

Die Entwicklung der Biometrie von den 1960er Jahren bis heute



Jürgen Hedderich, Prof. Dr. Lothar Sachs
Institut für Medizinische Informatik und Statistik
Christian-Albrechts-Universität Kiel

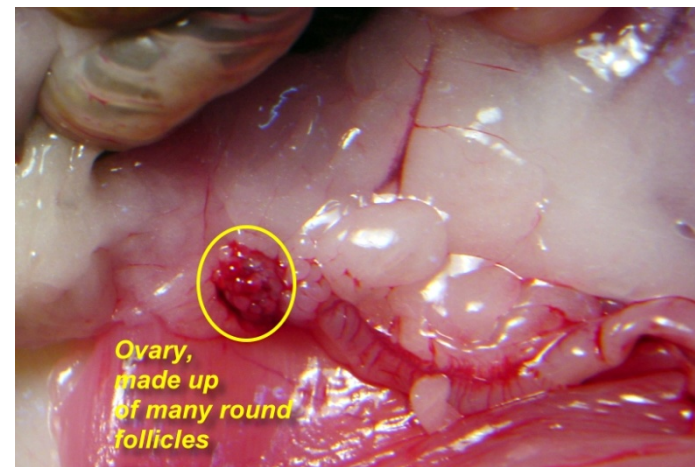
Statistisches Denken in der Medizin

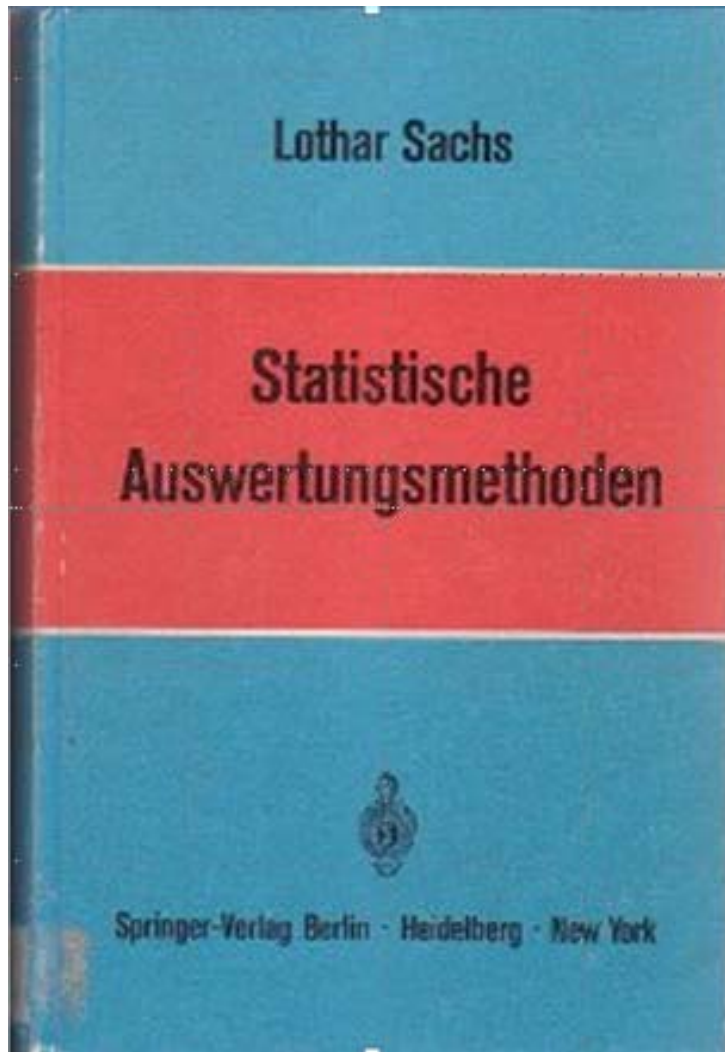
Theorie und Praxis im Wandel der Zeit

- Diskursives Denken
- Fortschritte in der Wissenschaft
- Angewandte Statistik
- Prägende Faktoren
 - Zeitgeist (Mode und Trend)
 - Motivation
 - Technische Möglichkeiten
 - Erfahrungen

