

Contents

Contents	3
Introduction	7
Information and Entropy	13
Entropy	14
Entropy of a Continuous Random Variable	18
Partitioning a Continuous Variable for Entropy	19
An Example of Improving Entropy	24
Joint and Conditional Entropy	26
Code for Conditional Entropy	31
Mutual Information	32
Fano's Bound and Selection of Predictor Variables	34
Confusion Matrices and Mutual Information	37
Extending Fano's Bound for Upper Limits	39
Simple Algorithms for Mutual Information	43
The TEST_DIS Program	48
Continuous Mutual Information	51
The Parzen Window Method	52
Adaptive Partitioning	61
The TEST_CON Program	76
Asymmetric Information Measures	77
Uncertainty Reduction	77
Transfer Entropy: Schreiber's Information Transfer	81
Screening for Relationships	91
Simple Screening Methods	93
Univariate Screening	93
Bivariate Screening	93
Forward Stepwise Selection	94
Forward Selection Preserving Subsets	94
Backward Stepwise Selection	95

Criteria for a Relationship	96
Ordinary Correlation	96
Nonparametric Correlation	98
Accommodating Simple Nonlinearity	101
Chi-Square and Cramer's V	104
Mutual Information and Uncertainty Reduction	107
Multivariate Extensions	108
Permutation tests	109
A Modestly Rigorous Statement of the Procedure	109
A More Intuitive Approach	112
Serial Correlation Can Be Deadly	114
Permutation Algorithms	115
Outline of the Permutation Test Algorithm	116
Permutation Testing for Selection Bias	117
Combinatorially Symmetric Cross Validation	119
The CSCV Algorithm	125
An Example of CSCV OOS testing	132
Univariate Screening for Relationships	133
Three Simple Examples	138
Bivariate Screening for Relationships	140
Stepwise Predictor Selection Using Mutual Information	148
Maximizing Relevance While Minimizing Redundancy	148
Code for the Relevance Minus Redundancy Algorithm	152
An Example of Relevance Minus Redundancy	156
A Superior Selection Algorithm for Binary Variables	160
FREL for High-Dimensionality, Small Size Datasets	166
Regularization	171
Interpreting Weights	172
Bootstrapping FREL	173
Monte-Carlo Permutation Tests of FREL	174
General Statement of the FREL algorithm	176
Multi-Threaded Code for FREL	181
Some FREL Examples	193

Displaying Relationship Anomalies	195
Marginal Density Product	200
Actual Density	200
Marginal Inconsistency	201
Mutual Information Contribution	202
Code for Computing These Plots	203
Comments on Showing the Display	214
Fun With Eigenvectors	215
Eigenvalues and Eigenvectors	218
Principal Components (if you really must)	220
The Factor Structure is More Interesting	221
A Simple Example	222
Rotation Can Make Naming Easier	224
Code for Eigenvectors and Rotation	226
Eigenvectors of a Real Symmetric Matrix	226
Factor Structure of a Dataset	228
Varimax Rotation	231
Horn's Algorithm for Determining Dimensionality	234
Code for the Modified Horn Algorithm	236
Clustering Variables in a Subspace	247
Code for Clustering Variables	251
Separating Individual From Common Variance	255
Log Likelihood the Slow, Definitional Way	263
Log Likelihood the Fast, Intelligent Way	265
The Basic Expectation Maximization Algorithm	268
Code for Basic Expectation Maximization	270
Accelerating the EM Algorithm	273
Code for Quadratic Acceleration with DECME-2s	278
Putting It All Together	283
Thoughts On My Version of the Algorithm	294
Measuring Coherence	295
Code for Tracking Coherence	298
Coherence in the Stock Market	301

Using the DATAMINE Program	303
File / Read Data File	304
File / Exit	304
Screen / Univariate Screen	305
Screen / Bivariate Screen	307
Screen / Relevance Minus Redundancy	309
Screen / FREL	311
Analyze / Eigen Analysis	313
Analyze / Factor Analysis	313
Analyze / Rotate	314
Analyze / Cluster Variables	315
Analyze / Coherence	316
Plot / Series	317
Plot / Histogram	317
Plot / Density	317
Index	321