

# Januar 2023

Vor 60 Jahren starb

**RONALD A. FISHER** (17.02.1890 - 29.07.1962)

Ronald A. Fisher (1890 - 1962)



Mathematica

RONALD AYLMER FISHER gilt als der „Vater der modernen Statistik“; seine zahlreichen und vielfältigen Beiträge zur mathematischen Theorie wie auch zu statistischen Anwendungen, insbesondere zur Versuchsplanung, die er von den 1920er Jahren an bis zu seinem Tod veröffentlichte, gaben wesentliche Impulse zur Entwicklung dieser Wissenschaft.

RONALD AYLMER wächst als jüngstes Kind zusammen mit vier älteren Geschwistern im Londoner Stadtteil East Finchley auf. Der Vater verdient den Unterhalt als Auktionator und Kunsthändler. Seinen zweiten Vornamen verdankt RONALD AYLMER einem ungewöhnlichen Einfall seiner Mutter; diese

hatte den frühen Tod ihres dritten Kindes mit Namen ALAN darauf zurückgeführt, dass in dessen Vorname kein „y“ vorkommt (die ersten beiden Kinder hatten die Vornamen GEOFFREY und EVELYN).

Als der Junge 14 Jahre alt ist, stirbt die Mutter, und kurze Zeit danach verliert der Vater seine Arbeit. Jedoch dank eines Stipendiums kann RONALD AYLMER ein Studium der Mathematik und Astronomie am *Caius and Gonville College* in Cambridge aufnehmen. Er interessiert sich auch für biologische Themen, u. a. regt er die Gründung einer *Eugenics Society* an der Universität an.

1912 schließt FISHER seine Prüfung mit Auszeichnung ab; nach der Meinung seines Tutors hätte das Ergebnis noch besser ausfallen können, wenn er sich angestrengt hätte. Eine Zeitlang setzt er sein Studium fort; insbesondere fasziniert ihn eine Schrift zur Fehlertheorie von GEORGE BIDDELL AIRY (1801-1892), dem langjährigen Leiter des Observatoriums in Greenwich.

FISHERS erste Veröffentlichung erfolgt noch im selben Jahr; sie befasst sich mit der sog. *Maximum-Likelihood-Methode* zur optimalen Schätzung eines unbekannten Parameters, wenn der Typ der zugrundeliegenden Verteilungsfunktion bekannt ist.



MO	DI	MI	DO	FR	SA	SO
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31					

Beispiel für die Anwendung der Maximum-Likelihood-Methode:

In einer Urne befinden sich schwarze und weiße Kugeln, deren Anteil durch Ziehen einer gewissen Anzahl von Kugeln geschätzt werden soll. Wenn beispielsweise in einer Stichprobe von zehn Kugeln (Ziehen mit Zurücklegen) zwei schwarze Kugeln enthalten sind, dann erweist sich ein Anteil von  $p = 0,2$  schwarzen Kugeln in der Urne als der mit

der größten Wahrscheinlichkeit:  $P_{p=0,19}(X = 2) = \binom{10}{2} \cdot 0,19^2 \cdot 0,81^8 \approx 0,3010$ ;

$P_{p=0,2}(X = 2) = \binom{10}{2} \cdot 0,2^2 \cdot 0,8^8 \approx 0,3020$ ;  $P_{p=0,21}(X = 2) = \binom{10}{2} \cdot 0,21^2 \cdot 0,79^8 \approx 0,3011$ .

Um Geld zu verdienen, arbeitet FISHER einige Monate auf einer Farm in Kanada, danach als Statistiker in einer Londoner Investment-Firma. Als 1914 der Weltkrieg ausbricht, meldet er sich begeistert zum Dienst in der Armee, wird aber wegen einer angeborenen Sehschwäche ausgemustert. Danach unterrichtet FISHER eine Zeitlang Mathematik und Physik an verschiedenen Schulen, spielt aber auch mit dem Gedanken, eine eigene Farm zu kaufen und diese zu bewirtschaften.

1917 heiratet er heimlich RUTH EILEEN GRATTON GUINNESS, die wenige Tage zuvor 17 Jahre alt geworden ist; in der lange Zeit glücklichen Ehe werden neun Kinder geboren.



Im Jahr 1918 macht FISHER mit der Veröffentlichung des Beitrags *The Correlation Between Relatives on the Supposition of MENDELian Inheritance* (Korrelation zwischen Verwandten unter der Annahme der MENDEL'schen Vererbung) auf sich aufmerksam; u. a. zeigt er durch eine Modellrechnung, dass sich die Allelhäufigkeiten in einer

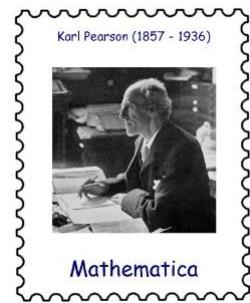
Population durch natürliche Auslese verändern können.

