Index

Bold page numbers indicate a major reference, usually a definition.

Acausal, 10, 487, 569
AIDS, 4
Analysis of variance, 543, 544
importance of the normal assumption, 562
multiple treatments, 550
relationship to regression analysis, 555
ANOVA. See Analysis of variance
Association, 123, 143, 148, 150, 488, 570
measures of, 135, 137, 143, 151, 163
Bayes theorem, D15
Bayesian interval, D15
Between sum of squares, 546
Bias. See Moments, bias in
Binomial probability distribution. See Distribution
Bivariate data, 120, 135
categorical data, 144
Bivariate distributions. See Distributions, bivariate
Bootstrap, C17
Bonferroni inequality, 553
Box-and-whisker plot, 45, 127, 430, 482, 543
BSS. See Between sum of squares
Cardinal measurement. See Measurement
categorical data, 23, 144
Category(ies), 151
Causal(ity), 3, 9, 10, 28, 130, 139, 149, 487, 488, 504, 539
CDF. (cumulative distribution function). See Distribution
Cell, 51, 121
Cell mark, 54, 82, 106
Central Limit Theorem, 288, 307, 310, 628
Challenger, 5
Chebyshev inequality, 299
Chi-square. See Distribution
Class mark. See Cell mark
Coefficient of association. See Association, measures of
Coefficient of determination, 500, 502
equal to square of correlation coefficient, 503
Combination, 219
Combinatorial coefficients. See Binomial coefficients
Complement of an event, 181
Composite probability. See Probability
Compound event, 177, 273
Conditional mean, 484, 487
Conditional normal. See Distribution, conditional
Conditional probability. See probability
Conditioning variable, or event, 487
Confidence interval, 352, 358, 360, 368, 374, 384, 388, 507, 510, 515
effect of asymmetric distributions on, 354, 388
Confidence level, 352
Consistency, 341, 342, 344, 358, 385
Contingency table, 144, 628
Continuous variable, 23, 199, 271, 278, 281, C3
Contour plots, 462
Convolution sum, 222
Correlation, 3, 135, 429
and multiple variables, 163
nonsense, 489
Correlation coefficient, 135, 138, 460, 482, 491, 570
and slope of regression line, 140
as standardized covariance, 132, 491
Spearman’s Rank. See Spearman’s rank correlation
zero for nonlinear association, 139, 488, 570
Estimate, 340
confidence interval. See Confidence interval
point, 345
Estimation, 339, 346, 394
binomial probability, 356
difference in means, 387
distribution free, C19
large sample results, 339
mean, 347, 363, 369
mean of the Poisson, 360
regression coefficients, 495, 496, 498, 523
variance, 372, 374
Estimator, 340, 571
bias in, 343
consistency of, 341, 342, 344
Exhaustive. See Sample event
Expectation, 236, 262, 263, 571
Expected value. See Expectation
Experiment, 28, 29, 58, 60, 172, 176, 220, 278, 281, 315, 331, 338, 575
Explanation, 8, 10
F. See Distribution
First cross-product moment. See Covariance; Moments
Forecasting. See Predictor
Frequency
absolute, 48, 82, 145, 174
cumulative, 49, 52, 83, 174
relative, 48, 53, 82, 145, 174
two-sided, 171
Functional relationship, 22, 27
Guassian. See Distribution
Global warming, 3
Goodness of fit, C8
Grouped data, 76, 83, 106, 115
higher moments, 106
Histogram, 55, 77, 121, 136, 171, 269, 270, 289, 310, 315
cell (mark), 51, 54
class mark. See Cell mark
two-dimensional, 121
HIV, 4
Hooke’s law, 21, 26
Hypothesis test(s), 396, 438, 444, 553
acceptance region, 398
alternate hypothesis, 396, 433, 441
composite, 412
critical bound, 395, 412
critical region, 395, 398, 412, 415, 420, 424
difference between two means, 421, 446, 563
null hypothesis, 393, 441
special role of, 409, 437
one-sided, 413
operationally significant, 426, 504
power, 401, 411, 413, 441
pretesting bias, 553
of proportions, 417
regression coefficients, 520
rejection region. See critical region
simple versus composite, 412
simple versus simple, 411, 437
states of the world, 395, 399
statistical significance, 425
two sided, 415, 424
Type I error, 397, 400, 408, 438, 442
familywise error rate, 553
why small values chosen, 407
Type II error, 397, 400, 408, 438, 442
using normal, one-sided, 410
using normal, two-sided, 417
using Student’s T, 423, 428
variances, equality of, 440
variance unknown, 418
Hypothesis(ses), 8, 396
comparing, D15
Ignorance, representing, D9, D12
Independence, 149, 242, 282, C10
not ‘mutually exclusive’, 207
statistical, 197
Independent variable. See Regressor
Induction, D3, D4
Inference, 171, 314, 338, 571
Interquartile range, 44
Intersection of events, 179
Invariance to transformations. See Moments
IPCC, 4
Joint events, 186
Joint probability, 183
Kurtosis. See Shape, peakedness

