DOCUMENT RESUME

ED 451 242 TM 032 470

AUTHOR	Inoue, Yukiko
TITLE	Educational Research and Statistics: Examples of Questions and Answers.
PUB DATE	2001-00-00
NOTE	10p.
PUB TYPE	Reports - Descriptive (141)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*College Students; *Educational Research; Higher Education; *Research Methodology; *Statistical Analysis

ABSTRACT

Some questions that might be asked of students with respect to educational research and related statistical analysis are presented with answers. The questions focus on: (1) the definition of operational variables; (2) null and research hypotheses; (3) the use of analysis of variance (ANOVA); (4) when ANOVA is appropriate; (5) whether statistical assumptions are reasonable; (6) the use of computer software in statistical analysis; (7) a hypothetical table of means; (8) the use of post hoc procedures; (9) significant results and causality; (10) research design and analysis for greater power; (11) regression analysis; (12) alternatives to ANOVA; and (13) appropriate analysis when the dependent variable can only be measured at the nominal level. Some additional questions are attached. (SLD)

Educational Research and Statistics: Examples of Questions and Answers

Yukiko Inoue, PhD
Foundations and Educational Research, College of Education
University of Guam
UOG Station, Mangilao, Guam 96923
Phone: 671-735-2447 Fax: 671-734-3651
Email: yinoue@uog.edu

Educational Research and Statistics: Examples of Questions and Answers

1. *Operationally define the variables being used and describe how you plan to control for confounding variables.*

Definition of operational variables:
DEPENDENT VARIABLES
SCORE: test performance on English (effectiveness of the remedial program)
INDEPENDENT VARIABLES:
Focusing variables:
PROGRAM: have taken the remedial program or have not taken the remedial program.
ETHNICITY: the student is white or black
GENDER: the student is male or female.
Controlling variable:
IQ: the student's IQ score

The effect of PROGRAM on student performance (SCORE) will be tested using analysis of covariance (ANCOVA). Including IQ score as a covariance in this design is a method of controlling the potential confounding of initial group ability differences.

In this case, nominal variables are GENDER (male/female) and ETHNICITY (white/black), and PROGRAM (with/without).

2. *State the null and research hypotheses to be tested.*

Null hypothesis: there will be no differences in the effect of PROGRAM 1) between males and females and 2) between blacks and whites on student English performance.

Research hypothesis: there will be differences in the effect of PROGRAM 1) between males and females and 2) between blacks and whites on student English performance.

3. *Describe how ANOVA is to be used to test the hypotheses.*

A. *Its nature.*
B. *Its assumptions and how they are provided for.*
C. *How it tests the stated hypotheses.*

A. ANOVA should be chosen over regression, particularly because of the interest in testing the interactive effects among PROGRAM, GENDER, and ETHNICITY in the situation.

B. Three major assumptions in ANOVA are:
Independence of errors. The error for one observation should not be related to the error for any other observation. This assumption is often violated when data are collected over a period of time. If we assume such things as data were collected properly, data were collected at the same period of time and no twin subjects in this study, this assumption appears to be reasonable in this situation.

Normality. The values in each group are normally distributed. Just as in the case of *t* tests, ANOVA is a "robust" against departures from normal distribution. As long as the distribution is not extremely different from the normal distribution, the level of significance of ANOVA is not greatly affected by the lack of normality, particularly for large sample size. It is assumed that sample size was large and this assumption would be reasonable in this situation. It is also assumed that an entire population of scores was obtained in a particular condition, the scores would be normally distributed.

Homogeneity of variance. The variance within each population should be equal for all populations. If there is the equal sample size in each group, inferences based upon the F-distribution may not be seriously affected by the unequal variances. Further, tests of homogeneity of variance (Cochran's C test, for example, would indicate the reasonableness of this assumption).

C. The analysis can be done with ANOVA using SCORE as the dependent variable, and PROGRAM, GENDER, and ETHNICITY as the independent variables, along with IQ as the covariate. Accordingly, seven hypotheses will be examined.

4. *State why this ANOVA design and analysis are appropriate (as opposed to some other use of ANOVA).*

ANCOVA is appropriate because of an analysis with the statistical influence of covariate (IQ score) removed from the dependent variable. Theoretically speaking, ANCOVA is what we would get if we could do the ANOVA with the level of the covariates controlled. ANCOVA decreases the error variance by extracting variance that is due to the relationship between the covariate and the dependent variable; therefore, the independent variables may be more likely to show significant effects.

