

# INDEX

## A

actual measurement frequencies,  
130, 131  
alpha value ( $\alpha$ ), 159, 163  
alternative hypothesis  
accuracy of, 166  
considerations, 174  
Cramer's coefficient, 186  
examples of, 161, 171–173  
overview, 170–174  
P-value and, 175–179  
test of difference between popula-  
tion ratios, 173  
arithmetic mean, 43, 73, 74  
average (mean). *See* mean  
AVERAGE function, 196  
average savings, 46–47

## C

calculations. *See* Excel calculations  
categorical data, 14–29  
correlation ratio, 121  
creating tables, 60–64  
cylinder charts, 114  
defined, 19  
examples of, 20, 23–26  
indexes, 117  
overview, 14–19  
as result of survey, 60–64  
scatter charts, 114  
charts  
converting to graphs, 33–39  
correlation ratio, 126  
cylinder, 114  
degree of relation and, 115  
expenditure, 116–120  
scatter. *See* scatter charts  
CHIDIST function, 107  
CHIINV function, 107, 205–206  
chi-square distribution, 99–105  
calculating, 130–133  
degrees of freedom, 99–108  
described, 99  
examples of, 99–105, 152  
points on horizontal axis, 205–206  
chi-square symbol, 103  
chi-square test of independence,  
151–169

CHITEST function, 210–211  
class midpoint, 36–39, 54, 56  
classes  
calculating with Sturges' Rule, 55,  
56, 58  
intraclass variance, 117, 123,  
124, 126  
range of, 39, 54–57, 84  
coefficient  
correlation, 116–120, 206–207  
Cramer's. *See* Cramer's coefficient  
CORREL function, 207  
correlation, 115, 119  
correlation coefficient, 116–120,  
206–207  
correlation ratio, 117, 121–127, 207  
COUNTIF function, 197–198  
Cramer's coefficient, 127–138  
accuracy of, 147  
alternative hypothesis, 186  
calculating, 130–135, 141  
examples of, 127–136  
Excel and, 207  
indexes, 117, 129  
informal standard, 136  
making informed decision about,  
147–148  
null hypothesis, 168, 186  
ratio of preference, 155  
variances in population, 145–150,  
157, 186  
Cramer's V. *See* Cramer's coefficient  
critical region, 159, 165–167, 187  
cross tabulation, 62–64, 128, 130,  
135, 151, 153, 197–198  
curve, grading on. *See* standard scores  
cylinder charts, 114

## D

data  
categorical. *See* categorical data  
collection of, 186  
immeasurable. *See* categorical data  
numerical. *See* numerical data  
"scattering of," 49, 58, 69, 70, 80  
unsuitable for correlation  
coefficient, 120  
data point, 80

data types, 13–29, 117  
degree of relation, 115, 116–120  
degrees of freedom (df), 99–108  
descriptive statistics, 57–58  
deviation, standard, 48–53, 70–79  
deviation scores, 74–80, 199–203  
df (degrees of freedom), 99–108  
distributions  
chi-square. *See* chi-square  
distribution  
Excel and, 107–109  
F, 106–107  
normal, 86–91  
standard normal, 89–98, 204–205  
t, 106

## E

estimation theory, 57–58  
Euler's number, 86  
Excel calculations, 191–211  
chi-square distribution, 205–206  
correlation coefficient, 206–207  
cross tabulation, 197–198  
deviation scores, 74–80, 199–203  
distributions and, 107–109  
frequency tables, 192–195  
mean, 195–196  
median, 195–196  
standard deviation, 195–196  
standard normal distribution,  
204–205  
standard scores, 199–202  
tests of independence, 208–211  
Excel files, downloading, 192  
Excel functions  
AVERAGE, 196  
CHIDIST, 107  
CHIINV, 107, 205–206  
CHITEST, 210–211  
CORREL, 207  
COUNTIF, 197–198  
FDIST, 107  
FINV, 107  
FREQUENCY, 193–194  
NORMDIST, 107  
NORMINV, 107  
NORMSDIST, 107, 204  
NORMSINV, 107

Excel functions, *continued*

STANDARDIZE, 199–201

TDIST, 107

expected frequencies, 130, 131

expenditure chart, 116–120

## F

F distribution, 106–107

FDIST function, 107

FINV function, 107

freedom, degrees of, 99–108

frequency

actual, 130, 131

described, 36

distribution tables, 32–39

expected, 130, 131

relative, 36–37, 39

FREQUENCY function, 193–194

frequency tables

creating with Excel, 192–195

range of class of, 54–56

## G

geometric mean, 43

grading on a curve. *See* standard scores

graphs

converting price charts to, 33–39

converting surveys to, 62–64

shape of, 100–101

slope of, 101

## H

harmonic mean, 43

histograms

advantages of, 83

examples of, 39, 83, 84, 154

overview of, 38–39

probability density function, 83–84

range of class and, 84, 85

variables, 39

homogeneity, test of, 184–186

horizontal axis, 39, 102, 107, 109, 125

calculating points on, 107

hypothesis tests, 143–189. *See also*

tests of independence

alternative hypothesis. *See* alternative hypothesis

chi-square test of independence,

151–169

conclusions, 187

defined, 149

examples of, 149, 168–174

null hypothesis. *See* null hypothesis

overview of, 144–150

population considerations, 149, 186

procedure for, 150, 175–179

P-value, 163, 175–179, 189

tests of correlation, 149, 171

tests of correlation ratio, 149, 171

tests of difference between population means, 149, 171, 173

tests of difference between population ratios, 149, 171, 173

tests of homogeneity, 184–186

tests of independence, 149, 171

types of, 149, 171

## I

immeasurable data. *See* categorical data

independent coefficient. *See* Cramer's coefficient

indexes

correlation coefficient, 120

Cramer's coefficient, 117, 129

numerical data, 117

intraclass variance, 117, 123, 124, 126

## L

linear relationships, 120

## M

mean (average)

