**Conceitos Fundamentais**

1.1- **Estatística**: É a ciência que estuda as técnicas necessárias para coletar, resumir, organizar, apresentar, analisar e interpretar os dados de um experimento aleatório, a fim de extrair informações sobre a população.

Se divide em duas partes:

* **Estatística Descritiva**: Se preocupa com a coleta, resumo, organização e apresentação dos dados.
* **Estatística Indutiva ou Inferencial**: Se preocupa com a análise e interpretação dos dados, com a finalidade de obter informações que possam ser generalizadas para a população.

**- Aleatório:** ao acaso, todos os dados têm a mesma chance de ocorrência.

**- População:** é o conjunto de **todos** os indivíduos, seres, objetos que satisfaçam uma ou mais características que se pretende estudar.

**-** **Amostra:** é qualquer subconjunto (parte) extraído da população.

Características ou variáveis:

1.2- **Variável**: Uma variável corresponde a uma característica de um objeto ou de um indivíduo.

1.2.1- **Tipos de variável**

* **Variável Qualitativa ou Categórica**: É uma variável que não pode ser medida por um número.

Ex: Nacionalidade; Grupo sanguíneo; Naturalidade; Estado Civil; Escolaridade; Raça, religião, Mão de obra; ...

**Variável Qualitativa Nominal**: A característica é dada pelo nome.

Ex: Cor dos Olhos; Nacionalidade; Raça, religião, naturalidade, mão de obra; ..

**Variável Qualitativa Ordinal**: Quando existe certa ordem p/ a característica.

Ex: Escolaridade; Estado Civil; patente militar; Plano de Carreira, cargos religiosos; Certificação de Habilitação; ...

* **Variável Quantitativa:** É uma variável que pode ser medida numericamente.

**Ex:** Temperatura; Idade; Massa; Imc; Altura; Nº de filhos; Natalidade; Velocidade, Produtividade; Salário; PIB; Taxa de juros; Número de indivíduos contaminados; número de mortes; número de acidentes; deslocamento ou comprimento;...

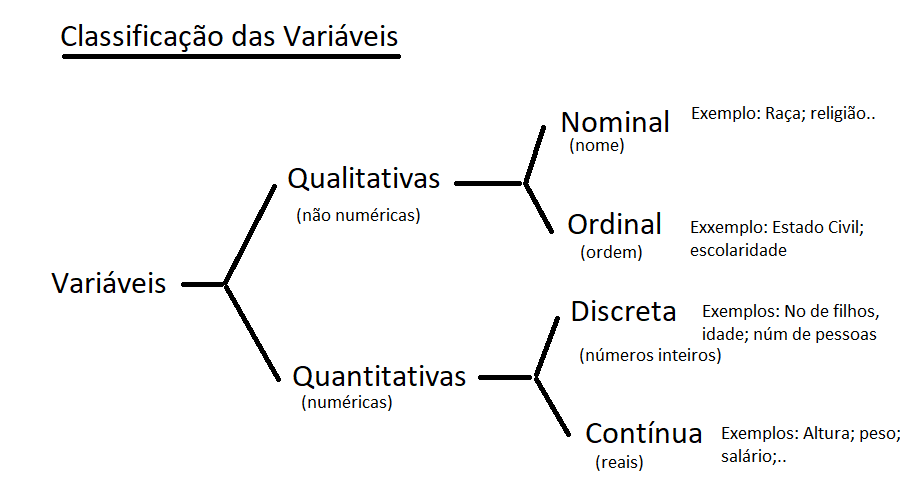
**Variável Quantitativa Discreta:** Expressa somente por números inteiros ou naturais.

**Ex:** Idade; Nº de filhos; Nº de carros vendidos, número de celulares, número de pessoas infectadas, número de barris de petróleo produzidos.

**Variável Quantitativa Continua:** Expressa por nº reais (podem ser não inteiros).

**Ex:** Temperatura; PIB; Salário; IMC; Altura; Massa; Taxa de juros; Impostos;

Taxa de natalidade**,** ...

****

**1.4 – Métodos de estudo e coleta de dados**

* Censo: Coleta de dados da população

Características:

- É Lento

- É Caro

- Confiabilidade 100%

Para que realizar um censo?

Para obter dados confiáveis, com vistas à orientação e planejamento de políticas públicas.

* Estimação ou Amostragem: É o levantamento de informações a partir de uma amostra.

Características:

- É Rápido

- É mais barato

- Não é 100% de confiabilidade.

Obs – Parâmetro: É qualquer dado numérico obtido em toda a população através do estudo de um fenômeno aleatório.

* Estimador: Dado numérico obtido em uma amostra.

Exe:

* Eleição p/ governador

População: 20 milhões de eleitores

Votação:

Candidato A -> 40% dos votos => **Parâmetro**

Pesquisa: Amostra: 1000 Pessoas

Candidato A -> 38 % das intenções de voto => **Estimador**

1.5 Métodos de coleta de dados Primários

A análise de dados geralmente envolve a manipulação ou o trabalho com os dados que obtidos em uma pesquisa, por esse motivo é importante garantir que a coleta de dados seja executada de maneira correta, cão contrário os dados podem ser tendenciosos.

* **Questionários**
* **Pontos Positivos**
* São enviados, geralmente pelo correio ou e-mail, para uma amostra previamente selecionada.
* Seleciona-se a região (Geográfica) em que o questionário será aplicado.
* Relativamente barato.
* **Problemas**
* Não ocorre devolução de muitos questionários
* Falta de esclarecimento sobre as perguntas.
* Submeter a uma comissão de ética, para avaliação das questões e do procedimento que será utilizado para obtenção dos dados.
* **Entrevistas Pessoais**
* **Pontos Positivos**
* Selecionar a região, o grupo de interesse, enfim a amostra a ser estudada. Elaborar todos os passos da entrevista.
* Interação do pesquisador com os entrevistados
* **Problemas**
* Treinamento da equipe entrevistadora
* Perguntas ambíguas
* Estimulação de resposta
* Submissão ao comitê de ética

**Ricardo Feltrin**

04/03/2021 10h17

Uma pergunta que muita gente se faz —e faz a mim também— é:

Como a Kantar Ibope decide quais os domicílios no Brasil que podem ter instalada uma caixinha de medição de audiência de TV?