5. *Discuss the reasonableness of the statistical assumptions in this situation.*

The ANOVA can be done as follows:

Step 1 *Null hypotheses:*

There is no PROGRAM effect.
There is no GENDER effect.
There is no ETHNICITY effect.
There is no PORGRAM-GENDER interaction
There is no PROGRAM-ETHNICITY interaction.
There is no GENDER-ETHNICITY interaction.
There is no PROGRAM-GENDER-ETHNICITY interaction.

Step 2 *Alternative hypotheses:*

There is PROGRAM effect.
There is GENDER effect.

There is ETHNICITY effect.
There is PORGRAM-GENDER interaction
There is PROGRAM-ETHNICITY interaction.
There is GENDER-ETHNICITY interaction.
There is PROGRAM-GENDER-ETHNICITY interaction.

Step 3 Set Alpha = .05.

Step 4 Set rejection rules. If F-critical is larger than F-observed, we reject the null hypothesis in each of the hypotheses. If F-critical is smaller than F-observed, we do not reject.

6. *Assuming the data has been collected appropriately, describe how you would use the computer in performing this analysis and reporting your results.*

With the SPSS subprogram ANOVA, the analysis can be done with ANOVA using SCORE as the dependent variable, PROGRAM, ETHNICITY and GENDER as the independent variables and IQ as a covariate. Command for the computer is:
ANOVA SCORE by PROGRAM (1, 2,), GENDER (1, 2), ETHNICITY (1, 2) with IQ

7. *Make up a hypothetical table of means (associated with rejection of at least one null hypothesis) and interpret them.*

Source	df	MS	F
IQ covariate			.045*
PROGRAM (A)			.050
GENDER (B)			.045*
ETHNICITY (C)			.045*
A x B			.090
A x C			.080
B x C			.059
A x B x C			.001*
Error Total			

*$p <. 05$

8. *What post hoc procedure would you utilize to follow-up significant F-ratios. Why?*

Assuming the optimal conditions (equal sample sizes, and no violation of assumption), I will use Tukey method to test which difference in group means contributed to the interaction effect. This is often called the HSD (honestly significant difference) test, and is designed to make all pairwise comparisons while maintaining the experiment wise error rate at the pre-established alpha level for equal number in a subsample ($\alpha_E = .05$, for example, in this case).

9. *Assuming significant results, can this researcher claim causality? Why?*

No, the researcher cannot claim causality. Relationships between two variables are not necessary "cause-and-effect." A causal relationship implies that the independent variable is the cause and the dependent variable is the effect. Some relationships are causal while others may be consequential or functional. In this case, a functional relationship exists when variables are related in some functional way. Thus the researcher's primary concern is with these relationships themselves and not with the reasons they exist.

10. *Describe how you would change the design and analysis to achieve greater power and to control for additional variation in the dependent variable.*

Power is the complement of the probability of rejecting the null hypothesis, and β is accepting the null hypothesis when error occurs if the null hypothesis is not rejected when it is false and should be rejected. Unlike the Type I error rate α, the magnitude of the Type II error rate β is dependent on the actual population value of mean.

There are many ways to improve power, but one way to improve power in this situation is the use of "repeated measure design." This design is a particularly helpful, because all individual differences due to the average response of subjects is removed from the error term, and individual differences are the main reason for within group variability.

11. *Discuss regression analysis in relation to this problem.*

   A. *Describe how regression analysis could have been done to answer the research question.*
   B. *Indicate why regression analysis or the ANOVA described above is better in this situation.*
   C. *Are the assumptions for regression analysis appropriate in this situation? Why?*

A. Multiple regression is the most widely used of the multivariate statistical methods. In this case, we can explore the strength of relationship between the dependent and the independent variables.

B. The ANOVA described above is better than regression analysis. One difficulty with multiple regression is that of multicollinearity.

C. Three assumptions are:
The conditional probability distribution of the dependent variable for given independent variables follows the normal pattern. Assuming that the sample size is large, this assumption is appropriate.

The conditional distribution of the dependent variable for each combination of the independent variables has an identical variance. It will be assumed that tests of homogeneity of variance rejected the null hypothesis of variance.