Infolge bietet KARL PEARSON ihm die Stelle als Chefstatistiker an den GALTON Laboratories an, was FISHER aber aufgrund von (vermeintlich) schlechten Erfahrungen mit PEARSON ablehnt.

Stattdessen nimmt er eine befristete Stelle an der Rothamsted Experimental Station in Hertfordshire an - er wird 14 Jahre lang dort bleiben. Ausschlaggebend für seine Entscheidung ist vor allem die Tatsache, dass dort eine gewaltige Menge von Erntedaten zur Verfügung steht, die seit 1842 im Rahmen von *Classical Field Experiments* gesammelt wurden.

Bereits 1921 veröffentlicht FISHER die *Studies in Crop Variation* mit einer ersten Anwendung der von ihm entwickelten Varianzanalyse (Analysis of Variance = ANOVA), ein Prüfverfahren, bei dem untersucht wird, ob die Grundgesamtheiten von zwei durchgeführten Stichproben die gleiche Varianz besitzen.

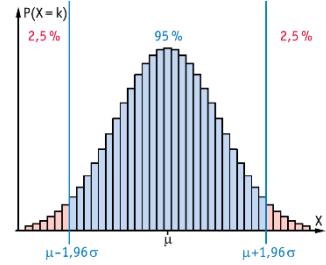
FISHER nutzt die nahezu unbegrenzten Möglichkeiten in Rothamsted, um Erfahrungen im Hinblick auf die Versuchsplanung zu gewinnen. Ziel seiner Experimente ist es, mit möglichst wenigen Versuchen den Einfluss möglichst vieler Faktoren herauszufinden. Im Prinzip können Experimente nach dem Prinzip *Versuch und Irrtum* (trial and error) durchgeführt werden, oder indem man einzelne Parameter abwechselnd und schrittweise verändert (*one factor at a time*), was allerdings sehr aufwendig ist und sehr kostspielig sein kann. Seine Erfahrungen veröffentlicht er zusammenfassend unter dem Titel *The design of experiments* (erst im Jahr 1935).



1924 folgt sein Beitrag *On a distribution yielding the error functions of several well known statistics*, in der er die (von ihm so bezeichnete)  $z$ -Verteilung einführt, die in Sonderfällen mit der STUDENT'schen  $t$ -Verteilung und der PEARSON'schen  $\chi^2$ -Verteilung übereinstimmt.

Der Statistiker WILLIAM SEALY GOSSET hatte - da sein Arbeitgeber, die Guinness-Brauerei in Dublin, seinen Angestellten keine eigenen Publikationen gestattete - im Jahr 1908 unter dem Pseudonym STUDENT einen Beitrag veröffentlicht, in dem er zeigte, dass die standardisierte Schätzfunktion des Stichproben-Mittelwerts normalverteilter Daten selbst nicht normalverteilt, sondern  $t$ -verteilt ist, wenn die Varianz nicht bekannt, sondern durch die Stichprobenvarianz geschätzt werden muss.

1925 veröffentlicht FISHER das Werk *Statistical Methods for Research Workers*, eines der wohl einflussreichsten Statistikbücher des 20. Jahrhunderts. Unter anderem führt er in diesem Buch den  $p$ -Wert von  $p = 0,05$  als „Grenzwert“ für statistische Signifikanz ein. Bei einem zweiseitigen Test ist dies die (Maximal-)Wahrscheinlichkeit dafür, dass ein Versuchsergebnis zufällig außerhalb der  $1,96\sigma$ -Umgebung des Erwartungswerts  $\mu$  liegt.



In dem o. a. 1935 veröffentlichten Werk zur Versuchsplanung verwendet FISHER dann die Geschichte der tea testing lady, um die Vorgehensweise eines randomisierten Experiments und die Strategie eines sog. Hypothesentests zu erläutern:

Die betreffende Dame behauptet, sie könne am Geschmack erkennen, ob zuerst der Tee oder zuerst die Milch in die Tasse gegeben wurde. FISHERS Versuchsanleitung sieht vor, die angebliche Fähigkeit der Dame dadurch zu überprüfen, dass ihr in zufälliger Reihenfolge hintereinander acht Tassen Tee serviert werden, je vier von jeder Sorte. Die Wahrscheinlichkeit, mit der eine gewisse Anzahl von Tassen einer Sorte zufällig erraten wird, kann mithilfe eines hypergeometrischen Ansatzes berechnet werden:

Es gibt  $\binom{8}{4} = 70$  Möglichkeiten, die vier Tassen einer Sorte auf die acht Plätze zu verteilen; davon gibt es genau  $\binom{4}{0} \cdot \binom{4}{4} = 1$  Möglichkeit dafür, dass keiner bzw. dass jeder der Rateversuche richtig ist,  $\binom{4}{1} \cdot \binom{4}{3} = 16$  Möglichkeiten für einen bzw. für drei Treffer und  $\binom{4}{2} \cdot \binom{4}{2} = 36$  Möglichkeiten für zwei Treffer. Bei einem Signifikanzniveau von 5 % kann die Nullhypothese Die Dame verfügt nicht über die von ihr behauptete Fähigkeit nur dann verworfen werden, wenn sie alle Tassen richtig zuordnet (denn die Wahrscheinlichkeit für das zufällige 8-fach richtige Raten beträgt  $1/70 \approx 1,4\%$ ).

