**ESTATÍSTICA**

**2º Ano**

**CIÊNCIAS CONTÁBEIS**

**ADMINISTRAÇÃO**

**ANÁLISE DE SISTEMAS**

**Profa. Ms. Sandra Mara Neri Vidotto**

**2019**

**Introdução à Estatística**

1.1 A Ciência Estatística

O conceito de Estatística pode ser considerado de duas maneiras. O primeiro conceito, logo relaciona a Estatística com tabelas e gráficos nos quais os dados obtidos são representados, ou melhor, relaciona à números específicos. Ouvimos, assim, falar em estatísticas do IBGE, estatísticas relacionadas à saúde e educação, índices econômicos, pesquisas de opinião, etc. Um segundo conceito refere-se ao conjunto de processos ou técnicas empregadas na investigação e análise de fenômenos. Neste caso, a Estatística é a ciência ou método científico que estuda os fenômenos aleatórios e, procura inferir as leis que os mesmos obedecem. Assim, um conceito mais abrangente e absoluto deve englobar tanto o primeiro conceito, o qual é o mais popular, quanto o segundo, o qual normalmente escapa à noção corrente.

Definição 1.1 (Estatística). A Estatística é uma ciência que se preocupa com a coleta, organização, descrição, análise e interpretação dos dados, a fim de extrair informações a respeito de uma população.

Dentro dessa idéia, podemos considerar a Ciência Estatística como dividida basicamente em duas partes:

1. Estatística Descritiva - que se preocupa com a organização e descrição dos dados experimentais;

2. Estatística Inferencial - que, a partir da observação de alguns dados experimentais, realiza a análise e interpretação de dados com o objetivo de generalizar e prever resultados, utilizando-se para isto da Teoria das Probabilidades.

Nesta disciplina, serão abordados tópicos referentes à estatística descritiva, conceitos fundamentais de probabilidade e os modelos probabilísticos mais importantes para o estudo da inferência estatística.

1.2 Estatística: Uma Visão Sistêmica (Desenhar figura representando uma visão sistêmica da estatística)

1.3 Conceitos Fundamentais Um dos principais conceitos utilizados na estatística é o de população.

1.3.1 População e Amostra

Definição 1.2 (População). A população é um conjunto de todos os elementos

(pessoas, objetos, etc) que possuem pelo menos uma característica em comum, a(s) qual(is) os relacionam ao problema que está sendo estudado.

Exemplo 1. Se o problema a ser pesquisado está relacionado com a qualidade de um certo produto produzido numa indústria, a população pode ser composta por todas as peças produzidas numa determinada hora, turno, dia ou mês, dependendo dos objetivos;

Exemplo 2. Se o objetivo de um estudo é pesquisar o nível de renda familiar de uma certa cidade, a população seria todas as famílias desta população. Mas, se o objetivo fosse pesquisar apenas a renda mensal do chefe da família, a população a ser pesquisada seria composta por todos os chefes de família desta cidade.

A População pode ser:

1. Finita - quando o número de unidades de observação pode ser contado e é limitado; 2. Infinita - quando a quantidade de unidades de observação é ilimitada;

Podemos citar como exemplo de população finita o conjunto formado pelos alunos que cursam a disciplina de estatística num determinado semestre da UFCG. Um exemplo de população infinita seria o conjunto formado por todos os alunos de estatística do Brasil, pois este conjunto é composto por um número incontável de elementos.

Definição 1.3 (Amostra). A amostra é apenas uma parte da população, ou seja, é um subconjunto da população.

Vários motivos levam a necessidade de se observar apenas uma parte da população, como, por exemplo: a falta de tempo, recursos financeiros e/ou humanos. A amostra deve ser obtida através de técnicas de amostragem, as quais tem como objetivo principal garantir a representatividade da população, ou seja, fazer com que a amostra seja um retrato fiel da população.

Exemplos de amostra podem ser conjuntos formados por apenas uma parte dos elementos populacionais descritos nos Exemplos 1 e 2.

1.3.2 Parâmetro e Estatística

Dois novos conceitos estreitamente relacionados com os de população e amostra são os de Parâmetro e Estatística, tendo em vista que:

Definição 1.4 (Parâmetro). é uma medida numérica que descreve uma característica da população, ou ainda, que é obtida a partir de todos os dados populacionais (através de um censo).

Definição 1.5 (Estatística). é uma medida numérica que descreve uma característica da amostra, ou ainda, que é obtida a partir de dados amostrais (de uma parte da população).

Exemplos de algumas medidas numéricas são: proporção, média, moda, índices, etc.

1.3.3 Variáveis (ou Dados) e Tipos de Variáveis

Definição 1.6 (Variável). Uma Variável nada mais é que uma característica (ou dado) associada a cada elemento da população ou amostra. A variável apresenta diferentes valores, quando sujeita a mensurações sucessivas, e, em geral, é denotada pelas letras maiúsculas: X, Y ou Z.

Antes de realizar qualquer tratamento estatístico de um conjunto de dados, é importante identificar qual é o tipo de dado (ou variável) que será analisado, pois, é mediante a este conhecimento que o pesquisador poderá ou não adotar determinadas técnicas estatísticas para a resolução de problemas. Por exemplo, será que é possível calcular o peso médio de lutadores de boxe, quando os dados são coletados segundo a categoria de peso: Leve, Médio ou Pesado?

Tipos de Variáveis

Basicamente, as variáveis podem ser classificadas como sendo Qualitativas ou Quantitativas.

1. Variáveis Qualitativas - quando os valores que elas podem receber são referentes à qualidade, atributo ou categoria. Exemplos são:

• Raça: podendo assumir os valores Branco ou Negro; • Sexo: Masculino ou Feminino;

• Escolaridade: 1◦ grau completo, 2◦ grau completo, superior, pós-graduado;

• Conceito de qualidade: péssima qualidade, regular ou boa qualidade.

As variáveis qualitativas podem, ainda, ser classificadas como: Nominais ou Ordinais.

(a) As variáveis qualitativas nominais - são caracterizadas por dados que se apresentam apenas sob o aspecto qualitativo (Ex: raça e sexo).

(b) As variáveis qualitativas ordinais - são caracterizadas por categorias que aprentam uma ordenação natural. Por exemplo: escolaridade e conceito de qualidade.

2. Variáveis Quantitativas - quando os valores que ela pode assumir são numéricos, os quais podem ser obtidos através de uma contagem ou mensuração.

As variáveis quantitativas podem ser classificadas de acordo com o processo de obtenção; podendo ser: Discreta ou Contínua.

(a) As variáveis quantitativas discretas - são variáveis numéricas obtidas a partir de procedimento de contagem. Por exemplo: Quantidade de pessoas numa família, quantidade de acidentes numa indústria, etc.

(b) As variáveis quantitativas contínuas - são variáveis numéricas cujos valores são obtidos por um procedimento de mensuração, podendo assumir quaisquer valores num intervalo dos números reais, como por exemplo, a temperatura, altura, salário, etc..

Observação 1. O fato de uma variável poder ser expressa por números não significa que ela seja necessariamente quantitativa, por que a classificação da variável depende de como foi medida. Por exemplo, para a variável peso de um lutador de boxe, se for anotado o peso marcado na balança, a variável é quantitativa contínua; por outro lado, se esse peso for classificado segundo as categorias do boxe, a variável é qualitativa ordinal.

1.4 Fases do Método Estatístico

Assim como qualquer ciência, a estatística utiliza o método científico, que consiste das cinco etapas básicas seguintes:

1. Definir cuidadosamente o problema.

Nesta etapa o pesquisador deve certificar-se de que é clara a finalidade de um estudo ou análise. Ao definir o que se quer estudar, ou seja, o problema, é necessário que se faça um levantamento sobre quais estudos já realizados no campo de pesquisa abordado. Deve-se também especificar quem ou o quê será observado no estudo, ou seja, a população a ser pesquisada.

Nesta fase, o pesquisador deverá listar as variáveis (características ou dados) que sejam relevantes para se atingir os objetivos propostos pela pesquisa. Além disso, deve-se decidir se a coleta dos dados será realizada através de um censo ou amostragem, ou seja, se todos os elementos da população serão observados ou se apenas uma parte da população é que será observada e neste último caso deve-se decidir por alguma técnica de amostragem, podendo ser probabilística ou não.

