

Apostila de probabilidade e estatística

Contents

1	Sobre	5
2	Um pouco da história	7
2.1	Primeiros levantamentos, estudos e publicações & Demografia e aritmética política	7
2.2	Visualização de dados & Estudos e primeiras publicações	14
2.3	Nomes notáveis	17
2.4	Revista Biometrika	18
2.5	Eugenio	18
3	Estatística descritiva	23

Chapter 1

Sobre

Apostila com alguns tópicos de estatística e probabilidade.

Chapter 2

Um pouco da história

2.1 Primeiros levantamentos, estudos e publicações & Demografia e aritmética política

1086

O *Domesday Book* (link) foi encomendado em dezembro de 1085 por Guilherme, o Conquistador (*King William I*), que invadiu a Inglaterra em 1066.

O primeiro esboço foi concluído em agosto de 1086 e continha registros de 13.418 assentamentos nos condados ingleses ao sul dos rios Ribble e Tees (a fronteira com a Escócia) com informações sobre terras, proprietários, uso da terra, empregados e animais cujo propósito básico era fundamentar a taxação.

1602

O dramaturgo inglês William Shakespeare usou a palavra **statists** (estadistas e, portanto, num sentido não relacionado com números ou matemática) no diálogo da Cena II de Hamlet (link).

“Hamlet: Cercado assim por tantas vilanias, mesmo antes de eu poder dizer o prólogo, representava o cérebro. Sentei-me e escrevi com capricho nova carta. Já pensei, como os nossos estadistas, que é feio escrever bem, tendo insistido, até, em desaprendê-lo; mas, nessa hora muito bom me foi isso. Quererias saber qual o conteúdo da mensagem? [...]”

1603

O negociante inglês John Graunt (1620-1674) substituiu a crença pela evidência em *Natural and Political Observations Mentioned in a Following Index and Made upon the Bills of Mortality* (Observações naturais e políticas feitas sobre as notas de mortalidade).



Figure 2.1: Domesday Book

Nesse trabalho, realizado com dados coletados das paróquias de Londres entre 1604 e 1660, Graunt tirou as seguintes conclusões: que havia maior nascimento de crianças do sexo masculino, mas havia distribuição aproximadamente igual de ambos os sexos na população geral; alta mortalidade nos primeiros anos de vida; maior mortalidade nas zonas urbanas em relação às zonas rurais.

1660

Herman Conring (1606-1681), professor de filosofia, medicina e política da Universidade de Helmstadt (atual Alemanha), criou um curso de Ciência política em 1660, que descrevia e examinava as questões fundamentais do Estado. Nele a **estatística** passou a ser considerada como uma disciplina autônoma que tinha por objetivo a descrição das coisas do Estado.

1687

Em 1687 o economista e filósofo inglês William Petty (1623-1687) publicou *Five Essays on Political Arithmetic* (Cinco ensaios sobre aritmética política), sugerindo ao governo inglês a criação de um departamento para registro de **estatísticas** vitais.

O Capitão John Graunt e William Petty instituiram na Inglaterra um novo ramo de estudos denominado de *Political arithmetic* (Aritmética política)

1693

2.1. PRIMEIROS LEVANTAMENTOS, ESTUDOS E PUBLICAÇÕES & DEMOGRAFIA E ARITMÉTICA POLÍTICA

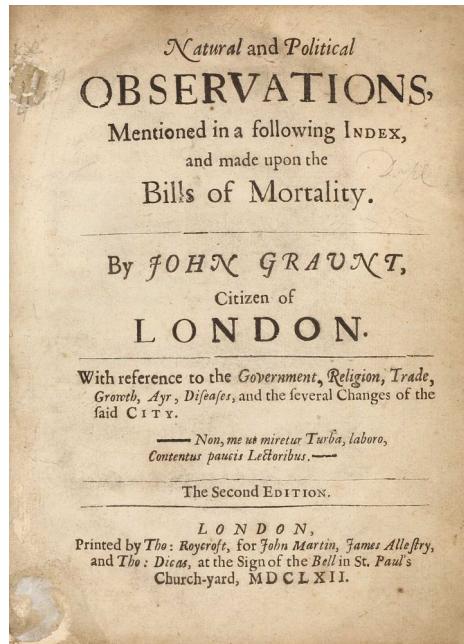


Figure 2.2: Natural and Political Observations Mentioned in a Following Index and Made upon the Bills of Mortality (ed. de 1662)

