

CURSO DE SPSS

AULA 1

UFBA/FACED

José Albertino Lordello

Sheila Regina Pereira

Introdução

O *Statistical Package for Social Science for Windows* (SPSS) - O SPSS é um software estatístico que permite a utilização de dados em diversos formatos para gerar relatórios, calcular estatísticas descritivas, conduzir análises estatísticas complexas e elaborar gráficos.

Principais tópicos Abordados

Introdução à Estatística

1. Definição de variável
2. Tipos de variáveis
3. Escala de medidas

Estatística

É uma ciência baseada na Teoria das Probabilidades, cujo objetivo principal é nos auxiliar a tomar decisões ou obter conclusões em situação de incerteza, a partir de informações numéricas.

Inclui um conjunto de técnicas que permitem, de forma Sistemática, organizar, descrever, analisar e interpretar conjuntos de dados provenientes de estudos ou experiências, realizadas em qualquer área do conhecimento.

Por que estudar Estatística?

- ✓ Para saber como apresentar e descrever informações de forma apropriada;
- ✓ Para saber como tirar conclusões a partir de grandes populações, com base somente na informação obtida a partir de amostras;
- ✓ Para saber como melhorar os processos;
- ✓ Para saber como obter previsões confiáveis;

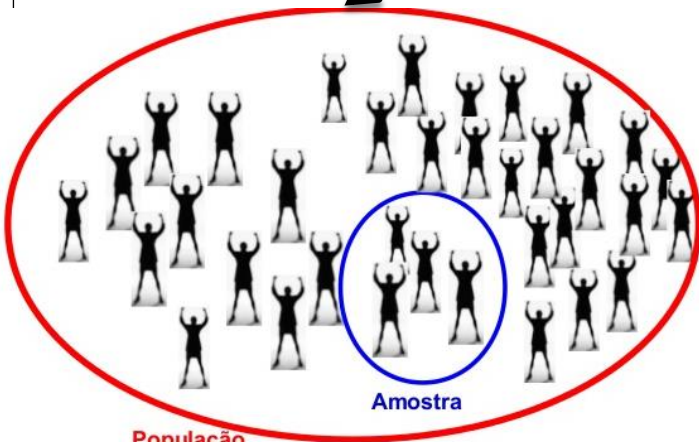


Grandes áreas da Estatística

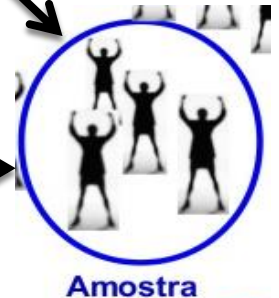
- Estatística Descritiva = se preocupa com a coleta, organização, apresentação, interpretação e análise dos dados.
- Probabilidade = é uma teoria matemática utilizada para estudar a incerteza de fenômenos de caráter aleatório.
- Estatística Inferencial = consiste em inferir (deduzir ou tirar conclusões) a respeito das propriedades de um universo a partir de uma amostra.

Estatística Descritiva:

- 1 – Consistência dos Dados.
- 2 – Interpretações Iniciais.



AMOSTRAGEM



Inferência Estatística:

- 1 – Estimação de Quantidades Desconhecidas.
- 2 – Extrapolação dos Resultados.
- 3 – Teste de Hipóteses.

Definições Importantes

Em estatística utilizaremos extensivamente os termos população e amostra. Assim, definiremos esses termos no contexto da estatística:

- ✓ **População:** conjunto de todos os indivíduos, objetos ou informações que apresentam pelo menos uma característica (observável) comum cujo comportamento interessa-nos analisar.

Em geral, como os universos são grandes, investigar todos os elementos populacionais para determinarmos alguma característica necessita muito de tempo, ou o custo é elevado, ou o processo de investigação leva a destruição do elemento observado, ou, como no caso de populações infinitas, é impossível observar a totalidade da população. Assim, estudar parte da população constitui-se um aspecto fundamental da Estatística.

Amostra: qualquer subconjunto finito de elementos extraídos da população, sobre o qual se faz as observações.

Censo é uma coleção de dados relativos a todos os elementos de uma população.

