



## Index

- 68–95–99 rule, 137
- African American homeownership, 114, 121–123
- aggregation, 38, 80, 81
- Allport, Gordon, 57
- alternative hypotheses, 190–193
- American National Election Study (ANES), 93
- American Sociological Association (ASA) (*Code of Ethics*), 20–23
- analysis
- alongside writing, 255
  - and variable types, 118–119
- analysis of variance (ANOVA), 148, 168–171
- ANES (American National Election Study) *see* American National Election Study (ANES)
- anonymity, 23–24
- ANOVA (analysis of variance) *see* analysis of variance (ANOVA)
- assumptions, of two-variable regression, 194–198
- autocorrelation, 196
- bar graphs, 119
- bell curve *see* normal distribution
- Berelson, Bernard, 7
- Berk, Richard A., 73
- Between Women* (Rollins), 22
- bias, 217
- assumption of zero bias, 194–195
  - measurement bias, 106–107
  - omitted-variables bias, 217–220, 240
- binomial logit (BNL), 235–238
- binomial probit (BNP), 235–238
- bivariate hypothesis testing
- analysis of variance, 148, 168–171
  - and causal relationships, 147–148
  - choosing appropriate test, 148–149
  - correlation coefficients, 149, 162–168
  - difference of means test, 148, 159–162
  - exercises, 173–176
  - regression models *see* bivariate regression
  - tabular analysis, 148, 152–158
- bivariate theories, 47
- bivariate regression *see* two-variable regression
- BNL (binomial logit), 235–238
- BNP (binomial probit), 235–238
- Bobo, Lawrence, 254
- boundaries *see* parameters
- box-whisker plots, 122–123, 159
- categorical variables, 116, 118
- and bivariate hypothesis testing, 148–149, 159, 168
  - describing, 119
  - dummy independent variables, 221–228
- CATI (computer-assisted telephone interviewing), 89
- causal approach, 3
- causal claims
- evaluating, 52–56
  - presenting, 253–254
- causal explanations, 9, 39

280 **Index**

- causal hurdles, 49–52
  - in experimental research, 70
  - in observational research, 78–79
  - in tabular analysis, 158, 202–205
- causal hurdles scorecard, 52, 72
- causal mechanisms
  - identifying, 50, 54, 58–60, 71, 78, 158, 202
  - stated in theory section, 258
- causal relationships, 8–16
  - and bivariate hypothesis testing, 147–148
  - criteria for establishing *see* causal hurdles
  - exercises, 62–64
  - and experimental research, 67–72
  - and observational research, 78–79
  - sociological debates about, 57–61
  - and variables, 8–16, 47
- causal theories, 4, 17, 39
- causality, notions of, 47–49
- census, 135
- central limit theorem, 135–142
- central tendency, 119, 121, 123–127
- charts and graphs, 115
  - bar graphs, 119
  - box-whisker plots, 122–123, 159
  - histograms, 125–127
  - kernel density plots, 126–127
  - pie graphs, 119
  - scatter plots, 162–163, 164, 165, 178–180
  - see also* tables
- chi-squared test, 157–158
- citations, 251–252
- clarity, 255
- classification tables, 239
- cluster sampling, 96–97
- code, documenting, 255–256
- coding rules, 106
- coefficient *see* parameter estimates
- comparison, 65–66
  - presented in regression tables, 262–264
- competitive threat, 254
- computer-assisted telephone interviewing (CATI), 89
- conceptual clarity, 104–105
- conclusions section, 259
- confidence intervals, 140–142
  - in regression models, 189–190, 192
  - and sample size, 143–144
- confidentiality, 23–24
- confounding (Z) variables (control of), 51, 55–241
  - in experiments, 69–72
  - failure to control, 56–57, 216–220
  - in multiple regression *see* multiple regression
  - in observational studies, 79–82
  - and regression model assumptions, 197
  - in tabular analysis, 158, 203–205
- consent, 21–23
- construct validity, 108
- contact hypothesis, 57–58
- content validity, 107–108
- context of experiments, 76–77
- context of research, 250–251
- continuous variables, 117–118
  - and bivariate hypothesis testing, 148–149, 159, 162, 168
  - describing, 119–127
- control groups, random assignment to, 68–72
- Converse, Philip E., 250
- Copernicus, Nicolaus, 6
- correlation, 17
  - autocorrelation, 196
  - see also* covariation/correlation
- correlation coefficients, 149, 162–168
- correlational studies, 78
- covariance, 163–168
- covariance tables, 166
- covariation, 17, 162–163
- covariation/correlation, causality and
  - presence of, 51, 54, 55, 58, 59, 61, 71, 78–79, 158, 203
- critical values, 157–158, 161–162
- cross-sectional measures, 32–33
- cross-sectional observational studies, 79–80, 265–266
- Currie, Janet, 61
- data
  - exercises, 129–130
  - populations and samples, 131–133
  - and theory development, 17–18, 32, 40
- data section of project, 258
- data set dimensions, 80
- data sets, 79, 92–94, 101
- datum, 79
- deception, 22–23

