Rev Bras Anest
1990;40:2:119-132

Elementos de Estatística

Antonio Leite Oliva Filho, TSA[1]

OLIVA FILHO AL - Statistical elements.

Nota do Autor

As informações constantes neste artigo resultam de uma compilação de informações estatísticas orientadas à medicina expostas nas referências 1,2, 3 e 4, com exemplos adaptados à anestesiologia, Desta forma, na exposição do texto apenas serão destacadas as referências para conceitos ou expressões eminentemente interpretativa.

Como os conceitos e explicações estão abordados de forma seqüencial (cada tópico dependente das explicações que o precedem),é sugerida uma leitura da mesma forma: só avançar ao tópico seguinte com o entendimento do assunto que o precedeu.

Para o Médico Clínico, independente da especialidade, a afinidade com os números não é muito comum. Durante a graduação, com um currículo essencialmente constituído de informações descritivas, inexiste estímulo a análises e interpretações quantitativas. Se tanto, o mínimo que se requer de conhecimetos numéricos diz respeito a doses de medicamentos e suas relações *com* outros parâmetros biométricos: peso, altura, freqüências, pressões ou superfície corporal (este último parcamente utilizado porque requer cálculos com fórmulas mais complexas). Alguns poucos que se encaminham à pesquisa acabam por perceber, a duras penas, que, também em Medicina, as ciências dos números, principalmente a Estatística, são essenciais para análises de informações e para inferir conclusões. Os que praticam a clínica, na leitura das revistas médicas, pela absoluta falta de contato com tais conceitos, ou se enfadam com os artigos e deles não alcançam o contingente pleno de informações, ou simplesmente relegam a descrição da metodologia e dos resultados, apegando-se às frações meramente descritivas dos trabalhos: Introdução e Conclusões. Mas, são as Conclusões aplicáveis ao dia-a-dia?

Muitos podem partir do princípio que, se os artigos foram aprovados para publicação em revistas tão importantes, suas conclusões devem estar absolutamente confiáveis, o que não é verdadeiro. Poucos são os periódicos que dispõem de assessoria estatística. Tal providência é duplamente díficil: são raros os estatísticos que dominam paralelamente os conhecimentos médicos, ou vice-versa, e a dependência desta análise tornaria o cronograma editorial extremamente oneroso e lento[1]. Algumas revisões têm sido publicadas apontando falhas grosseiras de interpretação estatística em artigos de revistas internacionalmente reputadas como inatacáveis[1].

Por estes motivos, Colton[2] aponta as seguintes justificativas para o aprendizado de noções de estatística por um clínico:

a) “A Medicina vem se tornando altamente quantitativa. Com o progresso tecnológico, os clínicos deparam-se, cada vez mais, com dados quantitativos. A Estatística é a ciência adequada para a organização e análise destes tipos de dado.”

b) “O planejamento, conduta e a interpretação da pesquisa médica valem-se freqüentemente da metodologia estatística. Também, vários dos questionamentos do clínico têm extensa conotação estatística: este novo método, ou droga, é melhor que o tradicional? Quanto melhor? Qual é o risco de efeitos colaterais associados ao seu uso? Em um ensaio clínico, quantos pacientes são necessários estudar, e de que maneira para extrapolar os resultados à prática diária? Qual é a dose adequada para esta droga? Qual é a variação de dose aceita como normal? Como interpretar insucessos? As próprias perguntas dos pacientes exigem explanação estatística: Qual é o risco, doutor?”.

c) “Como conseqüência natural do aumento de dados numéricos na Medicina e do uso de metodologia estatística para analisá-los, os artigos científicos estão repletos desta terminologia. Se o clínico lê tais artigos,

1 Membro do Serviço de Anestesiologia da Clínica de Fraturas e Ortopedia XV Ltda., em Curitiba, PR. Editor Chefe da Revista Brasileira de Anestesiologia

*Correspondência para Antonio Leite Oliva Filho
Rua Padre Anchieta, 1500- ap. 407
80430- Curitiba - PR*

necessita um mínimo de conhecimento para avaliação crítica das evidências numéricas apresentadas. Tem que conhecer como e por que usar a terminologia."

Estatística descritiva

É a parte da estatística que permite organizar e sumarizar os dados obtidos nas observações de qualquer fato, sintetizando grande número de informações em expressões numéricas simples e compreensíveis, que facilitam a sua interpretação.

Supondo a avaliação de uma nova droga de indução (droga X, na busca de uma padronização de dose, estabelece-se como critério a quantidade suficiente da droga X, em $mg.kg^{-1}$, para abolir o reflexo corneano em adultos, de ambosos sexos, estado físico ASA l, usando-se solução a 1 %, injetada, sob velocidade constante, em veia de prega do cotovelo. Após uma série de 11 observações, poder-se~a apresentar uma nota Prévia equivalente à Tabela l.

Tabela I - Apresentação tabular dos resultados encontrados na pesquisa da droga X am 11 pacientes, na ordem da experimentação

Obs. nº	Dose ($mg.kg^{-1}$)
1	5,25
2	3,75
3	4,75
4	4,50
5	4,25
6	4,25
7	3,75
8	2,50
9	3,75
10	4,00
11	3,25

Fica claro que a simples apresentação dos dados na ordem da experimentação não expressa totalmente o objetivo. Se a série de dados fosse maior, perder-se-ia muito tempo consultando toda a tabela apresentada para deduzir alguma coisa. Organizando os valores, em ordem crescente, já ficaria mais fácil (Tabela Il).

Tabela II -Apresentação tabular dos mesmos resultados na pesquisa da droga X em 11 pacientes, ordenados pela dose encontrada, em forma crescente.

Obs. nº	Dose ($mg.kg^{-1}$)
8	2,50
11	3,25
2	3,75
7	3,75
9	3,75
10	4,00
5	4,25
6	4,25
4	4,50
3	4,75
1	5,25

Nota-se, facilmente, que a dose na qual todos os pacientes atingiram o critério especificado ficou entre 2,5 e 5,25 $mg.kg^{1}$. Mas este é, ainda, um intervalo de dose muito grande para ser clinicamente adotado. Se houvesse a organização em intervalos menores de doses, ordenados por freqüência de ocorrência, simplificar-se-ia, ainda mais, a análise (tabela lll).

Tabela III - Distribuição tabular, por freqüência, dos intervalos de doses encontradas na pesquisa da droga X, em 11 pacientes

Doses ($mg.kg^{-1}$)	Freqüência
2,50 a 3,55	2
3,56 a 4,44	6
4,45 a 5,25	3

Demonstra-se uma tendência entre 3,56 e 4,44 $mg.kg^{1}$' para a dose de indução com a droga X.

Também chegar-se-ia a um valor similar (4,0 $mg.kg^{1}$) dividindo-se o somatório dos valores encontrados pelo número de observações.

A organização e síntese das informações obtidas neste exemplo constituem formas de estatística descritiva que facilitam a análise e interpretação das informações.