Apresentação dos Dados

Os dados estatísticos se obtêm mediante um processo que envolve a observação ou outra mensuração de características de uma população ou amostra tais como:

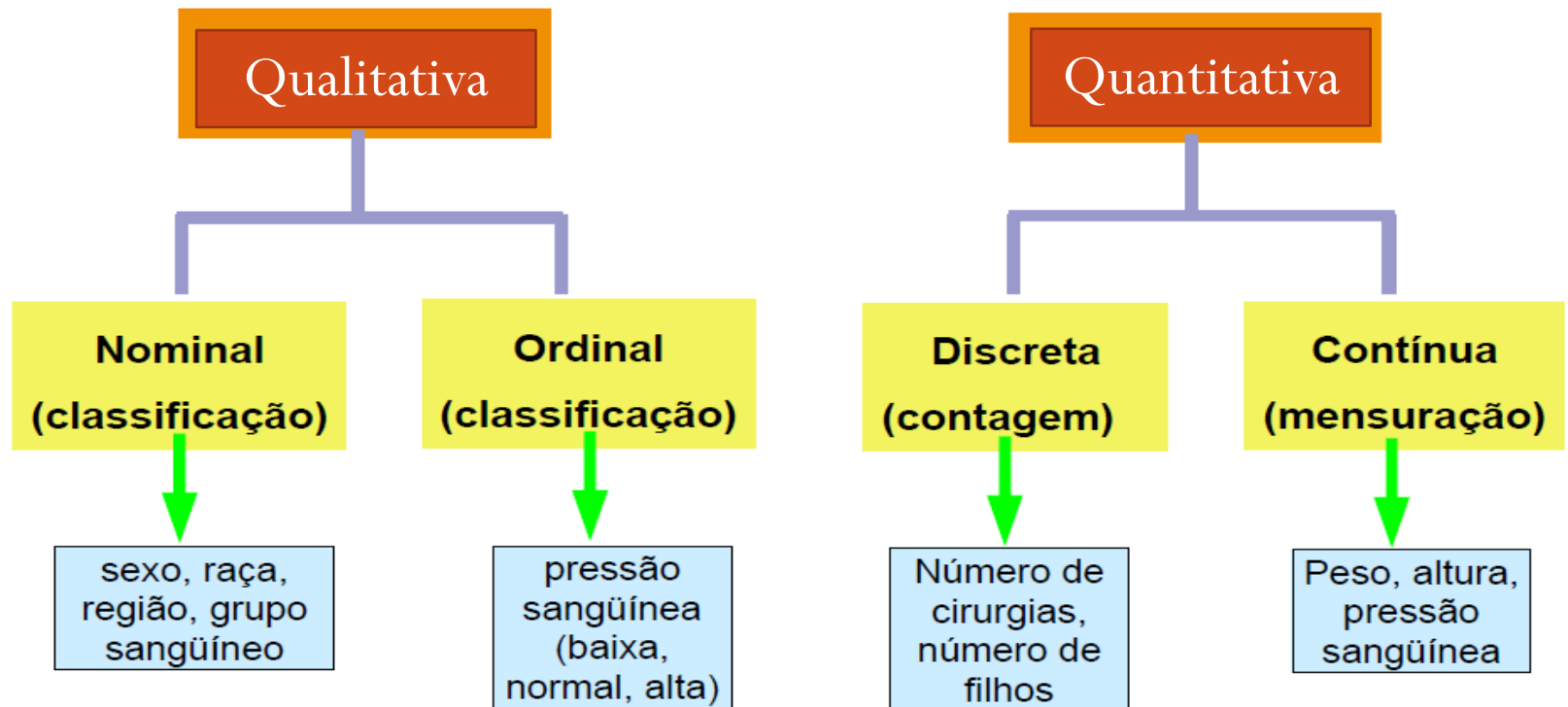
- ✓ Renda anual numa comunidade,
- ✓ Sexo dos indivíduos de uma tribo indígena,
- ✓ Percentagem de açúcar em cereais.
- ✓ Número de peças defeituosas
- ✓ Tempo de vida de um dispositivo eletrônico

Cada uma dessas características é chamada de *variável*, porque originam valores que tendem a exibir certo grau de variabilidade quando se fazem mensurações sucessivas.