The values of the dependent variable are independent of one another. Assuming that the data were collected at the same period of time and conducted properly, this assumption is appropriate.

12. *Name two other statistical procedures that could have been used in place of ANOVA. For each:*

  A. *Indicate why this procedure or the ANOVA described above is better in this situation.*
  B. *Describe the essential characteristics of this statistical procedure, being sure to include its primary purpose.*

t test

A. ANCOVA is basically an analysis of variance with the statistical influence of one or more variables (called covariate) removed from the dependent variable. ANCOVA is better, because by including IQ as the covariate we can control the potential confounding of initial group differences. In addition, we can test the interactive effects.

B. Researchers use the *t* test most often to compare the means of two groups. If the two sample means are far enough apart, the *t* test will yield a significant difference.

Chi-square test

A. In this situation we are mainly interested in the SCORE and PROGRAM relationship, but we suspect that another variables (GENDER and ETHNICITY) may influence. In terms of the purpose of the major question, ANCOVA will be better for this situation.

B. Another frequently used probability distribution is the chi-square distribution. The technique is of the goodness-of-fit type in which we test for significant differences between the observed distribution of data among categories and the expected distribution based on the null hypothesis. Chi-square is useful in cases of one-sample analysis, two independent samples, or *K* independent samples. It must be calculated with actual counts rather than percentage.

13. *Assume the dependent variable could only be measured at the nominal level of measurement. Describe an appropriate analytic procedure for this situation.*

Log-linear analysis will be appropriate. Using a cross-classification table and the chi-square statistics, we can test whether two variables (PROGRAM and SCORE) are related. We also want to know the effect of additional variable (GENDER and ETHNICITY) on the relationships that we are examining. We can always make a cross tabulation table of all of the variables, yet this would be very difficult to interpret. With a log-linear analysis we try to predict the number of cases in a cell of cross tabulation, based on the values of the individual variables and on their combinations. We see whether certain combinations of values are more likely or less likely to occur than others, and this tells us about the relationships among the variables.

Additional Questions

1 *Describe a research question that would call for a one-way analysis of variance (ANCOVA) as the statistical procedure appropriate for addressing the question. Define the independent and dependent variables.*

Suppose that three different methods are used for teaching English as a second language (ESL) at the University of Evergreen. Each method is applied to one of three groups of ESL students of the university selected at random. At the end of the course, the students are given a standardized test. The *research question* we wish to answer here is:

> On the basis of these sample data, shall we reject the null hypothesis that the three teaching methods are equally effective?

Let μ_1, μ_2, and μ_3 be, respectively, the means of test scores earned by all the students taught by the three teaching methods, then null hypothesis that is to be tested is μ_1, μ_2, and μ_3 against the alternative that they are all unequal. The analysis can be done with a one-way analysis of variance, using teaching method as the *independent variable* and the standardized test score as the *dependent variable.*

2. *Describe a research question that would call for a one-way multivariate analysis of variance (MANOVA) as the statistical procedure appropriate for addressing the question. Define the independent and dependent variables.*

Suppose that Virtual Reality (VR) is used for teaching College Algebra at the University of Evergreen. In order to assess the students' perceived usefulness of VR, at the end of the course, all the students, who used VR as a learning tool, are required to complete the student virtual reality experience questionnaire. We want to determine whether the four groups (namely, White, Black, Asian, and Hispanic Origin) differ on the average on a set of dependent variables. The *research question* we wish to answer here is:

> On the basis of these sample data, shall we reject the null hypothesis that the four groups of students do not differ in their perception of the VR support?

The analysis can be done with a one-way multivariate analysis of variance, using a subset of the questionnaire (for example, satisfaction, effectiveness, and accessibility) as the *dependent variables* and the three groups of students as the independent variables.

3. *There are several possibilities for follow-up analyses following a significant multivariate test in MANOVA. (1) Some go immediately from the multivariate setting to the univariate. (2) Others remain in a multivariate context and only at the final stage examine individual variables. (3) Still another always remains in a multivariate context. Select one of each type (one approach representing each of the three general approaches), and for each describe the procedure and its rationale.*

Using four groups of students of the University of Evergreen (their majors are Business, Education, Arts and Sciences, and Agriculture), a study is conducted to determine whether there are differences in their perceptions of growth and development in the four areas: 1) mathematics and science, 2) writing skills, 3) personal development, and 4) perspectives of the world. The students are selected at random.

