



Sir **Ronald Aylmer Fisher**, FRS (17 de fevereiro de 1890 — 29 de julho de 1962) foi um estatístico, biólogo evolutivo e geneticista inglês.

Foi descrito por Anders Hald como "um gênio que criou praticamente sozinho as fundações para a moderna ciência estatística" e Richard Dawkins, que o descreveu como "o maior dos sucessores de Darwin".

Complementando a medida de informação própria para medir incerteza sobre *espaços desordenados* — proposta com sucesso pela **teoria da informação**, e publicada em 1949 por Claude Shannon (1916-2001) e Warren Weaver (1894-1978) no livro *Teoria matemática da comunicação* (*The Mathematical Theory of Communication*) —, **Fisher** criou uma medida alternativa de informação apropriada para medir incerteza sobre *espaços ordenados*.

FONTE: <http://pt.wikipedia.org/wiki/Brasil>

- ⇒ História
- ⇒ Conceitos Fundamentais



## 1 - INTRODUÇÃO

### 1.1 - HISTÓRIA

O termo **Estatística** é proveniente do latim *Status*, cujo significado é “Estado”. Utilizada a princípio no tratamento de dados e gráficos que descreviam aspectos de um estado ou país.

O surgimento e desenvolvimento da Estatística podem ser observados paralelamente à evolução da civilização e do conhecimento humano. Os egípcios, persas e demais povos antigos utilizavam-se da estatística para avaliar as condições econômicas do Estado, o movimento das populações e a produção de cada província. Um exemplo do uso da Estatística pode ser observado no livro de Confúcio onde são apresentados os resultados de censos realizados nos anos de 2.275 a.C. e de 2.238 a.C..

No século XVI a estatística passou a ser estudada pelos pensadores da época, principalmente os matemáticos. A estatística teve acelerado desenvolvimento a partir do século XVII, com os estudos de Bernoulli, Fermat, Pascal, Laplace, Gauss, Galton, Pearson, Fisher, Poisson e outros que estabeleceram suas características atuais.

