42. É descontada em um banco, 4 mêses antes do vencimento, uma letra cujo valor nominal é de Cr$ 48 000,00. Qual a importância recebida do banco, sabendo-se que a taxa de desconto é de 10 % ao ano?
43. Reduzir 300 dólares a moeda brasileira ao câmbio de 62.
44. Quanto valem, em moeda brasileira, £ 124-8-6, ao câmbio de 52,5?
45. Preciso remeter para Lisboa uma importância eqüivalente a 20 000 escudos, ao câmbio de 1 978. Quanto devo enviar em moeda nacional?

RESPOSTAS:

1. 1.º) Cr$ 72,60; 2.º) Cr$ 47,50; 3.º) 0,73g; 4.º) 75t; 5.º) 4 500hb.
2. 1.º) 20 %; 2.º) 12,5 %; 3.º) 0,002 %.
3. 1.º) 5 ‰; 2.º) 31,25 ‰.
4. 1.º) Cr$ 625,00; 2.º) 40 000hb.
5. 12 %.
6. Cr$ 18 000,00.
7. 500.
8. 1 200.
9. 50 000.
10. 35 %.
11. 3 %.
12. Cr$ 12 478,24.
13. 7 ‰.
14. 428l.
15. 21 %.
16. Cr$ 648,00.
17. Cr$ 738,00.
18. Cr$ 270 000,00.
19. Cr$ 741,00.
20. Cr$ 1 800,00.
21. Cr$ 96 000,00.
22. Cr$ 300 000,00.
23. Cr$ 120 000,00.
24. 10 %.
25. Cr$ 18 360,00.
26. Cr$ 9 180,00.
27. 12 %.
28. $8\frac{1}{8}$ %.
29. 3a 3me 15d.
30. 11.
31. 10 anos.
32. 5 %.
33. 20 anos.
34. Cr$ 432 000,00.
35. Cr$ 18 000,00 a 6%.
36. Cr$ 22 800,00.
37. Cr$ 12 000,00.
38. 6 750.
39. $9\frac{1}{2}$ %.
40. 50 dias.
41. Cr$ 1 000,00.
42. Cr$ 1 600,00.
43. Cr$ 18 532,00.
44. Cr$ 6 532,00.
45. Cr$ 39 560,00,

NOTA: Como "*curiosidade*", já que não podemos fixar taxas cambiais e como confronto com a tabela da pág. 176, damos abaixo as *taxas cambiais "livres"* para pagamento "ad valorem" no mês de Outubro de 1958 na Alfândega do ponto de Santos (Publicação d'"O Estado de S. Paulo, de 19/10/1958).

PAÍSES	OFICIAL	LIVRE
A. do Norte	18,82	161,11
Alemanha..	4,50 91	38,35
Argentina .	1,04 56	3,77
Áustria....	0,72 77	6,38
Bélgica....	0,37 91	3,25
Canadá ...	—	164,13
Dinamarca.	2,72 51	22,20
Espanha ..	—	3,13 28
França......	0,04 49	0,38 18
Holanda....	4,98 86	41,85
Inglaterra...	52,51 62	452,70
Itália......	0,03 03	0,26 33
Portugal ...	—	5,65 31
Suécia	3,64 14	28,71 94
Suíça.......	4,42 69	37,69
Uruguai	—	21,02 37

CAPÍTULO SEGUNDO

Geometria Prática

INTRODUÇÃO

O objetivo primeiro do estudo da *Geometria Prática*, no Curso de Formação de Professôres Primários, é o de fazer com que o futuro mestre possa, ao apresentar às crianças as principais propriedades das figuras geométricas, ter meios para realçar essas mesmas propriedades, recorrendo, tanto quanto possível, às noções intuitivas fornecidas pelos objetos e corpos que nos rodeiam.

Evitar-se-ão, dêste modo, os excessos de justificações racionais que, dificilmente, os alunos, na idade em que se encontram, podem acompanhar. Outro ponto importante que o professor primário deve ter sempre presente é o de não dar definições geométricas sem sentido, a fim de não permitir a construção de uma *falsa geometria*. Assim, não podemos dizer, por exemplo, que "reta é a *linha* que passa por dois pontos", quando sabemos que por dois pontos podem passar tantas *linhas* quantas quisermos. Como a reta não é definida diretamente devemos deixar o seu conceito como *primitivo*, adquirido através de inúmeros exemplos (fios, raio de luz, etc..) e frisar, como propriedade fundamental, a *unicidade* da reta que passa por dois pontos — propriedade aliás de fácil verificação experimental.

É conveniente, de início, acostumar os alunos a se familiarizarem com as figuras planas, fazendo-os desenhar no fim de cada lição: quadrados, retângulos, losangos, paralelogramos, triângulos, trapézios, circunferências, etc., deixando, em seguida, um pouco de tempo à sua própria fantasia. Mais tarde, com relação aos sólidos, é sempre interessante orientá-los no sentido de fazerem coleções de modelos em cartolina. Propiciar-se-á, assim, aos jovens alunos, uma espécie de emulação artística ao lado do conhecimento de fato que passam a ter com as figuras, com os sólidos e as respectivas propriedades fundamentais.

I) Equivalência entre figuras geométricas planas. Áreas. Teorema de Pitágoras e suas aplicações

§ 1. Generalidades.

1. Noção intuitiva de equivalência plana. Sabemos que dois polígonos *iguais*, isto é, sobreponíveis, têm *superfícies iguais* (*). Também dois círculos iguais ou dois setores iguais, têm as suas superfícies respectivamente iguais, e, assim por diante. Na vida prática, porém, encontramos, muitas vêzes, figuras que têm superfícies iguais tendo, contudo, *formas diferentes*. É o que acontece, por exemplo, quando, para assoalhar o piso de duas salas, uma de forma quadrada e outra de forma retangular, empregamos o *mesmo número* de tacos de um dado tipo. Dizemos, simplesmente, que os dois pisos têm *superfícies iguais*, embora tenham formas diferentes.

A relação que exprime a igualdade entre as superfícies de figuras geométricas planas, de mesma forma ou não, recebe o nome de *equivalência plana*. Logo: *dois polígonos, ou duas figuras de contornos quaisquer, são equivalentes quando têm superfícies iguais.* Vemos, assim, que duas figuras iguais (sobreponíveis) são sempre equivalentes, enquanto que duas figuras equivalentes *podem deixar de ser iguais.*

Estendendo o conceito a mais de uma superfície, consideremos, por exemplo, o caso de dois terrenos contíguos que estão separados por uma cêrca. Suprimindo-se a cêrca comum obteremos um terreno único denominado *soma* dos dois primeiros. Por outro lado se no terreno único colocarmos uma cêrca de separação, que o desmembre em outros dois, obteremos a *diferença* entre o terreno todo e uma de suas partes.

Com relação às figuras geométricas planas equivalentes, valem os seguintes *critérios de equivalência*, que não podem deixar dúvidas quanto à sua evidência :

(*) Ver *Matemática*, Curso Ginasial, 3.ª Série, pág. 96, do mesmo autor.

1.º) *Duas figuras equivalentes a uma terceira são equivalentes entre si.*

2.º) *Se de duas figuras iguais ou equivalentes se somam (ou se subtraem) figuras iguais ou equivalentes, obtêm-se figuras equivalentes.*

§ 2. Áreas das principais figuras planas.

2. Medida dos polígonos: *área.* Êste estudo é feito não levando em conta a *forma* dos polígonos e sim as *superfícies* dos mesmos. Dêsse modo cada polígono pode ser substituído por um seu eqüivalente e a sua *medida* é expressa por um número denominado *área.*

Por ser pràticamente impossível a determinação direta da área de um polígono é usado o *quadrado* como elemento de confronto. No sistema métrico decimal, que é o adotado por nós, a unidade fundamental é o *metro quadrado* (m^2) — área do quadrado de um metro de lado —. Na Inglaterra, nos Estados Unidos, usa-se o *yard quadrado* (sq.yd.).

3. Retângulo. *A área do retângulo é igual ao produto da base (*) pela altura.*

$$\boxed{S = bh}$$

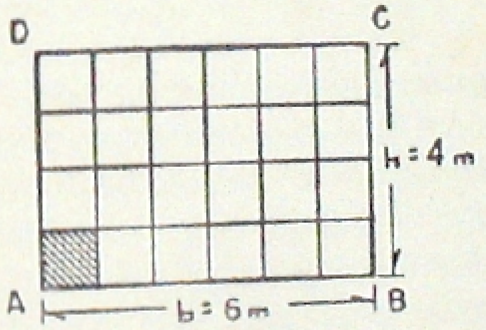


FIG. 12

De fato, consideremos o retângulo $ABCD$ (fig. 12), onde a base AB mede, por exemplo, 6m e a altura DA, 4m. Assinalados os quatro metros em AD e, traçando pelos pontos de divisão as paralelas a AB, o retângulo $ABCD$ aparece dividido em quatro retângulos iguais. Fazendo o mesmo em relação à base AB obteremos, pelas intersecções das paralelas traçadas, 24 quadrados, cada um com 1m² de área. Logo, a área do retângulo $ABCD$ é igual a 24m², isto é, número que se obtém pelo produto das medidas da base e da altura ($AB \times CD$).

É evidente que, conhecida a área de um retângulo e uma de suas dimensões, a outra é obtida dividindo a área pela dimensão conhecida.

4. Quadrado. *A área do quadrado é igual ao quadrado do lado.*

(*) Estamos nos referindo à *medida* da base (como também à *medida* da altura).

Com efeito, o quadrado é um retângulo que tem as duas dimensões iguais ($b = h$). É essa a razão porque a segunda potência de um número é chamada de quadrado.

$$\boxed{S = b^2}$$

5. Paralelogramo. *A área de um paralelogramo é igual ao produto da base pela altura.*

$$\boxed{S = b \cdot h}$$

Isto devido ao fato do paralelogramo $ABCD$ (fig. 13) ser *eqüivalente* ao retângulo $MNCD$, como é fácil de se verificar, aplicando-se os *critérios de eqüivalência.*

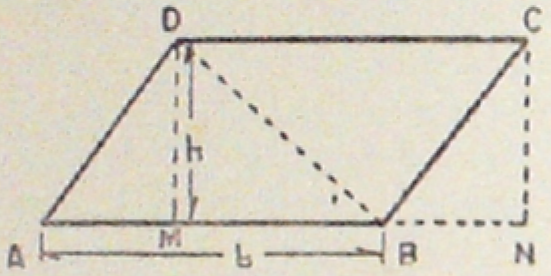


FIG. 13

6. Triângulo. *A área de um triângulo é igual ao semi-produto da base pela altura.*

$$\boxed{S = \frac{bh}{2}}$$

Realmente, o triângulo ABD (fig. 13) *é eqüivalente à metade do paralelogramo $ABCD$,* de mesma base e altura.

No caso do *triângulo retângulo,* um cateto pode ser considerado como base e outro como altura. A área do triângulo retângulo será, portanto, igual ao *semi-produto dos catetos.*

7. Trapézio. *A área do trapézio é igual ao produto da semi-soma das bases pela altura.*

$$\boxed{S = \frac{(b+b')h}{2}}$$

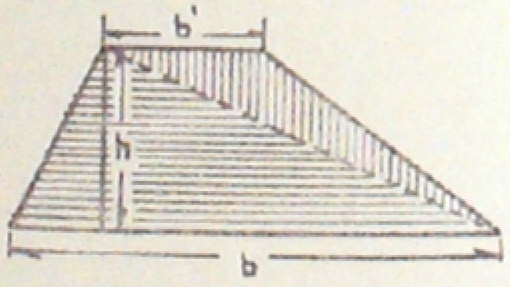


FIG. 14

De fato, o trapézio (fig. 14) é eqüivalente à soma de dois triângulos, sendo a base do primeiro dêles a base maior do trapézio e a base do segundo, a base menor, e, tendo ambos a mesma altura do trapézio

8. Losango. *A área de um losango é igual ao semi-produto das diagonais.*

De fato, o losango (fig. 15) *é eqüivalente a quatro triângulos retângulos iguais,* cujos catetos são, respectivamente, as metades das diagonais.

$$\boxed{S = \frac{d \times d'}{2}}$$

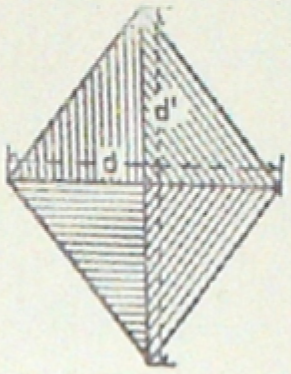


FIG. 15

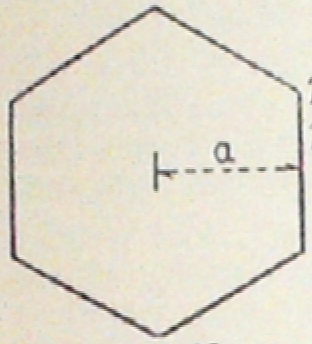


FIG. 16

9. Polígono regular. *A área de um polígono regular é igual ao semi-produto do perímetro pelo apótema.*

$$\boxed{S = \frac{2p \times a}{2} = p \times a}$$

Basta verificar que, um polígono regular (fig. 16), *é eqüivalente a um triângulo, tendo por base o perímetro do polígono e, por altura, o seu apótema*

10. Polígono qualquer. É determinada decompondo-se o polígono em figuras de áreas já conhecidas.

§ 3. Teorema de Pitágoras e suas aplicações.

11. Teorema de Pitágoras. Para os triângulos retângulos existe um teorema de equivalência, que é um dos mais importantes de tôda a Geometria Elementar. Atribuído ao geômetra grego *Pitágoras* (VI Séc. a.C.) o seu enunciado é o seguinte: *o quadrado construído sôbre a hipotenusa de um triângulo retângulo é equivalente à soma dos quadrados construídos sôbre os catetos.*

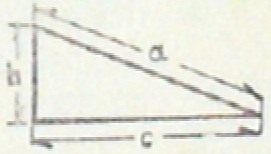


FIG. 17

Seja, de fato, o triângulo retângulo ABC (fig. 17), onde b e c são os catetos e a, a hipotenusa. A *verificação* do Teorema de Pitágoras é feita da seguinte maneira :

1) construímos o quadrado $MNPQ$ (fig. 18-I) de lado $b+c$ e adaptamos em cada um de seus ângulos retos o triângulo ABC, formando assim o quadrado $XYZT$, cujo lado é a hipotenusa a ;

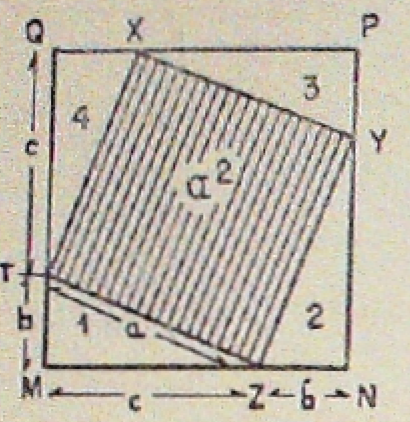


FIG. 18 - I

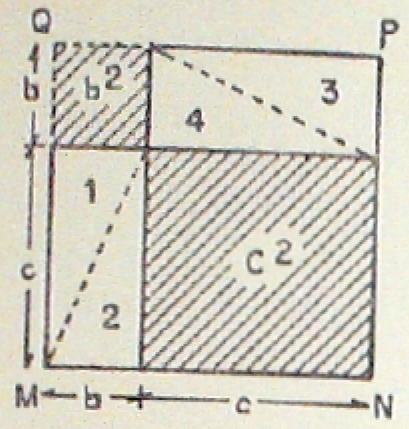


FIG. 18 - II

2) fazemos, agora, outra distribuição dos quatro triângulos retângulos no quadrado $MNPQ$ (fig. 18-II), de modo que fiquem individuados os quadrados, cujos lados são os catetos b e c, respectivamente.

Observamos, assim, que, se da fig. 18-I subtrairmos quatro vêzes o triângulo retângulo ABC, obteremos o *quadrado de lado a*, e, fazendo a mesma operação na fig. 18-II, obteremos a *soma dos quadrados de lados b e c*, respectivamente. Logo, êstes resultados são *eqüivalentes*, como quer o Teorema de Pitágoras.

12. Aplicações. As diversas aplicações métricas do Teorema de Pitágoras foram estudadas na 4.ª Série Ginasial, de acôrdo com os programas vigentes (*). Limitar-nos-emos aos enunciados das mais importantes:

1. determinação do comprimento da hipotenusa de um triângulo retângulo do qual se conhecem os comprimentos dos catetos;
2. determinação do comprimento de um dos catetos de um triângulo retângulo, do qual são conhecidos os comprimentos da hipotenusa e do outro cateto;
3. determinação do comprimento da diagonal de um retângulo conhecidas as suas dimensões; idem para o quadrado;
4. determinação dos comprimentos do lado, altura e base de um triângulo isósceles conhecido dois dêsses elementos;
5. determinação do lado e da altura de um triângulo equilátero conhecido um dêsses elementos.

(*) Ver *Matemática*, Curso Ginasial, 4.ª Série, pág. 103, do mesmo autor.

II) Comprimento da circunferência. Área do círculo

§ 1. Determinação do comprimento de uma circunferência.

13. Noção intuitiva. Consideremos, por exemplo, uma roda de bicicleta. Contornêmo-la com um barbante que fique bem ajustado à sua periferia e sôbre uma régua procuremos ler, com a melhor aproximação possível, o resultado dessa medida.

Dividindo-se êsse resultado pelo diâmetro ($2R$) da circunferência, representada por essa roda, obteremos para quociente um número não exato (irracional), de valor aproximado a 3,1415926... Repetindo-se a experiência com outras circunferências representadas por outras rodas, ou arcos de barris, notaremos que os quocientes entre as medidas de seus contornos e dos respectivos diâmetros é *sempre o mesmo*, valendo aproximadamente 3,1415926...

Indicando por C o comprimento de qualquer circunferência e por $2R$ o seu diâmetro, temos que

$$\frac{C}{2R} = 3{,}1415926\ldots..$$

Êsse número não exato, já conhecido dos antigos, é indicado com a letra π, que se lê "pi", e pertence ao alfabeto grego.

Logo: $$\frac{C}{2R} = \pi$$

donde $$\boxed{C = 2R\,.\,\pi} \quad \text{ou} \quad \boxed{C = 2\pi R}$$

isto é, *o comprimento de uma circunferência é dado pelo produto de seu diâmetro por* π.

Nos cálculos práticos, o valor de π é tomado com um êrro menor que 0,0001 por excesso, ou seja com o valor 3,1416. Exemplos:

1.º) Calcular o comprimento de uma circunferência que tem 5cm de raio. Aplicando a fórmula acima, temos:

$$C = 2 \times \pi \times R = 2 \times 3{,}1416 \times 5\text{cm}$$

ou $$C = 31{,}416\text{cm}.$$

2.º) Determinar o valor do raio de uma circunferência, cujo comprimento é 12,5664cm.

Como $C = 2 . \pi . R$, segue-se que, dividindo o valor do comprimento (C = 12,5664m) por π (3,1416) encontramos o diâmetro ($2R$). Dividindo o diâmetro por 2 encontramos o valor do raio. Cálculos:

$$12{,}5664\text{cm} : 3{,}1416 = 4\text{cm}$$
$$4\text{cm} : 2 = 2\text{cm}.$$

Logo, o raio mede 2cm.

§ 2. Determinação da área do círculo.

14. Noção intuitiva. Seja a circunferência (fig.19-I) de centro 0 e raio R. Para a obtenção da *área do círculo* limitado por essa circunferência, pode-se, primeiramente, dividi-la em 4 partes iguais, depois em 8, em 16 e assim sucessivamente.

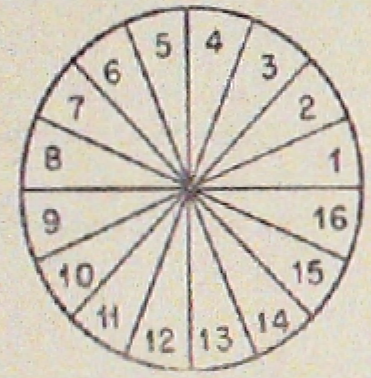


FIG. 19-I

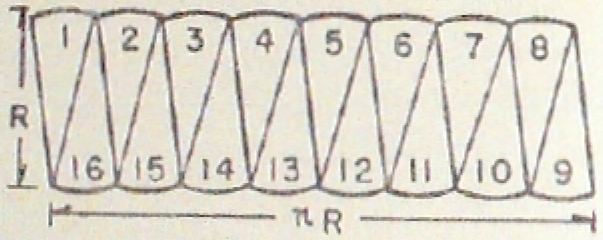


FIG. 19-II

Estas operações mostram-nos que os pequenos arcos iguais, nascidos das divisões sucessivas, tendem a se confundir com a corda respectiva.

Considerando os setores correspondentes a êstes arcos e dispondo-os como indica a fig. 19-II obtemos (no caso dos arcos se confundirem com as cordas) um *paralelogramo* que tem por base a semi-circunferência (πR) e por altura o raio R da circunferência. Logo, aplicando a fórmula que dá a área de um paralelogramo, vem:

$$\boxed{S = \pi R^2}$$

NOTA: Conhecida a área do círculo o seu raio é dado pela expressão:

$$R = \sqrt{\frac{S}{\pi}}$$

III) Equivalência entre figuras geométricas sólidas. Definições. Áreas das superfícies lateral e total. Volumes

§ 1. Generalidades.

15. Superfície poliédrica. Poliedros. Chama-se *superfície poliédrica* (convexa fechada) tôda figura constituída de polígonos situados em planos diversos e dispostos de tal modo que cada lado seja comum a dois dêstes polígonos e o plano de cada um dêles deixa todos os outros situados numa mesma região do espaço. Êstes polígonos dizem-se *faces*, e os vértices, os lados e os ângulos, respectivos, denominam-se: *vértices*, *arestas* e *ângulos* da superfície poliédrica. As arestas que partem de um mesmo vértice determinam um ângulo sólido denominado *anguloide*. Chama-se *poliedro* a figura geométrica constituída por todos os pontos comuns aos angulóides de uma superfície poliédrica. Os prismas e as pirâmides são particulares *poliedros*. Um poliedro é designado pelo número de faces que possui (no mínimo 4). Assim, chamamos de:

tetraedro, ao poliedro de *quatro faces*;
pentaedro, ao de *cinco faces*;
hexaedro, ao de *seis*;
heptaedro, ao de *sete*;
octaedro, ao de *oito*;
decaedro, ao de dez;
dodecaedro, ao de *doze*;
icosaedro, ao de *vinte faces*.

Entre os poliedros destacam-se os *poliedros regulares*. Diz-se *regular* um poliedro que tem tôdas as faces regulares e iguais, e, todos os angulóides iguais. Logo, um poliedro regular tem iguais tôdas as arestas e todos os diedros.

Enquanto que no plano existem polígonos regulares possuindo um número qualquer de lados: 3, 4, 5, 6, 7, 8, ...

entre os poliedros existem sòmente **cinco** *que são regulares:* o *tetraedro regular*, o *hexaedro regular* (cubo), o *octaedro regular*, o *dodecaedro regular* e o *icosaedro regular* (fig. 20).

