obtained pairs of measures, that is, each pre-test score is logically paired with a post-test score on the same subject in the experimental group."

Right, and in H_2 we have two independent measures where no such logical pairing is present and neither have we any reason for connecting any given scores in one sample with any given score in the other (because the subjects in the two groups are different from each other).

Now that you have seen various methods of comparing two samples, perhaps you could design a study proposal to compare the effects of triple vaccine in rural children on the incidence of whooping cough; in one village children are vaccinated and another village children have not received any vaccine.

Another exercise might be to study the effectiveness of various contraceptive methods in the population you serve where the dependent measures could be incidence of pregnancy, vaginal bleeding, anaemia and/or other complications.

'I would like to know more about other parametric and non-parametric tests."

That is good. We hope that you will want to explore additional parametric and non-parametric tests for use in additional research.

Remember, t-test, and Chi-square tests are just examples of tests of significance. One must plan the data analysis while stating the hypothesis, keeping in mind the nature of the population, sampling method, size of sample and the type of measurement. There is always one test which is appropriate to your data. Good luck!

REFERENCES

1. Burroughs, C. E. R., **Design and Analyses in Educational Research** London: University of Birmingham, 1975, pp 43-45.
2. **Ibid.,** p. 44.
3. Farrell, M., Bhaduri, A., "A Pilot Study to Assess the Effectiveness of the Workshop Format for Teaching Research Methodology in Nursing", **NIHAE Bulletin** Vol. X. No. 4, 1977, p. 272.

DONATIONS : JUNE-JULY 1980

Nurses Welfare Fund

	Rs.
Mrs. Aminni Mathew, Nilgiri	25
Mrs. S. D. Bhobe, Dubai	1001
Catholic Nurses Guild, Bombay	420
Late Miss I. Dorabji, Bombay (Through Miss J. M. Benjamin)	60
Miss Florence H. Martyn, Canada	1000
Miss C. P. Rachel, Trichur	50
Mrs S. Sharma, Kelvin	55
Miss T. K. Neudoerffer, Ontario	200
Miss S. Chettri Siliguri	25
Mr. K. P. Venugopal, Canada	130
Miss Ruby L. Hobson, USA	80
ICN-TNAI Delegates Fund	
Mr K. P. Venugopal, Canada	200
Miss Ruby L. Hobson, USA	100
General Fund	
Mr. Nathi Ram New Delhi	300
Miss S. Chettri, Siliguri	25
Mrs. Padmini M. Nair, Madras	20
Miss A. Ernestine, Pithapuram	100
Mr. C. I. Abraham, Vallore	3
Mrs. Urmil Chopra, New Delhi	100
Flood Relief Fund	
Miss V. D. Baneanbehram, Kuwait	104
Staff Welfare Fund	
Late Miss I. Dorabji, Bombay (Through Miss J. M. Benjamin)	60
Miss Florence H. Martyn, Canada	1000