## INDEX

## A

Achievement Check, 13, 40, 52, 83, $112,124,149,155,183,201,215$, 231, 240, 246, 259, 261, 281, 288, 295, 299, 319, 326, 342, 355, 361, $376,386,405,409,440,450,458$, 473
Accidental relationship, 196-197
Action plan, 488-490
Addition rule for mutually exclusive events, 337-339
Addition rule for non-mutually exclusive events, 339
Additive counting principle, 228, 266-267, 337-338
Algorithm, 6
Arrow diagrams, 8
Average, 125-128, 131-133

## B

Bar graph, 92-93, 100
Bell curve, 418, 423-427, 435-436, 459
Bernoulli trials, 378-379
Bias, 119-122
intentional, 122, 202, 205, 208-209
leading question, 121-122
loaded question, 121-122
measurement, 120-122
non-response, 120, 122
response, 122
sampling, 119-122
statistical, 119
unintentional, 122, 208-209
Bimodal distribution, 415
Binomial distribution, 378-384, 445-448
normal approximation, 445-448
Binomial expansion, 289, 291-293
Binomial theorem, 289-293
factoring with, 293
Box-and-whisker plots, 141, 143-144

## C

CANSIM, 30
Career Connection
Actuary, 387
Cryptographer, 86
Economist, 211
Forensic Scientist, 272
Investment Broker, 356
Statistician, 150
Cause and effect, 195-199
reverse, 196-197, 199
Census, 30
Central tendency
measures of, 126-133, 501-502
Chapter Problem
Collecting Cards, 367, 376, 396
Contract Negotiations, 89, 110, 149
Genetic Probabilities, 301, 313, 318, 326, 335, 341, 354
Job Prospects, 157, 169, 181, 193, 200, 210
Radio Programming, 263, 272, 280, 288
Restless Earth, 411, 421, 440, 466, 469
Students' Council Elections, 223, 230, 240, 246
Via Rail Routes, 3, 77
Circle graph, 98-100
Circuit, 43, 48
Class, 94
Cluster sample, 116
Codes, 70-71
Coefficient of determination, 184-185, 190-191
Combinations, 273-278, 282-286, 322
all possible, 282-283, 290, 322-324
numbers of, 275-276
vs permutations, 274-275
Combinatorics, 225
Common-cause factor, 196-197, 199
Complement of an event, 309
Complete network, 43-45, 48
Compound events, 327
Conditional probability, 331-333, 344

Confidence intervals, 459-463
population proportion, 461
Confidence level, 314, 455-456, 459-463
Connected network, 42-45, 48
Consumer price index (CPI), 104-109
Continuity correction, 436-438
Continuous distributions, 414-418
Control group, 197-199
Convenience sample, 116
Correlation
coefficient, 161-167, 174-175, 177-179, 203, 206
direct, 160
inverse, 160
linear correlation, 159-167, 171-179, 203-204
moderate, 163, 166
negative, 160-161, 163
non-linear, 184-191
positive, 160-162, 173, 177-178, 204, 206
zero, 161, 177
Correlation coefficient, 161-167, 174-175, 177-179, 203, 206
Cost of living index, 107
Counting outcomes, 225-228, 235-236
additive counting principle, 228, 266-267, 337-338
counting subsets, 228
fundamental counting principle, 227-228, 235-237, 278, 284
indirect method, 227-228, 235-236
multiplicative counting principle, 227-228, 235-237, 278, 284
mutually exclusive actions, 228
ordered arrangements, 232-233
rule of sum, 228, 266-267, 337-338
tree diagram, 226, 507
Counting paths or routes, 254-256
Covariance, 161-162
Cubic regression, 184

Culminating Project, 481-495
action plan, 488-490
choice of , 482-487
critique, 494-495
evaluation, 491-492
hypothesis, 486
mind map, 482-483
presentation, 493-495, 598-605
research, 594-597
Cumulative frequency, 92-94, 100
Cumulative probability, 427
Curve of best fit, 184-185, 187-191

## D

Data
categorical, 98, 100
class, 94,100
dispersion, 136-147
grouped, 94, 132
interval, 94, 100
midrange, 150
percentiles, 145
quartiles, 144
range, 94
raw, 91
spread, 136-147
variance, 137
Data analysis, 91-100, 104-108
box-and-whiskers plot, 141, 143-144
graphs, 91-100
Databases, 28-31
field, 29
record, 29
relational, 14
software, 14-15
Data groups, 140
interquartile ranges, 140-144
percentiles, 145, 428-429
quartiles, 140-144
semi-interquartile range, 141-143
Dependent events, 327, 331-333, 398
product rule, 332
Deviation, 172
Dimensions of matrix, 53
Direct correlation, 160