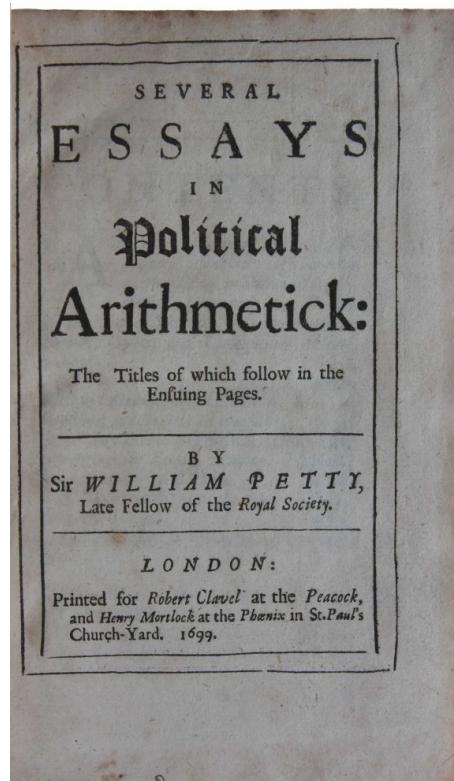


Figure 2.3: Several Essays in Political Arithmetick (ed. de 1699)

2.1. PRIMEIROS LEVANTAMENTOS, ESTUDOS E PUBLICAÇÕES & DEMOGRAFIA E ARITMÉTICA POLÍTICA

O matemático e astrônomo inglês Edmond Halley (1656-1742) construiu em 1693, baseado em dados coletados na cidade (à época) alemã de Breslau, uma *Life Table* (Tábua de sobrevivência), um estudo que analisa as probabilidades de sobrevivência e morte em relação à idade.

Age.	Per- sons.										
Curt.											
1	1000	8	680	15	628	22	585	29	559	36	481
2	855	9	672	16	622	23	579	30	531	37	471
3	796	10	661	17	610	24	573	31	523	38	463
4	750	11	655	18	610	25	567	32	515	39	454
5	734	12	646	19	604	26	560	33	507	40	445
6	710	13	640	20	598	27	553	34	499	41	436
7	692	14	634	21	592	28	549	35	490	42	427
											56 2194
											63 1694
											70 1264
42	419	50	346	57	272	64	202	71	181	78	58 277
44	497	51	335	58	262	65	192	72	120	79	49 84 253
45	397	52	324	59	252	65	182	73	105	80	41 100 107
46	387	53	313	60	242	67	172	74	93	81	34
47	377	54	302	61	232	68	162	75	82	82	28
48	367	55	292	62	222	69	152	76	78	83	23
49	357	56	282	63	212	70	142	77	69	84	201 Sum Total.

Figure 2.4: Halley's life table (1693)

1749

Com um sentido não relacionado com números ou matemática, a palavra **estatística** parece ter sido proposta pela primeira vez no século XVII, pelo historiador e professor alemão (à época Transilvânia) Martin Schmeitzel (1679-1747) da Universidade de Jena e, posteriormente adotada por seu aluno, (igualmente) historiador e jurista Gottfried Achenwall (1719-1772) em 1749, em *Abriß der neuen Staatswissenschaft der vornehmen Europäischen Reiche und Republiken* (Esboço da nova ciência política dos nobres impérios europeus e repúblicas).

1771

William Hooper usou a palavra **estatística** em sua tradução de *The Elements of Universal Erudition* (Elementos da Erudição Universal) escrita por Jacob Friedrich Freiherr von Bielfeld (1717-1770).

Nesse livro, a **estatística** foi definida como a ciência que nos ensina o arranjo político de todos os estados modernos do mundo conhecido (mais uma vez num sentido não associado a números ou matemática).

