

Aula 02 – Introdução à Estatística Descritiva – Gráficos

Luciana Rocha Pedro

GCC 1518 – Estatística e Probabilidade – CEFET Maracanã

USP – Projeto de Ensino

Aprender Fazendo Estatística

Gráficos

Definição

Gráfico é um recurso visual da estatística utilizado para representar um fenômeno.

Sua utilização em larga escala nos meios de comunicação social, técnica e científica deve-se tanto à sua capacidade de refletir padrões gerais e particulares do conjunto de dados em observação, como à facilidade de interpretação e a eficiência com que resume informações dos mesmos.

Embora os gráficos forneçam menor grau de detalhes que as tabelas, estes apresentam um ganho na compreensão global dos dados, permitindo que notemos imediatamente a sua forma geral, sem deixar de evidenciar alguns aspectos particulares que sejam de interesse do pesquisador.

Gráficos

Uma representação gráfica coloca em evidência as **tendências**, as **ocorrências ocasionais**, os **valores mínimos e máximos** e também as **ordens de grandezas** dos fenômenos que estão sendo observados.

Todo gráfico, em sua versão final, deve primar pela simplicidade, clareza e veracidade nas informações.

Gráfico de Barras

O **gráfico de barras** é um gráfico formado por retângulos horizontais de larguras iguais, em que cada um deles representa a intensidade de uma modalidade ou atributo. O objetivo deste gráfico é comparar grandezas e é recomendável para variáveis cujas categorias tenham designações extensas.

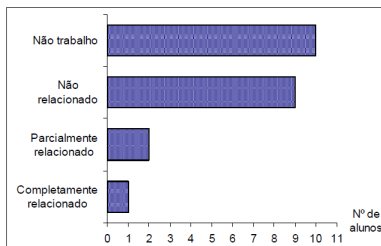


Figura 02 - Relação trabalho e curso dos alunos da disciplina Inferência Estatística do curso de Estatística da UEM, 21/03/2005.

Fonte: Tabela 01.

Gráfico de Colunas

O **gráfico de colunas** é o gráfico mais utilizado para representar variáveis qualitativas. Difere do gráfico de barras por serem seus retângulos dispostos verticalmente ao eixo das abscissas, sendo mais indicado quando as designações das categorias são breves.

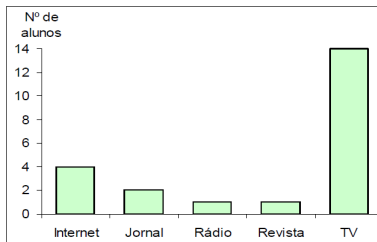


Figura 03 - Meios de informação utilizados pelos alunos da disciplina Inferência Estatística, curso de Estatística da UEM, 21/03/2005.

Fonte: Tabela 01.

Gráficos Comparativos

Ao descrevermos, simultaneamente, duas ou mais categorias para uma variável, é conveniente fazer uso dos gráficos de barras ou colunas justapostas (ou sobrepostas), chamados de **gráficos comparativos**.

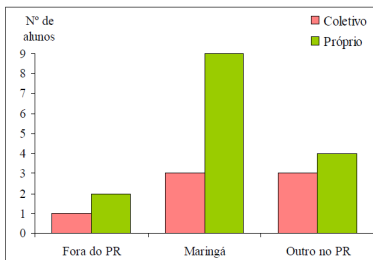


Figura 04 - Município de procedência segundo o tipo de transporte utilizado pelos alunos da disciplina Inferência Estatística do curso de Estatística da UEM, 21/03/2005.

Fonte: Tabela 01.

Gráfico de Setores

O **gráfico de setores** é um tipo de gráfico em que a variável em estudo é projetada num círculo, de raio arbitrário, dividido em setores com áreas proporcionais às **frequências** das suas categorias.

São indicados quando desejamos comparar cada valor da série com o total.

Recomenda-se o seu uso para o caso em que o número de categorias não é grande e não obedece a alguma ordem específica.

Gráfico de Setores

A Figura 05 mostra um gráfico de setores para a variável município de procedência da Tabela 01.

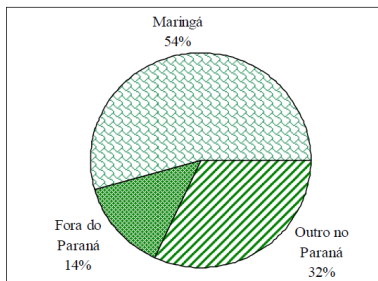


Figura 05 - Município de procedência dos alunos da disciplina Inferência Estatística do curso de Estatística da UEM, 21/03/2005.

