Who was Bonferroni?

Michael E Dewey

Trent Institute for Health Services Research
University of Nottingham

mailto:michael.dewey@nottingham.ac.uk
http://www.nottingham.ac.uk/~mhzmd/bonf.html

Applied statisticians are familiar with the name of Bonferroni in the field of simultaneous statistical inference where it is given to a method relying on a set of inequalities named after him. Of all the eponymous figures in modern statistical practice he remains perhaps the least well known, certainly as far as Anglo-Saxon audiences are concerned. In this talk I shall outline briefly his life and work, and give a personal account of my search for details about him.

======Slide 1======

Typical structure of a local group RSS talk

- Introduction to the problem (3 overheads)
- Some notation (2 overheads)
- Some mathematics (26 overheads)
- Return to the problem (2 overheads)
- Triumphant dénoument (1 overhead)

======Slide 2======

Structure of this talk

- Bonferroni adjustment and the inequalities
- My interest and the early search
- His life
- His other work
- The context

======Slide 3======

The method of adjustment named after Bonferroni is based firmly within the classical (frequentist) tradition in applied statistics.

In simultaneous inference if we are to make k tests and desire an overall test of size α we make each test at $\frac{\alpha}{k}$ or more generally such that $\sum_i \alpha_i = \alpha$. In fact nobody ever uses the more general form.

There is an analogue for confidence intervals which nobody uses either.

This procedure has become known as the Bonferroni method or Bonferroni adjustment.

======Slide 4======

There are of course other more recent methods which improve on the classical Bonferroni procedure. In Holm's method (Holm, 1979) we order the p_i

$$p_1 \leq p_2 \leq \cdots \leq p_n$$

We compare each p_i to $\frac{\alpha}{n-i+1}$ starting from p_1. At each stage if we reject (because $p_i \leq \frac{\alpha}{n-i+1}$) then we proceed to the next stage whereas once we fail to reject we stop without examining further.

This retains the conservative property of the Bonferroni method while being always at worst as powerful and usually better. By the end of 2000 this paper had received about 1500 citations.

======Slide 5======

The earliest reference to the use of Bonferroni's inequalities in simultaneous statistical inference appears to be a paper by Paulson (1952) This considered the problem of selecting the best of k categories when comparing $k-1$ experimental with one control.

> Making use of Bonferroni's Inequality ...(he cites Feller)

The modern usage seems to come from two papers by Dunn. In Dunn (1959) she considers confidence intervals for k means (and mentions 'a Bonferroni inequality') in Dunn (1961) she considers m contrasts among k means.

> The method given here is so simple and so general that I am sure it must have been used before this. I do not find it, however, so can only conclude that perhaps its very simplicity has kept statisticians from realizing that it is a very good method in some situations.

======Slide 6======

In the 1936 article we have

2 simultaneous probabilities In a set composed of m objects we consider n characteristics C_1 C_2 C_3 $\ldots C_n$. We denote by m_i the number having characteristic C_i, and by m_{ij} those having both C_i and C_j (and possibly others) Then

$$p_i = \frac{m_i}{m}, \quad p_{ij} = \frac{m_{ij}}{m}, \quad p_{ijk} = \frac{m_{ijk}}{m}, \ldots$$

3 contrary probabilities

$$S_0 = 1, \quad S_1 = \sum p_i, \quad S_2 = \sum p_{ij}, \quad S_3 = \sum p_{ijh}$$

7 r-exact multiplicity We shall call the multiplicity of an object the number of characteristics which it possesses. We shall determine the probability $p_{1\ldots n(r)}$ that an object has multiplicity r.

$$P_r = p_{12\ldots n(r)} \ ,$$

======Slide 7======

9 r-minimum multiplicity

$$P_{(r)} = p_{(12\ldots n)r}$$

Exactly r and at least r

29 limits on simultaneous probabilities

$$P_o \leq 1 \ , \ P_o \geq 1 - S_1 \ , \ P_o \leq 1 - S_1 + S_2 \ , \\ P_o \geq 1 - S_1 + S_2 - S_3 \ , \text{ etc} \qquad [27]$$

[27] are of course the inequalities of Bonferroni.