1790

O jurista e político escocês John Sinclair propôs que se realizasse uma detalhada pesquisa em 938 paróquias para elucidar a história natural e política de seu país (*Statistics Accounts*). Essa pesquisa fazia parte de um projeto muito mais ousado: *The Pyramid of Statistical Enquiry* (A Pirâmide da Pesquisa Estatística).

1854

O médico inglês (considerado por alguns como o “pai” da epidemiologia moderna) John Snow (1813-1858) estudou a dispersão espacial dos casos de cólera em Londres e concluiu que sua causa residia na contaminação da água consumida (poço localizado na *Broad Street*, no distrito do *Soho*): *Report to the*



Figure 2.5: Abriß der neuen Staatswissenschaft der vornehmen Europäischen Reiche und Republiken (1749)

2.1. PRIMEIROS LEVANTAMENTOS, ESTUDOS E PUBLICAÇÕES & DEMOGRAFIA E ARITMÉTICA POLÍTICA

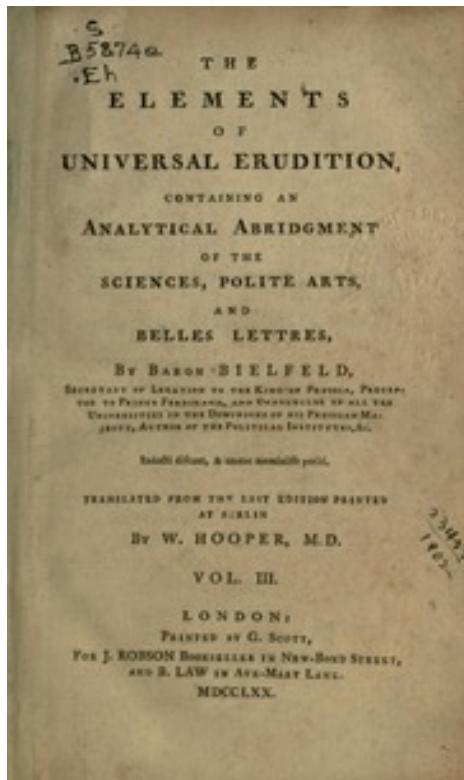


Figure 2.6: The Elements of Universal Erudition (1771)

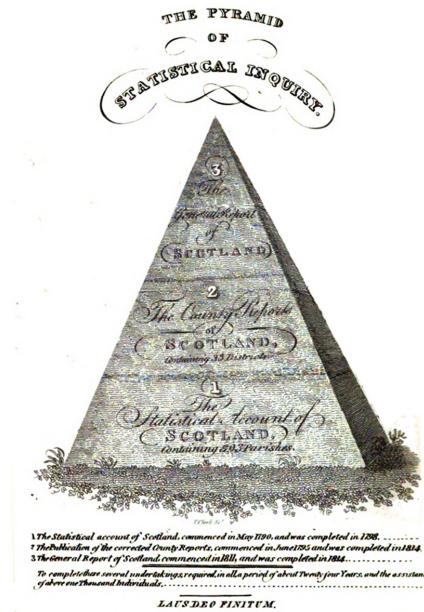


Figure 2.7: The Pyramid of Statistical Enquiry (1814)

Cholera Outbreak in the Parish of St. James, Westminster during the Autumn of 1854 (Relatório sobre o surto de cólera na paróquia de St. James, Westminster durante o outono de 1854).

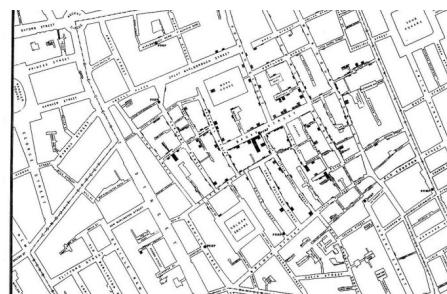


Figure 2.8: Mapa dos casos de cólera (1854)

2.2 Visualização de dados & Estudos e primeiras publicações

1765

O teólogo e filósofo inglês Joseph Priestley (1733-1804) introduziu como inovação

2.2. VISUALIZAÇÃO DE DADOS & ESTUDOS E PRIMEIRAS PUBLICAÇÕES

os primeiros gráficos com linha temporal, em que barras individuais eram usadas para visualizar o tempo de vida de uma pessoa e o todo pode ser usado para comparar a expectativa de vida de várias pessoas.