Least squares, 493, 495
Likelihood, 376, D6
Line chart, 49
Loss, expected loss function, 439, D14, D23

\( \mu_r \). See Moments, theoretical

\( m_i \). See Moments about the mean

\( m'_i \). See First cross product moment
Marginal probability. See Probability

Maximum likelihood estimation, 376, 377
Maximum likelihood estimator, 389
as modal estimator, 378
Maximum likelihood principle, 375, 376, 389
Mean, 78, 237
Mean squared error, 344, 345, 386
Measurement
cardinal, 23, 142
discrete, 24
ers, 455, 487
interval scale, 23
ratio scale, 23
categorical, 25, 144
cross-sectional, 26
index, 25
nominal, 25
ordinal, 24, 142
panel, 26
time series, 26
units of, 429, 490, 551
Median, 41, 78, 128, 482, C16, C17, C22
Meta-language, 9, 575
MLE. See Maximum likelihood estimation
Model coefficients, 226, 228, 456, 458
Modeling, 8
Moment-generating function, 242, 261, 296
Moments, 89, 135, 629
bias in, 343
first cross product, 126, 131, 132, 136,
473, 491
invariance to transformation, 95, 104, 115,
116
mean, 77, 237
method of, 375
nonexistence of, 298, 366, 518

\( \mu_r \). See Moments, theoretical

\( m_i \). See Moments about the mean

\( m'_i \). See First cross product moment
Marginal probability, See Probability

Maximum likelihood estimation, 376, 377
Maximum likelihood estimator, 389
as modal estimator, 378
Maximum likelihood principle, 375, 376, 389
Mean, 78, 237
Mean squared error, 344, 345, 386
Measurement
cardinal, 23, 142
discrete, 24
ers, 455, 487
interval scale, 23
ratio scale, 23
categorical, 25, 144
cross-sectional, 26
index, 25
nominal, 25
ordinal, 24, 142
panel, 26
time series, 26
units of, 429, 490, 551
Median, 41, 78, 128, 482, C16, C17, C22
Meta-language, 9, 575
MLE. See Maximum likelihood estimation
Model coefficients, 226, 228, 456, 458
Modeling, 8
Moment-generating function, 242, 261, 296
Moments, 89, 135, 629
bias in, 343
first cross product, 126, 131, 132, 136,
473, 491
invariance to transformation, 95, 104, 115,
116
mean, 77, 237
method of, 375
nonexistence of, 298, 366, 518
problem of, 299
sample, 322, 339, 571
about the mean, 89, 310
about the origin, 89
fourth, 93, 102, 310
second, 85, 100, 310, 342
standard deviation, 343, 389
third, 91, 99, 100, 310
variance, 343, 372, 389, 498
variance of a difference, 422, 431
variance, pooled estimate of, 423
standardized
third, fourth, 96, 234, 310, 385, 430
theoretical, 231, 235, 244, 277, 315,
340, 571
binomial, 232, 233
chi-square, 373
Laplace, 297
normal, 285, 385
Poisson, 250, 262
Student’s T, 365
uniform, 280, 305, 306
Monte Carlo sampling, C16
MSE. See Mean squared error
Multiple coefficient of determination.
See Coefficient of determination
Multiple regression, 503
Mutually exclusive, 173
implies dependence, 207

Nature’s experiments. See Experiment
Noise, 11, 162
Nonparametric, C19
Normal distribution. See Distribution, normal
Normal probability plots. See Q-Q plots
Nuisance parameters, C20
Null set. See Empty set

Obesity, 6
Operationally significant. See Hypothesis
test(s), operationally significant
Outliers, 45, 123, 128, 430
Order statistics, C3

P value, 426, 428
panel data. See Measurement
parameters. See model coefficients
INDEX 647

