

Department of Sociology  
American University  
Comprehensive Examination in Social Research Methods  
August 2005

This exam has four parts. You must answer one question in each part for a total of four questions. Your answers should: (1) demonstrate the full range and depth of your understanding of the relevant research techniques; (2) refer to appropriate academic studies or research traditions; (3) minimize the degree of repetition between responses; (4) number your answers to correspond to the questions and their different parts. Be sure to budget your time. Please double-space your answers or write on every other line of your blue book. Be sure that your exam number and the page number appear on each page of your answers. Do not write your name anywhere on the exam.

Part I. Answer one question.

1. Imagine that you are part of a team charged with designing a multi-method way to study the causes of and the impacts on communities of increasingly long commutes to work in the United States.
  - a. Divide your contributions to this design into quantitative and qualitative research.
  - b. Explain why both types of methodologies would be necessary to such a research project.
  - c. Name and describe one data-gathering strategy from each that you would consider best for such a project.
  - d. Explain why you consider them appropriate—that is, what information will they give you and how do they relate to each other.
  - e. Describe and justify the sequence in which you would do them.
2. Unobtrusive measures are a standard methodological tool.
  - a. What are unobtrusive measures?
  - b. Give two examples of techniques that are entirely unobtrusive.
  - c. What are the advantages and disadvantages of unobtrusive measures compared to (1) participant observation, (2) interviews, and (3) questionnaires?

Part II. Answer one question.

1. Zaller's RAS (Receive-Accept-Sample) model describes how survey respondents display their attitudes.
  - a. Define each of the three parts of the model.
  - b. Explain how this model accounts for question order effects.
  - c. Explain why key informants do not suffer from the same problem as survey respondents (two hints: van Haitsma and "off the top of their head"). Describe the characteristics that a researcher should look for in choosing a good key informant.
2. Mainstream social science research (e.g., Positivist) and critical social science research (e.g., Feminist) bring to the research project differing concerns about the ethical implications of the research activity for the people being studied.
  - a. Describe these two configurations of ethical concern.
  - b. What steps would each orientation to research suggest for mitigating harm and

exploitation of the persons whose lives form the basis of the research?

Part III. Answer one question.

1. Imagine that you are being asked to teach a unit on the "Logic of Inquiry" in a beginning methods course; that is, you are being asked to explain how a student would move from an area of general interest to a plan for gathering and analyzing data. Create an example that students could follow; that is, select a specific general interest area and outline (numbering as you go) the steps necessary to arrive at researchable questions and appropriate data for answering those questions in that area.
2. Discuss the concepts of a population, a sample, and a sampling frame in answering the following:
  - a. How are these 3 different? Use an example to illustrate the differences.
  - b. Discuss the cautions that should be applied when generalizing from the sampling frame to the sample and to the population.
  - c. Explain how probably sampling techniques provide estimates of population parameters and estimates of error. Note that you do not need to supply formulas in this explanation, but you do have to explain what the formulas estimate and why it is important.

Part IV. Answer one question. Only students who answer the regression question are eligible for a Ph.D. pass; students who answer the tables question are eligible for only an M.A. pass.

1. A regional sociological association would like to know why more sociologists are not members. As part of this research it conducts a study to find the characteristics of sociologists who are more likely to be members. It wishes to distinguish between sociologists employed teaching in higher education versus other kinds of employment. Figures 1-3 report some results. There are 3 variables:  
 TEACH\$: Does the sociologist work at a teaching institution? Yes or No  
 SIZE\$: Size of institution where employed. Small, Medium, or Large  
 MEMBER\$: Is the sociologist an association member? Yes or No

Figure 1

```
>xtab
>freq=count
>print none / freq rowpct chi sq gamma lambda
>tab size$ * member$
```

Case frequencies determined by value of variable COUNT.

Frequencies  
 SIZE\$ (rows) by MEMBER\$ (columns)

No	Yes	Total
Large	60 61	121
Medium	64 97	161
Small	22 33	55
Total	146 191	337

```

Row percents
SIZE$ (rows) by MEMBER$ (columns)
No Yes Total N
+-----+
Large | 49.587 50.413 | 100.000 121
Medium | 39.752 60.248 | 100.000 161
Small | 40.000 60.000 | 100.000 55
+-----+
Total 43.323 56.677 100.000
N 146 191 337
    
```

```

Test statistic Value df Prob
Pearson Chi-square 3.017 2.000 0.221
Coefficient Value Asymptotic Std Error
Goodman-Kruskal Gamma 0.146 0.093
Lambda (column dependent) 0.000 0.000
    
```

```

Figure 2
>select teach$ = 'No'
>tab size$ * member$
    
```

