

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 Eugenia	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).

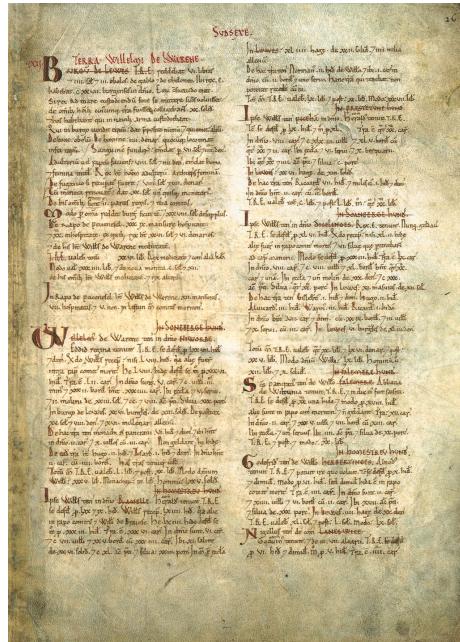


Figure 2.1: Domesday Book

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

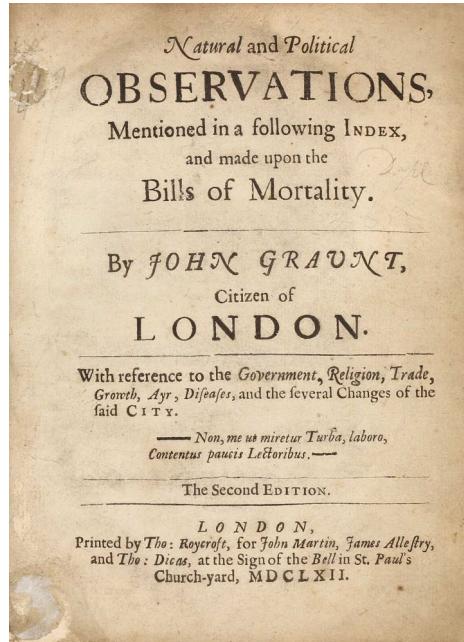


Figure 2.2: Natural and Political Observations Mentioned in a Following Index and Made upon the Bills of Mortality (ed. de 1662)

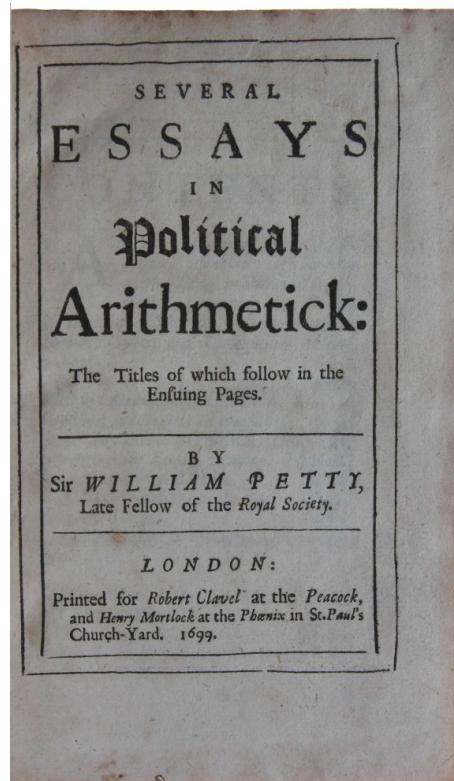


Figure 2.3: Several Essays in Political Arithmetick (ed. de 1699)

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 Bresláu, uma *Life Table* (Tábua de sobrevivência), um estudo que analisa as probabilidades de sobrevivência e morte em relação à idade.