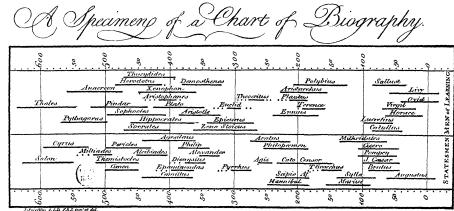


Figure 2.9: Expectativa de vida de diversas pessoas (1765)

1786

O engenheiro e economista escocês William Playfair (1759-1823) é considerado comumente como fundador dos métodos gráficos para apresentação de estatísticas. Playfair concebeu vários tipos de diagramas para visualização de dados:

- em 1786, o gráfico de barras; e,
- em 1801, o gráfico de setores.

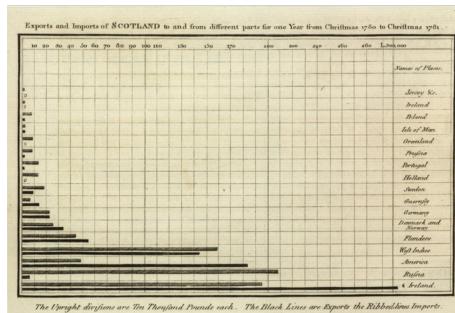


Figure 2.10: Commercial and Political Atlas (Atlas Comercial e Político de 1786): cada barra representa as exportações e importações da Escócia para 17 países em 1781

1856

A enfermeira inglesa Florence Nightingale (1820-1910) conduziu um trabalho pioneiro ao chegar no hospital militar britânico na Turquia em 1856, estabelecendo uma ordem e um método muito necessários aos registros médicos estatísticos e que indicaram serem as precárias práticas sanitárias o culpado da alta mortalidade (link).

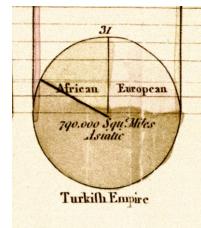


Figure 2.11: Statistical Breviary (Breviário Estatístico de 1801): proporção da extensão do Império Turco em diferentes regiões do mundo: Ásia, Europa e África, antes de 1789

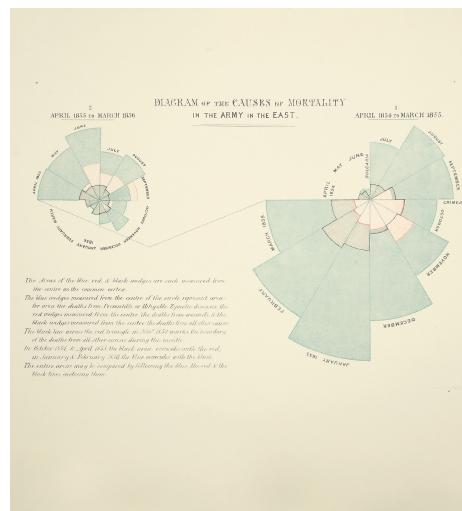


Figure 2.12: Esse diagrama (coxcomb) feito durante a Guerra da Crimeia foi dividido igualmente em 12 setores, representando os meses do ano, com a área sombreada do setor de cada mês proporcional à taxa de mortalidade naquele mês. Seu sombreamento com código de cores indicava a causa da morte em cada área do diagrama

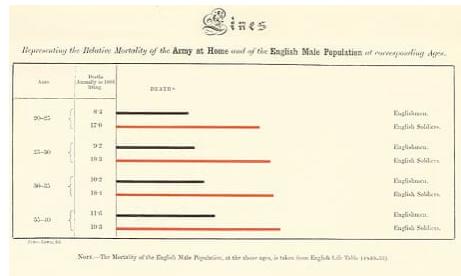


Figure 2.13: Gráfico de barras de Florence Nightingale mostrando as diferenças de mortalidade entre soldados britânicos e a população masculina inglesa geral (civis)

2.3 Nomes notáveis

Karl Pearson (1857-1936) é amplamente considerado o fundador da disciplina moderna de **estatística**, e também é famoso como um filósofo da ciência, como escritor sobre o darwinismo social e como um dos principais impulsionadores para instalar a eugenia como a ciência social chave. Uma breve biografia de cada um dos pesquisadores a seguir relacionados pode ser obtida em: [\(link\)](#).