They are referred to by Fréchet (1940) as his formula (212) and this is the formula given by Feller (1968) (as 5.7 in chapter IV) .

Section 1 of this paper closes a few sections later having taken 20 pages to get this far.

======Slide 8======

From the 1935 article it is perhaps necessary only to quote the abstract

> The author establishes above all a symbolic calculus which enables the expression in a rapid and uniform manner of the various probabilities of survival and death amongst a group of assured, expressed as a function of a particular type assumed as primary. This calculus does not require the hypothesis that the assured lives should be independent, as is usual in treatments of this problem. He establishes a noteworthy law of duality between the probability of survival and that of death, introducing as a consequence of the notation some new results.(Bonferroni, 1935)

======Slide 9======

My interest arose from a question in consultancy. I realised that I had no idea who Bonferroni was; whether s/he is/was male female, or even a place.

Reference works (Encyclopaedia of Statistics) and SCI (citation index) provide limited information

Even in 1990 electronic methods helped - allstat provided reference to an obituary by Pagni (1960)

Eventually tracked the two inequalities articles (Bonferroni, 1935, 1936)

The citations are usually wrong - probably secondary citation from Fréchet. The obvious implication is that nobody has read the articles.

======Slide 10======

Further search made difficult by poor standards of citation of references in the obituaries.

Personal contacts

- Grigoletto - more articles
- Seneta - Benedetti article

Web search for journals - Metron held at UCL library (in the now disbanded RSS collection) which also holds Bollettino UMI and Bollettino di Storia delle Scienze Matematiche an article in which (Barbieri and Pepe, 1992) on the history of mathematics in Italian contains 3053 articles of which only 2 relate to him - Pagni obituary, and a memoir by de Finetti (1964)

Genoa library held his books

How to disseminate this?

Web pages with improved bibliography

Grandson - son - more details about life

======Slide 11======

Bonferroni, Carlo Emilio

Born: 28 January 1892, Bergamo

Died: 18 August 1960, Firenze

Studied conducting and piano at the Conservatory in Torino

Studied for the degree of *laurea* in Torino under Peano and Segre, spent time abroad (Wien and Zürich) and then became *incaricato* (assistant professor) at the Polytechnic in Torino. He served in the 1914-18 war in the engineers.

In 1923 took up the chair of financial mathematics at the Economics Institute in Bari

In 1933 he transferred to Firenze where he held his chair until his death.

======Slide 12======

The obituary of him by Pagni (1960) lists his works under three main headings: actuarial mathematics (16 articles, 1 book); probability and statistical mathematics (30, 1); analysis, geometry and rational mechanics (13, 0).

His interests in statistics include the study of various types of mean, and of correlation. He also developed a concentration index.

======Slide 13======

Bonferroni's concentration index is a measure of income inequality. Let $x_{(i)}$ be the observed ith order statistic in a sample size n, so that $x_{(i-1)} \leq x_{(i)} \quad (i = 2, \ldots, n)$. Define m_i the sample partial means, and $m = m_n$ the sample mean.