Random interval, 352, 369
Random variable, 10, 31, 38, 171, 191, 320, 455, 569
continuous, 171, 268, C9
Random(ness), 9, 332, 432, 569
Range, 43
interquartile, 45
Rank correlation. See Spearman’s rank correlation
Regressand, 485
Regression equation, 485, 490, 535
regression coefficient estimators and residuals uncorrelated, 536
Regression model, 486
estimators for coefficients of, 495, 496, 498, 523
Regression sum of squares, 502
“Regression to the mean” explained, 525
Regressor, 485, 487
Relationship functional, 27, 137, 139, 140, 500
linear statistical, 135
structural, 31, 126, 137, 140, 162, 484
Residual variance estimator, 499, 515
unbiased, 536
Response bias, 323
Residuals, 493
minimize sum of squares of. See Least squares
RSS. See Regression sum of squares

S². See Moments, sample, variance
Sample, 320
Sample distribution. See distribution
Sample experiment, 173
Sample event, 173, 177, 179, 273
Sample space, 173, 174
Sampling, 316, 320, 335, 488, 571
from a finite population, 323
proportionate stratified random, 325, 327, 328
simple random, 320, 324
without replacement, 35, 323
Sampling theory, 316
Scatter diagram. See Scatter plot
Scatter plot, 122
Science
methodology of, 7
statistics as a, 9, 171, 575

Pascal’s triangle, 227, 591
Permutations, 216
PDF. See Probability distribution function
Plim, See probability limit
Poisson. See Distribution
Polling. See Sampling
Population, 330
finite, 321, 333
hypothetical, 321, 331, 333
Posterior odds ratio, D16
Posterior probability distribution, D6, D9
Prediction. See Predictor
Predictor, 508
variance, 509, 510
Pretesting bias, 553
Principle of insufficient reason. See Equally likely principle
Prior probability distribution, D6
conjugate, D13
uninformed, D13
Probability, 175, 184, 185, 276, 418, 432, 443, 570
composite, 190
conditional, 185, 186, 190
joint, 183, 190
marginal, 194
as a ratio of counts, 207
Probability distribution, 176, 193, 198, 214, 221, 275
conditional, 195, 209, 453
joint, 194, 223, 453
marginal, 194, 223, 453
predictive, D22
Probability distribution function. See Probability distribution
Probability function. See Probability distribution
Probability limit, 341
Probability mass function. See Probability distribution
Probability theory, 170, 175
Probability tree, 191

Q-Q plots, C2, C5, C6, C7
Quantiles, C3
Quartiles, 43

r. See Correlation coefficient
R². See Coefficient of determination
Scientific method, 7
Serial independence, C10
Shape, 59, 75, 116, 122, 134, 170, 171, 228, 258, 259, 268, 277, 282, 288, 315, 323, 483, 570
location, 77, 277
mean, 77, 235, 252
median, 41, 78, 128, 131
peakedness, 62, 93
fourth moment, 92, 102, 235, 252, 285, 365, 373
skewness, 62, 90
third moment, 91, 101, 235, 252, 285, 373
spread, 61, 85, 277
interquartile range, 45
quartiles, 44
range, 43, 85, 128
second moment, 85, 99, 128, 235, 252
standard deviation, 87, 99, 236
Skewness. See Shape, skewness
Slope, 130, 132, 482, 493
Slope coefficient, 139,
and correlation, 139
Spearman’s rank correlation coefficient, 143, C13
Standard deviation, 87, 99, 236, 286, 343
Standard Gaussian distribution function. See Standard normal distribution function
Standard normal distribution function, 285
Standardization of variables, 104, 125, 126, 132, 282, 287, 288, 291, 490
Standardized moments,
third and fourth, 97, 280, 285
Statistic, 320, 340
Statistical tables, 291, 636
Statistically significant. See Hypothesis tests, statistical significance
Statistics
acausal nature of, 10
and science, 9, 171, 575
to theory of. See Probability theory
Stochastic. See Random(ness)
Stratum. See Sampling, proportionate stratified
Structural relationship. See relationship
Student’s T. See Distribution
Survey. See Sampling

Total sum of squares, 500, 502, 546
Transformation of variables, 237, 263, 308, 310, C3
Trial. See sample experiment
TSS. See Total sum of squares
Unconsidered alternatives, D10
Uniform distribution function. See
Distribution
Union of events, 177, 180
Univariate data, 122

Variable(s), 21
Variance. See Moments, sample, variance;
Moments, theoretical
Venn diagram, 182

Wald-Wolfowitz runs test, 629, C11
Wilcoxon signed rank test, C21
Within sum of squares, 546
WSS. See Within sum of squares

\bar{x}. See Moments, mean