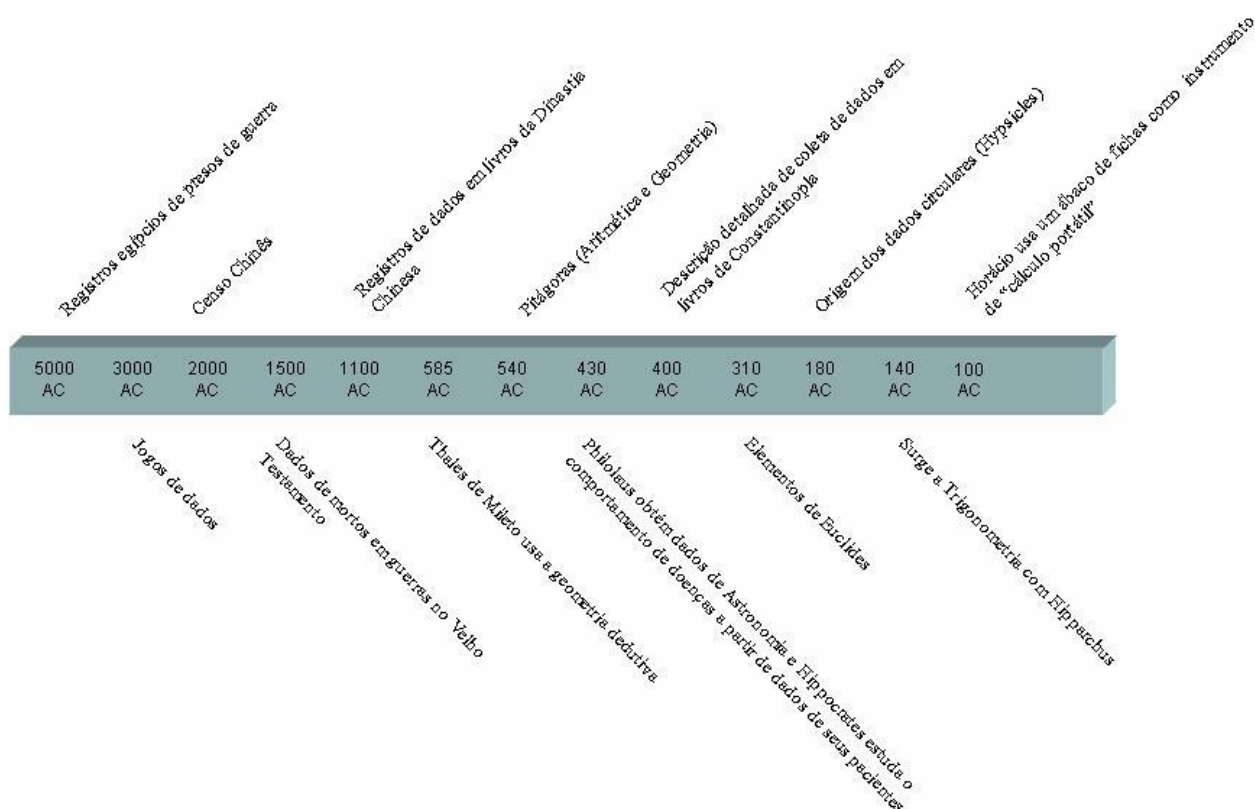


Figura 1.- Linha do tempo apontando conceitos e fatos importantes da Estatística

Muitas das informações veiculadas pelos meios de comunicação atuais são provenientes de pesquisas e estudos estatísticos. Como exemplo da aplicabilidade da Estatística tem-se os índices inflacionários, de emprego e desemprego anunciados e analisados pela mídia.

O Instituto Brasileiro de Geografia e Estatística - IBGE - é o órgão responsável pela produção das estatísticas oficiais que subsidiam estudos e planejamentos governamentais no país.

Os conceitos estatísticos têm exercido profunda influência na maioria dos campos do conhecimento humano. Métodos estatísticos vêm sendo utilizados no aprimoramento de produtos agrícolas, no desenvolvimento de equipamentos espaciais, no controle do tráfego, na previsão de surtos epidêmicos bem como no aprimoramento de processos de gerenciamento, tanto na área governamental como na iniciativa privada.

Na prática, a Estatística pode ser empregada como ferramenta fundamental em várias outras ciências.

Para o governo de um determinado país, a Estatística é ferramenta fundamental para que se possa traçar planos sociais e econômicos e projetar metas para o futuro.

Na pesquisa científica, a Estatística desempenha importante papel na obtenção de dados relevantes, em testes de hipóteses, estimação de parâmetros e interpretação dos resultados.

Na indústria, técnicas estatísticas extremamente simples são utilizadas para que a qualidade dos produtos possa ser mantida dentro de um determinado nível.

No mercado financeiro, os métodos estatísticos são empregados para previsões de taxas de juros e preços de diferentes bens e para desenvolvimento de estratégias de investimentos que maximizem os lucros.

No comércio, a Estatística pode ser usada para previsão de demandas, planejamento da produção e implantação de técnicas administrativas eficientes que garantam o melhor lucro.

Na Medicina, os princípios de planejamento de experimentos são utilizados em análises de drogas e em ensaios clínicos. A informação que é fornecida por um grande número de testes bioquímicos é acessada estatisticamente para diagnósticos e previsões de possíveis causas de doenças. A aplicação de técnicas estatísticas tornou o diagnóstico médico mais objetivo, combinando-se a sabedoria dos melhores *experts* com o conhecimento das diferenças entre doenças indicadas pelos testes clínicos.

Na Literatura, os métodos estatísticos podem ser usados para quantificar os estilos de diversos autores, o que pode ser útil para se decidir a autoria de determinada obra, em casos de disputa autoral.

Em alguns estudos arqueológicos, técnicas estatísticas de comparação entre diferentes objetos encontrados têm representado um eficiente método de se determinar a que culturas pertenciam antigos artefatos e de colocar tais artefatos em ordem cronológica.

Nas cortes de justiça, evidência estatística na forma de probabilidade de ocorrência de eventos pode ser uma importante informação trazida por uma das partes em um tribunal.

Em Administração, a análise estatística funciona como uma importante ferramenta para se diagnosticar problemas de gerenciamento em diferentes setores de uma empresa e para propor políticas de investimento mais eficientes dentro da própria empresa.