Fonte: Tabela 01.

Cálculo do Ângulo em Gráfico de Setores

O procedimento para o cálculo do ângulo correspondente a cada categoria é feito por meio de simples proporções: 360 graus, que corresponde a um círculo completo, está para o total de alunos entrevistados, 22, assim como x graus está para o total de alunos que pertencem à categoria desejada.

Por exemplo, aos 54% de alunos que residem no município de Maringá corresponderá um ângulo x resultante da expressão

$$\frac{360}{22} = \frac{x}{12},$$

cujo valor é aproximadamente 196 graus.

Gráfico de Linhas

A aplicação de um **gráfico de linhas** é mais indicada para representações de séries temporais sendo, por tal razão, conhecidos também como gráficos de séries cronológicas.

Sua construção é feita colocando no eixo vertical (y) a mensuração da variável em estudo e na abscissa (x) as unidades da variável em uma ordem crescente.

Este tipo de gráfico permite representar séries longas, o que auxilia detectar suas flutuações tanto quanto analisar suas tendências. Também podemos representar várias séries em um mesmo gráfico.

Gráfico de Linhas

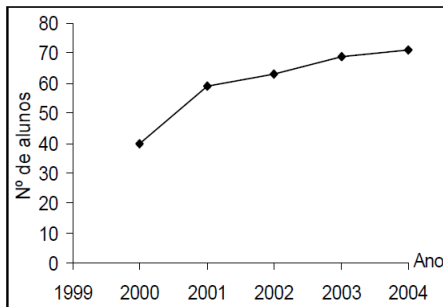


Figura 06 - Número de matrículas anuais na disciplina Probabilidade do curso de Estatística da UEM, 21/03/2005.

Fonte: Tabela 01.

Exercício 03

Considerando as informações sobre os alunos da disciplina de Inferência Estatística do curso de Estatística da UEM, contidas na Tabela 01, construa um gráfico adequado para representar as variáveis:

1. sexo;
2. estado civil;
3. transporte;
4. meios de informação segundo o sexo dos alunos.

Gráfico de Bastões

Um **gráfico de bastões** é formado por segmentos de retas perpendiculares ao eixo horizontal (eixo da variável), cujo comprimento corresponde à frequência absoluta ou relativa de cada elemento da distribuição.

Suas coordenadas não podem ser unidas porque a leitura do gráfico deve tornar claro que não há continuidade entre os valores individuais assumidos pela variável em estudo.

Gráfico de Bastões

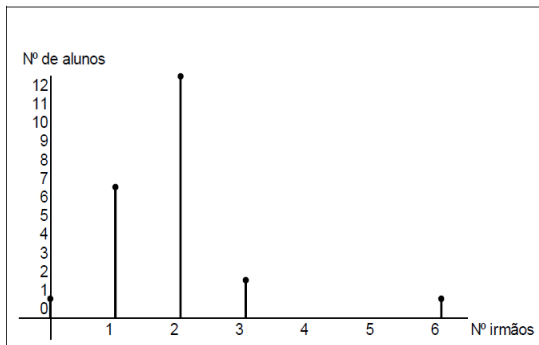


Figura 07 – Número de irmãos dos alunos da disciplina Inferência Estatística do curso de Estatística da UEM, 21/03/2005.

Fonte: Tabela 01.

Gráfico de Freqüência Acumulada

A Figura 08 mostra o **gráfico de frequência acumulada** de uma variável quantitativa discreta.

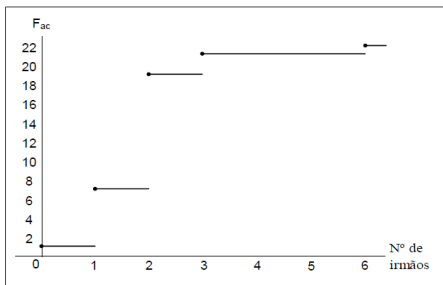


Figura 08 - Número acumulado de irmãos dos alunos da disciplina Inferência Estatística do curso de Estatística da UEM, 21/03/2005.

Fonte: Tabela 01.

Gráfico de Freqüência Acumulada

Na abscissa são alocados os valores assumidos pela variável número de irmãos e, no eixo das ordenadas, as suas freqüências acumuladas.