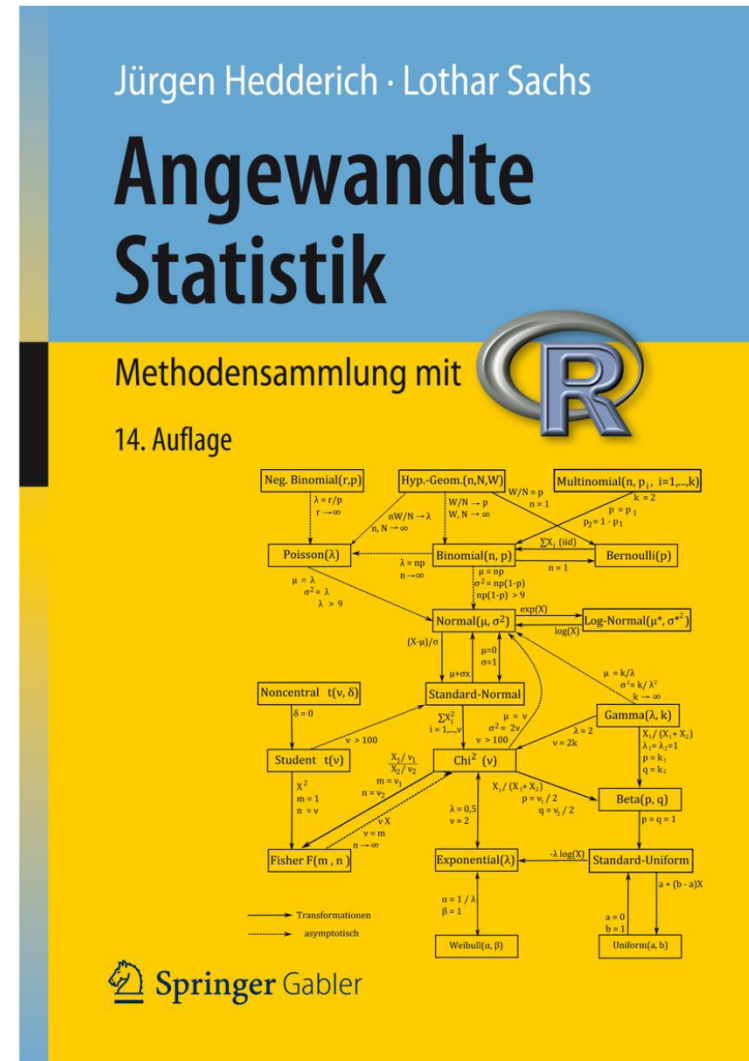
Endokrinologische Studien an Ratten-Ovarien

- Wistar-Ratten
 - unterschiedliche Charakteristika der Ovarien einzelner Tiere
- Varianzanalyse
- Holtzman-Ratten
 - praktisch "gleiche" Ovarien





1968



2012

Empfehlungen des Wissenschaftsrates (1960) zum Aufbau wissenschaftlicher Einrichtungen

- "**Medizinische Statistik** einschließlich zugehöriger Dokumentation ist für die medizinische Forschung unentbehrlich, bisher jedoch in den medizinischen Fakultäten fast nicht vertreten."
- "Jede Fakultät sollte daher einen Lehrstuhl erhalten, dessen Hauptaufgabe in **der Unterstützung der Kliniken** liegt; er könnte aber auch für die medizinische Statistik in den theoretischen Fächern zuständig sein."
- "Da die **Nachwuchslage unbefriedigend** ist, können Lehrstühle für medizinische Statistik zunächst nur an wenigen Hochschulen eingerichtet werden (Schwerpunktbildung)."

Abteilung für Medizinische Dokumentation und Statistik Klinikum der CAU Kiel (1965)



Prof. Dr. Gerd Griesser
(1918-2001)

Die statistische Arbeitsweise

- Sammeln von Daten
- Ordnen und Klassifizieren
- Vergleichen
- Interpretieren
- Schließen auf das
 - Besondere
 - Allgemeine

Gegenstandskatalog für den ersten Abschnitt der Ärztlichen Prüfung (Dezember 1973)

"Er (der Student) muss den **Unterschied zwischen funktionalen und statistischen Gesetzmäßigkeiten** verstanden haben und daher wissen, dass das Kriterium der ‚exakten‘ Naturwissenschaften, die Reproduzierbarkeit des gesetzmäßigen Verhaltens im Einzelfall, in den biologischen Wissenschaften wegen der **individuellen Variabilität** durch die Reproduzierbarkeit des gesetzmäßigen Verhaltens einer **Gesamtheit gleichartiger Fälle** ersetzt wird."

WHAT Statistics should be Taught! (Karachi 1979)

[...]

4. The **variability** of biological, clinical laboratory, psychological, social and environmental data.
5. Measures of **central tendency and location**: mean, median, mode, their uses and interpretation.
6. Measures of **variation**: range, percentiles, variance, standard deviation.
7. The concept of **probability**.

[...]

9. **Statistical inference**: tests of significance, their uses and interpretation.
10. The concepts of **association and causality**: measures of regression and correlation and their interpretation.

Beispiel 44. Die Werte des Beispiels 33 (Kartoffelerträge zweier Sorten an 30 verschiedenen Plätzen) sind nach der Größe angeordnet und durchnummeriert worden, wobei gleiche Werte in beiden Reihen den Mittelwert der Rangplätze erhielten.

Sorte A		Sorte B	
Ertrag	Rangnummer	Ertrag	Rangnummer
165	1	170	2
172	3		
176	4		
182	5		
		183	6
188	8	185	7
191,191	9,10		
193	11		
		194	12
195	13,5	195	13,5
197	15		
		199	16
		202	17
203,203	19,19	203	19
		207	21
208	22		
210	23,5	210	23,5
		212	25
		213	26
217	27		
		225	28
		231	29
		240	30
	190 = R_1		275 = R_2