Os dados podem ser classificados quanto à forma de coleta, como:

a. Dados primários - quando o próprio pesquisador é quem elabora e aplica os instrumentos necessários para a coleta dos dados, ou seja, quando a Coleta é Direta; b. Dados secundários - quando o pesquisador utiliza informações já colhidas por outrem, retirando-as de livros, revistas, mapas anuários, etc.

2. Coligir ou apurar os dados.

Esta fase consiste em resumir os dados, através de sua contagem e agrupamento. É possível que nesta fase seja identificado a presença de dados absurdos fazendo-se necessário a eliminação ou correção destes tipos de dados.

3. Analisar e interpretar os dados.

4. Relatar as conclusões de maneira que sejam facilmente entendidas por quem as for usar na tomada de decisões.

**1a LISTA DE EXERCÍCIOS**

1 ) Defina e/ou explique com suas próprias palavras, o que você entende por Ciência Estatística e quais os principais ramos (partes) da Estatística.

2)Através de um exemplo, defina: População e Amostra. 3 - Considere as seguintes situações:

3) Em uma pesquisa, feita pela EMPETUR com 1015 pousadas escolhidas aleatoriamente, 269 (ou 26,5%) possuíam Home-page na Internet para divulgação e prestação de serviços ao turista.

4) Outra pesquisa feita entre as 50 Agências de Viagens de uma certa localidade mostra que 42 (ou 84%) prestam serviços pela Internet.

Identifique em qual das situações nós temos um exemplo de Parâmetro e outro de Estatística (no sentido de medida). Justifique sua resposta.

5)O que você entende por variável? Justifique a sua resposta por intermédio de um exemplo.

6) Como você diferencia uma variável discreta de uma variável contínua? Utilize um exemplo para melhor ilustrar.

7)Defina e/ou explique com suas próprias palavras, o que você entende por amostragem.

8)Qual é o principal objetivo de qualquer plano de amostragem?

9) As estatísticas geradas por intermédio de uma amostra devem ser representativas desta amostra ou da população de origem? Justifique a sua resposta.

10) Para que uma amostra seja representativa, é necessário apenas que a mesma tenha um tamanho apropriado? Justifique a sua resposta.

11)A Revista dos Eventos, N 13, tentando sanar, ao menos parcialmente, a carência de informações precisas sobre a indústria de eventos, promoveu a 1a PESQUISA - O Mercado de Congressos no Brasil. Os resultados desta pesquisa se baseiam em 40 questionários respondidos sobre um total de 1000, os quais foram encaminhados por entrega pessoal a dirigentes de entidades integrantes do cadastro da própria Revista dos Eventos. Qual é o problema ou a limitação desta pesquisa? Pelo menos teoricamente, qual seria o melhor procedimento para este tipo de pesquisa, já que a empresa possui um cadastro das entidades?

ESTATÍSTICA DESCRITIVA

Constitui num conjunto de técnicas que objetivam descrever, analisar e interpretar os dados numéricos de uma população ou amostra.

1. **Dados**

Uma vez coletados os valores das variáveis, passa-se a ser tratados como dados, que são considerados Dados Brutos, quando não existem nenhuma ordenação entre eles. Exemplo:

28-36-17-54-98-62-9

Chama-se de **Rol**, ao ordenamento dos dados brutos em ordem crescente ou decrescente. Exemplo:

9-17-28-36-54-62-98

Uma vez coletados os dados, pode-se trabalhar no sentido de estudar o seu comportamento dentro daquele conjunto. Podendo calcular alguns parâmetros e mesmo representando-o através de tabelas e gráficos, que possibilitam uma melhor visualização geral da distribuição dos dados.

1. **Representação Tabular**

Consiste em apresentar os dados coletados através de tabelas, dando uma visão mais clara do que ocorre com os dados observados.

Existem várias maneiras de representar um conjunto de dados através de tabelas, depende para isso, do tipo de dados e da classificação que se queira dar, fixando uma das três modalidades que caracteriza um fato: tempo, local, e fato observado.

Podendo classifica-la em cinco tipos:

- Série Cronológica;

- Série Geográfica;

- Série Específica;

- Série Conjugada;

- Distribuição de Freqüência.

Sendo que a distribuição de freqüência é usada quando os dados pertencem a um intervalo numérico, e tem uma técnica específica parta construí-la, que será visto logo em seguida.

Para construir uma tabela estatística, também conhecida como série estatística. Existem algumas normas universais que devem ser respeitadas, tais como:

2.1– Elementos Essenciais

- Título: é a indicação que precede a tabela e contém a designação do fato observando, o local e a época em que foi obtido.

- Cabeçalho: é a parte superior da tabela que especifica o conteúdo das colunas.

- Corpo: é o conjunto de colunas e linhas que contém respectivamente, em ordem vertical e horizontal, as informações sobre o fato observado.

- Coluna indicadora: é a parte da tabela que especifica ao conteúdo das linhas.

2.2– Elementos Complementares

Geralmente se situam no rodapé da tabela sendo a:

- Fonte: Específica a entidade responsável pelo fornecimento dos dados ou pela sua construção.

- Notas: são informações gerais, para esclarecer o conteúdo da tabela ou indicar a metodologia adotada.

- Chamada: são informações específicas sobre determinada parte da tabela, ou esclarecer dados.

As tabelas são obrigatoriamente fechadas por traços horizontais, enquanto que os verticais só podem ser usados para separação de colunas.

Pode-se Empregar alguns sinais como:

Traço ( - ), quando o dado for nulo;

Três Pontos ( ... ), quando existem falta de dados;

Zero ( 0 ), quando o valor numérico for muito pequeno com relação aos que estão sendo expressos;

XIX ( x ), quando o dado foi emitido.

2.3– Série Cronológica

São tabelas em que os dados são distribuídos de acordo com o tempo, e deixando fixo o local do evento e o fato observado. Exemplo:

2.4– Série Geográfica

São as séries em que o fato observado e tempo permanecem constantes e os dados são distribuídos conforme o local. Exemplo:

2.5– Série Específica

Nesta ficam fixas o tempo e o local do evento, sendo od dados distribuídos de acordo com o fato observado. Exemplo:

2.6– Série Conjugada

São aquelas tabelas construídas com uma combinação entre os tipos apresentados anteriormente. Tabelas que apresentam dupla entrada. Exemplo:

3 – **Apresentação Gráfica**

Depois de sintetizados em tabelas, os dados podem ser apresentados em gráficos a fim de proporcionarem ao interessado uma visão global dos dados. Os gráficos são usados para que haja uma comunicação clara, rápida e efetiva.

**Tipos de Gráficos**

Gráfico Em Linha Gráfico de Bastão

Gráfico em Colunas Gráfico em Barras

Gráfico em setores

**Gráfico de Colunas ou Barras Múltiplas**

Permitem comprar diversas variáveis simultaneamente. Caracteriza-se por apresentar duas ou mais colunas ou barras representativas de variáveis num mesmo período de tempo, sem espaço entre si, formando conjunto de colunas ou barras, existindo o espaço entre os conjuntos.

Além dos gráficos de colunas ou barras múltiplas, tem-se também os seguintes gráficos de colunas ou barras remontadas, gráficos de colunas ou barras compostas ou superpostas, gráficos de colunas ou barras contrapostas.

Neste caso de representação de duas ou mais variáveis, pode-se utilizar também o gráfico.

Alguns exemplos:

Gráfico em Colunas Justapostas Gráficos em Colunas Remontadas

Gráfico de Colunas Superpostas Gráfico em Linhas

**Gráfico de Barras**

Nesse tipo de gráfico usamos retângulos com bases de mesma medida e separados por distâncias iguais. As freqüências dos fatos observados são dadas pelas alturas dos retângulos, anotadas no eixo **y**, se as barras forem verticais. Se as barras forem horizontais, as freqüências são dadas pelos comprimentos dos retângulos, anotados no eixo **x**.