Mit Hypothesentests von kleinen Stichproben beschäftigt sich auch der sog. Exakte Test von FISHER, bei dem Daten in Vierfeldertafeln (Kontingenztafeln) untersucht werden. Dabei werden die Randwerte als gegeben angenommen, und es wird die Wahrscheinlichkeit dafür bestimmt, dass die beobachtete oder eine noch extremere Besetzung der Tafel zufällig zustandekommt.

Beispiel: Vergleich des Behandlungsergebnisses durch zwei Medikamente:

Die Wahrscheinlichkeit, dass unter den insgesamt 20 Patienten zufällig  $k$  Patienten sind, die durch Anwendung eines neuen Medikaments geheilt werden, beträgt  $\binom{9}{k} \cdot \binom{11}{10-k} / \binom{20}{10}$ . Der Verwerfungsbereich der Nullhypothese

Das neue Medikament bietet keine höheren Heilungschancen als das bisher verwendete Medikament ist für ein Signifikanzniveau von 5 % durch  $k = 7, 8, 9$  bestimmt. Die Hypothese kann also aufgrund des Versuchsergebnisses nicht verworfen werden.

	erfolgreich	nicht erfolgreich	gesamt
bisher verwendetes Medikament	4	7	11
neues Medikament	6	3	9
gesamt	10	10	20

In seinem Werk *The Genetical Theory of Natural Selection*, das 1930 erscheint, geht FISHER auf die Frage ein, warum in den „zivilisierten Gesellschaften“ eher die schwachen als die starken Individuen gefördert werden: Das natürliche Prinzip *survival of the fittest* würde seiner Meinung nach künstlich zugunsten der Schwachen der Gesellschaft verändert; eher müssten die „gesunden Mitglieder der Gesellschaft“ finanziell unterstützt werden. Er geht sogar so weit, dass er den Niedergang von Zivilisationen in der Menschheitsgeschichte mit der geringen Fertilität der oberen Gesellschaftsschichten in Verbindung bringt. – Insofern ist es auch konsequent, dass der bekennende Eugeniker FISHER 1933 als Nachfolger PEARSONS als Leiter des *Department of Eugenics* am University College London ernannt wird.

Mit KARL PEARSON hatte er zwischenzeitlich heftige fachliche Auseinandersetzungen, u. a. hinsichtlich der Anwendung des  $\chi^2$ -Unabhängigkeitstests in Kontingenztafeln, wobei beide stets auch ihre gegenseitige persönliche Abneigung nicht verbargen.

Seine Ansichten über Eugenik ändert Fisher auch nach den Erfahrungen mit den Verbrechen des Nationalsozialismus nicht; vielmehr trägt sein Entlastungsschreiben mit dazu bei, dass OTMAR FREIHERR VON VERSCHUER, Leiter des Kaiser-Wilhelm-Instituts für Anthropologie, menschliche Erblehre und Eugenik in Berlin, nach dem Krieg nur als „Mitläufer“ eingestuft wird und seine Karriere als Professor für Humangenetik an der Universität Münster fortsetzen kann, obwohl er während des Kriegs engen wissenschaftlichen Kontakt zu seinem ehemaligen Doktoranden JOSEF MENGELE hatte, dem KZ-Arzt von Auschwitz.

Von 1943 bis 1956 übernimmt FISHER in Cambridge einen Lehrstuhl für Genetik; nach seiner Emeritierung emigriert der fachlich sehr angesehene und vielfach geehrte Wissenschaftler (darunter auch die Erhebung in den Adelsstand durch ELISABETH II) nach Adelaide (Australien), um dort seine Studien fortzusetzen.

Erwähnt werden sollte auch noch eine Episode aus dem Jahr 1950: Als eine Studie über den möglichen Zusammenhang zwischen Rauchen und Lungenkrebs-Erkrankungen veröffentlicht wird, widerspricht FISHER, da seiner Meinung nach die aufgezeigten Korrelationen nicht unbedingt auf kausale Bedingungen hinweisen. Es ist nicht auszuschließen, dass FISHER für seine Stellungnahme Geld von der Tabakindustrie erhalten hat. Es könnte aber auch sein, dass der leidenschaftliche Pfeifenraucher FISHER einfach nur deshalb widersprochen hat, weil er eine mögliche Einschränkung seiner Gewohnheiten durch die „Puritaner“ befürchtete.