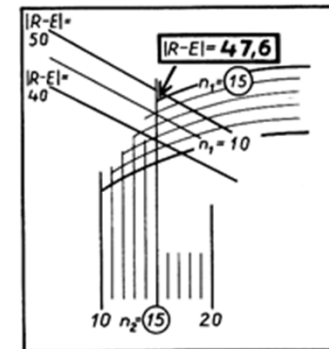


Abb. 21. Ablesegraphik zu Tafel 13 a, Beispiel 44

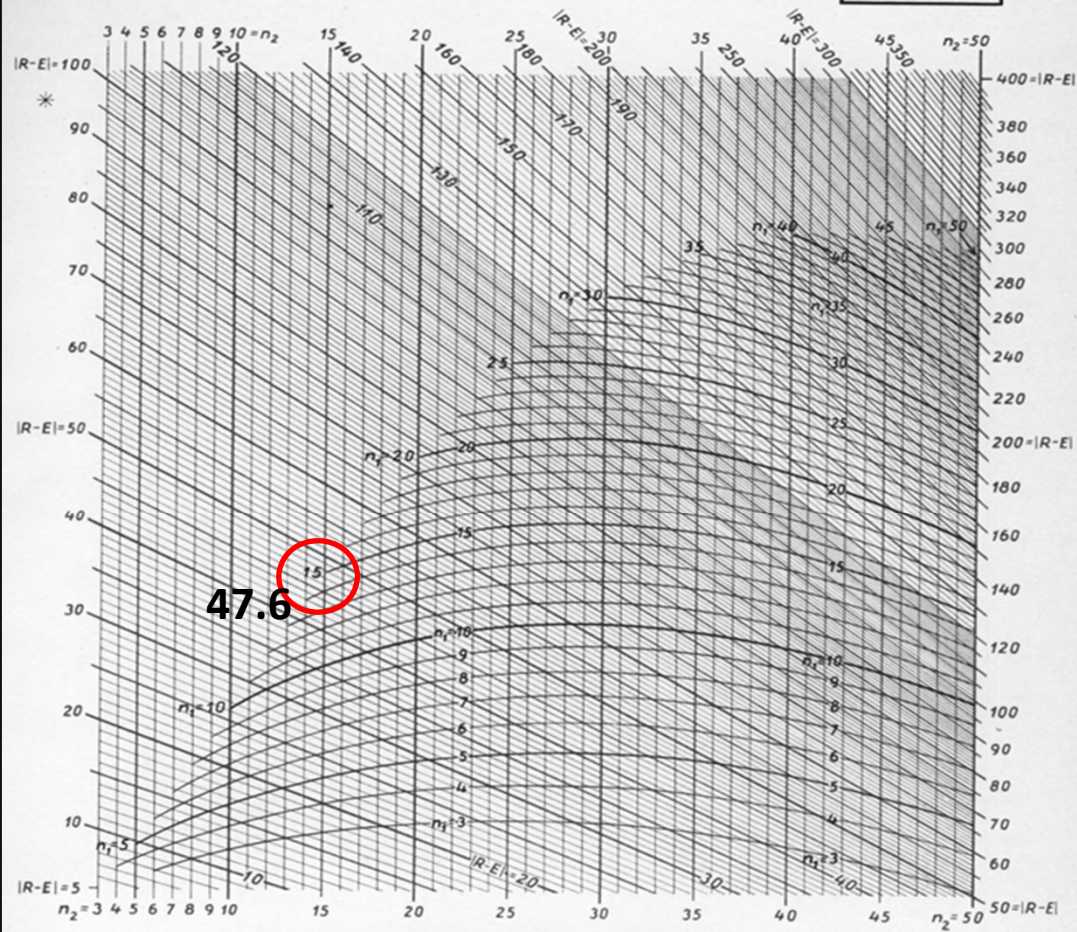
Die Kontrollsumme ist $R_1 + R_2 = \frac{1}{2} 30 \cdot 31 = 465$.

Die Testdifferenz wird $D = |R_1 - \frac{1}{2} 15 \cdot 31| = |190 - 232,5| = 42,5$

und erreicht nicht die in Tafel 13 a erlaubte Differenz von 47,6 (vgl. Abb. 21). Ein Unterschied ist nicht zu sichern. (Auch beim t-Test ergab sich kein auffälliger Unterschied.)

zugelassene
Abweichung

$\alpha = 5\%$

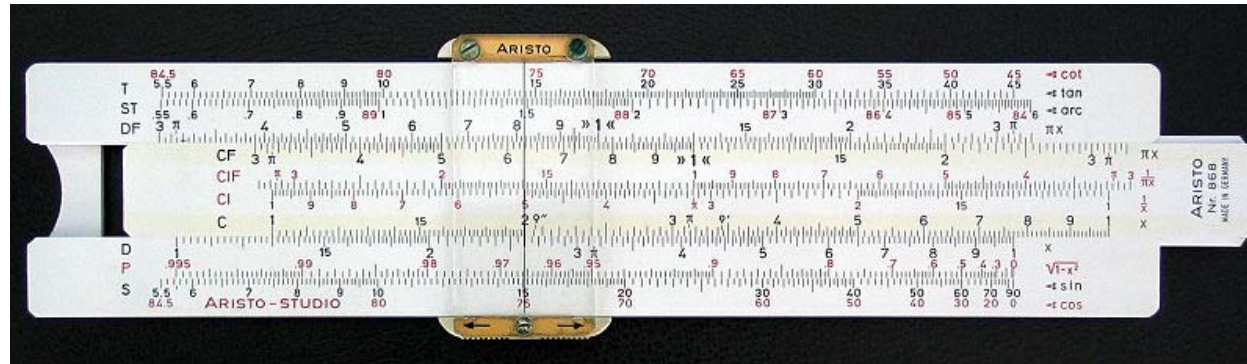


Tafel 13. Wilcoxon-Rangtest

für 2 unabhängige Beobachtungsreihen ($n_1 \leq n_2$)
 $|R - E|$ Größte zugelassene Zufallsabweichung (Ordinate)
 n_1 Umfang der kleineren Reihe (Kurven)
 n_2 Umfang der größeren Reihe (Abszisse)
 a) Irrtumswahrscheinlichkeit $\alpha = 5\%$