Distributions, 369-374
binomial, 380-384
continuous, 414-418
discrete, 372-373
equations, 415-418
expectation, 401-402
exponential, 416-418
geometric, 388-394, 416
uniform, 372-374, 415-416
Discrete variable, 91, 100, 370-371, 374
Disjoint events, 337
Dispersion, 136

## E

Edge, 41-46, 48
Elements of set, 266
Entry in matrix, 53
Equations, solving, 506-507
Evaluating expressions, 496
Evaluating laws, 496
Events, 305
dependent, 327, 331-333, 398
independent, 327-331, 333
mutually exclusive, 336-340, 345
non-mutually exclusive, 338-340
Expectation, 372, 374, 378, 380-384, 392
general formula, 402
Experimental group, 197-198
Exponential regression, 185-186, 189-191
Extraneous variables, 197-199, 205-206, 208
Extrapolation, 171, 174, 190-191

## F

Factorials, 232-234, 237
calculating, 234
evaluating factorial expressions, 235
Fair game, 373-374
Fibonacci sequence, 253
Field, 29
First-step probability vector
Fractal, 8
Fractions, percents, decimals, 497
Frequency, 91-100
cumulative, 92-94, 100
relative, 97-100

Frequency diagrams, 91-94, 100
histograms, 92, 94, 97-98
Pareto chart, 112
polygons, 92-94, 97-98
Frequency tables, 91-94, 100
Fundamental counting principle, 227-228, 235-237, 278, 284

## G

Geometric distributions, 388-394
expectation, 393
Graphs
bar, 91-93, 100
circle, 98-100
histograms, 92, 94, 97-98
pictograph, 98-100
pie, 98-100
polygons, 92-94, 97-98
time-series, 105-109, 207-208
Graphing
data, 498-494
exponential functions, 499-500
quadratic functions, 500-501
Graph theory, 41-48
adjacent vertices, 43, 48
circuit, 43, 48
complete, 43-45, 48
connected, 42-45, 48
edge, 41-46, 48
map colouring, 41-42, 46-48
neighbour, 41
nodes, 41-46, 48
path, 43-45, 48
planar, 45-48
traceable, 44-45, 48
vertex, 41-46, 48
Grouped data, 94-98
mean and median, 131-132
standard deviation, 140

## H

Hidden variable, 206-209
Histogram, 92, 94, 97
relative frequency, 98
Hypergeometric distributions, 397-403
expectation, 401-402
Hypothesis test, 453-456

Identity matrix, 67
Independent events, 327-331, 333
product rule for, 329-330
Independent variable, 160, 177, 197-198
Index, 104-109
base level, 104
consumer price (CPI), 104-109
cost of living, 109
economic, 106-107, 109
readability, 108
S\&P/TSE 300
Composite, 105-106
Indirect method, 227
Inflation, 106-107
Initial probability vector, 345
Initial value, 9
Interpolation, 171, 174, 190
Interquartile range, 140-143
Intersection of sets, 267
Interval, 94, 100, 131, 133
relative frequency of, 98
Iteration, 6-10, 247, 254-255
Inverse matrix, 67-69, 74

## L

Least-squares method, 172-174, 179
Leading question, 121-122
Linear correlation, 159-167, 171-179, 203-204
classification, 160-165
coefficient, 161-167, 174-175, 177-179, 203, 206
direct, 160
inverse, 160
moderate, 163, 166
negative, 160-161, 163
perfect, 160, 163
positive, 160-162, 173, 177-178, 204, 206
zero, 161, 177
Linear regression, 171-179, 203-208
Linear relationships, 171-179
Line of best fit, 160, 171-179, 184, 203-205
Loaded question, 121-122
Lurking variable, 206-209