Não é possível detalhar aqui neste texto toda a metodologia, mas tudo é baseado em dados do Censo do IBGE, do PNAD (Pesquisa Nacional de Amostra por Domicílio) e mais um monte de estudos, estatísticas e pesquisas da própria Kantar.

O objetivo dessa complexa metodologia é tentar colocar e representar todos os estratos da sociedade (social, econômica e culturalmente). Pobres, ricos, classe média, gente com faculdade e doutorado, ou só com primeiro grau ou nem isso. Todos precisam estar representados nesse pequeno universo mensurado 24 horas por dia.

Quero ser medido pelo Ibope. Posso?

Não, não pode, meu filho.

Para começar, você não pode pedir ou se inscrever para ter um aparelho do Ibope em casa. -

Como disse, a escolha é absolutamente impessoal e baseada em dados sólidos de pesquisas, e não em voluntariado. O segundo ponto é que a empresa não paga valores em dinheiro (exceto cerca de R$ 50 para compensar o gasto de energia elétrica) para que essas pessoas/ famílias aceitem ser monitoradas Sim, as famílias eventualmente ganham alguns presentes e regalos (como eletrodomésticos, por exemplo), mas não é algo trocado por dinheiro. É algo quase "altruísta", pois é trabalhoso, como verão a seguir. Isso faz da medição algo muito mais confiável. De certa forma, esses presentinhos e benefícios são uma forma de compensar um tormento: afinal, essas pessoas estão abrindo mão de sua privacidade ao indicar 24 horas o que estão consumindo em TV (e agora na internet e no streaming também).... -

Milhares de famílias monitoradas

No Brasil são 6.600 residências nas 15 maiores regiões metropolitanas que têm o aparelho instalado. Uma única residência pode ter um, dois , três ou até mais aparelhinhos instalados. Depende do tamanho da família. O endereço e as famílias onde os "peoplemeters" —em breve, em boa parte delas, o focal meter— estão instaladas são informações sigilosas.

As famílias também assinam um contrato e se comprometem a não divulgar (a vizinhos ou jornalistas) que têm o aparelho instalado.

Em 24 anos cobrindo TV, eu só descobri uma única residência. E foi por mero acaso. O aparelho liga junto com a TV e os moradores devem "informá-lo" (com seus botões) quem está no ambiente assistindo.

Se o morador receber uma visita, por exemplo, de um parente, ele deve "cadastrar" o visitante também, informando idade e sexo.

Na hora que alguém vai dormir, deve se "deslogar" do aparelho".

Ah, mais uma: se a pessoa ligar a TV (e consequentemente o aparelho) e não se "logar", se ela não informar quem está assistindo, o equipamento passa a fazer um barulho chatíssimo (apitar) até que o login seja feito.

É possível corromper a família? Uma "desconfiança" que muitos leitores têm e já me perguntaram é: "E se uma TV descobre onde tem uma família e vai lá e suborna essa família para só assistir a essa TV?" Bom, primeiro lugar, segundo a Kantar, isso jamais aconteceu. Mas, suponhamos que acontecesse. Acontece que os softwares da empresa monitoram não só os dados presentes de cada domicílio, mas também os do passado. Ou seja, se porventura (gosto dessa palavra) uma residência/família mudar de comportamento, mudar seu consumo televisivo de forma radical, esse "ponto" instalado ficará sob escrutínio muito maior por parte dos analistas da Kantar. Eventualmente ele pode até ser descredenciado e substituído por outro. O que não é incomum.

A Kantar eventualmente também pode trocar os domicílios sem aviso prévio. Está no contrato e ela pode fazer isso. Muita gente também acha que 6.600 aparelhos é um número pequeno para uma população de mais de 200 milhões de pessoas. Acontece que, estatisticamente, esse número é mais que suficiente para o estudo.

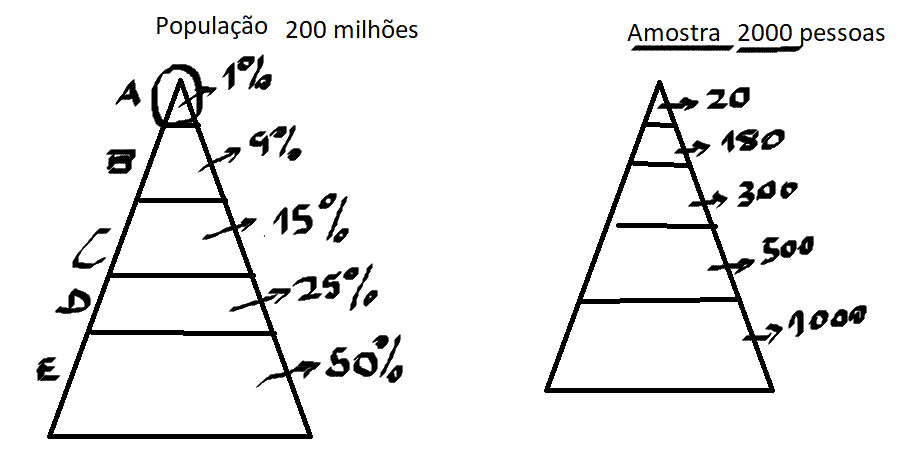
Veja mais em https://www.uol.com.br/splash/noticias/ooops/2021/03/04/como-sao-escolhidas-familias-que-terao-aparelhinhos-de-ibope.htm?cmpid=copiaecola

**1.6 Tipos de Amostragens**

* **Amostragem Aleatória Simples**
* A chave para este tipo de amostragem é que cada elemento da população de ter a mesma chance.
* Esse processo envolve normalmente um sorteio, ou seja, você atribui números a população, e sorteia uma amostra.
* Essa forma o processo fica não tendencioso, pois se for tendencioso (Preferência) deixa de ser um processo aleatório.
* Processo mais usado.
* O sorteio pode ser feito através do computador, pela geração de números aleatórios (randomizar).
* **Amostragem Aleatória Estratificada**
* **Às vezes, a população é constituída de subpopulações ou estratos.**

