

A estatística

A Estatística é uma ciência multidisciplinar, parte da Matemática Aplicada, que fornece métodos para coleta, organização, descrição, análise e interpretação de dados e para a utilização dos mesmos na tomada de decisões. Seu objetivo é o estudo da variabilidade, da incerteza, para a tomada de decisões frente a essa incerteza.

Como a variabilidade e a incerteza estão presentes em todas as áreas do conhecimento, a estatística é uma ciência crucial para resolver uma série de problemas com uso de metodologias científicas apropriadas a partir da análise dos dados coletados.

A utilização de técnicas, destinadas à análise de situações complexas ou não, tem aumentado e faz parte do nosso cotidiano. No passado, tratar uma grande massa de números era uma tarefa custosa e cansativa, que exigia horas de trabalho tedioso. Recentemente, no entanto, grandes quantidades de informações podem ser analisadas rapidamente com um computador pessoal e programas adequados. Desta forma o computador contribui positivamente na difusão e uso de métodos estatísticos. Por outro lado, o computador possibilita uma automação que pode levar um indivíduo sem preparo específico a utilizar técnicas inadequadas para resolver um dado problema. Assim, é necessária a compreensão dos conceitos básicos da Estatística, bem como as suposições necessárias para o seu uso de forma criteriosa.

História da estatística

História da Estatística no mundo

Contar, enumerar e recensear sempre foram preocupações nas mais antigas civilizações, tanto do ponto de vista econômico como do ponto de vista social. Os Imperadores, ou Governantes, ordenavam recenseamentos com o intuito de conhecer sua população para assim realizar cobranças de impostos e para o recrutamento militar.

A estatística de massa iniciou-se nos grandes Impérios da Antiguidade, como Grécia, Roma, Egito, Índia e China, entre outros, tendo como principal objetivo a administração dos bens, homens, armas e obras públicas do Estado. Os registros históricos mais antigos indicam que o primeiro censo foi realizado em 2238 a.C. pelo primeiro Imperador da China. Outras atividades que sempre despertaram a curiosidade foram os jogos de azar. Os jogos não foram objetos de estudo até a Idade Média. A abordagem matemática do

azar (ou acaso) iniciou-se há cerca de 500 anos atrás, dando início à Teoria das Probabilidades, com as tentativas de quantificação dos riscos dos seguros e da avaliação das possibilidades de se ganhar em jogos de azar.

A partir do século XVIII a estatística começa a caminhar para a ciência que conhecemos hoje. Foi quando começou a ocorrer a ligação entre probabilidade e os conhecimentos estatísticos, surgindo a Inferência Estatística. É nesta época também que originou-se a palavra “estatística” e o desenvolvimento da demografia.

Na segunda década do século XIX e principalmente no início do século XX, acelera-se o desenvolvimento da estatística, tendo como principal responsável, Sir Ronald A. Fisher, conhecido como o “Pai” da estatística moderna. É na primeira metade do século XX que desenvolveu-se e sedimentou-se a grande parte da metodologia estatística, desde as bases axiomáticas das probabilidades, passando pela inferência estatística clássica e bayesiana, análise de regressão, delineamentos e análise de experimentos, análise multivariada, de sobrevivência, análise não-paramétrica e de séries temporais, consolidando-se aplicações importantes nas áreas biológicas, agrárias, industriais, econômicas, além dos levantamentos populacionais.

A maior revolução nesta ciência, ocorreu por volta de 1970, mudando o foco da estatística para sempre: o rápido desenvolvimento e disponibilidade dos computadores e programas aplicativos mudou completamente o significado da realização de uma análise estatística. Influenciou também na facilidade com que cientistas e profissionais podem coletar e armazenar dados, surgindo uma nova era, com inúmeras possibilidades para a implementação de novas e antigas idéias, a partir de abordagens em maiores escalas e soluções computacionalmente intensivas. Como consequência, no final do século XX, foi possível uma explosão de novas possibilidades em termos de metodologias e aplicações, crescendo o grau de interdisciplinaridade.

História da Estatística no Brasil

No Brasil, a Estatística tem sua história associada à história do Instituto Brasileiro de Geografia e Estatística – IBGE. A Diretoria Geral de Estatística, criada em agosto de 1872, data do “primeiro Recenseamento Geral do Império do Brasil”, foi quem primeiro coordenou e sistematizou atividades ligadas a levantamentos censitários. No período anterior a esta data (1750 - 1872), a Coroa Portuguesa determinava levantamentos populacionais, realizados precariamente, com o objetivo maior de “conhecer a população livre e adulta apta a ser usada na defesa do território”. A partir da segunda metade do século XIX, esses levantamentos passaram a ser realizados por juízes de paz

e chefes de polícia dos municípios, mas com fins eleitorais, constituindo-se as paróquias, as bases para as informações.

