RESEARCH DESIGN AND STATISTICS FOR APPLIED LINGUISTICS

Evelyn Hatch and Hossein Farhady
University of California, Los Angeles
# CONTENTS

PREFACE ................................................................. iii
PRETEST ................................................................. ix
GLOSSARY OF SYMBOLS .............................................. xiii
INTRODUCTION ......................................................... xv

**CHAPTER**

1  *What Is Research?* ................................................. 1  
   Asking Appropriate Questions 1  
   What Is a "Systematic Way"? 4  
   Internal Validity 7  
   External Validity 8  
   Activities 10

2  *What Is a Variable?* .............................................. 12  
   Variable Scales 12  
   Nominal Scale 13  
   Ordinal Scale 13  
   Interval Scale 14  
   The Functions of Variables 14  
   Independent Variable 15  
   Dependent Variable 15  
   Moderator Variable 15  
   Control Variable 16  
   Intervening Variable 16  
   Activities 17

3  *Constructing Research Designs* ................................... 18  
   Pre-experimental Designs 19  
   True experimental Designs 22  
   Quasi-experimental Designs 22  
   Ex Post Facto Designs 26  
   Factorial Designs 28  
   Activities 31

4  *The Research Report Format* ..................................... 33

5  *Sorting and Displaying the Data* ................................. 39  
   Coding the Data 39  
   Nominal Data 40  
   Ordinal Data 40  
   Interval Data 41
CHAPTER

Simple Numerical Computations 42
  ratio, proportion, percent, relative frequency,
  rate, percent change, cumulative frequency, percentile
Displaying the Data 47
  tables, histograms and polygons
Activities 51

6 Describing the Data ........................................... 53
  Mode 54
  Median 54
  Mean 55
  Measures of Variability 56
    Range 57
    Standard Deviation 57
    Variance 60
Activities 61

7 Standardized Scores .......................................... 63
Activities 68
Using Computers to Do the Work 69

8 Probability and Hypothesis Testing ......................... 80
Activities 90
Working with the Computer 91

9 Distribution of Means and Differences between Means .... 95
Case I Studies 95
Sampling Distribution of Means 96
Case II Studies 102
Sampling Distribution of Differences
  between Means 103
Activities 107

10 t-tests—Comparing Two Means .............................. 108
Case I Studies 108
Case II Studies 111
Assumptions Underlying t-tests 114
Matched t-test 114
Activities 120
Working with the Computer 122

11 Comparing Means—ANOVA .................................... 128
One-way ANOVA 129
A priori and Post hoc Comparisons of Means 140
  Planned or a priori Comparisons 140
  Post hoc Comparisons 143
  Scheffé Test 143
## CHAPTER

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>Factorial Designs and ANOVA</td>
<td>151-158</td>
</tr>
<tr>
<td></td>
<td>Working with the Computer</td>
<td>160-161</td>
</tr>
<tr>
<td>13</td>
<td>Statistical Tests of Nominal Data—Chi-Square</td>
<td>165-171</td>
</tr>
<tr>
<td></td>
<td>Correction Factors</td>
<td>171-172</td>
</tr>
<tr>
<td>14</td>
<td>Implicational Scaling</td>
<td>174-178</td>
</tr>
<tr>
<td></td>
<td>Scaling</td>
<td>176-181</td>
</tr>
<tr>
<td></td>
<td>Problems with Implicational Scaling</td>
<td>181-187</td>
</tr>
<tr>
<td></td>
<td>Working with the Computer</td>
<td>188-189</td>
</tr>
<tr>
<td>15</td>
<td>Correlational Analyses</td>
<td>192-212</td>
</tr>
<tr>
<td></td>
<td>Computing the Correlation Coefficient</td>
<td>196-197</td>
</tr>
<tr>
<td></td>
<td>Computing $r$ from Raw Scores</td>
<td>197-198</td>
</tr>
<tr>
<td></td>
<td>Interpreting the Correlation Coefficient</td>
<td>201-203</td>
</tr>
<tr>
<td></td>
<td>Other Correlation Coefficients</td>
<td>203-207</td>
</tr>
<tr>
<td></td>
<td>Point Biserial Correlation</td>
<td>204-205</td>
</tr>
<tr>
<td></td>
<td>Spearman Rank Order Correlation</td>
<td>205-206</td>
</tr>
<tr>
<td></td>
<td>Factors Affecting Correlation Coefficients</td>
<td>208-210</td>
</tr>
<tr>
<td></td>
<td>Working with the Computer</td>
<td>210-212</td>
</tr>
<tr>
<td>16</td>
<td>Simple Linear Regression</td>
<td>215-225</td>
</tr>
<tr>
<td></td>
<td>Predicting $X$ from $Y$</td>
<td>225-226</td>
</tr>
<tr>
<td></td>
<td>The Relationship between ANOVA and Regression</td>
<td>225-226</td>
</tr>
<tr>
<td></td>
<td>Working with the Computer</td>
<td>226-227</td>
</tr>
<tr>
<td></td>
<td>Simple Linear Regression</td>
<td>227-228</td>
</tr>
<tr>
<td></td>
<td>Multiple Regression</td>
<td>228-236</td>
</tr>
<tr>
<td>17</td>
<td>Reliability and Validity</td>
<td>243-252</td>
</tr>
<tr>
<td></td>
<td>Reliability</td>
<td>243-246</td>
</tr>
<tr>
<td></td>
<td>Ways to Estimate Reliability</td>
<td>246-249</td>
</tr>
<tr>
<td></td>
<td>test-retest, parallel tests, internal consistency methods, Spearman-Brown prophecy formula, Kuder-Richardson 20, Kuder-Richardson 21</td>
<td>249-250</td>
</tr>
<tr>
<td></td>
<td>Reliability of Criterion-Referenced Tests</td>
<td>249-250</td>
</tr>
<tr>
<td></td>
<td>Factors Affecting Reliability</td>
<td>250-252</td>
</tr>
<tr>
<td></td>
<td>Validity</td>
<td>250-251</td>
</tr>
<tr>
<td></td>
<td>Content Validity</td>
<td>251-251</td>
</tr>
<tr>
<td></td>
<td>Criterion-Related Validity</td>
<td>251-252</td>
</tr>
<tr>
<td></td>
<td>Construct Validity</td>
<td>252-252</td>
</tr>
</tbody>
</table>