A estatística descritiva, no entanto, vai mais Ionge, procurando avaliar com mais detalhes as expressões numéricas que sintetizam o universo de resultados das observações. Para se compreender mais adequadamente esta síntese é necessário explicar alguns conceitos básicos utilizados em estatística

Tipos de dado

Os métodos de estatística descritiva variam com o tipo de dado observado. Em medicina manipulam-se parâmetros aos quais se pode estabelecer, de forma exata, um valor numérico: altura em metros; peso em kg; pressão em kiloPascal, mm Hg ou cm H_2O; débito em $L.h^{-1}$, doses em $mg.kg^{-1}$ etc. Também manipulam-se variáveis cujos critérios de medida são, no mínimo, dependentes de interpretação do observador, com algum grau de subjetividade: estado físico segundo a American Society of Anesthesiologists, Escala de vitalidade neonatal de Apgar; Escala de recuperação da anestesia de Aldrete e Kroulik, ou de Saraiva; grau de relaxamento muscular segundo Bromage; extensão do bloqueio peridural em número de segmentos bloqueadores etc. Dois, então, são os tipos de dados:

1) *Dados paramétricos:* grandezas escalares cujos intervalos de medida são unidades constantes. A dife-

rença entre uma pressão de 100 e outra de 101 mm Hg é exatamente igual àquela obtida entre 180 e 181 mm Hg, porque o intervalo constante entre todas as grandezas desta unidade de medida é 1 mm Hg.

2) *Dados não paramétricos:* informações classificáveis dentro de critérios apenas conceituais, ordinais ou nominais, onde não há constância nos intervalos. Mesmo que se queira, não se pode dizer que a diferença de estado físico entre dois pacientes EF ASA I e II, respectivamente, seja exatamente igual à diferença entre dois outros EF ASA IV e V. O intervalo desta escala de valores não é constante porque Estado Físico ASA é apenas uma classficação ordinal. Distribuição de ocorrência por sexo envolve dados não paramétricos dentro de critérios nominais: masculino e feminino. Extenso de bloqueio espinhal em número de metâmeros envolvidos é, também, grandeza não paramérica de ordem conceitual: o intervalo entre as medidas, um segmento metamérico, embora unitário, não constitui medida objetiva constante. Se aplicarmos uma dose de anestésico local, no espaço peridural torácico, e bloquearmos 5 segmentos, não quer dizer que, no mesmo paciente, bloquearemos 5 segmentos usando a mesma dose através do hiato sacro.

Por estas razões, a estatística descritiva de dados não paramédicos deve se restringir a tabelas de distribuição perceptual ou de freqüência.

A estatística descritiva de dados paramédicos, por seu lado, pode ser feita através da redução de todas as observações para expressões numéricas mais simples, como média, desvio padrão, variância, conceitos estes que serão apresentados na seqüência.

População, Amostra Observação aleatória, Estudo encoberto

População, em terminologia estatística, tem um significado diverso daquele aplicado em expressão coloquial. Não quer dizer exatamente pessoas ou seres animados de uma determinada região do país. Em estatística utiliza-se a palavra para designar pessoas, objetos, situações, observações, fatos como: cirurgias plásticas no Hospital Z; medidas de substâncias redutoras na urina; anestesica gerais inalatórias em Crianças de 1 a 12 anos no Hospital W. O que se deseja exprimir, de fato, quando se fala em população em estatística, é o universo possível de elementos envolvidos em uma determinada análise. Embora se possa descrever exatamente a população da qual trata uma análise, é freqüentemente difícil quantificar, exatamente, o número de elementos que estariam envolovidas. Quando se descreveu a população no exemplo drogaX, dizia-se *adultos,* de *ambos os sexos, estado físico ASA I submetidos à indução venosa* para anestesia geral. De quantos elementos esta população, de fato, se constitui, no universo de anestesia e gerais praticadas?

Como é praticamente impossível envolver toda a população abrangida pelos critérios propostos, seja pelo custo ou pelo tempo que seria dispendido, é lícito fazer a análise sobre observações de uma ou mais de suas frações.

Cada uma das frações estudadas constitui uma Amostra da População. Os cálculos de estatística descritiva foram desenvolvidos no sentido de procurar extrapolar, à população, os valores obtidos na observação da amostra.

Por isso, uma amostra deve representar criteriosamente a população alvo da análise. Cada elemento da população deve ter exatamente a mesma probabilidade de participar da amostra. Em outras palavras, seus elmentos, para serem efetivamente representativos, devem ser escolhidos ao acaso, de forma a garantir o respeito A probabilidade. O ideal portanto é uma Observação Aleatória dos casos (em inglês, *diz-serandomic study, o* que induz muitos autores nacionais a usar o anglicismo “estudo randômico” ou “randomizado”).

Observação aleatória, cientificamente, também não quer dizer ao azar. Para este tipo de seleção, normalmente segue-se uma tabela de números aleatórios, previamente preparada (Anexo I).

Nos estudos comparativos, o ideal, ainda, é que o observador desconheça a tabela de distribuição aleatória para os grupos, fazendo seus apontamentos sem saber a que grupo pertence cada paciente, para não se deixar induzir. Faz, então, um Estudo Encoberto. Diz-se Duplamente Encoberto o estudo em seres humanos, quando nem os pacientes, nem o observador, sabem a que grupo cada caso foi alocado (a expressão inglesa double *blind* gerou a tradução literal - “duplo cego”).

Se as amostras não seguem tais critérios de seleção para sua composição, toma-se obrigatória a ampliação do número de observações para ter garantida uma re presentação mais acurada da população.

Com todos estes critériosde definição da população, de seleção dos elementos da amostra, e cuidados tomados na observação, é possível inferir conclusões para a população a partir dos resultados obtidos com a amostragem.

Ainda assim, para garantir isenção, o pesquisador deve descrever seus resultados e conclusões vinculando-os à população definida de forma precisa e, eventualmente, apenas para a amostra cujos critérios de seleção deixou claramente descritos, sem qualquer pretensão de generalizar seus conceitos. :

Distribuição normal- Curva de Gauss

No exemplo, já utilizado, da droga X, poder-se-ia ter aumentado a amostra para um número bem maior de observações: 300 pacientes. Completadas estas observações e arredondadas as doses até uma casa decimal, encontrar-se-ia a distribuição exposta na Tabela IV.

Tabela IV - Distribuição, por freqüencia, das doses encontradas para indução com a droga X, em mg.kg^{-1},em 300 pacientes adultos, de ambos os sexos, estado físico ASA I

	Tabulação de freqüência (Droga X)		
Classes	Doses	Freqüência	Percentual
1	2.4	1	0,33
2	2.5	1	0,33
3	2.6	1	0,33
4	2.7	1	0,33
5	2.9	1	0,33
6	3.0	1	0,33
7	3.1	1	0,33
8	3.2	3	1,00
9	3.4	8	2,67
10	3.5	11	3,67
11	3.6	13	4,33
12	3.7	20	6,67
13	3.8	27	9,00
14	3.9	40	13,33
15	4.0	42	14,00
16	4.1	39	13,00
17	4.2	26	8,67
18	4.3	18	6,00
19	4.4	14	4,67
20	4.5	11	3,67
21	4.6	8	2,67
22	4.7	5	1,67
23	4.8	3	1,00
24	4.9	1	0,33
25	5.0	1	0,33
26	5.2	1	0,33
27	5.3	1	0,33
28	5.4	1	0,33
			100,00

$n = 300$
$x = 1200,00$
$x = 4,00$
$Var: = 0,156$
$DP = 0,39$
$EP=0,23$

4,0 mg.kg^{-1}, e uma redução progressiva nas freqüências das doses acima e abaixo daquela. Esta mesma distribuição poderá ser representada graficamente, em um histograma, como o da Fig. 1.

