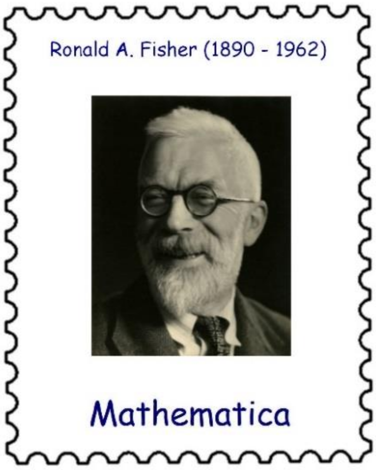
RONALD AYLMER FISHER (February 17, 1890 – July 29, 1962)

by HEINZ KLAUS STRICK, Germany



RONALD AYLMER FISHER is considered the "father of modern statistics". His numerous and varied contributions to mathematical theory as well as to statistical applications, especially to experimental design, which he published from the 1920s until his death, provided essential impulses for the development of this science.

RONALD AYLMER grew up as the youngest child together with four older siblings in the London district of East Finchley. His father earned a living as an auctioneer and art dealer.

RONALD AYLMER owed his middle name to an unusual idea of his mother's; she had attributed the early death of her third child, named ALAN, to the fact that there was no "Y" in his first name (the first two children had the first names GEOFFREY and EVELYN).

When the boy was 14 years old, his mother died and shortly afterwards his father lost his job. However, thanks to a scholarship, RONALD AYLMER was able to study mathematics and astronomy at *Caius and Gonville College* in Cambridge. He was also interested in biological topics, including the founding of a *Eugenics Society* at the university.

In 1912, FISHER passed his exams with distinction; according to his tutor, the result could have been even better if he had tried harder. He continued his studies for a while; in particular, he was fascinated by a paper on error theory by GEORGE BIDDELL AIRY (1801-1892), the long-time director of the *Greenwich Observatory*.

FISHER's first publication was in the same year; it dealt with the so-called *Maximum Likelihood Method* for the optimal estimation of an unknown parameter when the type of the underlying distribution function was known.

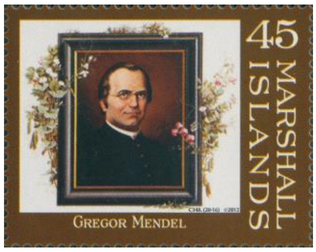
Example of the application of the *Maximum Likelihood Method*:

In an urn there are black and white balls, the proportion of which is to be estimated by drawing a certain number of balls. If, for example, there are two black balls in a sample of ten balls (drawn with replacement), then a proportion of p = 0.2 black balls in the urn turns out to be the one with the greatest probability: $P_{p=0.19}(X=2) = \binom{10}{2} \cdot 0.19^2 \cdot 0.81^8 \approx 0.3010$;

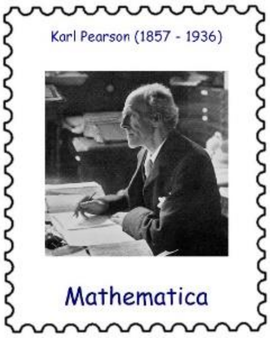
$P(X=2) = \binom{10}{2} \cdot 0.2 \cdot 0.8 \approx 0.3020$; $P_{p=0.21}(X=2) = \binom{10}{2} \cdot 0.21^2 \cdot 0.79^8 \approx 0.3011$.

To earn money, FISHER worked for a few months on a farm in Canada and then as a statistician in a London investment firm. When the World War broke out in 1914, he enthusiastically signed up for service in the army, but was invalided out because of a congenital visual impairment. Afterwards, FISHER taught mathematics and physics at various schools for a while, but also toyed with the idea of buying his own farm and cultivating it.

In 1917 he secretly married RUTH EILEEN GRATTON GUINNESS, who had turned 17 a few days earlier. Nine children were born in the long happy marriage.



In 1918, FISHER attracted attention with the publication of the paper *The Correlation Between Relatives on the Supposition of MENDELian Inheritance*. Among other things, he showed by means of a model calculation that the *allele* or *gene* frequencies in a population can change through natural selection.



As a result, KARL PEARSON offered him the position of chief statistician at GALTON Laboratories, which FISHER turned down due to (supposedly) bad experiences with PEARSON.

Instead, he accepted a temporary position at *Rothamsted Experimental Station* in Hertfordshire where he would stay for 14 years. The decisive factor for his decision was the fact that a huge amount of crop data, which has been collected since 1842 within the framework of *Classical Field Experiments*, was available there.

As early as 1921, FISHER published *Studies in Crop Variation* with a first application of the *Analysis of Variance* (ANOVA) developed by him, a test procedure which examines whether the populations of two samples carried out have the same variance.

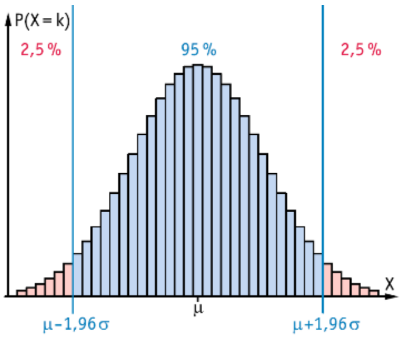
FISHER used the almost unlimited possibilities at Rothamsted to gain experience with regard to experimental design.

