**DEFINIÇÕES BÁSICAS DA ESTATÍSTICA**

**.**

**FENÔMENO ESTATÍSTICO**: é qualquer evento que se pretenda analisar, cujo estudo seja possível a aplicação do método estatístico. São divididos em três grupos:

***Fenômenos de massa ou coletivo:*** são aqueles que não podem ser definidos por uma simples observação. A estatística dedica-se ao estudo desses fenômenos. *Ex: A natalidade na Grande Vitória, O preço médio da cerveja no Espírito Santo, etc.*

***Fenômenos individuais:*** são aqueles que irão compor os fenômenos de massa. *Ex: cada nascimento na Grande Vitória, cada preço de cerveja no Espírito Santo, etc.*

***Fenômenos de multidão:*** quando as características observadas para a massa não se verificam para o particular.

**DADO ESTATÍSTICO:** é um **dado numérico** e é considerado a **matéria-prima sobre a qual iremos aplicar os métodos estatísticos.**

**POPULAÇÃO**: é o **conjunto total de elementos portadores** de, pelo menos, **uma característica comum**.

**AMOSTRA:** é uma **parcela representativa da população** que **é examinada** com o propósito de tirarmos conclusões sobre a essa população.

**PARÂMETROS:** São **valores singulares que existem na população** e que **servem para caracterizá-la**. Para definirmos um parâmetro devemos examinar toda a população. *Ex: Os alunos do 2º ano da UNC têm em média 1,70 metros de estatura.*

**ESTIMATIVA:** é um **valor aproximado do parâmetro** e é calculado com o uso da amostra.

**ATRIBUTO:** quando os dados estatísticos apresentam um caráter qualitativo, o levantamento e os estudos necessários ao tratamento desses dados são designados genericamente **de estatística de atributo**.

**VARIÁVEL:** É o **conjunto de resultados possíveis de um fenômeno**.

**VARIÁVEL QUALITATIVA:** Quando seu valores são expressos por **atributos**: sexo, cor da pele,etc.

**VARIÁVEL QUANTITATIVA:** Quando os **dados são de caráter nitidamente quantitativo**, e o conjunto dos resultados **possui uma estrutura numérica**, trata-se portanto da estatística de variável e se dividem em :

**VARIÁVEL DISCRETA OU DESCONTÍNUA**: Seus valores são expressos geralmente através de números inteiros não negativos. Resulta normalmente de contagens. *Ex: Nº de alunos presentes às aulas de introdução à estatística econômica no 1º semestre de 1997: mar = 18 , abr = 30 , mai = 35 , jun = 36.*

**VARIÁVEL CONTÍNUA**: Resulta normalmente de uma **mensuração**, e a **escala numérica de seus possíveis valores corresponde ao conjunto R** dos números Reais, ou seja, podem assumir, teoricamente, qualquer valor entre dois limites. *Ex.: Quando você vai medir a temperatura de seu corpo com um termômetro de mercúrio o que ocorre é o seguinte: O filete de mercúrio, ao dilatar-se, passará por todas as temperaturas intermediárias até chegar na temperatura atual do seu corpo.*

Exemplos **-**

. Cor dos olhos das alunas: **qualitativa**

. Índice de liquidez nas indústrias capixabas: **quantitativa contínua**

. Produção de café no Brasil: **quantitativa contínua**

. Número de defeitos em aparelhos de TV: **quantitativa discreta**

. Comprimento dos pregos produzidos por uma empresa: **quantitativa contínua**

. O ponto obtido em cada jogada de um dado: **quantitativa discreta**

**AMOSTRAGEM**

**MÉTODOS PROBABILÍSTICOS**

 Exige que cada elemento da população possua determinada **probabilidade de ser selecionado**. Normalmente possuem a mesma probabilidade. Assim, se **N** for o tamanho da população, a probabilidade de cada elemento ser selecionado será **1/N**. Trata-se do método que garante cientificamente a aplicação das técnicas estatísticas de inferências. Somente com base em amostragens probabilísticas é que se podem realizar inferências ou induções sobre a população a partir do conhecimento da amostra.

* + - É uma técnica especial para recolher amostras, que garantem, tanto quanto possível, o acaso na escolha.

.

**AMOSTRAGEM CASUAL** ou  **ALEATÓRIA SIMPLES**

 É o processo mais elementar e freqüentemente utilizado. É **equivalente a um sorteio lotérico**. Pode ser realizada numerando-se a população de **1** a **n** e sorteando-se, a seguir, por meio de um dispositivo aleatório qualquer, **x** números dessa seqüência, os quais corresponderão aos elementos pertencentes à amostra.

