

Contents

1. The New Era of Analytics

- 1.1 The Evolution of Analytics
 - 1.1.1 Analytics 1.0 - Era of “Business Intelligence”
 - 1.1.2 Analytics 2.0 – “Era of Big Data”
 - 1.1.3 Analytics 3.0 - Era of ‘Data-Enriched Offerings’
 - 1.1.4 The Next Big Thing, in Beta
- 1.2 Ten Requirements for Capitalizing on ‘Analytics 3.0’
- 1.3 Process Analytics - Which Metrics Matter Most?

2. Can you Measure Variation?

- 2.1 Historical Perspective
- 2.2 Scientific Inference
- 2.3 Correlation Statistics for Linear Relationships
- 2.4 Variation is Cumulative (Quincunx Experiment)
- 2.5 Advanced Analytics
- 2.6 Can you predict (with models) within the Natural Variation of the process?
- 2.7 Box Plots
- 2.8 Difference between “Statistical Estimates” and “Population Parameters”
- 2.9 Enumerative and Analytical Statistical Studies
- 2.10 Select References

3. Control Charts

- 3.1 Control Chart Guide

4. Data Quality Assessment (Often overlooked)

- 4.1 “Descriptive Statistics” (Good Starting Point for Data Quality Assessment)
 - 4.1.1 JMP Exercise - Develop “Descriptive Statistics”
 - 4.1.2 Minitab Exercise - Develop “Descriptive Statistics”
- 4.2 Data Quality Verification - “Differences Charts”
- 4.3 Process Data and Data Quality
- 4.4 Trend Charts (MatTempCL, BlenderAmpsSL)
- 4.5 X-Y Scatter Plots

5. Correlation Coefficient

- 5.1 Is r statistically significant?
- 5.2 Hypothesis Testing
- 5.3 Level of Significance
 - 5.3.1 Exercise in JMP and Minitab for correlation coefficient
- 5.4 Autocorrelation coefficient (lag 1) – r_1

5.4.1 Example Data Set: “Line Moisture”

6. Control Charts for ‘Autocorrelated’ Data

6.2.1 Exercise: Control Chart adjusted for lag-1 autocorrelation

6.2.2 JMP - Generating Autocorrelated Control chart using “Residuals”

6.2.3 Minitab – Generating Autocorrelated Control Chart using Residuals

7. Exponentially Weighted Moving Average (EWMA)

7.1 Comparing the Individuals Control Chart to the EWMA Control Chart

7.2 Calculating the EWMA

7.3 Calculating the Control Limits for the EWMA Control Chart

7.5 EWMA – Minitab Exercise

7.6 EWMA Control Limits when the Data are Autocorrelated

7.6.1 EWMA Autocorrelated Data Exercise

8. Multivariate Control Charts -“Hotelling’s T2 Statistic”

8.1 Hotelling’s T2 Statistic and Control Limits for Individual Observations

8.2 Recall F- Distribution

8.3 Hotelling T2 Exercise

8.3.1 Hotelling’s T2 Control Chart - JMP Example

8.3.2 Hotelling T2 Control Chart – Minitab Example

8.3.3 Select Multivariate SPC References

9. Probability Density Functions (pdfs) ~ “Distributions”

9.1 Discrete Density Functions

9.1.1 “Binomial Distribution”

9.1.2 Binomial Distribution Exercise – Bead Box

9.1.3 Poisson Distribution

9.2 Continuous Density Functions

9.2.1 Normal Distribution

9.2.2 Standard Normal Distribution

9.2.3 Weibull Distribution

9.2.4 Chi-Square Distribution

9.2.5 t-distribution or “Student’s T Distribution”

9.2.6 F Distribution

9.2.7 Lognormal Distribution

9.3 Distribution Fitting

9.4 Distribution Fitting in JMP

9.5 Distribution Fitting in Minitab

10. Sampling

- 10.1 Random Sample
 - 10.1.1 Simple Random Sampling
 - 10.1.2 Sampling Without Replacement
 - 10.1.3 Sample Size n for Estimating Population Mean
 - 10.1.4 Example – Estimating the Sample Size n for Estimating Population Mean
 - 10.1.5 Sample Size for Estimating a Proportion (p)
- 10.2 Stratification and Stratified Random Sampling
 - 10.2.1 Excel Exercise for Sample Size for Estimating Sample Mean
 - 10.2.2 Excel Exercise Stratification and Stratified Random Sampling
 - 10.2.3 Sample Size Determination with OC Curves, Minitab and JMP

11. Simple Linear Regression

- 11.1 Scatter Plot and Linear Fit
- 11.2 Curvilinear Statistical Relations
- 11.3 Regression Model
- 11.4 Theoretical Regression Model – Simple Linear Regression (SLR)
 - 11.4.1 SLR – Meaning of Parameters

12. Method of Least Squares - Estimation of Regression Function

- 12.1 Estimation of Regression Function
- 12.2 Least Squares Example
 - 12.2 SLR Example using JMP 14 or Minitab 18
 - 12.3 SLR Example using Minitab 18
- 12.4 SLR – Case Study (For Some you had in SPC Training)
- 12.5 Review - What are Confidence Intervals?
- 12.6 Review - What Are Prediction Intervals? (Recall a Control Chart)
 - 12.6.1 Confidence interval of the prediction
 - 12.6.2 Prediction interval
- 12.7 Analysis of Variance (ANOVA)
 - 12.7.1 F Distribution
- 12.8 R^2 or “Coefficient of Determination”
- 12.9 Diagnostics for Residuals
 - 12.9.1 Plots for Residuals
- 12.10 SLR Case Study Analysis in JMP and Minitab
 - 12.10.1 SLR Example in JMP 14
 - 12.10.2 SLR Example in Minitab 18