Vejamos alguns exemplos:

a) A tabela e o gráfico representam os percentuais de reprovação em determinada disciplina no ano letivo:

|  |  |
| --- | --- |
| **Bimestre** | **Percentual** |
| 1º | 35% |
| 2º | 50% |
| 3º | 20% |
| 4º | 30% |

b) A tabela e o gráfico representam a avaliação da União Nacional dos Estudantes (UNE), em porcentagem, feita pelos estudantes:

|  |  |
| --- | --- |
| **Avaliação – UNE** | |
| Ótima | 4% |
| Boa | 25% |
| Regular | 27% |
| Ruim | 9% |
| Péssima | 13% |
| Abstenção | 22% |

O número à frente das barras horizontais, indicando o percentual, torna mais claro o gráfico.

c) Um gráfico de barras também pode ser usado na forma comparativa, apresentando dois ou mais conjuntos de dados:

|  |  |  |
| --- | --- | --- |
| **Utilidades Vendidas** | | |
| **Mês** | **1993** | **1994** |
| Janeiro | 15 | 12 |
| Fevereiro | 10 | 18 |
| Março | 12 | 20 |
| Abril | 18 | 25 |

Outros exemplos podem ser vistos nos jornais, como gráficos estatísticos que eles publicam diariamente.

**Gráfico de Setores**

Os dados são apresentados em setores circulares proporcionais aos valores. Para representar os setores, fazemos corresponder a uma volta do circulo (360º) o total (100%) dos dados e estabelecemos com uma regra de três o ângulo relativo ao setor de acordo com cada valor.

***Exercício Resolvido***

Construa o gráfico de setores de acordo com os dados da tabela.

|  |  |
| --- | --- |
| **Produtos** | **Quantidade Produzida (unidades)** |
| A | 32,4 |
| B | 21,6 |
| C | 43,2 |
| D | 10,8 |
| Total | 108,0 |

- Produto **C**

108 ----- 360º C =  = 144º

43,2 ---- C

- Produto **D**

108 ---- 360º  D =  = 36º

**Solução:** 10,8 --- D

- Produto **A**

108 ------ 360º  A =  = 108º

32,4 ----- A

- Produto **B**

108 ----- 360º  B =  = 72º

21,6 ---- B

**Histograma**

O histograma é formado por retângulos justapostos, sendo o número de retângulos igual ao número de intervalos de classe. A largura de cada retângulo é igual a amplitude do intervalo de classe, enquanto a sua altura representa a freqüência do intervalo de classe. A área do histograma é proporcional a soma das freqüências.

Vejamos o histograma para a distribuição de freqüências que segue:

|  |  |
| --- | --- |
| **Peso (Kg.)** | **Freqüência** |
| 40  50 | 6 |
| 50  60 | 10 |
| 60  70 | 18 |
| 70  80 | 12 |
| 80  90 | 4 |

**Polígono de Freqüências**

O polígono de freqüências é um gráfico de linha em que cada ponto é obtido considerando-se como valor de **x** o ponto médio do intervalo de classe e como valor de **y** a respectiva freqüência do intervalo. Consideramos também uma classe anterior à primeira e outra posterior à última. Ligando todos os pontos, temos o polígono de freqüências.

Vamos observar o polígono de freqüências para a distribuição anterior, em que os pontos médios dos intervalos são dados por 45, 55, 65, 75 e 85; acrescentamos os valores 35 e 95 para as classes anterior à primeira e posterior à última, com freqüência nula:

4. Uma distribuição de freqüência é uma série de estatísticas onde os dados são agrupados em intervalos, também chamados por classes, de tal forma que possa determinar o número de elementos em cada classe. Isso proporciona a visualização de um conjunto de números sem levar em conta os valores individualmente.

**Construção de uma Distribuição de Freqüência**

Na construção de uma distribuição de freqüência deve-se seguir as seguintes etapas:

1º) Determinar a amplitude total (At) através da seguinte relação:

**At = n.º maior – n.º menor**

2º) Determinar o número de classes (**K**)

A definição do número de classes deve ser seguida pelos objetivos que melhor representa o fato ocorrido. Existem várias técnicas desenvolvidas com este objetivo porém, **deve ficar bem claro que estas técnicas devem servir como indicação e não como regra fixa**.

Aconselha-se que o “número de classes não deva ser inferior a 5 e nem superior a 15”, pois a eficiência da sumarização estaria comprometida, assim menos de 5 classes pode ocultar detalhes relevantes e mais de 15 torna a compreensão detalhada em demasia.

Então para se determinar o número de classes pode-se usar um dos seguintes critérios:

a) critério raiz

**K =** 

b) fórmula de Sturges

**K = 1 + 3,3 log n**

onde **n** é o número de elementos do rol.

3º) Determinar o intervalo de classe (**C**)

Estabelecido o número de classes a se utilizar, determina-se o intervalo de classe:

**C = **

4º) Construção das classes:

Na construção das classes deve-se tomar cuidado com a sua representação para que não haja dúvidas, ou seja, não se exclua nenhuma observação e não provoque uma superposição de classes.

Podemos usar umas das seguintes representações:

a) Li |---- Ls inclui o extremo inferior e exclui o extremo superior

b) Li ----| Ls exclui o extremo inferior e inclui o extremo superior

c) Li ---- Ls exclui os dois extremos

d) Li |----| Ls inclui os dois extremos

5º) Freqüência Absoluta (Fi)

É o número de observações do rol que se localiza em cada classe.

Outros tipos de freqüências são normalmente calculadas numa tabela, tais como:

- Freqüência Relativa

- Freqüência Relativa Percentual

- Freqüência Absoluta Acumulada

- Freqüência Relativa Acumulada

6º) Freqüência Relativa (Fri)

A freqüência relativa apresenta a proporção de observações de um valor de uma classe, em relação ao número total de observações. Para calcular a freqüência relativa, basta dividir a freqüência absoluta da classe pelo número total de observações.

Fri =  = 

A freqüência relativa geralmente é expressa em porcentagem, para isso, basta multiplicar o quociente obtido por 100. Assim:

Fri% =  x 100 %

7º) Freqüência Acumulada (Fa)

A freqüência acumulada de uma classe é a soma da freqüência absoluta dessa classe com as freqüências absolutas das classes anteriores ou posteriores, obtendo assim as freqüências acumuladas crescente ou decrescentes, respectivamente.

A freqüência acumulada pode ser tanto na freqüência absoluta quanto na freqüência relativa.

8º) Ponto Médio (Xi)

É o ponto intermediário de cada classe e é determinado por:

Xi = 

9º) Representação Gráfica da Distribuição de freqüências

Histograma

É um gráfico formado por retângulos justapostos, onde cada retângulo representa a freqüência absoluta de cada classe. Sendo construído num par de eixos ortogonais, em que no eixo das abscissas destacamos os intervalos de classes e no eixo da ordenadas, uma escala que representa as freqüências absolutas de cada classe. Para construir o histograma podemos utilizar qualquer tipo de freqüência.

EXERCÍCIOS DE fFIXAÇÃO

1) O seguinte conjunto de dados é referente ao número de acidentes por dia em certo trecho de rodovia no mês de setembro de certo ano:

http://s3.amazonaws.com/magoo/ABAAABTdYAI-9.jpg

Responda as seguintes questões:

a) Qual o número mínimo de acidentes, num certo dia? E o número máximo? b) Freqüêntemente, ocorreram quantos acidentes por dia? E o que isso representa em termos de percentuais? c) Represente graficamente a distribuição de frequência da variável número de acidentes por dia, no mês de setembro.d) Faça um gráfico de colunas para o percentual acumulado.

2)Feito um levantamento sobre os salários de 40 empregados de uma firma, pode-se elaborar o seguinte quadro de distribuição conforme os dados:

420 - 690- 480- 1100- 678 -980-435- 876- 980- 567- 1100- 1090- 889- 1038- 778 -456- 456-876-789 -667- 987-1234- 567 -789- 457- 887- 980- 678- 879- 887- 556- 557-667-667-890- 789-987-1009-1100-450

1. Elabore um quadro de distribuição de freqüências, acumuladas, relativas, relativas acumuladas e ponto médio, inicie por R$ 400,00 e intervalo de R$ 100,00.
2. Quantos empregados dessa firma ganham menos que R$ 1.000,00 mensais?
3. Qual o índice em % de empregados dessa firma quer ganham mais de R$ 1.000,00 mensais?
4. Quantos empregados dessa firma ganham entre R$ 800,00 e R$ 1.200,00?
5. Qual é o índice em % de empregados dessa firma que ganham menos de R$ 1.000,00?
6. Construir o histograma.

3)Montar uma série cronológica para representar dados:

Os valores das exportações de café , fornecidas pelo Instituto de Café do Brasil, nos anos de 1995 a 2000 em milhares de dólares :60.564,-82.137, –43.267,-112. 765,-453.345,-251.234.