Com o advento da República, a produção das estatísticas dispersou-se nos âmbitos Federal, Estadual e Municipal, quase impossibilitando a unificação dos resultados e dificultando as análises. Em 1907 foi criado o Conselho Superior de Estatística, com o objetivo de padronizar conceitos e apurar resultados em todo o território nacional.

Em 1934, foi criado o Instituto Nacional de Estatística, que passou a existir de fato apenas em 1936, mudando seu nome em 1938 para Instituto Brasileiro de Geografia e Estatística, no qual ficaram vinculados os serviços geográficos. Hoje o IBGE é chamado de Fundação Instituto Brasileiro de Geografia e Estatística, sendo integrante da Administração Federal, subordinado diretamente à Secretaria de Planejamento e Coordenação Geral da Presidência da República. O IBGE coordena o Sistema Estatístico Nacional - SEN, tendo a seu encargo a orientação, a coordenação e o desenvolvimento, em todo o Território Nacional, das atividades técnicas do Plano Geral de Informações Estatísticas e Geográficas - PGIE.

O primeiro curso de Inferência Estatística dado no Brasil, ocorreu em 1947, baseado no livro de Cramer, muito embora somente em 1953 duas Escolas iniciaram o ensino de Estatística no Brasil: a Escola Nacional de Ciências Estatísticas - ENCE, criada pelo IBGE com vistas a contribuir para o cumprimento de sua missão institucional e a Escola de Estatística da Bahia mantida pela Fundação Visconde de Cairú.

Em julho de 1955, Sir Ronald Aylmer Fisher, veio ao Brasil em companhia de C. Radhakrishna Rao, para participar do 2º Congresso Internacional de Biometria, realizado em Campinas.

Em 1961, Jerzy Neyman permanece por um mês em São Paulo, onde propôs a criação de um Departamento de Estatística na Universidade de São Paulo – USP. A sua proposta foi concretizada onze anos depois, em 1972, com a criação do Departamento de Estatística e o Curso de Bacharelado em Estatística, formando sua primeira turma em 1975. A finalidade básica do Bacharelado era “formar o profissional em Estatística para atuar junto a empresas públicas e privadas ou que pretendam prosseguir estudos acadêmicos nos cursos de pós-graduação”.

Em 1970, o Instituto de Matemática Pura e Aplicada - IMPA (Rio de Janeiro), a Universidade Estadual de Campinas e a Universidade Federal do Rio de Janeiro iniciaram a formação de grupos de pesquisa em probabilidade, constituindo-se num dos grandes passos para a criação de outros cursos nessa

área.

Definições de estatística

A noção de “Estatística” foi originalmente derivada da mesma raiz da palavra “Estado”, já que foi a função tradicional de governos centrais no sentido de armazenar registros da população, nascimentos e mortes, produção das lavouras, taxas e muitas outras espécies de informação e atividades. A contagem e mensuração dessas quantidades gera todos os tipos de dados numéricos que são úteis para o desenvolvimento de muitos tipos de funções governamentais e formulação de políticas públicas.

Dados numéricos são de fato uma parte da Estatística, mas são apenas a matéria-prima, que precisa ser transformada pelos “métodos estatísticos” para posterior análise. A Estatística, como um método científico, refere-se ao projeto de experimentos e a descrição e interpretação de observações que são feitas. De um ponto de vista moderno, a Estatística é freqüentemente definida como um método de tomada de decisão em face da aleatoriedade dos fenômenos. Em uma mais vasta perspectiva, o escopo da estatística pode ser pensado em termos de três áreas diferentes de estudos: (1) a Estatística Descritiva (2) A Estatística Indutiva e (3) A Teoria da Decisão Estatística.

AQUI TENTAR COLOCAR UMA CHAVE MOSTRANDO A ESTATÍSTICA DESCRITIVA E UMA COM INFERÊNCIA E SUAS TÉCNICAS IGUAL TENHO A CANETA NO MATERIAL DA DISCIPLINA.

A estatística é um conjunto de técnicas que permite, de forma sistemática, organizar, descrever, analisar e interpretar dados oriundos de estudos ou experimentos, realizados em qualquer área do conhecimento.

A estatística é uma coleção de métodos para planejar experimentos, obter dados e organizá-los, resumi-los, analisá-los, interpretá-los e deles extrair conclusões (Triola, 1998).

“Estatística é a Ciência que permite obter conclusões a partir de dados” (Paul Velleman).

Estatística: Transforma dado em informação, ou seja, implica que seu emprego se dá em qualquer área. Logo há oportunidade de trabalho em qualquer área. Por exemplo, trabalha na melhoria ou desenvolvimento de novos produtos, para isso avalia variáveis.