Observamos que a leitura do gráfico exige alguns cuidados básicos: caso o valor da variável esteja ou não incluído, sua freqüência acumulada difere.

Se for de interesse saber quantos alunos tem dois ou menos irmãos (incluímos dois irmãos), a freqüência acumulada é de 19 alunos. Caso queiramos apenas saber quantos alunos têm menos de dois irmãos (portanto o número dois não está incluso), sua freqüência acumulada é de 7 alunos.

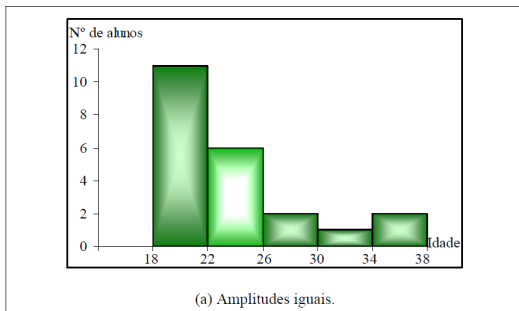
Histograma

O **histograma** é um gráfico de colunas justapostas que representa uma distribuição de freqüência para dados contínuos ou uma variável discreta, quando esta apresentar muitos valores distintos.

No eixo horizontal são dispostos os limites das classes segundo as quais os dados foram agrupados, enquanto que o eixo vertical corresponde às freqüências absolutas ou relativas das mesmas.

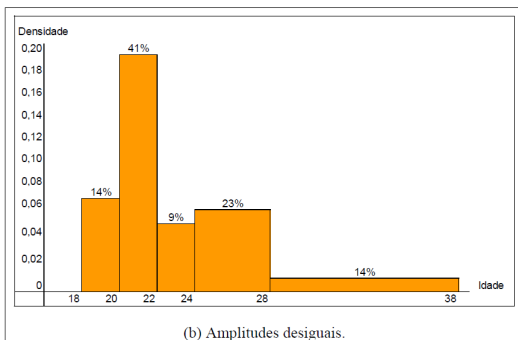
Histograma

Quando os dados são distribuídos em classes de mesma amplitude, Figura 09 (a), todas as colunas apresentam bases iguais com alturas variando em função das suas frequências absolutas ou relativas. Neste caso, a área de cada retângulo depende apenas da sua altura.



Histograma

No caso dos dados serem agrupados em classes de amplitudes diferentes, como mostra a Figura 09 (b), a área de cada coluna já não é mais proporcional à sua altura.



Histograma

Portanto, podemos dizer que, no primeiro caso, o eixo dos valores informa a frequência relativa de cada classe e, no segundo caso, tal procedimento perde todo significado e é necessário comparar as áreas para interpretar as informações que são expostas.

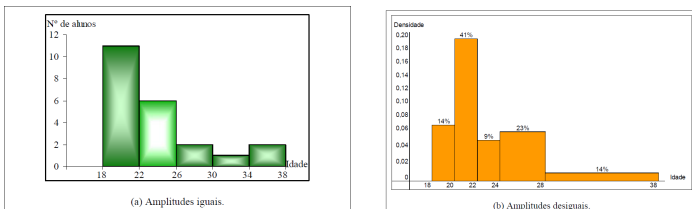


Figura 09 - Idade dos alunos da disciplina Inferência Estatística do curso de Estatística da UEM, 21/03/2005.

Fonte: Tabela 01.

Polígono de Freqüência

O **polígono de freqüência** é um gráfico de linha cuja construção é feita unindo os pontos de coordenadas de abscissas correspondentes aos pontos médios de cada classe e as ordenadas, às freqüências absolutas ou relativas dessas mesmas classes.

O polígono de freqüência é um gráfico que deve ser **fechado no eixo das abscissas**. Então, para finalizar sua elaboração, devemos acrescentar à distribuição uma classe à esquerda e outra à direita, ambas com freqüências zero.

Tal procedimento permite que a área sob a linha de freqüências seja igual à área do histograma.

Polígono de Freqüência

Uma das vantagens da aplicação de polígonos de freqüências é que, por serem gráficos de linhas, permitem a comparação entre dois ou mais conjuntos de dados por meio da superposição dos mesmos.