## M

Margin of error, 462-463
Markov chains, 344-345, 350-352
Matrices, 53-59, 63-74, 344-352
addition and subtraction, 56, 59
coding, 70-71
column, 54
column sum, 53
dimensions, 53
entry, 53
graphing calculator, 58, 65-66
identity, 67
inverse, 67-69, 74
scalar multiplication, 56, 59
Markov chains, 344-345, 350-352
multiplication, 63-67, 74
network, 71-74
representation of data, 54-55
row, 54
row sum, 53
spreadsheet, 58-59, 66-67
square, 54
transition, 345-351
transpose, 53, 59
Mean, 126-128, 131-133, 501-502
Mean absolute deviation, 150
Measurement bias, 120-122
Measures of central tendency
mean, 126-128, 131-133, 501-502
median, 126-128, 131-133, 501-502
midrange, 150
mode, 126-127, 133, 501-502
population vs sample, 127, 128, 136
weighted mean, 130-131, 133
Measures of spread or dispersion, 136-147
interquartile ranges, 140-143
mean absolute deviation, 150
percentiles, 145, 428-429
quartiles, 140-144
semi-interquartile range, 141-142
standard deviation, 137-140
variance, 137, 139-140
z-scores, 146-147, 426, 433, 453
Median, 126-128, 131-133, 501-502
Member of set, 266
Midrange, 150
Mind map, 482-483

Modal interval, 135
Mode, 126-127, 133, 501-502
Multiplicative counting principle, 227-228, 235-237, 278, 284
Multi-stage sample, 116
Mutually exclusive events, 336-340, 345
addition rule, 337-338
disjoint, 337
non-overlapping, 337

## N

Negative skew, 415
Networks, 41-48
adjacent vertices, 43, 48
circuit, 43, 48
complete, 43-45, 48
connected, 42-45, 48
edge, 41-46, 48
map colouring, 41-42, 46-48
neighbour, 41
path, 43-45, 48
planar, 45-48
traceable, 44-45, 48
vertex, 41-46, 48
Network matrix, 71-74
Node, 41-46, 48
Non-linear regression, 184-191
Non-mutually exclusive events, 338-340
Non-response bias, 120, 122
Normal distribution, 418, 422-429, 432-438, 459
discrete data, 436-438
equations, 425-429
plotting, 422-423
standard, 426
table, 427, 433-434, 436-437, 446-447, 606-607
Normal probability plots, 442-444
Null hypothesis, 454-456
Number patterns, 502

Odds, 314-317, 330-331
Ogive, 92

Oral presentations, 598-604
audience, 598-599
delivery, 603-605
planning, 598-601
rehearsing, 494, 602-603
visuals, 602
Order of operations, 503
Organized counting
additive counting principle, 228,
266-267, 337-338
counting subsets, 228
fundamental counting principle, 227-228, 235-237, 278, 284
indirect method, 227-228, 235-236
multiplicative counting principle, 227-228, 235-237, 278, 284
mutually exclusive actions, 228
ordered arrangements, 232-233
rule of sum, 228, 266-267, 337-338
tree diagram, 226, 507
Outcomes, 305
Outliers, 126-128, 133, 141,
177-179, 204, 208

## P

Pareto chart, 112
Pascal's formula, 290-291
Pascal's method, 254-256, 290
counting paths or routes, 254-256
Pascal's triangle, 247-251, 254-255,
289-293
divisibility, 249
perfect squares, 250
row sums, 248-249, 251
triangular numbers, 249-251
Path, 43-45, 48
Pearson's correlation coefficient, 162, 185
Percentiles, 145, 428-429
Perfect linear correlations, 160, 163
Permutations, 236-238, 241-244,
263, 321, 324
vs combinations, 274-275
formula, 237-238
with identical items, 241-244
Pictograph, 98-100
Pie graph, 98-100
Planar network, 45-48
Plotting data, 92-99, 159-167

Polls, opinion, 459-463
Polynomial regression, 189-191
Population, 113-116, 133, 137-138, 203
compared with sample, 128
identifying, 113
proportion, 461, 463
strata, 115
Positive linear correlatrion, 160-162, 173, 177-178, 204, 206
Positive skew, 415
Power regression, 187-189
Presentations, 493-494, 598-605
audience, 598-599
delivery, 603-605
planning, 598-601
rehearsing, 494, 602-603
visuals, 602
Presumed relationship, 196-197
Principle of inclusion and exclusion, 267-270
Probability, 304-311, 369-374, 378-384
classical (a priori), 306, 308, 311, 314, 321-324, 338-339
compound, 332-333
conditional, 331-333, 344
counting techniques, 320-324, 329-330, 332
cumulative, 427
density, 415
empirical (experimental), 306, 308, 311, 314, 320-321
independent events, 329-331
matrices, 344-352
multiple events, 321-324
mutually exclusive events, 326-340
product rule, 330-331
from relative frequency, 306,308 , 311, 314, 320-321
subjective, 306, 310-311, 314
successive events, 321-324
theoretical, 306, 308, 311, 314, 321-324, 338-339
tree diagrams, 307, 320, 328, 344
vectors, 344-352
Probability distributions, 369-374
binomial, 380-384
continuous, 414-418
discrete, 372-373
equations, 415-418
expectation, 401-402
exponential, 416-418
geometric, 388-394, 416
simulations, 370, 374, 379, 388, 397-398
uniform, 372-374, 415-416
Probability experiment, 305, 311, 320
events, 305-307, 311
outcomes, 305-307, 324
sample space, 305
simulation, 369
trials, 305
Probability vectors, 345-352
initial, 345-346, 348-350
steady-state, 349-352
Product rule
dependent events, 332-333
independent events, 329-330, 380, 389
Project Prep, 20, 29, 47, 66, 97, 107, 108, 122, 129, 140, 167, 176, 179, 190, 199, 208, 227, 237, 244, 277, 307, 315, 323, 332, 339, 350, 370, 392, 423, 454, 459