**Ex: Camadas Sociais.**

* **Nesse caso para que a amostra seja representativa, ela deve representar a mesma estratificação do universo em estudo.**
* **Procedimento:**
* **Quais são os estratos presentes na população.**
* **Calcular seus tamanhos relativos (Proporções).**
* **Determinar o tamanho da amostra mantendo as mesmas proporções.**
* **Obter aleatoriamente os elementos de cada estrato.**

****

* **Amostragem Aleatória Sistemática**
* **Se os elementos da população estiverem ordenados, por exemplo em listas, arquivos ou outros.**
* **Processo:**
* **Você escolhe uma constante previamente.**
* **Sorteia-se o primeiro indivíduo.**
* **Evitam-se tantos indivíduos quantos forem indicados pela constante e seleciona-se o próximo.**
* **Repete-se o processo até obter o tamanho Amostral desejado.**
* **Amostragem Aleatória por conglomerados (clusters)**
* **Se a população se encontra dividida em grupos ou conglomerados (frequentemente no espaço), então podemos selecionar diretamente no grupo.**

**2.1 – Distribuições de Frequências – Variável Quantitativa Discreta**

Def.

**Frequência(f)**: Número de vezes que um mesmo dado aparece em uma pesquisa aleatória

Uma distribuição de frequência é uma representação tabular de um conjunto de dados.

Ex – Durante 30 dias foram registrados o nº de acidentes diários em uma rodovia do estado de SP.

Os dados obtidos são:

Dados Brutos:

0;2;2;1;1;0;0;3;2;3;

1;0;1;2;1;2;3;2;2;0;

4;1;3;5;5;2;1;1;0;1;

Representação:

**Fórmulas:**

* **Frequência Relativa (fi rel)**

**Fórmula:**

**n =**

* **Frequência Acumulada (Fi )**
* **Frequência Acumulada Relativa (Firel)**
* Firel = 

**Significados:**

* fI = 9 => Tivemos no período de 30 dias, 9 dias em que ocorreu apenas 1 acidente. ( Na maioria dos dias ocorre apenas 1 acidente)
* fi rel => 20% dos dias observados não houveram acidente.
* Fi => 23 Dos 30 dias observados, em 23 deles ocorreram 2 acidentes ou menos.
* Fi rel => 90% dos dias analisados ocorreram 3 acidentes ou menos.
* **2.2 - Distribuição de frequências – Variável Quantitativa Contínua**

**2.2.1 – Construção da Variável:**

Ex: Notas dos 30 Alunos de uma turma na disciplina de cálculo

Notas: 0,5 ; 1,4 ; 6,3 ; 5,0 ; 7,0 ; 8,0 ; 9,5 ; 9,8 ; 4,0 ; 3,2;

0,2 ; 2,5 ; 3,8 ; 6,0 ; 5,5 ; 6,2 ; 7,5 ; 7,3 ; 8,1 ; 5,8;

3,0 ; 4,6 ; 5,3 ; 6,5 ; 6,5 ; 7,2 ; 4,0 ; 1,0 ; 2,0 ; 5,6;

Contagem: cont.ses(Intervalo;”critério1”;Intervalo;”critério2”)

=CONT.SES(A43:A72;">=2";A43:A72;"<4")

1º Passo: Calcule a Amplitude total (AT );

AT = Xmáx – Xmin

Ou seja: AT = 9,8 – 0,2 = 9,6

2º Passo: Número de classes ou intervalos (k):

- Critério da Raiz:

k =  onde n é o tamanho da população ou amostra. Obs: K E IN\*

No exemplo: k =   5,47 => K= 5 intervalos

Se n =10.000 🡪 K =

- Fórmula de Sturges:

K = 1 + 3,22\*log (n) ; onde n é o tamanho da amostra.

No exemplo:

K = 1 + 3,22\*log (30) ~ 5,756 -> k = 6 intervalos

Se n = 10.000 🡪 k =

3º Passo: Amplitude do Intervalo de Classe (h):

h =

Ex: h = 1,92 => h = 2

4º Passo: Intervalo de Classe:

É designado por dois números reais, sendo o primeiro chamado de limite inferior (I) e o segundo de limite Superior (L)

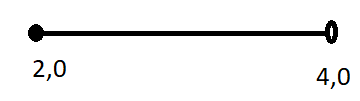
H = I – L

Obs: Nomenclatura: |---- Fechado à esquerda e aberto à direita.

---- Aberto em ambos os lados.

Exemplo: 2 |----- 4 🡪 Significa que o intervalo inicia no número 2,0 e vai até 3,999

Intervalo fechado em 2,0 e aberto em 4,0



5º Passo: Resumo

Notas fi fi rel Fi Fi REL

0 |---- 2 4 13,33% 4 13,33%

2 |---- 4 5 16,67% 9 30%

4 |---- 6 8 26,67% 16 56,67%

6 |---- 8 9 30% 26 86,66%

8 |---- 10 4 13,34% 30 100%

Total 30 100%

**Histograma**

É um conjunto de retângulos justapostos.

**Diagrama de Pareto:** Aula por vídeo: Link: https://www.youtube.com/watch?v=VMEhNx0jzOs

**3. Medidas de Tendência Central**

Ex: Médias; mediana e moda.

**3.1 – Somatório – Notação Sigma** ()

Para expressarmos uma soma, de forma sintética, usamos a notação de somatório.