4)O gráfico abaixo mostra a distribuição dos conteúdos de Química exigidos nos vestibulares de 1996, 1995 e 1994 em três universidades brasileiras.

Considerando os dados apresentados no gráfico pode-se dizer que:



1. a Química Orgânica foi a mais exigida em todos os anos
2. a exigência de Química Geral de 1995 para 1996 aumentou
3. os conteúdos de Química foram exigidos igualmente
4. decresceu a exigência em Físico-Química de 94 a 96
5. aumentou a exigência em Orgânica de 94 a 9

5)O gráfico abaixo mostra a situação dos alunos de uma escola no período 92-95



Podemos afirmar que:

1. O número de alunos aprovados cresceu de 92 a 93
2. O número de alunos aprovados em 92 é igual ao de 94
3. Em 93 houve o maior número de alunos aprovados
4. O número de alunos aprovados diminui de 93 para 94
5. Em 95, o número de alunos aprovados é igual ao de reprovados
6. Considere o gráfico abaixo:



Examinando o gráfico, é correto afirmar que:

1. os "mais pobres" são a maioria da população
2. os "mais ricos" são a maioria da população
3. quem tem renda mensal de R$ 1.000 não está representado
4. quem tem renda mensal de R$ 50 não está representado
5. quem tem renda mensal de R$ 1.200 pertence à região branca do gráfico

7)Um sensor instalado em uma rodovia conta o número de automóveis que passa por minuto. Os dados obtidos em um período de 1 hora, foram:

25 40 13 17 34 30 25 20 28 43 23 21 32 38

20 16 45 52 26 38 30 35 46 49 11 19 78 53

65 54 45 48 16 23 27 14 63 53 41 17 24 33

62 36 33 50 27 10 29 58 44 32 23 39 46 32 31

COMPLETE O QUADRO DE DISTRIBUIÇÃO DE FREQUÊNCIAS E FAÇA UM HISTOGRAMA.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| CLASSES | FI | FA | FR | FRA | XI |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

## ***8). A massa (em quilogramas) de 20 trabalhadores de uma empresa com 100 funcionários está registrada a seguir.***

## ***65 52 73 80 65 50 70 75 80 65 70 77 82 91 75 52 68 86 70 80***

## ***Com base nos dados obtidos,responda:***

## ***Qual a população e a unidade estatística dessa pesquisa?***

## ***Qual é a sua amostra?***

## ***Qual é a variável nessa pesquisa?Ela é discreta ou contínua?***

## ***Com base na massa dos funcionários da empresa elabore um quadro de distribuição de frequências absolutas e relativas.***

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **CLASSES** | **CONTAGEM** | **FI** | **FA** | **FR** | **FRA** | **XI** |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |

9). Complete a tabela abaixo :

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **...CLASSE.** |  | **.....ffa.fa . fa** | **....frfr.fr.... fr** | **%**  **%xi5xi5%%** |
| 50 |-------- 54 |  | 4 |  |  |
| 54 |-------- 58 |  |  | 0,225 |  |
| 58 |-------- 62 |  | 11 |  |  |
| 62 |-------- 66 |  |  | 0,200 |  |
| 66 |-------- 70 |  | 5 |  |  |
| 70 |-------- 74 |  |  | 0,075 |  |
| Total |  | 40 |  |  |
|  |  |  |  |  |

10) Segue abaixo os dados da variável taxa de mortalidade infantil de 34 municípios:

http://s3.amazonaws.com/magoo/ABAAABTdYAI-13.jpg

Obtenha uma distribuição de frequências com 7 classes, começando do valor 0 (incluso) e com amplitudes de classe iguais a 10. Apresente alguns comentários sobre a taxa de mortalidade infantil dos 34 municípios.

11)Em uma pesquisa foram anotados os tempos decorridos entre a incidência de uma certa doença e sua cura, em 50 pacientes. Estes tempos são os seguintes, em horas:



Construa um histograma e comente sobre alguns aspectos relevantes desta distribuição

**MEDIDAS DE POSIÇÃO**

**Introdução**

* São as estatísticas que representam uma série de dados orientando-nos quanto à posição da distribuição em relação ao eixo horizontal do gráfico da curva de freqüência.
* As medidas de posições mais importantes são as **medidas de tendência central ou promédias** (verifica-se uma tendência dos dados observados a se agruparem em torno dos valores centrais).
* As medidas de tendência central mais utilizadas são**: média aritmética**, **moda** e **mediana**. Outros promédios menos usados são as médias: **geométrica**, **harmônica**, **quadrática, cúbica e biquadrática**.

**1) MÉDIA**

**MÉDIA ARITMÉTICA =** XX

🡺 É igual ao quociente entre a soma dos valores do conjunto e o número total dos valores.

...... XXX

onde **xi** são os valores da variável e **n** o número de valores.

.***Dados não-agrupados:*** Quando desejamos conhecer a média dos dados não-agrupados em tabelas de freqüências, determinamos a **média aritmética simples**.

*Ex: Sabendo-se que a venda diária de arroz tipo A, durante uma semana, foi de 10, 14, 13, 15, 16, 18 e 12 kilos, temos, para venda média diária na semana de:*

***Dados agrupados:***

**Sem intervalos de classe 🡺** Consideremos a distribuição relativa a 34 famílias de quatro filhos, tomando para variável o número de filhos do sexo masculino. Calcularemos a quantidade média de meninos por família:

|  |  |
| --- | --- |
| **Nº de meninos** | **freqüência = fi** |
| 0 | 2 |
| 1 | 6 |
| 2 | 10 |
| 3 | 12 |
| 4 | 4 |
| **total** | **34** |

* Como as freqüências são números indicadores da intensidade de cada valor da variável, elas funcionam como fatores de ponderação, o que nos leva a calcular a **média aritmética ponderada**, dada pela fórmula:

XXXX

|  |  |  |
| --- | --- | --- |
| **..xi.** | **..fi.** | **..xi.fi .** |
| 0 | 2 |  |
| 1 | 6 |  |
| 2 | 10 |  |
| 3 | 12 |  |
| 4 | 4 |  |
| **total** | **34** |  |

**Com intervalos de classe 🡺** Neste caso, convencionamos que todos os valores incluídos em um determinado intervalo de classe coincidem com o seu ponto médio, e determinamos a média aritmética ponderada por meio da fórmula:

XXXX..

onde **Xi** é o ponto médio da classe.

*Ex: Calcular a estatura média de bebês conforme a tabela abaixo.*

|  |  |  |  |
| --- | --- | --- | --- |
| Estaturas (cm) | freqüência = **fi** |  |  |
| 50 |------------ 54 | 4 |  |  |
| 54 |------------ 58 | 9 |  |  |
| 58 |------------ 62 | 11 |  |  |
| 62 |------------ 66 | 8 |  |  |
| 66 |------------ 70 | 5 |  |  |
| 70 |------------ 74 | 3 |  |  |
| Total | 40 |  |  |

**2) MEDIANA ( Md )**

Pelas suas propriedades e características é um dos mais importantes valores médios utilizados em Estatística. É também chamado de valor mediano, separatriz. A mediana de uma série de valores ordenados num rol crescente ou decrescente é o valor que divide esse rol em dois subconjuntos com o mesmo número de elementos .

Segundo a forma na qual os dados estão disponíveis, seus estimadores podem assumir diferentes caracterizações, assim :

1. Se os dados estão dispostos num :

a) ROL IMPAR: A mediana, é valor que ocupa a posição central do rol, ou seja, é o elemento que está na posição dada por: i = n + 1

2

1. ROL PAR: A mediana é igual a média dos valores dos elementos que

ocupam as seguintes posições :

i = n i = n + 1

1. 2

2) Se os dados estão em uma *DISTRIBUIÇÃO DE FREQUÊNCIAS:*

Se os dados estão agrupados numa distribuição de frequências , a

identificação da mediana é feita através de :

Md = Li + [ Em - Fa ant ] . C

fi

Onde :

Li limite inferior da classe mediana

Fa ant frequência acumulada crescente da classe anterior a classe mediana

fi frequência absoluta simples da classe mediana

Em amplitude da classe mediana

Para localizar a classe mediana , usa- se o elemento mediano.