- degrees of freedom, 158, 161, 167, 189
- dependent variables
- in causal relationships, 8–15
  - displayed in tables, 152, 154
  - dummy variables, 232–240
  - in formulation of research projects, 246–251
  - identifying variation in, 30
- depression, measurement of, 103
- descriptive graphs, 115, 119, 122–123, 125–127
- descriptive statistics, 115
- for categorical variables, 119
  - for continuous variables, 119–127
  - limitations of, 127–128
  - rank statistics, 121–123
  - statistical moments, 123–127
- designs *see* research designs
- deterministic relationships, 48–49
- dice rolling, 134
- difference of means test, 148, 159–162
- directional hypothesis, 192
- dispersion, 122
- distribution
- assumptions in two-variable regression, 194
  - normal distribution, 136–141, 194
  - representations of *see* descriptive statistics
  - sampling distribution, 139
- diversionary use of force, 38
- Duflo, Esther, 73
- dummy dependent variables, 232–240
- dummy independent variables, 221–229
- dummy-variable trap, 223
- dummy variables in interactive models, 229–232
- economic accountability, 35
- economic performance, 8–16
- economic voting, 232–240, 251
- economics, measurement in, 102–103
- education
- gender, income and, 221–232, 264–265
  - private school choice programs, 53–56
- educational attainment
- and race, 159–162
  - and support for government assistance, 211–213
- empathy, and gender, 154–158, 202–205
- empirical evidence, 18
- EPSEM samples, 95
- equal unit differences, 117, 118
- Erikson, Robert S., 253
- error
- assumed absence in measurement, 196–197
  - proportionate reduction of, 239–240
  - root MSE, 185
  - sample and population error terms, 180
  - sampling error, 97
  - see also* standard error
- ethics, 19–24
- ethnicity *see* race/ethnicity
- evidence
- empirical evidence, 18
  - scientific approach to, 5
- expected value, 124
- experiments
- definition and features, 66–72
  - drawbacks, 74–77
  - random assignment and random sampling, 72–73
  - varieties, 73–74, 90
- external validity, 75–77, 91
- face-to-face interviews, 87–88
- face validity, 107
- factual approach, 2–3
- Feldman, Stanley, 110
- field experiments, 73–74
- figures
- use in research projects, 260, 265–266
  - see also* graphs
- financial satisfaction, 116, 117
- findings section, 258
- focus groups, 67
- frequency data, presentation of, 119
- frequency distribution, 137
- frequency tables, 119, 265
- functional form, as problem of LPM, 235
- Gaudet, Hazel, 7
- gender
- education, income and, 221–232, 264–265
  - empathy and, 154–158, 202–205
  - science education, race and, 91–92