Atuação do Estatístico

A Regulamentação da Profissão de Estatístico foi publicada no Diário Oficial da União (Seção I – Parte I) de 05 de abril de 1968, pág. 2.802/5. O registro profissional, obrigatório a todo Estatístico, de acordo com o disposto no artigo 2o Da Lei no. 4.739, de 1965, deve ser feito junto aos Conselhos Regionais de Estatística, CONRE, que são por sua vez organizados pelo Conselho Federal de Estatística, CONFE. Segundo a regulamentação oficializada no Decreto no 62.497, de 1o de abril de 1968, Capítulo III, Art. 3o, as atividades da profissão são definidas como:

- Planejar e dirigir a execução de pesquisas ou levantamentos estatísticos;
- Planejar e dirigir os trabalhos de controle estatístico de produção e de qualidade;
- Efetuar pesquisas e análises estatísticas;
- Elaborar padronizações estatísticas;
- Efetuar perícias em matéria de estatística e assinar os laudos respectivos;
- Emitir pareceres no campo da estatística;
- Assessorar e dirigir órgãos e seções de estatística;
- Escrever os livros de registro ou de controle estatístico criados em lei.

As principais áreas de atuação no mercado de trabalho são: Instituições Públicas como IBGE, EMBRAPA, Ministérios, Bancos e Seguradoras, Operadoras de plano de Saúde, Mercado Financeiro, Indústrias, Hospitais e Instituições de Pesquisa na área médica e biológica, Universidades, Centros de Pesquisa, entre outros. A direção de uma empresa de qualquer tipo, incluindo as estatais e governamentais, exige de seu administrador a importante tarefa de tomar decisões, e o conhecimento e o uso da Estatística facilitarão seu tríplice trabalho de organizar, dirigir e controlar a empresa. Por meio de sondagem, de coleta de dados e de recenseamento de opiniões, podem conhecer a realidade social, os recursos naturais, humanos e financeiros disponíveis, as expectativas da comunidade sobre a empresa, estabelecer suas metas, seus objetivos com maior possibilidade de serem alcançados a curto, médio ou longo prazo. A Estatística ajudará em tal trabalho, como também

na seleção e organização da estratégia a ser a dotada no empreendimento e, ainda, na escolha das técnicas de verificação e avaliação da qualidade e da quantidade do produto e mesmo das possíveis lucros e/ou perdas.

De forma resumida, o estatístico é o profissional responsável pela indicação do melhor modelo de previsão, pela modificação de desempenho de tarefas executadas por diversos setores apontando caminhos mais viáveis e práticos. Na leitura de um banco de dados é capaz de minerar os dados para uma tomada de decisão mais assertiva. Na área de seguro/planos de saúde indica modelos de custo/preços dos seguros que garantem o equilíbrio financeiro das seguradoras, elabora pesquisas dos serviços ofertados, revisa contratos firmados evitando prejuízos, entre outras atividades. Em órgãos públicos, participa de todas etapas de levantamento de dados para tomadas de decisões de investimentos, melhorias, destino de verbas a locais mais necessários, verifica regiões que ocorrem mais crimes, identifica causas e propõe ações para eliminar ou sanar os problemas, além de atuar em diversos e amplos trabalhos de grande responsabilidade. Na área de saúde, pode participar tanto de testes de novos medicamentos como indicar caminhos para prevenir o aumento de custos das patologias no mais diversos aspectos. Em bancos, pode atuar no suporte de modelos econômicos, caracterizar os clientes bancários em diferentes formas fornecendo ao gestor o melhor meio de se relacionar com o clientes. Em lojas de varejo, o estatístico é capaz de sugerir promoções para atrair o cliente pois consegue determinar o perfil de consumo do mesmo, aumentando dessa forma o faturamento da empresa. Na área de marketing é capaz de determinar a melhor maneira de elaborar uma propaganda (peça publicitária) e até indicar os melhores pontos de mercados para atuação da empresa e melhores nichos de mercado comprador do produto. Em uma indústria, o estatístico aperfeiçoará os processos produtivos, irá medir entre outras variáveis a produtividade de uma máquina por hora, irá avaliar a qualidade dos produtos, checará as informações sobre itens fabricados com defeitos, entre outros.

Fato é que daqui em diante e cada vez mais existirá uma demanda por profissionais que apontem caminhos para aumentar os lucros das empresas. Um profissional que cruze uma pilha de papéis, documentos, gráficos, banco de dados de clientes, fornecedores, negócios e gere informações que possibilitem a otimização dos processos da empresa, que ofereça menores riscos aos investimentos e aumente o ganho de produtividade. Esse profissional é o Estatístico.

Algumas definições

População e Amostra

Chama-se população ou universo estatístico ao conjunto de todos os elementos que têm pelo menos uma característica comum. A população pode ser finita ou infinita.

São exemplo de populações finitas:

- Os alunos que estudam matemática;
- Os Educadores de Infância;
- Os jogadores profissionais de futebol;
- etc.