Variável

- É algo que varia, ou seja, uma propriedade que pode assumir valores diferentes (Kerlinger, 1979)
- Característica associada a cada indivíduo pertencente à amostra ou a população.

Esquemáticamente



Variáveis Qualitativas

Se os resultados das observações serão expressos através de categorias ou atributos, que se distinguem por alguma característica não-numérica.

Nominal

- ✓ Caracteriza-se por dados que consistem apenas em nomes, rótulos ou categorias.
- ✓ Quando **não** existe qualquer ordenação dos resultados obtidos.

Exemplo: Sexo, cor da pele, (sim ou não), estado civil, ...

Ordinal

- Quando existe uma certa possível ordenação nos resultados das observações.
- Envolve dados que podem ser dispostos em alguma ordem, mas as diferenças entre os valores dos dados não podem ser determinadas ou não tem sentido.

Exemplo: Nível de escolaridade, grau de satisfação, classe social, porte de uma empresa , grau de uma lesão

Variáveis Quantitativas

Se os resultados das observações forem expressos sempre através de números, que representem contagens ou medidas.

Discreta

Quando os resultados possíveis da observação formam um conjunto finito ou infinito enumerável de números e que resultam, frequentemente, de uma contagem.

Exemplo: Número de filhos, número de alunos, idade (em anos completos), etc.

Contínua

Pode assumir qualquer valor pertencente a um determinado intervalo do conjunto dos números reais e que resultam, normalmente, de uma mensuração (medição).

Exemplo: Salário mensal, altura, peso, comprimento,

Exemplo

Classificar as variáveis do dicionário de variáveis do Enade 2011 (qualitativa nominal ou ordinal; quantitativa discreta ou contínua).

co_regiao_curso = Código da região de funcionamento do curso

nu_idade = **Idade do inscrito em 21/11/2010**

ano_in_gra = Ano de início da graduação

tp_semest = Semestre de graduação

nt_ce = Nota bruta no componente específico - Média ponderada da parte objetiva (85%) e discursiva (15%) no componente específico (0 a 100).

co_rs_i1 = Qual o grau de dificuldade desta prova na parte de Formação Geral?

co_rs_s2 = Como você se considera?

co_rs_s4 = Quantas pessoas, da sua família, moram com você na mesma casa?

co_rs_s5 = Somando a sua renda com a renda dos familiares que moram com você, quanto é, aproximadamente, a renda familiar?

co_rs_s13 = Até que nível seu pai estudou?

NÍVEIS DE MEDIDAS

- Medir é atribuir números a quantidades do atributo dos objetos segundo determinadas regras.
- Regra é um guia, um método ou um comando que diz ao investigador como medir o atributo.

As escalas mais comuns são:

- ✓ Nominal
- ✓ Ordinal
- ✓ Intervalar
- ✓ Razão

Escala Nominal

- A escala nominal de medidas é a mais primitiva das escalas. Sua natureza é apenas classificatória.
- São fixadas categorias bem definidas e delineadas, cujos elementos têm como prioridade fundamental a equivalência ou igualdade.
- As categorias compondo uma variável nominal são mutuamente excludentes, não tendo qualquer outra relação entre si.

Exemplo:

Religião: (1) Católico, (2) protestante, (3) espírita, (4) candomblé.

região do país: (1) Norte, (2) Sul, (3) Leste, (4) Oeste.

Tipo de escola: (1) Particular, (2) Pública.

OBS1 : O tipo de escola Particular é atribuído ao número 1 e à Pública ao número 2, isso apenas significa que escola Particular e escola Pública são diferentes em relação ao atributo medido.

OBS 2: Os números, numa escala nominal, não permitem inferir sobre as diferenças na quantidade do atributo.