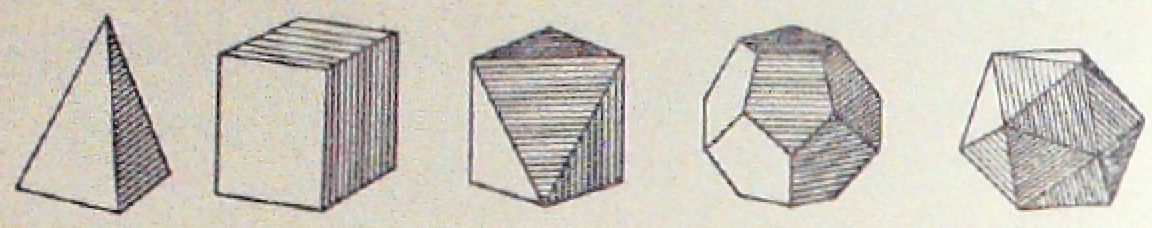


FIG. 20

16. Noção intuitiva de equivalência espacial. Assim como os polígonos (e mais geralmente as superfícies planas de contôrno qualquer), também os *poliedros* (e em geral as figuras sólidas de contôrno qualquer), podem ser confrontadas, levando-se em conta a *forma* e a *extensão*. Dêsse modo, podemos reconhecer, experimentalmente, se dois sólidos são *iguais*, ou com mais precisão, *equivalentes*, considerando-os como recipientes e notando se contêm a mesma quantidade de líquido (ou de areia finíssima) ou também, se modelados com a mesma argila, resultam com o mesmo pêso.

Geomètricamente, podem-se aplicar os mesmos critérios estudados na equivalência plana. Assim, reconhecem-se como *equivalentes*, duas figuras sólidas que se obtenham a partir de figuras iguais ou equivalentes somando ou subtraindo figuras iguais ou equivalentes.

§ 2. Principais sólidos geométricos (poliedros). Desenvolvimentos das respectivas superfícies. Volumes.

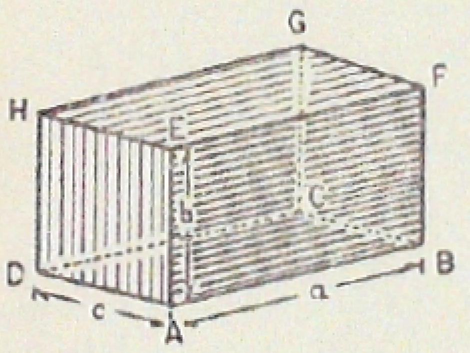
17. Desenvolvimento das superfícies lateral e total. A superfície plana que se obtém com as faces laterais de um poliedro, quando desenvolvidas em um plano, é denominada *superfície lateral* do poliedro. Acrescentando-se a essa superfície, as faces tomadas como bases, obtém-se a *superfície total*. Chama-se *área de superfície lateral*, e indica-se por S_l, a soma das áreas de tôdas as faces laterais do poliedro, e, *área da superfície total* (S_t) a soma das áreas de tôdas as suas faces (laterais e bases).

18. Medida dos poliedros: volume. A medida dos poliedros e de todos os sólidos em geral é chamada particularmente de *volume*. Logo, *volume de um sólido*, é o número que exprime a sua medida. A determinação direta do volume de um sólido é geralmente dificílima e pràticamente impossível. Daí a sua determinação indireta usando para o confronto o *volume de um cubo*.

No sistema métrico decimal adota-se como fundamental o *metro cúbico* (m^3), que é o volume de um cubo de 1m de aresta. No sistema inglês, usa-se a *jarda cúbica* (cu.yd.).

PRINCIPAIS SÓLIDOS GEOMÉTRICOS POLIÉDRICOS

19. Paralelepípedo retângulo. 1.°) Definição: *é o sólido geométrico* (poliedro), *cujas faces são retângulos* (fig. 21).



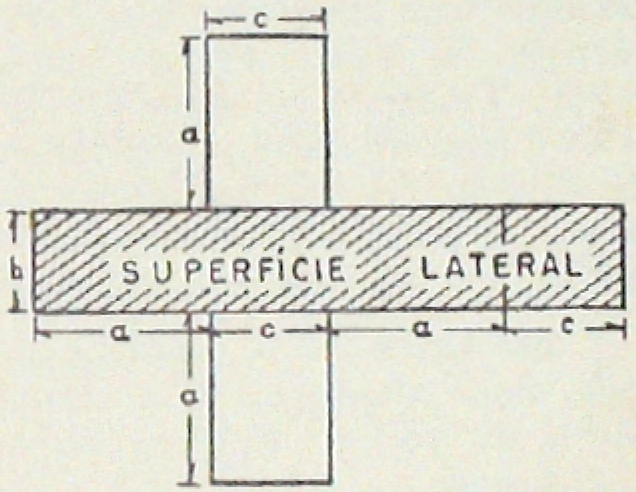


FIG. 21 FIG. 22

2.°) Desenvolvimento da superfície: Suponhamos ter um paralelepípedo retângulo, limitado por cartolina de dimensões: $a=4$m, $b=3$m e $c=2$m. Cortando-o segundo as arestas EA, EF, FB, FG, GC, CH e HD, e estendendo o cartão sôbre uma fôlha de desenho (fig. 22), obteremos o desenvolvimento da superfície do paralelepípedo. Reciprocamente, recortando essa superfície segundo as arestas assinaladas e com oportunas ligaduras, chega-se a construir o sólido. A área da superfície lateral é dada pela soma dos produtos 4×3 e 2×3, tomados duas vêzes, ou seja, $[(4\times2)\times3]2$, isto é: *a área da superfície lateral de um paralelepípedo retângulo é igual ao produto do perímetro da base pela altura.* Somando a essa área o dôbro da área da base, obtemos a *superfície total*. Em fórmulas, temos:

$$\boxed{S_l = 2(a + c)b} \qquad \boxed{S_t = 2(a + c)b + 2ac}$$

No exemplo, temos: $S_l = 36\text{m}^2$ e $S_t = 52\text{m}^2$.

3.°) Volume: *O volume de um paralelepípedo retângulo é igual ao produto de suas três dimensões.* Se a, b e c, são as três dimensões, temos: $\boxed{V = a \times b \times c}$.

Substituindo o produto $a \times b$, que indica a área do retângulo da base, por B e a outra dimensão c (altura) por h, o volume do paralelepípedo retângulo pode também ser expresso por

$$\boxed{V = B \times h}$$

De fato, o paralelepípedo retângulo $ABCDEFGH$ (fig. 21), onde as três dimensões caracterizadas pelas arestas que partem de um mesmo vértice medem, respectivamente, 4m, 3m e 2m. Traçando pelos pontos de divisão de AB planos paralelos à face $AEHD$ e procedendo anàlogamente com os pontos de divisão das outras arestas com relação às demais faces obtemos, pelas intersecções dêsses planos, 24 cubos de 1m de aresta, ou seja $4\text{m} \times 3\text{m} \times 2\text{m} = 24\text{m}^3$.

20. Cubo. 1.°) Definição: *É um paralelepípedo retângulo, cujas arestas são tôdas iguais* (fig. 23-I). O cubo é, portanto, um *poliedro regular*, tendo, por faces, quadrados iguais.

2.°) Desenvolvimento da superfície: Na fig. 23-II: vemos o desenvolvimento da superfície do cubo. Temos, agora,

$$\boxed{S_l = 4a^2} \qquad \boxed{S_t = 6a^2}$$

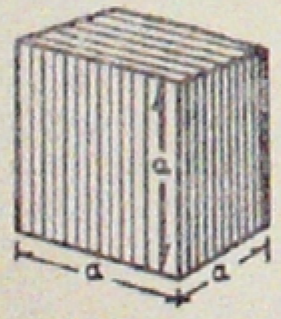


Fig. 23 - I

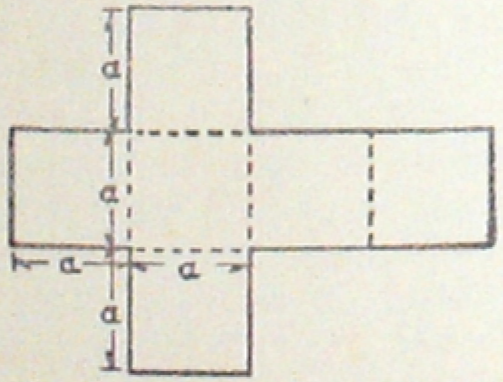


Fig. 23 - II

3.°) Volume: *O volume de um cubo é igual ao cubo* (medida) *da aresta.*

É o caso particular do paralelepípedo retângulo que possui as três dimensões iguais entre si ($a = b = c$).

$$\boxed{V = a \times a \times a = a^3}$$

21. Prisma reto. 1.°) Definição: *É o sólido geométrico* (poliedro), *cujas faces laterais são retângulos de mesma altura e as bases são polígonos iguais situados em planos paralelos* (fig. 24). A distância entre as bases é a *altura* do prisma.

2.°) Desenvolvimento da superfície: é dado pelos polígonos das duas bases e pelo retângulo de altura igual à altura do prisma e base igual ao perímetro da base do prisma. Portanto, *a área da superfície lateral de um prisma reto é igual ao perímetro da base pela altura.* Somando a essa área o dôbro da área da base, obtemos a *área da superfície total*. Vêm, assim, as fórmulas:

$$\boxed{S_l = 2p \times h} \qquad \boxed{S_t = S_l + 2B} \qquad \text{onde} \begin{cases} 2p \text{ — perímetro da base} \\ h \text{ — altura} \\ B \text{ — área da base.} \end{cases}$$

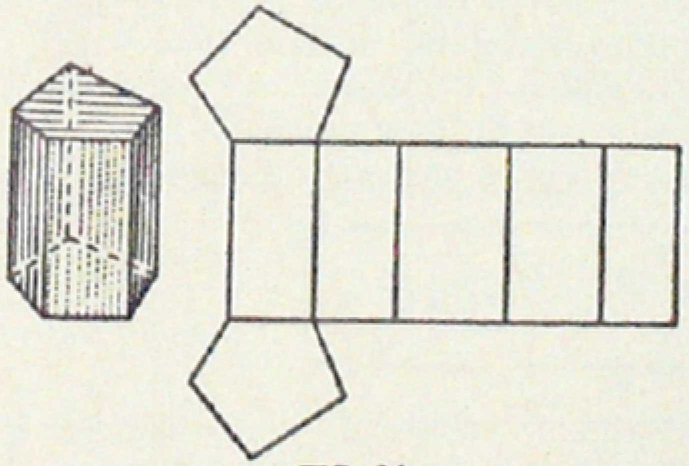


FIG. 24

3.°) Volume: *O volume de um prisma reto qualquer é igual ao produto da área da base pela altura.*

$$\boxed{V = B \times h}$$

A regra enunciada para a determinação do volume de um paralelepípedo retângulo pode ser estendida a um prisma reto quando se usam os *critérios de equivalência.*

22. Pirâmide reta. 1.°) Definição: *É o sólido geométrico* (poliedro) (fig. 25), *cujas faces laterais são triângulos isósceles, possuindo todos um vértice comum, e, dois a dois, lados comuns. A base é um polígono qualquer* (convexo).

2.°) Desenvolvimento da superfície: Se a pirâmide é reta de base regular, as faces laterais são triângulos isósceles iguais, tendo por altura o *apótema* da pirâmide. O seu desenvolvimento consta do polígono da base e de tantos triângulos

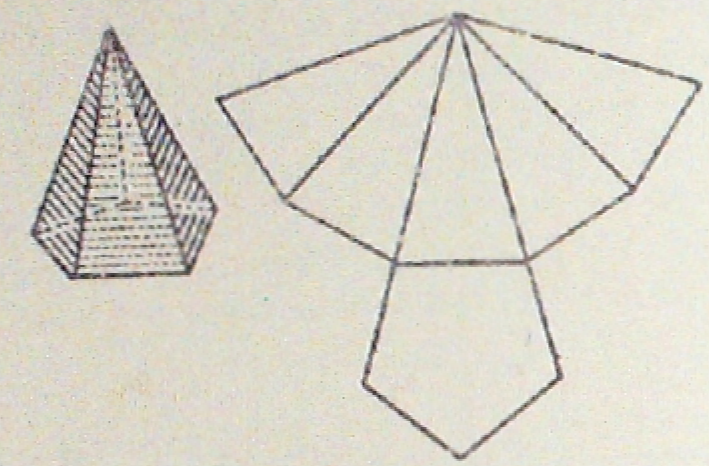


FIG. 25

isósceles iguais quantos forem os lados do polígono da base. A área da superfície lateral é dada pela soma das áreas de todos êsses triângulos. Portanto: *a área da superfície lateral de uma pirâmide reta de base regular é igual ao semi-produto do perímetro da base pelo apótema.* Temos, assim, as fórmulas:

$$\boxed{S_l = \frac{2p \times a}{2} = p \times a} \qquad \boxed{S_t = S_l + B}\,.$$

3.°) Volume: *O volume de uma pirâmide é igual a um têrço do produto da área da base pela altura.*

$$\boxed{V = \frac{1}{3}\,Bh}$$

A regra para calcular o volume da pirâmide pode ser verificada experimentalmente da seguinte maneira: toma-se uma caixinha de forma de um paralelepípedo retângulo e constrói-se com o mesmo material uma pirâmide que tenha a base

e a altura, respectivamente, iguais às da caixinha. Enchendo-a com areia bem fina e, em seguida, vertendo o seu conteúdo na caixinha, notar-se-á que serão necessárias *três dessas operações* para a encher exatamente. Isto vem mostrar, experimentalmente, que o volume da pirâmide é um têrço do volume do paralelepípedo retângulo de mesma base e altura.

§ 3. Principais sólidos geométricos redondos. Desenvolvimento das respectivas superfícies. Volumes.

23. Cilindro circular reto. 1.°) Definição: *É o sólido geométrico gerado pela rotação completa de um retângulo em tôrno de um de seus lados* (fig. 26). O lado que permaneceu fixo diz-se *eixo* e o lado que girou, *geratriz*.

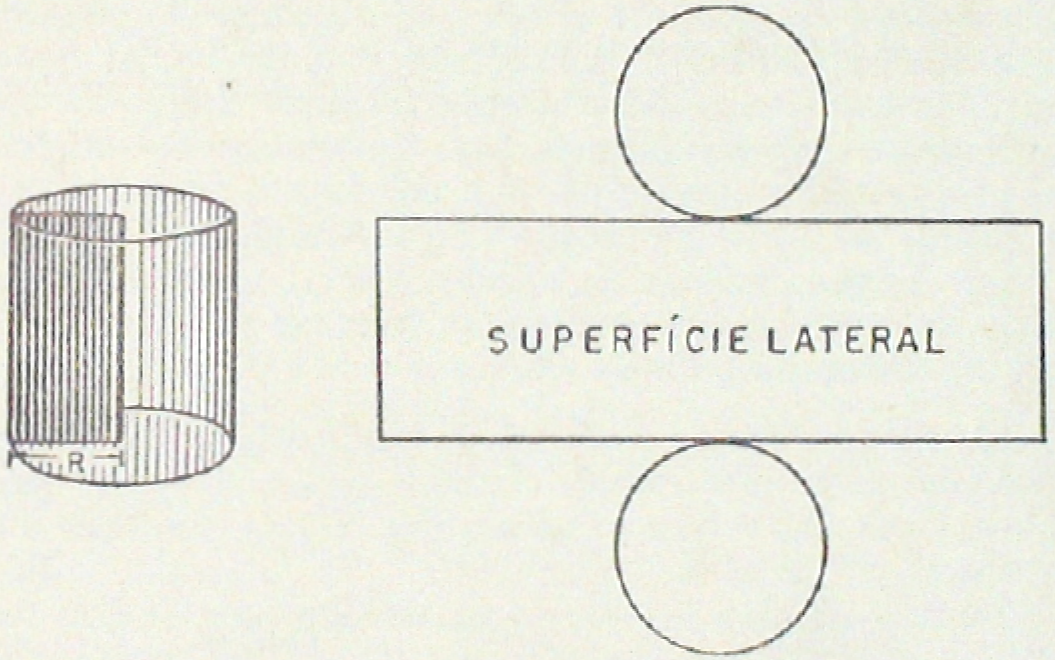


FIG. 26

2.°) Desenvolvimento da superfície: Suponhamos um cilindro de cartolina (fig. 26). Depois de cortado, segundo as circunferências das bases e uma geratriz, estendamos a cartolina obtida sôbre uma fôlha de desenho. A superfície desenvolvida do cilindro constará de dois círculos (bases) e de um retângulo (cuja altura é a altura do cilindro e cuja base é igual ao comprimento da circunferência da base do cilindro). O retângulo constitui a superfície lateral, e, o retângulo, com os dois círculos da base, a superfície total do cilindro. Logo: *a área da superfície lateral de um cilindro circular reto é igual ao produto do comprimento da circunferência da base pela altura.* Acrescentando-se a essa área o dôbro da área do círculo da base, obtemos a área da superfície total. Portanto,

$$\boxed{S_l = 2\pi Rh} \qquad \boxed{S_t = 2\pi Rh + 2\pi R^2 = 2\pi R(h+R)}$$

3.°) Volume: *O volume de um clindro circular reto é igual ao produto da área da base pela altura.*

$$\boxed{V = B \, . \, h} \qquad \boxed{V = \pi . \, R^2 \, . \, h} \, .$$

Para encontrar êsse resultado, pensemos inscrito numa das bases um polígono, e, consideremos o prisma reto que tem por base êste polígono e por altura a mesma do cilindro. É evidente que o volume do prisma é menor que o volume do cilindro e que a diferença entre os dois sólidos é tanto menor quanto menor fôr a diferença entre as superfícies do círculo da base, do cilindro e a do polígono da base do prisma. Tal diferença será pràticamente desprezível se o polígono tiver um número tão grande de lados que possa ser confundido com o círculo da base. Logo, o volume do cilindro é igual ao volume de um prisma de mesma altura e tendo por base um polígono de área igual ao círculo da base do cilindro.

24. Cone circular reto. 1.°) Definição: *É o sólido geométrico gerado pela rotação completa de um triângulo retângulo em tôrno de um de seus catetos* (fig. 27). A sua base é um

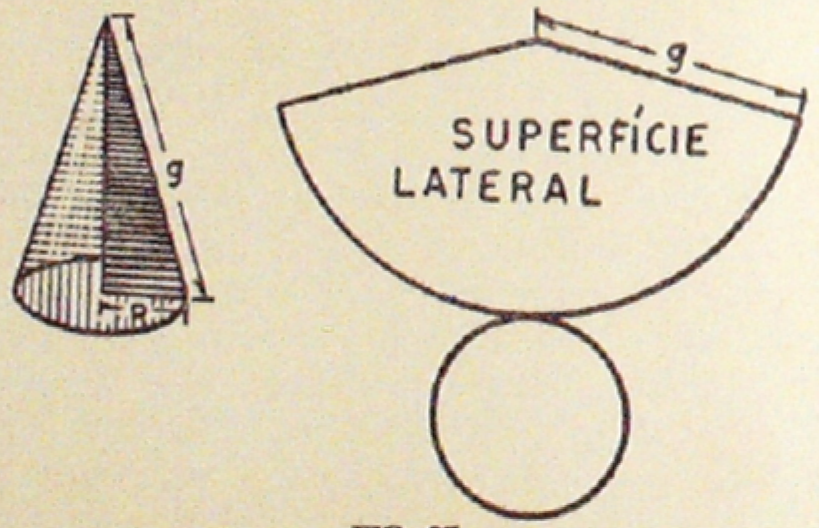


FIG. 27

círculo de raio igual ao cateto que girou. O cateto fixo é denominado *eixo* e a hipotenusa, *geratriz* ou *apótema*.

2.°) Desenvolvimento da superfície: Usando processo análogo ao estudado anteriormente, temos que o cone desenvolvido sôbre uma fôlha de desenho constará (fig. 27) de um círculo (base) e de um setor circular, cujo raio é igual à geratriz do cone e o arco é igual ao comprimento da circunferência da base do cone. O setor circular constitui a superfície lateral do cone, e, êsse setor mais o círculo da base, a superfície total. Assim, *a área da superfície lateral de um cone circular reto é igual ao semi-produto da geratriz pela circunferência da base.* A área da superfície total será obtida somando-se a S_l a área do círculo da base do cone. As fórmulas, são, pois:

$$\boxed{S_l = \pi Rg} \qquad \boxed{S_t = \pi Rg + \pi R^2 = \pi R(R+g)}$$

3.°) Volume: *O volume de um cone é igual a um têrço do produto da área da base pela altura.*

$$\boxed{V = \frac{1}{3} Bh} \quad \text{ou} \quad \boxed{V = \frac{1}{3} \pi R^2 h}$$

A justificação é a mesma feita para o volume do cilindro desde que se inscreva na base do cone um polígono e considera-se a pirâmide que tenha êsse polígono por base e o mesmo vértice do cone.

25. Esfera. 1.°) Definição: *É o sólido geométrico gerado pela rotação completa de um semi-círculo em tôrno de seu diâmetro* (fig. 28). Os pontos da semi-circunferência descrevem, no mesmo movimento, uma superfície denominada *superfície esférica*. Qualquer das circunferêcias de centro O e raio R diz-se *máxima* da superfície esférica.

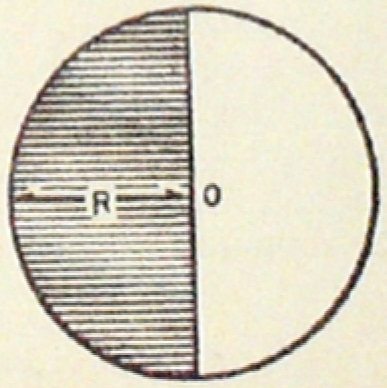


FIG. 28

2.°) Área da superfície esférica: O problema de medir uma superfície esférica é mais delicado do que os problemas análogos relativos ao cilindro e ao cone, em virtude da impossibilidade de se estender a esfera sôbre um plano, nem mesmo recortando-a. Êste fato permite dizer que *a esfera não é uma superfície desenvolvível*. Estudos que, dificilmente poderiam ser aqui explicados, levam-nos a concluir que: *a área de uma superfície esférica é igual a quatro vêzes a área do círculo limitado pela sua circunferência máxima.*

$$\boxed{S = 4\pi R^2}\ .$$

Experimentalmente, considerando uma esfera, cuja superfície esférica seja de latão (de mínima espessura) e cortando de uma fôlha de latão, de mesma espessura, quatro discos circulares de diâmetros iguais ao da esfera, obteremos o equilíbrio dos pratos de uma balança quando colocamos num deles a esfera (superfície esférica) e no outro, os quatro discos. Êste fato mostra que a superfície esférica e os quatro discos têm a mesma área.