$$m_i = \frac{1}{i} \sum_{j=1}^{i} x_{(j)}$$

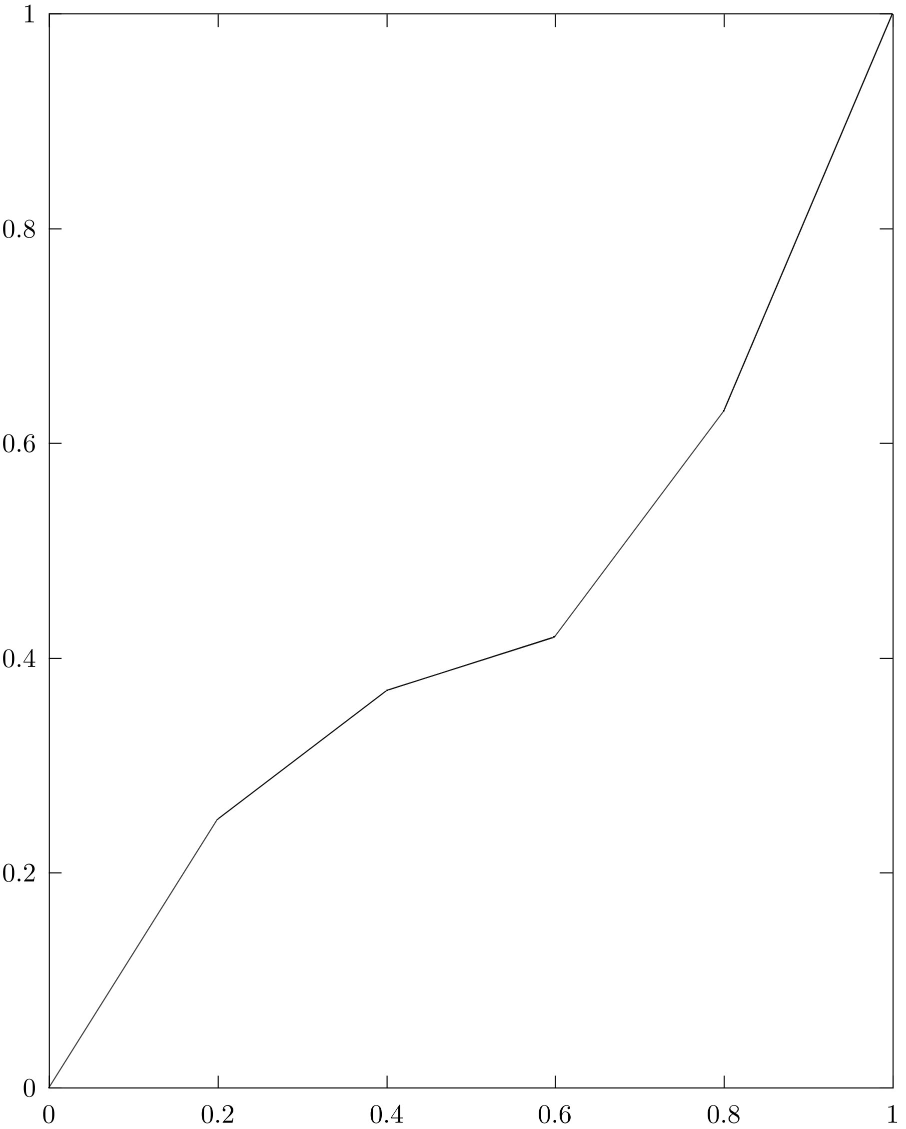
Then Bonferroni's index B_n is

$$B_n = 1 - \frac{1}{n-1} \sum_{i=1}^{n-1} \frac{m_i}{m}$$

This is one of a number of suggested measures of income inequality (Giorgi, 1998). Its minimum value 0 is attained for equidistribution of income and its maximum 1 when one person has all the income. It is more sensitive then other measures to changes in distribution amongst the poor.

======Slide 14======

x_i	m_i	$\frac{m_i}{m}$	$\frac{i}{n}$
10	10.0	0.25	0.2
20	15.0	0.37	0.4
20	16.7	0.42	0.6
50	25.0	0.63	0.8
100	40.0	1.00	1.0



In this example $B = 0.58$

In his textbook he discusses not just the usual arithmetic mean and the slightly less usual geometric mean and harmonic mean, but also exponential means, algebraic means, and weighted forms of all of them.

Arithmetic

$$\frac{\sum_i x_i}{n}$$

Geometric

$$\sqrt[n]{x_1 x_2 \dots x_n}$$

Quadratic

$$\sqrt{\frac{\sum_i x^2}{n}}$$

Harmonic

$$\frac{n}{\frac{1}{x_1} + \dots + \frac{1}{x_n}}$$

======Slide 15======

Algebraic means

$$\sqrt[k]{\frac{\sum_i x_i^k}{n}}$$

k	
-1	Harmonic
$\lim_{k \to 0}$	Geometric
0.5	Root mean
1	Arithmetic
2	Quadratic

======Slide 16======

Paper on algebraic means (Bonferroni, 1950)