Quadratic regression, 184, 188-191
Quantiles, 443-444
Quartiles, 140-144

## R

Random variable, 370-371, 374, 389, 394
Range, 94
Ranking data, 6
Ratios of areas, 503
Raw data, 91
Record, 29
Recursive formula, 9
Regression
exponential, 187-187
linear, 171-179
polynomial, 189-190
power, 187-189
quadratic, 184, 188-191
Relational database, 14
Relative frequency, 97-100, 418
diagrams, 97-98

Repeated sampling, 451-453
simulating, 451
Research skills, 594-597
accessing sources, 595
citing sources, 597
evaluating sources, 596
Residuals, 172, 179
Response bias, 122
Reverse cause-and-effect, 196-197, 199
Row matrix, 54
Row sum, 53
Rule of sum, 228, 266-267, 337-338

## S

Sample, 113-116, 133, 137-138
bias, 119-122
central value, 137
cluster, 116
convenience, 116
extrapolating from, 113
frame, 114
interval, 114-115
multi-stage, 116
vs population, 128-130
simple random, 114
size, 203-205, 209, 462-463
stratified, 115-116
systematic, 114
techniques, 203-204, 209
voluntary-response, 116
Sample space, 305
Scalar, 56
Scatter plots, 159-161, 164-167, 172-173, 175, 177, 188-190, 203, 206, 414, 504-505
Seed value, 9
Self-similar shape, 8
Semi-interquartile range, 141-2
Sets, 265-270
common elements, 266
elements, 266
intersection, 267
members, 266
null set, 283
principle of inclusion and
exclusion, 267-270
subset, 266, 283
union, 267

Sierpinski triangle, 7-8
Sigma notation, 126, 505
Significance level, 456
Similar triangles, 505-506
Simplifying expressions, 506
Simulation, 33-40, 320, 369, 378, 388, 397, 416-418, 451
Skewed distributions, 415
Solving equations, 506-507
Spread, measures of, 136-147
Square matrix, 54
Standard deviation, 137-140
Standard normal distribution, 426
Statistical fluctuation, 308
Statistics, 202
Statistics Canada, 28-30, 104, 106
Steady state, 349-352
Stratified sample, 115-116
Subset, 266, 283
Substituting into equations, 507
Survey
bias, 119-122
control group, 197-199
experimental group, 197-198
Systematic sample, 114

## T

Tetrahedral numbers, 252
Time-series graphs, 105-109, 207-208
Traceable network, 44-45, 48
Transition matrix, 345
Tree diagrams, 226, 507
Trials, 305, 378-379
Triangular numbers, 249-251

## U

Uniform distribution, 372-374, 415-416
Unimodal distribution, 415
Union of sets, 267
Universal set, 265-269

## V

Variables, 91, 100, 104, 197
continuous, 91-92, 100, 370-371, 374
correlation between, 161
dependent, 160, 197
discrete, 91, 100, 370-371, 374
explanatory, 160
extraneous, 197-199, 204-206, 208-209
hidden, 206-209
independent, 160, 177, 197-198
lurking, 206-209
random, 370-371, 374, 389, 394
relationships between, 159,
171-179, 184
response, 160
Variance, 137, 139-140
Venn diagrams, 265-270, 337, 339
Vertex, 41-46, 48
adjacent, 43, 48
degree, 43-45, 47-48
Voluntary-response sample, 116

## W

Waiting time (waiting period),
388-389, 392-394, 416-418
Weighted mean, 130-131, 133

## Z

Z-scores, 146-147, 426, 433, 453