3.°) Volume: *O volume de uma esfera é igual a quatro têrços do produto de π pelo cubo do raio.*

$$\boxed{V = \frac{4}{3}\pi R^3}\ ,$$

A justificação dêsse resultado pode ser feito da seguinte maneira: tomemos uma parte bem pequena de uma superfície esférica, e que seja limitada por pequeninos arcos de circunferências máximas. Unindo os vértices desta parte da superfície esférica com o seu centro obteremos um sólido que muito se aproxima de uma pirâmide, visto a base ser aproximadamente plana. Dêste modo, tôda a esfera pode ser imaginada como obtida somando um grande número de pirâmides construídas da maneira exposta. Logo: o volume de uma esfera é igual ao volume de uma pirâmide, cuja base tenha a mesma área que a superfície esférica e cuja altura seja igual ao raio da esfera, ou seja:

$$V = \frac{1}{3}Bh = \frac{1}{3}4\pi R^2 \times R = \frac{4}{3}\pi R^3.$$

EXERCÍCIOS SÔBRE ÁREAS E VOLUMES

1. Calcular a área do retângulo cujas dimensões são: base 4,5m; altura 2,3m.
2. O perímetro de um retângulo é igual a 32dm e a base vale o triplo da altura. Qual é a sua área?
3. Calcular, em dam², a área das seguintes figuras:
   1.°) retângulo (base: 12,32dam; altura: 8dam).
   2.°) quadrado (lado: 4,21dm).
   3.°) paralelogramo (base: 18,36m; altura $= \frac{1}{3}$ do valor da base).
4. A área de um retângulo é igual a 12dm². O dôbro de sua base vale 8dm. Qual é o valor de sua altura?
5. Um quadrado tem 36dm por perímetro. Qual é o valor de sua área?
6. Calcular a base de um retângulo sabendo-se que sua altura mede 9m e sua área é a *mesma* que a de um quadrado de 12m de lado.
7. Um losango tem as suas diagonais medindo respectivamente 12,35dm e 8,4dm. Calcular o valor de sua área em cm².
8. A área de um losango é igual a 72dm² e uma de suas diagonais mede 60cm. Quanto mede a outra?
9. Calcular, em dam², a área de um triângulo de base igual a 48,30m e de altura igual a 12m.
10. Um triângulo tem 64m² de área e a sua altura é igual a 80dm. Qual é o valor de sua base?
11. Calcular a área de um trapézio, sabendo-se que a base maior mede 3,8m, a base menor 2,6m e a altura 3,2m.
12. A área de um trapézio é de 150cm² e as suas bases são, respectivamente, 18cm e 12cm. Calcular o valor de sua altura.
13. Qual é a área de um círculo de raio igual a 6cm? (usar π com o valor 3,14).
14. Calcular a área de um semi-círculo pertencente a uma circunferência de 20dm de diâmetro.
15. Quanto se gastou para ladrilhar uma sala de 7,5m de comprimento por 4,8m de largura, sabendo-se que os ladrilhos usados são de forma quadrada, de 0,20m de lado, e custaram Cr$ 300,00 o cento.
16. João tem uma propriedade em forma de trapézio, medindo as bases 718m e 484m, respectivamente, e a altura 520m. No centro do terreno há um tanque de forma circular de 5m de raio. A residência de João e um bosque ocupam nesse terreno uma área igual a 11 500m². Qual é a área do terreno disponível para se plantar?
17. Num paralelepípedo retângulo cujas dimensões são: 6dm, 4dm e 3dm, respectivamente, calcular: as áreas da superfície lateral e total; o volume.

18. Qual é a área da superfície total de um cubo de 2m de aresta? Qual é o volume?
19. Calcular a S_l, S_t e o volume de uma pirâmide de base quadrada, cujo perímetro mede 24m, a altura 4m e o apótema 5m.
20. Calcular a S_l, a S_t e o volume de um cilindro de 5cm de raio e 16cm de altura. ($\pi = 3{,}14$).
21. Determinar a S_l, a S_t e o volume de um cone que possui 20dm de diâmetro e 15dm de geratriz.
22. Qual é a superfície da esfera de 3dm de raio? Qual é o volume?
23. Conhecendo-se de um paralelepípedo retângulo o seu volume que é de 144dm³ e a sua altura que mede 9dm, calcular o valor da área da base dêsse paralelepípedo.
24. A soma de tôdas as arestas de um cubo é 36m. Calcular, em dm³ o seu volume.
25. A base de um prisma é um trapézio, cujas bases medem, respectivamente, 12dm e 8dm e a altura 5dm. A altura do prisma é igual a 28dm. Calcular o seu volume.
26. Uma pirâmide de 12dm de altura tem por base um retângulo cujas dimensões são: 5dm e 3dm, respectivamente. Calcular o volume dessa pirâmide.
27. Calcular a altura de uma pirâmide de volume igual a 93dm³ e cuja área da base é de 31dm².
28. Num cilindro circular reto temos: volume 9 420cm³ e área da base 314cm². Calcular a área da superfície lateral.
29. O raio de uma esfera é igual a 6cm. Calcular o volume dessa esfera.
30. Determinar o volume de um cone de 10dm de altura, sabendo-se que a circunferência de sua base mede 28,26dm.
31. Pagaram-se Cr$ 13 500,00 pela construção de um muro de 3m de altura por 0,30m de espessura. Qual é o seu comprimento, se o preço do m³ foi de Cr$ 300,00?
32. As dimensões de uma árvore jequitibá, de forma cilíndrica, são: altura 15m e raio da base 0,70m. Sabendo-se que o m³ dessa árvore, foi vendido à razão de Cr$ 900,00, pergunta-se quanto rendeu tôda a árvore.
33. Antão tem um sapo de borracha que cheio de ar ocupa um volume igual ao volume de uma esfera de 3dm de raio. Qual é o volume do sapo?
34. Um vagão de estrada de ferro medindo 18m de comprimento por 3m de largura e 2,5m de altura está cheio de areia. Qual é o preço total do transporte dessa areia se o preço do transporte de $\frac{1}{3}$ de m³ de areia custa Cr$ 30,00?
35. Um reservatório de forma cilíndrica, cujas dimensões são: raio 2m e altura 10m está cheio de uma certa substância. Qual é o valor dessa substância, sabendo-se que 10m³ valem Cr$ 12 000,00?

RESPOSTAS:

1. 10,35m².
2. 48dm².
3. 1.º) 98,56dam²; 2.º) 0,00177241dam²; 3.º) 1,123632dam².
4. 3dm.
5. 81dm².
6. 16m.
7. 5 187cm².
8. 24dm.
9. 2,8980dam².
10. 16m.
11. 10,24m².
12. 10cm.
13. 113,04cm².
14. 157dm².
15. Cr$ 2 700,00.
16. 300 941,50m².
17. S_l = 60dm²; S_t = 108dm²; V = 72dm³.
18. 24m²; 8m³.
19. S_l = 60m²; S_t = 96m²; V = 48m³.
20. S_l = 471m²; S_t = 628m²; V = 1 177,500m³.
21. S_l = 471m²; S_t = 785m²; V = 1 570m³.
22. 113,04dm²; 113,04dm³
23. 16dn².
24. 27 000dm³.
25. 1 400dm³.
26. 60dm³.
27. 9dm.
28. 1 884cm².
29. 904,320cm³.
30. 211,950dm³.
31. 50m.
32. Cr$ 20 771,10.
33. 113,040dm³.
34. Cr$ 12 150,00.
35. Cr$ 150 720,00.

CAPÍTULO TERCEIRO

Noções de Estatística

INTRODUÇÃO

1. Origem e natureza dos dados estatísticos. Todo progresso do conhecimento humano é dominado modernamente pela noção de *medida*. Essa tendência, como não poderia deixar de ser, também fez-se notar na Psicologia, na Sociologia, na Educação e, de um modo geral, onde a *medida dos fenômenos* que se estudam é ponto fundamental na legitimidade dos resultados encontrados. A medição da inteligência, por exemplo, onde se destacam o experimentalismo de Wundt (1879) e a revisão da Escala de Binet (redistribuição dos Q.I. pela Universidade de Stanford, 1937), foi possível graças a decisiva contribuição recebida pela estatística.

Desta maneira, nos campos científico, econômico, estatal, etc., ... estudo algum deixa de apresentar *numéricamente* os seus fatos, empregando geralmente a palavra *estatística*. O médico mostra as suas *estatísticas* de cura; o higienista, as suas *estatísticas* de mortalidade; o educador, as suas *estatísticas* de crescimento das matrículas escolares ou das crianças separadas em grupos de inteligentes, médias e retardadas; os políticos (ou as organizações específicas), as suas estatísticas da opinião pública sôbre as preferências do povo, etc. São assim, as informações estatísticas, de grande valia para o homem que deseja curar, educar, fazer ciência ou dirigir a política de acôrdo com os reais desejos do povo.

2. Definição de Estatística. A palavra *Estatística*, de origem latina (*), introduzida nos meados do século XVIII por Godofredo Achenwall, foi considerada por muito tempo como *ciência dos negócios do estado*. Os que governavam, sentindo a necessidade de informações sôbre os respectivos países, organizavam departamentos aos quais eram confiadas essas investigações. Hoje, porém, não sòmente os governos,

(*) Ver *Métodos estatísticos em Psicologia* — F. M. Urban (Boletim CLXII, da Faculdade de Filosofia, Ciências e Letras, da Universidade de São Paulo).

como também as emprêsas particulares e genèricamente *todos os ramos de estudos*, lidam com tôda espécie de investigações estatísticas, que têm, por assim dizer, um característico em comum: *tratam de fatos que são expressos numéricamente.*

Habitualmente toma-se por "Estatística" qualquer tabela ou gráfico que apresente os resultados numéricos de uma observação. De um modo mais rigoroso, chamando de *fenômenos coletivos* aquêles que dependem de uma multiplicidade de causas, e, portanto, sujeitos a uma tendência imprecisa, podemos definir Estatística da seguinte maneira:

> ***Estatística** é o método que tem por objeto o estudo dos fenômenos coletivos traduzidos nas suas expressões numéricas.*

É, pois, a Estatística um dos ramos da matemática aplicada a dados que se observam e que procura, sob forma analítica ou gráfica, estudar as tendências da variação dêsses dados. Os processos usados pela Estatística (*), filiam-se todos êles, ao chamado *método indutivo*, isto é, aquêle que, partindo dos fatos verificados por meio de observações e experiência, procura chegar aos princípios que os regem.

Na Educação, o método estatístico é empregado com grande freqüência para estudos dos problemas quer pedagógicos quer administrativos. Nos problemas pedagógicos enquadram-se, por exemplo, os relativos às capacidades físicas (distribuições de alunos por estaturas, por influências de efeitos na aprendizagem, etc....); aos rendimentos mentais das crianças; aos diversos processos de ensino, e assim por diante. Nos problemas administrativos, destacam-se os concernentes aos planejamentos e restruturações de departamentos educacionais; aos estudos sôbre vencimentos, qualificações e eficiência dos professores; aos estudos sôbre o custo do ensino, capacidade de prédios escolares, etc.

Chama-se *universo estatístico* ou *população estatística* ao conjunto de entes portadores de, pelo menos, uma característica em comum. Assim, por exemplo, os *estudantes* constituem um universo estatístico, pois, possuem pelo menos uma característica em comum: *são os que estudam.*

Denomina-se *amostra* a uma parte representativa do universo estatístico que se estuda. Dêsse modo, se quisermos estu-

(*) Ver *Elementos de Estatística Geral* — Milton da Silva Rodrigues (Publicação da Companhia Editora Nacional).

dar a variação da *estatura* dos alunos que cursam os Grupos Escolares, podemos recolher *amostras* representativas de todos os Grupos Escolares para proceder tal estudo. Devemos frisar, neste instante, que existe uma técnica especializada (técnica de amostragem) para escolher amostras, no mínimo 10% do universo estatístico, que garante, tanto quanto possível, o *acaso* na escolha.

Exprimindo por *números* as observações feitas a entes de um mesmo universo estatístico, obteremos os chamados *dados estatísticos* relativos a êsse universo.

3. Séries ou distribuições estatísticas. Série estatística é o nome que se atribui a um conjunto de dados estatísticos distribuídos segundo as diversas modalidades do fenômeno que representam (no exemplo acima o fenômeno é a *estatura*). Os dados de uma série estatística são também denominados *têrmos* da série. As séries são classificadas atendendo-se aos três elementos principais: *tempo*, *local* e *categoria*, que todos nós estamos familiarizados a ver nas tabelas estatísticas constantes na maioria dos livros didáticos, anuários, jornais, etc. Podemos, assim, destacar quatro tipos de *séries* ou *distribuições*:

1.ª) *Série cronológica* (ou *distribuição temporal*): Quando os dados forem distribuídos de acôrdo com o *tempo* em que se produziram. Nessas séries os elementos local e categoria não variam. Exemplo:

MOVIMENTO DA POPULAÇÃO DO ESTADO DE SÃO PAULO* (1940-1960)**

ANOS	POPULAÇÃO	ANOS	POPULAÇÃO	ANOS	POPULAÇÃO
1940	7 155 000	1947	8 512 200	1954	10 081 000
1941	7 334 800	1948	8 726 100	1955	10 330 000
1942	7 519 100	1949	8 945 300	1956	10 585 000
1943	7 708 000	1950	9 142 000	1957	10 847 000
1944	7 901 600	1951	9 368 000	1958	11 115 000
1945	8 100 100	1952	9 600 000	1959	11 390 000
1946	8 303 600	1953	9 837 000	1960	12 930 000

(*) *Fontes:* Boletim n.º 1 (1953) do Departamento de Estatística do Estado de São Paulo. *Anuários Estatísticos do IBGE* (1951-1961).

(**) Ver quadro da população recenseada do Brasil, por Unidades da Federação (Recenseamento de 1960), na pág. 253.

2.ª) *Série geográfica* (ou *distribuição territorial*): No caso dos dados serem distribuídos de acôrdo com os *locais* onde foram obtidos. Agora, os elementos que não variam são o tempo e a categoria.

ANALFABETISMO NAS AMÉRICAS (1952) (*)
(a partir de 15 anos)

PAÍSES	PORCENTAGEM
Canadá	2,5
Estados Unidos	3,0
Argentina	16,6
Cuba	22,0
Chile	27,0
Panamá	37,9
Colômbia	44,0
México	53,9
Brasil	56,0
Peru	57,6
Venezuela	58,5
Honduras	65,7
El Salvador	72,4

NOTA: Faltam os dados relativos ao Uruguai, República Dominicana, Equador, Nicarágua, Bolívia, Guatemala e Costa Rica.

Meditemos, diante dêsse quadro, com a *realidade brasileira*: *em cada grupo de cem cidadãos, maiores de* 15 *anos, apenas* 44 *sabem ler, escrever e contar.* É imprescindível, aos que têm a felicidade de estudar, contribuir para a alteração dêsse quadro tão contristador para nós.

3.ª) *Série categórica* (ou *distribuição específica*): Quando os dados são distribuídos de acôrdo com a sua *espécie*. São fixos o tempo e o local. Exemplo:

COMÉRCIO EXTERIOR PELO PÔRTO DE SANTOS (**)
(Exportação — Julho de 1952)

MERCADORIAS	QUANTIDADE (kg)	VALOR (Cr$)
Animais vivos	—	—
Matérias-primas	8 041 726	101 171 837
Gêneros alimentícios	62 173 481	908 562 625
Manufaturas	283 116	4 150 669
TOTAL	70 498 323	1 013 885 131

(*) Fonte: Serviço de Educação de Adultos – S. Paulo.
(**) Fonte: Serviço de Estatística Econômica e Financeira.

4.ª) *Seriação (ou distribuição por freqüência)*: Quando os dados são distribuídos de acôrdo com a sua *grandeza* em ordem crescente (ou decrescente) obedecendo gradações convenientes, pois, o tempo, o local e a categoria permanecem fixos. Nestas séries, aparece o importante problema da *tabulação*, de grande aplicação no campo da estatística, e onde os dados são dispostos em *classes* oportunas, como veremos nos parágrafos seguintes. Exemplo:

COMPOSIÇÃO DEMOGRÁFICA DO DISTRITO FEDERAL (1949) (*)

IDADES (Classes)	PONTOS MÉDIOS (das classes)	FREQÜÊNCIA (por 100 000 hb.)
De 0 a 10 anos (**)	5	20 208
De 10 a 20 anos ...	15	20 306
De 20 a 30 anos ...	25	21 939
De 30 a 40 anos ...	35	16 588
De 40 a 50 anos ...	45	10 433
De 50 a 60 anos ...	55	6 056
De 60 a 70 anos ...	65	2 996
De 70 a 80 anos ...	75	1 082
De 80 a 90 anos ...	85	295
De 90 a 100 anos ...	95	79
De 100 e mais anos..	—	18
TOTAL.........	....	100 000

§ 1. Levantamento estatístico.

4. Definição. Fases de um levantamento. *Levantamento estatístico* é a operação que possibilita estudar os universos estatísticos, determinando a tendência característica de seus dados. As principais fases de um levantamento estatístico são:

a) Coleta de dados;

b) Disposição dos dados, tabulação e distribuição por freqüência;

(*) Exclusive.
(**) Fonte: Boletim do M. T. I. C.

c) Análise das distribuições por freqüência;

d) Conclusões sôbre resultados obtidos.

Vamos descrever, inicialmente, as duas primeiras fases e, a seguir, depois da introdução de novos conceitos e de um *exemplo-modêlo*, as duas últimas.

***a*) Coleta de dados.**

É o primeiro trabalho estatístico do levantamento. A coleta pode ser feita diretamente ou indiretamente, segundo os dados sejam obtidos pelo próprio organizador ou por intermédio de outros. Em qualquer dos casos, os dados devem ser em totais que possam justificar a divulgação que se queira dar aos resultados (o tamanho mínimo da amostra representativa deve ser de 10% do universo que se estuda).

***b*) Disposição dos dados, tabulação e distribuição por freqüência.**

Coligidos os dados, êstes são dispostos num quadro ou tabela inicial denominada *tabela primitiva*. Em seguida, procura-se imprimir uma certa ordem aos dados que facilite o estudo dos mesmos. Se os dados forem *medidas* é comum dispô-los por ordem de grandeza (crescente ou decrescente) constituindo, assim, uma nova tabela, agora, denominada *rol*. Os dados que compõem o rol passam a ser seus *têrmos*.

Logo depois vem a fase da *tabulação*. Tabular significa registrar o número de vêzes que cada têrmo aparece na tabela primitiva ou no rol. Chamando *freqüência absoluta* ou simplesmente *freqüência* ao número de vêzes que cada têrmo figura no rol, podemos dizer, também, que *tabular é registrar a freqüência de cada têrmo do rol.*

Com essa finalidade procuramos reunir todos os têrmos do rol em convenientes grupos denominados *classes*, a fim de melhor *distribuir as freqüências*.

A tabulagem manual (*) é feita, quase sempre, por duas pessoas: a primeira enunciando os dados que contam da tabela primitiva ou do rol e a segunda registrando-os, por meio de *sinais* (geralmente pausinhos), na frente de cada classe cujos totais serão no final substituídos por números que representam a freqüência de cada classe, como veremos no exemplo-modêlo que será estudado logo mais.

(*) A tabulagem mecânica é feita, geralmente, por máquinas de cálculo ou aparelhos dos tipos Powers ou Hollerith.

Cada classe tem por extremo dois números que são denominados *limites da classe.* O menor dêles *limite inferior*, cuja indicação é l_i e o maior, *limite superior*, de indicação l_s.

Chama-se *amplitude de classe* ou *intervalo unitário* de uma classe a diferença entre os limites superior e inferior dessa classe. Indicação: h.

Logo: $$h = l_s - l_i$$

Amplitude total ou *intervalo total* de uma distribuição de freqüências é diferença entre o *maior* l_s e o *menor* l_i figurantes na distribuição. Isto significa que o l_i e o l_s que definem a amplitude total, podem deixar de constar como dados do rol. Indicação: A.

É evidente que dividindo o valor da *amplitude total* (A) pelo valor que representa a *amplitude de classe* (h), comum a tôdas, obtemos o *número de classes* da distribuição.

Para melhor compreendermos as duas primeiras fases de um levantamento estatístico consideremos um exemplo, que será *modêlo* para o desenvolvimento de nosso curso.

Exemplo-modêlo: *Efetuar o levantamento estatístico das* ***estaturas*** *das alunas do 1.º Ano Normal, de um Instituto de Educação, usando como* ***amostras*** *as estaturas de 40 alunas, constantes da tabela primitiva ao lado.* As medidas foram aferidas em *cm* e se aplicaram as convenções usuais de *arredondamento*: se o algarismo seguinte à casa a que se quer aproximar fôr inferior a cinco, despreza-se êsse algarismo, conservando-se apenas os que vêm antes; se êsse algarismo fôr igual ou superior a cinco, deve-se aumentar uma unidade na casa anterior. Assim, por exemplo, uma estatura de 153,4cm será arredondada para 153cm; uma outra de 153,7cm será arredondada para 154cm e, em ambos os casos, diremos que *as medidas foram efetuadas com êrro inferior a meio centímetro.* O *rol* correspondente a essa tabela primitiva, ou seja a disposição dêsses dados por ordem de grandeza crescente é:

TABELA PRIMITIVA

160	163	165	151
152	156	178	158
155	162	152	166
154	161	157	169
161	161	167	170
162	171	160	158
162	160	170	160
161	156	156	168
150	155	164	164
160	153	155	163

150	156	161	164
151	156	161	165
152	157	161	166
152	158	161	167
153	158	162	168
154	160	162	169
155	160	162	170
155	160	163	170
155	160	163	171
156	160	164	178

Entremos, agora, na *tabulação*. Como a menor estatura é de 150cm e a maior 178cm, podemos *distribuir* os dados relativos a essas 40 estaturas, constantes da tabela primitiva ou do rol, num intervalo total, cujos limites sejam 150cm e 180cm, respectivamente, isto é, de *amplitude total* (A) igual a 180cm − 150cm = 30cm.

Tomando para as classes que vão compor essa distribuição, uma amplitude (h) igual a 5cm, teremos um total de 30cm : 5cm = 6 (seis) *classes*, cada uma de amplitude igual a 5cm, as quais compreenderão todos os dados da tabela primitiva.

As seis classes: de 150cm a 155cm; de 155cm a 160cm; de 160cm a 165cm; ..... ; de 175cm a 180cm, serão representadas por intervalos fechados à esquerda e abertos à direita, da seguinte maneira:

$$150 \vdash\!\!\!- 155 \qquad 155 \vdash\!\!\!- 160 \qquad 160 \vdash\!\!\!- 165 \qquad 165 \vdash\!\!\!- 170 \qquad 170 \vdash\!\!\!- 175 \qquad 175 \vdash\!\!\!- 180$$

Surge então a pergunta: A que classe pertence o dado 155cm, à primeira ou à segunda classe? Com a notação usada podemos, desde já, situar o valor 155cm (que é o limite superior da 1.ª classe) *na* 2.ª *classe*, como seu limite inferior. Estamos, agora, aptos para a construção do quadro relativo às *distribuições por freqüências* das 40 estaturas em estudo. Para isso, dispomos as *classes* numa 1.ª coluna encimada por um X que simboliza as classes (de estaturas) da distribuição; na 2.ª coluna é feita a *tabulação*, em que um dos operadores "canta" os dados da tabela primitiva (ou do rol) e o outro vai registrando por meio de "pauzinhos", marcação aliás conhecida em diversos jogos; na 3.ª coluna, encimada por F são escritos os totais de valores (freqüência) de cada classe. Temos, assim, o quadro:

h = 5	Classes (X)	Tabulação	Freqüências (F)
	150 ⊢ 155	☒ |	6
	155 ⊢ 160	☒ □	9
	160 ⊢ 165	☒ ☒ ☒ |	16
	165 ⊢ 170	☒	5
	170 ⊢ 175	⊔	3
	175 ⊢ 180	|	1
			$\Sigma F = 40$

NOTA:

$h = 5$ (amplitude de cada classe: 5cm).