Data for the following results were selected according to:  
teach\$ = 'No'

Case frequencies determined by value of variable COUNT.

```

Frequencies
SIZE$ (rows) by MEMBER$ (columns)
No Yes Total
+-----+
Large | 10 27 | 37
Medium | 15 28 | 43
Small | 5 5 | 10
+-----+
Total 30 60 90
    
```

```

Row percents
SIZE$ (rows) by MEMBER$ (columns)
No Yes Total N
+-----+
Large | 27.027 72.973 | 100.000 37
Medium | 34.884 65.116 | 100.000 43
Small | 50.000 50.000 | 100.000 10
+-----+
Total 33.333 66.667 100.000
    
```

```

N 30 60 90
Test statistic Value df Prob
Pearson Chi-square 1.959 2.000 0.376
Coefficient Value Asymptotic Std Error
Goodman-Kruskal Gamma -0.253 0.188
Lambda (column dependent) 0.000 0.105
    
```

```

Figure 3
>select teach$ = 'Yes'
>tab size$ * member$
Data for the following results were selected according to:
teach$ = 'Yes'
    
```

Case frequencies determined by value of variable COUNT.

```

Frequencies
SIZE$ (rows) by MEMBER$ (columns)
    
```

```
No Yes Total
+-----+
Large | 50 34 | 84
Medium | 49 69 | 118
Small | 17 28 | 45
+-----+
Total 116 131 247
```

```
Row percents
SIZE$ (rows) by MEMBER$ (columns)
No Yes Total N
+-----+
Large | 59.524 40.476 | 100.000 84
Medium | 41.525 58.475 | 100.000 118
Small | 37.778 62.222 | 100.000 45
+-----+
Total 46.964 53.036 100.000
```

```
N 116 131 247
Test statistic Value df Prob
Pearson Chi-square 8.246 2.000 0.016
Coefficient Value Asymptotic Std Error
Goodman-Kruskal Gamma 0.291 0.102
Lambda (column dependent) 0.138 0.073
```

2. Below are the results of a state-level study of the effects of job growth on the poverty rate.

Note the use of regional dummy variables, South is the omitted category. Since the District of

Columbia is not included, there are 50 cases. The variables are:

POOR: Poverty rate, %

JOBGROW: Job growth, %

INCOME: Median annual household income, 1,000s of dollars

NEAST: dummy variable where 1 = Northeast, 0 = other.

SOUTH: dummy variable where 1 = South, 0 = other

NCENT: dummy variable where 1 = North central, 0 = other

WEST: dummy variable where 1 = West, 0 = other

Answer the following questions about the regression results and the plots:

a. Explain what you learned from the scatterplot matrix (SPLOM) in Figure 4.

b. Discuss the coefficients and p-values in Figure 6. What do you conclude?

c. The box plot in figure 5 shows that region is probably not a significant predictor of

POOR, but in Figure 6 the coefficient for WEST is significant. Explain this apparent contradiction.

d. What additional test involving dummy variables do you need in order to draw firm conclusions from these results? Describe the test and the information it would give you.

Figure 4

```
>splom jobgrow income poor / smooth lowess half
JOBGROW
INCOME
JOBGROW
POOR
INCOME POOR
```

Figure 5

```
>box poor * region$
NCent NEast South West
REGION
5
```

10  
15  
20  
POOR

Figure 6

>model poor = constant + jobgrow + income + ncent + neast + west  
>est

Dep Var: POOR N: 50 Multiple R: 0.835 Squared multiple R: 0.697  
Adjusted squared multiple R: 0.662 Standard error of estimate: 0.995  
Effect Coefficient Std Error Std Coef Tolerance t P(2 Tail)

CONSTANT	26.801	1.983	0.000	.13.519	0.000
JOBGROW	-0.887	0.288	-0.429	0.357	-3.084 0.004
INCOME	-0.013	0.003	-0.496	0.462	-4.064 0.000
NCENT	-0.077	0.416	-0.019	0.626	-0.184 0.855
NEAST	-0.249	0.515	-0.063	0.409	-0.484 0.631
WEST	0.279	0.111	0.070	0.602	2.658 0.024

Analysis of Variance

Source	Sum-of-Squares	df	Mean-Square	F-ratio	P
Regression	100.042	5	20.008	20.224	0.000
Residual	43.531	44	0.989		

---

-----  
-----  
Durbin-Watson D Statistic 1.254  
First Order Autocorrelation 0.361  
<Additional output omitted>