- Niccolò Fontana Tartaglia (Veneza à época, hoje Itália: 1499-1557)
- Girolamo Cardano (Pávia à época, hoje Itália: 1501-1576)
- Galileu Galilei (Florença à época, hoje Itália: 1564-1642)
- Pierre de Fermat (França: 1607-1665)
- Blaise Pascal (França: 1623-1662)
- Jakob Bernoulli (Suíça: 1655-1705)
- Abraham de Moivre (França: 1667-1754)
- Thomas Bayes (Inglaterra: 1702-1761)
- Pierre-Simon Laplace (França: 1749-1827)
- Johann Carl Friedrich Gauss (Alemanha: 1777-1856)
- Lambert Adolphe Jacques Quêtelet (França à época, hoje Bélgica: 1796-1874)
- Pafnuti Lvovitch Chebyshev (Rússia: 1821-1894)
- Francis Galton (Inglaterra: 1822-1911)
- Wilhelm Lexis (Alemanha: 1837-1914)
- Thorvald Nicolai Thiele (Dinamarca: 1838-1910)
- Friedrich Robert Helmert (Saxônia: 1843-1917)
- Francis Ysidro Edgeworth (Inglaterra: 1845-1926)
- James Douglas Hamilton Dickson (Escócia: 1849-1931)
- Andrei Andreyevich Markov (Rússia: 1856-1922)
- Aleksandr Mikhailovich Lyapunov (Rússia: 1857-1918)
- Walter Frank Raphael Weldon (Inglaterra: 1860-1906)
- Karl Pearson (Inglaterra: 1857-1936)
- William Seally Gosset (Inglaterra: 1876-1937)

- Ronald Aylmer Fisher (Inglaterra: 1890-1962)
- Andrei Nikolaevich Kolmogorov (Rússia: 1903-1987)

2.4 Revista Biometrika

“Pretende-se que a *Biometrika* sirva como um meio não apenas de coletar ou publicar, sob um título, dados biológicos de um tipo não coletados sistematicamente ou publicados em outro lugar em qualquer outro periódico, mas também de disseminar um conhecimento de tal teoria estatística para o seu tratamento científico[...]”

Em outubro de 1901 foi fundada a *Biometrika, the Journal for the Statistical Study of Biological Problems* (Biometrika, o Jornal para o Estudo Estatístico de Problemas Biológicos) com o propósito de promover a análise estatística de fenômenos biológicos, isto é, a matematização da biologia.

Os fundadores da *Biometrika* foram Sir Francis Galton (primo de Charles Darwin), Walter Frank Raphael Weldon e Karl Pearson. A maior parte do trabalho foi feita por Pearson e Weldon, este último focando na edição do conteúdo (ou seja, o aspecto biológico) e o primeiro nos detalhes, incluindo correções de prova. Galton e o eugenista americano Charles Davenport atuaram, respectivamente, como consultor e editor.

Alguns dos tópicos abordados na revista incluem criminologia, botânica, zoologia, epidemiologia e outros aspectos da saúde humana. Na década de 1930, o caráter da *Biometrika* mudou, e “representou a vanguarda internacional da pesquisa em métodos estatísticos e sua aplicação na ciência e tecnologia”, ao invés de focar a hereditariedade.

Sir Francis Galton, que serviu como editor da primeira edição (1901), escreveu a Introdução, que incluiu uma declaração de propósito para a revista ([link](#)).

2.5 Eugenia

Em 16 de maio de 1883 Sir Francis Galton cunhou o termo “eugenia”, posteriormente descrevendo-o como “o estudo das agências sob controle social que podem melhorar ou reparar as qualidades raciais das gerações futuras, seja fisicamente ou mentalmente”.

Galton detalha o conceito em seu livro *Inquiries into Human Faculty and its Development*, e recomenda que indivíduos de famílias altamente classificadas em seu sistema de mérito sejam encorajados a se casar cedo e receber incentivos para ter filhos. Ele também condenou os casamentos tardios dentro desse mesmo grupo como “desgênicos” ou desvantajosos para a espécie humana.