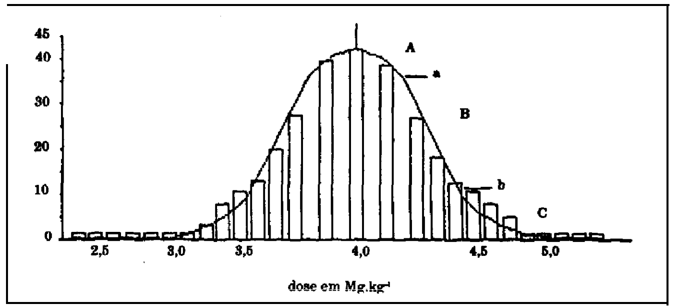

Fig. 1 Histograma das freqüências das doses de droga X, em mg.kg^{-1}, encontrada na amostra de 300 pacientes adultos, de ambos os sexos, estado físico ASA 1.A linha contínua do gráfico representa a Distribuição Normal (curva de Gauss) esperada para a população.

A projeção regular das colunas representativas das freqüências em uma linha contínua estabelece um padrão que, para ambos os lados, a partir de seu ponto mais alto, apresenta, de início, uma convexidade (A) seguida, após uma inflexão (a), de um segmento quase retilíneo (B) e, após outra inflexão (b), por uma concavidade (C) que se estende até a linha de base, com a qual se confunde nos extremos. Este padrão de curva é denominado de Curva de Distribuição Normal, ou curva de Gauss. É um padrão característico para todos os parâmetros cujos valores dependem de múltiplos fatores. Tem um ápice (cabeça) e duas caudas. O ápice, central, representa os valores mais freqüentes para o parâmetro, e as caudas, equivalentes no grau de curvatura e extensão, representam os valores progressivamente menos freqüentes.

Outros parâmetros podem seguir curvas menos regulares, com cabeças tendentes à esquerda, ou à direita, ou mesmo com cabeças múltiplas. Para efeito didático, considerar-se-ão apenas as distribuições normais.

Medidas da tendência central

Entende-se por medida da tendência central qualquer dos métodos de estatística descrita capaz de detectar o valor correspondente ao ponto central da distribuição de um parâmetro, na população, a partir de uma amostra. Três são as medidas de tendência central mais comumente utilizadas em estatistica: a moda, a mediana e a média.

Moda: denomina-se moda o valor mais freqüente encontrado para o parâmetro estudado. No exemplo didático da Tabela II, a moda foi a dose 3,75 mg.kg^{-1}, que apareceu três vezes no estudo. No exemplo da Tabela IV, a dose de 4,0 mg.kg^{-1} é a moda com 42 ocorrências (14%).

Mediana: é o valor encontrado para o parâmetro estudado, que, na distribuição, tem metade das demais freqüências à sua esquerda (menores) e a outra metade à sua direita (maiores). No exemplo da Tabela II, é a dose de 4 $mg.kg^{-1}$, que tem 5 ocorrências menores (45,45%) e 5 ocorrências maiores (45,45%). No exemplo da Tabela IV, pelo grande número de elementos da amostra, e pelo fato da distribuição ser normal, a mediana coincide com a moda (4 $mg.kg^{-1}$). Quando a curva de distribuição é assimétrica, nem sempre há coincidência entre a moda e a mediana.

Média: é a medida de tendência central mais comumente utilizada, particularmente a média aritmética. Corresponde ao resultado do somatário de todos os valores encontrados para o parâmetro estudado (Rx), dividida pelo número de elementos da amostra (n) (fórmula 1).

$$MA = \bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n} \quad \text{ou} \quad \frac{\Sigma x}{n} \qquad (1)$$

Aplicando a fórmula, nos exemplos das Tabelas II e IV, a média aritmética é 4 $mg.kg^{-1}$.

Como se viu, em grandes amostras, com distribuição normal, há a tendência em se igualarem a moda, a mediana e a média.

A média aritmética só é válida para dados paramétricos. Com dados ordinais (estado físico, Apgar, extensão de bloqueio espinhal) a utilização da moda é mais adequada, ou eventualmente a mediana, ou melhor ainda, a apresentação das freqüências encontradas para cada ordem.

Para alguns parâmetros, no lugar da média aritmética, utiliza-se a média geométrica para cálculo da tendência central, principalmente quando seus valores são apresentados em notação científica, com expoentes positivos ou negativos, como é o caso da concentração hidrogeniônica. Para obtê-la extrai-se a raiz enésima do **produto nos n valores (fórmula 2).**

$$MG = \sqrt[n]{x_1.x_2.\cdots.xn} \quad \text{ou} \quad (x_1.x_2.\cdots.x_n)^{1/n}$$

Medidas da variação (ou da dispersão)

Com o cálculo da média, é possivel saber o valor do elemento mais freqüente em uma amostra com curva de distribuição normal (ponto central). Mas só isto não basta para expressar quantitativamente detalhes da curva de distribuição esperada para a população. É necessário calcular outros valores para se quantificar a sua forma (se mais aguda ou mais obtusa). Quanto mais aguda menor a dispersão ou menor é a distância entre os vários pontos da curva com seu eixo central (a média). Se, por outro lado, a curva é achatada, os valores dos elementos estão muito dispersos, havendo maiores distâncias entre os pontos da curva e seu eixo central (Fig. 2).

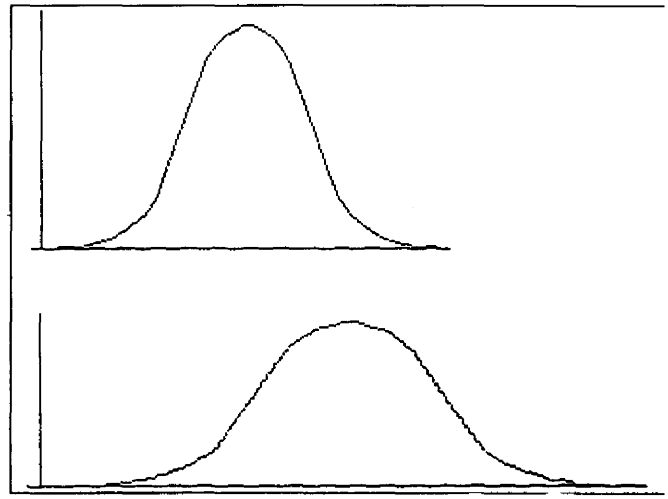

Fig. 2 Histograma e curva de Gauss estimada para duas amostras com médias idênticas, porém com distribuição diferente. A curva mais aguda denota menor variação nos valores do parâmetro estudado.

Vários são os métodos de estatística descritiva que permitem uma estimativa da dispersão (ou variação) da curva.

Como exemplo didático, imagine-se, agora, outro estudo em 11 pacientes, aos quais se aplicou a droga Y, para cálculo da dose de indução, utilizando-se os mesmos critérios apontados para o exemplo da droga X.

Extremos (Máximo e Mínimo) - são o menor e maior valores encontrados para o parâmetro, dentro da amostra. Fornecem os valores que constituem os pontos extremos da curva. No exemplo da Tabela II, seriam as doses de 2,5 (minima) e 5,25 $mg.kg^{-1}$ (máxima). Com a média, permitem avaliar o ponto central e a largura da base de cada um dos lados da curva de distribuição.