*Ex: Calcule a mediana da série { 1, 3, 0, 0, 2, 4, 1, 3, 5, 6 }*

**A mediana em dados agrupados 🡺**

**a) Sem intervalos de classe:** Neste caso, é o bastante identificar a freqüência acumulada imediatamente superior à metade da soma das freqüências. A **mediana** será aquele valor da variável que corresponde a tal freqüência acumulada.

*Ex.: conforme tabela abaixo:*

|  |  |  |
| --- | --- | --- |
| **Variável xi** | **Freqüência fi** | **Freqüência acumulada** |
| 0 | 2 | 2 |
| 1 | 6 | 8 |
| 2 | 9 | 17 |
| 3 | 13 | 30 |
| 4 | 5 | 35 |
| **total** | **35** |  |

* uando o somatório das freqüências for **ímpar** o valor mediano será o termo de ordem dado pela fórmula :

|  |
| --- |
| .MED1 |

* Quando o somatório das freqüências for **par** o valor mediano será o termo de ordem dado pela fórmula:

|  |
| --- |
| MED2 |

*Ex: Calcule Mediaa da tabela abaixo:*

|  |  |  |
| --- | --- | --- |
| **Variável xi** | **Freqüência fi** | **Freqüência acumulada** |
| 12 | **1** |  |
| 14 | **2** |  |
| 15 | **1** |  |
| 16 | **2** |  |
| 17 | **1** |  |
| 20 | **1** |  |
| **total** | **8** |  |

**Emprego da Mediana**

* Quando desejamos obter o ponto que divide a distribuição em duas partes iguais.
* Quando há valores extremos que afetam de maneira acentuada a média aritmética.
* Quando a variável em estudo é salário.

**3 )MODA ( m o )**

A moda de uma série de **n**  valores é o valor que se repete maior número de vezes .

Para identificarmos um valor modal, devemos verificar se os dados estão ordenados :

a) EM UM ROL

A identificação da moda é facilitada pela simples observação do elemento que apresenta maior frequência .

3.1) EM DISTRIBUIÇÃO DE FREQUÊNCIAS

**a) Sem intervalos de classe:** Uma vez agrupados os dados, é possível determinar imediatamente a moda: **basta fixar o valor da variável de maior freqüência.**

*Ex: Qual a temperatura mais comum medida no mês abaixo:*

|  |  |
| --- | --- |
| **Temperaturas** | **Freqüência** |
| 0º C | 3 |
| 1º C | 9 |
| 2º C | 12 |
| 3º C | 6 |

*Resp:* ***2º C*** *é a temperatura modal, pois é a de maior freqüência*.

.**b) Com intervalos de classe:**

*Ex: Calcule a estatura modal conforme a tabela abaixo.*

|  |  |
| --- | --- |
| **Classes (em cm)** | **Freqüência** |
| 54 |------------ 58 | 9 |
| 58 |------------ 62 | 11 |
| 62 |------------ 66 | 8 |
| 66 |------------ 70 | 5 |

**pela fórmula de CZUBER**: **Mo = Li + (d1/(d1+d2)) x C**

**Li =** limite inferior da classe modal..... e..... = limite superior da classe modal

**d1** = freqüência da classe modal - freqüência da **classe anterior** à da classe modal

**d2** = freqüência da classe modal - freqüência da **classe posterior** à da classe modal

**C** = amplitude da classe modal

**Obs:**  A **moda** é utilizada quando desejamos obter uma medida rápida e aproximada de posição ou quando a medida de posição deva ser o valor mais típico da distribuição. Já a **média aritmética** é a medida de posição que possui a maior estabilidade.

# EXERCÍCIOS

1)Os dados abaixo referem-se às porcentagens de aprovação, por parte das populações de 10 cidades ,de certo projeto governamental.

**15% 12% 15% 8% 86% 13% 13% 83% 11% 13%**

a)Em média, qual é a porcentagem da população favorável ao projeto? E a mediana e moda. b)ELIMINE as duas observações discrepantes e calcule novamente a média.

2) As classes A, B,C tiveram, respectivamente, as seguintes médias na prova de estatística: 6.5, 6,0 e 7,0. Sabendo que a classe A é formada por 28 alunos, B é formada por 25 alunos e C, por 22 alunos, calcule a nota média de 75 alunos.

3) Uma faculdade está conduzindo uma campanha por telefone para levantar fundo para a construção de um Centro de Artes. Os dados a seguir representam as quantias (em mil reais) prometidas por todos os alunos que receberam ligações durante as primeiras nove noites da campanha.

**16, 18, 11, 17, 13, 10, 22, 15, 16.**

Calcule a média, mediana e moda em reais de fundos conseguido nas ligações.

4)As idades dos jogadores de dois times de futebol são:

Time A: 16, 15, 18, 15, 16, 16, 17, 18, 19, 17, 16.

Time B: 15, 17, 19, 19, 17, 18, 19, 18, 18, 17, 16.

Responda:

a)qual o time que apresenta maior idade média?

b)Qual a idade que mais se repete em cada time?(MODA)

c)Qual a idade média de todos os jogadores em campo?

d)Qual a mediana de todos os jogadores?

5)Os dados abaixo referem-se ao número de cigarros que os funcionários de uma fumam por dia.

0 5 6 0 2 0 8 0 15 0 0 0 0 8 0 0 6 0 5 5 4 0 0 5 0 0 6

a) calcule a média , a moda e a mediana desses dados.

b) calcule a média , a moda e a mediana se retirarmos os não fumantes do estudo.

6)Um comerciante atacadista vende determinado produto em sacas que deveriam conter 16,50 kg .A pesagem de 40 sacas revelou os resultados apresentados na tabela abaixo .

|  |  |
| --- | --- |
| PESOS Nº DE SACAS |  |
| 14,55 15,05 01 |  |
| 15,05 15,55 03 |  |
| 15,55 16,05 08 |  |
| 16,05 16,55 09 |  |
| 16,55 17,05 10 |  |
| 17,05 17,55 06 |  |
| 17,55 18,05 03 |  |

Pede-se:

1. média da distribuição
2. mediana
3. moda

7)Num teste aplicado a 20 alunos, obteve-se a seguinte distribuição de pontos:

|  |  |
| --- | --- |
|  | **Nº de Alunos** |
| **30[----40**  **40[----50**  **50[----60**  **60[----70**  **70[----80**  **80[----90** | 1  3  8  3  3  2 |

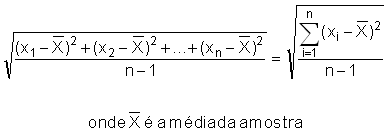
 Pede-se:a)média da distribuição

b)mediana

c) moda

**Desvio padrão**

O **desvio padrão** é uma medida de dispersão usada com a média. Mede a variabilidade dos valores à volta da média. O valor mínimo do desvio padrão é 0 indicando que não há variabilidade, ou seja, que todos os valores são iguais à média.

A fórmula de cálculo do desvio padrão para os valores x1, x2, x3,…, xn de uma amostra é a seguinte:

 Por convenção, usa-se a letra grega (sigma) para o desvio padrão da população e **s** para o desvio padrão da amostra

NOTA: Por razões matemáticas que não estão no âmbito deste manual, no caso do cálculo do desvio padrão da população deve-se usar como quociente da fórmula n em vez de n-1, e a média da população em vez da média da amostra.

**Exemplo:** Consideremos os seguintes dados:

|  |  |  |  |
| --- | --- | --- | --- |
| **Nome** | **Idade** | **Nome** | **Idade** |
| Paula | 22 | Gonçalo | 22 |
| Manuel | 24 | Pedro | 20 |
| Carla | 26 | Cristina | 24 |
| Maria | 23 | Sofia | 28 |
| João | 21 | Susana | 30 |

A média das idades é:

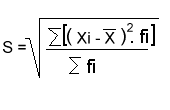
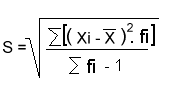
(22+24+26+23+21+22+20+24+28+30) /10 = **24 anos.**

O desvio padrão é:

dpadrao_2

A variância tende a ser um número grande e de difícil manejo e o seu valor sai dos limites dos valores observados em um conjunto de dados. Portanto, o desvio padrão, que é a raiz quadrada da variância, pode ser usado para descrever a quantidade de dispersão na distribuição da freqüência.