A Estatística é definida no dicionário como um “ramo da Matemática que trata da coleta, análise, interpretação e apresentação de massas de dados numéricos”.

Atualmente, a estatística é definida da seguinte forma:

- ***Estatística:*** *é um conjunto de métodos e processos quantitativos que servem para estudar e medir os fenômenos coletivos.*

Ou ainda

- ***Estatística:*** *é um conjunto de métodos para planejar experimentos, obter dados e organiza-los, resumi-los, analisa-los, interpretá-los e deles extrair conclusões.*

A estatística tem basicamente duas finalidades:

- Descrever os fenômenos e suas características;
- Fazer previsões sobre as ocorrências futuras de certo fenômeno em condições semelhantes às aquelas em que ele ocorreu no passado.

## 1.2 - CONCEITOS FUNDAMENTAIS

### 1.2.1 - POPULAÇÃO, CENSO E AMOSTRA.

Ao escutarmos a palavra “Estatística” pensamos imediatamente em populações, censos e amostras.

**População** é o conjunto de todos os itens que interessam ao estudo de um fenômeno segundo alguma(s) característica(s). Por exemplo, num processo de controle de qualidade de uma fábrica a população seria composta por todos os aparelhos produzidos nessa fábrica, ao se estudar o desempenho dos candidatos num vestibular a população seria composta por todos os candidatos inscritos no vestibular que compareceram a ele. Isto é, valores, medidas, pessoas, outros seres ou objetos que constituem uma população são todos aqueles que estão ligados diretamente ao fato que desejamos estudar. É óbvio que as diversas populações são diferentes, podendo seus elementos apresentar apenas uma característica ou várias, em alguns casos sendo possível até dividir a população em populações menores separando seus elementos de acordo com as características que eles apresentam. Temos, ainda, populações finitas ou infinitas e ainda aquelas que, apesar do número de elementos ser um número finito este número pode ser tão grande que a população pode ser considerada infinita.

Um **censo** é uma coleção de dados relativos a todos os elementos de uma população. Isto é o censo é um estudo que envolve todos os elementos relacionados a uma determinada população. É certo que em alguns casos um estudo que abrange toda uma população é inviável, ou porque o número de elementos desta população é muito grande ou porque o processo para a obtenção das informações necessárias é um processo dispendioso ou muito demorado ou, ainda, porque esse processo é um processo destrutivo, isto é, só conseguiremos a informação se o item a ser testado for destruído. Neste caso deve-se proceder de outra maneira.

Uma **amostra** é qualquer subconjunto, não vazio, de uma população. Assim uma amostra é um conjunto formado por alguns elementos de uma população, sendo que o processo para a determinação de quais elementos pertencerão à amostra deve obedecer a certas considerações que serão estudadas mais adiante. Geralmente uma amostra é utilizada nos casos em que não é possível realizar um censo.

Diretamente ligados aos conceitos de população e amostra estão os conceitos de *parâmetro* e *estatística*. Estes conceitos são utilizados para diferenciar se um valor obtido num estudo está ligado a toda uma população ao à apenas uma parte dela (amostra).

- **Parâmetro**: é uma medida numérica que descreve uma característica de toda uma população.
- **Estatística**: é uma medida numérica relacionada a uma característica de uma amostra.

Assim ao afirmarmos que 35% do faturamento de uma empresa é decorrente de apenas um cliente, este dado é um *parâmetro*, pois esta informação é baseada em todos os clientes da empresa. Já se uma pesquisa, realizada com 800 alunos, indica que apenas 12% deles possuem condução própria, esta medida é uma estatística devido ao fato de que ela ter sido baseada numa amostra e não em toda a população de alunos.

### 1.2.2 - FENÔMENOS ESTATÍSTICOS

**Fenômeno Estatístico** é qualquer evento que se pretenda analisar, cujo estudo seja passível da aplicação das técnicas estatísticas. Os fenômenos podem ser classificados em três tipos:

- **Fenômeno Coletivo** ou **Fenômeno de Massa** são aqueles que não pode ser definido por uma simples manifestação. O volume médio de vendas diárias numa loja de departamentos, a taxa de natalidade ou de mortalidade de uma cidade são exemplos de fenômenos coletivos.
- **Fenômenos Individuais** ou **Particulares** compõem os fenômenos coletivos. Cada venda diária numa loja de departamentos, cada nascimento ou óbito numa cidade são exemplos de fenômenos individuais.
- **Fenômenos de Multidão** diferenciam-se dos fenômenos coletivos pelo fato de que as características observadas para a coletividade não são verificadas para cada elemento isoladamente.