Desvio da média - diferença, em números relativos, entre o valor de um elemento da amostra e a média:

$$\text{desvio da média} = (x_i - \bar{x})$$

Permite estimar a distância entre o valor de um elemento qualquer e a média. Pode resultar em valores negativos (para elementos menores que a média) ou positivos (maiores). O 9º elemento da Tabela II dista -0,25 $mg.kg^{-1}$ da média enquanto o 1º elemento fica a +1,25 $mg.kg^{-1}$ da média.

Somatória dos quadrados dos desvios - é a soma de todas as diferenças, entre cada um dos elementos e a média, elevadas ao quadrado. Estatisticamente tem o nome simplificado para Somatória dos Quadrados *(Sum of Squares- SS).*

Como a simples somatória dos desvios tende a zero, pois há uma distribuição regular de valores positivos e negativos (Tabela V), a opção é elevá-los ao quadrado, obtendo-se a Somatória dos Quadrados, que permite quantificar a dispersão de uma amostra, numericamente.

$$SS = \sum (X-\bar{X})^2$$

É um dado bastante utilizado como base de outros cálculos estatísticos, Pode-se chegar ao seu valor, utilizando-se a fórmula 3:

$$SS = \sum x^2 - \frac{(\sum x)^2}{n} \quad (3)$$

Tabela V - Estudo comparativo entre droga X e droga Y quanto à dose de indução em mg.kg^{-1}, onde foram observados 11 pacientes estado físico ASA I. Note-se a tendência a 0 da somatória dos desvios e um valor representativo da dispersão obtido com a sormatória dos quadrados dos desvios

Droga X			Droga Y		
(mg.Kg^{-1})	x-x/	(x-x/)^2	(mg.kg^{-1})	x-x/	(x -x/)^2
2,50	-1,50	2,2500	1,50	-2,50	6,2500
3,25	-0,75	0,5625	2,75	-1,25	1,5625
3,75	-0,25	0,0625	3,00	-1,00	1,0000
3,75	-0,25	0,0625	3,00	-1,00	1,0000
3,75	-0,25	0,0625	3,00	-1,00	1,0000
4,00	0,00	0,0000	4,00	0,00	0,0000
4,25	0,25	0,0625	4,75	0,75	0,5625
4,25	0,25	0,0625	4,75	0,75	0,5625
4,50	0,50	0,2500	5,50	1,50	2,2500
4,75	0,75	0,5625	5,75	1,75	3,0625
5,25	1,25	1,5625	6,00	2,00	4,0000
∑ 44,00	0,00	5,5000	44,00	0,00	21,2500
n 11			11		
x/ 4,00			4,00		

Grau de liberdade - Se uma amostra tem apenas dois (2) elementos, é possível fazer uma (1) comparação entre eles. Se tiver 3 elementos só 2 comparações são possíveis, desde que, estatisticamente, não se utiliza fazer todas as combinações possíveis entre os elementos para comparação; assim, compara-se, apenas, o primeiro elemento com o segundo e este com o terceiro e assim sucessivamente. Desta maneira, para qualquer amostra, é possível obter *n-1* comparações entre seus elementos. Este número de comparações possíveis ente os *n* elementos de uma amostra estatisticamente, recebe o nome de Grau de Liberdade, e sua notação, em alguns livros, é a letra N. Outros autores utilizam a própria expressão n-1. Utiliza-se ainda GL ou DF *(degrees of freedom).*

$$N = n-1 \quad (4)$$

Variância (média dos quadrados dos desvios da amostra) - A distância média entre cada um dos elementos e o eixo central podem fornecer o subsídio para avaliar o grau de dispersão da cuma de distribuição. O resultado da divisão da somatória do quadrado dos desvios (SS) pelo grau de liberdade (N) fornece o valor médio dos quadrados dos desvios. O dado assim obtido recebe, em estatística, o nome de Variância. Sua notação nos manuais de estatística é Var.

$$Var = \frac{SS}{N} = \frac{\sum (x-x)^2}{n-1} \quad (5)$$

Desvio padrão - Extraindo-se a raiz quadrada da variância, obtém-se o valor médio dos desvios, que em estatística é denominado Desvio Padrão Médio, ou simplesmente Desvio Padrão *(Standard Deviation). E* um dos métodos de estatística descritiva mais comumente utilizado nos estudos de dados paramétricos.

$$DP = var \text{ ou } (var)^{1/2} \quad (6)$$

Computacionalmente, calcula-se avariância da qual se extrai a raiz quadrada para obtenção do padrão. A fórmula global para o desvio padrão, substituindo-se, na fórmula 6, os conteúdos das fórmulas 3,4 e 5, resulta em:

$$SD = \left[\frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1} \right]^{1/2}$$

Coeficiente de variação - É a relação entre o desvio padrão e a média de uma amostra que, expressada percentualmente, quantifica o grau de variação obtido nos valores do parâmetro em estudo:

$$CV = \frac{DP \times 100}{x}$$

Calculados o desvio padrão e o coeficiente de variação, para as doses da droga X e droga Y, teriamos dados mais adequados para comparar as duas amostras (Quadro I).

Quadro I - Número, média, desvio padrão, extremos e variância obtidospara as doses das drogas X e Y

	Droga X	DrogaY	
Número	11	11	casos
Média ± DP	4,00 ± 0,74	4,00 ± 1,46	$mg.kg^{-1}$
Extremos	2,50<->5,25	1,50<->6,00	
E P	0,22	0,44	
Coef. var.	18,5	36,5	%

Ficou bem mais sintética e organizada a apresentação dos reultados dos estudos das duas amostras, quando se conseguiu apresentar o número de casos estudados, a média obtida (eixo da curva de distribuição), os extremos encontrados (base da curva), o desvio padrão e o coeficiente de variação (grau dispersão). A média igual para as duas amostras sugere um eixo central igual para as duas curvas, embora o coeficiente de variação maior para a droga Y sugira que a curva é bem mais larga (achatada) para esta. Pode-se supor que a droga X tem uma dose de indução menos variável que a droga Y.

Erro padrão - É uma correção do desvio padrão (da amostra) estendendo-o a toda a população. Como a média é a representação do valor central da curva de distribuição da população, necessita-se do ajuste do desvio padrão da amostra para o mesmo alcance. E obtido com a divisão do Desvio Padrão pela raiz quadrada do número de elementos da amostra. Denomina-se, estatisticamente, Erro Padrão da Média da População, ou simplesmente, Erro Padrão *(Sfandard Error):*

$$EP = \frac{DP}{\sqrt{n}}$$

A denominação Erro não quer dizer propriamente que há erro no cálculo. É apenas para diferenciar de Desvio Padrão, uma vez que o DP corresponde à média dos desvios da amostra (desvio padrão da média da amostra) e o EP, à média dos desvios da população (desvio padrão da média da população).

É um dado muito utilizado nos cálculos dos testes estatísticos (veja Limites de Confiança e teste “t” de Student). Para o estudo com a droga X, o EP é igual a **0,22 (0,74 / $\sqrt{11}$). No estudo da droga Y, encontra-se um SE de 0,44 (1,46 / $\sqrt{11}$) (Quadro I).**

Relação entre o desvio padrão e a curva de Gauss - Limites de confiança

Calculados a média e o desvio padrão a partir de uma amostra pode-se estimar, com dois valores, a tendência central e a variação do valor do parâmetro estudado. Se a amostra tem uma distribuição normal *(gaussiana),* é lícito esperar que a maior parte dos elementos tem um valor próximo da média e uma menor parte fica mais longe dela.

