

# Index

## A

- Analysis of variance
  - factors, 409
  - vs. regression with indicators, 395
  - treatments, 409
- Assignment 3-3 Latte, 63
- Assignment 13-1 Big Drug Co, 251
- Association between categorical variables, Excel, 353–354

## B

- “Bottom line” title, 7
- Bounded dependent variable, 237

## C

- Case 3-3 Polaski Package, 68
- Case 3-3 Polaski Taste, 67
- Case 7-2 Frosted Flakes, 377
- Categorical, 9
- Central tendency, 9
- CHIDIST(*chisquare, df*), 358
- Chi square distributions, 345
- Chi square ( $\chi^2$ ) statistic, 343
  - Excel, 355
- Chi square test of association, 343
- Conditional probabilities, 341–343
  - Excel, 353
- Confidence intervals, 35–68
  - for the difference between alternate scenarios or pairs, 53–56
  - for difference between segment means, 47–48, 57–60
  - for a population mean, 42–43
  - for a population proportion, 48–49
- Confidence level, 42
- Conjoint analysis, 379
  - orthogonal array, 379
  - part worth utilities, 380
    - Excel, 398
- Conservative confidence intervals, 49–50
- Contingency analysis, 341–378
  - hypotheses, 342
  - sparse cells, 344–346
  - with summary data, 358
- Correlation, 114
  - regression, 116, 117
- Critical Student *t*, 42
- Crosstabulation, 341
  - Excel, 353–354
- Cumulative distribution plot, 6

## D

- DC Hotels, 64
- Decide Whether Insignificant Drivers Matter, 263–264
- Describing Your Data, 5–33
- Descriptive statistics, 13
- Difference between alternate scenarios or pairs, 36–37
- Difference between segment means, 57
- Difference between two segments, 43–47
- Dispersion, 9
- Durbin Watson (*DW*) statistic, 108

## E

- Empirical Rule, 36, 40
- Excel 2.1 Executive Compensation, 13
- Excel 3.9 SD Portfolio, 53
- Excel 4.1 HitFlix Movie Rental Revenues, 119
- Excel 6.1 Portfolio3, 145
- Excel 7.1 American Cars, 353
- Excel 11.1 Executive Compensation, 209
- Excel 13.1 Pampers Concept Test, 243
- Excel shortcuts, 24–26

## F

- Forecasting, 293–340
  - Durbin Watson (*DW*) statistic, 108
  - Excel, 311–312
  - illustrate fit and forecast, Excel, 328
  - inertia, 296–297
  - leading indicators, 296–297
  - lengths of lags, 298–305
  - recalibration, 305
  - trend, 298
  - validation, 96, 296, 298

## H

- Histograms, 5–7, 13, 14, 16, 17
- Hypothesis tests, 35–68

## I

- Illustrate fit and forecast, Excel, 328
- Inference, 35
- Interquartile range, 6

## J

- JMP
  - for ANOVA, 393–395, 406, 411, 413
  - for regression, 95–133, 172, 209–213, 237, 238, 253–256, 266, 385
- Joint probability, 341–343

## L

- Lab 3 Dell PDA Inference, 63
- Lab 7 Skype Appeal, 365
- Lab Practice 3 Nationals, 57, 61
- Limited dependent variable, 237
  - Excel, 243–248
- Logit, 237
  - Excel, 243–248
  - odds, Excel, 243

## M

- Mean, 8
- Median, 8
- Memos, 190
- Mode, 9
- Model building process, cross sectional vs. time series, 98, 294, 295
- Multiple regression
  - bounded dependent variable, 237
  - conjoint analysis, 379
    - part worth utilities, 380
  - F* test, 256
  - goals, 253
  - identify drivers, 253
  - illustrate fit and forecast, Excel, 328
  - indicator, 155
  - joint impact of multiple drivers, 253
  - limited dependent variable, 237
    - Excel, 243–248
  - logit, 237
    - Excel, 243–248
    - odds, Excel, 243
  - marginal influence of drivers, 254
  - marginal slope
    - hypotheses, 258
    - t* statistic, 258
  - model building process, cross sectional vs. time series, 98, 294, 295
  - models, 253–291, 311–312
    - Excel, 266
    - hypotheses, 256
  - multicollinearity, 255–256
    - remedies, 259
    - symptoms, 258–259
  - nonlinear, 195
    - Excel, 209
    - gains vs. linear, 208
    - rescaling variables, 196
    - skew, 196
  - one tail test, 258
  - predict performance under alternative scenarios, 253
  - rescaling variables, synergies, 199
  - sensitivity analysis, 262–263
    - Excel, 270–274
  - vs. simple regression, 253
  - time series
    - vs. cross sectional models, 293

Durbin Watson (*DW*) statistic, 108

- leading indicators, 296–297
- lengths of lags, 298–305
- recalibration, 305
- trend, 298
- validation, 298
- variable selection
  - lengths of lags, 298–305
  - logic, 253–255

## N

- Nominal, 9
- Normally distributed, 35

## O

- Odds ratio, 237
- Ordinal, 9
- Orthogonal array, 379

## P

- Paired *t* test, 53
- Pareto* chart, 9
- Plot a cumulative distribution, 19–20
- Population proportion, confidence interval, 48–49
- Population standard deviation estimate, 40
- Portfolio analysis
  - beta, 135
  - correlation, 137
    - Excel, 145
  - regression estimate, 137
- Efficient Frontier, 141
  - Excel, 146–149
- expected rate of return, 135, 140
- prediction intervals
  - individual prediction intervals, 123, 125, 160, 161, 163, 172, 204, 207, 216–218, 220–222, 227–231, 235, 247, 249, 305, 309, 322, 323, 325–328, 332
  - rate of return, Excel, 145
  - risk, 140
- PowerPoints
  - colors, 190
  - design, 189
  - font, 189
  - notes, 188
  - presentations, 183
  - slide sorter, 187

## R

- Range, 5, 8
- Recalibration, 305
- Regression
  - correlation, 116, 117
  - equation, standard format, 113
  - Excel, 119
  - forecasting, 95
  - F* test, 105–108
  - F* test of model, 107

- intercept estimate, 96
- Mean Square Error, 104
- Model Sum of Squares, 105
- prediction intervals
  - individual prediction intervals, 123, 125, 160, 161, 163, 172, 204, 207, 216–218, 220–222, 227–231, 235, 247, 249, 305, 309, 322, 323, 325–328, 332
- quantifying the influence of a driver, 95
- residuals, 102
- Rsquare, 105–108
- simple linear, equation, 96
- slope
  - estimate, 96
  - hypotheses, 97, 100, 106
  - 95% confidence interval, 102
  - one-tail test, 100
  - standard error, 102–104
  - Sum of Squared Errors, 103
  - Sum of Squares, 105
- Round descriptive statistics, 7
- S
- Sample means, 39–42
- Scale, 8–9
- Simpson's Paradox, 346–351
- Skewness, 9
- Standard deviation (SD), 8
  - of difference, 53
  - of the sample proportions, 48
- Standard error
  - of difference between segment means, 46
  - estimate, 40
  - of sample means, 39
  - of sample proportion, 48
- Standardized sample means, 40
- Student *t*, 40
- Summary statistics, 5–7
- T
- Trend, 298
- t* test
  - of difference between alternate scenarios or pairs, 36–37
  - of difference between segment means, 47
  - of matched pairs, 37
  - of repeated samples, 37
- TTEST(array1,array2,tails,type), 53, 57
- V
- Validated, 96, 296
- Variance, 8