arithmetic, 43, 73, 74

calculating with Excel, 195–196

defined, 43

examples, 40–44

geometric, 43

harmonic, 43

normal distribution and, 87–89

standard normal distribution and, 89–90

median

calculating with Excel, 195–196

defined, 45

examples of, 45–47

uses for, 44

Microsoft Excel. *See* Excel calculations, Excel files, *and* Excel functions

midpoint, class, 36–39, 54, 56

multiple-choice answers, 28

## N

Napier's constant, 86

negative correlation, 119

non-linear relationships, 120

normal distribution, 86–91

normalization, 71–72

NORMDIST function, 107

NORMINV function, 107

NORMSDIST function, 107, 204

NORMSINV function, 107

null hypothesis

considerations, 174

Cramer's coefficient, 168, 186

difficulty of proving, 174

examples of, 167–174

failing to reject, 150, 167, 178, 179, 187

overview, 170–174

P-value and, 175–179

rejecting, 158, 159, 178

for tests of correlation, 172

for tests of correlation ratio, 172

for tests of difference between

population ratios, 173

for tests of independence, 172

numerical data, 14–29

correlation ratio, 121

defined, 19

descriptive statistics, 57–58

estimation theory, 57–58

examples of, 21–23, 26

frequency tables, 32–39, 54–56, 58

histograms, 38–39, 54, 58

indexes, 117

mean (average), 40–43

median, 44–47

overview, 31–58

scatter charts, 114

standard deviation, 48–53, 70–79

## P

Pearson's chi-square test statistic, 132, 152–155, 158

percentage, 5, 37, 62, 64

population

Cramer's coefficient, 145–150,

157, 186

defined, 6

hypothesis tests and, 149, 186

vs. sample, 52

standard deviation, 52

- status of, 4, 7, 57
- variances in, 145–150, 157, 186
- population ratios, 149, 171, 173
- positive correlation, 119
- price charts, 33–39
- probability, 81–109
  - associated, 104
  - chi-square distribution, 99–105, 205–206
  - defined, 82
  - degrees of freedom (df), 99–108
  - distributions and Excel, 107–109
  - F distribution, 106–107
  - normal distribution, 86–89
  - standard normal distribution, 89–98, 204–205
  - t distribution, 106
  - test results, 83–84
- probability density function, 82–85, 99, 107, 109
- P-value
  - alternative hypothesis and, 175–179
  - hypothesis tests, 163, 175–179, 189
  - null hypothesis and, 175–179
  - tests of independence, 175

## Q

- qualitative data. *See* categorical data
- quantitative data. *See* numerical data
- questionnaires, 15–19

## R

- range, class, 39, 54–57, 84
- relationships
  - correlation ratio, 117, 121–127
  - degree of, 115, 116–120
  - linear, 120
  - non-linear, 120
  - variables, 112–115
- relative frequency, 36–37, 39

## S

- samples, 6, 7, 52, 57
- scatter charts
  - correlation ratio, 122, 126
  - examples of, 114, 116
  - monthly expenditures, 116–120
- scattering, of data, 49, 58, 69, 70, 80
- scores
  - deviation, 74–80
  - evaluating, 71
  - standard, 65–80, 73, 199–202

- significance level ( $\alpha$ ), 159, 163
- slope, graph, 101
- standard deviation
  - calculating with Excel, 195–196
  - normal distribution and, 87–91
  - numerical data, 48–53, 70–79
  - population, 52
  - standard normal distribution and, 89–90
- standard normal distribution, 89–98, 204–205
- standard scores, 65–80, 73, 199–202
- standardization, 71–72, 80
- STANDARDIZE function, 199–201
- statistical hypothesis testing. *See* hypothesis tests
- statistical significance, 187
- statistics
  - defined, 4
  - descriptive, 57–58
  - estimation theory, 4–7
- STEP test, 23–25
- Sturges' Rule, 55, 56, 58
- surveys, 4–7
  - categorical data, 60–64
  - converting to graphs, 62–64
  - limitations of, 4–7
  - tests of independence, 137, 208–211

## T

- t distribution, 106
- tables
  - categorical data, 60–64
  - chi-square distribution, 102–105, 205–206
  - cross tabulation, 128, 130, 135, 151, 153
  - frequency. *See* frequency tables
  - normal distribution and, 107
  - standard normal distribution, 92–93, 104, 108
- TDIST function, 107
- test results
  - normal distribution, 86–89
  - probability density function, 83–84
  - standard normal distribution, 89–98
- tests of correlation, 149, 171, 172
- tests of correlation ratio, 149, 171, 172
- tests of difference between population means, 149, 171, 173

- tests of difference between population ratios, 149, 171, 173
- tests of homogeneity, 184–186
- tests of independence, 208–211. *See also* hypothesis tests
  - chi-square, 151–169
  - examples of, 149, 171, 184–186
  - P-value, 175
  - vs. tests of homogeneity, 186
  - uses for, 137, 149
- TINV function, 107

## V

- values
  - median, 44–47
  - P-value, 163, 175–179, 189
- variables, 111–142
  - correlation coefficient, 116–120
  - correlation ratio, 121–127
  - Cramer's coefficient, 127–138, 141, 142
  - degree of relation, 115, 116–120
  - histograms, 39
  - relationships, 112–115
- vertical axis, 39

## W

- weather forecasts, 82

## Z

- zero correlation, 119
- z-score. *See* standard scores