Exponencialmente pode-se precisar que o primeiro ponto de inflexão de cada lado da curva de Gauss da Amostra corresponde a um valor igual à média ± 1 desvio padrão (DP). O ponto da segunda inflexão corresponde a um valor semelhante à média ± 2 DP e os extremos da curva (fim do terceiro segmento) equivalem ao valor da média ± 3 DP (Fig. 3).

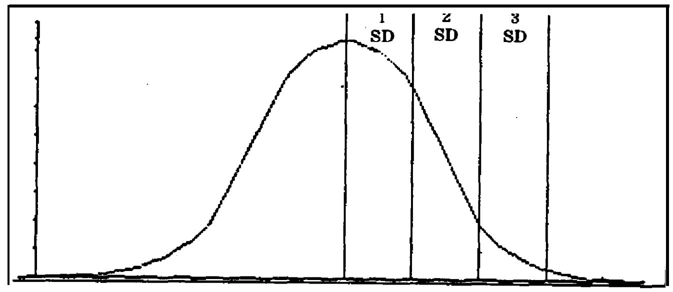

Fig.3 Curva de Gauss mostrando a relação entre os pontos de inflexão da curva e os valores de média ± 1,2 e 3 desvios padrões.

Os limites da curva de Gauss da Amostra precisados pelo valor da média ± 1 DP abrangem próximo de 66% da superfície total da curva, correspondente portanto a cerca de 66% dos valores do parâmetro (≈ 34% de cada lado da média) (Fig. 4A). A média ± 2 DP delimita próximo de 95% da superfície (≈ 47,5% de cada lado da média) (Fig. 4B). A média ± 3 DP cobre próximo de 99% da superfície (≈ 49,5% de cada lado) (Fig. 4C).

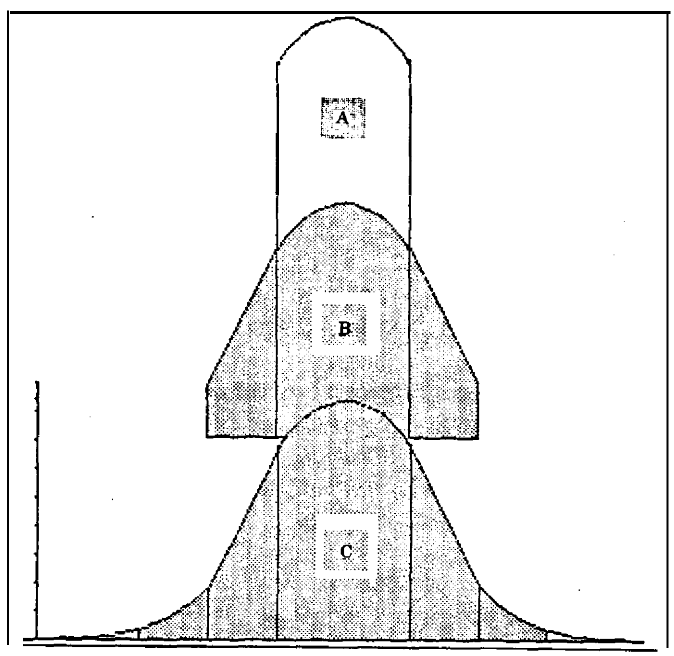

Fig. 4A Superfície coberta pela média ± 1 DP, equivalente a 68%. B - Superfície coberta pela média ± 2 DP, equivalente a 95%. C- Superfície coberta pela média ± 3 DP, equivalente a 99%, A média ± 2 DP corresponde a, exatamente, 95,44%. Portanto, 95% é a superfície abrangida, de fato, pelos valores correspondentes à média ± 1,96 DP.

Da mesma forma a média ± 3 DP corresponde a, exatamente, 99,7% da superfície da curva. logo, 99% da superfície, equivale, de fato, à média ± 2,58 DP.

Limite de Confiaça - Extrapolando o mesmo raciocínio para a População, a média ± 1,96 EP (erro padrão da média da população) fornece um intervalo dentro do qual há 95% de chance de se encontrar a média para a população. Este intervalo é denominado Limite de Confiança. Para as ciências exatas, ou mesmo circunstancialmente na área biológica, utiliza-se a média ± 2,58 EP, o que da um Limite de Confiança de 99%. Se um estudo é adequadamente desenvolvido e encontram-se valores para a média e para o EP de um parâmetro, qualquer outro estudo que encontre média, para o mesmo parâmetro, fora destes limites de confiança, permite inferir que se está frente a uma população diferente.

Estatística inferencial

Depois de concluir uma análise de estatística descritiva evolui-se para a segunda parte da estatística que é denominada estatística inferencial. Dentro desta área é que se utilizam os chamados Testes Estatísticos.

Teste "z" *ou* Transformação "z" - É um método estatístico permite avaliar em quanto um dado isolado se distancia da média da população, em números ab solutos. Em outras palavras, quantos EP o dado isolado dista da média da população? Baseia-se, exatamente, no conceito de Limite de Confiança. Sua expressão matemática é:

$$z = \frac{x_i - \mu}{E}$$

onde x_1 é o dado em teste, μ é a média da população, e EP é o erro padrão (desvio padrão da média da população).

Utilizando-se das conclusões do estudo de 300 casos para a droga X Tabela IV), se considerarmos um valor qualquer para a dose da droga X, equivalente à média + 1,96 EP, teríamos 4,4508 mg.kg^{-1}. Se aplicarmos o teste Z, encontraremos:

$$z = \frac{4{,}4508-4}{0{,}23} = \frac{0{,}4508}{0{,}23}$$

$$Z = 1{,}96$$

Ou seja, o teste "z" mede, exatamente, quantos EP, aquele valor dista da média da população. Permite inferir, então, para um valor isolado, se o mesmo está dentro ou fora do limite de confiança.

Menor que 1,96 = dentro dos 95%

Menor que 2,58 = dentro dos 99%

Maior que 1,96 = fora dos limites de 95%

Maior que 2,58 = fora dos limites de 99%

O teste "z", no entanto, só é válido para análise de dados expressivos para uma população de fato (grandes amostras). Nos manuais de estatística existem as tabelas completas para todos os percentuais e o resultado do teste "z" (consulte a referência 2).

Tamanho da Mostra: Sabendo-se que EP = DP @$_+$ é possível, matematicamente, calcular o tamanho necessário da amostra para inferir conclusões à população. A partir do limite de confiança, da média ideal estimada para a população e dos dados encontrados em uma nota prévia, teríamos:

$$z = \frac{x_i - \mu}{EP}$$

$$z = \frac{x_i - \mu}{DP/\sqrt{n}} \text{ logo}$$

$$n = \left[\frac{z \cdot DP}{x_i - \mu}\right]^2$$

Em nossa nota prévia com a droga X (Quadro I), encontramos a média 4,00 e um desvio padrão 0,74 mg.kg^{-1}. Partindo da suposição de que a média da população deva respeitar uma variação máxima de 5% sobre o valor que encontramos na amostra, dentro do limite de confiança de 95%, quantos elementos deve ter o nosso estudo para ser extrapolável à população?