Ex:=

+

1+2+3+4+5+6 =

* 1+3+5+7+9+11 =

Se i =1 🡪 2i -1 = 2.1 – 1 = 1

Se i =2 🡪 2i -1 = 2.2 -1 = 3

Se i =3 🡪 2.i – 1 = 2.3 -1 = 5

* 2+4+6+8+10+12+14 =

**3.2 – Médias**

**3.2.1 – Média Aritmética Simples (** **)**

Seja a sequência:

Xi: X1 ; X2 ; X3 ;....; Xn

; onde n é o número de elementos

;

**3.2.2 – Média Aritmética Ponderada ( )**

Seja uma sequência de valores

Xi : X1 ; X2 ; X3 ; ... ; Xn com respectivas frequências fi : f1 ; f2 ; f3 ; ... ; fn ; então:

Ou seja:

**Exemplo 1**

**Notas x Pesos (fi)**

**1ºbim 3,0 | 1**

**2ºbim 5,0 | 2**

**3ºbim 4,0 | 2**

**4ºbim 7,0 | 3**\_\_\_\_



Exemplo 2: Suponha um processo seletivo: Contratar uma pessoa na área de TI

Os candidatos precisaram passar por algumas atividades:

|  |  |  |
| --- | --- | --- |
| Atividades | Notas (0 a 10) | fi |
| Aval Conhecimentos na área | 7 | 4 |
| Dinâmica de grupo | 8 | 2 |
| Conhecimentos gerais | 9 | 2 |
| Entrevista pessoal | 5 | 2 |

**3.2.3 Média Geométrica Simples (**g **)**

Seja a sequência xi : x1 ; x2 ; x3 ;...; xn

Então g = =

Ou

g = ( produtório)

Ex:

* Notas: 4 ; 6 ; 8 ; 2

g =   4,426727679..... Média geométrica

 =  Média aritmética

Observações:

1. ( a média geométrica é sempre menor ou igual a média aritmética)
2. A média geométrica, em geral, possui muitas casas decimais, possuindo, portanto, maior precisão.

Produtor de Leite:

João compra leite de dois produtores:

Jan Fev Mar Abr

Produtor A: 50 L 60 L 40 L 50 L

Produtor B: 80 L 50 L 40 L 30 L

Média de produção nestes meses?

Média Aritmética:

Conclusão: Observando as médias, os produtos têm valores iguais, logo é indiferente.

Média Geométrica:

Conclusão: O Produtor A possui uma média maior, logo é mais regular, em termos da produção de leite. Devemos escolher ele como fornecedor.!

Como fazer, na calculadora: 50\*60\*40\*50 = 6.000.000^(1/4) = 49,49232004..

Quais as diferenças entre média aritmética e geométrica?

Note que:

1. g ≤  ( a média geométrica é sempre menor ou igual a média aritmética)
2. A média geométrica, em geral, possui maior precisão, isto é, um grande número de casas após a virgula;
3. Se um valor for nulo (zero), a média geométrica será nula.
4. A média geométrica permite observar a regularidade, enquanto a média aritmética despreza qualquer irregularidade.

* Nota concurso: 0 ; 8 ; 10 ; 10

g = 0 => Eliminado concurso

 =  = 7

* Notas:

Aluno A: 4 ; 2 ; 8 ; 6

Aluno B: 5 ; 5 ; 6 ; 4

Qual o aluno mais regular?

1º: Média Aritmética:

A = 5,0

B = 5,0 empate!

2º: Média geométrica:

gA = 4,426727679

gB = 4,9492321

Resp: O aluno B é mais regular do que A.

**3.2.4 Média Geométrica Ponderada (****g)**

Seja Xi: X1 ; X2 ; X3 ; ... ; Xn com as

Respectivas frequências (pesos) fi: f1 ; f2 ; f3 ; ... ; fn

Então: g =

Sendo n =

Exemplo

* Sejam as notas:

|  |  |  |
| --- | --- | --- |
|  | Nota | Peso (fi) |
| 1º bim | 4,0 | 1 |
| 2º bim | 5,0 | 2 |
| 3º bim | 3,0 | 3 |
| 4º bim | 7,0 | 4 |

n =

g = ~ 4,7992745...

Interpretação p/ média:

* É o ponto de equilíbrio do conjunto de dados.
* É o valor esperado p/ a distribuição de dados.
* É o valor em torno do qual os dados se distribuem.

**Média Móvel Simples**

<https://www.youtube.com/watch?v=iYjLOBfeIqc>

Em estatística a **média móvel** simples, ou MMS, ou simplesmente **média móvel**, é um recurso utilizado para se identificar a tendência de um conjunto de dados dispostos em uma série temporal.

**3.3 Mediana (md)**

É o valor que divide o conjunto de dados em duas partes iguais, isto é, 50% dos dados são menores ou iguais a mediana e 50% dos dados são maiores ou iguais a ela.

**3.3.1 Váriavel Quantitativa Discreta**

1º Caso: n Ímpar

Idades

Ex : 1º passo: colocar em Rol (ordem crescente ou decrescente)

Xi : 18 ; 21 ; 22 ; 24 ; 26 ; 30 ; 35 ; 38 ; 39 ; 40 ; 42

n = 11 elem. 6º elemento é a mediada => md = 30 anos

então md =

2º Caso: n é Par

Idades

Ex: Xi : 18 ; 21 ; 22 ; 24 ; 26 ; 30 ; 35 ; 38 ; 39 ; 40

n = 10 elem. A mediana está entre o 5º e o 6º elemento md ==

então md = ; para n par.

**3.3.2 Variável Quantitativa Contínua**

Ex:

|  |  |  |  |
| --- | --- | --- | --- |
| Tabela1 - Faixa Salarial dos 60 funcionários da fundição da empresa A | | | |
| Classes | Faixa Salarial | fi | Fi |
| 1 | 500 |- 1000 | 15 | 15 |
| 2 | 1000 |- 1500 | 20 | 35 |
| 3 | 1500 |- 2000 | 10 | 45 |
| 4 | 2000 |- 2500 | 8 | 53 |
| 5 | 2500 |- 3000 | 7 | 60 |
|  | Soma (n) = | 60 |  |

Fórmula da Mediana:

*med* =

Onde:

*Limd* = Limite inferior da classe mediana

n = (nº total de elementos )

Fant = freq acumulada da classe anterior à classe mediana

fmd = freq simples da classe mediana

h = amplitude do intervalo de classe.