* Quando os dados estão agrupados (temos a presença de freqüências) a fórmula do desvio padrão ficará :

  ou  quando se trata de uma amostra

*Ex: Calcule o desvio padrão populacional da tabela abaixo:*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Xi** | **f i** |  |  |  |  |  |
| 0 | 2 |  |  |  |  |  |
| 1 | 6 |  |  |  |  |  |
| 2 | 12 |  |  |  |  |  |
| 3 | 7 |  |  |  |  |  |
| 4 | 3 |  |  |  |  |  |
|  |  |  |  |  |  |  |
| Total | 30 |  |  |  |  |  |

***Obs****: Nas tabelas de freqüências* ***com intervalos de classe*** *a fórmula a ser utilizada é a mesma do exemplo anterior.*

**VARIÂNCIA - S2**

🡺 É o desvio padrão elevado ao quadrado. A variância é uma medida que tem pouca utilidade como estatística descritiva, porém é extremamente importante na inferência estatística e em combinações de amostras.

**MEDIDAS DE DISPERSÃO RELATIVA**

**Coeficiente de Variação de Pearson - CVP**

* Na estatística descritiva o **desvio padrão** por si só tem grandes limitações. Assim, um desvio padrão de 2 unidades pode ser consderado pequeno para uma série de valores cujo valor médio é 200; no entanto, se a média for igual a 20, o mesmo não pode ser dito.
* Além disso, o fato de o **desvio padrão ser expresso na mesma unidade dos dados** limita o seu emprego quando desejamos comparar duas ou mais séries de valores, relativamente à sua dispersão ou variabilidade, quando expressas em unidades diferentes.
* Para contornar essas dificuldades e limitações, podemos caracterizar a dispersão ou variabilidade dos dados em termos relativos a seu valor médio, medida essa denominada de **CVP: Coeficiente de Variação de Pearson (é a razão entre o desvio padRão e a média referentes a dados de uma mesma série).**

**CVP = (S / XX ) x 100**

o **resultado** neste caso é **expresso em percentual**, entretanto pode ser expresso também através de um fator decimal, desprezando assim o valor 100 da fórmula.

*Ex: Tomemos os resultados das estaturas e dos pesos de um mesmo grupo de indivíduos:*

|  |  |  |
| --- | --- | --- |
| Discriminação | M É D I A | DESVIO PADRÃO |
| ESTATURAS | 175 cm | 5,0 cm |
| PESOS | 68 kg | 2,0 kg |

- Qual das medidas (Estatura ou Peso) possui maior homogeneidade ?

**Resposta:** Teremos que calcular o **CVP** da **Estatura** e o **CVP** do **Peso**. O **resultado menor** será o de **maior homogeneidade ( menor dispersão ou variabilidade).**

CVP estatura = ( 5 / 175 ) x 100 = 2,85 %

CVP peso = ( 2 / 68 ) x 100 = 2,94 %.

Logo, nesse grupo de indivíduos, as estaturas apresentam menor grau de dispersão que os pesos.

1. Em cinco tentativas, uma pessoa levou **12 18 14 11 15** minutos para trocar o óleo de determinada marca de automóvel. Calcule o tempo médio da troca de óleo, o desvio padrão, a variância e o coeficiente de variação dessa amostra.
2. Os seguintes valores representam a venda diária de Segunda à Sexta de jornais em uma banca:**231 228 244 240 236**
3. Calcule o desvio padrão
4. Coeficiente de variação
5. Dá-se a seguir o número de erros cometidos por 200 estudantes submetidos a um teste de múltipla escolha na língua inglesa:

|  |  |
| --- | --- |
| **Número de erros** | **Número de alunos** |
| 6-10 | 12 |
| 11-15 | 73 |
| 16-20 | 52 |
| 21-25 | 39 |
| 26-30 | 24 |
| Total | 200 |

1. Determine a média
2. Determine o desvio padrão
3. Determine a variância

d)Coeficiente de variação

OUTROS EXERCÍCIOS

 1) Calcule o desvio padrão da distribuição:  4,45

|  |  |
| --- | --- |
| **CLASSES** | 2 ι— 6 ι— 10 ι— 14 ι— 18 ι— 22 |
| **fi** | 5     12      21     15      7 |

2) Em um exame final de Matemática, o grau médio de um grupo de 150 alunos foi 7,8 e o desvio padrão, 0,80. Em Estatística, entretanto, o grau médio final foi 7,3 e o desvio padrão, 0,76. Em que disciplina foi maior a dispersão?  Estatística

3) Medidas as estaturas de 1.017 indivíduos, obtivemos X = 162,2 cm e s = 8,01 cm. O peso médio desses mesmos indivíduos é 52 kg, com um desvio padrão de 2,3 kg. Esses indivíduos apresentam maior variabilidade em estatura ou em peso?  estatura

4) Obtenha o desvio padrão de cada um dos jogadores A e B, de basquete, em relação aos pontos por partida, conforme a tabela abaixo:

|  |  |
| --- | --- |
| **A** | 26   32   28   30   27   31 |
| **B** | 15   45   19   42   31   22 |

 5) Um grupo de 85 moças tem estatura média de 160,6 cm, com um desvio padrão igual 5,97 cm. Outro grupo de 125 moças tem uma estatura média de 161,9 cm, sendo o desvio padrão igual a 6,01 cm. Qual é o coeficiente de variação de cada um dos grupos? Qual o grupo mais homogêneo?

       3,72% e 3,71%, respectivamente; o segundo grupo

6) Numa competição de salto triplo, três atletas disputavam apenas uma vaga para uma olimpíada entre faculdades de uma cidade. Cada atleta fez  4 tentativas obtendo os seguintes resultados:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Atleta I**  **Atleta II**  **Atleta III** | 16,50 m  13,90 m  15,70 m | 15,81 m  17,01 m  16,02 m | 16,42 m  16,82 m  16,95 m | 16,12 m  15,10 m  17,00 m |

      a) Qual deles obteve melhor média?

      b) Qual deles foi o mais regular nessas quatro tentativas?

7) A tabela a seguir mostra o número de acertos numa prova com 10 questões aplicadas numa turma com 50 alunos.

|  |  |
| --- | --- |
| **Nº da questão** | 1     2     3     4     5     6     7     8     9     10 |
| **Quantidade de acertos** | 15    20   12    25   48   40    35   10    30    40 |

     Obtenha:

     a) a média de acertos por questão;

     b) o desvio padrão dessa distribuição.

   8) Sabendo que um conjunto de dados apresenta para média aritmética e para desvio padrão, respectivamente, 18,3 e 1,47, calcule o coeficiente de variação

9)As companhias de seguro pesquisam continuamente as idades na morte e as causas de morte. Os dados se baseiam no estudo levado a efeito da revista OLHA sobre as pessoas que morreram vitimadas por armas de fogo durante uma semana.

|  |  |
| --- | --- |
| IDADE NA MORTE | FREQUÊNCIA |
| 15I 25  25I 35  35I 45  45I 55  55I 65  65I 75  75I 85 | 22  10  6  2  4  5  1 |

Determinar o desvio padrão da causa dessa morte.

10)Num teste aplicado a 40 alunos, obteve-se a seguinte distribuição de pontos:

|  |  |
| --- | --- |
| **Pontos** | **Nº de Alunos(fi)** |
| **10 [----20**  **20 [---30**  **30 [---40**  **40 [----50**  **50 [----60**  **60 [----70** | **3**  **8**  **12**  **6**  **7**  **4** |

1. calcular a variância amostral
2. determinar o desvio padrão
3. calcular o coeficiente de variação (explique)

Introdução à Probabilidade

O estudo da **probabilidade** vem da necessidade de em certas situações, prevermos a possibilidade de ocorrência de determinados fatos.

Ao começarmos o estudo da probabilidade, normalmente a primeira ideia que nos vem à mente é a da sua utilização em jogos, mas podemos utilizá-lo em muitas outras áreas. Um bom exemplo é na área comercial, onde um site de comércio eletrônico pode dela se utilizar, para prever a possibilidade de fraude por parte de um possível comprador.

Para iniciarmos o estudo da probabilidade, vamos a seguir definir alguns conceitos importantes sobre a matéria.

Objetivo: Definir um modelo matemático probabilístico que seja conveniente a descrição e interpretação de fenômenos aleatórios.