5% de 4,00 é igual a 0,2 logo $(x_i - \mu) = 0{,}2$

Na Tabela z, 95% correspondem a 1,96,

então

$$n = \left[\frac{1{,}96 \cdot 0{,}74}{0{,}2}\right]^2$$

$$n = 52{,}59 \cong 53$$

Serão necessários, portanto, 53 dementos no estudo para, eventualmente, se inferir as conclusões para a população. Ao cabo desta amostragem, o cálculo deve ser refeito, com base nos novos dados de média e devio padrão encontrados para aferir se a amostra já é suficiente.

Hipótese nula (HO) e probabilidades (p)

Quando se faz uma análise estatística sempre se parte de uma hipótese, e as conclusões jamais podem garantir que o resultado encontrado é absolutameme verdadeiro, pelas dificuldades já expostas em definição correta de uma população e garantia de uma amostra adequada. Aventam-se hipóteses e afirmam-se probabilidades.

Digamos que um paciente que tomou a droga X, em outro serviço, exigiu uma dose de 5,2 mg.kg^{-1}. Ao tomar conhecimento da informação, vem a pergunta: Será que esta dose se enquadra na variação da média de nossa amostra? O sensato é achar que sim! A hipótese que estabelecemos, portanto, é que há 95% de chance desta dose estar dentro dos limites de confiança da distribuição encontrada em nosso estudo. Por princípio, achamos que não há diferença entre nossa variação de dose e o valor encontrado no outto serviço. Esta hipótese inicial, em estatística, chama-se hipótese nula (não há diferença).

Usando os dados da Tabela IV, vemos que os limites de 95% para a amostra ficam entre 3,22 e 4,78 mg.kg^{-1} (média ± 2 DP). Logo, 5,2 mg.kg^{-1} cai fora dos limites de 95% de nossa curva de Gauss.

Então hÁ, apenas, uma probabilidade menor que 5% (100-95) ou $p < 0,05$ de que aquela dose esteja dentro da distribuição de nossos resultados. .

Quando, em estatística, ocorre tal fato, $p < 0,05$, diz-se que há diferenga significativa entre os valores comparados. Portanto, nossa hipótese nula não foi confirmada. Ou o paciente testado tem sensibilidade diferente à droga (toxicomania, alcoolismo), fugindo, portanto, do critério que utilizamos para compor a nossa amostra, ou o outro serviço estabeleceu outros elementos para definir dose de indução, também diferindo de nossos critérios.

$$p \leq 0,05 = \text{diferença significativa}$$

Se quiséssemos aumentar o limite de confiabilidade para 99%, o intervalo ficaria entre 2,83 e 5,17 mg.kg^{-1}. Nesta circunstância, o valor testado também cairia fora do intervalo. Então, $p < 0,01$ (probabilidade menor que 1% de corfirmação da hipótese). Quando ocorre este fato ($p < 0,01$), diz-se que há diferença altamente significativa. Para pesquisas na área biológica, no entanto, raramente utiliza-se tal intervalo de confiança.

$$p \leq 0,01 = \text{diferença altamente significativa}$$

Se a dose que estivéssemos testando fosse de 5 mg.kg^{-1}, estaria fora dos limites de 95% e dentro dos 99%.

$$0,01 < p < 0,05 = \text{diferença significativa}$$

Se o limite de confiança inicialmente estabelecido for de 95%, qualquer análise que encontre $p > 0,05$ (probabilidade maior que 5%), ou seja, o valor testado caindo dentro do intervalo, confirma-se a hipótese nula, e diz-se que a diferença não é significativa (N.S.).

$$p > 0,05 = \text{NS}$$

Testes estatísticos mais usados

Existem testes específicos para dados Paramétricos como o "t" de Student para uma e duas amostras, Teste do "t" para amostras pareadas, Análise de Variância para amostras múltiplas, intra e extragrupos, Análise de Regressão e Coeficiente de correlação. Como este artigo não pretende esgotar o assunto, far-se-á a elucidação, apenas, da aplicação do teste "t" de Student.

Para dados Não Paramétricos utilizam-se outros testes: x^2= (Qui-quadrado), com ou sem correção de Yates, teste "U" de Mann-Whitney, o teste de Kruskal-Wallis. Da mesma forma, far-se-á alguma consideração sobre o teste do x^2.

A razão da esolha do teste do "t" de Student e o x^2 parte do princípio de que tanto nas publicações nacionais, na área médica, como nas intemacionais, mais de 80% dos trabalhos utilizam um deles ou mesmo os dois testes em suas análises"[1,4].

Teste do "t"

Este teste foi desenvolvido e publicado no início da revolução industrial, em 1908, por WS Gosset, visando subsídios para o controle de qualidade da empresa onde trabalhava. Como havia segredos industriais, utilizou o pseudônimo *Student* na publicação do artigo, fazendo com que seu teste se tornasse internacionalmente conhecido como Teste do "T" de Student[4].

Constitui uma extensão do que foi explanado para o teste "z" (relação entre urn dado isolado e a média da população em números de EP). Permite testar, também, a relação entre duas médias (comparação de duas amostras).

$$t = \frac{x_i - /x}{EP}$$ para uma amostra

ou

$$t = \frac{/x}{EP}$$ para amostra pareada

e, como o EP da diferença entre duas médias equivale a:

$$EP_{(x1-x2)} = (EP_1^2 + EP_2^2)^{1/2}$$

temos:

$$t = \frac{/x_1 - /x_2}{(EP_1^2 + EP_2)^{1/2}}$$ para comparação entre duas amostras não pareadas

O mérito do autor baseia-se na identificação do seguinte fato: para pequenas amostras os valores encontrados não são tão precisos como seria lícito prever com o teste "z", Exatamente o que acontece com o tamanho das amostras na área biológica.

Gosset, então, calculou os valores reais do "t" para um grande número de elementos variáveis e desenvolveu uma tabela que leva em conta os valores do "t" para cada grau de liberdade, dentro de cada faixa de probabilidade (p) (Tabela VI). Nota-se que quanto maior o valor do t, maior é o limite de confiança e menor é a probabilidade estimada (os valores iniciam pequenos para p = 0,05 e vão aumentando para a direita em direção a = 0,01).

Como usar este teste?

a) *Para uma amostra:*

Vamos avaliar a condição que estabelecemos quando se tratou de Hipótese nula e probabilidade. O paciente do outro serviço que apresentou dose de 5,2 mg.kg^{-1}em comparação com os nossos dados obtidos com os 300 pacientes.