Resp:

1º passo:  <-> 30º elemento

2º passo:

med =

Interpretação: 50% dos funcionários recebem salários menores ou iguais à R$ 1.375,00 e 50% recebem salários maiores ou iguais ao valor.

**3.4 Moda (mo)**

É o valor de maior frequência.

**3.4.1 Váriavel Discreta**

Ex: 1

|  |  |
| --- | --- |
| Tabela – 1 – Notas dos Alunos | |
| Notas | fi |
| 2 | 5 |
| 3 | 8 |
| 5 | 10 |
| 6 | 4 |
| 9 | 1 |

Mo = 5;

2. Seq: 0 ; 1 ; 2 ; 2,5 ; 3 ; 4 ; 4,1 ; 4,2 ; 10 ; 10

mo = 10 ; média = 4,53

3.

|  |  |
| --- | --- |
| Tabela – 2 – Nº de filhos dos alunos | |
| #Filhos | fi |
| 0 | 10 |
| 1 | 5 |
| 2 | 10 |
| 3 | 6 |
| 4 | 4 |
| 5 | 1 |

Mo = 0 ou mo = 2 filhos. A moda é Bimodal.

4.

|  |  |
| --- | --- |
| Tabela – 3 | |
| # Variável | fi |
| 2 | 4 |
| 3 | 4 |
| 5 | 4 |
| 6 | 4 |

mo = não existe (Amodal)

**3.4.2 Variável Continua**

Existem diversos processos para o cálculo da moda, embora o mais usado seja a moda de Czuber. (Fórmula de Czuber)

ou

**Legenda:**

Limo: Limite inferior do intervalo modal

fmo: freq. do intervalo modal

fant: freq anterior ao intervalo modal

h: Amplitude do intervalo de classe

|  |  |
| --- | --- |
| Tabela 1: Faixa salarial dos funcionários da emp. A | |
| **Salário (R$)** | fi |
| 500,00 |- 800,00 | 15 |
| 800,00 |- 1100,00 | 30 |
| 1100 | - 1400 | 10 |
| 1400 |- 1700 | 5 |
| 1700 |- 2000 | 2 |
| Total | 62 |

Qual o valor da moda?

1º. Passo: **O intervalo modal:** 800 à 1100 reais

2º. Passo: Utilizar a Fórmula de Czuber para a moda:

Mo = 800 +

**Interpretação: O salário mais frequente nesta empresa é de R$ 928,57.**

**4. Medidas Separatrizes**

Note que:

D5 = Q2 = P50 = md (mediana

Q1 = P25 ; 13 = P75

D9 = P 90

Intervalos inter-quartílicos

Q2-Q1 = 25%

Q3-Q1 = 50%

**Cálculo do Percentil**

P50 = md = 

Generalizando:

Pi = 

**Aplicação**

Uma pesquisa de mercado revelou que a distribuição de salários para uma função de TI segue a distribuição.

|  |  |  |
| --- | --- | --- |
| Tabela: Distr. de Salários de TI em 20 Empresas de Pesq | | |
| Faixa Salarial (R$) | fi | Fi |
| 2000 |- 2500 | 5 | 5 |
| 2500 |- 3000 | 8 | 13 |
| 3000 |- 3500 | 12 | 25 |
| 3500 |- 4000 | 10 | 35 |
| 4000 |- 4500 | 8 | 43 |
| 4500 |- 5000 | 7 | 50 |
| Total (n)= | 50 |  |

Qual deve ser o salário do profissional de TI, de tal forma que ele receba acima de 90% da sua categoria? 🡪 P90 = ?

Resolução

1º. Passo: Calcular 90% de total(n): 90%\*50 = 45º salário

2º. Passo: Em qual classe (intervalo) está o 45º. Salário? O 45º. Salário está na última faixa salarial.

3º. Passo: Utilizar a fórmula:

Interpretação: 90% dos funcionários de TI recebem R$ 4.642,86 ou menos e 10% dos funcionários dessa área recebem este valor ou mais.

Calcule P75%

1º. Passo: 75% \* 50 = 37,5

2º. Passo: P75% =

**5 Medidas de Dispersão**

**5.1 Desvio Médio Simples (DMS)**

Toda medida possui uma incerteza (erro). Neste caso, convém determinarmos com essas medidas ( valores ) se dispersam em relação à média.

Exemplo

* Medidas referentes ao diâmetro de um fio obtidas com um paquímetro

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Tab 1.: d(mm) do fio | | | | | | |
| d (mm) | 2,03 | 1,98 | 2,00 | 2,05 | 2,10 |  |

1º Passo: (Média) =  = 2,033 mm = 2,03 mm

2º Passo: Desvios:

D1 = 2,03 – 2,03 = 0  
 D2 = 1,98 – 2,03 = 0,05 ( em Módulo)  
 D3 = 2,00 – 2,03 = 0,03  
 D4 = 2,05 – 2,03 = 0,02  
 D5 = 2,10 – 2,03 = 0,07

Desvio Médio Simples

DMS =  mm

**5.2 Variância Populacional (σ²)**

Sejam os dados

X1 ; X2 ; X3 ; ... ; Xn

Logo

Var(σ2) = ( Erro médio quadrático)

**5.2.1 Desvio padrão Populacional (σ)**

O desvio padrão é igual a raiz quadrada da variância

σ = ( medida de dispersão dos dados)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Tab 1.: 0(mm) do fio | | | | | | |
| 0 (mm) | 2,03 | 1,98 | 2,00 | 2,05 | 2,10 |  |

S² = sqrt(0,001736)

= 0 ,04

OBS: Se os dados possuírem uma distribuição de frequências, teremos:

Significado:

|\_\_\_\_\_\_\_\_\_\_\_\_|\_\_\_\_\_\_\_\_\_\_\_|

Variância Amostral ( S2):

Xi:x1;x2;x3;...;xn com fi:f1;f2;f3;...;fn

Variância (S²) = ; n =

Desvio Padrão Amostral (S)

S = ;

Interpretação: Desvio Padrão é uma medida da dispersão média dos dados em relação à

média.