*Ex:* *Vamos obter uma amostra, de 10%, representativa para a pesquisa da estatura de 90 alunos de uma escola:*

*1º - numeramos os alunos de 1 a 90.*

*2º - escrevemos os números dos alunos, de 1 a 90, em pedaços iguais de papel, colocamos na urna e após mistura retiramos, um a um, nove números que formarão a amostra.*

**OBS:** quando o número de elementos da amostra é muito grande, esse tipo de sorteio torna-se muito trabalhoso. Neste caso utiliza-se uma Tabela de números aleatórios, construída de modo que os algarismos de 0 a 9 são distribuídos ao acaso nas linhas e colunas.

.

**.AMOSTRAGEM PROPORCIONAL ESTRATIFICADA:**

 Quando a população se divide em estratos (sub-populações), convém que o sorteio dos elementos da amostra leve em consideração tais estratos, daí obtemos os elementos da amostra proporcional ao número de elementos desses estratos.

*Ex: Vamos obter uma amostra proporcional estratificada, de 10%, do exemplo anterior, supondo, que, dos 90 alunos, 54 sejam meninos e 36 sejam meninas. São portanto dois estratos (sexo masculino e sexo feminino). Logo, temos:*

|  |  |  |  |
| --- | --- | --- | --- |
| SEXO  | POPULACÃO  | 10 %  | AMOSTRA  |
| MASC. | 54 | 5,4 | 5 |
| FEMIN. | 36 | 3,6 | 4 |
| Total | 90 | 9,0 | 9 |

Numeramos então os alunos de 01 a 90, sendo 01 a 54 meninos e 55 a 90, meninas e procedemos o sorteio casual com urna ou tabela de números aleatórios.

.

**AMOSTRAGEM SISTEMÁTICA:**

 Quando os elementos da população já se acham ordenados, não há necessidade de construir o sistema de referência. São exemplos os prontuários médicos de um hospital, os prédios de uma rua, etc. Nestes casos, a seleção dos elementos que constituirão a amostra pode ser feita por um sistema imposto pelo pesquisador.

*Ex: Suponhamos uma rua com 900 casas, das quais desejamos obter uma amostra formada por 50 casas para uma pesquisa de opinião. Podemos, neste caso, usar o seguinte procedimento: como 900/50 = 18, escolhemos por sorteio casual um número de 01 a 18, o qual indicaria o primeiro elemento sorteado para a amostra; os demais elementos seriam periodicamente considerados de 18 em 18. Assim, suponhamos que o número sorteado fosse 4 a amostra seria: 4ª casa, 22ª casa, 40ª casa, 58ª casa, 76ª casa, etc.*

**AMOSTRAGEM POR CONGLOMERADOS (ou AGRUPAMENTOS)**

 Algumas populações não permitem, ou tornam extremamente difícil que se identifiquem seus elementos. Não obstante isso, pode ser relativamente fácil identificar alguns subgrupos da população. Em tais casos, uma amostra aleatória simples desses subgrupos (conglomerados) pode se colhida, e uma contagem completa deve ser feita para o conglomerado sorteado. Agrupamentos típicos são quarteirões, famílias, organizações, agências, edifícios etc.

*Ex: Num levantamento da população de determinada cidade, podemos dispor do mapa indicando cada quarteirão e não dispor de uma relação atualizada dos seus moradores. Pode-se, então, colher uma amostra dos quarteirões e fazer a contagem completa de todos os que residem naqueles quarteirões sorteados*.

**MÉTODOS NÃO PROBABILÍSITCOS**

 São amostragens em que há uma **escolha deliberada dos elementos da amostra**. Não é possível generalizar os resultados das pesquisas para a população, pois as ***amostras não-probabilísticas não garantem a representatividade da população***.

**AMOSTRAGEM ACIDENTAL**

 Trata-se de uma **amostra formada por aqueles elementos que vão aparecendo**, que são possíveis de se obter até completar o número de elementos da amostra. **Geralmente utilizada em pesquisas de opinião**, em que os entrevistados são acidentalmente escolhidos.

*Ex: Pesquisas de opinião em praças públicas, ruas de grandes cidades*;

**AMOSTRAGEM INTENCIONAL**

 De acordo com determinado critério, **é escolhido intencionalmente um grupo de elementos que irão compor a amostra**. O investigador se dirige intencionalmente a grupos de elementos dos quais deseja saber a opinião.

*Ex: Numa pesquisa sobre preferência por determinado cosmético, o pesquisador se dirige a um grande salão de beleza e entrevista as pessoas que ali se encontram.*

**AMOSTRAGEM POR QUOTAS**

 Um dos métodos de **amostragem mais comumente usados em levantamentos de mercado** e em **prévias eleitorais**. Ele abrange três fases:

1ª - classificação da população em termos de propriedades que se sabe, ou presume, serem relevantes para a característica a ser estudada;

2ª - determinação da proporção da população para cada característica, com base na constituição conhecida, presumida ou estimada, da população;

3ª - fixação de quotas para cada entrevistador a quem tocará a responsabilidade de selecionar entrevistados, de modo que a amostra total observada ou entrevistada contenha a proporção e cada classe tal como determinada na 2ª fase.