Age- Curt.	Per- sons														
1 1000	8 585	15 628	22 685	29 739	33 769	35 816	41 881	48 911	54 954	57 981	61 1000	64 1014	68 1027	72 1037	75 1047
2 855	9 670	16 623	23 739	30 731	33 723	37 743	42 753	47 753	52 748	56 743	60 736	64 736	68 736	72 736	75 736
3 798	10 651	17 616	24 626	31 573	33 573	38 523	43 519	49 515	55 454	59 454	63 394	67 394	71 394	75 394	79 394
4 760	11 653	18 610	25 567	33 516	36 516	40 466	44 409	49 401	54 366	58 366	62 316	66 316	70 316	74 316	78 316
5 738	12 646	19 564	26 466	34 366	37 366	42 316	46 266	50 216	55 166	59 166	63 116	67 116	71 116	75 116	79 116
6 710	13 640	20 598	27 533	35 491	38 491	43 441	47 391	51 341	56 291	60 291	64 241	68 241	72 241	76 241	80 241
7 692	14 634	21 592	28 535	36 454	39 454	44 404	48 354	52 304	57 254	61 254	65 194	69 194	73 194	77 194	81 194
8 674	15 628	22 592	29 535	37 454	40 454	45 404	49 354	53 304	58 254	62 254	66 194	70 194	74 194	78 194	82 194
9 656	16 623	23 592	30 535	38 454	41 454	46 404	50 354	54 304	59 254	63 254	67 194	71 194	75 194	79 194	83 194
10 638	17 616	24 592	31 535	39 454	42 454	47 404	51 354	55 304	59 254	63 254	67 194	71 194	75 194	79 194	83 194
11 620	18 610	25 592	32 535	40 454	43 454	48 404	52 354	56 304	60 254	64 254	68 194	72 194	76 194	80 194	84 194
12 602	19 604	26 592	33 535	41 454	44 454	49 404	53 354	57 304	61 254	65 254	69 194	73 194	77 194	81 194	85 194
13 584	20 604	27 592	34 535	42 454	45 454	49 404	53 354	57 304	61 254	65 254	69 194	73 194	77 194	81 194	85 194
14 566	21 604	28 592	35 535	43 454	46 454	50 404	54 354	58 304	62 254	66 254	70 194	74 194	78 194	82 194	86 194
15 548	22 604	29 592	36 535	44 454	47 454	51 404	55 354	59 304	63 254	67 254	71 194	75 194	79 194	83 194	87 194
16 530	23 604	30 592	37 535	45 454	48 454	52 404	56 354	60 304	64 254	68 254	72 194	76 194	80 194	84 194	88 194
17 512	24 604	31 592	38 535	46 454	49 454	53 404	57 354	61 304	65 254	69 254	73 194	77 194	81 194	85 194	89 194
18 494	25 604	32 592	39 535	47 454	50 454	54 404	58 354	62 304	66 254	70 254	74 194	78 194	82 194	86 194	90 194
19 476	26 604	33 592	40 535	48 454	51 454	55 404	59 354	63 304	67 254	71 254	75 194	79 194	83 194	87 194	91 194
20 458	27 604	34 592	41 535	49 454	52 454	56 404	60 354	64 304	68 254	72 254	76 194	80 194	84 194	88 194	92 194
21 440	28 604	35 592	42 535	50 454	53 454	57 404	61 354	65 304	69 254	73 254	77 194	81 194	85 194	89 194	93 194
22 422	29 604	36 592	43 535	51 454	54 454	58 404	62 354	66 304	70 254	74 254	78 194	82 194	86 194	90 194	94 194
23 404	30 604	37 592	44 535	52 454	55 454	59 404	63 354	67 304	71 254	75 254	79 194	83 194	87 194	91 194	95 194