Ou, uma medida da incerteza ou erro em relação ao valor esperado.

Exemplo 2: Calcule a média o desvio padrão populacional da notas de alunos de uma turma

|  |
| --- |
| Tabela 1: |
|  |

**5.4 Coeficiente de Variação**

Quando desejamos comparar as dispersões de dois conjuntos de dados, o desvio padrão não nos permite avaliar corretamente qual das dispersões é maior ou menor, pois fornece apenas números absolutos. Neste caso, o coeficiente de variação pode nos fornecer uma comparação mais efetiva.

**5.4.1 População:**

CV pop = (σ/µ) \* 100%

**5.4.2 Amostra**

|  |  |
| --- | --- |
| **Amostra** | |
| **Tabela 1: Relação entre Peso e estatura** | |
| **Peso(kg)** | **Estatura (cm)** |
| **69** | **180** |
| **57** | **172** |
| **74** | **183** |
| **57** | **168** |
| **85** | **188** |
| **103** | **194** |
| **68** | **182** |

**1º Passo**

X’ p   
S’ p  

**2º Estatura**

X’ p   
S’ p  

Como podemos observar, a variável peso possui maior dispersão do que a variável é o peso!

**5.5** **Aplicações do desvio padrão**

No argumento que levou à definição do desvio padrão, notamos que se a dispersão for pequena, significa que os dados estão concentrados em torno da média, e se a dispersão for grande estes dados se distanciam muito da média.  
Esta ideia é expressa mais formalmente pelo matemático russo P.L. Tchebychev

Teorema de Tchebychev

Para qualquer conjunto de dados (População ou amostra) e qualquer constante *K* >1, *K* E IN\*, a proporção de dados a *K* desvios padrões à direita e a esquerda da média é dada por no mínimo:

Exemplo: k = 2 desvios padrões, então:

🡪 ao menos 75%

Significa que:

No mínimo 75% dos dados estão a dois desvios padrões à direita e à esquerda da média.

Note: μ = média

\_\_\_\_\_\_|\_\_\_\_\_\_\_│\_\_\_\_\_\_\_\_|\_\_\_\_\_\_\_\_\_|\_\_\_\_\_\_\_\_\_|\_\_\_\_\_\_\_\_

μ – 2.DP μ – 1.DP μ μ+1.DP μ + 2.DP

|\_\_\_\_\_\_\_\_\_\_\_\_\_no mínimo 75%\_\_\_\_\_\_\_\_\_\_\_|

Se k = 3 desvios padrões, teremos:

Significa que, no mínimo 88,9% de todos os dados estão entre a média e três desvios padrões à direita e à esquerda dela.

Note: μ = média

\_\_|\_\_\_\_\_\_|\_\_\_\_\_\_\_│\_\_\_\_\_\_\_\_|\_\_\_\_\_\_\_\_\_|\_\_\_\_\_\_\_\_\_|\_\_\_\_\_\_\_|\_

μ – 3.DP μ – 2.DP μ – 1.DP μ μ+1.DP μ + 2.DP μ + 3.DP

|\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_no mínimo 89%\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_|

Exemplo Aplicado:

* Se todas as latas de 1 libra (lb) de café enchidas por um processador de alimentos têm peso médio de 16,00 onces (oz), com desvio-padrão de 0,02 onces, qual a porcentagem mínima das latas que devem conter entre 15,80 e 16,20 oz de café?

Obs: 1 lb(pound) ~ 453,6 g   
 1 oz(once) ~ 28 g

Note: 1 oz-troy = 31,10 g ( para negociar ouro/metais preciosos)

* Resp: 1º. passo: μ = 16,00 oz.  
   S = 0,02 oz.  
     
   2º. passo: \_\_\_\_\_|\_\_\_\_\_\_\_\_\_\_|\_\_\_\_\_\_\_\_\_\_|\_\_\_\_\_

15,80oz 16,00oz 16,20oz

16,00 – 15,80 = 0,20 oz ou 16,20 – 16,00 = 0,20 oz

3º. Passo: *K*. 0,02 oz = 0,20 oz 🡪 k = = 10 desvios padrões.

4º. Passo: Pelo teorema de Tchebychev

Então:

Conclusão: No mínimo 99% de todas as latas de 1 lb de café enchidas por este processador tem peso entre 15,80 e 16,20 oz.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Ex 2: Um estudo do valor nutritivo de certo tipo de pão mostra que, em média, uma fatia contém 0,260 mg | | | | | |
| de tiamina(vitamina B1) com desvio padrão de 0,005 mg. Com base no teorema de Tchebychev | | | | |  |
| entre quais valores deve estar o conteúdo de tiamina de | |  |  |  |  |
| a-) ao menos 35/36 (97,2%) de todas as fatias desse pão; |  |  |  |  |  |
| b-) ao menos 80/81 (~ 98,8%) de todas as fatias desse pão? |  |  |  |  |  |

Resolução:

1. Dados: Média (μ) = 0,260 mg e Desvio Padrão(σ) = 0,005 mg

1º. Passo: Pelo Teorema de Tchebychev, temos:

Logo k = 6 desvios padrões

2º. Passo: |\_\_\_\_\_\_\_\_\_\_|\_\_\_\_\_\_\_\_\_\_\_|

μ - 6σ μ μ + 6σ

logo: μ - 6σ = 0,260 mg – 6.(0,005 mg) = 0,230 mg

μ + 6σ = 0,260 mg + 6.(0,005 mg) = 0,290 mg

**Conclusão:**

**Ao menos 35/36 (97,2%) de todas as fatias desse pão, possuem conteúdo de tiamina entre 0,230 mg e 0,290 mg.**

1. 1º. Passo: Pelo Teorema de Tchebychev, temos:

2º. Passo: |\_\_\_\_\_\_\_\_\_\_|\_\_\_\_\_\_\_\_\_\_\_|

μ - 9σ μ μ + 9σ

logo: μ - 9σ = 0,260 mg – 9.(0,005 mg) = 0,215 mg

μ + 9σ = 0,260 mg + 9.(0,005 mg) = 0,305 mg

**Conclusão:**

**Ao menos 80/81 (99%) de todas as fatias desse pão, possuem conteúdo de tiamina entre 0,215 mg e 0,305 mg.**

**7. Probabilidades**

**7.1 Contagem – Análise combinatória.**

**Existem diversos tipos de agrupamentos.**

**7.1.1 Arranjo Simples (An,p)**

**São agrupamentos que diferem pela natureza dos elementos e pela ordem que estão dispostos.**

**Fórmula:**

**Exs:**

1. **Suponha que tenhamos 20 times, e queremos saber de quantos modos teremos 3 times em 1º, 2º e 3º lugar?**

1. **Queremos organizar uma equipe de trabalho, com 3 pessoas, de forma que tenhamos um supervisor, um que execute o planejamento e um que ministre curso. Dispomos de 5 pessoas no escritório. De quantas maneiras podemos ter essa equipe?**

**Resp:**

**7.1.2 Permutação Simples (Pn)**

**É um arranjo simples de todos os elementos**

**Fórmula: , obs: 0! = 1**

**Ex: Quantas permutações ou anagramas tem a palavra amor?**

**Resp: Amor possui 4 letras**



**7.1.3 Combinação Simples** (

**Os agrupamentos diferem pela natureza dos seus elementos mas não diferem pela ordem.**

**Fórmula:**

**Exs:**

1. **De quantas maneiras podemos combinar 20 times de 3 em 3 ?**



1. **De quantas maneiras podemos combinar 5 pessoas de 3 a 3?**
2. **Quantas possibilidades de resultados da Megasena existem?**

A ordem não importa! 🡪 Combinação simples

**Obs: Se uma pessoa joga um cartão =>**

**Probabilidade(ganhar) =**

**Obs: A chance de cair um pedaço de lata do espaço na sua cabeça é de 1 em 25.000.000. Note, tem duas vezes mais chances de cair um pedaço de lata na sua cabeça do que você ganhar na Megasena.**

**Se você participa de um bolão. Por exemplo, vocês jogam 10 números**

**Prob (ganhar) =**

**Mas, 1 jogo custa 4,50 reais**

**Logo se você jogar 10 números, o valor do jogo será: 210 \*4,50 = 945,00 reais**

**7.2 Postulados da Probabilidade**

**1º Postulado: Seja um evento a qualquer P(a)**  **0**

**2º Postulado: Seja S um espaço amostral, que é o conjunto de todas possibilidades de eventos, então: P(s) = 1 =100%**

**3º Postulado: Se dois eventos A e B são mutuamente exclusivos, então:**

**P(AUB) = P(A) + P(B)**

**7.3 Regra da adição:**

**n(AUB) = n(A) + n(B) – n(A****B)**

**P(AUB) = P(A) + P(B) – P(A****B)** se AB = 0 (Mutuamente exclusivos)

**Ex:**

* Extraindo-se aleatoriamente uma carta de um baralho comum de 52 cartas qual é a probabilidade de se obter uma carta de paus ou uma figura (Rei, Dama, Valete)?

Resposta:

Evento: A= Carta de Paus ={ Ás; 2; 3; ....;10; valete; dama; rei} -> 13 cartas

Evento: B =Figura: { Valete; Dama; Rei(Paus); Valete;...(copas);...} -> 12 cartas

A Ո B = Paus Ո Figura = { Valete; Dama; Rei} de paus -> n(A Ո B) = 3 cartas

S = { todas as cartas do baralho} = Espaço amostral -> n (S) = 52 cartas

Então:

n (A U B) = n(A) + n(B) – n( A Ո B)

Note: Se os eventos forem mutuamente exclusivos, teremos: A Ո B = Φ

Neste caso, temos:

No exemplo:

P(paus) = 

P(figura) = 

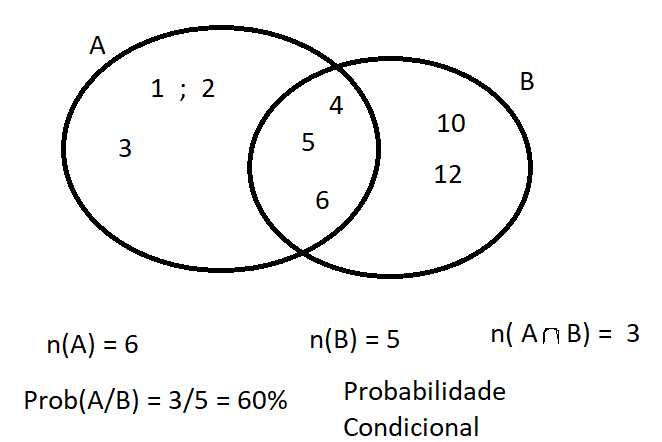
P(paus Ո figura) = 

P(paus OU figura) = 

**7.4 Probabilidade Condicional (P(A|B))**

É a probabilidade de ocorrer o evento A, tendo ocorrido o evento B

n(A U B ) = 8 elementos



Então:

Obs: P(A|B) =

Ou

P(B|A) =

logo:

= P(B|A)\*P(A)

**7.5 Regra da Multiplicação**

Note que: Se os eventos A e B são independentes, então: P (A|B) = P(A)

Exemplo de Eventos independentes

1. Qual a probabilidade de um casal ter 1 filho homem e 1 filha mulher, em duas gestações?

Evento sexo das crianças é independente!

Prob(menino) = 50% =1/2

Prob(menina) = 50% = ½

Resposta:

1º passo: h.m ou m.h





2º.passo: Portanto: P ( h e m ou m e h) =

1. Se tirarmos duas cartas de um baralho comum de 52 cartas, com reposição, qual a probabilidade de obtermos a primeira carta de paus e a segunda carta um dez?