A set of dependent variables: Mathematics and science, writing skills, personal development, and perspectives of the world. And four groups of students are independent variables: Group 1: Business (n = 48); Group 2: Education (n = 48); Group 3: Arts and Sciences (n = 48); and Group 4: Agriculture (n = 48).

(1) Hotelling's T^2 and Univariate *t* test

Procedure:
We follow a significant ominous multivariate result by all pairwise multivariate tests (T^2s) to determine which pairs of groups differ significantly on the set of variables. Then, we use univariate *t* tests at the .05 level in this case to determine which of the individual variables is contributing to the significant multivariate pairwise differences. For the above four groups of students, there will be six Hotelling's T^2s; we do each T^2, for example, at the .10/6 = .167.

Rationale:
Compared to both Hotelling's T^2 and Tukey Confidence Interval and Roy-Bose Simultaneous Confidence Intervals, this procedure has the best power. In other words, it is least conservative in terms of protecting against type I error but still has fairly good control on type I error. Further, it has merit as long as we recognize that the individual variables identified must be treated somewhat tenuously.

(2) Hotelling's T^2 and Tukey Confidence Interval

Procedure:
We follow a significant omnibus multivariate result by all pairwise multivariate tests, and then we apply the Tukey simultaneous confidence intervals to determine which of the individual variables are contributing to each pairwise significant multivariate result. This procedure has better protection against type I error, especially if we set the experimentwise error rate for each variable that we are applying the Tukey to such that the overall α is at maximum .10 in this case. In the Tukey procedure, the studentized range statistics is used, and the critical values for it are in the table (Percentile Points of Studentized Range). If the interval does not cover "0," the population means of the groups are significantly different.

Rationale:
Using the Tukey procedure, we can examine all pairwise group differences on a variable with experimentwise error rate held in check. The confidence interval tells whether the means differ; it also gives a range of values within which the mean differences lies. This tells us the precision with which we have captured the mean difference and can be used in with which we have captured the mean difference and can be used in judging the practical significance of

a result. That is, the confidence interval approach is more "informative." This procedure assumes equal group sizes and no problem with the above situation.

(3) Discriminant analysis

Procedure:
In discriminate analysis, we compute "discriminant scores" for each case to predict what group it is in. These scores are obtained by fining linear combinations of the independent variables; a linear combination is formed by multiplying each variable by some constant, and then adding up the products. Particularly, the SPSS subprogram DISCRIMINANT does all the computations and quite useful. Examining the printout, we 1) use the discriminant function/variable correlations to define the dimensions, 2) plot the group centroids in the space defined by the discriminant functions, and 3) determine which pairs of groups differ in distance between their centroids through the use of the pairwise F tests resulting from the discriminant analysis.

Rationale:
In this case, the purpose of this procedure is to discover the underlying dimensions, which differentiate among the four groups of students. Because the discriminate functions are uncorrelated, they yield and additive partitioning of the between association. About 20 subjects per variable are needed for reliable results, and this case has it.

TM032470

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

Reproduction Release

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Educational Research and Statistics: Examples of Questions and Answers

Author(s): Yukiko Inoue

Corporate Source: University of Guam

Publication Date: 2001

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced i ERIC system, Resources in Education (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, an Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following noti

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign in t

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker s
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY ___ SAMPLE ___ TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY ___ SAMPLE ___ TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)	PERMISSI DISSEMIN MICROFICHE O ___ SAMPLE ___ TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
Level 1	Level 2A	Level 2B
X		
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g. electronic) *and* paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche, or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Signature: Yukiko Inoue

Printed Name/Position/Title: Yukiko Inoue, PhD/Assist. Professor, College of Education

Organization/Address:
College of Education
University of Guam
UOG Station, Mangilao, Guam 96923

Telephone: 1-671-735-2447

Fax: 1-671-734-4914

E-mail Address: yinoue@uog.edu

Date: 2/6/2001

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
4483-A Forbes Boulevard
Lanham, Maryland 20706
Telephone: 301-552-4200
Toll Free: 800-799-3742
e-mail: ericfac@inet.ed.gov
WWW: http://ericfac.piccard.csc.com

EFF-088 (Rev. 9/97)