### 1.2.3 - VARIÁVEIS

Durante a observação de um fenômeno, o estatístico pode coletar dados de diversas características deste fenômeno. Para cada característica (que passaremos a chamar *variável*) podem ser observados vários valores. Dependendo destes valores a variável pode ser “**Qualitativa**” ou “**Quantitativa**”.

- **Variável Qualitativa:** *é uma variável onde seus valores (ou dados) podem ser agrupados em categorias (ou atributos).*
- **Variável Quantitativa:** *é uma variável onde seus valores são numéricos, podendo representar enumerações ou medições..*

Uma variável Qualitativa é uma variável que assume valores que ilustram categorias ou atributos de uma característica do fenômeno que está sendo estudado. Exemplificando, se estamos interessados em saber a opinião de consumidores sobre a qualidade de certo produto as possíveis respostas seriam, por exemplo, “Excelente”, “Bom”, “Ruim” ou “péssimo”. Assim a qualidade do produto em questão estaria dividida em categorias. Uma variável qualitativa pode ser de dois tipos: “**Nominal**” ou “**Ordinal**”. Uma variável qualitativa Nominal é aquela que apresenta apenas valores categorizados onde não é evidente uma ordenação lógica. Como exemplo, temos as variáveis SEXO, COR, LOCAL, etc. Uma variável qualitativa Ordinal é aquela que apresenta valores categorizados, porém é evidente uma ordenação lógica dos valores. Exemplificando, temos variáveis que assumem valores tipo o grau de instrução de uma pessoa (1º Grau, 2º Grau ou Nível Superior), a qualificação de um item (Ruim, Bom ou Ótimo), etc.

- **Variável Qualitativa Nominal:** *é uma variável que assume apenas valores categorizados não ordenáveis.*
- **Variável Qualitativa Ordinal:** *é uma variável que assume apenas valores categorizados ordenáveis.*

Uma variável Quantitativa é uma variável que assume valores numéricos, podendo representar uma enumeração ou uma medição. Dependendo do tipo de valores que uma variável quantitativa pode assumir ela ainda pode ser classificada em “**Discreta**” ou “**Contínua**”. Uma variável quantitativa será Discreta se ela assumir apenas valores inteiros. Geralmente os valores provenientes de uma enumeração são valores discretos. Uma variável quantitativa será Contínua se ela assumir qualquer valor pertencente ao conjunto dos números Reais num intervalo dado. Geralmente valores obtidos através de medições são valores contínuos.

- **Variável Quantitativa Discreta:** *é uma variável que assume apenas valores inteiros.*
- **Variável Quantitativa Contínua:** *é uma variável que assume qualquer valor pertencente ao conjunto dos números reais num intervalo dado.*

O esquema a seguir ilustra resumidamente os tipos de variáveis.