282 **Index**

- General Social Survey (GSS), 92  
 generality, 19, 33–36, 40  
 generalization from samples *see* statistical inference  
 geographic context of research, 251  
 Gini coefficients, 32–33  
 goodness-of-fit  
   assessing, 187  
   dummy dependent variables, 238–240  
   R-squared statistic, 185–187  
   root mean-squared error, 185  
 government assistance, unemployment and support for  
   as correlation coefficient example, 162–168  
   as multiple regression example, 210–214, 263  
   as two-variable regression example *see* two-variable regression  
 graphs *see* charts and graphs  
 group interviews, 88  
 GSS (General Social Survey) *see* General Social Survey (GSS)
- Hanson, Sandra L., 91–92  
 harm (potential), 20–21, 23  
 Head Start program, 60–61  
 heteroscedasticity, 195, 234–235  
 histograms, 125–127  
 homoscedasticity, 195  
 Huddy, Leonie, 110  
 human rights violations, 20  
 Humphreys, Laud, 23  
 Hutchings, Victor, 254  
 hypotheses, 4  
   and theory, 10  
 hypothesis testing, 5–6  
   *see also* bivariate hypothesis testing
- immigrants, attitudes toward, 8–16  
 implications section of project, 259  
 income  
   gender, education and, 221–232, 264–265  
   measuring, 104–105  
 income inequality, 32–33  
 independent outcomes, 134  
 independent variables  
   in causal relationships, 8–15  
   controlled in experiments, 68, 74–75  
   displayed in tables, 152, 154  
   dummy variables in interactive models, 229–232  
   dummy variables in OLS, 221–229  
   in formulation of original research projects, 246–251  
 indicator variables, 221  
 individual-level dynamics, 38  
 informed consent, 21–23  
 Institutional Review Boards (IRBs), 20, 24  
 interactive models, with dummy variables, 229–232  
 intergroup contact, and racial tolerance, 57–58  
 internal validity, 72, 75, 91  
 international conflict, 34–35, 38  
 International Social Survey Programme (ISSP), 93–94  
 interquartile range (IQR), 122, 123  
 interval variables *see* continuous variables  
 interviews  
   face-to-face, 87–88  
   in groups, 88  
   by telephone, 88–89  
 introduction section, 256–257  
 IQR (interquartile range) *see* interquartile range (IQR)  
 ISSP (International Social Survey Programme) *see* International Social Survey Programme (ISSP)
- Jennings, Kent, 8
- Kellstedt, Paul M., 265  
 kernel density plots, 126–127  
 Kinder, Donald, 248  
 Knowledge Networks, Inc. (KN), 89–90  
 Kuhn, Thomas, 6  
 kurtosis, 125
- Lazarsfeld, Paul, 7  
 least-squares property, 124  
 levels of aggregation, 38  
 Levendusky, Matthew, 251  
 Lewis-Beck, Michael, 35–36  
 linear probability model (LPM), 232–235, 237–240  
 lines *see* regression lines  
 link functions, 236

## Index

283

- literature, examining previous research, 36–38, 248–249, 251–254  
 literature review section, 257  
 logit model, 148, 235–238
- MacKuen, Michael B., 253  
 macro level of aggregation, 38  
 marijuana legalization example, 66–67, 69–70, 76–77  
 mean value, 123–124  
   in ANOVA, 168–170  
   in central limit theorem, 138–142  
   difference of means test, 159–162  
   in multiple comparison procedures, 171–172
- measurement  
   assumed absence of error in, 196–197  
   conceptual clarity, 104–105  
   consequences of poor, 111  
   exercises, 112–113  
   operationalization example, 11–15  
   and *p*-values, 151  
   problematic in social sciences, 101–104  
   of racial tolerance, 109–111  
   reliability, 105–109  
   spatial/cross-sectional, 31–33  
   and theory, 111  
   time dimension, 31–32  
   validity, 107–109
- measurement bias, 106–107  
 measurement metric, 115–119, 213–214  
 median value, 121  
 micro level of aggregation, 38  
 Milgram, Stanley, 21  
 mode, 119  
 model sum of squares (MSS), 186  
 models, 16  
 moment statistics, 123–127  
 Mueller, John, 34–35, 37  
 multicollinearity, 220  
 multiple comparison procedures, 171–172
- multiple regression  
   building on two-variable method, 205–209  
   comparison of effects, 213–214  
   constructing regression tables, 260–264  
   dummy dependent variables, 232–240  
   dummy independent variables, 221–229  
   exercises, 242–245  
   failure to control for *Z*, 216–220  
   interactive models with dummy variables, 229–232  
   interpretation example, 210–213  
   mathematical requirements, 220–221  
   population regression model, 205  
   statistical and substantive significance, 214–216  
   writing about regression tables, 264–265
- multistage cluster sampling, 96–97  
 multivariate analysis, 201–202  
   and causality, 60  
   multiple regression *see* multiple regression  
   tabular analysis, 202–205
- multivariate theories, 47
- Nagler, Jonathan, 256  
 natural experiments, 74  
 Nazi medical experiments, 20  
 negative relationship, 15, 164  
 new/original theory, 40, 246–251  
 Niemi, Richard, 8  
 nominal variables *see* categorical variables  
 non-obvious theory, 41  
 normal distribution, 136–141, 194  
 normal science, 6  
 normative statements, 18–19  
 null hypotheses, 4, 6, 152, 190–193
- obedience study (Milgram), 21  
 observational data, 79–80  
 observational studies  
   cross-sectional, 80  
   definition and features, 77–79  
   drawbacks, 81–82  
   time-series, 80–81
- OLS *see* ordinary least-squares (OLS) regression  
 omitted-variables bias, 217–220, 240  
 one-tailed hypothesis tests, 192–193  
 operationalization *see* measurement  
 ordinal variables, 116–118  
 ordinary least-squares (OLS) regression, 182  
   dummy dependent variables in, 232–235  
   dummy independent variables in, 221–229

