2745-8716

MODELING AND DATA TREATMENT IN THE PHARMACEUTICAL SCIENCES

J. T. Carstensen, M.Ch.E., Ph.D.
Professor Emeritus
The University of Wisconsin

TECHNOMIC PUBLISHING CO., INC.
LANCASTER · BASEL

Table of Contents

Introduction xiii

1. PURPOSE OF PHARMACEUTICAL RESEARCH **1**

1.1 Literature Searches 1
1.2 Research Motivation 1
1.3 Negative Modeling 2
1.4 Hypothetical Models 2
1.5 Strict Modeling 3
1.6 Pseudomodeling 3
1.7 Curve Fitting 3
1.8 One-Point Experimentation 4
1.9 References 4

2. THE SINGLE EXPERIMENT **5**

2.1 Data in Tables 5
2.2 Correlation 5
2.3 Formula Considerations 6
2.4 Specifications 8

3. REPRESENTATION OF NUMBERS AND DATA **9**

3.1 Tabular Presentations 9
3.2 Means 9
3.3 The Dispersion 10
3.4 Measures of Dispersion 11
3.5 Degrees of Freedom 12
3.6 Special Cases of Standard Deviations 13
3.7 Populations and Samples 13
3.8 The Pooled Standard Deviation 15

3.9 Problem 17
3.10 Answer 17

4. COMPUTER PROGRAMS **19**

4.1 Data Generation from Equations, BASIC 19
4.2 StatWorks™ 21
4.3 SigmaPlot® 22
4.4 Graphical Program (CricketGraph™) 23

5. CURVE FITTING AND PHENOMENOLOGY **25**

5.1 Curve Fitting of Continuous Curves 26
5.2 Curve Fitting 26
5.3 Curve Fitting by Computer Graphing Programs 26
5.4 Domains 28
5.5 Selection of Ordinate and Abscissa 28
5.6 Common Considerations 31
5.7 General Curve Forms 34
5.8 Second-Order Curve, First-Order Curve with Equilibrium 35
5.9 The Sigma-Minus Plot 36
5.10 The Logarithmic Plot 38
5.11 Polynomial Plots 38
5.12 Inverse Plots 39
5.13 Deduction and Curve Fitting 41
5.14 Are the Data Monophasic or Biphasic? 45
5.15 Curve-Fitting the Normalized Frequency Function 48
5.16 Probability Plots 51
5.17 Weibull Plotting 53
5.18 Problems 53
5.19 Answers 55
5.20 References 60

6. NORMALIZED FREQUENCY DISTRIBUTIONS **61**

6.1 Histograms 61
6.2 Making Frequency Distributions Comparable 63
6.3 Normal (Gaussian) Distribution 66
6.4 Probability Calculations from the Normal Distribution 68
6.5 Cumulative Presentation: Distribution Function 68
6.6 Construction of Probability Paper 70
6.7 Particle Sizes: Log-Normal Distributions 71

6.8 Statistical Moments 73
6.9 References 74

7. OTHER DISTRIBUTIONS **75**

7.1 Populations 75
7.2 The χ^2-Distribution, Confidence Intervals 77
7.3 Outliers 79
7.4 The Square Distribution: Random Numbers 82
7.5 The Student *t*-Distribution 82
7.6 Confidence Limits: Significant Numbers 83
7.7 The Binomial Distribution 84
7.8 Problem 86
7.9 Answer 86
7.10 References 87

8. SIGNIFICANCE TESTING **89**

8.1 Null Hypothesis 89
8.2 Testing by Comparing Variances 90
8.3 The *F*-Test 91
8.4 Factorials 92
8.5 Differences between Estimates of *s* 95
8.6 The Unpaired *t*-Test: Samples at Hand the Only Source of Variance 96
8.7 Checking the Meaning of Parameters in Programs: Program Validation 98
8.8 Unequal Variances 99
8.9 Paired *t*-Test 100
8.10 χ^2-Test 101
8.11 The Wilcoxon/Friedman Test 104
8.12 Equivalence Testing 105
8.13 The Power of a Test 106
8.14 Test for Normality 107
8.15 Problems 110
8.16 Answers 111
8.17 References 114