4.1 Introdução

Ao jogarmos uma moeda para o ar, de modo geral, não podemos afirmar se ocorrerá cara ou coroa. Da mesma forma, quando lançamos um dado não sabemos qual das faces 1, 2, 3, 4, 5, ou 6 ocorrerá. No campo dos negócios e do governo há numerosos exemplos de tais situações. Por exemplo, a incerteza existe quando desejamos realizar uma previsão sobre a procura de um novo produto, a opinião pública em relação a determinado assunto, o sucesso de um novo plano econômico, etc - tudo isso contém algum elemento de acaso.

Na Estatística, a incerteza existe quando se quer fazer alguma afirmação a respeito de alguma característica populacional baseada em informações extraídas de dados amostrais. Neste caso, a aplicação da Teoria das Probabilidades é de fundamental importância para a solução de problemas de Inferência Estatística.

Independente de qual seja a aplicação em particular, a utilização das probabilidades indica que existe um elemento de acaso, ou de incerteza, quanto à ocorrência ou não de um evento futuro. Assim é que em muitos casos, pode ser impossível afirmar com antecipação o que ocorrerá. No entanto, é possível dizer o que pode ocorrer.

O ponto central em todas essas situações é a possibilidade de quantificar quão provável é determinado evento.

Em suma, podemos dizer que, as probabilidades são utilizadas para exprimir a chance de ocorrência de determinado evento.

4.2 Definições Básicas

Definição 4.1 (Experimentos Aleatórios ou Fenômenos Aleatórios). são aque- les onde o processo de experimentação está sujeito a influências de fatores casuais e conduz a resultados incertos.Notação: E.

Exemplos:

E1 : Jogar uma moeda e observar a face superior. E2 : Lançar um dado e observar o número da face superior.

E3: Uma lâmpada é fabricada. Em seguida é testada a duração da vida útil dessa lâmpada.

Observações:

a) Cada experimento poderá ser repetido um grande número de vezes sob as mesmas condições; b) Não podemos afirmar que resultado particular ocorrerá, porém podemos descrever o conjunto de todos os possíveis resultados do experimento; c) Quando o experimento é repetido um grande número de vezes, surgirá uma regularidade nos resultados. Esta regularidade, chamada de regularidade estatística, é que torna possível construir um modelo matemático preciso com o qual se analisará o experimento.

Definição 4.2 (Espaço Amostral). é o conjunto de todos os possíveis resultados de um experimento aleatório.

Notação: S ou Ω.

Exemplos:

Os espaços amostrais associados aos experimentos aleatórios E1, E2 e E3 são: S1 = S2 = S3 =

Definição 4.3 (Evento). Dado um espaço amostral S associado a um experimento E, definimos como evento, qualquer subconjunto desse espaço amostral, ou seja, é qualquer coleção de resultados do experimento E.

Notação: A, B, C, D, etc.

Exemplos:

1 - Considerando o espaço amostral S2, exemplos de eventos seriam: A: Ocorre face par =B: Ocorre um número menor que 4 = C: Ocorre um número maior que 0 = D: Ocorre o número 10 =

2 - Considerando o espaço amostral S3, um exemplo de evento seria: A: A vida útil de uma lâmpada é menor que 10 horas =

## Evento

Quando lançamos um dado ou uma moeda, chamamos a ocorrência deste fato de **evento**. Qualquer subconjunto de um espaço amostral é um evento.

Em relação ao **espaço amostral** do lançamento de um dado, veja o conjunto a seguir:

**A = { 2, 3, 5 }**

Note que http://www.matematicadidatica.com.br/MEx.ashx?QVxxdWFkXHN1YnNldFxxdWFkIFM= ( A está contido em S, A é um subconjunto de S ). O conjunto **A** é a representação do evento do lançamento de um dado, quando temos a face para cima igual a um número primo.

Classificação de Eventos

Podemos classificar os eventos por vários tipos. Vejamos alguns deles:

### Evento Simples

Classificamos assim os eventos que são formados por um único elemento do espaço amostral.

**A = { 5 }** é a representação de um **evento simples** do lançamento de um dado cuja face para cima é divisível por **5**. Nenhuma das outras possibilidades são [divisíveis por 5](http://www.matematicadidatica.com.br/CriteriosDeDivisibilidade.aspx).

### Evento Certo

Ao lançarmos um dado é certo que a face que ficará para cima, terá um [número divisor](http://www.matematicadidatica.com.br/Divisores.aspx) de **720**. Este é um **evento certo**, pois **720 = 6! = 6 . 5 . 4 . 3 . 2 . 1**, obviamente qualquer um dos números da face de um dado é um divisor de**720**, pois **720** é o produto de todos eles.

O conjunto **A = { 2, 3, 5, 6, 4, 1 }** representa um evento certo pois ele possui todos os elementos do espaço amostral**S = { 1, 2, 3, 4, 5, 6 }**.

Evento Impossível

No lançamento conjunto de dois dados qual é a possibilidade de a soma dos números contidos nas duas faces para cima, ser igual a **15**?

Este é um **evento impossível**, pois o valor máximo que podemos obter é igual a doze. Podemos representá-lo por http://www.matematicadidatica.com.br/MEx.ashx?QVxxdWFkPVxxdWFkXGVtcHR5c2V0, ou ainda por **A = {}**.

### Evento União

Seja **A = { 1, 3 }** o evento de ocorrência da face superior no lançamento de um dado, ímpar e menor ou igual a **3** e **B = { 3, 5 }**, o evento de ocorrência da face superior, ímpar e maior ou igual a **3**, então **C = { 1, 3, 5 }** representa o evento de ocorrência da face superior ímpar, que é a união dos conjuntos **A** e **B**, ou seja, http://www.matematicadidatica.com.br/MEx.ashx?Q1xxdWFkPVxxdWFkIEFccXVhZFxiaWdjdXBccXVhZCBC.

Note que o evento **C** contém todos os elementos de **A** e **B**.

### Evento Intersecção

Seja **A = { 2, 4 }** o evento de ocorrência da face superior no lançamento de um dado, par e menor ou igual a **4** e **B = { 4, 6 }**, o evento de ocorrência da face superior, par e maior ou igual a **4**, então **C = { 4 }** representa o evento de ocorrência da face superior par, que é a intersecção dos conjuntos **A** e **B**, ou seja, http://www.matematicadidatica.com.br/MEx.ashx?Q1xxdWFkPVxxdWFkIEFccXVhZFxiaWdjYXBccXVhZCBC.

Veja que o evento **C** contém apenas os elementos comuns a **A** e **B**.

Eventos Mutuamente exclusivos

Seja **A = { 1, 2, 3, 6 }** o evento de ocorrência da face superior no lançamento de um dado, um número divisor de **6** e**B = { 5 }**, o evento de ocorrência da face superior, um divisor de **5**, os eventos **A** e **B** são **mutuamente exclusivos**, pois http://www.matematicadidatica.com.br/MEx.ashx?QVxxdWFkXGJpZ2NhcFxxdWFkIEJccXVhZD1ccXVhZFxlbXB0eXNldA==, isto é, os eventos não possuem elementos em comum.

Evento Complementar

Seja **A = { 1, 3, 5 }** o evento de ocorrência da face superior no lançamento de um dado, um número ímpar, o seu**evento complementar** é **A = { 2, 4, 6 }** o evento de ocorrência da face superior no lançamento de um dado, um número par.

Os elementos de **A** são todos os elementos do espaço amostral **S** que não estão contidos em **A**, então temos que**A = S - A** e ainda que **S = A + A**.

## Probabilidade de Ocorrência de um Evento

Os três irmãos **Pedro**, **João** e **Luís** foram brincar na rua. Supondo-se que as condições de retorno para casa são as mesmas para cada um deles, qual é a probabilidade de **Luís** voltar para casa primeiro?

Como **3** é o número total de irmãos, então **Luís** tem **1** chance em **3** de voltar para casa primeiro, por isto a**probabilidade** de **Luís** voltar para casa antes dos seus irmãos é igual a **1/3**.

Definição

A **probabilidade de um evento ocorrer** (Luís voltar para casa primeiro) considerando-se um **espaço amostral** (Pedro, João e Luís) é igual a [razão](http://www.matematicadidatica.com.br/razao.aspx) do **número de elementos do evento** (1, apenas Luís) para o **número de elementos do espaço amostral** (3, o número de irmãos que foram brincar na rua), desde que espaço o amostral seja um **conjunto equiprovável**, ou seja, todos os seus elementos tenham a mesma possibilidade de ocorrer (as condições de retorno para casa são as mesmas para os três irmãos).