13. Multiple Linear Regression (MLR)

- 13.1 Model Building - Multiple Linear Regression (MLR)
 - 13.1.1 Training and ‘Validation’ (or ‘Test’) Data Sets
 - 13.1.2 Using Validation in JMP

- 13.1.3 Make Validation Column
- 13.2 Stepwise Regression
 - 13.2.1 Multicollinearity
 - 13.2.2 VIF - Statistic to assess severity of multicollinearity
 - 13.2.3 Criteria for Good Candidates - "Best Subset"
 - 13.2.4 A Note on Sample Size
- 13.3 Select References
- 13.4 MLR and Problems with Industrial Data
 - 13.4.1 Statistics for identifying undue leverage
 - 13.4.2 Effects of Autocorrelation on MLR
 - 13.4.3 Stringent Assumptions of MLR
- 13.5 MLR Exercise in JMP
- 13.6 MLR Exercise in Minitab

14. Model Building

- 14.1 Variable Pre-Selection
- 14.2 Model Screening Generalized Regression
- 14.3 All Possible Models

15. Logistic Regression

- 15.1 Simple Logistic Regression
- 15.2 Logistic Response Function
 - 15.2.1 Parameter Estimation using JMP (Binary Logistic.xlsx)
 - 15.2.2 Parameter Estimation in Minitab
- 15.3 Multiple Logistic Regression
 - 15.3.1 Multiple Logistic Regression in JMP
 - 15.3.2 Multiple Logistic Regression in Minitab
- 15.4 Multinomial Logistic Regression
 - 15.4.1 Multinomial using JMP
 - 15.4.2 Multinomial in Minitab
- 15.5 Poisson Regression
 - 15.5.1 Poisson Regression in Minitab
 - 15.5.2 Poisson Regression in JMP
- 15.6 Ordered Logistic Regression (Optional)

16. Decision Trees (DT)

- 16.1 Regression Trees ("Greedy Algorithms ~ Overfit")
- 16.2 Decision Tree - Structure
- 16.3 Decision Tree - Induction
- 16.4 Decision Tree - Software
- 16.5 Building a "Regression Tree" Model in JMP

- 16.6 'Boosted' Regression Trees
 - 16.6.1 Boosting
 - 16.6.2 Boosted Regression Trees (BRT)
 - 16.6.3 BRT Method
 - 16.6.4 The Partition platform in JMP
 - 16.6.5 BRT Example in JMP
- 16.7 'Bootstrap Forest'
 - 16.7.1 Example of Bootstrap Forest with a Continuous Response
 - 16.7.2 The Bootstrap Forest Report

17. Introduction to Multivariate Analysis

- 17.1 Introduction
 - 17.1.1 Univariate versus Multivariate
 - 17.1.2 Matrix Representation
 - 17.1.3 Variance, Covariance and Correlation
 - 17.1.4 Hidden Data Structures
 - 17.1.5 Objectives of Multivariate Analysis Techniques
 - 17.1.6 Projection Methods
- 17.2 Principal Component Analysis (PCA)
 - 17.2.1 First Principal Component (PC1)
 - 17.2.2 Higher-Order Principal Components (PC2, PC3, ...)
 - 17.2.3 Principal Components Model Scores – Loadings
 - 17.2.4 Objectives of PCA
 - 17.2.5 Score Plot and Interpretation
 - 17.2.6 Loading Plot (line) and Interpretation
 - 17.2.7 NIPALS Algorithm
 - 17.2.8 Lab Exercise – PCA for Exploratory Analysis 1
 - 17.2.9 Lab Exercise – PCA for Exploratory Analysis 2
- 17.3 Multivariate Regression (PCR/PLS)
 - 17.3.1 Calibration –Prediction – Validation
 - 17.3.2 Cross–Validation
 - 17.3.3 Multiple Linear Regression (MLR)
 - 17.3.4 Principal Component Regression (PCR)
 - 17.3.5 PLS (Partial Least Squares Regression /Projection to Latent Structures)
 - 17.3.6 PLS Discriminant Analysis (PLS-DA)
 - 17.3.7 Outliers
 - 17.3.8 Lab Exercise – PLS Regression 1
 - 17.3.9 Lab Exercise – PLS Regression 2
 - 17.3.10 Lab Exercise – PCA for Exploratory Analysis 3
 - 17.3.11 Variance Inflation Factor
 - 17.3.12 Lab Exercise - MLR
 - 17.3.13 Lab Exercise – Standardize

17.3.14 Lab Exercise – Calculate VIF

18. Introduction to Neural Networks

18.1 Introduction

18.2 Neuron Model and Transfer Functions

18.3 Feedforward Network

18.4 Considerations

18.5 Neural Networks Outline

18.6 Training Algorithm

18.7 Lab Exercise: Neural Networks 1

18.8 Lab Exercise: Neural Networks 2

19. Additional Lab Exercises

19.1 Lab Exercise: PLS on Jam Dataset

19.2 Lab Exercise: Validation Feature in JMP

20. Lab Exercises with Minitab

20.1 Lab Exercise – PCA for Exploratory Analysis 1

20.2 Lab Exercise – PLS Regression 1

20.3 Lab Exercise – PLS Regression 2

20.4 Lab Exercise – PCA for Exploratory Analysis 3

20.5 Lab Exercise - MLR

20.6 Lab Exercise – Standardize