24 386	31 604	38 592	45 535	53 454	56 454	60 404	64 354	68 304	72 254	76 254	80 194	84 194	88 194	92 194	96 194
25 368	32 604	39 592	46 535	54 454	57 454	61 404	65 354	69 304	73 254	77 254	81 194	85 194	89 194	93 194	97 194
26 350	33 604	40 592	47 535	55 454	58 454	62 404	66 354	70 304	74 254	78 254	82 194	86 194	90 194	94 194	98 194
27 332	34 604	41 592	48 535	56 454	59 454	63 404	67 354	71 304	75 254	79 254	83 194	87 194	91 194	95 194	99 194
28 314	35 604	42 592	49 535	57 454	60 454	64 404	68 354	72 304	76 254	80 254	84 194	88 194	92 194	96 194	100 194
29 296	36 604	43 592	50 535	58 454	61 454	65 404	69 354	73 304	77 254	81 254	85 194	89 194	93 194	97 194	101 194
30 278	37 604	44 592	51 535	59 454	62 454	66 404	70 354	74 304	78 254	82 254	86 194	90 194	94 194	98 194	102 194
31 260	38 604	45 592	52 535	60 454	63 454	67 404	71 354	75 304	79 254	83 254	87 194	91 194	95 194	99 194	103 194
32 242	39 604	46 592	53 535	61 454	64 454	68 404	72 354	76 304	80 254	84 254	88 194	92 194	96 194	100 194	104 194
33 224	40 604	47 592	54 535	62 454	65 454	69 404	73 354	77 304	81 254	85 254	89 194	93 194	97 194	101 194	105 194
34 206	41 604	48 592	55 535	63 454	66 454	70 404	74 354	78 304	82 254	86 254	90 194	94 194	98 194	102 194	106 194
35 188	42 604	49 592	56 535	64 454	67 454	71 404	75 354	79 304	83 254	87 254	91 194	95 194	99 194	103 194	107 194
36 170	43 604	50 592	57 535	65 454	68 454	72 404	76 354	80 304	84 254	88 254	92 194	96 194	100 194	104 194	108 194
37 152	44 604	51 592	58 535	66 454	69 454	73 404	77 354	81 304	85 254	89 254	93 194	97 194	101 194	105 194	109 194
38 134	45 604	52 592	59 535	67 454	70 454	74 404	78 354	82 304	86 254	90 254	94 194	98 194	102 194	106 194	110 194
39 116	46 604	53 592	60 535	68 454	71 454	75 404	79 354	83 304	87 254	91 254	95 194	99 194	103 194	107 194	111 194
40 98	47 604	54 592	61 535	69 454	72 454	76 404	80 354	84 304	88 254	92 254	96 194	100 194	104 194	108 194	112 194
41 80	48 604	55 592	62 535	70 454	73 454	77 404	81 354	85 304	89 254	93 254	97 194	101 194	105 194	109 194	113 194
42 62	49 604	56 592	63 535	71 454	74 454	78 404	82 354	86 304	90 254	94 254	98 194	102 194	106 194	110 194	114 194
43 44	50 604	57 592	64 535	72 454	75 454	79 404	83 354	87 304	91 254	95 254	99 194	103 194	107 194	111 194	115 194
44 26	51 604	58 592	65 535	73 454	76 454	80 404	84 354	88 304	92 254	96 254	100 194	104 194	108 194	112 194	116 194
45 8	52 604	59 592	66 535	74 454	77 454	81 404	85 354	89 304	93 254	97 254	101 194	105 194	109 194	113 194	117 194
46 0	53 604	60 592	67 535	75 454	78 454	82 404	86 354	90 304	94 254	98 254	102 194	106 194	110 194	114 194	118 194