Rechenkunst: frühe Werkzeuge

Rechenstab



elektrischer Tischrechner



"Handkurbel"

Rechenkunst: frühe Werkzeuge

Taschenrechner (HP)



Personal-Computer (PC)



"Give us the tools and we will finish the job."
(Sir Winston Churchill, radio broadcast 1941)

Statistik-Software "konfektioniert"

- 1961 **BMDP** - Biomedical Computer Programs; entwickelt an der University of California (UCLA) unter W. Joseph Dixon
- 1968 **SPSS** - Statistical Package for the Social Sciences; entwickelt an der University of Stanford unter Norman Nie, Hadlan Hull und Dale Bent (1983 1. PC-Version, 2009 unter PASW bei IBM)
- 1971 **SAS** - Statistical Analysis System; Firmengründung 1966 durch Anthony Barr, James Goodnight und John Sall. Wurde in den 80er Jahren in C vollständig neu programmiert und war unter verschiedenen Betriebssystemen lauffähig.

Computational Statistics



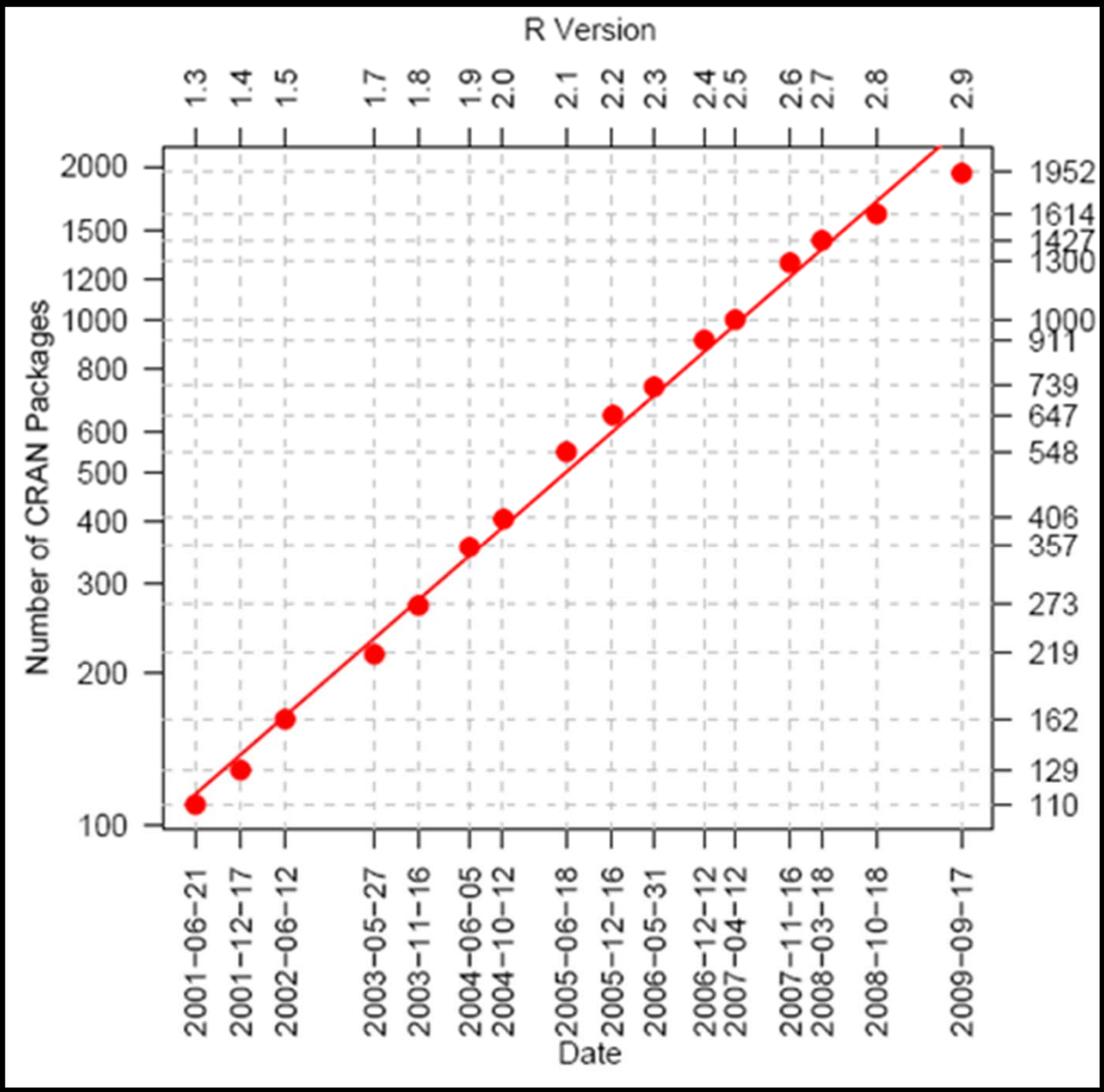
Schloss Reisenburg bei Ulm

Wissenschaftliche Anwendung
und Weiterentwicklung von
Statistiksoftware

- Anforderungen der Fachwissenschaften
- Methodensammlung
- Vergleich der Programmsysteme
- Anforderungen an Rechenzentren

Statistik-Software "frei ..."