São exemplo de populações infinitas:

- Os resultados obtidos (cara ou coroa) em sucessivos lançamentos de uma moeda;
- A temperatura em cada ponto da cidade de Bragança;
- Os pontos de uma reta;
- etc.

Um método para definir se uma população, de tamanho N , é finita ou infinita também pode ser obtido em função do tamanho da amostra (n). Quando a relação $\frac{n}{N} < 0.05$ a população é considerada infinita.

Chama-se unidade estatística a cada elemento da população.

Chama-se amostra a um subconjunto finito da população. Razões para estudar uma amostra:

- * econômicas;
- * comodidade;
- * tempo;

Existem técnicas científicas para a seleção correta de amostras. Dentre essas técnicas, as mais conhecidas são a amostragem aleatória simples, a amostragem sistemática e a amostragem estratificada. Estas são probabilísticas.

Citar técnicas não probabilística ...

Colocar o exemplo do peso em libras (115, 135, 185 e 205) e construir a distribuição da média para a técnica de amostragem aleatória simples e estratificada. Mostrar a margem de erro e calcular uma probabilidade do erro ser menor do que 15, por exemplo. Citar o que é o erro, diferença do estimador com parâmetro, e grau de confiança. Dar o exemplo de pesquisas políticas.

Comentar o que é parâmetro, o que é estimador, o que é estimativa. Depois precisão de estimadores.

Exemplos do dia a dia de amostra: experimentar uma sopa tomando uma pequena quantidade em uma colher; retirada de sangue, fezes ou urina para análise; ao abrir o chuveiro tomar percepção se a água está fria ou quente. Antes de entrar em uma piscina, colocar o pé na água. Chamar a atenção que em todos os casos uma pequena parcela foi analisada para concluir para o todo.

População versus População alvo

População objetivo(Alvo): é formada pelo conjunto de indivíduos (ou elementos) que queremos abranger em nosso estudo e para os quais desejamos que as conclusões da pesquisa. Os indivíduos da população têm pelo menos uma característica em comum.

População amostral(população acessível ou população): conjunto de indivíduos da população que estão de fato acessível para serem amostrados.

Exemplo: Supor o estudo sobre a ocorrência de sobrepeso em crianças de 7 a 12 anos no Município de São Paulo.

População alvo – todas as crianças nesta faixa etária deste município.

População de estudo – crianças matriculadas em escolas.

Unidades amostrais: são unidades de análise; podem ser pessoas, domicílios, escolas, creches, células ou qualquer outra unidade. Uma unidade amostral é formada dos elementos da população ou por grupos que contém os elementos da população.

Amostra: é uma parte da população de estudo.

Amostragem: processo para obtenção de uma amostra. Tem como objetivo estimar parâmetros populacionais.

Parâmetro: Quantidade fixa de uma população.

Ex: peso médio ao nascer de crianças que nascem no município de São Paulo ($\mu = 3100$ g); Proporção de crianças de 7 a 12 anos classificadas como obesas, no município de São Paulo ($\pi = 12\%$).

Estimador: é uma fórmula matemática que permite calcular um valor (estimador por ponto) ou com um conjunto de valores (estimador por intervalo) para um parâmetro.

XERCÍCIOS DE POPULAÇÃO, COLOCAR ...

Variáveis

Definição geral de variável: Do latim variabilis, uma variável é aquilo que varia ou pode variar. Trata-se de algo instável, inconstante e sujeito a alterações. Por outras palavras, uma variável é um símbolo que representa um elemento não especificado de um determinado conjunto. Este conjunto é denominado conjunto universal da variável ou universo da variável, e cada elemento do conjunto é um valor da variável.

Uma variável é um elemento de uma fórmula, de uma proposição ou de um algoritmo, podendo ser substituído ou podendo adquirir um valor qualquer dentro do seu universo. Os valores de uma variável podem ser definidos dentro de um intervalo ou estar limitados por condições de pertença.

Na matemática ...

No âmbito da programação (informática), as variáveis são estruturas de dados que podem mudar de conteúdo ao longo da execução de um programa. Estas estruturas correspondem a uma área armazenada na memória principal do computador.

Na astronomia, as estrelas **variáveis** são aquelas que sofrem variações significativas de luminosidade. Aqui a palavra variável é uma qualidade da estrela.

Definição de variável na estatística: é a característica de interesse que é medida **em cada elemento** da amostra ou da população. Como o nome diz, seus valores variam **de elemento para elemento**. As variáveis podem ter valores numéricos ou não numéricos.