$\Sigma F = 40$ (*Freqüência total:* soma de tôdas as freqüências que se distribuem).

A letra grega Σ, que se lê *sigma* maiúsculo ou somatório, é o símbolo matemático para indicar *soma*. Logo: ΣF indica a *freqüência total* de uma distribuição.

5. Ponto-médio de uma classe. É de muito interêsse, para os estudos que se seguem, *concentrar* os valores pertencentes a uma mesma classe no valor médio dessa classe. Surge, assim, um novo elemento denominado *ponto-médio* de uma classe, que é definido como o número que se obtém quando se soma ao limite inferior de uma classe a metade de sua amplitude. Indicando o ponto-médio por P_m, temos:

$$P_m = l_i + \frac{h}{2}.$$

Na prática o P_m de uma classe pode ser obtido efetuando-se a semi-soma dos l_i e l_s. Assim, por exemplo, na classe $150 \vdash 155$ o seu P_m é 152,5 $\left(\frac{150 + 155}{2}\right)$.

6. Outros tipos de freqüências. Numa distribuição destacamos, ainda, os seguintes tipos de freqüências:

1) Freqüências acumuladas. Acumular freqüências, numa distribuição, significa somar cada freqüência (absoluta) com a soma das que lhe são anteriores na distribuição. Ao

resultado dessa soma damos o nome de *freqüência acumulada* e indicaremos no quadro por F_a. É lógico que, ao acumular a freqüência da última classe de uma distribuição com as freqüências já acumuladas pela penúltima, obteremos a freqüência total (ΣF) da distribuição.

2) Freqüências relativas. Freqüência relativa de uma distribuição é o quociente da divisão de cada freqüência (absoluta) pela freqüência total. Indicação: F_r.

Logo:
$$F_r = \frac{F}{\Sigma F}.$$

Multiplicando-se cada freqüência relativa por 100, obteremos a *freqüência relativa percentual*. Indicação: $F_{\%}$.

Logo:
$$F_{\%} = 100 \times F_r.$$

Já podemos, neste instante, escrever no quadro de distribuições por freqüências, correspondente ao *exame-modêlo*, novas colunas destinadas aos elementos ora definidos: P_m, F_a, F_r e $F_{\%}$. Temos, então:

$h = 5$	X	P_m	F	F_a	F_r	$F\%$
	150 ⊢ 155	152,5	6	6	0,150	15,0
	155 ⊢ 160	157,5	9	15	0,225	22,5
	160 ⊢ 165	162,5	16	31	0,400	40,0
	165 ⊢ 170	167,5	5	36	0,125	12,5
	170 ⊢ 175	172,5	3	39	0,075	7,5
	175 ⊢ 180	177,5	1	40	0,025	2,5
			$\Sigma F = 40$		$\Sigma F_r = 1{,}000$	$\Sigma F\% = 100{,}0$

NOTA:

É evidente que:

1. A soma de tôdas as freqüências (ΣF_r) é igual à *unidade*.
2. A soma de tôdas as freqüências relativas percentuais ($\Sigma F\%$) é igual a 100.

É possível, agora, responder, por exemplo, às seguintes perguntas:

1.ª) *Quantas* alunas de estatura inferior a 165cm existem entre as 40 alunas?

Resposta: 31 (Basta procurar a F_a correspondente à classe $160 \vdash 165$).

2.ª) Que *porcentagem* representa sôbre as demais alunas, aquelas que possuem estaturas superiores a 175cm?

Resposta: 2,5% (Basta procurar a $F_{\%}$ correspondente à classe $175 \vdash 180$).

3.ª) Que porcentagem representa na distribuição, as alunas de estaturas menores que 1,70m?

Resposta: 90 % (Por quê?)

*c) **Análise das distribuições por freqüências.***

É uma das partes mais importantes do levantamento estatístico, pois, permite estudar as tendências características de cada distribuição, isoladamente ou em conjunto com outras. Com êsse objetivo são introduzidos *símbolos numéricos* que permitem:

1.º) Representar as distribuições por um *valor central*, mediante as *medidas de posição*: médias (aritmética, geométrica, harmônica), mediana e moda;

2.º) Reconhecer se há *dispersão* na distribuição, mediante as *medidas de dispersão* ou de *variabilidade*: desvio médio, desvio padrão e coeficiente de variação;

3.º) Verificar qual o *grau de assimetria* de uma distribuição, mediante a *medida de assimetria*: índice de assimetria.

Todos êsses símbolos numéricos, que são também denominados *elementos típicos* de uma distribuição, serão estudados separadamente na proporção da importância que representam num levantamento.

*d) **Conclusões sôbre os resultados obtidos.***

Os resultados obtidos, após as três fases iniciais de um levantamento estatístico, devem permitir *conclusões* que, além de estabelecerem um balanço da realidade, forneçam, principalmente, elementos para a indução e previsão futuras. Daí a necessidade de se avaliar o grau de precisão em que foram tomados os dados, corrigindo tanto quanto possível erros cometidos a fim de que o levantamento estatístico possa realmente preencher a sua finalidade precípua: *servir ao homem e à sociedade.*

§ 2. Representações gráficas.

7. Generalidades. Além da representação dos dados estatísticos por tabelas e quadros de distribuição de freqüências é muito comum apresentar êsses mesmos dados mediante formas ilustradas denominadas *gráficos*. Os gráficos têm a vantagem de produzir ao grande público uma impressão mais rápida e viva do que as tabelas comuns. Para isso êles devem ser *simples* e de *fácil compreensão.*

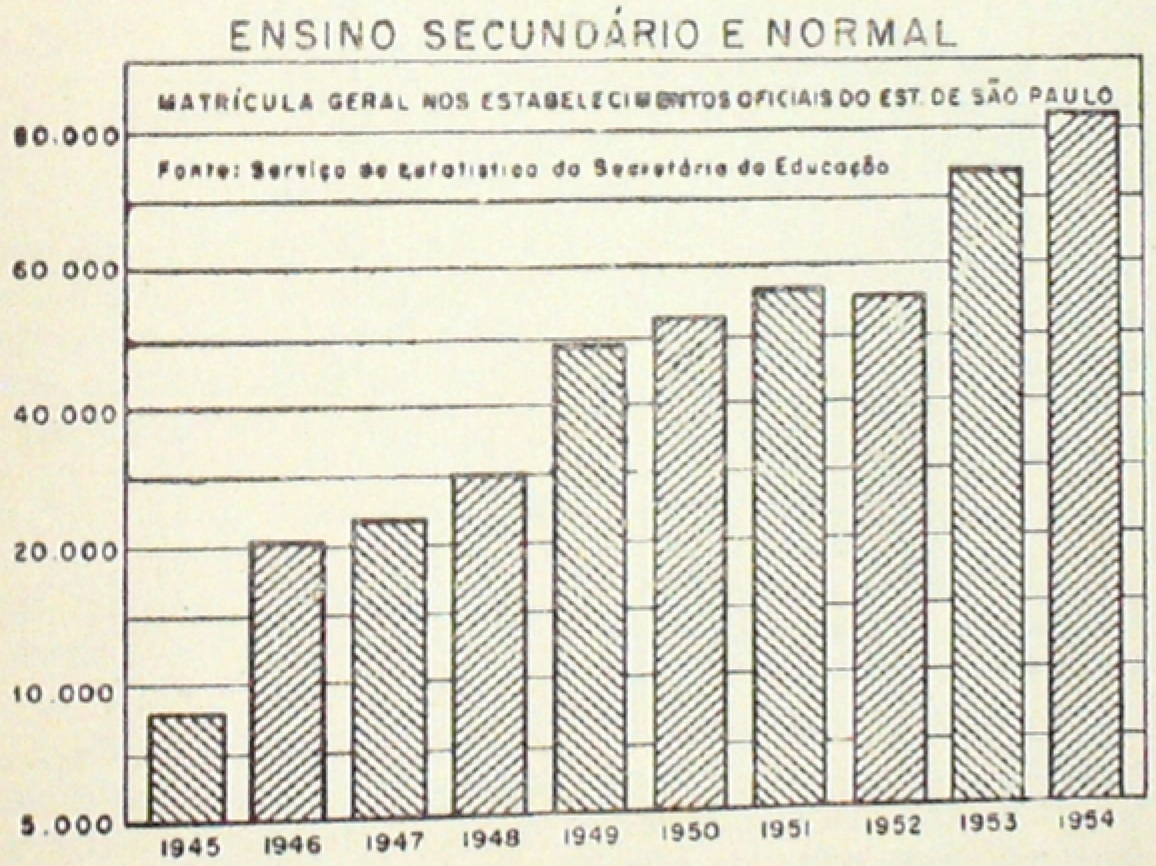


FIG. 29

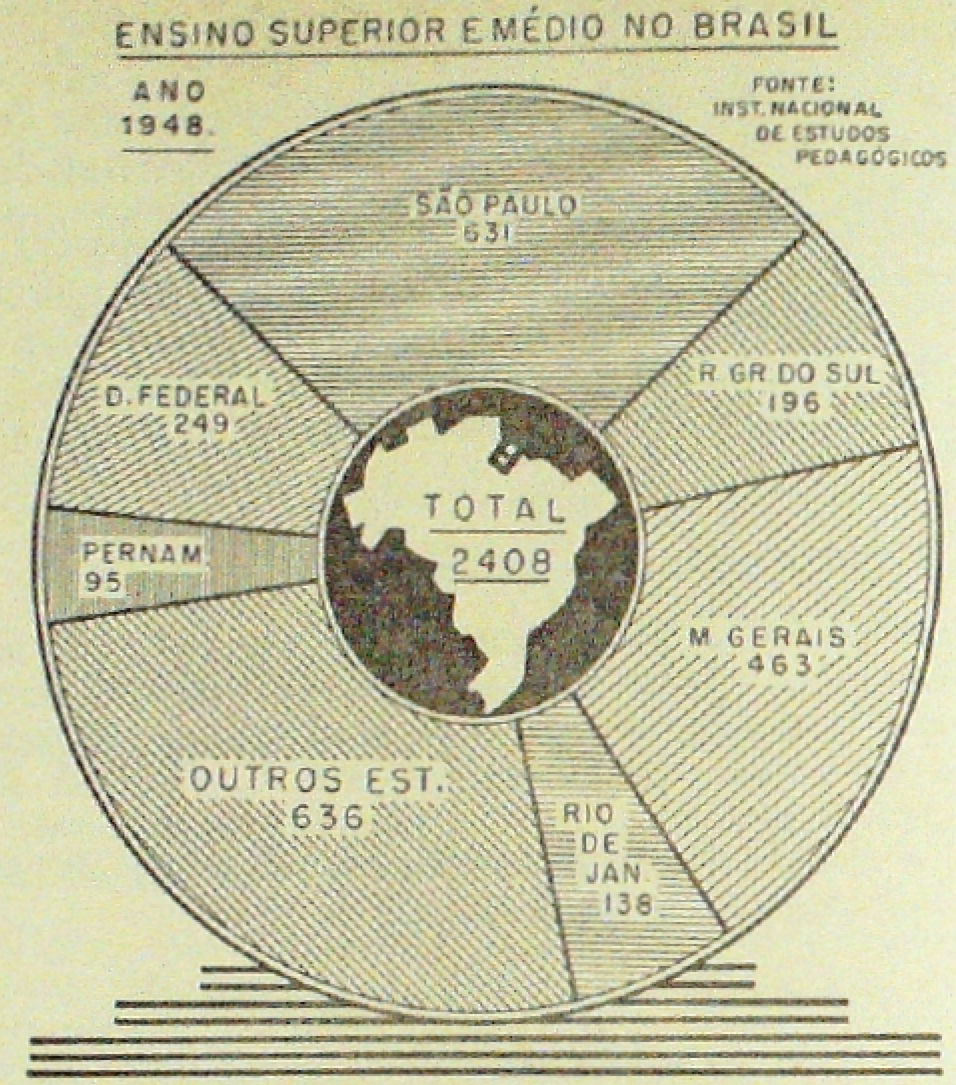


FIG. 30

Dividiremos os gráficos em duas categorias: os de *informação* e os de *análise*. Os primeiros se destinam ao grande público e, geralmente, são feitos com a finalidade de "chamar a atenção", enquanto que os gráficos de análise, tendo em vista os estudiosos, são baseados na precisão e no rigor matemático.

8. Alguns gráficos de informação.

a) Gráfico de barras ou de colunas: Compõe-se de retângulos de mesma base (arbitrária) e de alturas proporcionais aos valores (dados estatísticos) (fig. 29).

b) Gráfico em setores: Os valores são dispostos num círculo, onde o valor total equivale a uma amplitude de 360º. O cálculo do setor correspondente a cada valor é feito por uma regra de três (fig. 30).

c) Cartogramas: Têm por objetivo representar os dados estatísticos relacionados com áreas geográficas ou políticas (fig. 31).

d) Gráficos pictóricos: São os preferidos para esclarecimento do grande público, pois, os dados são representados por figuras semelhantes diretamente proporcionais aos valores que se representam (fig. 32).

9. Alguns gráficos de análise.

a) Diagramas cartesianos: São os mais usados em Estatística. Enquanto que os gráficos de informação se caracterizam na representação de valores que não se apresentam com continuidade, os diagramas cartesianos, como de resto todos os sistemas que usam *coordenadas*, representam a *varia-*

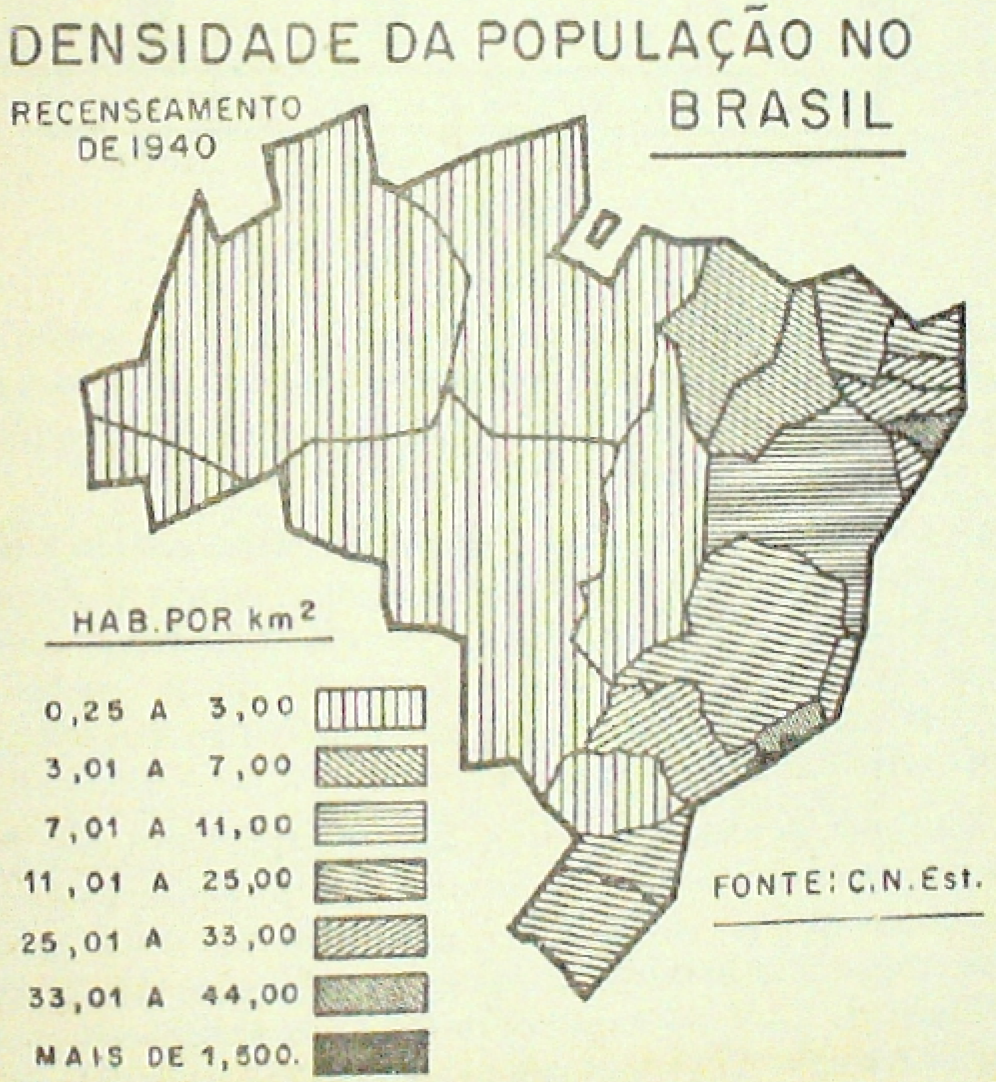


FIG. 31

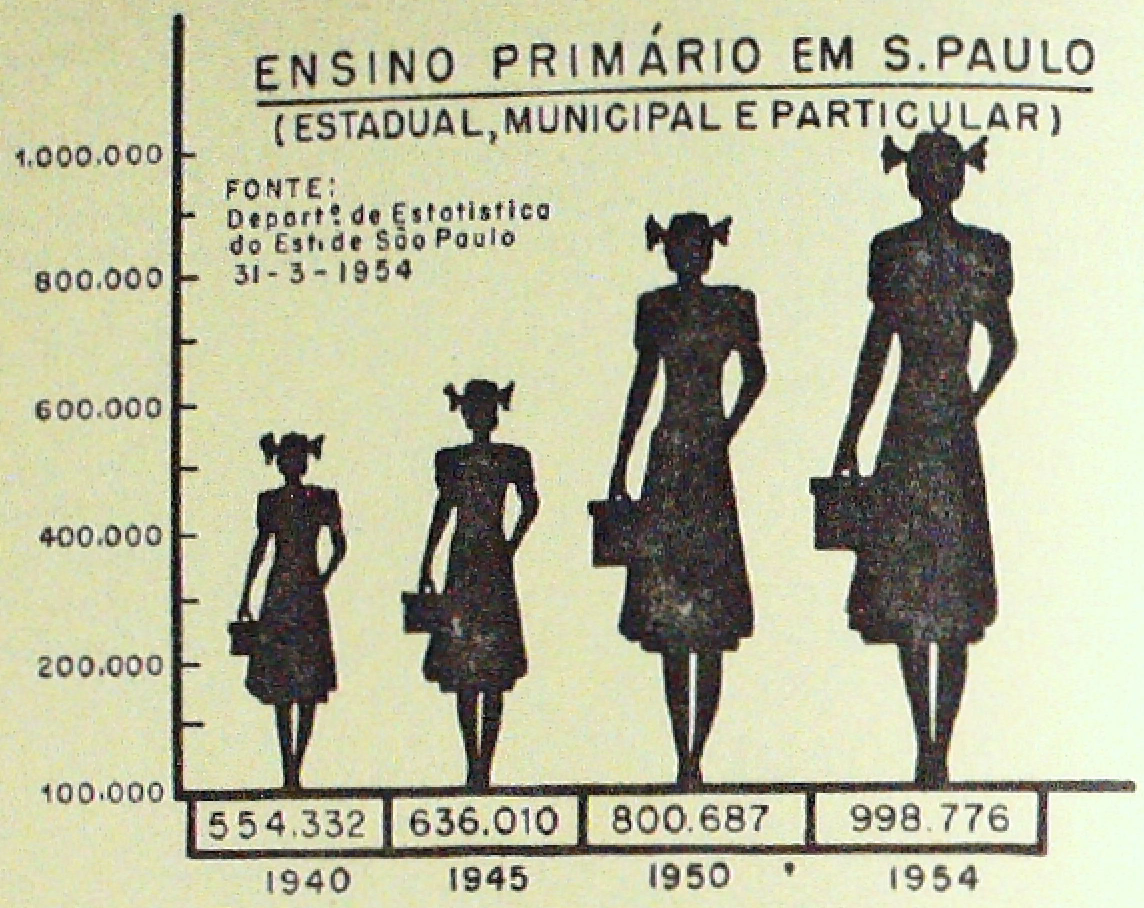


FIG. 32

ção contínua de um mesmo fenômeno. Pode-se, assim, ao variar uma das grandezas, estudar atentamente a variação dos valores correspondentes de uma outra que lhe esteja ligada funcionalmente, mediante um diagrama cartesiano.

Os elementos de referência dêsse diagrama são dois eixos (retas orientadas) que se interceptam num ponto 0 denominado origem. De preferência os eixos são tomados perpendiculares entre si e nesse caso o diagrama diz-se cartesiano ortogonal (fig. 33). O eixo $0x$ é denominado *eixo das abscissas* e o eixo $0y$, das *ordenadas*.

Qualquer ponto P do plano $x0y$, determinado por êsses eixos, é representado por dois números a e b que representam, respectivamente, as medidas dos segmentos $0A$ e $0B$ quando se traçam por P as paralelas aos eixos. O número a é a *abscissa* e o número b, a *ordenada* do ponto P. Ambos dizem-se *coordenadas cartesianas* de P em homenagem a Descartes (Cartesius em latim), seu introdutor. Indicação: $P(a,b)$.