1997 **R** - Programming Language and Software Environment for Statistical Computing; entwickelt ursprünglich durch John Chambers an den Bell Laboratories, portiert durch Ross Ihaka und Robert Gentleman an der University of Auckland (GNU-Basis: general public licence)



Fox, J. (2009) Aspects of the Social Organization and Trajectory of the R Project. The R Journal 1/2



Karl Pearson (1857-1936)

X. *On the Criterion that a given System of Deviations from the Probable in the Case of a Correlated System of Variables is such that it can be reasonably supposed to have arisen from Random Sampling.* By KARL PEARSON, F.R.S., University College, London*.

THE object of this paper is to investigate a criterion of the probability on any theory of an observed system of errors, and to apply it to the determination of goodness of fit in the case of frequency curves.



William S. Gosset (1876-1937)

VOLUME VI

MARCH, 1908

No. 1

BIOMETRIKA.

THE PROBABLE ERROR OF A MEAN.

By STUDENT.

Introduction.

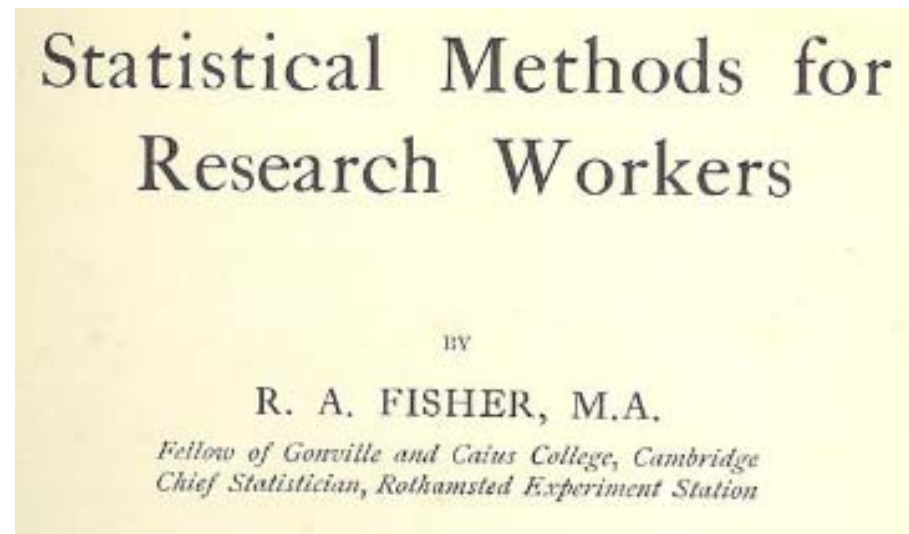
ANY experiment may be regarded as forming an individual of a "population" of experiments which might be performed under the same conditions. A series of experiments is a sample drawn from this population.

XV.—**The Correlation between Relatives on the Supposition of Mendelian Inheritance.** By **R. A. Fisher, B.A.** *Communicated by Professor J. ARTHUR THOMSON.* (With Four Figures in Text.)

(MS. received June 15, 1918. Read July 8, 1918. Issued separately October 1, 1918.)

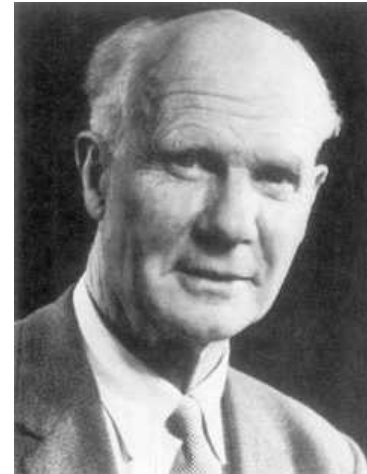


Sir Ronald A. Fisher (1890-1962)





Jerzy Neyman (1894-1981)