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

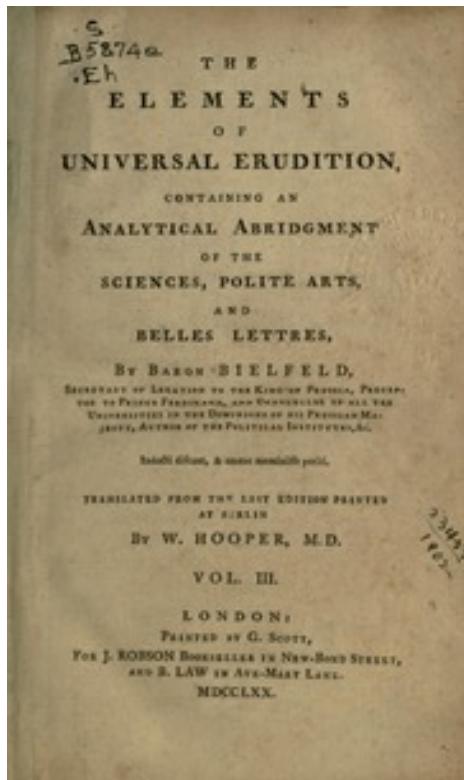


Figure 2.6: The Elements of Universal Erudition (1771)

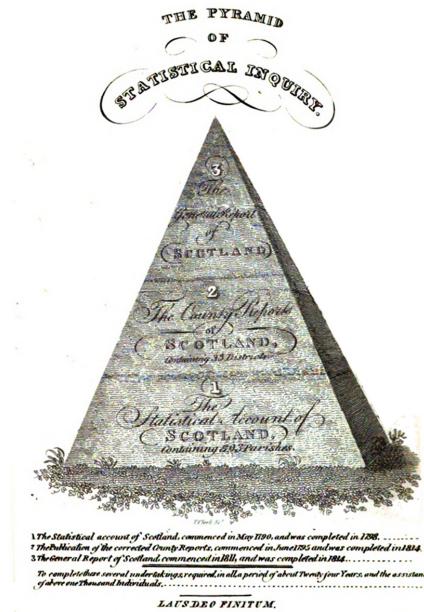


Figure 2.7: The Pyramid of Statistical Enquiry (1814)

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).

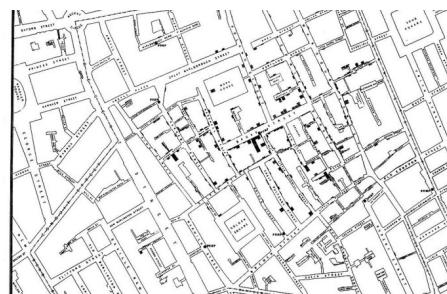


Figure 2.8: Mapa dos casos de cólera (1854)

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.

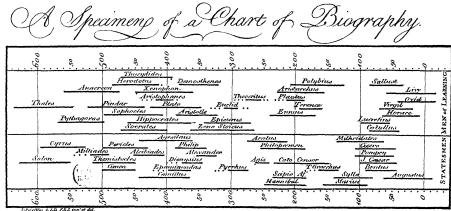


Figure 2.9: Expectativa de vida de diversas pessoas (1765)

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.

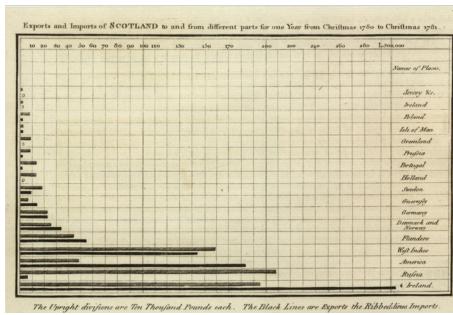


Figure 2.10: Commercial and Political Atlas (Atlas Comercial e Político de 1786): cada barra representa as exportações e importações da Escócia para 17 países em 1781

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).

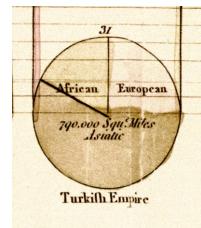


Figure 2.11: Statistical Breviary (Breviário Estatístico de 1801): proporção da extensão do Império Turco em diferentes regiões do mundo: Ásia, Europa e África, antes de 1789

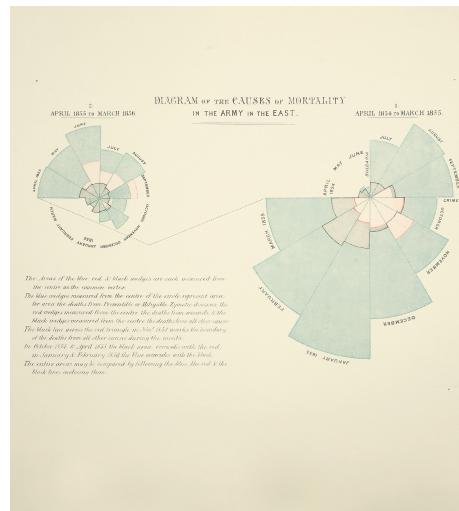


Figure 2.12: Esse diagrama (coxcomb) feito durante a Guerra da Crimeia foi dividido igualmente em 12 setores, representando os meses do ano, com a área sombreada do setor de cada mês proporcional à taxa de mortalidade naquele mês. Seu sombreamento com código de cores indicava a causa da morte em cada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