Nossa hipótese nula é sim, ou seja provavelmente a dose daquele paciente está dentro dos limites de confiança de 95% e 99%. Então vamos testar (acompanhe os valores pelas Tabelas IV e VI).

dose a ser testada=5,2
média=4,0
EP=0,23
DF = 300-1 = 299 »¥

então,

$$t_0 = \frac{5,2-4}{0,23} = 5,21$$

Tabela VI - Tabela do "t" de Student resumida para as probabilidades de 5% e 1% e graus de liberdade(DF) de 1 a 11,20,30,40,60, 120 e para um número infinito de elementos da amostra.Os Valores abaixo de cada uma das probabilidades apontam o "t" a ser encontrado para cada um dos graus de liberdade.Verifique como usar, no texto

		P	
DF		**0,05**	**0,01**
1		12,71	63,66
2		4,30	9,92
3		3,18	5,84
4		2,78	4,60
5		2,57	4,03
6		2,45	3,71
7		2,36	3,50
8		2,31	3,35
9		2,26	3,25
10		2,23	3,17
11	t	2,20	3,11
20		2,09	2,84
30		2,04	2,75
40		2,02	2,70
60		2,00	2,66
120		1 98	2,62
$		1,96	2,58

Consultando a Tabela VI, para DF = $, e para os limites de confiança preestabelecidos (p = 0,05 e p = 0,01), encontramos:

		P			
DF		0,05	0,01		
$	t	1,96	2,58	>	t_0

Nosso valor (t_0) está à direita do valor previsto para 0,01, logo nossa probabilidade é...

p <0,01 ou, inferindo,

Há diferença altamente significativa ou seja, a dose obtida para o paciente testado não está prevista em nossa população.

Imagine, agora ,outro paciente que obteve uma dose menor: 4,5 mg.kg^{-1}, comparado ao mesmo grupo. Nossa hipótese nula é que ele faz parte da população definida em nosso estudo.

dose a ser testada =4,5
média = 4,00
EP = 0,23
DF=300-1 = 299 $\approx \infty$

então,

$$t_0 = \frac{4,5-4}{0,23} = 2,27$$

Consultando a Tabela Vl, para DF = \$, e para os limites de confiança preestabelecidos (p = 0,05 e p = 0,01) encontramos:

DF		0,05	P			0,01
\$	t	1,96	>	t_0	>	2,58

O valores testado está a direita de p = 0,05 e à esquerda de p = 0,01, logo...

0,01 < p < 0,05

Como para as ciências biológicas é aceito o limite de confiança de 95%, e como a amostra utilizada para comparação é composta de um grande número de elementos, representando, de fato, a população, podemos inferir que este paciente não está previsto em nossa população, desde que...

p < 0,05 (significativa)

2) *Para duas amostras:*

Vamos super que tenhamos testado a droga X em pacientes estado físico ASA II e Ill, no sentido de avaliar se outras condições mórbidas e terapias em uso interfere na dose de indução. Encontramos para um grupo inicial de 10 pacientes, os seguintes resultados: média =3,5 mg.kg^{-1} DP = 1,45 EP = 0,46. Nessa hipótese nula é que não há diferença. Então temos:

x_1 = 4,00 x_2= 3,5
EP_1 = 0,23 EP_2 = 0,46
DF = (300-1) + (10-1)= 299 + 9=308 $\approx \infty$

$$t_0 = \frac{4-3,5}{(0,23^2+0,46^2)^{1/2}} = \frac{0,50}{0,51} = 0,98$$

Consultando a Tabela Vl, para DF = ∞ encontramos...

DF				0,05	P	0,01
∞	t	t_0	>	1,96		2,58

Logo, p >0,05 N.S.

Portanto, a hipótese inicial é confirmada, ou seja, a existência de outras doenças e terapias em uso, aparentemente, não interferem com a dose de indução da droga X (logicamente é apenas uma conclusão didática).

3) *"t" de Student para amostras pareadas*

A situação prática ocorre quando se quer testar um parâmetro, no mesmo grupo de pacientes, antes e depois de um determinado acontecimento. Freqüência cardíaca antes e após a intubação traqueal, p. ex. Calculam-se as diferenças entre os valores antes e depois da intubação para cada paciente. Com as diferenças, calcula-se a média, desvio padrão e erro padrão. Aplica-se aos resultados o teste "t" para uma amostra, considerando-se a segunda fórmula.

Usando a situação proposta, após analisar 10 pacientes encontramos uma média das diferenças de freqüência cardíaca de 12 ± 3 bpm entre o momento antes e 2 minutos após a intubação traqueal. O erro padrão calculado é 0,095. Temos então...

$$t_0 = \frac{12}{0,95} = 12,631 \qquad (p < 0,01)$$

Infere-se que há diferença altamente significativa entre a FC antes e depois da intubação.

Limitação do teste do "t" de Student

O teste "t" de Student tem valores exclusivamente para comparação entre duas amostras. E freqüente observar-se o erro de utiliza-lo entre vários momentos de observação de um grupo de pacientes: variação de parâmetros circulatórios: 1) antes da punção venosa 2) depois da droga de indução, 3) após aplicar um hipnoanalgésico, 4) após o relaxante muscular e 5) após a laringoscopia e intubação. São múltiplas comparações, dentro da mesma população. Neste tipo de objetivo, o teste correto é o da Análise de Variância (ANOVA), que foge ao escopo deste artigo.

A única maneira possível de se utilizar o "t" para aquela situação é aplicando-se o corretor de Bonferroni[1], que corresponde à divisão dos valores da tabela do "t" pelo número correspondente às comparações desejadas (na situação descrita no parágrafo anterior, os números da tabela do "t" deveriam ser divididos por 5 para tornar possível o uso deste teste, sem incorrer em erro na inferência).

Teste do X^2 (leia-se teste do Qui-quadrado)

O teste do V^2 é utilizado para comparação de dados não paramétricos: ordinais, conceituais ou nominais.

Imagine-se a situação na qual se pretenda avaliar a incidência de náusea e vômito entre pacientes submetidos à indução com a droga X e com a droga Y, mantendo-se iguais todas as demais variáveis possíveis que pudessem interferir na observação: idade, peso, distribuição por sexo, droga de manutenção, relaxante muscular, tipo de ventilação etc. Os resultados observados, em termos de freqüência de ocorrência, seriam apresentados no que se denomina tabela de contingência das freqüências observadas (Tabela VII).

Tabela VII - Tabela de contingência das observações de 90 pacientes quanto à incidência de náusea e/ou vômito, após o uso em 40, da droga X e em 50, da droga Y

	Número de Pacientes		
	Náusea/ vômito	Sem náusea/ vômito	Total
DrogaX	24	16	40
DrogaY	32	18	50
Total	56	34	90

A incidência OBSERVADA de náusea/vômito para a amostra X foi de 60%(24/40) enquanto na amostra Y foi de 64%, aparentando que há alguma diferença entre os grupos. Para as duas amostras juntas encontram-se 56 casos para um total de 90 observações, o que dá um percentual de 62,22%

Se os resultados das duas amostras fossem iguais (não houvesse a diferença), as frequências para as observações seriam encontradas aplicando-se a incidência global de 62,22% para cada um dos grupos (62,22% de 40 e 62,22% de 50, respectivamente). A freqüência ESPERADA para um resultado que fosse igual poderia ser expresso em outra tabela de contingência (TabelaVIII).

Tabela VIII - Tabela de contingência do que seria esperado,se não houvesse diferença de incidência quanto à náusea e vômito para os 90 pacientes testados com a droga X e Y

	Número de pacientes		
	Náusea/ vômito	Sem náusea/ vômito	Total
Droga X	24,89	15,11	40
Droga Y	31,11	18,89	50
Total	5 6	34	90

A fórmula geral do teste do V^2, baseia-se na comparação entre as freqüências Observadas e as Esperadas, dentro da seguinte expressão:

$$X^2 = \sum \frac{(0 - E)^2}{E}$$

onde se lê: somatória dos quadrados das diferenças entre o Observado e o Esperado divididos pelo Esperado. A tabela de valores possível para o X^2, exposta na Tabela IX, apresenta os valores para o teste, para os vários graus de liberdade e para as probabilidades p = 0,05 e p = 0,01. Quanto maior o valor encontrado para o qui-quadrado, menor é a probabilidade de que haja semelhança entre os grupos comparados.