Escala Ordinal

- A escala ordinal ou escala de postos reflete a posição ou a importância relativa da medida de um atributo.
- A escala ordinal é uma escala de ordenação, designando uma posição relativa das classes segundo uma direção. Qualquer conjunto de valores que preservem a ordem são válidos para essa variável tornando, dessa forma, a escala ordinal invariante sob transformações que preservem a ordem. Ou seja, uma escala ordinal pode ser transformada em outra escala ordinal. Isto implica que, se eventualmente forem empregados números para representar as classes, apenas a propriedade de ordem deve ser respeitada. Assim, as operações aritméticas (somadas, diferenças, etc.) entre esses valores não têm sentido.

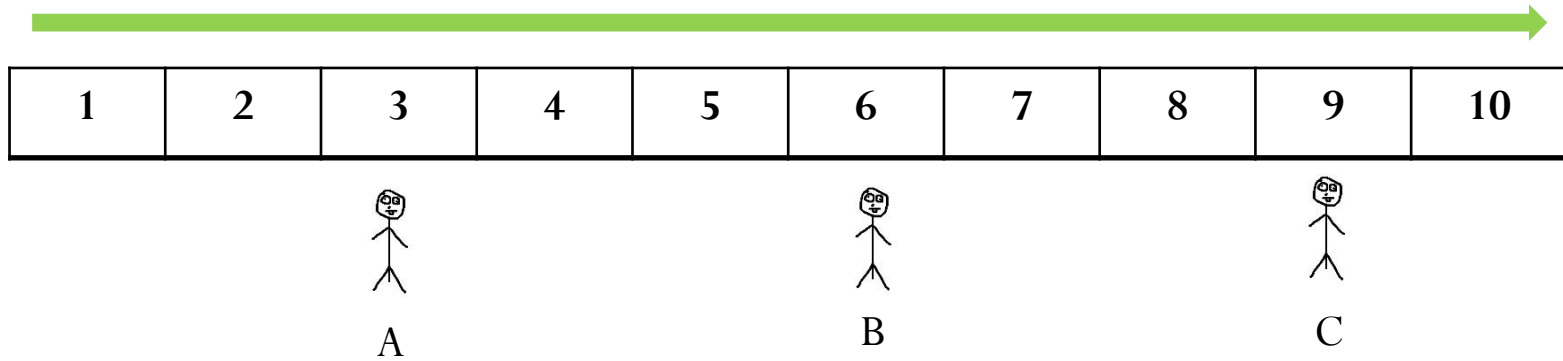
Exemplo: **Classe social:** (1) Alta (2) Média (3) Baixa

A classe (1) é maior do que a classe (3) e, portanto, que a classe (3) é menor que a classe (1).

A propriedade de transitividade é preservada na escala ordinal: se a classe (1) é maior ou mais elevada que a classe (3), qualquer unidade particular da classe (1) é maior ou mais elevada que qualquer unidade específica da classe (3).

Nada é possível dizer sobre o número de vezes que um atributo é maior ou menor do que o outro.

Exemplo: Dez estudantes foram classificados segundo sua habilidade numérica e receberam posições que variam de 1 (maior habilidade) a 10 (menor habilidade).



- Quando se tem uma escala ordinal, os números não permitem inferências sobre a quantidade da diferença entre um atributo e outro.
- Uma das deficiências das escalas ordinais refere-se ao fato de não proporcionarem informações sobre a dispersão da variável.

- Sempre que temos variáveis quantitativas os dados estão pelo menos em escala intervalar. Muitas metodologias estatísticas mais sofisticadas são legítimas com dados em escala intervalar. Mas claro que continua a ser legítimo usar as metodologias desenvolvidas para dados nominais.
- Os escores de um teste educacional, são considerados como constituindo uma escala intervalar. Sem esse posicionamento, seria impossível, dada a natureza de sua escala (ordinal), estabelecer medidas de dispersão, como, por exemplo, a variabilidade e o desvio padrão, que são indispensáveis para a definição de normas e a verificação do funcionamento efetivo do teste como instrumento de medida educacional.

Escala Razão

- A escala de razão é a mais elaborada das escalas de medida, no sentido de permitir todas as operações aritméticas. Essa escala possui um **ponto zero único**, além de unidade de medida constante.
- O zero absoluto significa ausência total do atributo mensurado
- É a escala de medida mais comum nas ciências físicas, tais como as escalas para a medida de comprimento, peso, altura, etc.

