

**Raymond W. Falk, Ph.D.**  
**3985 Park Avenue (#23)**  
**Fairfield, Connecticut 06825-1262**  
**(203) 520-3793 RayFalk@OptOnline.net**

## SUMMARY

**Statistician applying complex modeling techniques in business and scientific decision support and enhancing the conceptual framework for their use:** Through a process of progressive internal refinement, I have combined sound analytical judgment with symbolic computational finesse to maintain sufficient "analytical bandwidth" to generate in parallel substantial numbers of high-quality predictive models (for response, amount, and time-to-event) saving some \$300K per month over years of on-going use. Furthermore, I have abstracted and simplified the "hyper-parameters" defining and managing the workflow associated with this powerful and robust methodology to empower colleagues involved with similar Basel II modeling of probability of default, excess loss, and the loss given default ratio.

## PROFESSIONAL EXPERIENCE

- Royal Bank of Scotland Card Services**                      **Bridgeport, CT**                      **2005-2009**  
Experience modeling acquisition and transaction data for response, amount, and attrition.  
Gamma Regression Modeling: Introduced general methodology for Gamma regression models for heavy-tailed time-to-event and dollar-amount data; developed routines for automated modeling and deployment.  
SwapSet Modeling: Developed methodology and software for analysis of the effect of repeated contact on response and origination rates.  
Call Volume Forecasting: Developed models for forecasting in-bound call volume as a function of serial correlation, seasonal effects, and historical events.  
Data Preparation: Developed systematic scripts for comprehensive processing of predictive data for general modeling purposes, including: conversion of ordinal codes and numeric symbols, detection and correction of special missing value codes, summarization of historical time series, detection and elimination of excessively small categories, detection of sources of multicollinearity, identification of representative and non-representative members of factor-analytic variable clusters, and automatic nonlinear transformation of predictors via preliminary modeling.  
Validation: Developed systems for comprehensive detection of defects and assessment of distributional change in monthly syndicated data sources.  
Integration: Unified workflow to build and deploy rafts of Logistic and Gamma regression models of high quality (%Error<0.6; KS>0.5) in parallel in record time under harrowing accelerating deadlines.
- RedShirtImaging, LLC**    **Fairfield, CT**    **2001-2005**  
Experience designing, developing, and testing software/methodology for specialty scientific instrumentation.  
Spatial Filtering: Introduced general parameterized kernels for spatial filtering.  
CardioCCD: Designed and adapted methodology and software for analysis of optical images of heart surfaces and developed user interface for specifying and navigating among electrophysiological endpoints.  
System Set-up and Testing: Experimented with software product to identify and correct defects.
- RxRemedy, Inc.**    **Westport, CT**    **1999-2001**  
Experience designing, developing, and delegating processes for quantitative modeling, analysis, and reporting.  
Weight-ratio Estimation: Introduced unsaturated log-linear modeling for estimation of post-stratification weights relative to U.S. Census Bureau CPS data. Automated production and validation procedures.  
Comorbidity Clustering: Clustering of patterns of combinations of concurrent health conditions; profiling of the database with respect to condition combinations, and identification of representative syndromes.  
Change in Market Share: Design, simulation, and deployment of studies for change in market share using stratified two-sample multinomial repeated measures model (tripling statistical efficiency).  
Persistence Analysis: Episoding of product use histories based on compliance rates, followed by formal survival analysis (eliminating negative bias) in the modeling of the persistence of individual and combination product use, transitions in product use, and evaluation of product use status at selected times.  
General: Developed methods for automated comparison and validation of diverse databases.

**Raymond W. Falk, Ph.D.**  
**3985 Park Avenue (#23)**  
**Fairfield, Connecticut 06825-1262**  
**(203) 520-3793 RayFalk@OptOnline.net**

**Metropolitan Life Insurance Company**                      **New York, NY**                      **1997-1999**

Experience as a statistician participating in business, marketing, and operations data mining in PC environment.

Individual Business: Response modeling to identify sales campaign targets and clustering into market segments.