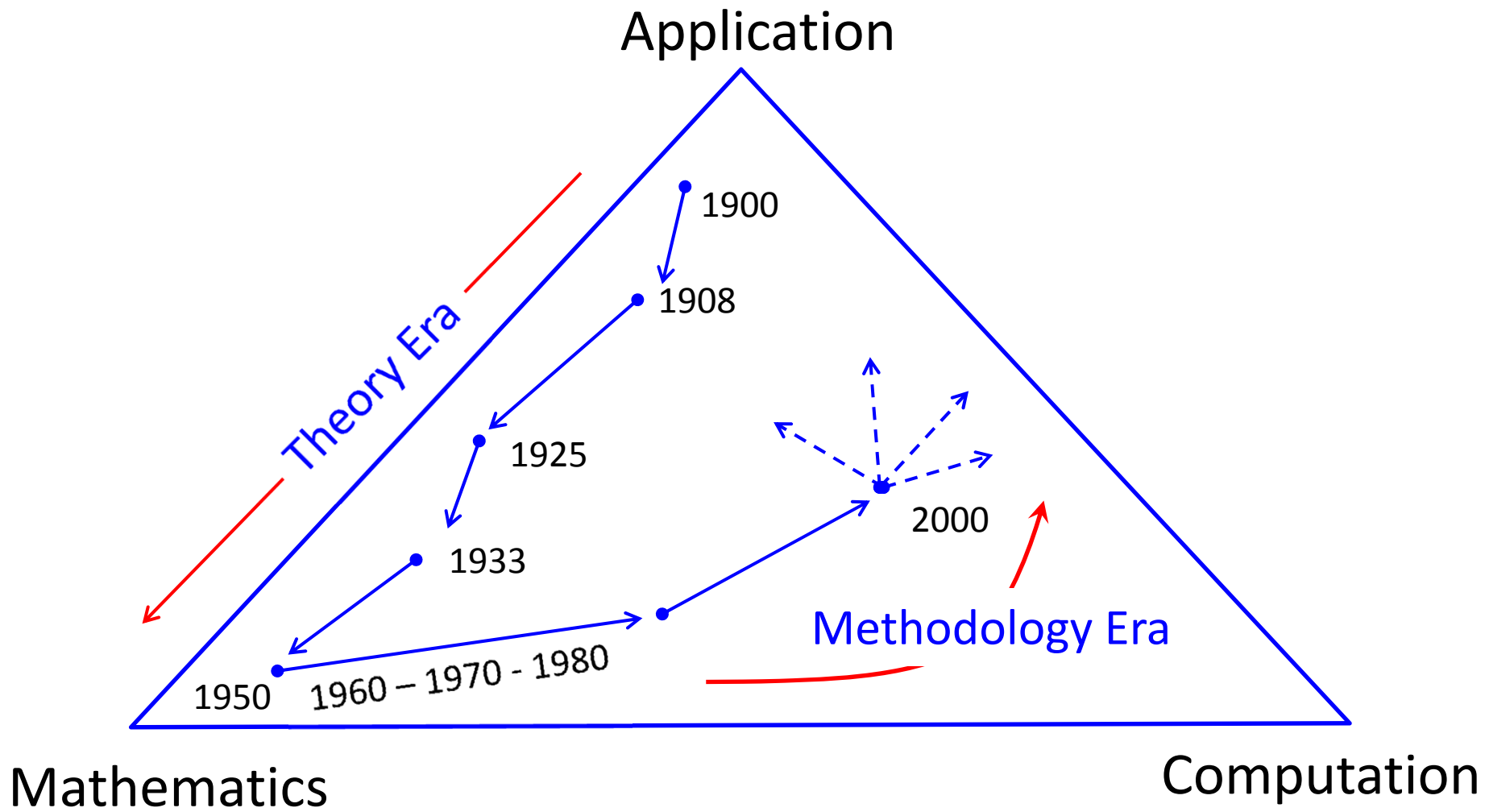
Egon Pearson (1895-1980)

IX. On the Problem of the most Efficient Tests of Statistical Hypotheses.

By J. NEYMAN, Nencki Institute, Soc. Sci. Lit. Varsoviensis, and Lecturer at the Central College of Agriculture, Warsaw, and E. S. PEARSON, Department of Applied Statistics, University College, London.

(Communicated by K. PEARSON, F.R.S.)

(Received August 31, 1932.—Read November 10, 1932.)



nach B. Efron (2001)

Nichtparametrische (robuste) Verfahren

INDIVIDUAL COMPARISONS BY RANKING METHODS

FRANK WILCOXON
American Cyanamid Co.



Frank Wilcoxon (1892-1965)

The comparison of two treatments generally falls into one of the following two categories: (a) we may have a number of replications for each of the two treatments, which are unpaired, or (b) we may have a number of paired comparisons leading to a series of differences, some of which may be positive and some negative. The appropriate methods for testing the significance of the differences of the means in these two cases are described in most of the textbooks on statistical methods.

The object of the present paper is to indicate the possibility of using ranking methods, that is, methods in which scores 1, 2, 3, . . . n are substituted for the actual numerical data, in order to obtain a rapid approximate idea of the significance of the differences in experiments of this kind.

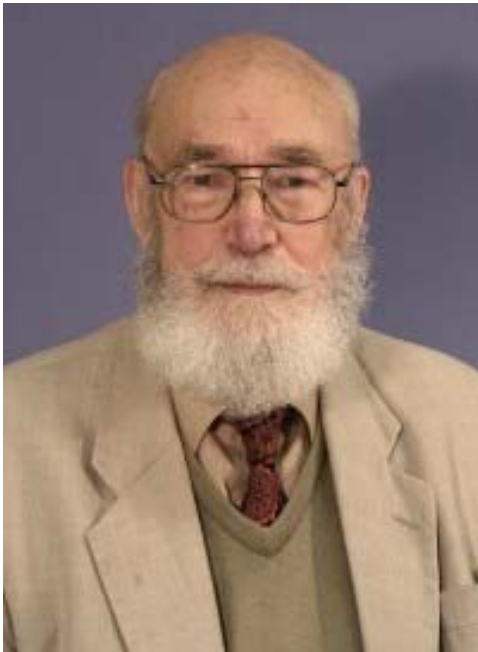
Explorative Datenanalyse ("numerical detective work")



John W. Tukey (1915-2000)

"In our opinion, it is more natural to begin with data analysis and end with models, for this is the order in which things usually arise in the real world, as opposed to the world of the mathematician. In scientific work, one almost always begins with the data, and when structures in the data are found to be statistically significant, a model or theory is sought to explain these patterns."