9. SAMPLING **115**

9.1 Sampling by Variance 115
9.2 Sampling by Attributes 117
9.3 A Pharmaceutical Example 118
9.4 Blending Validation 119
9.5 Control Charts 120

9.6 Arbitrary and Tabulated Sample Sizes 124
9.7 Problems 124
9.8 Answers 125
9.9 References 126

10. LEAST SQUARES FITTING 127

10.1 Equation for a Straight Line 127
10.2 Concept of Least Squares Fitting 128
10.3 Derivation of the Least Squares Fit Equation 132
10.4 A Long-Hand Example 133
10.5 The Slope 134
10.6 The Intercept and Extrapolated Estimations 135
10.7 Prediction Intervals 136
10.8 The Correlation Coefficient, *R* 136
10.9 Upper and Lower Confidence Bounds: Programming the Least Squares Fit 137
10.10 Expiration Periods 138
10.11 Programs for Regression with Confidence Bounds 140
10.12 Assumptions Made in the Least Squares Fit 144
10.13 ANOVA of Least Squares Fit 145
10.14 Weighted Least Squares 147
10.15 Significance of Correlation and Curvature 148
10.16 The Significance of the Correlation Coefficient 148
10.17 Durbin-Watson Statistics 149
10.18 Multiple Regression 151
10.19 Transformations 152
10.20 Nonlinear Regression 154
10.21 Weighting of Transformed Variables 154
10.22 Polynomials 155
10.23 Least Squares Fits of Distribution Functions 157
10.24 Reference 159

11. ITERATION 161

11.1 One Iterant 161
11.2 Nonlinear Fitting 164
11.3 Curvature 166
11.4 Multiple Parameter Curve Fitting 167
11.5 A Pharmaceutical Example (Shear Cell) 172
11.6 Multiple Iterants 174
11.7 Deconvolution 177
11.8 The Least Squares Wrist 180

11.9 Modeling by Comparison and Statistics 181
11.10 Zero-Point Problems 181
11.11 An Example of Multiple Iterants 182
11.12 Appendix 185
11.13 References 188

12. FACTORIALS AND PHENOMENOLOGY 189

12.1 The Simple Factorial 190
12.2 Screening for Variables 192
12.3 Multiple Specifications 193
12.4 Dimensionless Analysis 195
12.5 The Differentiation Approach 198
12.6 Examples of Differential Fitting 199
12.7 Monophasic versus Biphasic Profiles 202
12.8 Optimization 202
12.9 References 204

13. MONTE CARLO METHOD AND SIMULATION 205

13.1 Description of the Steps of the Process 205
13.2 The Initial Thought-Process 205
13.3 Monte Carlo Method 207
13.4 Simulation 211
13.5 References 215

14. PSEUDOMODELING 217

14.1 Relying on Previous Models 217
14.2 Pseudomodeling 219
14.3 Going Backwards 220
14.4 General Equations Imitated in Reverse Deduction 221
14.5 A Pharmaceutical Example: Auto-Oxidation 222
14.6 Curves with Extrema 222
14.7 Models with Functional Assumptions 224
14.8 Models with Input Functions: A Pharmaceutical Example 225
14.9 Arriving at Differential or Other Equations Not Solvable in Closed Form 229
14.10 References 233

15. MODELING 235

15.1 Definition of Modeling 235
15.2 Getting Started 235

15.3 Literature Search 236
15.4 The Basic Idea 236
15.5 Graphing 237
15.6 The Saturday Afternoon Experiment: The Quick-and-Dirty 240
15.7 Initial Conclusions from the Graph 241
15.8 The Initial Modeling Thoughts: The Doodle 241
15.9 Stressing the Model 243
15.10 The Finishing Touches 246
15.11 Scrutinizing the Shortcuts Taken 247
15.12 It Never Ends 248
15.13 Problems 248
15.14 References 250

16. CONCLUDING REMARKS **251**

Appendices 255

Index 265