Exemplo: se uma unidade tem 3 m e a outra 1 m, pode-se dizer que a primeira unidade tem altura 3 vezes superior a da segunda.

- Se as alturas das duas unidades forem transformadas em centímetros, suas medidas serão, respectivamente, 30 cm e 10 cm, que estão na mesma razão 3:1.
- **Se o máximo que pode ser dito que um objeto é tantas vezes maior, mais pesado, etc., então tem-se uma escala de razão.**

Variáveis Dependentes e Variáveis independentes

Variável Dependente- uma variável que é influenciada por outra variável

Variável Independente - é uma variável que se supõe influenciar outra variável .

Exemplo:

1) O reforço escolar aumenta a aprendizagem;

$$Y = a + bX$$

Y é a variável dependente (aprendizagem)

X é a variável independente (explicativa – reforço escolar)

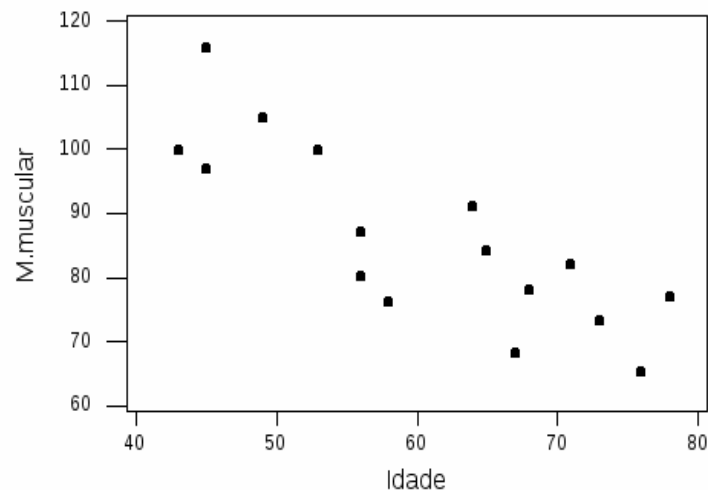
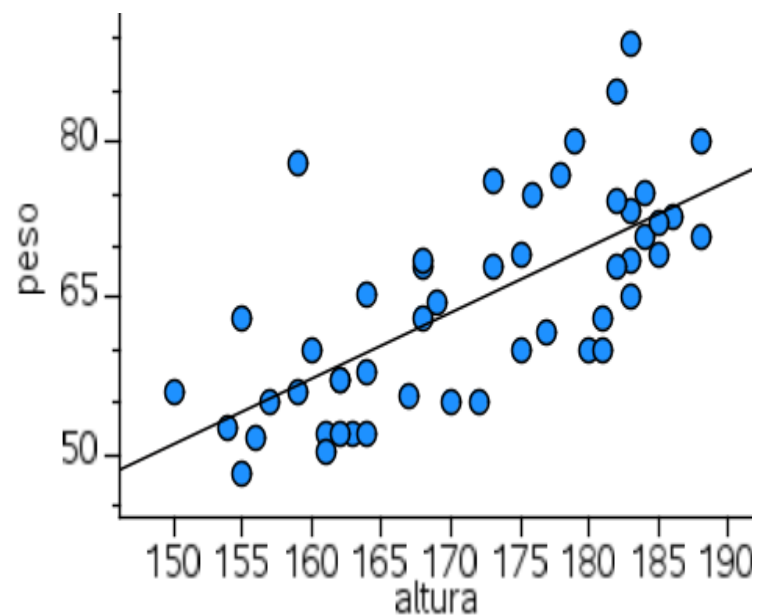
a e b são constantes cujos valores são determinados pela pesquisa.

2) Quanto maior o tempo dedicado aos estudos maior é a aprendizagem ou a nota ou rendimento escolar;

3) Números de filhos x idade do indivíduo;

4) Numero de filhos e gasto com educação;

5) Nota em uma prova de inglês x peso do indivíduo.



Próxima aula

Medidas de tendência central.

Medidas de Variabilidade.

Medidas de Associação.