Verallgemeinerte lineare Modelle



John A. Nelder (1924-2010)

Generalized Linear Models

By J. A. NELDER and R. W. M. WEDDERBURN

Rothamsted Experimental Station, Harpenden, Herts

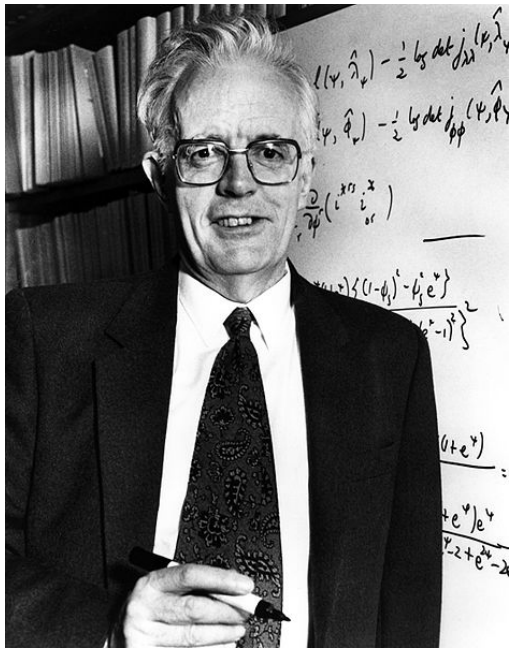
SUMMARY

The technique of iterative weighted linear regression can be used to obtain maximum likelihood estimates of the parameters with observations distributed according to some exponential family and systematic effects that can be made linear by a suitable transformation. A generalization of the analysis of variance is given for these models using log-likelihoods. These generalized linear models are illustrated by examples relating to four distributions; the Normal, Binomial (probit analysis, etc.), Poisson (contingency tables) and gamma (variance components).

The implications of the approach in designing statistics courses are discussed.

J. R. Statist. Soc. A,
(1972), **135**, Part 3, p. 370

Proportional Hazard Regression



Sir David Cox (*1924)

Regression Models and Life-Tables

BY D. R. COX

Imperial College, London

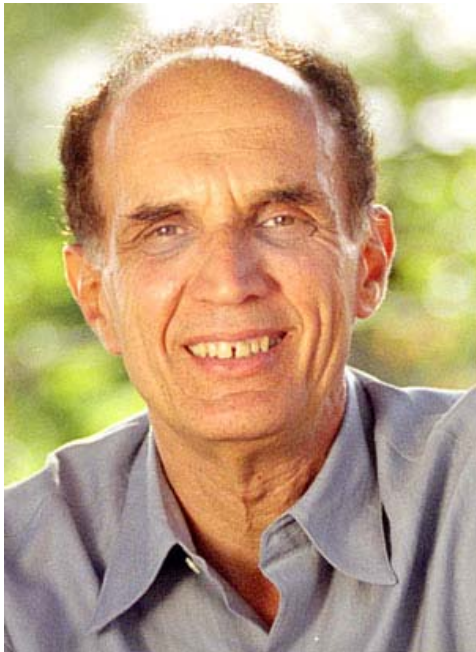
[Read before the ROYAL STATISTICAL SOCIETY, at a meeting organized by the Research Section, on Wednesday, March 8th, 1972, Mr M. J. R. HEALY in the Chair]

SUMMARY

The analysis of censored failure times is considered. It is assumed that on each individual are available values of one or more explanatory variables. The hazard function (age-specific failure rate) is taken to be a function of the explanatory variables and unknown regression coefficients multiplied by an arbitrary and unknown function of time. A conditional likelihood is obtained, leading to inferences about the unknown regression coefficients. Some generalizations are outlined.

Journal of the Royal Statistical Society.
Series B, Vol. 34, No. 2(1972), pp. 187-220

Resampling Methoden (Bootstrap)



Bradley Efron (*1938)

BOOTSTRAP METHODS: ANOTHER LOOK AT THE JACKKNIFE

BY B. EFRON

Stanford University

We discuss the following problem: given a random sample $\mathbf{X} = (X_1, X_2, \dots, X_n)$ from an unknown probability distribution F , estimate the sampling distribution of some prespecified random variable $R(\mathbf{X}, F)$, on the basis of the observed data \mathbf{x} . (Standard jackknife theory gives an approximate mean and variance in the case $R(\mathbf{X}, F) = \theta(\hat{F}) - \theta(F)$, θ some parameter of interest.) A general method, called the “bootstrap,” is introduced, and shown to work satisfactorily on a variety of estimation problems. The jackknife is shown to be a linear approximation method for the bootstrap. The exposition proceeds by a series of examples: variance of the sample median, error rates in a linear discriminant analysis, ratio estimation, estimating regression parameters, etc.

The Annals of Statistics, Vol. 7, No. 1 (1979), 1-26

Daten			Tests für				
			1 Stichprobe	2 Stichproben		mehr als 2 Stichproben	
				unabhängige	abhängige	unabhängige	abhängige
Häufigkeiten	diskret	folgen speziellen diskreten Verteilungen	Berechnungen anhand der: Poissonverteilung, Binomialverteilung, hypergeometrischen Verteilung χ^2 -Anpassungstest $2I$ -Anpassungstest	Fisher-Test χ^2 -Tests G-Test Kontingenzkoeffizient Folgetestpläne	Vorzeichen-tests	χ^2 -Tests $2I$ -Test Kontingenzkoeffizient	Q-Test
Rangzahlen	stetig	nicht normalverteilt	Iterationstests	Siegel-Tukey-Test Mosteller-Test Tukey-Test Median-Test U-Test	Maximumtest Wilcoxon-Test Quadrantenkorrelation Eckentest Spearman-Rangkorrelation	Erweiterte Mediantests H-Test	Friedman-Test Multiple Vergleiche nach WILCOXON und WILCOX
Meßwerte		angenähert normalverteilt, z.T. nach Transformation	Kolmogoroff-Smirnoff-Test Cox-Stuart-Trend-Test	Kolmogoroff-Smirnoff-Test	t -Tests Produktmomentkorrelation, Lineare Regression	Levene-Test Cochran-, Hartley- und Bartlett-Test Link-Wallace-Test Nemenyi-Vergleiche Varianzanalyse Student-Newman-Keuls-Test	Varianzanalyse Scheffé-Test Student-Newman-Keuls-Test Kovarianzanalyse, Partielle Korrelation Multiple Korrelation Multiple Regression Diskriminanzanalyse

Sachs, L. (1968): Statistische Methoden in der Medizin.