Na estatística, as variáveis podem ser classificadas da seguinte forma:

1. Variáveis Quantitativas: são as características que podem ser medidas em uma escala quantitativa, ou seja, apresentam valores numéricos que fazem sentido. Podem ser contínuas ou discretas.
 - Variáveis discretas: características mensuráveis que podem assumir apenas um número finito ou infinito contável de valores

e, assim, somente fazem sentido valores inteiros. Geralmente são o resultado de contagens. Exemplos: número de filhos, número de bactérias por litro de leite, número de cigarros fumados por dia.

- Variáveis contínuas, características mensuráveis que assumem valores em uma escala contínua (na reta real), para as quais valores fracionais fazem sentido. Usualmente devem ser medidas através de algum instrumento. Exemplos: peso (balança), altura (régua), tempo (relógio), pressão arterial, idade.
2. Variáveis Qualitativas (ou categóricas): são as características que não possuem valores quantitativos, mas, ao contrário, são definidas por várias categorias, ou seja, representam uma classificação dos indivíduos. Podem ser nominais ou ordinais.
- Variáveis nominais: não existe ordenação dentre as categorias. Exemplos: sexo, cor dos olhos, fumante/não fumante, doente/sadio.
 - Variáveis ordinais: existe uma ordenação entre as categorias. Exemplos: escolaridade (1º, 2º, 3º graus), estágio da doença (inicial, intermediário, terminal).

Observações:

1. Os valores de uma variável podem depender dos valores de outra variável, neste caso a chamamos de variável dependente. Quando uma determinada variável não sofre influência dos valores de outra variável, dizemos que a primeira variável é independente da outra variável. Observação, uma variável qualitativa pode exercer influência em uma variável do tipo qualitativa ou quantitativa, e o mesmo vale para a variável quantitativa.
2. Uma variável originalmente quantitativa pode ser coletada de forma qualitativa. Por exemplo, a variável idade, medida em anos completos, é quantitativa (contínua); mas, se for informada apenas a faixa etária (0 a 5 anos, 6 a 10 anos, etc...), é qualitativa (ordinal). Outro exemplo é o peso dos lutadores de boxe, uma variável quantitativa (contínua) se trabalharmos com o valor obtido na balança, mas qualitativa (ordinal) se o classificarmos nas categorias do boxe (peso-pena, peso-leve, peso-pesado, etc.).

3. Outro ponto importante é que nem sempre uma variável representada por números é quantitativa. O número do telefone de uma pessoa, o número da casa, o número de sua identidade. Às vezes o sexo do indivíduo é registrado na planilha de dados como 1 se macho e 2 se fêmea, por exemplo. Isto não significa que a variável sexo passou a ser quantitativa!

Níveis de mensuração das variáveis

COLOCAR EXEMPLO DE CADA TIPO DE NÍVEL DE MENSURAÇÃO

Existem níveis crescentes de mensuração. As operações admissíveis para um conjunto numérico de dados experimentais dependem do nível ou escala de mensuração alcançada por tal conjunto que, em ordem crescente são apresentados a seguir:

- Escala nominal: Quando classes ou símbolos são usados para identificar os grupos a que vários objetos pertencem, essas classes, não ordenadas, constituem uma escala nominal.
- Escala ordinal: Quando, além de classificar as unidades de acordo com as classes, a mensuração permite ordenar essas classes relativamente ao grau de classificação da variável, atinge-se o nível seguinte de mensuração denominado de escala ordinal.
- Escala intervalar: Atinge-se esse nível quando além de ordenar as classes de uma variável, podemos dizer quanto valem exatamente as diferenças entre as classes.
- Escala das razões: A escala da mensuração mais elevada, com origem zero não arbitrária e onde é possível a realização de todas as operações aritméticas.

Definição de variável aleatória (utilizada a partir da probabilidade em diante): Dado um experimento aleatório, descrito pelo espaço de probabilidades (Ω, \mathcal{E}, P) , uma função numérica $X : \Omega \rightarrow \mathbb{R}$ será dita uma variável aleatória (do experimento). De forma simples, é uma função que associa os elementos do espaço amostral aos números reais.

Exemplo: E: lançamento de um dado. X = número obtido na face superior é uma variável aleatória. Mais precisamente, $X : \Omega = \{1, 2, \dots, 6\} \rightarrow \mathbb{R}$ tal que $X(\omega) = \omega$ é uma função numérica do experimento, e logo é uma variável aleatória.

EXERCÍCIOS DE VARIÁVEIS, COLOCAR ...