Tabela IX - Tabela do X^2 resumida para as probabilidades de 5% e 1%, graus de liberdade da 1 a 10. Os valores abaixo de cada uma das probabilidades apontam o valor do X^2 a ser encontrado para cada um dos graus de liberdade. Verifique como usar no texto

		P	
DF		0,05	0,01
1		3,84	6,63
2		5,99	9,21
3		7,81	11,34
4		9,49	13,28
5		11,07	15,09
6	x^2	12,59	16,81
7		14,07	18,48
8		15,51	20,09
9		16,92	21,67
10		18,31	23,21

O grau de liberdade, no caso deste teste, baseia-se na relação entre o número de linhas e colunas da tabela de contingência. É obtido multiplicando-se número de colunas menos 1 (C-1), pelo número de linhas menos 1 (L-1).

DF = (C-1). (L-1)

Aplicando-se as fórmulas, em nosso exemplo, teríamos.

DF = (2-1).(2-1) = 1.1 = 1

$$X^2_0 = \frac{(24-24,89)^2}{24,89} + \frac{(32-31,11)^2}{31,11} + \frac{(16-15,11)^2}{15,11} + \frac{(18-18,89)^2}{18,89}$$

x^2_0 = 0,032 + 0,025+ 0,052+ 0,042= 0,151

Na tabela de valores do X^2 (Tabela IX) para um grau de liberdade = 1 temos:

DF		P 0,05	0,01
1	x^2 | $X^2_0 <$	3,84	6,63

Logo, p >0,05 N.S.

Não há, portanto, diferença entre a incidência de náusea e vômito quando se usa a droga X ou a Y.

Neste teste, sempre são montadas as tabelas de contingência para as freqüências observadas e as esperadas, para facilitar a aplicação da fórmula.

Correção de Yates

Quando o número total de observações é muito pequeno, abaixo de 200, para evitar erros deve-se aplicar à fórmula do X^2 o fator de correção de Yates:

$$X^2 = \sum \frac{[(O - E) - 0,5]^2}{E}$$

De cada diferença entre observado e esperado extrai-se 0,5 e depois eleva-se ao quadrado. No exemplo citado, como o número de observações foi apenas 90, o fator de correção de Yates deveria ter sido aplicado; portanto:

$$x^2_0 = \frac{[(24-24,9)-0,05]^2}{24,9} + \frac{[(32-31,1)-0,5]}{31,1} + \frac{[(16-15,1)-0,5]}{15,1} + \frac{[(18,9)-0,5]^2}{18,9}$$

x^2_0= 0,078+ 0,005+ 0,01 + 0,103= 0,196

Em nosso caso, não houve mudança substancial no resultado final, que continua p > 0,05, portanto demonstrando que não há diferença significativa entre as duas amostras.

O exemplo didático apresentado levou em conta apenas uma observação em dois grupos de pacientes. O teste do X^2, porém, pode ser aplicado em situações não paramétricas com mais observações e para mais grupos. Só a título de exemplo, poderíamos avaliar para as drogas X, Y e Z, três possibilidades distintas: presença de vômito, presença de náusea ou ausência de qualquer um dos sinais e sintomas relacionados. Nossa tabela de contingência teria três linhas e três colunas. O grau de liberdade passaria a ser:

DF = (3-1).(3-1) = 2.2=4

e a fórmula geral precisaria somar o resultado da expressão para as nove casas das tabelas de contingência

$$X^2 = \frac{(O_1 - E_1)^2}{E_1} + \ldots + \frac{(O_9 - E_9)^2}{E_9}$$

Limitações do teste do X^2

Existem algumas restrições ao uso do teste do X^2. Além da correção de Yates, já relatada, outras situações limitam o seu uso:

a) Quando a tabela de contingência contiver apenas duas linhas e duas colunas (2 x 2), os valores esperados para cada uma das células não podem ser inferiores a 5, ou o teste resultará em erro (sugestão de diferença quando elas de fato não existem). Nesta circunstância, obriga-se o uso do teste exato de Fisher (consulte referências 2 e 3).

b) Em tabelas de contingência mais amplas (além de 2 x 2), os valores esperados para cada célula não podem ser inferiores a 1, e não mais que 20% das células podem conter valores menores que 5. Se este fato ocorrer, será necessária a associação de dados entre colunas e linhas no sentido de diminuir a tabela, ou aumentar a amostra das linhas que estejam com valores reduzidos.

CONCLUSÕES

Mesmo que haja um grande número de detalhes e fórmulas matemáticas envolvidos em Estatística, é fundamental que elementos do assunto sejam conhecidos.

Na eventualidade do desenvolvimento de um trabalho de pesquisa, mesmo clínico, com estes elementos, poder-se-á atentar para os cuidados necessários no sentido de não se pecar com generalizações indevidas. Porsegurança, será importante que haja o concurso de um estatístico. Ainda assim, é fundamental que se conheça o mínimo para discutir o protocolo do trabalho com aquele profissional. Seu conhecimento de Medicina e dos parâmetros com os quais convivemos, podem não ser profundos o suficiente para orientar o tipo de teste a ser aplicado. Se nada se conhece, poderá haver indução a um teste errado. A observação que se pretende é de dado paramétrico? ou não paramétrico? O estatístico não sabe se a extensão do bloqueio espinhal é um parâmetro exato. O anestesiologista, pelo menos, deve reconhecer que não.

Por outro lado, só com a compreensão do que sejam probabilidade, $p < 0,05$, variância, desvio padrão, erro padrão, e, principalmente, da necessidade de estabelecer criteriosamente as características da população e da amostra a serem testadas, haverá o direito à crítica sobre as informações lidas. Assim, mesmo que o trabalho publicado seja de origem importante, eventualmente suas contribuições não sejam totalmente extrapoláveis ao trabalho clínico diário. Ou não sejam confiáveis.

Fica claro, também, que trabalhos com pequenas amostragens (1 O ou 20 pacientes) só têm validade quando a distribuição dos dados tem uma dispensão muito estreita (pequeno coeficiente de variação). Caso contrário, será melhor considerar o artigo, apenas como uma nota prévia e que os resultados só tenham validade para a amostra do autor.

Seja para publicar ou ler o que está publicado, são necessários alguns conhecimentos elementares de estatística.

REFERÊNCIAS

1. Fisher DM-Statistics in Anesthesia,em Miller RD,Anesthesiology,2ª Ed., Churchil & Livingstone, New York, 1986: 185-221.
2. Colton T - Statistics in Medicine, 1ª Ed., Little Brown and Co., Boston, 1978: 344 pgs. e apêndices.
3. Winscow TDV - Statistics at Square One, 3.ª ed., British Medical Association, London, 1978:84.
4. Burley DM - Data Handling and Statistics, em Scurr C & Feldman S - Scientific Foundations of Anaesthesia, 1.ª ed., William Heinemann Medical Books Ltda., London, 1970:432-447.