A palavra “eugenia” foi extraída da palavra grega *eu*, que significa bem, e *genos*, que significa prole. Juntos, significa bem-nascido.

Este livro caiu em domínio público e pode ser lido na íntegra online. A caracterização original de eugenia de Galton pode ser encontrada na página 17 desta edição de domínio público (Parte 1 do pdf):

“uma breve palavra para expressar a ciência de melhorar o rebanho, que não está de modo algum confinado a questões de acasalamento criterioso, mas que, especialmente no caso do homem, toma conhecimento de todas as influências que tendem, mesmo que em grau remoto, a dar ao raças ou linhagens de sangue mais adequadas uma melhor chance de prevalecer rapidamente sobre os menos adequados do que teriam de outra forma [...]”(Galton, 1883, p.17)

Há poucos anos alguns grupos sociais viram no trabalho e opiniões de Fisher endossos ao colonialismo, à supremacia branca e à eugenia.

Outros grupos, todavia, afirmam que Fisher não era racista e eugenista, embora ele achasse que havia diferenças comportamentais e de inteligência entre os grupos humanos.

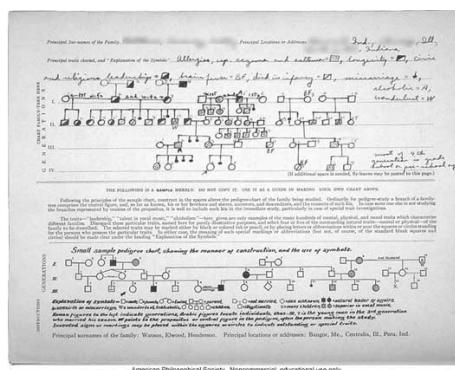


Figure 2.14: Gráfico de linhagens para alergias

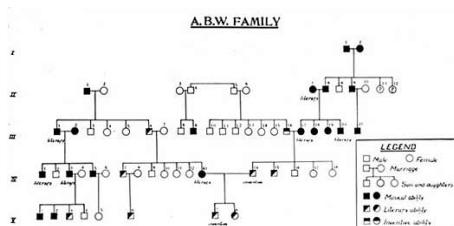


Figure 2.15: Gráfico de linhagens para aptidão musical

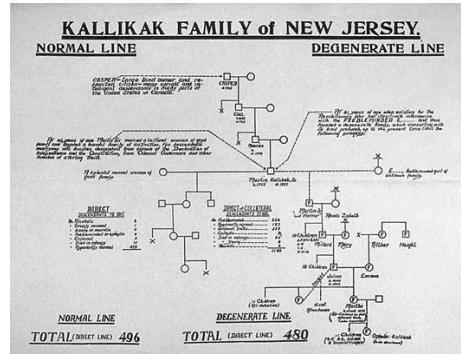


Figure 2.16: Linhas "normais" e "degeneradas" da família Kallikak (New Jersey)

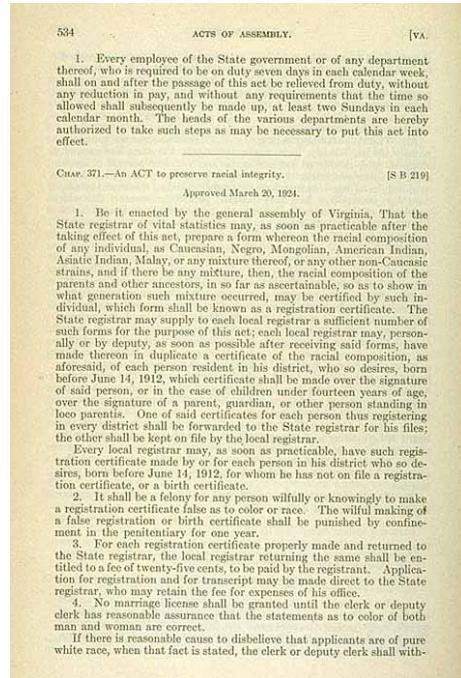


Figure 2.17: Lei da Inegridade Racia (Virginia, EUA, 1924)



Figure 2.18: Licença para casamento

Chapter 3

Estatística descritiva