Algumas citações sobre estatística

- "Há três tipos de mentiras: mentiras, mentiras descabeladas, e estatísticas.". Benjamin Disraeli (primeiro-ministro britânico, falecido em 1881)
- "No futuro, o pensamento estatístico será tão necessário para a cidadania eficiente como saber ler e escrever.". H.G. Wells (escritor, autor de "A Guerra dos Mundos" e "A Máquina do Tempo")
- "Se há duas pessoas e duas laranjas e uma das pessoas comeu ambas as laranjas, na média cada pessoa comeu uma laranja."
- "Use a estatística como o bêbado utiliza o poste, mais pelo apoio do que pela luz."
- "Estatística, a menos exata das ciências exatas!"
- Apaixone-se sempre pelos seus dados. Nunca pelo seu modelo!
- "Torture seus dados até que eles confessem o que você deseja comprovar"
- "Dados são como filhos: você não sabe no que eles vão se tornar"
- "As Estatísticas não mentem, mas existem muitos mentirosos fazendo Estatística."

Pensamento Estatístico

<http://estatbr.blogspot.com.br/2007/11/o-pensamento-estatstico.html>

Pensamento Estatístico: Necessidade Futura. Veja a famosa frase de H.G. Wells: "No futuro, o pensamento estatístico será tão necessário para a cidadania eficiente como saber ler e escrever." (Statistical thinking will one day be as necessary for efficient citizenship as the ability to read and write)

Interessante notar que, embora Wells tenha dito esta frase há mais de 70 anos, continuamos a citá-la como sendo um fato que ainda haveremos de ver. Veremos mesmo?

O que é o Pensamento Estatístico? Como definir? Todo estatístico define da mesma forma? Não é na faculdade que devemos aprender a ter um raciocínio estatístico. Wells colocou o ensino do Pensamento Estatístico no mesmo instante da Leitura e da Escrita, ou seja, o cidadão deve crescer aprendendo estas três habilidades concomitantemente.

Para o exercício pleno da cidadania, ou seja, para que o cidadão seja capaz participar da sociedade mais efetivamente, é preciso que ele tenha condições de analisar uma série de dados informados pelos muitos órgãos (oficiais ou não) e polir o seu senso crítico.

A grande maioria da população não tem condições de analisar nada: engole notícias (isso quando as recebe) mas não forma opiniões. Como é possível ser cidadão?

Uma forma (por Doris Satie Fontes) de definir Pensamento Estatístico é a capacidade de analisar um conjunto de informações, obtido através de experiência própria, por leitura, ouvindo de alguém, etc., e poder tomar uma decisão. Cada um deve poder raciocinar dentro de sua capacidade, dos mais leigos aos mais estudiosos.

Se levarmos em conta que muitas decisões devem ser tomadas lendo informações numéricas, é muito importante que o cidadão conheça algumas ferramentas matemáticas para processá-las.

Pensando assim, é essencial melhorarmos o ENSINO DA MATEMÁTICA desde a pré-escola E introduzir consistentemente conceitos ESTATÍSTICOS às crianças, desde muito jovem.

Decidir diante de incertezas e de informações diversas é uma arte. Quanto mais preparados estivermos, melhores as chances de acerto. Quanto mais complexa for a decisão a ser tomada, mais informações precisaremos e melhores ferramentas estatísticas precisaremos. Seguindo a mesma linha de raciocínio, quanto mais investimento houver no ensino da matemática e do pensamento estatístico desde criança, maior será a sua capacidade de absorver técnicas mais sofisticadas de análises estatísticas.

Como ensinar Pensamento Estatístico?

“Sorte ou Azar”. Se eu lançar um dado diante de vocês inúmeras vezes, suponha umas 50 vezes, e o número 3 sair em 40 das vezes lançadas, o que vocês diriam? Ok, algo normal ou que o dado está com algum problema? Possivelmente devem concluir que o dado está viciado por sair um grande número de vezes o número 3, mas por qual razão chegariam nessa conclusão? O que os leva a concluir por esse caminho? Vejam que a ideia é muito simples, mas suscitou possivelmente uma série de raciocínios novos a partir

das observações: 1) um número tem saído muito mais que outros: isso é normal? 2) a partir de quantas “coincidências” (sair o mesmo número) posso declarar que o dado é viciado?

Aos alunos do Ensino Médio costumo perguntar sobre decisões que tomamos no dia-a-dia:

“Amanhã vou sair de casa 15 minutos mais cedo para não chegar atrasada na escola.” Como é que eu decidi isso? Adivinhei que o trânsito estará pior?

“Vou pegar um outro caminho alternativo hoje.” Ontem peguei o caminho A e o trânsito estava ruim. Anteontem peguei o B e estava igualmente ruim mas era uma segunda-feira. Devo tentar o B novamente (já que hoje é quarta-feira), ou tentar um terceiro caminho alternativo?

Nas duas situações acima, nada mais fizemos que raciocinar diante de fatos conhecidos, buscando correlacionar informações, ponderando-as segundo algum critério que julgamos importante e, finalmente, tomamos a decisão. Se acertamos, repetimos a fórmula. Se erramos, buscamos outras soluções. Melhoramos nossos processos.

O pensamento estatístico – não se resume em – é também aplicar corretamente as teorias da probabilidade, ou traduzir técnicas estatísticas em aplicações compreensíveis.