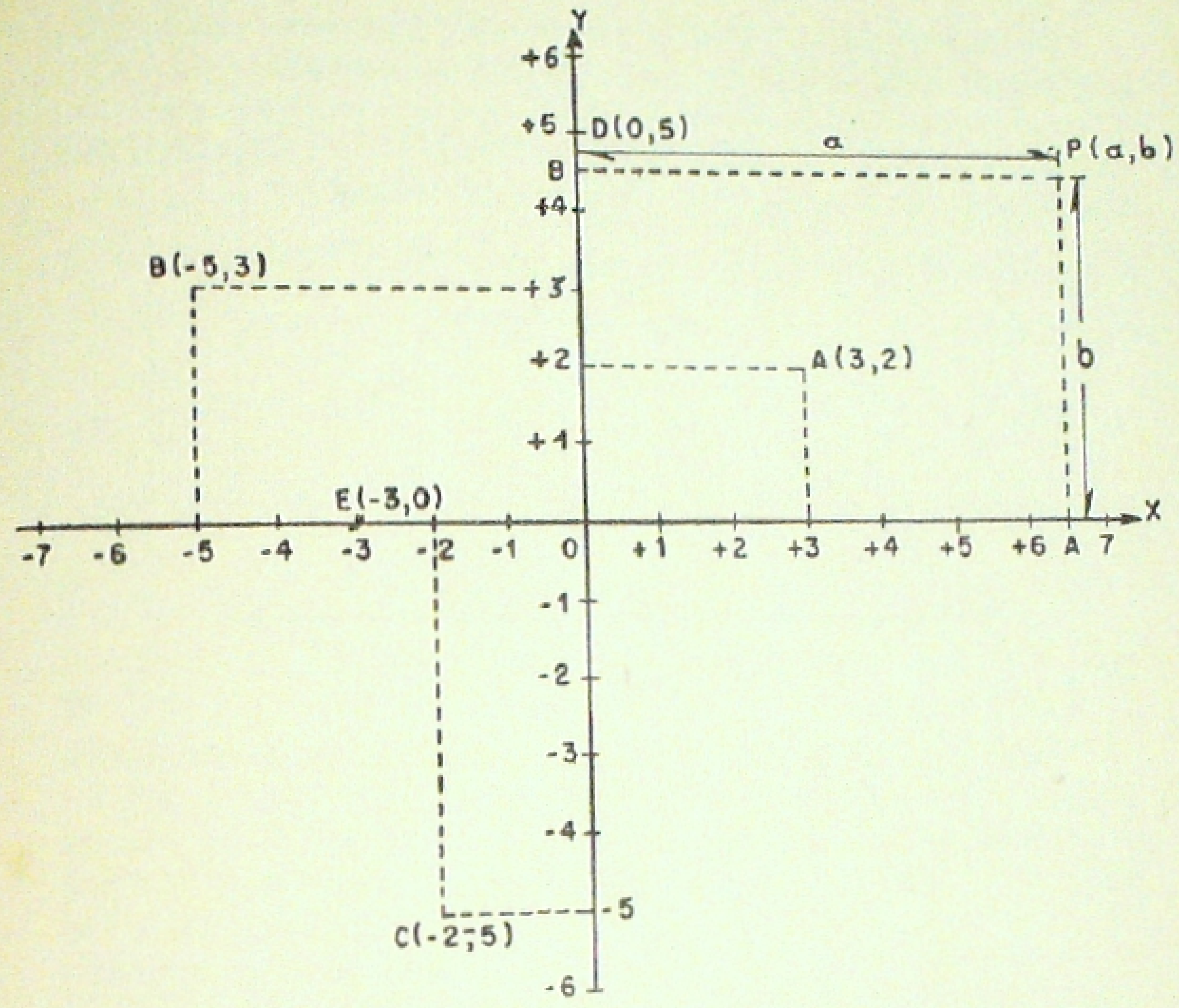


FIG. 33

Os eixos dividem o plano em quatro quadrantes, tais que:

I Qdte.	abscissa positiva ordenada positiva	III Qdte.	abscissa negativa ordenada negativa
II Qdte.	abscissa negativa ordenada positiva	IV Qdte.	abscissa positiva ordenada negativa

Os pontos de abscissas nulas estão sôbre o eixo das ordenadas ($0y$) e os de ordenadas nulas estão sôbre o eixo das abscissas ($0x$). A origem tem as coordenadas nulas. Na fig. 33 temos representados os pontos: $A(3,2)$; $B(-5,3)$; $C(-2, -5)$; $D(0,5)$ e $E(-3,0)$.

b) ***Diagramas polares:*** Se ao invés de dois números para representar um ponto de um plano usarmos *um número* e *um ângulo*, estaremos usando as coordenadas polares que caracterizam o diagrama polar (fig. 34).

Fixa-se sôbre uma reta orientada, tomada como eixo, um ponto 0 denominado *polo*; a reta é chamada *eixo polar*. A todo ponto P, do mesmo plano, fazemos corresponder um número a, que é a medida do segmento $0P$ e um ângulo α, que é a medida do ângulo $x0a$. Indicação: $P(a,\alpha)$.

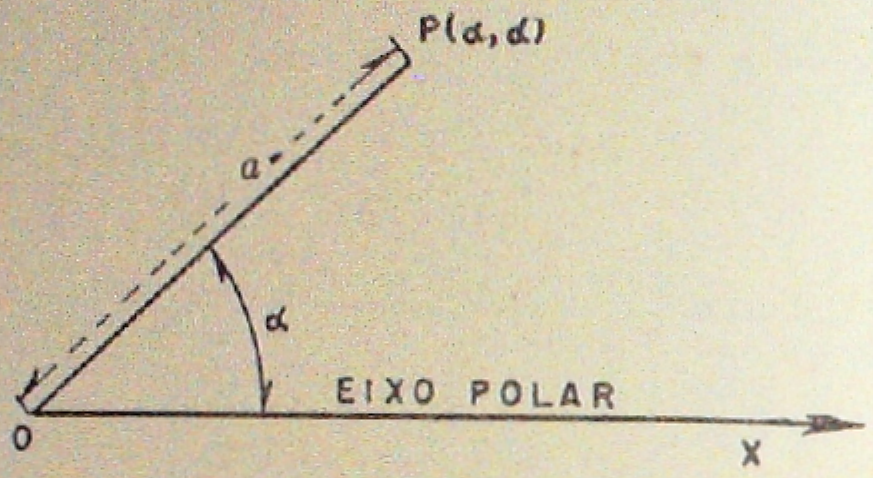


FIG. 34

Êste tipo de gráfico é usado quando o fenômeno que se estuda varia em certos intervalos constantes. Por exemplo, o *comparecimento dos alunos* durante o ano letivo (9 mêses) pode ser muito bem representado gràficamente (fig. 35) por um diagrama polar, onde o ano todo é indicado por uma circunferência (360°) e os mêses, por arcos de 40° (360°: 9). Sôbre a bissetriz de cada um dêsses ângulos (40°), a partir do vértice comum (no centro), marca-se um segmento proporcional à respectiva freqüência média mensal. Unindo-se os extremos livres de tais segmentos obtemos o *diagrama polar* correspondente.

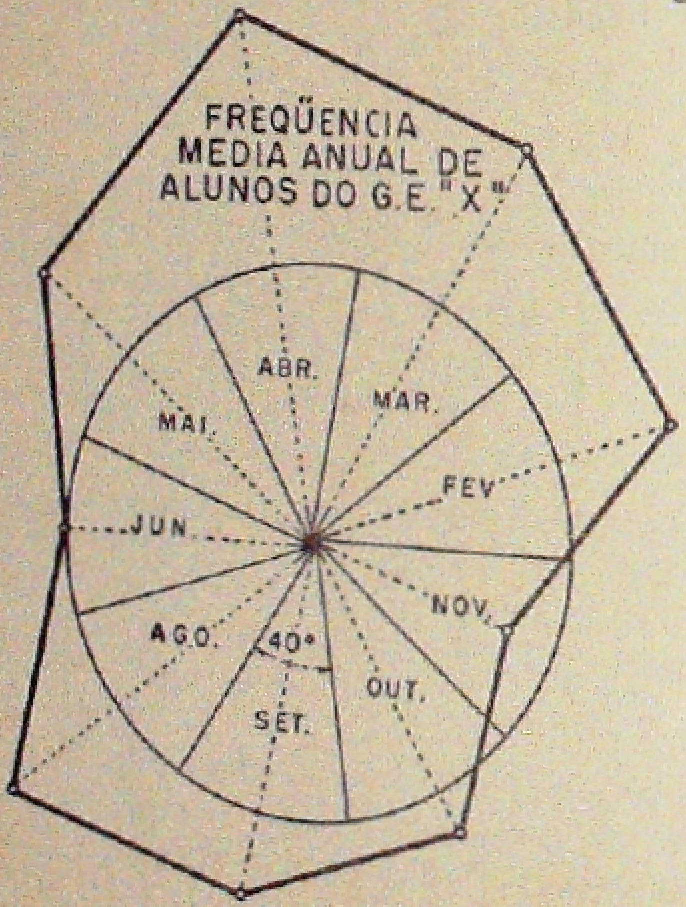


FIG. 35

c) Diagramas de curvas: No caso particular de variação sempre

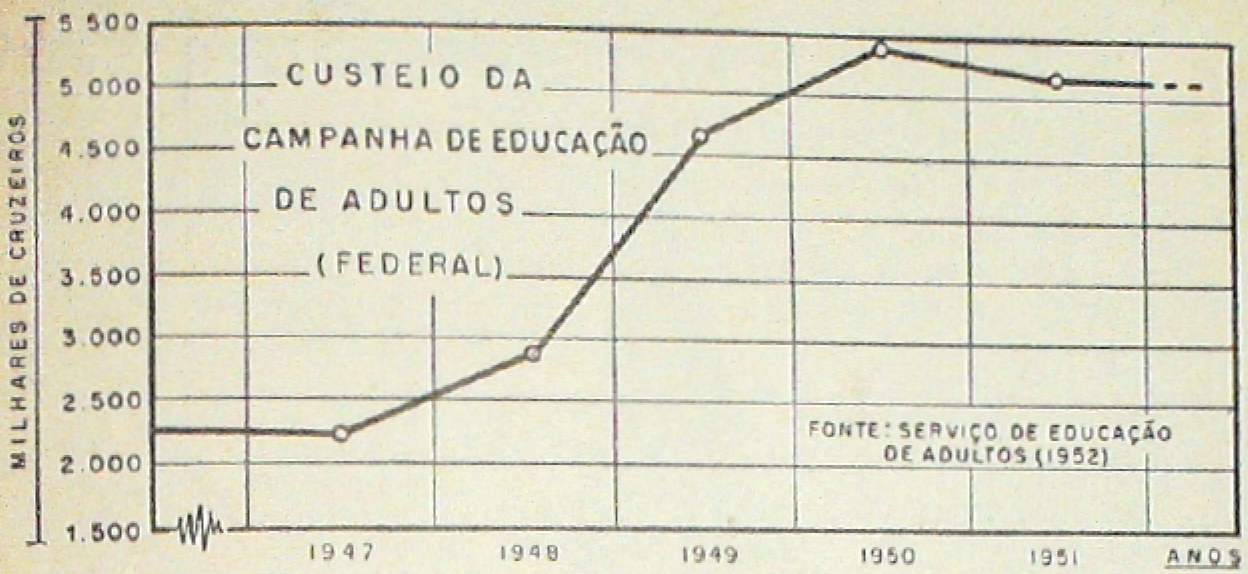


FIG. 36

contínua, e sòmente nesse caso, poderemos unir por segmentos de retas os diversos pontos obtidos por pares de coordenadas cartesianas e encontraremos um diagrama constituído por uma linha poligonal denominado *diagrama poligonal*. Costuma-se na prática chamar a essa diagonal de *curva*, e, portanto, impròpriamente, *diagrama de curvas*. Êsse fato provém de que, aumentados, cada vez mais, os fatos observados, a linha poligonal tende, cada vez mais, para uma *curva* que seria, assim, a *posição limite* do diagrama poligonal (fig. 36).

10. Representação gráfica das distribuições. As distribuições por freqüências construídas, até agora, por tabelas, também podem ser representadas por gráficos que permitam ao observador uma impressão geral da variação das mesmas.

a) Polígonos de freqüência: Num sistema de coordenadas cartesianas ortogonais tomamos sôbre o eixo das abscissas segmentos proporcionais aos valores dos pontos-médios das classes de valores assumidos pelo atributo, que se estuda, e, sôbre o eixo das ordenadas, segmentos proporcionais às freqüências respectivas. Unindo-se os pontos obtidos determinamos um diagrama poligonal que, convencionalmente, é *fechado* no eixo das abscissas pelo ponto-médio da classe imediatamente inferior à classe inicial e pelo ponto-médio da classe imediatamente superior à classe final. Dessa forma formamos um polígono que é denominado *polígono de freqüência*. Na figura 37 temos o *polígono de freqüência* relativo à distribuição, que corresponde ao nosso exemplo-modêlo.

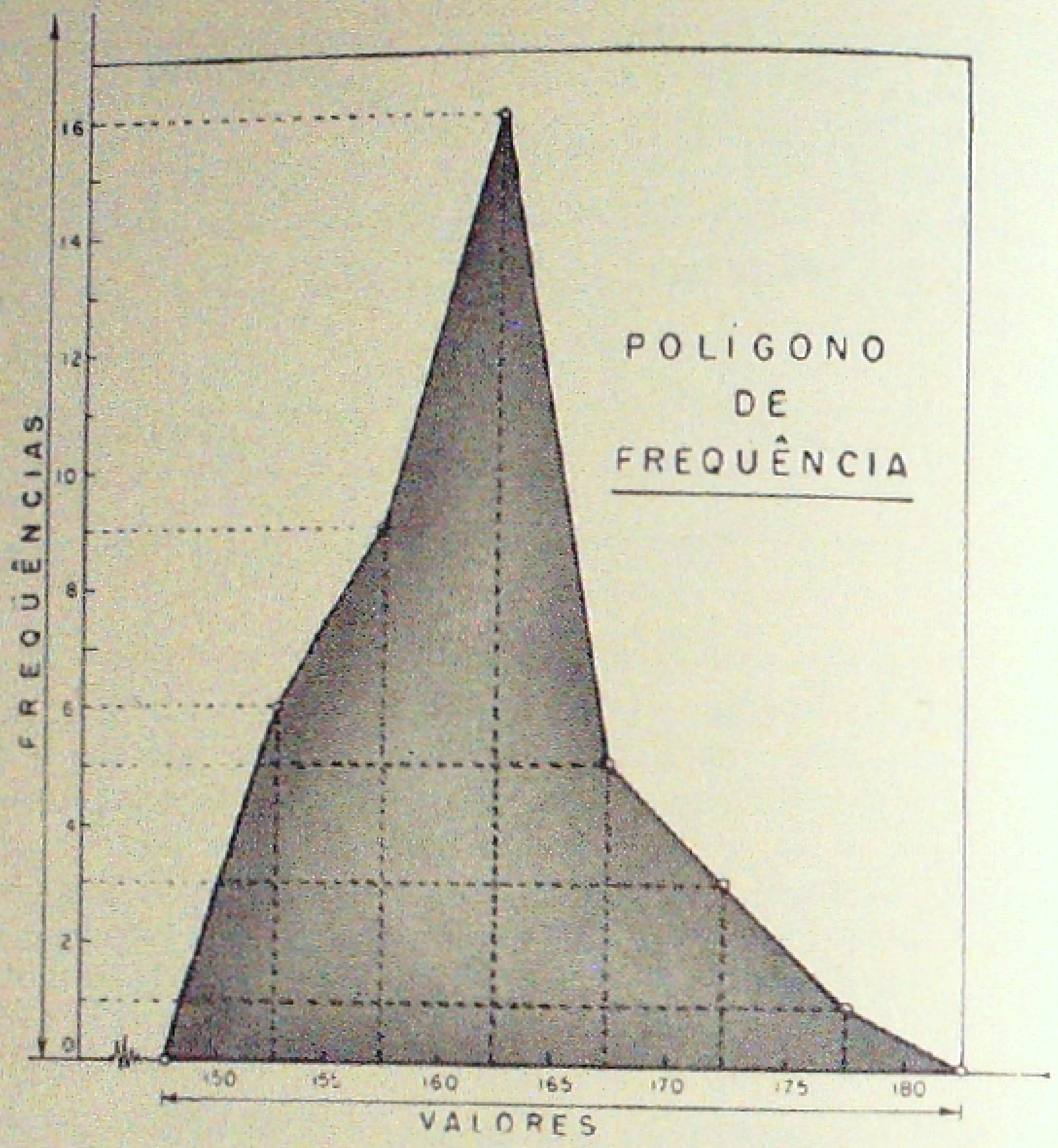


FIG. 37

OBSERVAÇÃO: No eixo das abscissas, entre a origem 0 e o limite inferior da 1.ª classe, consideraremos, sempre que fôr necessário, um intervalo interrompido que quer significar as unidades existentes entre 0 e êsse mesmo limite. O mesmo acontecerá com o eixo das ordenadas, quando preciso.

Entre as *vantagens* oferecidas pelos *polígonos de freqüência*, destacamos:

1.ª) dão ao observador uma idéia da curva que representaria o fenômeno estudado;

2.ª) servem para confrontar diversas distribuições entre si.

É muito comum representar, num mesmo sistema de coordenadas cartesianas, duas ou mais distribuições. Os gráficos respectivos são destacados por linhas cheias, interrom-

pidas, pontilhadas, etc., permitindo comparar, em pouco tempo, a marcha das variações das diversas distribuições.

b) ***Histogramas:*** É o gráfico construído mediante um número de retângulos contíguos, igual ao número das classes da distribuição. As suas bases, iguais para todos, têm por medida comum a amplitude usada para as classes e as suas alturas são diretamente proporcionais às freqüências das respectivas classes (fig. 38).

PEARSON denominou êsse gráfico de *histograma*, em virtude das freqüências serem representadas por *áreas*. Sendo a área de cada retângulo, que compõe o histograma, proporcional a uma certa freqüência da distribuição é evidente que a área de todo o histograma é proporcional à soma de tôdas as freqüências ou seja ao número total de casos estudados pela

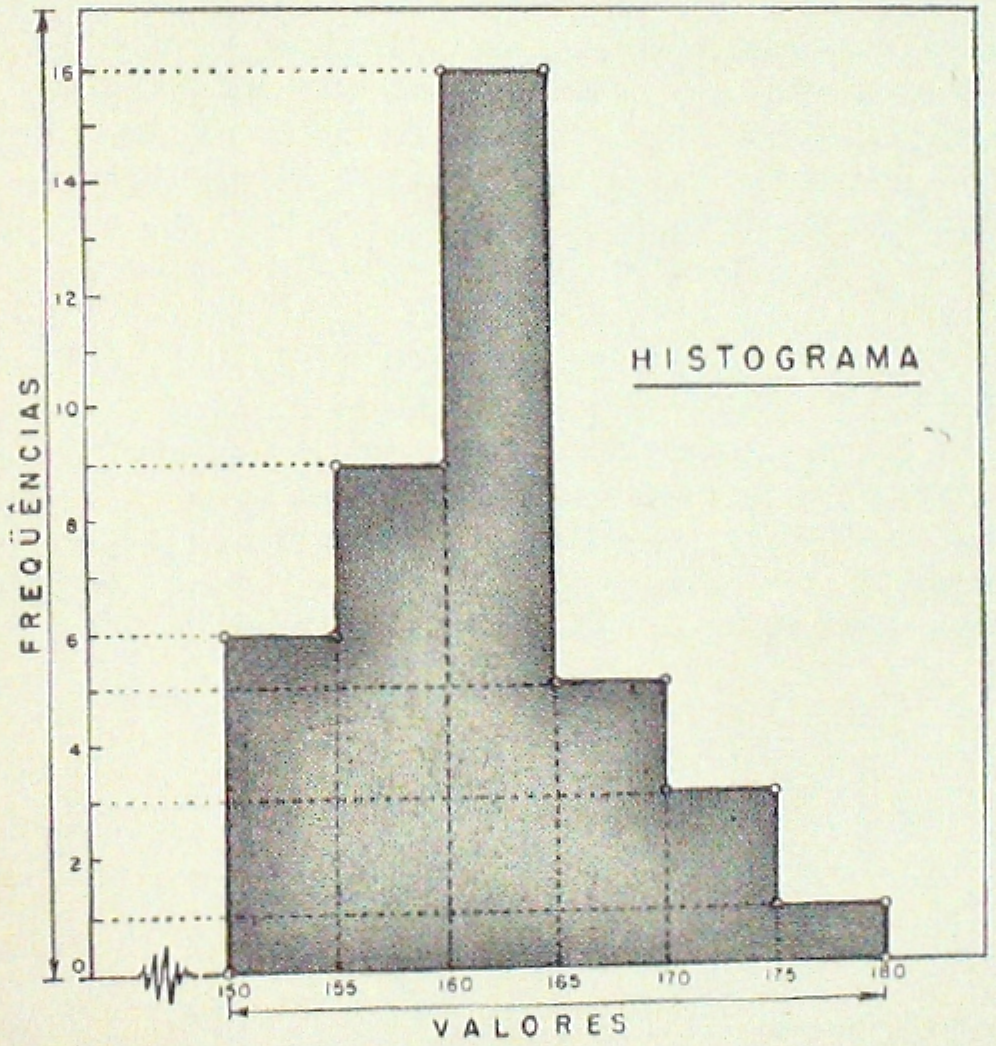


FIG. 38

distribuição. Na figura 38 temos o histograma correspondente à distribuição do *exemplo-modêlo.*

OBSERVAÇÕES:

1.ª) No caso da amplitude de classe não ser a mesma para tôdas as classes, a altura do correspondente retângulo seria obtida pelo quociente da divisão da freqüência da classe, pela respectiva amplitude.

2.ª) Comumente não se fazem figurar no histograma as linhas auxiliares, a fim de melhor ressaltar o fenômeno que se estuda. O contôrno externo do histograma é denominado *poligonal característica.*

A *vantagem* oferecida pelo histograma está no fato de que, qualquer que seja a fração do intervalo, tomado sôbre o eixo das abscissas, a *área do retângulo* construído é *proporcional* à mesma fração da área representativa da freqüência total.

*c) **Curvas de freqüência:*** Se o número total de casos observados *aumentar* cada vez mais, isto é, tender ao infinito, ao mesmo tempo em que as amplitudes de classe se tornem cada vez *menores*, a poligonal característica (contôrno do histograma) tenderá a se *confundir* com uma curva denominada, geralmente, *curva de freqüência.* A figura 39 representa o gráfico relativo a um grande número de estaturas e que se aproxima da curva de freqüência ora definida. É notório que a área da superfície compreendida entre a curva e duas ordenadas quaisquer está em rigorosa correlação com o número de observações (freqüência) correspondente ao intervalo considerado (área hachuriada na fig 39).

As curvas de freqüência mais ou menos regulares, que traduzem séries de observações numerosas, apresentam grande número de *formas* diferentes que dependem da distribuição ser *simétrica, moderadamente assimétrica* ou *extremamente assimétrica.* Para cada uma das formas as curvas recebem deno-

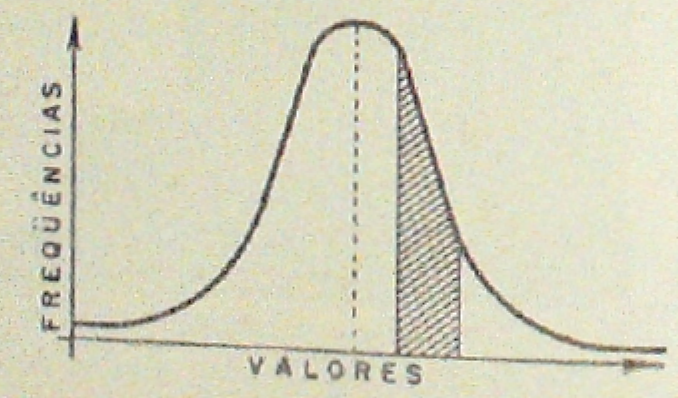


FIG. 39

minações já consagradas com os nomes: curvas em sino, curvas em J (jota), curvas em U, etc. ... Vejamos algumas das mais importantes curvas de freqüência:

1. CURVAS EM SINO. São gráficos relativos a distribuições simétricas, isto é, àquelas que apresentam a freqüência máxima no centro e, diminuindo gradativamente, à medida que se atingem as classes correspondentes aos menores e aos maiores valores da distribuição. São numerosos os fenômenos que apresentam formas de distribuição em sino ou campanular: a *estatura*, por exemplo (fig. 39), é um atributo que se distribui simètricamente, não exigindo, portanto, um grande número de observações para se terem curvas bastante regulares; os *pesos dos adultos*, as *distribuições dos quocientes de inteligência* (Q.I) apresentam-se, da mesma maneira, por curvas em sino.