284 **Index**

- ordinary least-squares (OLS) regression  
 (cont.)  
 measures of uncertainty in, 184–193  
 original/new theory, 40, 246–251  
 outcomes, 134  
 outliers, 120, 122–123
- p*-values, 149–152, 158, 161–162  
 panels, in survey research, 90  
 paradigm shifts, 7–8  
 paradigms, 6  
 parameter estimates, 179, 182, 187–193,  
 207, 216, 228–229, 262  
 parameters, 178  
 parametric linearity, 197, 235  
 parsimony, 19, 40  
 partisan sort, 251  
 party identification, 117, 153–154,  
 250–251  
 Pearson, Karl, 157, 160, 165  
 Pearson's *r*, 165–168  
*The People's Choice* (Lazarsfeld et al.), 7  
 perfect multicollinearity *see*  
 multicollinearity  
 Pettigrew, Thomas, 58  
 pie graphs, 119  
 placebos, 68  
 policing and race, 89–90  
 policy attitudes, 250–251  
 political behavior, 7–8  
*The Political Character of Adolescence*  
 (Niemi and Jennings), 8  
 political control, and economic  
 accountability, 35  
 political participation, and race, 58–60  
 political party identification, 117,  
 153–154, 250–251  
 population data, 132  
 population error term, 180  
 population regression model, 178, 179,  
 205  
 population sampling *see* sampling  
 positive relationship, 13, 164, 165  
 Posner, Daniel N., 74  
 potential harm, 20–21, 23  
 poverty, measurement of, 102–103, 111  
 Powell, G. Bingham Jr., 251  
 predicted probability, 234, 236–238  
 prejudice, measurement of, 107  
 presidential approval, 31–32, 34–35,  
 37–38, 141–142  
 presidential election, 232–240  
 primary data collection, 92  
 primary sampling units (PSUs), 96  
 principle–policy gap, 109–111  
 prison experiment (Zimbardo), 20–21  
 privacy *see* confidentiality  
 private school choice programs, 53–56  
 probabilistic relationships, 48–49  
 probability, 15  
 definitions and properties, 134–135  
 linear probability model (LPM),  
 232–235, 237–240  
*see also p*-values  
 probability sampling, 94–97  
 probability theory  
 central limit theorem, 135–142  
 example, 133–134  
 probit model, 148, 235–238  
 proofreading, 259  
 proportionate reduction of error,  
 239–240  
 proportionate stratified random  
 sampling, 96  
 PSUs (primary sampling units) *see*  
 primary sampling units (PSUs)  
 psychology, 103–104  
 public opinion  
 use of surveys, 86–87  
*see also* presidential approval
- quartile ranks, 122  
 questionnaires, self-administered, 88
- R*-squared statistic, 185–187, 212  
 race/ethnicity  
 and educational attainment, 159–162,  
 169–172  
 and policing, 89–90  
 and political participation, 58–60  
 and residential segregation, 2–3  
 science education, gender and, 91–92  
 as variable in multiple regression,  
 225–228  
 racial policies, 265  
 racial prejudice, 254  
 racial resentment, 110–111  
 racial tolerance, 57–58, 109–111  
 racism (symbolic), 248  
 random assignment, 68–73