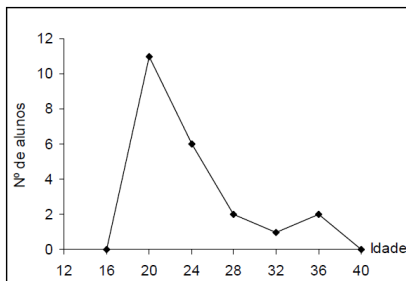


Figura 10 - Idade dos alunos da disciplina Inferência Estatística do curso de Estatística da UEM, 21/03/2005.

Fonte: Tabela 01.

Gráfico de Frequência Acumulada ou Ogiva

O **gráfico da frequência acumulada** ou **ogiva** é um gráfico que permite descrever dados quantitativos por meio da frequência acumulada.

A ogiva é um gráfico de linha que une os pontos cujas abscissas são os limites superiores das classes e as ordenadas são suas respectivas frequências acumuladas.

Gráfico de Freqüência Acumulada ou Ogiva

Convém observarmos que o ponto inicial desse gráfico é o limite inferior do primeiro intervalo, com freqüência acumulada zero, pois não existe qualquer valor inferior a ele.

Quando os dados contidos em cada classe são **distribuídos uniformemente**, podemos estimar, a partir da ogiva, o número de elementos pertencentes a qualquer uma das classes que compõe a distribuição de freqüência dos dados e a quantidade ou porcentagem de elementos que estão abaixo de certo valor pertencente ao conjunto de dados.

Exemplo

Pela Figura 11, notamos que não existem alunos com idade inferior a 18 anos enquanto que abaixo de 34 anos existem vinte alunos.

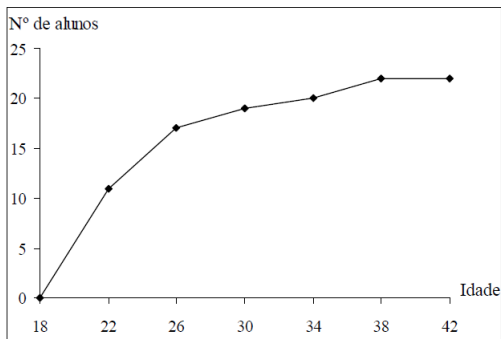


Figura 11 - Idade acumulada dos alunos da disciplina Inferência Estatística do curso de Estatística da UEM, 21/03/2005.

Exercício 04

Considerando as informações sobre os alunos da disciplina de Inferência Estatística do curso de Estatística da UEM, contidas na Tabela 01, construa os gráficos adequados para as variáveis:

1. peso;
2. altura;
3. número de reprovações no primeiro ano de curso.

Diagrama Ramo-e-Folhas

O **diagrama ramo-e-folhas** é um procedimento utilizado para armazenar os dados sem perda de informação.

É utilizado para termos uma idéia visual da distribuição dos dados.

Cada valor observado, x_i , da variável X , deve consistir de, no mínimo, dois dígitos e a variável pode ser tanto quantitativa discreta como contínua.

Diagrama Ramo-e-Folhas

Para construí-lo, dividimos cada número em duas partes. A primeira é denominada **ramo** e a segunda, **folhas**.

O ramo consistirá de um ou mais dígitos iniciais, se o valor da variável for um número inteiro, e do número inteiro, se o valor da variável for um número com decimais.

Nas folhas, colocamos os dígitos restantes, se o valor observado for número inteiro, ou os decimais, caso contrário.

Diagrama Ramo-e-Folhas

A Figura 12 (a) apresenta o ramo-e-folhas correspondente à variável idade do aluno da Tabela 01.

Observamos que o ramo correspondente ao dígito 2 tem muitas folhas. Neste caso, a opção é dividir este ramo em dois: as folhas de 0 a 4 pertencerão a uma linha e as folhas de 5 a 9 pertencerão à outra linha.

Diagrama Ramo-e-Folhas

Os ramos são discriminados por um sinal no seu expoente, como na Figura 12 (b).

Ramo	Folha	Frequência
1	8 8 9	3
2	0 0 0 0 0 0 0 1 1 2 3 4 5 5 6 9	16
3	5 7	2
(a) Sem divisão de ramos.		
Ramo	Folha	Frequência
1	8 8 9	3
2	0 0 0 0 0 0 0 1 1 2 3 4	12
2*	5 5 6 9	4
3	5 7	2
(b) Com divisão de ramos.		

Figura 12 - Idade dos alunos da disciplina Inferência Estatística do curso de Estatística da Universidade Estadual de Maringá, 21/03/2005.

Fonte: Tabela 01.

Exercício 05

01. Construir o diagrama ramo-e-folhas para a variável altura, da Tabela