The aim of his experiments was to find out the influence of as many factors as possible with as few experiments as possible. In principle, experiments could be carried out according to the principle of *trial and error*, or by changing individual parameters alternately and gradually (*one factor at a time*), which, however, was very time-consuming and could be very costly. He published a summary of his experiences under the title *The design of experiments* (though not until 1935).

In 1924 he published his paper *On a distribution yielding the error functions of several well known statistics*, in which he introduced the *z*-distribution (as he called it), which in special cases corresponds to the so called STUDENT's *t*-distribution and to PEARSON's c^2-distribution.

The statistician WILLIAM SEALY GOSSET had published a paper in 1908 under the pseudonym STUDENT, since his employer, the Guinness Brewery in Dublin, did not allow his employees to publish their own papers. In this he showed that the standardised estimator of the sample mean of normally distributed data is not itself normally distributed but *t*-distributed if the variance is not known but must be estimated by the sample variance.

In 1925, FISHER published *Statistical Methods for Research Workers*, probably one of the most influential statistics books of the 20th century.



Among other things, in this book he introduced the *p*-value of $p = 0.05$ as a "limit" for statistical significance.
In a two-sided test, this is the (maximum) probability that an experimental result lies outside the 1.96σ environment of the expected value μ by chance.

In the above-mentioned work on experimental design published in 1935, FISHER then used the story of the *tea testing lady* to explain the procedure of a randomised experiment and the strategy of a so-called hypothesis test:

The lady in question claims that she can tell from the taste whether the tea or the milk was put into the cup first. FISHER's instructions for the experiment are to test the lady's alleged ability by serving her eight cups of tea, four of each variety, in random order.

The probability of randomly guessing a certain number of cups of a variety can be calculated using a hypergeometric approach:

There are $\binom{8}{4} = 70$ ways of distributing the four cups of one kind among the eight places;

there is exactly $\binom{4}{0} \cdot \binom{4}{4} = 1$ possibility for none or all of the guesses to be correct;

$\binom{4}{1} \cdot \binom{4}{3} = 16$ possibilities for one or three hits and $\binom{4}{2} \cdot \binom{4}{2} = 36$ possibilities for two hits.

At a significance level of 5%, the null hypothesis *"The lady does not have the ability she claims"* can only be rejected if she correctly assigns all cups (because the probability of randomly guessing 8 times correctly is $\frac{1}{70} \approx 1.4\,\%$).

FISHER's so-called *exact test* is also concerned with hypothesis tests of small samples, in which data is examined in four-field tables (contingency tables). The marginal values are taken as given and the probability is determined that the observed or an even more extreme entry of the table occurs by chance.

Example: Comparison of the results of treatment by two drugs: The probability that among the total of 20 patients there are by chance *k* patients who are cured by the application of a new drug is $\binom{9}{k} \cdot \binom{11}{10-k} \Big/ \binom{20}{10}$.

	successful	unsuccessful	total
old drug	4	7	11
new drug	6	3	9
total	10	10	20

The *critical region* of the null hypothesis *The new drug does not offer a higher chance of cure than the previously used drug* is determined for a significance level of 5 % by $k = 7, 8, 9$.

The hypothesis can therefore not be rejected on the basis of the experimental result.

In his work *The Genetical Theory of Natural Selection*, published in 1930, FISHER addressed the question of why weak rather than strong individuals are promoted in "civilised societies". In his opinion, the natural principle of survival of the fittest ought to be artificially altered in favour of the weak in society; rather, the "healthy members of society" should be financially supported.

He even goes so far as to link the decline of civilisations in human history to the low fertility of the upper strata of society. In this respect, it is consistent that the avowed eugenicist FISHER was appointed in 1933 as PEARSON's successor as head of the *Department of Eugenics* at University College London.

In the meantime, he and KARL PEARSON had heated professional disputes, among other things concerning the application of the *χ^2-independence test* in contingency tables, whereby both always did not hide their mutual personal dislike.

Fisher did not change his views on eugenics even after his experiences with the crimes of National Socialism; on the contrary, his letter of exoneration contributed to the fact that OTMAR FREIHERR VON VERSCHUER, director of the *Kaiser Wilhelm Institute for Anthropology*, *Human Heredity and Eugenics* in Berlin, was only classified as a "fellow traveller" after the war and could continue his career as professor of human genetics at the University of Münster, although he had close scientific contact with his former doctoral student JOSEF MENGELE, the concentration camp doctor at Auschwitz, during the war.

From 1943 to 1956, FISHER held a chair in genetics at Cambridge and after his retirement, the scientist, who was highly respected and honoured many times (including being raised to the peerage by ELIZABETH II), emigrated to Adelaide (Australia) to continue his studies there.

An episode from 1950 should also be mentioned:

When a study on the possible connection between smoking and lung cancer diseases was published, FISHER objected, since in his opinion the correlations shown did not necessarily point to causal conditions. It cannot be ruled out that FISHER received money from the tobacco industry for his opinion. But it could also be that as the passionate pipe smoker FISHER objected simply because he feared a possible restriction of his habits by the "puritans".

First published 2023 by Spektrum der Wissenschaft Verlagsgesellschaft Heidelberg

https://www.spektrum.de/wissen/ronald-aylmer-fisher-der-vater-der-statistik/2091714

Translated 2023 by John O’Connor, University of St Andrews