The usual algebraic mean of order or degree p of $x_1 \dots x_n$ (positive), defined as

$$M_p = \sqrt[p]{\frac{S_p}{n}}, \quad S_p = x_1^p + \dots + x_n^p$$

can be generalised to double means, or of multiplicity 2, of partial degree p, q and total $p+q$, defined as

$$M_{p,q} = \sqrt[p+q]{\frac{S_{p,q}}{n(n-1)}}, \quad S_{p,q} = x_1^p x_2^q + x_2^p x_1^q + \dots$$

to triple or of multiplicity 3

$$M_{p,q,r} = \sqrt[p+q+r]{\frac{S_{p,q,r}}{n(n-1)(n-2)}}, \quad S_{p,q,r} = x_1^p x_2^q x_3^r + x_2^p x_1^q x_3^r + \dots$$

and so on for higher multiplicity

======Slide 17======

Exponential mean

For some base a mean η ($a > 0$ and $a \neq 1$)

$$a^\eta = \frac{\sum_i a^{x_i}}{n}$$

What application does this have?

Suppose we have x_i as a sample of positive times for which equal sums of money accumulate at compound interest. Then η is the time it takes to accumulate an average sum of money

======Slide 18======

The 1936 article uses classical probability (sample space of equally likely events - often attributed to Laplace). Was this really his view?

His view on probability may be illustrated with a quote from an address given to mark the academic year 1925-26 (Bonferroni, 1927)

> A weight is determined directly by a balance. And a probability, how is that determined? What is, so to say, the probability balance? It is the study of frequencies which gives rise to a specific probability.

He explicitly denies that subjective beliefs can form the basis for mathematical analysis.

======Slide 19======

Some conclusions

Why is he so little known?

Most works in Italian (but his reaction to the fame of his inequalities was that only foreigners took notice of him)