Informação é poder. Mas a informação não digerida não vale nada. Já o cidadão que não sabe nada...

Abordagens para a estatística: Clássica/Tradicional e Bayesiana

A estatística clássica associa probabilidades apenas a variáveis aleatórias, enquanto a bayesiana permite a associação de probabilidade a qualquer grau de crença ou incerteza sobre uma quantidade aleatória, evento ou hipótese.

A abordagem bayesiana leva à definição de probabilidades à priori e à posteriori. A primeira advém das informações que cada indivíduo traz consigo de experiências passadas. Já a outra é obtida à medida que experimentos são realizados e novos dados coletados. Assim, as probabilidades associadas a um evento são definidas previamente e atualizadas na medida em que se recebem novas informações. Quando os dados são mais informativos que as suposições iniciais, ou seja, quando a amostra aumenta, as probabilidades convergem independentes das distribuições iniciais.

Em defesa da estatística bayesiana: Descreva um experimento físico que não envolva a utilização de conhecimento a priori em sua análise de dados.

Uma Breve História

1. Estatística Bayesiana é assim chamada por ter sido elaborada por Thomas Bayes (1763).
2. Laplace (1812) aplicou esta abordagem a problemas práticos.
3. A abordagem frequentista foi desenvolvida entre 1850 e 1950 por Fisher, Pearson, Neyman e outros, tornando-se o método dominante.
4. Métodos Bayesianos passaram a ser mais utilizados devido às suas vantagens e desenvolvimento de computadores.
5. Há diferenças nos conceitos e métodos das abordagens frequentistas e Bayesianas.

Efrom (2005): Século 19 como essencialmente Bayesiano, o Século 20 como essencialmente frequentista e o Século 21 como uma combinação ideias frequentistas e bayesianas para resolver problemas com grandes conjuntos de dados. Qual paradigma usar no ensino de cursos básicos de Estatística?

Existem argumentos de todos os lados. Bayesianos dizem que algumas metodologias clássicas violam o princípio da verossimilhança. Os clássicos alegam que as metodologias bayesianas podem ser usadas para provar qualquer coisa (e.g., existência de entidades metafísicas). O que podemos afirmar, na verdade, é que ambas abordagens nos trazem coisas boas e ruins. O importante é que a abordagem escolhida seja utilizada de forma consciente e honesta.

INÍCIO CAIXA: Diferença entre estatística Bayesiana e Frequentista

Se você ainda não sacou a diferença entre os paradigmas, segue um comentário.

A Bayesian is one who, vaguely expecting a horse, and catching a glimpse of a donkey, strongly believes he has seen a mule.

Em português, ficaria algo como:

→ Um adepto da estatística Bayesiana é aquele que, vagamente esperando um cavalo, e vislumbrando um burro, acredita fortemente que viu uma mula.

Eu vi essa frase enquanto olhava o twitter do John Cook, mais especificamente a conta Stat Fact.

Mas não demorou muito e ele recebeu uma resposta:

A Frequentist, hoping to tell between a horse and a donkey, and glimpsing a mule, concludes he may as well have seen a horse.

Que continuando com minha tradução se utilizando de meu inglês nórdico é:

→ Um adepto da estatística frequentista é aquele que, querendo diferenciar entre um cavalo e um burro, ao vislumbrar uma mula, pode também concluir que viu um cavalo. FIM CAIXA

Estatística e Computação

- Avanços em Estatística diretamente relacionados com avanços na área computacional.
- 1960: máquinas de calcular manuais, elétricas, eletrônicas.
- 1960 → 1980: "grandes computadores": IBM 1620, CDC 360, VAX etc; cartões e discos magnéticos; FORTRAN.
- 1980 →: computadores pessoais; supercomputadores; computação paralela; "clouds"; C, C+, S.
- Pacotes estatísticos: S-Plus, SPSS, Minitab etc. Repositório R.
- Data Mining: redes neurais, support vector machines.
- Era do "Big Data": Análise de *microarrays* em Bioinformática; Dados de alta frequência em finanças; taxas de câmbio e taxa de juros; Dados meteorológicos, oceanográficos e astronômicos; Simulações.