Institutional Business: Validation of segmentations, with detailed attention to quality of discrimination; modeling transitions in roles and providers of financial services to institutional clients; clustering of preferred product combinations; and analysis of long term disability sales as a function of size of corporate customer and attitude toward price. (Syndicated data sources included SRI MacroMonitor and MoneyMarketDirectory surveys.)

Compliance: Modeling of complaint categorizations with respect to operational, administrative, product, and promotional features.

Auditing: Applied digit distribution analysis to dental claims data, testing deviations against both theoretical (Benford's Law) and empirical distributions, adapting commercially available SAS routines.

Year 2000 Compliance: Evaluations of SAS, S-Plus, Model 1/PRW, KnowledgeSEEKER, MS-Excel.

General: Developed Evenness Profiling (for Individual Business) and Cluster Silhouette Profiling (for Institutional Business) into fully documented, flexible applications available for external use.

**Nets Inc.**    **Cambridge MA**    **1997-1997**

Internet Usage Analysis: Regression Tree analysis of clustered usage patterns to separate distinct users. Survival analysis of duration of cumulative site usage as a function of date of initial contact and characteristics of user and host page. Multidimensional Scaling of host pages based on joint usage by individual users, with page usage density in the resulting coordinate system estimated via sums of pairwise smoothed functions. Analyses were conducted using S-Plus on a Sun (Solaris) Workstation on data extracted from an Oracle database.

**Statistical Consulting / SAS Programming**                      **Stamford CT**    **1996-1997**

Business Applications Programming: SAS programming against Sybase RDBMS on a Sun (Solaris) Workstation in a UNIX environment.

Credit Risk Analysis: Logistic regression modeling of account delinquency as a function of credit bureau data with imputation of responses for rejected applications based on preliminary modeling. Analyses were conducted using SAS in IBM-Mainframe (VM/CMS,ISPS) and networked PC environments.

**Bayer Corporation**    **West Haven CT**    **1987-1995**

Experience as a statistician participating in clinical trials research in IBM-Mainframe (VM/CMS) environment

Cerebrovascular Indications: Re-analysis of background studies for the purpose of designing future trials involving sequelae of stroke and Alzheimer's disease, including selection of response scales, identification and organization of endpoints, and sample-size determination. Critical influence on successful NDA in subarachnoid hemorrhage.

Devised computational methods for incrementally extracting relevant, but not consistently labeled information from a database so as to put critical information into machine-readable form. Matched several databases on time of day and continuous response variables in order to identify mislabeled items and exotic patterns of mismatches.

Cardiovascular Indications: Analysis of clinical trial for anti-anginal indication. Complete programming, analysis, and reporting of results. Pharmacodynamic analysis of change from baseline in blood pressure series using the SAS/ETS/MODEL, SYSLIN, and SYSNLIN procedures.

Anti-infective Indications: Analysis and reporting of results for selected studies. Methodological research into testing specified differences between proportions for purposes of establishing equivalence to an active control. Implemented novel methodology through SAS/Macro routines (allowing interchangeable numeric and character factors for contingency tables).

Pharmacokinetics: Automated analyses to establish equivalence or dose-proportionality between regimens using paired one-sided tests for parameter ratios under parallel, crossover, and ascending-dose study designs with log-scale and arithmetic-scale (Fieller theorem-based) analyses of non-compartmental parameter estimates. Integrated calculations for contrasts across multiple error terms in mixed effects analysis of

**Raymond W. Falk, Ph.D.**  
**3985 Park Avenue (#23)**  
**Fairfield, Connecticut 06825-1262**  
**(203) 520-3793 RayFalk@OptOnline.net**

variance. Achieved significant increase in methodological generality without delay in workflow. Reduced response time for typical PK statistical reports from 2 weeks to 8 hours.

Pharmacodynamics: Nonlinear mixed effects analyses (implemented via iterative application of the SAS/MIXED procedure) of a composite pharmacodynamic model for the response as a logistic function of effective drug concentration at the site of action, estimated as a parameterized convolution of the concentration series. Displayed cascade of physiological/biochemical effects concisely through a series of dual vertical axis graphs of successive responses against time; developed highly visible color display using SAS/GRAPH.