Books difficult to access

Influence of Gini

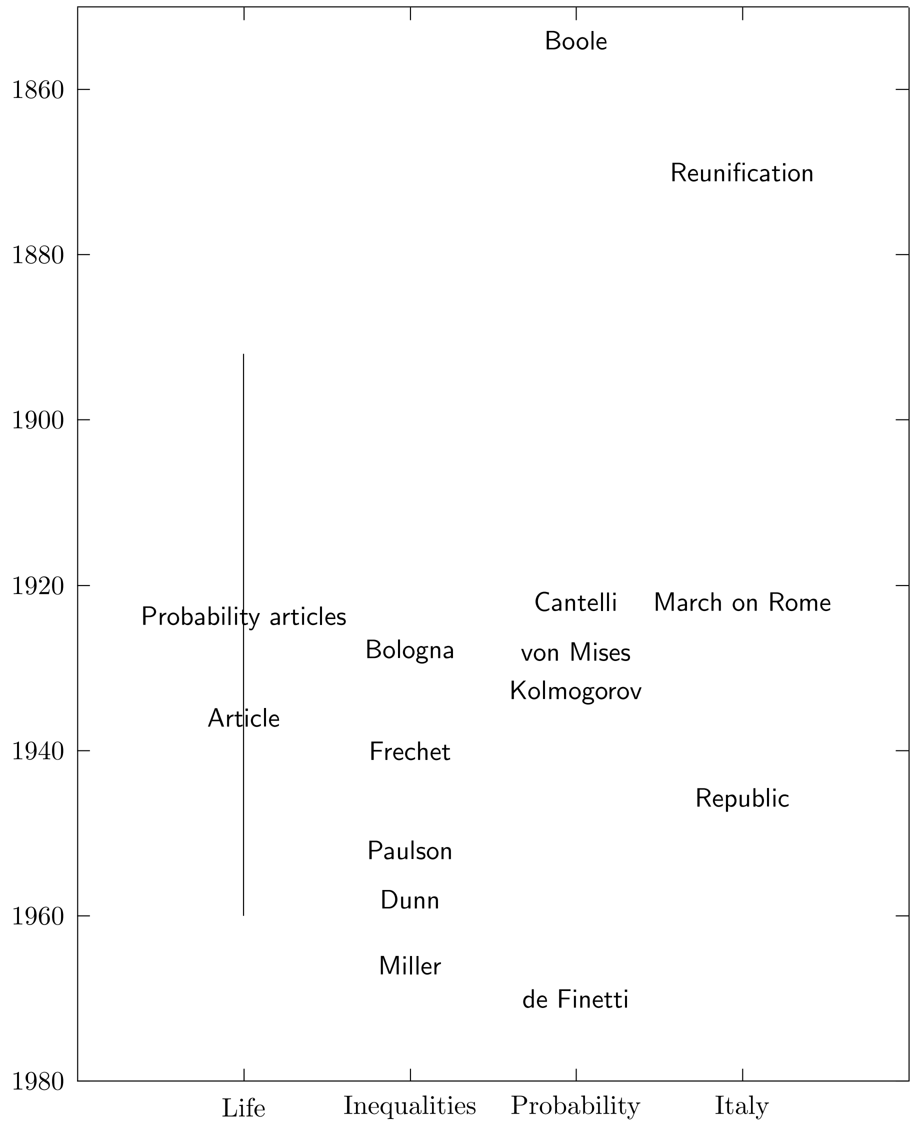
Bayesians think it is all a Bad Thing

The Bonferroni method is, of course, an example of Stigler's Law of Eponymy (Stigler, 1980) as we are really using only Boole's inequality.

$$P_o \geq 1 - S_1$$

Not a key figure in probability theory

======Slide 20======



Bologna refers to the 8th International Mathematical Congress; Cantelli to the appointment of Cantelli as the first university teacher of the calculus of probability and its applications; von Mises, Kolmogorov, Fréchet, Miller and de Finetti to the publication of books

References

F Barbieri and L Pepe. Bibliographia italiana di storia delle matematiche 1961-1990. *Bollettino di storia delle scienze matematiche*, XII(1):3–181, 1992.

C E Bonferroni. Teoria e probabilità. In *Annuario del R Istituto Superiore di Scienze Economiche e Commerciali di Bari per L'anno Accademico 1925-1926*, pages 15–46. Bari, 1927. Given as the inaugural lecture for the academic year 1925-1926 on November 22, 1925.

C E Bonferroni. Il calcolo delle assicurazioni su gruppi di teste. In *Studi in Onore del Professore Salvatore Ortu Carboni*, pages 13–60. Rome, 1935.

C E Bonferroni. Teoria statistica delle classi e calcolo delle probabilità. *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze*, 8:3–62, 1936.

C E Bonferroni. Sulle medie multiple di potenze. *Bollettino dell'Unione Matematica Italiana*, 5 third series: 267–270, 1950.

B de Finetti. Commemorazione del Prof. C. E. Bonferroni. *Giornale di Matematica Finanziaria*, 46:5–24, 1964.

O J Dunn. Confidence intervals for the means of dependent, normally distributed variables. *Journal of the American Statistical Association*, 54:613–621, 1959.

O J Dunn. Multiple comparisons among means. *Journal of the American Statistical Association*, 56:52–64, 1961.

W Feller. *An introduction to probability theory and its applications*, volume 1. Wiley, New York, 3rd edition, 1968. First edition published 1950.

M Fréchet. *Les probabilités associées a un système d'événements compatibles et dépendants. Première partie: d'événements en nombre fini fixe*. Number 859 in Actualités scientifiques et industrielles. Hermann, Paris, 1940.

G M Giorgi. Concentration index, Bonferroni. In S Kotz, C B Read, and D L Banks, editors, *Encyclopedia of statistical sciences, Update vol*, pages 141–146. Wiley, New York, 1998.

S Holm. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6:65–70, 1979.

P Pagni. Carlo Emilio Bonferroni. *Bollettino dell'Unione Matematica Italiana*, 15:570–574, 1960.

E Paulson. On the comparison of several experimental categories with a control. *Annals of Mathematical Statistics*, 23:239–246, 1952.

S M Stigler. Stigler's law of eponymy. *Transactions of the New York Academy of Sciences*, 39:147–157, 1980.