OBSERVAÇÃO: A curva em sino, também denominada *normal* ou de Gauss, tem uma importância bem declarada no estudo da *probabilidade*, por ser a curva de distribuição dos fenômenos que ocorrem *por acaso* (estatura, inteligência, pressão atmosférica, etc..).

2. CURVAS MODERADAMENTE ASSIMÉTRICAS. (fig. 40-*a* e 40-*b*). Quando a distribuição é moderadamente assimétrica, isto é, afasta-se ligeiramente do tipo de simetria, já descrito, há um decréscimo mais acentuado da freqüência, de um lado do máximo da curva do que do outro. Esta é a forma que mais ocorre nas distribuições de freqüência, sendo encontrada na maioria dos fenômenos estudados. No campo educacional, as notas de aproveitamento obtidas por meio de *testes* (Testes de Binet), apresentam-se numa distribuição moderadamente assimétrica, havendo uma diminuição gradativa da freqüência para as notas baixas, enquanto que, para as notas altas, há uma diminuição bem mais acentuada.

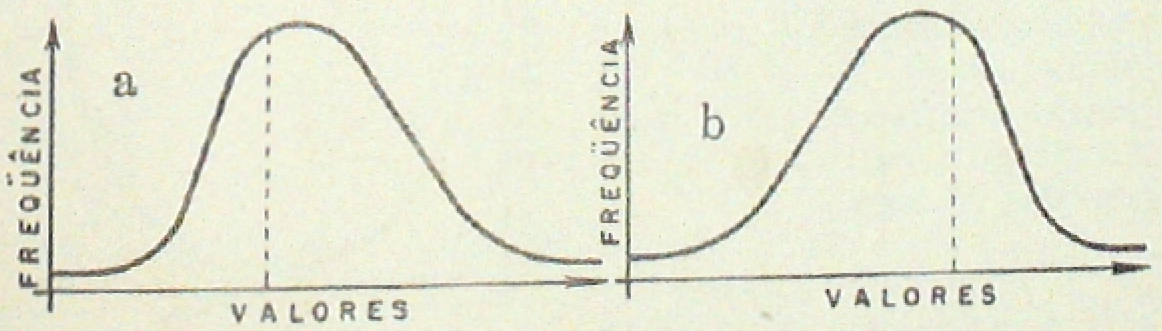


FIG. 40

3. Curvas em J. São os gráficos relativos a distribuições extremamente assimétricas, isto é, àquelas em que as freqüências aumentam ràpidamente até o máximo em uma das extremidades (fig. 41-*a*) ou diminuem ràpidamente até o mínimo (fig. 41-*b*).

Êstes tipos de distribuições são mais comuns nos fenômenos econômicos, como, por exemplo, na distribuição de rendas individuais, onde a freqüência vai diminuindo, à medida que o seu valor aumenta. Outro exemplo, nos é dado pelos gráficos concernentes às alturas dos edifícios, onde a freqüência vai rareando com o aumento da altura, a partir do máximo de um só andar.

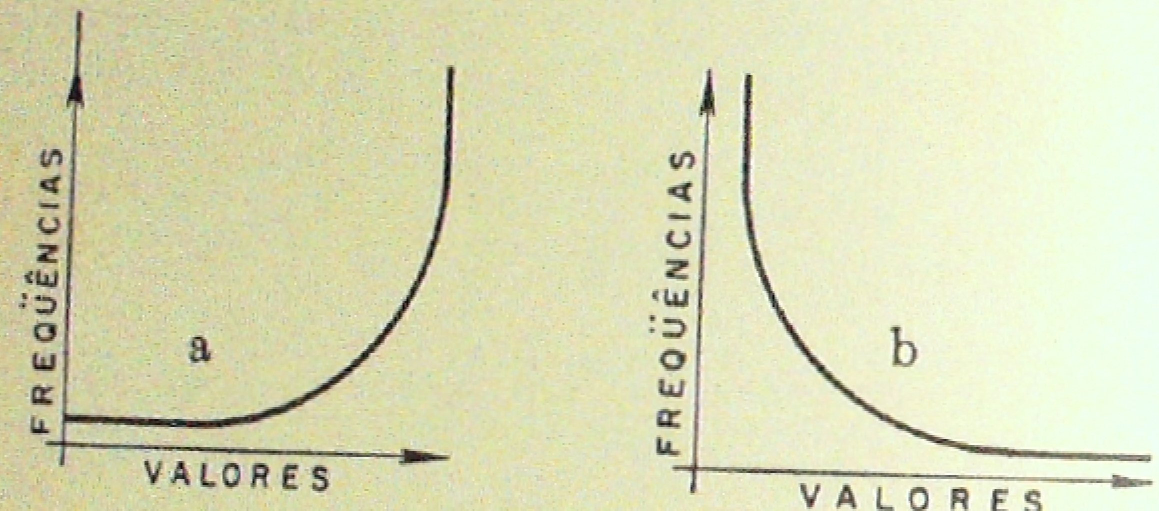


FIG. 41

4. Curvas em U (fig. 42). São gráficos correspondentes a distribuições, que apresentam o mínimo de freqüência na parte central, e, dois máximos nos extremos. Estas curvas diferem substancialmente dos tipos até agora estudados. Por exemplo, a forma da distribuição da mortalidade por idades é em U, pois, a freqüência é máxima logo depois do nascimento, vai diminuindo progressivamente até atingir o mínimo entre 10 e 15 anos, para depois recomeçar a subir até atingir o outro máximo, depois dos 55 anos.

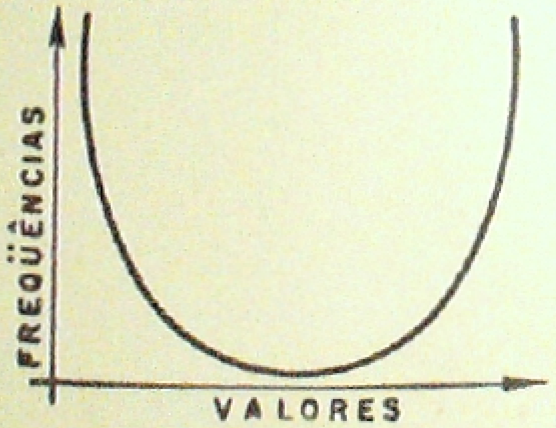


FIG. 42

d) ***Ogiva de Galton:*** É o gráfico, imaginado por Galton,

relativo às ***distribuições de freqüências acumuladas***. Daí o fato de ser também chamada *curva de acumulação de freqüências*. As freqüências acumuladas figuram como ordenadas e os limites superiores de cada classe, como abscissas. Podem-se acumular as freqüências no sentido dos valores crescentes (ogiva crescente) ou no sentido dos valores decrescentes (ogiva decrescente). Na figura 43, temos a Ogiva de Galton (crescente) relativa à distribuição do nosso *exemplo-modêlo*.

Observação: Pode-se, para melhor estudar uma distribuição, construir uma *escala relativa de freqüência*, sob forma percentual, indicando por 100 a ordenada máxima que acumula tôdas as freqüências (100%). Assim, a ogiva pode determinar, imediatamente: o valor da distribuição, que ocupa a posição do meio, isto é, aquêle que supera 50 % dos valores e é superado pelos restantes 50 %, denominado *mediana*; o valor que tem 25 % dos valores antes de si e os 75 % restantes depois, denominado *primeiro quartil*; o valor que tem 75 % dos valores antes de si e os restantes 25 % depois, denominado *terceiro quartil*. Êstes valores típicos, de uma distribuição serão estudados, pormenorizadamente, no próximo parágrafo

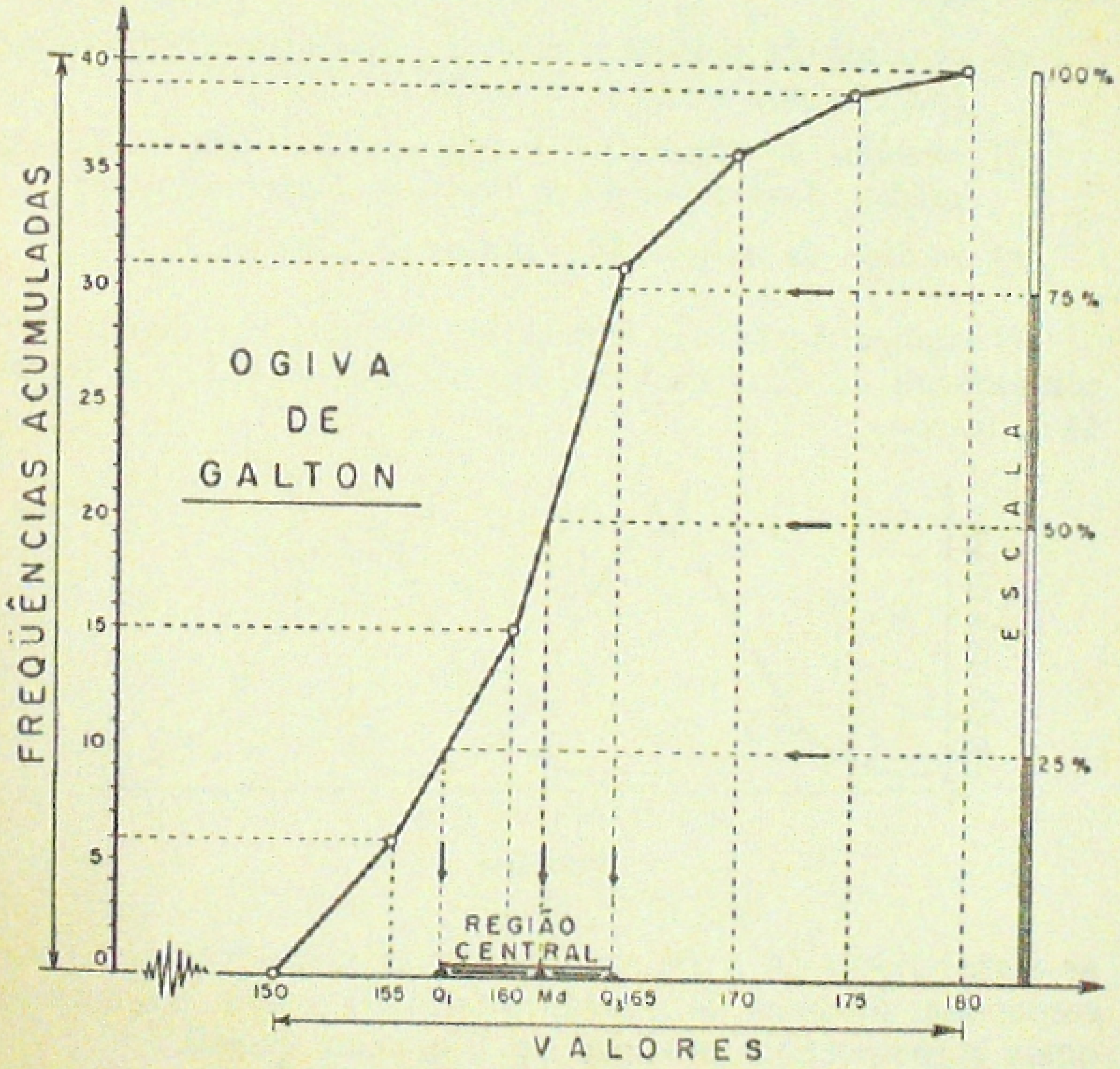


FIG. 43

§ 3. Elementos típicos de uma distribuïção de freqüência: medidas de posição, de dispersão e de assimetria.

11. Generalidades. O estudo das distribuições de freqüência, até êste instante, tem-nos permitido descrever, de um modo geral, os grupos de valores representativos da variação de um fenômeno. Assim, por exemplo, podemos saber se a maior *concentração* de valores de uma distribuição se situa no início, no meio ou no final, ou ainda, se se distribui igualmente. A fim de poder ressaltar as tendências características de cada distribuição, isoladamente, ou em confronto com outras, temos necessidade de introduzir símbolos, que possam traduzir numèricamente essas tendências, denominados *elementos típicos* da distribuição. São elementos típicos de uma distribuição de freqüência:

a) *medidas de posição:* médias (aritmética, geométrica, harmônica), mediana e moda;

b) *medidas de dispersão:* amplitude semi-quartil, desvio médio, desvio padrão e coeficiente de variação;

c) *medida de assimetria:* índice de assimetria.

O conhecimento dos elementos típicos é decisivo para a comparação de duas distribuições de freqüência. Na figura 44-*a*, temos:

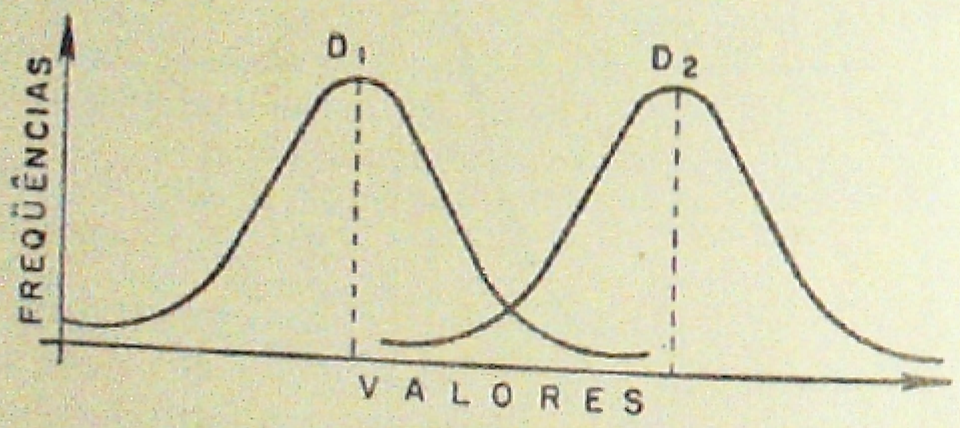


FIG. 44-a

as distribuições D_1 e D_2 em *posições diferentes* (às quais correspondem medidas de posição *diferentes*) e *formas iguais* (às quais correspondem medidas de dispersão *iguais*).

Na figura 44-*b*, temos: as distribuições D_3 e D_4 na *mesma posição* (às quais correspondem medidas de posição *iguais*) e *formas diferentes* (às quais correspondem medidas de dispersão *diferentes*).

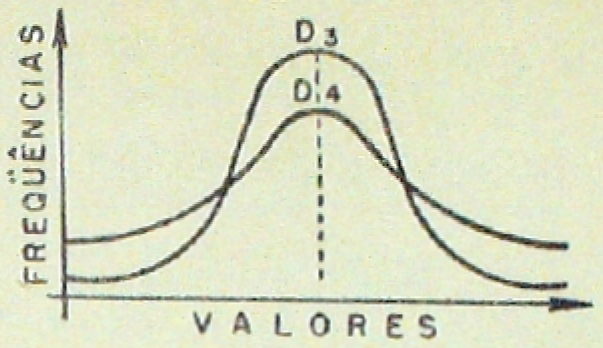


FIG. 44-b

MEDIDAS DE POSIÇÃO

I) A média aritmética e o seu cálculo

12. Média aritmética simples. Chama-se *média aritmética* de uma série de valores ao quociente da divisão da soma dêsses valores pelo seu número. Indicando os diversos valores (n), que uma variável x da série pode assumir por: $x_1, x_2, x_3, \ldots, x_n$, e, por M_a, a média aritmética, temos:

$$M_a = \frac{x_1 + x_2 + x_3 \ldots + x_n}{n}$$

ou

$$M_a = \frac{\sum_1^n{}_i\, x_i}{n}.$$

onde a letra i (ou outra qualquer) indica o índice de cada x ao variar de 1 a n. Exemplo: A média aritmética dos valores: 35, 40, 50, 60 e 65, que representam as notas de um aluno numa certa disciplina, é:

$$M_a = \frac{35 + 40 + 50 + 50 + 65}{5} = \frac{240}{5} = 48.$$

13. Média aritmética ponderada. No caso dos valores virem aferidos por *pesos*, que são números indicadores da intensidade do valor no conjunto, a média aritmética diz-se *ponderada*. A média aritmética ponderada é igual ao quociente da divisão, cujo dividendo é constituído da soma dos produtos dos valores pelos respectivos pesos, e, cujo divisor é a soma dos pesos. Exemplo: Calcular a média aritmética obtida por um aluno que obteve as seguintes notas em Matemática, constantes na relação abaixo, juntamente com os pesos respectivos:

1ª exame: 5 (pêso 2); 2ª exame: 4 (pêso 3); exame oral: 6 (pêso 3); média mensal: 5 (pêso 2).

$$M_{a.p.} = \frac{5\times2 + 4\times3 + 6\times3 + 5\times2}{2+3+3+2} = \frac{50}{10} = 5{,}0 .$$

14. Média aritmética de uma distribuição de freqüência. Quando os valores de uma série estão *agrupados em classes*, como aliás acontece mais freqüentemente, a média aritmética pode ser calculada fàcilmente. Êsse cálculo pode ser feito de dois modos: *processo longo* e *processo breve*.

a) Processo longo.

Usa-se a hipótese fundamental da tabulagem: *os valores de uma classe são todos concentrados no ponto-médio dessa classe.* Portanto, a média aritmética da distribuição será dada dividindo-se a soma dos produtos, cujos fatôres são o ponto-médio da classe e a respectiva freqüência, pela soma total das freqüências. Em símbolos, temos:

$$\boxed{M_a = \frac{\Sigma P_m \times F}{\Sigma F}}$$

Como exercício, calculemos a média aritmética do nosso *exemplo-modêlo*, pelo *processo longo*:

$h=5$	X	P_m	F	$P_m \times F$
	150 ⊢ 155	152,5	6	915,0
	155 ⊢ 160	157,5	9	1 417,5
	160 ⊢ 165	162,5	16	2 600,0
	165 ⊢ 170	167,5	5	837,5
	170 ⊢ 175	172,5	3	517,5
	175 ⊢ 180	177,5	1	177,5
			$\Sigma F = 40$	$\Sigma P_m \times F = 6465{,}0$

$\boxed{M_a = \frac{\Sigma P_m \times F}{\Sigma F}}$ $\quad M_a = \frac{6\,465{,}0}{40} = 161{,}625.$ Portanto:

$$M_a = 161{,}625\text{cm}$$

b) Processo breve.

Usa-se a seguinte propriedade fundamental da média aritmética de uma distribuição: *"a soma algébrica dos afastamentos dos valores de uma série em relação a M_a é nula".* Estamos chamando de *afastamento* (outros chamam *desvio*), de um valor da série, a diferença entre êsse valor e um outro valor qualquer da série. Exemplo: Na série de notas

35, 40, 50, 50 e 65

cuja $M_a = 48$, temos os seguintes afastamentos dêsses valores em relação à M_a:

$$\begin{aligned} 35 - 48 &= -13 \\ 40 - 48 &= -\ \ 8 \\ 50 - 48 &= \ \ \ \ 2 \\ 50 - 48 &= \ \ \ \ 2 \\ 65 - 48 &= \ \ 17 \end{aligned}$$

cuja soma algébrica $(-21+21)$ é *nula*.

Dêsse modo, pode-se calcular a M_a de uma distribuição de valores, evitando-se os cálculos demorados do processo longo. Para isso indicamos um valor qualquer da distribuição como média aritmética. Estando os valores *agrupados em classes*, escolhemos o P_m de uma delas, sendo de preferência tomado o P_m, que corresponde à freqüência (F) mais alta. Se o P_m escolhido coincidir com a M_a da distribuição, o problema está resolvido. Caso contrário, estaremos na dependência de uma possível *correção C* (para mais ou para menos). Logo:

$$M_a = P_m + C .$$

Vejamos, agora, a determinação da correção C: como os afastamentos (a), em relação ao valor escolhido como média aritmética, são múltiplos da amplitude (h) da classe, segue que, dividindo os afastamentos (a) pela amplitude (h), obteremos, para quociente $\left(\frac{a}{h}\right)$, a série dos números inteiros que será positiva (+1, +2, +3, ...), abaixo do valor da média e negativa (−1, −2, −3, ...), acima. Indicando êsses quocientes por α e multiplicando, a seguir, cada α pela freqüência (F) da classe correspondente, temos que, a soma algébrica dos produtos obtidos ($\Sigma\alpha F$), dividida pela soma de tôdas as freqüências (ΣF), dá um número que, multiplicado pela amplitude h (pela qual dividimos no início), fornece a correção a ser feita na média aritmética escolhida. Logo, em símbolos, vem:

$$\boxed{M_a = P_m + h \times \frac{\Sigma\alpha F}{\Sigma F}}$$

Temos, assim, a seguinte *regra* para o cálculo da M_a pelo *processo breve*:

1. Escolhe-se o P_m de uma das classes (de preferência a de mais alta freqüência) como média imaginária;
2. No quadro das distribuições de freqüência, numa coluna relativa aos quocientes α, escreve-se: um *zero* na linha correspondente à classe onde se encontra o P_m escolhido; a série −1, −2, −3, ... logo acima do zero e a série +1, +2, +3, ... logo abaixo;
3. Efetuam-se os membros de cada α pela freqüência (F) correspondente à mesma linha; somam-se algèbricamente êsses produtos ($\Sigma\alpha F$) e divide-se o total pela soma das freqüências (ΣF); o resultado obtido $\left(\frac{\Sigma\alpha F}{\Sigma F}\right)$, multiplicado pela amplitude (h), dá a correção C;
4. A soma algébrica do P_m escolhido com a correção C é a M_a procurada, isto é:

$$M_a = P_m + h \times \frac{\Sigma\alpha F}{\Sigma F}.$$

Como prática, calculemos a média aritmética do *exemplo-modêlo* pelo *processo breve*.

$h=5$	X	P_m	F	α	$\alpha . F$
	150 ⊢ 155	152,5	6	−2	−12
	155 ⊢ 160	157,5	9	−1	−9
	160 ⊢ 165	162,5	16	0	0
	165 ⊢ 170	167,5	5	+1	+5
	170 ⊢ 175	172,5	3	+2	+6
	175 ⊢ 180	177,5	1	+3	+3
			$\Sigma F = 40$		$\Sigma\alpha F = -7$

$$\boxed{M_a = P_m + h \times \frac{\Sigma\alpha F}{\Sigma F}}$$

$$M_a = 162{,}5 + 5 \times \frac{(-7)}{40}.$$

$$M_a = 162{,}5 + 5 \times (-0{,}175).$$

$$M_a = 162{,}5 - 0{,}875 = 161{,}625.$$

Portanto:

$$M_a = 161{,}625\,\text{cm}$$

NOTA: Limitar-nos-emos a definir e dar exemplos simples das outras médias: *geométrica* e *harmônica*, em virtude da pouca aplicação que teriam essas médias aos iniciantes de um curso de estatística.