*Ex: Numa pesquisa sobre o "trabalho das mulheres na atualidade", provavelmente se terá interesse em considerar: a divisão cidade e campo, a habitação, o número de filhos, a idade dos filhos, a renda média, as faixas etárias etc.*

*A primeira tarefa é descobrir as proporções (porcentagens) dessas características na população. Imagina-se que haja 47% de homens e 53% de mulheres na população. Logo, uma amostra de 50 pessoas deverá ter 23 homens e 27 mulheres. Então o pesquisador receberá uma "quota" para entrevistar 27 mulheres. A consideração de várias categorias exigirá uma* ***composição amostral*** *que atenda ao* ***n*** *determinado e às proporções populacionais estipuladas.*

.

**GRÁFICOS ESTATÍSTICOSG**

 São representações visuais dos dados estatísticos que devem corresponder, mas nunca substituir as tabelas estatísticas.

Características: Uso de escalas, sistema de coordenadas, simplicidade, clareza e veracidade.

***Gráficos de informação:*** São gráficos destinados principalmente ao público em geral, objetivando proporcionar uma visualização rápida e clara. São gráficos tipicamente expositivos, dispensando comentários explicativos adicionais. As legendas podem ser omitidas, desde que as informações desejadas estejam presentes.

***Gráficos de análise:*** São gráficos que prestam-se melhor ao trabalho estatístico, fornecendo elementos úteis à fase de análise dos dados, sem deixar de ser também informativos. Os gráficos de análise freqüentemente vêm acompanhados de uma tabela estatística. Inclui-se, muitas vezes um texto explicativo, chamando a atenção do leitor para os pontos principais revelados pelo gráfico.

* **Uso indevido de Gráficos:** Podem trazer uma idéia falsa dos dados que estão sendo analisados, chegando mesmo a confundir o leitor. Trata-se, na realidade, de um problema de construção de escalas.

.

**Classificação dos gráficos**: Diagramas, Estereogramas, Pictogramas e Cartogramas.

.

**1 - Diagramas:**

 São gráficos geométricos dispostos em duas dimensões. São os mais usados na representação de séries estatísticas. Eles podem ser :

* 1. **Gráficos em barras horizontais**.

1.2- **Gráficos em barras verticais** **( colunas ).**

* Quando as legendas não são breves usa-se de preferência os gráficos em barras horizontais. Nesses gráficos os retângulos têm a mesma base e as alturas são proporcionais aos respectivos dados.
	+ - A ordem a ser observada é a cronológica, se a série for histórica, e a
		- **decrescente**, se for geográfica ou categórica.
	1. **Gráficos em barras compostas**.

1.4- **Gráficos em colunas superpostas**.

* Eles diferem dos gráficos em barras ou colunas convencionais apenas pelo fato de apresentar cada barra ou coluna segmentada em partes componentes. Servem para representar comparativamente dois ou mais atributos.

1.5- **Gráficos em linhas ou lineares**.

* São freqüentemente usados para representação de séries cronológicas com um grande número de períodos de tempo. As linhas são mais eficientes do que as colunas, quando existem intensas flutuações nas séries ou quando há necessidade de se representarem várias séries em um mesmo gráfico.
* Quando representamos, em um mesmo sistema de coordenadas, a variação de dois fenômenos, a parte interna da figura formada pelos gráficos desses fenômenos é denominada de ***área de excesso.***

1.5- **Gráficos em setores**.

* Este gráfico é construído com base em um círculo, e é empregado sempre que desejamos ressaltar a participação do dado no total. O total é representado pelo círculo, que fica dividido em tantos setores quantas são as partes. Os setores são tais que suas áreas são respectivamente proporcionais aos dados da série. O gráfico em setores só deve ser empregado quando há, no máximo, sete dados.
* **Obs:** As séries temporais geralmente não são representadas por este tipo de gráfico.

.**2 - Estereogramas:**

 São gráficos geométricos dispostos em três dimensões, pois representam volume. São usados nas representações gráficas das tabelas de dupla entrada. Em alguns casos este tipo de gráfico fica difícil de ser interpretado dada a pequena precisão que oferecem.

.

**3 - Pictogramas:**

 São construídos a partir de figuras representativas da intensidade do fenômeno. Este tipo de gráfico tem a vantagem de despertar a atenção do público leigo, pois sua forma é atraente e sugestiva. Os símbolos devem ser auto-explicativos. A desvantagem dos pictogramas é que apenas mostram uma visão geral do fenômeno, e não de detalhes minuciosos. Veja o exemplo abaixo:



**4- Cartogramas:**

**** São ilustrações relativas a cartas geográficas (mapas). O objetivo desse gráfico é o de figurar os dados estatísticos diretamente relacionados com áreas geográficas ou políticas.