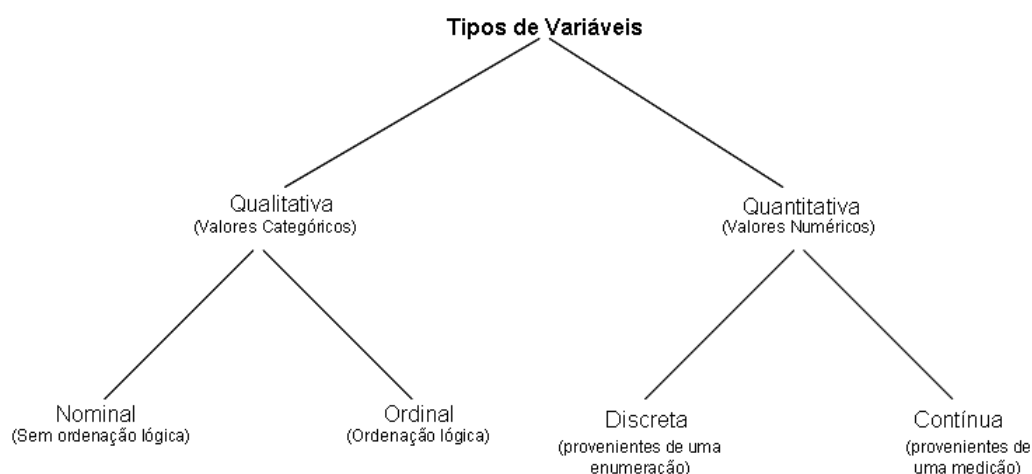


Figura 2.- Esquema representativo dos tipos de variáveis

A tabela a seguir ilustra os diversos tipos de variáveis.

**TABELA 1.- CLASSIFICAÇÃO DO CAMPEONATO DE FORMULA 1 - TEMPORADA DE 2009**

PILOTO	NACIONALIDADE	EQUIPE	TEMPORADA 2009		CARREIRA					
			Posição	Pontos	Títulos	Vitórias	Pontos	GPs	Pole position	Pódios
Jenson Button	ING	Brawn-Mercedes	1º	95	1	7	321	169	7	23
Sebastian Vettel	ALE	RBR-Renault	2º	84	0	4	115	42	5	8
Rubens Barrichello	BRA	Brawn-Mercedes	3º	77	0	11	602	283	14	68
Mark Webber	AUS	RBR-Renault	4º	69,5	0	2	161,5	137	1	9
Lewis Hamilton	ING	McLaren-Mercedes	5º	49	1	11	256	51	17	27
Kimi Raikkonen	FIN	Ferrari	6º	48	1	18	579	154	16	62
Nico Rosberg	ALE	Williams-Toyota	7º	34,5	0	0	75,5	69	0	2
Jarno Trulli	ITA	Toyota	8º	32,5	0	1	244,5	215	4	11
Fernando Alonso	ESP	Renault	9º	26	2	21	577	137	18	53
Timo Glock	ALE	Toyota	10º	24	0	0	51	36	0	3
Felipe Massa	BRA	Ferrari	11º	22	0	11	320	114	15	28
Heikki Kovalainen	FIN	McLaren-Mercedes	12º	22	0	1	105	51	1	4
Nick Heidfeld	ALE	BMW Sauber	13º	19	0	0	215	167	0	12
Robert Kubica	POL	BMW Sauber	14º	17	0	1	137	56	1	9
Giancarlo Fisichella	ITA	Ferrari	15º	8	0	3	275	228	4	19
Sebastien Buemi	SUI	STR-Ferrari	16º	6	0	0	5	16	0	0
Adrian Sutil	ALE	Force India-Mercedes	17º	5	0	0	6	51	0	0
Kamui Kobayashi	JAP	Toyota	18º	3	0	0	0	1	0	0
Sebastien Bourdais	FRA	STR-Ferrari	19º	2	0	0	6	27	0	0
Nelson Piquet Jr.	BRA	Renault	20º	0	0	0	19	28	0	1
Kazuki Nakajima	JAP	Williams-Toyota	21º	0	0	0	9	35	0	0
Jaime Alguersuari	ESP	STR-Ferrari	22º	0	0	0	0	7	0	0
Romain Grosjean	FRA	Renault	23º	0	0	0	0	6	0	0
Luca Badoer	ITA	Ferrari	24º	0	0	0	0	50	0	0
Vitantonio Liuzzi	ITA	Force India-Mercedes	25º	0	0	0	5	43	0	0

FONTE: Site Terra.com seção de esportes

As variáveis “Nacionalidade”, “Equipe” e “Posição” são variáveis qualitativas. As demais variáveis são variáveis quantitativas. As variáveis “Nacionalidade” e “Equipe” são variáveis qualitativas nominais; a variável “posição” é qualitativa ordinal. As variáveis “pontos na temporada” e “Pontos na Carreira” são variáveis contínuas, uma vez que no regulamento da competição os pontos podem ser fracionários, dependendo da porção concluída da corrida disputada. Isto é, em certos casos previstos no regulamento um piloto poderia receber uma pontuação 5,5, se a corrida for interrompida por algum motivo e seu prosseguimento for impossível. As outras variáveis são quantitativas discretas.