Chama-se *média geométrica* de uma série de n valores a raiz de índice n do produto dêsses n valores. Indicação:

$$M_g = \sqrt[n]{x_1 \times x_2 \times \ldots\ldots x_n}.$$

Se os valores forem sòmente dois números, como, por exemplo, 4 e 16, o cálculo da M_g se resume na extração de uma raiz quadrada, isto é:

$$M_g = \sqrt{4 \times 16} = \sqrt{64} = 8.$$

Chama-se *média harmônica* de uma série de n valores o inverso da média aritmética de seus inversos. Assim, por exemplo, a M_h de 5 e 8, onde a M_a de seus inversos é $\dfrac{\frac{1}{5}+\frac{1}{8}}{2}$, é igual a $\dfrac{80}{13}$. Indicação: $M_h = \dfrac{n}{\Sigma \frac{1}{x_n}}$.

II) A mediana e o seu cálculo. Quartis, decis e centis.

15. Definição. Chama-se *mediana* de uma série de valores dispostos em ordem crescente (ou decrescente) ao valor que ocupa a *posição do meio* dessa série. Indicando-se a série de valores por: $x_1, x_2, x_3, \ldots, x_n$ e, a mediana por M_d, temos que essa nova medida M_d é um valor *contado*, que vem precedido e seguido pelo mesmo número de valores.

Se existir um número ímpar $(2n+1)$ de valores na série, só teremos um valor $(n+1)$, que preenche a definição. Exemplo: Sejam as 15 notas seguintes, já dispostas em rol:

$$\underbrace{25,\ 30,\ 45,\ 45,\ 50,\ 50,\ 50}_{\textit{sete valores}},\ \underset{\substack{\downarrow \\ M_d}}{60},\ \underbrace{60,\ 65,\ 70,\ 70,\ 80,\ 85,\ 90}_{\textit{sete valores}}.$$

A M_d dessa série é a *nota* 60, que ocupa a posição do meio, pois, existem sete notas que a precedem e sete que a sucedem.

Se o número de valores fôr par $(2n)$, todo valor entre as posições n e $n+1$ satisfaz à definição. Nesse caso, convenciona-se como mediana a média aritmética dos dois valores centrais. Exemplo: Nas notas seguintes:

$$20,\ 25,\ 30,\ 45,\ 45,\ 50,\ 50,\ \underset{\substack{\downarrow \\ M_d}}{\underbrace{50,\ 60}},\ 60,\ 65,\ 70,\ 70,\ 80,\ 85,\ 90$$

a M_d é a *nota* 55 (média aritmética das notas 50 e 60).

16. Cálculo da mediana de uma distribuição por freqüência. Se os valores de uma série estão *agrupados em classes*, o cálculo da M_d é feito mediante a seguinte *regra*:

1. Determinam-se as freqüências acumuladas (coluna F_a);
2. Divide-se a soma das freqüências (ΣF) por 2, obtendo-se o quociente N;

3. Procura-se onde N está incluído na coluna F_a, marcando, com uma flecha, a classe correspondente à freqüência acumulada, imediatamente superior a N. No caso de se encontrar uma F_a exatamente igual a N, a mediana será o limite superior dessa classe. Caso contrário subtrái-se de N a freqüência acumulada que lhe é imediatamente inferior e que será indicada por F_a^*. Obtém-se assim uma diferença que será indicada por d. Logo: $d = N - F_a^*$;
4. Arma-se a seguinte proporção: $F : h :: d : x$ que se lê: a freqüência simples F da classe marcada está para a amplitude h assim como a diferença d está para x;
5. A mediana será o número que se obtém somando x, da proporção acima ao limite inferior (que será indicado por l_i^*), da classe marcada, isto é,

$$M_d = l_i^* + x \qquad \text{onde} \quad x = \frac{h \cdot d}{F}.$$

Querendo estabelecer uma fórmula para o cálculo da M_d de uma distribuição de freqüência, vem, de acôrdo com o exposto na regra:

$$M_d = l_i^* + \frac{h \cdot d}{F} \text{ e como } d = N - F_a^*, \quad \boxed{M_d = l_i^* + \frac{h\,(N - F_a^*)}{F}}$$

Como aplicação, calculemos a mediana do nosso *exemplo-modêlo*:

$h=5$	X	P_m	F	F_a
	150 ⊢ 155		6	6
	155 ⊢ 160		9	15
* (M_d)	160 ⊢ 165		16	31 → (M_d)
	165 ⊢ 170		5	36
	170 ⊢ 175		3	39
	175 ⊢ 180		1	40
			$\Sigma F = 40$	

CÁLCULOS:

$40 \div 2 = 20$ (quociente N, que na coluna F_a está entre 15 e 31).
$20 - 15 = 5$ (diferença d).
$16 : 5 :: 5 : x$ $(F : h :: d : x)$.

$$\therefore x = \frac{5 \times 5}{16} = \frac{25}{16} = 1{,}5625$$

Logo:
Querendo aplicar a fórmula

$$M_d = l_i^* + \frac{h\,(N - F_a^*)}{F}$$

vem:

$$M_d = 160 + \frac{5\,.(20-15)}{16} = 160 + \frac{25}{16} = 161{,}5625$$

ou

$$M_d = 161{,}5625\text{cm}.$$

17. Quartis, decis e centis. Recebem êsses nomes outras medidas que separam os valores de um rol em *quatro partes iguais* (quartis), ou em *dez* (decis), ou ainda em *cem partes iguais* (centis ou percentis).

18. Cálculo dos quartis. Os *quartis* são os três valores que separam uma série em *quatro partes iguais*. Costumeiramente são indicados, respectivamente, por Q_1, Q_2 e Q_3. É evidente que o *primeiro quartil* Q_1 é o valor precedido por $\frac{1}{4}$ do total e seguido pelos restantes $\frac{3}{4}$ dos valores; o *segundo quartil* Q_2 é o valor que ocupa a posição do meio e portanto é a própria mediana da distribuição $(Q_2 = M_d)$; e o *terceiro quartil* Q_3 é o valor precedido por $\frac{3}{4}$ de todos os valores e seguido por $\frac{1}{4}$.

Regra para o cálculo do primeiro quartil Q_1:

1. Determinam-se as freqüências acumuladas (coluna F_a);
2. Divide-se a soma das freqüências (ΣF) por 4, obtendo-se o quociente N;
3. Procede-se, a seguir, de modo análogo ao usado na determinação da mediana.

Regra para o cálculo do terceiro quartil Q_3:

1. Determinam-se as freqüências acumuladas (coluna F_a);

2. Divide-se a soma das freqüências por 4 e multiplica-se o resultado por 3;
3. Procedimento análogo aos anteriores.

Como prática, calculemos o Q_1 e o Q_3 do *exemplo-modêlo*:

$h=5$	X	P_m	F	F_a
	150 ⊢ 155		6	6
* (Q_1)	155 ⊢ 160		9	15 → (Q_1)
* (Q_3)	160 ⊢ 165		16	31 → (Q_3)
	165 ⊢ 170		5	36
	170 ⊢ 175		3	39
	175 ⊢ 180		1	40
			$\Sigma F = 40$	

CÁLCULO DE Q_1

$40 \div 4 = 10$ (N)
$10 - 6 = 4$ (d)
$F : h :: d : x$

ou

$9 : 5 :: 4 : x$

$$x = \frac{20}{9} = 2{,}222$$

$$\therefore Q_1 = 155 + 2{,}222$$

ou

$$Q_1 = 157{,}222\text{cm}$$

CÁLCULO DE Q_3

$40 \div 4 = 10$ (N)
$3 \times 10 = 30$
$30 - 15 = 15$ (d)
$F : h :: d : x$

ou

$16 : 5 :: 15 : x$

$$x = \frac{75}{16} = 4{,}6875$$

$$\therefore Q_3 = 160 + 4{,}6875$$

ou

$$Q_3 = 164{,}6875\text{cm}$$

19. Cálculo dos decis. Os *decis* são os nove valores que separam uma série em *dez partes iguais*. Indicação: D_1, D_2, D_3, ..., D_9. Naturalmente o D_5 é a própria mediana. O cálculo de um decil é feito dividindo-se a soma das freqüências por 10 e multiplicando-se o quociente obtido por 1, 2, 3, ... 9 conforme o decil que se queira. No resto segue a mesma marcha dos cálculos anteriores.

20. Cálculo dos centis. Os *centis* ou *percentis* são os noventa e nove valores que separam uma série em *cem partes iguais*. Indicação: $C_1, C_2, C_3, \ldots, C_{99}$. É evidente, agora, que o C_{50} é a própria mediana. O cálculo de um centil guarda analogia com o estudado na determinação dos decis, desde que se divida a soma das freqüências por 100 e multiplique-se o quociente obtido por 1, 2, 3, ..., 99 conforme o centil que se procura.

Como aplicação, calculemos o D_7 e o C_{25} do *exemplo-modêlo*:

$h=5$	X	P_m	F	F_a
	150 ⊢ 155		6	6
	155 ⊢ 160		9	15
	160 ⊢ 165		16	31
	165 ⊢ 170		5	36
	170 ⊢ 175		3	39
	175 ⊢ 180		1	40
			$\Sigma F = 40$	

Cálculo de D_7

$40 \div 10 = 4$

$7 \times 4 = 28$

$28 - 15 = 13$

$F : h :: d : x$

ou

$16 : 5 :: 13 : x$

$$x = \frac{65}{16} = 4{,}0625$$

$\therefore\ D_7 = 160 + 4{,}0625$

ou

$D_7 = 164{,}0625\,\text{cm}$

Cálculo de C_{25}

$40 \div 100 = 0{,}4$

$25 \times 0{,}4 = 10$

$10 - 6 = 4$

$F : h :: d : x$

ou

$9 : 5 :: 4 : x$

$$x = \frac{20}{9} = 2{,}222$$

$\therefore\ C_{25} = 155 + 2{,}222$

ou

$C_{25} = 157{,}222\,\text{cm}$

Nota: É evidente que o $C_{25} = 157{,}222$ tem o mesmo valor do Q_1.

21. Interpretação gráfica. Gràficamente podemos representar a mediana, os quartis, bem como os decis e, mais raramente, os centis de uma distribuição de freqüência, mediante a Ogiva de Galton (fig. 43). A mediana será a abscissa que corresponde à ordenada equivalente a 50% da distribuição total; os quartis Q_1 e Q_3 serão as abscissas que correspondem às ordenadas equivalentes, respectivamente, a 25% e 75% da distribuição total.

III) A moda e o seu cálculo

22. Definição. A *moda*, também chamada *norma* de uma distribuição, é o valor da série que se apresenta com *maior freqüência*. No sentido lato da palavra *o que está na moda* significa o *valor dominante* da distribuição. Indicação: M_o.

A rigor, a moda não é uma medida empregada em um número pequeno de observações, como acontece com a média aritmética e com a mediana, e, por essa razão, contentâmo-nos, nas aplicações à Educação, com valores aproximados, já que o cálculo de seu valor exato não é feito por processos elementares.

23. Cálculo empírico. Fórmula de Pearson. Existem fórmulas, como a de Czuber e a de King (*), que calculam a M_o com certa precisão, sendo, porém, de deduções complicadas. Na prática, determinamos a moda de uma distribuição, procurando o valor ou a classe de valores que apresente a máxima freqüência. Essa classe é denominada *modal* e o seu ponto-médio, *moda bruta*, pois, representa, na verdade, uma aproximação grosseira da moda verdadeira.

Em o nosso *exemplo-modêlo*, temos:

classe modal: 160 ⊢ 165 e *moda bruta:* 162,5cm.

(*) Fórmula de *Czuber:* $Mo = l_i^* + \dfrac{h\,(Fmax - Fant)}{2\,Fmax - (Fant + Fpost)}$

onde
- l_i^* é o limite inferior da classe de maior freqüência
- $Fmax$ é a freqüência máxima
- $Fant$ é a freqüência imediatamente inferior a $Fmax$
- $Fpost$ é a freqüência imediatamente superior a $Fmax$

Fórmula de *King:* $Mo = l_* + \dfrac{h\,.\,Fpost}{Fant + Fpost}$

Para as distribuições levemente assimétricas pode-se aferir melhor a moda usando a curva de freqüência, onde a M_o é sempre a abscissa da ordenada máxima (que corresponde à freqüência máxima). KARL PEARSON verificou, experimentalmente (fig. 45), que, nas curvas de freqüência moderadamente assimétricas (que correspondem a distribuições de freqüência levemente assimétricas), *a distância entre a média aritmética e a moda ($M_a - M_o$) é sempre cêrca de três vêzes maior que a distância entre a média aritmética e a mediana*, isto é:

$$M_a - M_o = 3(M_a - M_d).$$

Dessa igualdade, tiramos o valor da M_o,

$$\boxed{M_o = 3M_d - 2M_a}$$

que é a *fórmula empírica de Pearson.*

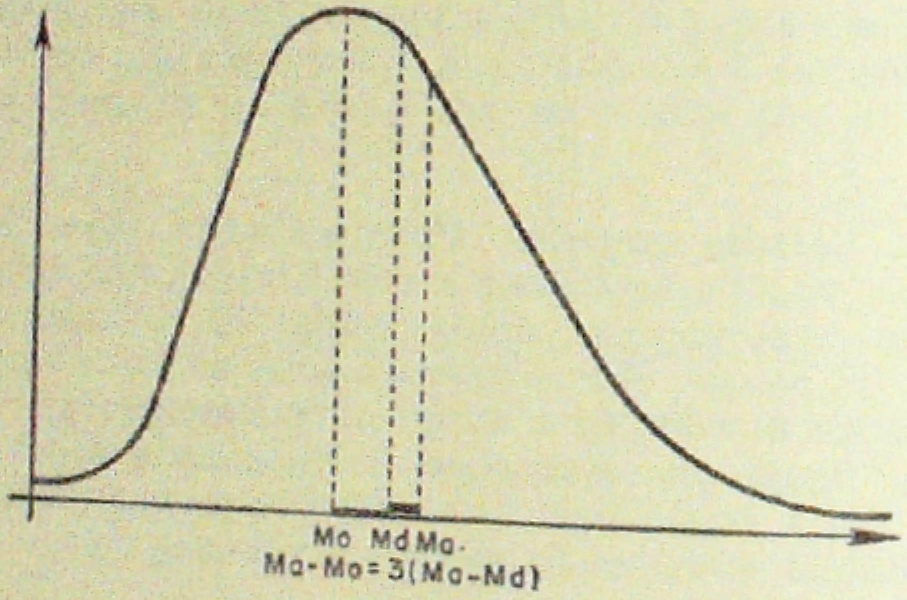


FIG. 45

É lógico que, para as distribuições exatamente simétricas, as medidas de posição: M_a, M_d e M_o, *coincidem*.

Aplicando a fórmula de Pearson na determinação da moda de nosso *exemplo-modêlo*, vem:

$$M_o = 3 \times 161{,}5625 - 2 \times 161{,}625$$
$$M_o = 484{,}6875 - 323{,}250$$

ou $\qquad \boldsymbol{M_o = 161{,}4375}$**cm.**

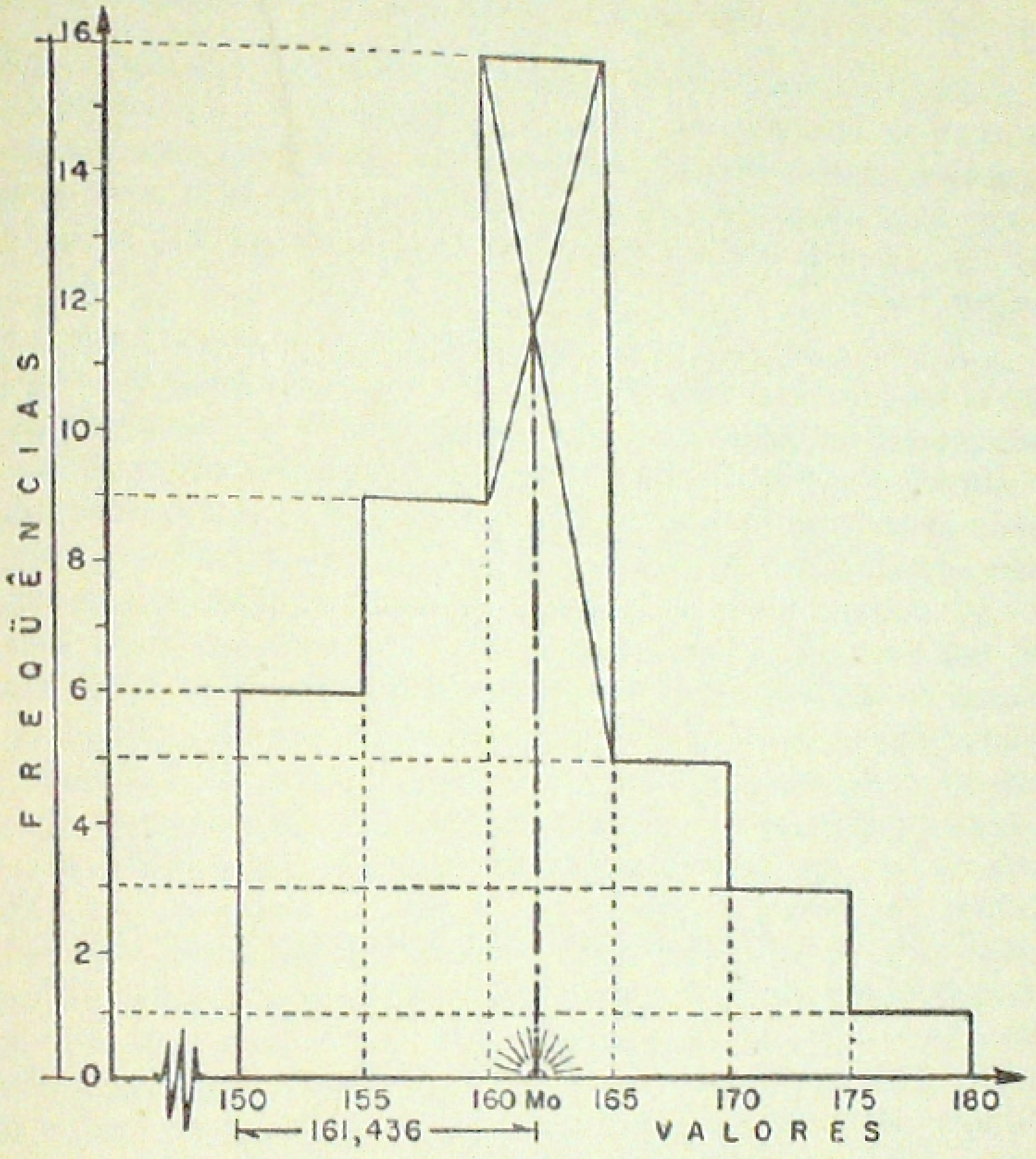


FIG. 46

24. Interpretação gráfica. Gràficamente, pode-se obter a moda de uma distribuição a partir do histograma correspondente a essa distribuição. Basta considerar a abscissa do ponto de intersecção (fig. 46) dos segmentos, cujos extremos são: do primeiro, os vértices representativos dos limites superiores do retângulo de maior área e do retângulo contíguo anterior; do segundo, os vértices representativos dos limites inferiores do retângulo de maior área e do retângulo contíguo posterior.

Uso das medidas de posição. Região central

25. Generalidades. Tôdas as medidas de posição que acabamos de expor (M_a, M_d, M_o) pretendem ser o *valor representativo* da distribuição que se estuda. Outrossim, visam favorecer as comparações das distribuições entre si e com êsse objetivo podemos dizer que tôdas elas têm o mesmo grau de importância.

A média aritmética é, geralmente, a mais empregada por ser o seu cálculo o mais fácil dos três e por ser também a mais compreensível pelos que têm algum estudo de matemática. Todavia, é preciso lembrar que, nos problemas educacionais, onde os valores extremos das distribuições, que se estudam, não apresentam a mesma importância que os valores centrais, é mais usada a mediana em lugar da média aritmética. Assim, se, por exemplo, quiséssemos saber o aproveitamento de uma classe de alunos, numa determinada disciplina, da qual são conhecidas as notas obtidas pelos alunos, o uso da média aritmética pode dar-nos uma idéia pouco precisa do aproveitamento da classe, pois, bastaria a existênciá de uma nota bem alta ou bem baixa para modificar sensìvelmente a média aritmética das notas, enquanto que o uso da mediana, por não sofrer esta da influência direta dos valores extremos, refletirá melhor o andamento da classe bem como os processos de ensino em vigor. Diga-se de passagem que a mediana tem sido a média mais usada nas nossas Escolas Militares, como aferição do aproveitamento de seus alunos e dos métodos de ensino a que estão os mesmos sujeitos.

A moda, apesar de maior instabilidade que as outras duas, tem a sua significação, como valor representativo, bem compreensível, por ser o valor dominante da série.

26. Região central. Atribui-se o nome de *região central* ou *região normal* de uma distribuição de freqüência, ao conjunto de valores compreendidos desde o primeiro quartil Q_1 até o terceiro quartil Q_3. Realmente, entre Q_1 e Q_3 se encontra a *metade* de todos os valores da distribuição, pois, antes de

Q_1 estão situados $\frac{1}{4}$ dos valores e depois de Q_3 mais outro $\frac{1}{4}$, ou seja fora do intervalo Q_1Q_3 está situada a metade $\left(\frac{1}{4} + \frac{1}{4} = \frac{1}{2}\right)$ dos valores que compõem a distribuição e, por conseguinte, entre Q_1 e Q_3 estará a outra metade. Convencionalmente, os valores pertencentes à região central são denominados *normais* e os que estão fora, *não normais*, sendo *supernormais* os valores posteriores a Q_3 e *subnormais*, os anteriores a Q_1.

Para a distribuição de nosso *exemplo-modêlo*, como:

$$Q_1 = 157{,}222\text{cm} \quad \text{e} \quad Q_3 = 164{,}6875\text{cm},$$

temos para a amplitude da região normal:

$$Q_3 - Q_1 = 164{,}6875\text{cm} - 157{,}222\text{cm} = 7{,}4655\text{cm}$$

e dizemos que

50 % das alunas observadas (que pertencem à região normal) medem de 157,222cm a 164,6875cm;

25 % das alunas observadas medem menos de 157,222cm;

25 % das alunas observadas medem mais de 164,6875cm.

MEDIDAS DE DISPERSÃO

27. Dispersão ou variabilidade de uma distribuição por freqüência. Duas distribuições de freqüência, embora com a mesma média aritmética, podem apresentar uma *variação de valores* bem diferente em tôrno dessa média. Por exemplo, consideremos como amostras representativas as notas obtidas por dois grupos de 10 alunos, um do 1.º Normal "A" e o outro do 1.º Normal "B". Estas notas dispostas em ordem crescente originam as seguintes séries:

1.º Normal "A": 35, 40, 45, 50, 50, 55, 60, 65, 70, 80.

1.º Normal "B": 10, 20, 30, 40, 50, 60, 70, 80, 90, 100.

Estas duas séries, que possuem o mesmo número de têrmos, admitem a mesma média aritmética ($M_a = 55$), apesar de apresentarem variações bem diferentes em tôrno dessa média, como é fácil de se observar. Assim notamos que a primeira série apresenta notas *mais uniformes* com uma *pequena variação* em tôrno da M_a, enquanto que a segunda apresenta notas *bem variadas* indicando uma grande *dispersão* em tôrno da M_a.