Stereotypen in der Beratung

- Den **Probabilisten** interessiert nur ein signifikanter P-Wert.
- Der **Datensammler** kommt in die Beratung mit Daten aus unzähligen Experimenten (EXCEL).
- Der **sporadische "Blutsauger"** nutzt die Gelegenheit einer beiläufigen, freundlichen Unterhaltung.
- Der **Amateurstatistiker** glaubt, dass jeder intelligente Mensch auch ein guter Statistiker sein kann.

Studienplanung

[Sir Ronald A. Fisher \(1938\)](#): "To consult the statistician after an experiment is finished is often merely to ask him to conduct a post mortem examination. He can perhaps say what the experiment died of."

[David A. Freedman \(1991\)](#): "Statistical technique can seldom be an adequate substitute for good design, relevant data, and testing predictions against reality in a variety of settings."

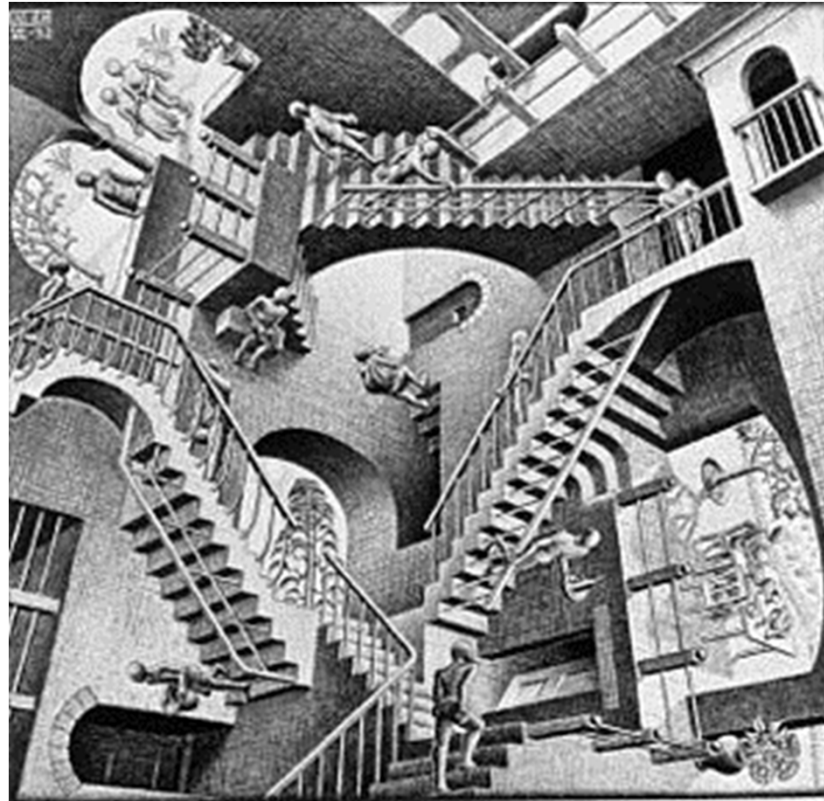
The Applied Statistician's Creed

Treatments	All treatments differ!
Factors	All factors interact!
Correlations	All variables are correlated!
Populations	No two populations are identical in any respect!
Normality	No data are normally distributed!
Variances	Variances are never equal!
Models	All models are wrong!
Equality	No two numbers are the same!
Size	Many Numbers are very small!

Ausblick

- Suchen nach den richtigen Fragen
- Exploration vs. hypothesenbasierter Forschung
- Erkennen der Variation (Messungen / Beobachtungen)
- Kontrolle der Variation
- Validieren von Modellen
- Bewerten der Unsicherheit bei Entscheidungen

Statistisches Denken in der Medizin



"Die Gewissheit ist eine Illusion" (Gerd Gigerenzer)



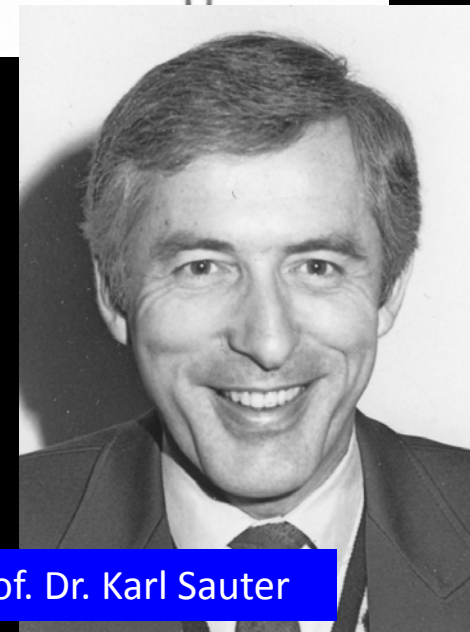
Prof. Dr. Gerd Griesser



Prof. Dr. Lothar Sachs



Abteilung für Medizinische
Statistik und Dokumentation



Prof. Dr. Karl Sauter