General: Programmed SAS/Macro tools for database development, review, and revision, and managed control specifications flexibly, so as to remove bottlenecks and encourage anticipation of future demands and transfer of technical solutions from similar situations. Developed SAS/Macro to merge collections of SAS formats. Simultaneously enhanced precision/accuracy and flexibility of operations. Achieved *instant turnaround* in some situations.

Incorporated confidence bands for *pegged* linear regression functions, fitting  $\alpha + \beta X$  and then estimating  $\beta(X-x)$  or  $(\alpha + \beta X)/(\alpha + \beta x)$ , into a general purpose S-Plus function. (Pegged linear regression functions have triangular (rather than hyperbolic) confidence bands, a desirable feature in appropriate settings.)

**G.D. Searle & Company**

**Skokie IL**

**1985 -1987**

Experience as a statistician participating in pre-clinical research in VAX/VMS environment

Radioimmunoassay: Comparison of inhibition curves using standard logistic multiple regression analysis implemented using the SAS/NLIN procedure.

Genetic Toxicology: Development of screening protocol for the UDS assay.

Drug Design: Identification of sequences of compounds representing progressive improvement simultaneously across multiple criteria; reprogrammed PL/I routine in SAS.

**Raymond W. Falk, Ph.D.**  
**3985 Park Avenue (#23)**  
**Fairfield, Connecticut 06825-1262**  
**(203) 520-3793 RayFalk@OptOnline.net**

## **HARDWARE/SOFTWARE SKILLS**

**SAS** on IBM-Mainframe (VM/CMS), VAX(VMS), Sun(Solaris), and IBM-PC (Windows95/98), 25 years,  
(SAS/STAT,ETS,QC,GRAPH,SQL, extensive MACRO programming)

Developed rank-based scoring facility for diverse types of data for purposes of quantifying evenness across classification systems with diverse modes of valuation within or between individual categories.

**S-Plus** on IBM-PC and Sun(Solaris), 10 years

Contributed *merge.data.frame* to StatLib@lib.stat.cmu.edu: Devised S-Plus function to perform many-to-many merges between data frames. A novel feature of the function is its automated identification of like-named columns as the key for matching. The algorithm can also be implemented using the SAS/IML procedure

Extended the S-Plus silhouette-based clustering functions to generate and display silhouette profiles for external data, including selective generation of cluster assignments where not explicitly provided.

**Model 1/PRW** on IBM-PC, 2 years

Response modeling and segmentation/clustering using multiple modeling techniques (logistic regression, neural networks, and tree models), with consistency among best fitting models applied to develop hybrid models and/or down-weight consistently scored records for further modeling.

**Q&A** database on networked PC – solved ‘multiple groups’ problem (associated attributes as successive records).

**KnowledgeSEEKER** on IBM-PC, 2 years

**PL/I** and Fortran on IBM-Mainframe, Lisp(PC-Scheme), **IDL** on IBM-PC

**MS-Office** - Word, Excel, Access; WordPerfect (5.2 & 6.1)

## **EDUCATION**

### **University of North Carolina at Chapel Hill**

1985 **Ph.D.** Biostatistics

1980 **M.S.** Biostatistics

1976 **B.S.** Zoology

## **PUBLICATIONS**

Falk, R.W. (1985) “*L’-Superadditive Function and Concepts of Multivariate Dependence*” (Doctoral dissertation written under Prof. P.K. Sen).

Oshiro, Y., Balwierz, P.S., Falk, R.W., and Piper, C.E. (1987) Decision criteria for the *in-vitro* rat hepatocyte UDS assay” *J. Applied Toxicology* **7(6)**: 379-385.

Falk, R.W. (1989) “Hommel’s Bonferroni-type inequality for unequally spaced levels” *Biometrika* **76(1)**: 189-191.

Falk, R.W. (1990) “Basic guidelines for inference from clinical trial data” *Drug Information Journal* **24**: 507-512.

Falk, R.W. and Koch, G.G. (1998) “Testing a specified difference between proportions” *Biometrics* **54(4)**: 1602-1614.

Falk, R.W. “The General Hommel-Rüger Inequality” (Submitted 1996).

Natural Languages: French, German  
**MEMBER - ASA, IMS, IEEE, INFORMS**