Pelo fato da média aritmética, assim como as demais médias de posição, não possibilitarem exprimir essa variabilidade, temos necessidade de estabelecer novos símbolos numéricos que permitam medir êste outro aspecto de uma série estatística: o da *variabilidade* ou *dispersão* de seus têrmos. Entre as principais *medidas de dispersão* destacamos: a amplitude semi-quartil, o desvio médio, o desvio padrão e o coeficiente de variação, que serão estudados a seguir.

28. Amplitude semi-quartil. A diferença entre os quartis Q_3 e Q_1 constitui uma medida de dispersão denominada *amplitude quartil*. Comumente se usa a metade dessa amplitude como medida de dispersão, com o nome de *amplitude semi-quartil* ou também de *afastamento provável*. Indicação:

$$q = \frac{Q_3 - Q_1}{2}.$$

A amplitude semi-quartil do *exemplo-modêlo* é igual a:

$$q = \frac{164{,}6875 - 157{,}222}{2} = \frac{7{,}4655}{2} = 3{,}7327 \text{ ou } \boldsymbol{q = 3{,}7327cm}.$$

NOTA: A *amplitude total*, definida no início dêste curso, pode constituir também uma medida de dispersão, embora não muito rigorosa. No *exemplo-modêlo*, a sua amplitude total de: 180cm − 150cm, significa que tôda a variabilidade da distribuição se dá em 30cm.

29. Desvio médio ou afastamento médio. Dá-se o nome de *desvio médio* de uma série à média aritmética dos valores absolutos dos desvios (ou afastamentos) dos têrmos dessa série, em relação à sua média aritmética (M_a). Tomamos o valor absoluto das diferenças para evitar que seja nula a soma dos citados desvios, de acôrdo com a propriedade fundamental da média aritmética de uma distribuição (n.º 14-*b*). Indicação: d_m.

30. Cálculo do desvio médio de dados não tabulados. Para dados não tabulados o cálculo do desvio médio é feito aplicando a fórmula:

$$\boxed{d_m = \frac{\Sigma |d|}{n}}$$

onde $\begin{cases} |d| & \text{representa, em valor absoluto, o desvio de cada valor em relação a } M_a. \\ n & \text{n.º de valores que compõem o rol.} \end{cases}$

Exemplo: Calcular o d_m do seguinte rol de 10 notas (1.º Ano "A"):

35, 40, 45, 50, 50, 55, 60, 65, 70, 80.

Temos:

$n = 10$	*Notas* (x)	*desvio* $\|d\|$	$M_a = 55$
	35	20 (−)	
	40	15 (−)	
	45	10 (−)	
	50	5 (−)	
	50	5 (−)	
	55	0	
	60	5 (+)	
	65	10 (+)	
	70	15 (+)	
	80	25 (+)	
	$\Sigma x = 550$	$\Sigma \|d\| = 110$	

CÁLCULOS:

$$M_a = \frac{\Sigma x}{n} = \frac{550}{10} = 55$$

$$\Sigma |d| = 110$$

portanto

$$d_m = \frac{\Sigma |d|}{n} = \frac{110}{10} = 11$$

ou

$$d_m = 11$$

NOTA: O desvio médio das notas do 1.º Ano "B" é 25, isto é, estas notas são bem *mais dispersas*, em tôrno de M_a, do que as notas do 1.º Ano "A". Deixaremos ao encargo do aluno essa verificação.

31. Cálculo do desvio médio de dados tabulados Se os dados estão tabulados o desvio médio da distribuição é calculado pela seguinte fórmula:

$$\boxed{d_m = \frac{\Sigma|d|.F}{\Sigma F}} \quad \text{onde, agora, } |d| = |P_m - M_a|$$

Como exercício de aplicação, calculemos o d_m da distribuição constante no *exemplo-modêlo*:

$h=5$	X	P_m	F	$\lvert d\rvert$	$\lvert d\rvert.F$	$M_a = 161{,}625$
	150 ⊢ 155	152,5	6	9,125	54,750	
	155 ⊢ 160	157,5	9	4,125	37,125	
	160 ⊢ 165	162,5	16	0,875	14,000	
	165 ⊢ 170	167,5	5	5,875	29,375	
	170 ⊢ 175	172,5	3	10,875	32,625	
	175 ⊢ 180	177,5	1	15,875	15,875	
			$\Sigma F = 40$		$\Sigma\lvert d\rvert.F = 183{,}750$	

$|d| = |P_m - M_a|$ (em valor absoluto)

$\Sigma\,|d|.F = 183{,}750$

$$d_m = \frac{\Sigma\,|d|.F}{\Sigma F} = \frac{183{,}750}{40} = 4{,}59375$$

portanto,

$$d_m = 4{,}59375\,\text{cm}$$

32. Desvio-padrão ou afastamento-padrão. Chama-se *desvio-padrão* ou *desvio-quadrático-médio* de uma série de valores ao valor positivo da raiz quadrada da média aritmética dos quadrados dos desvios dêsses valores.

Os desvios ou afastamentos, que podem ser tomados em relação a qualquer média de posição, são tomados, geralmente, em relação à média aritmética. Indicação: d_p ou σ (sigma minúsculo). O fato de se elevarem os desvios ao quadrado (*) visa estudar a dispersão por um processo que tratará êsses mesmos desvios como se tivessem o mesmo sinal.

33. Cálculo do desvio-padrão de dados não tabulados. Êsse cálculo é feito aplicando a seguinte fórmula:

$$\boxed{\sigma = \sqrt{\frac{\Sigma d^2}{n}}}$$

onde d e n têm os mesmos significados da fórmula do desvio médio.

Exemplo: Calcular o desvio-padrão do rol de 10 notas:

35, 40, 45, 50, 50, 55, 60, 65, 70, 80.

$n=10$	*Notas* (x)	$\lvert d\rvert$	d^2	$M_a = 55$
	35	20	400	
	40	15	225	
	45	10	100	
	50	5	25	
	50	5	25	
	55	0	0	
	60	5	25	
	65	10	100	
	70	15	225	
	80	25	625	
	$\Sigma x = 550$		$\Sigma d^2 = 1750$	

CÁLCULOS:

$$M_a = \frac{\Sigma x}{n} = \frac{550}{10} = 55$$

$$\Sigma d^2 = 1750$$

(*) YULE and KENDALL: *An introduction to the theory of Statistics.*

$$\sigma = \sqrt{\frac{\Sigma d^2}{n}} = \sqrt{\frac{1750}{10}} = \sqrt{175} = 13{,}22$$

portanto

σ = **13,22cm** ou (com aproximação de 0,01).

34. Cálculo do desvio-padrão de dados tabulados. No caso de dados tabulados o cálculo do desvio-padrão, a exemplo do que já foi estudado no cálculo da média aritmética, pode ser feito por dois processos: o *longo* e o *breve.*

Para o *processo longo* usamos a fórmula:

$$\sigma = \sqrt{\frac{\Sigma d^2 F}{\Sigma F}}$$

e para o *processo breve,* a fórmula:

$$\sigma = h.\sqrt{\frac{\Sigma \alpha^2 F}{\Sigma F} - \left(\frac{\Sigma \alpha F}{\Sigma F}\right)^2},$$

onde os desvios são calculados em relação a uma média arbitrária (de preferência o P_m da classe de maior freqüência), e, é usado o quociente $\alpha = \frac{d}{h}$, por nós já conhecido no cálculo da M_a, quando falámos dea fastamento $\left(\alpha = \frac{a}{h}\right)$. Da mesma forma, se escreve numa coluna relativa aos α, a série dos números inteiros negativos e positivos, respectivamente, acima e abaixo do zero (0) correspondente à classe, à qual pertence a média escolhida. É lógico que, pelo processo breve, serão evitados enormes cálculos que, sòmente seriam facilitados com o uso de *máquinas.*

Como aplicação calculemos, pelos dois processos, o *desvio-padrão* do *exemplo-modêlo*:

PROCESSO LONGO:

$h = 5$	X	P_m	F	$d = P_m - M_a$	d^2	d^2F	$M_a = 161{,}625$
	150 ⊢ 155	152,5	6	− 9,125	83,265 625	499,593 750	
	155 ⊢ 160	157,5	9	− 4,125	17,015 625	153,140 625	
	160 ⊢ 165	162,5	16	+ 0,875	0,765 625	12,250 000	
	165 ⊢ 170	167,5	5	+ 5,875	34,515 625	172,578 125	
	170 ⊢ 175	172,5	3	+ 10,875	118,265 625	354,796 875	
	175 ⊢ 180	177,5	1	+ 15,875	252,015 625	252,015 625	
			$\Sigma F = 40$			$\Sigma d^2F = 1\ 444{,}375\ 000$	

Cálculos: $\Sigma d^2F = 1\ 444{,}375\ 000$

$\Sigma F = 40$

$$\sigma = \sqrt{\frac{\Sigma d^2F}{\Sigma F}} = \sqrt{36{,}109375}$$

$\sigma = 6{,}009$

Portanto: σ = **6,009 cm** (c/ aproximação de 0,001).

Processo breve:

$h=5$	X	P_m	F	α	$\alpha.F$	$\alpha^2.F$
	150 ⊢ 155	152,5	6	−2	−12	24
	155 ⊢ 160	157,5	9	−1	−9	9
	160 ⊢ 165	162,5	16	0	0	0
	165 ⊢ 170	167,5	5	1	5	5
	170 ⊢ 175	172,5	3	2	6	12
	175 ⊢ 180	177,5	1	3	3	9
			$\Sigma F=40$		$\Sigma\alpha.F=-7$	$\Sigma\alpha^2F=59$

$M_a=161,625$

Cálculos

$\Sigma\alpha F=-7$

$\Sigma\alpha^2F=59$

$$\sigma=h.\sqrt{\frac{\Sigma\alpha^2F}{\Sigma F}-\left(\frac{\Sigma\alpha F}{\Sigma F}\right)^2}$$

$$\sigma=5.\sqrt{\frac{59}{40}-\left(\frac{-7}{40}\right)^2}=5.\ \sqrt{1,4443}=5\times1,20$$

Portanto:

$\sigma=$ **6,009 cm** (c/ aproximação de 0,001).

Nota: Êste quadro permite calcular ràpidamente a M_a, pois:

$$M_a=P_m+h.\frac{\Sigma\alpha F}{\Sigma F}$$

ou $$M_a=162,5+5.\left(\frac{-7}{40}\right)=162,5-0,875=161,625\ \therefore\ M_a=\mathbf{161,625\,cm}$$ (já determinada).

O *desvio-padrão* é a medida de dispersão mais usada no estudo da variabilidade das distribuições simétricas ou moderadamente assimétricas. Basta observar (fig. 47) que, tomado com o duplo sinal ($\pm\sigma$) em tôrno da média aritmética, abrange cêrca de 68,26 % dos valores da distribuição, enquanto que, a região central, como já vimos, abrange sòmente 50 %. O triplo do desvio-padrão, com o duplo sinal($\pm 3\sigma$) compreende pràticamente todo o conjunto de valores (99,72 %).

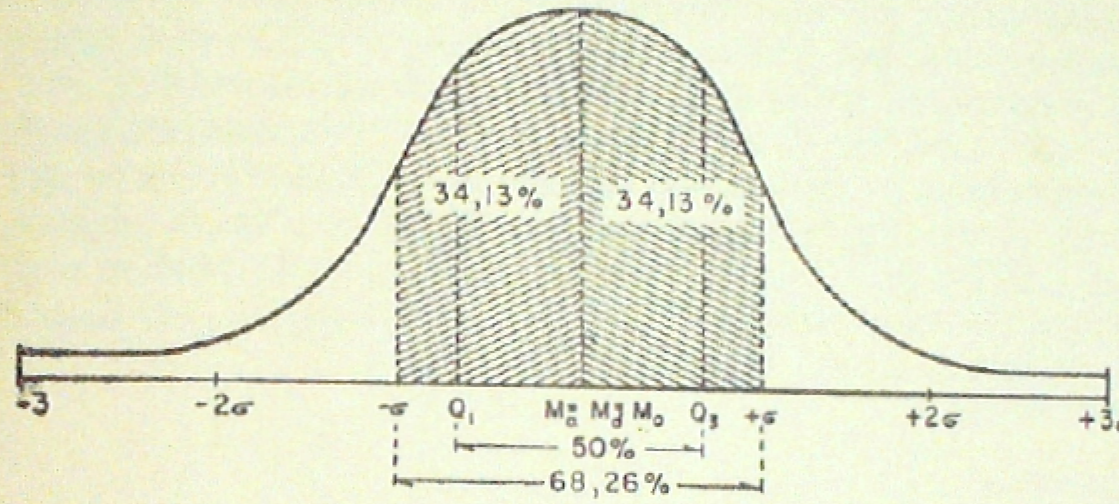


FIG. 47

35. Coeficiente de variação. Denominamos *coeficiente de variação* ao número que permite traduzir a maior ou menor variabilidade da distribuição, *independente da unidade* em que vêm expressos os seus valores. Indicação: $C.V.$ O coeficiente de variação é dado pela seguinte fórmula estabelecida por Pearson:

$$\boxed{C.V=\frac{100.\sigma}{M_a}}$$

Para o *exemplo-modêlo*, temos o seguinte coeficiente de variação:

$$C.V=\frac{100\times6,009}{161,625}=3,717\ \text{(c/ aprox. de 0,001)}$$

Para mostrar a importância do $C.V.$ no confronto de duas distribuições, vamos supor que, a distribuição de *pesos* das 40 alunas estudadas em nosso exemplo-modêlo apresente-se com um $C.V.$ igual a 11,22. Diríamos, então, que o pêso do mesmo grupo de alunas apresentou-se cêrca de *três vêzes mais variável* que a estatura.

MEDIDA DE ASSIMETRIA

36. Generalidades. Para as distribuições rigorosamente simétricas, isto é, para aquelas que teriam uma exata repartição de valores em tôrno do valor central, a M_a, a M_d e a M_o, seriam iguais (fig. 47). Como isso raramente pode acontecer, introduziu-se um novo número que permite medir a assimetria apresentada por uma distribuição de freqüência. Êsse novo número, que também *não vem expresso por nenhuma unidade*, é denominado *índice de assimetria* e indicado por $I \,.\, A$.

A assimetria de uma distribuição diz-se *positiva* quando predominam os valores mais altos e a tendência da curva que a representa (fig. 48) é para a direita. Nesse caso vale a relação: $M_o < M_d < M_a$. A assimetria é *negativa*, quando predominam os valores mais baixos e a tendência da curva (fig. 49) é para a esquerda. Agora, temos: $M_a < M_d < M_o$.

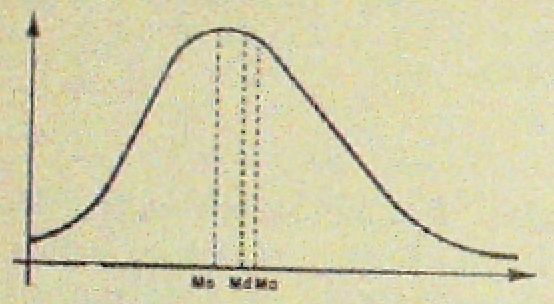


FIG. 48

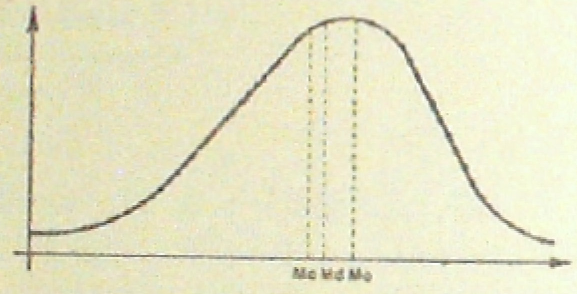


FIG. 49

O índice de assimetria, dado por Pearson, tem por expressão:

$$\boxed{I \,.\, A = \frac{3(M_a - M_d)}{\sigma}}$$

que representa um número, que varia entre -1 e $+1$.

É claro que se o $I \,.\, A \,. = 0$, estaremos diante de uma perfeita simetria.

Para o nosso *exemplo-modêlo*, temos o seguinte *índice de assimetria*:

$$I \,.\, A \,. = \frac{3(161{,}625 - 161{,}5625)}{6{,}009} = \mathbf{0{,}031} \text{ (c/ aprox. de 0,001).}$$

que revela ser a distribuição de *fraca assimetria positiva.*

RESUMO DOS RESULTADOS ENCONTRADOS NO LEVANTAMENTO ESTATÍSTICO RELATIVOS AO EXEMPLO-MODÊLO EM ESTUDO

Temos, agora, depois de estudadas tôdas as medidas mais importantes de *posição*, de *dispersão* e de *assimetria*, os seguintes resultados relativos ao atributo *estatura* das alunas do 1.° Ano Normal, de um Instituto de Educação, tendo usado, como amostras representativas, as estaturas de 40 alunas, constantes de uma tabela primitiva dada (pág. 208).

Medidas de *posição*	Média aritmética (M_a)	161,625cm
	Mediana (M_d)	161,5625cm
	Moda (M_o)	161,4375cm
	1.° Quartil (Q_1)	157,222cm
	3.° Quartil (Q_3)	164,6875cm
Medidas de *dispersão*	Amplitude total (A)	30,000cm
	Amplitude semi-quartil (q)	3,7327cm
	Desvio médio (d_m)	4,59375cm
	Desvio padrão (σ)	6,009cm
	Coeficiente de variação ($C.\,V.$)	3,717
Medida de *assimetria* —	Índice de assimetria ($I.\,A.$)	0,031

Região Normal (Amplitude: 7,4655cm) — 50 % das alunas observadas medem de 157,222cm (Q_1) a 164,6875cm (Q_3)

ALGUMAS CONSIDERAÇÕES SÔBRE A CURVA NORMAL DE GAUSS

Com o estudo que acabamos de fazer sôbre as medidas de posição, de dispersão e de assimetria, correspondentes a um conjunto de dados observados, pretendemos ter analisado as tendências características de tal conjunto. Todavia, as representações gráficas das distribuições por freqüências, aqui elaboradas com um *número limitado* de amostras, apesar da enorme contribuição que tiveram na apreciação de um fenômeno, revelaram por sua vez, que os polígonos e os histogramas apresentavam *irregularidades*, oriundas, geralmente, de erros de amostragem.

Após inúmeras verificações práticas descobriu-se que a maioria dos fenômenos que ocorrem *por acaso*, especialmente os de Educação (por ex., os quocientes de inteligência — *Q.I.*), e os de Biologia (por ex., os relativos a estaturas, pesos), apresentavam-se com distribuições aproximadamente *normais*, isto é, 50 % dos valores observados pertenciam ao intervalo limitado pelos quartis Q_1 e Q_3. É o que aconteceria se quiséssemos por exemplo, estudar o atributo estatura de tôda população escolar de São Paulo. Encontraríamos nesse levantamento uma distribuição de estaturas contendo uma série de medidas desde um mínimo valor coligido até um máximo e uma maior freqüência para as estaturas compreendidas entre o Q_1 e o Q_3 da distribuição. O gráfico correspondente seria *aproximadamente simétrico* em relação ao eixo que passasse pela abscissa relativa a moda.

Sendo o último escopo da Estatística *procurar leis que rejam os fatos verificados pela observação*, surgiu a idéia de se estabelecerem *curvas teóricas*, de equações determináveis, que pudessem atingir tal objetivo, já que as curvas obtidas pelos gráficos conhecidos apresentavam-se sem aquela continuidade que era de se desejar para estudos conclusivos. Uma das curvas tomada como *modêlo de normalidade*, equivalente a uma distribuição *rigorosamente simétrica*, é a conhecida com o nome de *Curva de Gauss* ou Curva Normal de Probabilidade, ou,

ainda, simplesmente como Curva Normal, cuja forma de "sino" (fig. 47) foi por nós ressaltada no estudo das curvas de freqüência (pág. 222).

A Curva de Gauss permite, por confronto ou ajustes, eliminar os erros de amostragem, tanto quanto possível, bem como determinar a probabilidade de acontecimentos dentro do fenômeno que se estuda.

É evidente que a determinação da equação da Curva de Gauss (*) e o emprêgo das operações de ajuste de dados observados ao seu traçado, exigem conhecimentos de cálculo das probabilidades e de matemática superior, que não se enquadram no caráter elementar dêste livro. Os estudiosos já devem ter travado conhecimentos de *tabelas* destinadas a obtenção do valor da probabilidade relativa ao bom ajustamento de uma dada distribuição à distribuição normal teórica correspondente à Curva de Gauss. Deixaremos êsse estudo para uma outra oportunidade.

POPULAÇÃO DO BRASIL

Segundo os últimos dados do Serviço Nacional de Recenseamento do IBGE, sujeitos, ainda, a revisão, são os seguintes os números de Municípios e a população recenseada em 1.º de setembro de 1960, por unidades da Federação (v. pág. seguinte).

(*) Equação reduzida da *Curva de Gauss*.

$$y = y_0 \cdot e^{-\frac{x^2}{2\sigma^2}} \quad \text{onde} \quad \begin{cases} y_0 \text{ – ordenada máxima correspondente a média;} \\ e = 2{,}718\,28 \ldots\ldots \text{ (constante)} \\ x \text{ – abscissa a partir da média} \\ \sigma \text{ – desvio padrão da distribuição} \end{cases}$$

POPULAÇÃO DO BRASIL

UNIDADES DA FEDERAÇÃO	POPULAÇÃO*		
	Número de Municípios	*Capital*	*Total*
BRASIL	2 768	—	70 528 625
Rondônia	2	51 049	70 783
Acre	7	47 882	169 208
Amazonas	44	175 343	721 215
Rio Branco	2	26 163	29 489
Pará	60	402 170	1 550 935
Amapá	5	46 905	68 889
Maranhão	91	159 628	2 492 139
Piauí	71	144 799	1 263 368
Ceará	142	514 818	3 337 856
Rio Grande do Norte	83	162 537	1 157 256
Paraíba	88	155 117	2 017 969
Pernambuco	102	788 580	4 120 000
Alagoas	69	170 134	1 271 062
Fernando de Noronha	1	1 389	1 389
Sergipe	62	115 713	760 273
Bahia	194	655 735	5 990 605
Minas Gerais	483	693 328	9 550 000
Serra dos Aimorés**	1	427 695	384 297
Espírito Santo	37	85 242	1 188 665
Rio de Janeiro	61	245 467	3 402 728
Guanabara	1	3 307 163	3 307 163
São Paulo	504	3 776 581	12 930 000
Paraná	162	361 309	4 110 000
Santa Catarina	102	98 520	2 146 909
Rio Grande do Sul	150	641 173	5 448 823
Mato Grosso	64	53 192	950 000
Goiás	179	153 505	1 954 862
Distrito Federal	1	141 742	141 742

(*) Publicado em "A Gazeta", de São Paulo, em 16-10-1961.
(**) Território em litígio entre os Estados de Minas Gerais e Espírito Santo.

30
18
110

30
12

São as professôras primárias as verdadeiras guardiãs da civilização

Bertrand Russell

Para os Institutos de Educação e Escolas Normais