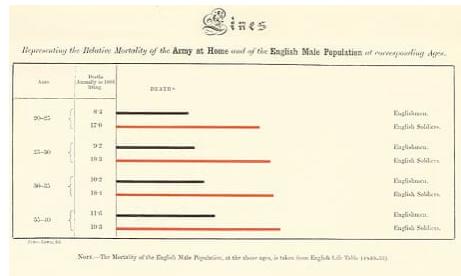


Figure 2.13: Gráfico de barras de Florence Nightingale mostrando as diferenças de mortalidade entre soldados britânicos e a população masculina inglesa geral (civis)

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)
- Abrahan 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.

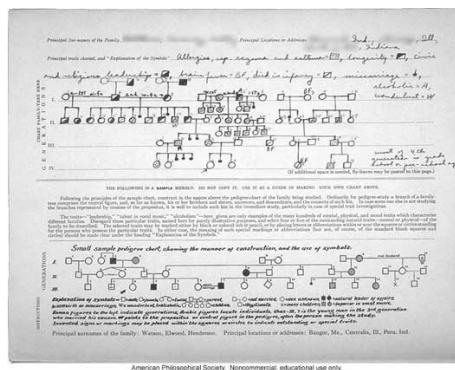


Figure 2.14: Gráfico de linhagens para alergias

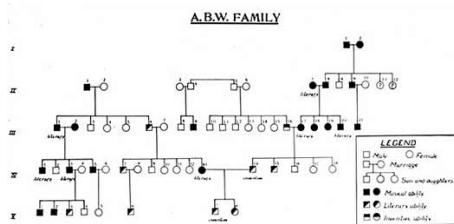


Figure 2.15: Gráfico de linhagens para aptidão musical

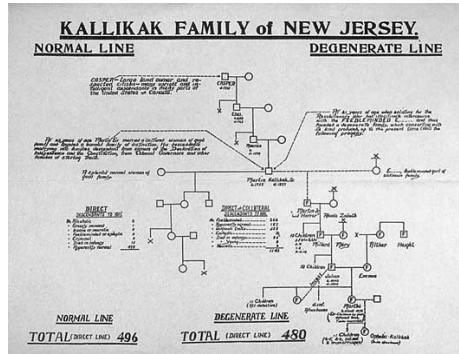


Figure 2.16: Linhas "normais" e "degeneradas" da família Kallikak (New Jersey)

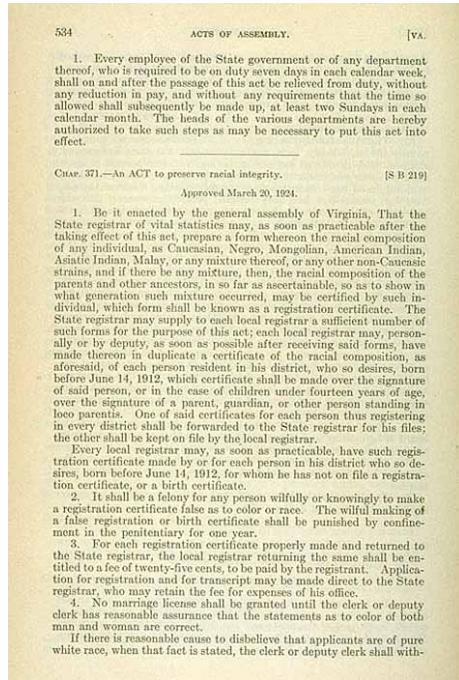


Figure 2.17: Lei da Inegridade Racia (Virginia, EUA, 1924)



Figure 2.18: Licença para casamento

Chapter 3

Estatística descritiva