## Index

285

- random digit dialing (RDD), 89, 90
- random samples/sampling, 132
  - compared with random assignment, 72–73
  - importance of, 143
  - and *p*-values, 151
  - simple, 95
  - stratified, 96
  - systematic, 95
  - unnecessary in experiments, 75
- rank statistics, 121–123
- reference category, 226–228
- regression *see* multiple regression; two-variable regression
- regression lines
  - estimation of, 180–183
  - fitting through scatter plots, 178–180
  - in interactive models, 230–232
  - measures of uncertainty, 184–193
  - in multiple regression, 224–225
- regression tables
  - constructing, 260–264
  - writing about, 264–265
- reliability, 105–109
- religiosity, 203
- replication, 76
- representativeness, 89
- research
  - examining previous research, 36–38, 248–249, 251–254
  - implications of, 37–38
  - rules of social research, 16–19
  - see also* literature review section
- research design section, 258
- research designs, 66
  - exercises, 84–85
  - see also* experiments; observational studies; survey research
- research ethics, 19–24
- research projects
  - effective writing, 255–260
  - examining previous research, 248–249, 251–254
  - exercises, 266
  - routes to formulating, 246–251
  - using tables and figures, 260–266
- research questions, 30, 39
- residual sum of squares (RSS), 186
- residual values, 180–182, 188, 207
- results section, 258
- reverse causality, ruling out, 50, 54, 58–60, 71, 78, 158, 202, 258
- Rollins, Judith, 22
- root mean-squared error (root MSE), 185
- RSS (residual sum of squares), 186
- sample of convenience, 76, 143
- sample data, 132–133
- sample error term, 180
- sample regression models, 179
- sample size, 150, 161, 188, 189, 215
- samples
  - generalizing from *see* statistical inference
  - importance of random samples, 143
  - and *p*-values, 151
- sampling
  - and experiments, 72–73, 76
  - multistage cluster sampling, 96–97
  - probability and non-probability, 94, 95
  - simple random sampling, 95
  - stratified random sampling, 96
  - systematic random sampling, 95
- sampling distribution, 139
- sampling error, 97
- satisfaction with financial situation, 116, 117
- scatter plots, 162–165
  - estimation of regression line, 180–183
  - fitting a line, 178–180
- Scheffe test, 171
- school choice programs, 53–56
- science education, race, gender and, 91–92
- scientific knowledge, 5–8
  - and original research projects, 246–251
- scientific revolutions, 6–7
- Sears, David, 248
- secondary data, 92–94, 101
- self-administered questionnaires, 88
- Sherif, Muzafer, 57
- Sherman, Lawrence W., 73
- simple random sampling, 95
- simplicity *see* parsimony
- skepticism, 5, 36, 257
- skewness, 125
- spatial dimension, 31
- spatial units, 31
- spell-checking, 259
- spurious relationships, 51, 55, 59, 61, 71, 82, 158, 203