Sendo **E** um evento, **n(E)** o seu número de elementos, **S** o espaço amostral não vazio e **n(S)** a quantidade de elementos do mesmo, temos que a probabilidade de **E** ocorrer é igual a:

http://www.matematicadidatica.com.br/MEx.ashx?UChFKVxxdWFkPVxxdWFkXGZyYWN7bihFKX17bihTKX0=, sendo **n(S)≠0**.

A probabilidade é um número entre zero e um, inclusive, o que significa que no mínimo não a nenhuma hipótese do evento acontecer e no máximo o evento sempre ocorrerá:

**0 ≤ P(E) ≤ 1**

Normalmente representamos probabilidades através de frações, mas também podemos representá-las por números decimais, ou até mesmo por porcentagens.

## Exemplos

**EnunciadoUm dado é lançado. Qual é a probabilidade de obtermos um número divisor de 6?**

Como vimos acima, o **espaço amostral** do lançamento de um dado é:

**S = { 1, 2, 3, 4, 5, 6 }**

Como estamos interessados apenas nos resultados [divisores](http://www.matematicadidatica.com.br/Divisores.aspx) de **6**, o evento **E** é representado por:

**E = { 1, 2, 3, 6 }**

Então **n(E) = 4** e **n(S) = 6**, portanto:

http://www.matematicadidatica.com.br/MEx.ashx?UChFKVxxdWFkPVxxdWFkXGZyYWN7bihFKX17bihTKX1ccXF1YWRcUmlnaHRhcnJvd1xxcXVhZCBQKEUpXHF1YWQ9XHF1YWRcZnJhY3s0fXs2fVxxcXVhZFxSaWdodGFycm93XHFxdWFkIFAoRSlccXVhZD1ccXVhZFxmcmFjezJ9ezN9

Podemos também apresentar o resultado na forma de uma [porcentagem](http://www.matematicadidatica.com.br/Porcentagem.aspx):

http://www.matematicadidatica.com.br/MEx.ashx?UChFKVxxdWFkPVxxdWFkXGZyYWN7Mn17M31ccXF1YWRcUmlnaHRhcnJvd1xxcXVhZCBQKEUpXHF1YWQ9XHF1YWRcZnJhY3syfXszfVxxdWFkXGNkb3RccXVhZDEwMCVccXF1YWRcUmlnaHRhcnJvd1xxcXVhZCBQKEUpXHF1YWQ9XHF1YWQ2Niw2NyU=

**RespostaA probabilidade de se obter um número divisor de 6 é 2/3 ou 66,67%.**

**EnunciadoUma moeda é lançada 4 vezes. Qual é a probabilidade de obtermos ao menos uma coroa?**

Recorrendo ao [princípio fundamental da contagem](http://www.matematicadidatica.com.br/PrincipioFundamentalContagem.aspx) podemos calcular o número de elementos do espaço amostral deste exemplo:

**n(S) = 2 . 2 . 2 . 2 = 16**

Agora precisamos saber o número de elementos do evento **E**, referente a quatro lançamentos de uma moeda, quando obtemos ao menos uma **coroa**.

Lembra-se do **evento complementar** explicado acima? Sabendo quantos são os resultados que não apresentam nenhuma **coroa**, ele nos permite descobrir o número dos que possuem ao menos uma.

E quantos são os eventos que não possuem nenhuma coroa? Apenas o evento **E = { cara, cara, cara, cara }**, ou seja, apenas **1**. Como o número total de eventos é **16** e **1** deles não apresenta qualquer coroa, então os outros **15**apresentam ao menos uma. Então:

http://www.matematicadidatica.com.br/MEx.ashx?UChFKVxxdWFkPVxxdWFkXGZyYWN7bihFKX17bihTKX1ccXF1YWRcUmlnaHRhcnJvd1xxcXVhZCBQKEUpXHF1YWQ9XHF1YWRcZnJhY3sxNX17MTZ9

Na forma de [porcentagem](http://www.matematicadidatica.com.br/Porcentagem.aspx) temos:

* **A probabilidade de obtermos ao menos uma coroa é 15/16, 0,9375 ou 93,75%.**

**EXERCÍCIOS**

1) No lançamento simultâneo de 2 dados, considere as faces voltadas para cima e determine a) espaço amostral S. b) evento E1 : números cuja soma á igual a 5.

c) evento E2: números iguais. d) evento E3: números cuja soma é um número par.

e) evento E4: números ímpares nos 2 dados. f) evento E5: número 2 em pelo menos 1 dos dados.

g) evento E6: números cuja soma é menor que 12. h) evento E7: números cuja soma é maior que 12. i) evento E8: números divisores de 7 nos 2 dados.

2) Um casal planeja ter 3 filhos. Determine os eventos:

a) os 3 são do sexo feminino. b) pelo menos 1 é do sexo masculino. c) os 3 do mesmo sexo.

3)Qual a probabilidade de ocorrer o número 5 no lançamento de um dado?

4)Qual a probabilidade de se obter um número par no lançamento de um dado?

5)Um disco tem uma face branca e a outra azul. Se o disco for lançado 3 vezes, qual a

probabilidade de a face azul ser sorteada pelo menos uma vez?

6) (Unesp) João lança um dado sem que Antônio veja. João diz que o número mostrado pelo dado é par. Qual a probabilidade de Antônio descobrir esse número?

7) (Funesp) Um baralho de 12 cartas tem 4 ases. Retiram-se 2 cartas, uma após a outra. Determine a probabilidade de a segunda ser um ás, sabendo que a primeira é um ás.

8) (UFSCar-SP) Uma urna tem 10 bolas idênticas, numeradas de 1 a 10. Se retirarmos uma bola da urna, qual a probabilidade de não obtermos a bola número 7 ?

9) Uma urna contém 2 bolas brancas e 5 bolas vermelhas. Retirando-se 2 bolas ao acaso e sem reposição, calcule a probabilidade de:

a) as bolas serem de cores diferentes.

b) as bolas serem vermelhas.

10)(Mauá-SP) Uma caixa contém 11 bolas numeradas de 1 a 11. Retirando-se uma delas ao acaso, observa-se que ela tem um número ímpar. Determine a probabilidade de esse número ser menor que 5.

11) Uma bola é retirada de um urna que contém bolas coloridas. Sabe-se que a probabilidade de ter sido retirada uma bola vermelha é 5/17. Calcule a probabilidade de ter sido retirada uma bola que não seja vermelha.

12)(Fuvest) A probabilidade de que a população atual de um país seja de 110 milhões ou mais é de 95%. A probabilidade de ser 110 milhões ou menos é de 8%. Calcule a probabilidade de ser 110 milhões.

13) Consultadas 500 pessoas sobre as emissoras de tevê que habitualmente assistem, obteve-se o seguinte resultado: 280 pessoas assistem ao canal A, 250 assistem ao canal B e 70 assistem a outros canais, distintos de A e B. Escolhida uma pessoa ao acaso, determine a probabilidade de que ela assista:

a) ao canal A. b) ao canal B. c) ao canal A ou ao canal B

14) (PUCCAMP-SP) Num grupo, 50 pessoas pertencem a um clube A, 70 pertencem a um clube B, 30 a um clube C, 20 pertencem aos clubes A e B, 22 aos clubes A e C, 18 aos clubes B e C e 10 pertencem aos 3 clubes. Escolhida ao acaso uma das pessoas presentes, a probabilidade de ela:

a) pertencer aos 3 clubes é 3/5. b) pertencer somente ao clube C é zero. c) pertencer a pelo menos dois clubes é de 60%. d) não pertencer ao clube B é 40%.

15)De uma reunião participam 200 profissionais, sendo 60 médicos, 50 dentistas, 32 enfermeiras e os demais nutricionistas. Escolhido ao acaso um elemento do grupo, qual é a probabilidade de ele ser médico ou dentista?

16) Escolhido ao acaso um elemento do conjunto dos divisores de 30, determinar a probabilidade de que ele seja primo?

17)Ao lançarmos dois dados não viciados, qual a probabilidade de obtermos faces voltadas para cima onde a soma entre elas seja 6?

18)No lançamento de uma moeda e um dado, determine a probabilidade de obtermos o resultado dado por (coroa, 1).