HABILIDADES E COMPETÊNCIAS PROFISSIONAIS REQUERIDA DE UM ESTATÍSTICO

De acordo com a Lei no 4.739, de 15 de julho de 1965, que disciplina a profissão do Estatístico, bem como contempla a proposta da Comissão de Especialistas de Ensino de Matemática e Estatística para as diretrizes curriculares para cursos de estatística no país e tendo em vista o perfil desejado para essa profissão, o Curso de Estatística deve priorizar o desenvolvimento das seguintes competências e habilidades:

1. capacidade crítica para analisar os conhecimentos adquiridos e assimilar novos conhecimentos científicos e/ou tecnológicos, necessários à atuação em equipe multidisciplinar;
2. capacidade de expressão e de comunicação;

3. habilidade numérica; raciocínio lógico e atenção concentrada; exatidão e rapidez de cálculo; meticulosidade e facilidade para associar, deduzir e sintetizar;
4. dominar a aplicação da matemática a problemas concretos;
5. conhecimento das formas de medição das variáveis de sua área de atuação e de organização e manipulação dos dados;
6. domínio da linguagem técnica estatística, aliada à capacidade de adequação dessa linguagem à interação com profissionais das diversas áreas do conhecimento e à comunicação com a sociedade;
7. dominar métodos e técnicas de informática, especialmente no tocante à utilização de softwares estatísticos;
8. saber produzir sínteses numéricas e gráficas dos dados, através da construção de índices, mapas e gráficos;
9. saber usar os modelos básicos de análise estatística, de forma especial os modelos de associação entre uma variável resposta e um conjunto de variáveis explicativas;
10. ser capaz de, a partir da análise dos dados, sugerir mudanças no processo, na política pública, na instituição, etc.;
11. dominar, ao menos em nível de leitura, uma língua estrangeira, preferencialmente, a inglesa;
12. capacidade para desenvolver ações de natureza empreendedoras e gerenciais.

O Excel e os softwares R, SAS, SPSS, SPLUS, MINITAB e outros

Todos estes fazem de alguma forma tudo o que é visto em estatística descritiva. Quando trabalha-se com amostra, para inferir, depende da ferramenta utilizada e o módulo ou tipo de licença comprada para que o software a realize.

Falando de Tabelas, gráficos e medidas.

Mostrar um conjunto de dados e ligá-lo a um questionário. Separar e recordar quais variáveis são do tipo qualitativa e quais são do tipo quantitativa. Dizer que as tabelas e gráficos são distintas para os dois tipos de variáveis.

Iniciando com o Excel. Diversas formas de fazer uma tabela. Mostrar como se faz uma forma, os tipos de operações, a função cont.se, comparar com a tabela dinâmica.

Texto da Doris sobre o CONE 2017

Ocorreu em BH nos dias 01 e 02 de junho o CONE-2017, Congresso de Estatística, organizado pela Brasil Produções e CONRE-6, com o apoio institucional do CONFE e todos os demais CONREs, da FGV, IBGE, UFMG, UFOP e UFJF, além do patrocínio da SAS. Foram dois dias de intensos debates acerca da nossa formação, atuação profissional e mercado de trabalho tanto hoje como num futuro próximo. Resumidamente, poderia dizer o seguinte: . Nossa formação segue um currículo básico antigo e, de certa forma, ultrapassado. Embora a formação deva continuar forte em teoria estatística, fundamentos de matemática e linguagem de programação, falta ao aluno mais vivência na manipulação de grandes massas de dados e entender um pouco mais desse mundo tecnológico que envolve Big Data e sua exploração. É necessário desenvolver esse lado prático durante o curso. A maioria concorda que o bacharelado em estatística não deve seguir modismos (por exemplo, Data Mining era o procedimento da moda até o ano passado, hoje não mais). . A maioria das técnicas de hoje têm nomes diferentes, parecem complicadas e sofisticadas, mas são, no fundo, as mesmas técnicas estatísticas com roupagem tecnológica. Portanto, vale a pena prestar atenção em todas as aulas. Ex: deep learning redes neurais regressão. . Ter uma boa formação teórica dá autonomia ao estatístico para pensar novos modelos e soluções. . Os alunos precisam se mobilizar para melhorar a formação desde o começo do curso. Buscar mais optativas (mesmo que seja como ouvinte), sobretudo nas áreas de COMPUTAÇÃO (linguagem de programação) e NEGÓCIOS enquanto estiver no bacharelado. Tem que treinar comunicação oral e escrita de boa qualidade. . Não dá para ficar esperando um bom emprego aparecer só porque tem um diploma de estatístico. Se o mercado exige conhecimentos de programação, corra atrás e se qualifique. . Tomar cuidado ao buscar qualificação extra em DATA SCIENCE: há muitos cursos que têm apenas um nome interessante, mas que nada ensina ao aluno:

curso para enganar os alunos. Prestar atenção na QUALIDADE DO CORPO DOCENTE: professores com doutorado em estatística, ou com forte formação em Computação, pelo menos. Melhor do que curso presencial ruim, é fazer um curso online de qualidade: Coursera, Johns Hopkins, Udemy, etc. . Todo estatístico precisa fazer um esforço para aprender sobre o negócio da empresa em que trabalha. Houve muitas palestras muito interessantes, debates em torno das nossas carreiras e funções na sociedade. Foram dois dias excelentes. 2018 será a vez de São Paulo — CONE-2018-SP