286 **Index**

- standard deviation, 124–125, 136, 138–140, 142, 213–214
- standard error, 139–144, 160–161, 188–189, 235  
*see also* root mean-squared error
- standardized coefficients, 213–214
- statistical analyses  
 alongside writing, 255  
 and variable types, 118–119
- statistical description *see* descriptive statistics
- statistical inference, 132  
 basics of probability, 133–135  
 central limit theorem, 135–142  
 exercises, 145–146  
 importance of random samples, 143  
 with multiple variables *see* multiple regression  
 and sample size, 143–144  
 with two variables *see* bivariate hypothesis testing
- statistical models, 178
- statistical moments, 123–127
- statistical significance, 151–152, 214–216, 262
- statistical software, 255–256, 261
- Stimson, James A., 253
- stochastic component, 178–180, 188, 194–197
- stratified random sampling, 96
- substantive significance, 216
- survey data sets, 92–94
- survey research organizations, 86–87, 89
- survey-based experiments, 73, 90–92
- surveys  
 exercises, 98  
 face-to-face interviews, 87–88  
 pre-existing data sets, 92–94  
 and probability sampling, 94–97  
 self-administered questionnaires, 88  
 telephone interviews, 88–89  
 ubiquity of, 86–87  
 web-based, 89–90
- symbolic racism, 248
- syntax files, 255–256
- systematic random sampling, 95
- t*-ratio, 191–193
- t*-test, 160–162, 189
- tables, 260–266  
*see also* frequency tables; regression tables
- tabular analysis, 148, 152–158, 202–205
- Tearoom Trade* (Humphreys), 23
- technology, and interviews, 89
- telephone interviews, 88–89
- theoretical models, 16
- theories  
 alternative applications, 37  
 causal, 4, 17, 39  
 definition, 4  
 elements of good theory, 29, 30, 38–41  
 exercises, 42–45  
 and hypotheses, 10–16  
 and measurement, 99–100, 111  
 new/original, 40, 246–251
- theory development  
 evaluating good theory, 38–41  
 examining previous research, 36–38  
 and generality, 33–36  
 identifying variation, 30–33  
 place of data in, 17–18
- theory of economic threat, 10–15
- theory section of project, 258
- Thomas, Duncan, 61
- time context of research, 250
- time dimension, 31
- time-series measures, 31–32, 196
- time-series observational studies, 79, 265–266
- time units, 31, 80
- tolerance (racial), 57–58, 109–111
- total sum of squares (TSS), 186
- trade union affiliation, 153–154
- treatment groups, random assignment to, 68–72
- Tuch, Steven A., 89
- Tuskegee syphilis study, 20
- two-tailed hypothesis tests, 190–192
- two-variable regression  
 assumptions, 194–198  
 as basis of multiple regression, 205–209  
 estimation of regression line, 180–183  
 exercises, 199–200  
 fitting a line, 178–180  
 mathematical requirements, 198  
 measures of uncertainty, 184–193



## Index

287

- uncertainty
  - measured in OLS regression, 184–193
  - see also* standard error
- unemployment and attitudes toward immigrants, 13–15
- unemployment and support for government assistance
  - correlation coefficient example, 162–168
  - as multiple regression example, 210–214, 263
  - as two-variable regression example *see* two-variable regression
- units of analysis, in observational research, 79–80
- unstandardized coefficients, 213–214
  
- validity, 107–109
- variable labels, 8, 115
- variable types
  - and bivariate hypothesis tests, 148–149
  - categorical, 116, 118–119, 148–149, 159, 168, 221–228
  - comparing variables in different metrics, 213–214
  - continuous, 117–127, 148–149, 159, 162, 168
  - exercises, 129–130
  - measurement metrics, 115–119
  - ordinal, 116–118
  - and statistical analyses, 118–119
  - see also* confounding (Z) variables; dependent variables; independent variables
- variable values, 8, 115
  - describing *see* descriptive statistics
- variables
  - and causal relationships, 8–16, 47
  - definition, 3
- variance, 124, 187–189, 195
- variation, 118–119
  - cross-sectional, 32–33, 79–80, 265–266
  - describing *see* descriptive statistics
  - identifying, 30–33
  - R-squared statistic, 185–187
  - as requirement of linear regression models, 198
  - time-series, 31–32, 79, 80–81, 196, 265–266
- Verba, Sidney, 59
  
- web-based surveys, 89–90
- Weitzer, Ronald, 89
- Whitten, Guy D., 251
- World Values Survey (WVS), 94
- writing strategies
  - documenting code, 255–256
  - examining previous research, 253–254
  - proofreading, 259
  - sections, 256–259
  - thinking and clarification, 255
- WVS (World Values Survey) *see* World Values Survey (WVS)
  
- Z variables *see* confounding (Z) variables
- zero bias assumption, 194–195
- zero-sum property, 124
- Zimbardo, Philip, 20–21