Resp:

1ª. carta: P(paus) =

2ª.carta: P(dez) =

Os eventos são independentes, logo:

Prob (paus e dez) =

Exemplo de eventos dependentes:

* Suponha que uma organização de pesquisa junto a consumidores tenha estudado os serviços prestados dentro da garantia por 200 comerciantes de pneus em uma grande cidade, obtendo os resultados:

|  |  |  |  |
| --- | --- | --- | --- |
| *Tab.:1* | *Bom serviço dentro da garantia* | *Serviço deficiente dentro da garantia* | *Total* |
| *Vendedores de determinada marca de pneus* | *64* | *16* | *80* |
| *Vendedores de qualquer marca de pneus* | *42* | *78* | *120* |
| *Total* | *106* | *94* | *200* |

Se selecionarmos aleatoriamente um vendedor, teremos:

Seja **N** o nº de vendedores de uma marca específica de pneus.

Seja **G** o nº de vendedores que prestam bom serviços em garantia.

Seja **N’** o nº de vendedores de uma marca qualquer de pneus.

Logo:





Qual a probabilidade de selecionarmos um vendedor aleatoriamente e ele prestar bom serviço, sendo um vendedor de marca específica de pneu?

Resp:

P (G|N) =

Ou ainda,

P(G|N) =

1. Qual a probabilidade de retirarmos um Ás e um dez de um baralho comum?

Resp:

P(Ás) =

P(dez) = ( se houvesse reposição)

Sem reposição: P(dez) =

Então: P( Ás e dez) =

**8. Esperanças e Decisões**

**8.1 Esperança Matemática**

Def: Seja p(x1); p(x2);...p(xn) as probabilidades de ganhar as quantias, x1; x2; ....; xn; respectivamente, então chamamos esperança matemática à:

Pois: Média aritmética ponderada =

Exemplos:

* Qual é a nossa esperança matemática, se adquirirmos um dos 2.000 bilhetes, para um primeiro prêmio constante de Aparelho de Tv, avaliado em R$640,00, um segundo prêmio constante de um Gravador, no valor de R$120,00, e, como terceiro prêmio, um Rádio no valor de R$40,00.

Vamos comprar um bilhete: 🡪 p(x) =



Probabilidade de ganhar c/ 1 bilhete será:

P(ganhar) =

Então:

E(x) = 640\*(0,0005) + 120\*(0,0005) + 40\*(0,0005) = (640 + 120+40)\*0,0005

E(x) = 800 \* 0,0005 = R$ 0,40

Este valor representa o valor esperado de ganho, logo, se comprarmos um bilhete por um valor acima de 40 centavos, estaremos jogando o dinheiro fora, pois, esperamos ganhar menos do que pagamos.

* As probabilidades de um investidor vender uma propriedade com lucro de U$$ 2.500; de U$$ 1.500; de U$$ 500 ou com prejuízo de U$$ 500 são 0,22; 0,36; 0,28; e 0,14, respectivamente.

Qual o lucro esperado pelo investidor?

Solução:  
Note que: 

 +





...................



Lucro Esperado = Esperança matemática

Lucro Esperado =

Lucro Esperado = 550 + 540 + 140 – 70 = US$ 1,160.00

Lucro Esperado = U$$ 1,160.00

* A agência de certa companhia aérea, em certo aeroporto, tem as probabilidades 0,06; 0,21; 0,24; 0,18; 0,14; 0,10; 0,04; 0,02; 0,01 de receber 0,1,2,3,4,5,6,7 ou 8 reclamações sobre desvios de bagagens por dia. Quantas reclamações a agência espera receber por dia?

Solução:

Núm. de reclamações esperadas = 0\*0,06 + 1\*0,21 + 2\*0,24+ 3\*0,18 + 4\*0,14 + 5\*0,10+ 6\*0,04 + 7\*0,02 + 8\*0,01 =

Núm. de reclamações esperadas = 2,75 reclamações por dia.

* Suponha que em um sorteio da Mega Sena o prêmio seja de R$ 100.000.000,00, neste caso, qual a esperança matemática de ganhar este prêmio com um único jogo de seis dezenas? Suponha que o jogo custe R$ 4,50 cada, é vantajoso jogar?

Solução

1º Passo:

Note:

Qual a probabilidade de ganharmos com uma única aposta de seis números?

P (ganhar) =

2º Passo: Valor esperado de ganho?

Valor esperado =

Então note, para este prêmio, o valor esperado de ganho é menor do que o valor investido (apostado = 4,50), logo, não é lógico apostar neste caso.

3º Passo: Não é vantajoso jogar, pois o custo do jogo é maior do que espero ganhar.

Note:

Existe um valor de prêmio que vale a pena jogar?

Valor Esp. =

Ex: 8 números 🡪

Nesse caso, sua prob (ganhar) =

Mas note, você poderia jogar , e aí você fecharia a Mega, e ganharia o prêmio.

Vamos supor que na Mega da virada o prêmio seja 300.000.000,00, e aí, vale a pena?

Vale a pena se ganhar sozinho.

**8.2 Tomada de decisões**

Quando nos defrontamos com problemas que envolvem incertezas, a esperança matemática torna-se uma ferramenta interessante para a tomada de decisões. Neste caso, a idéia é encontrar a opção de maior retorno, isto é, mais promissora em termos do que se espera/deseja-se obter.

Exemplos:

* Um fabricante de mobília deve decidir se amplia agora a capacidade de sua fábrica, ou se espera até o próximo ano. Seus consultores afirmam que, se ele ampliar agora, e se as condições econômicas permanecerem favoráveis haverá um lucro de R$ 328.000,00 no próximo ano fiscal; se ele ampliar agora, e houver uma recessão, haverá um prejuízo de R$ 80.000,00; no entanto, se ele esperar o próximo ano e as condições econômicas permanecerem boas, haverá um lucro de R$ 160.000,00; se esperar o próximo ano e houver uma recessão, haverá um pequeno lucro de R$16.000,00. Se o fabricante de mobília acha que se as probabilidades de que as condições econômicas permanecerem estáveis, ou de haver uma recessão, são de 1/3, e 2/3, a ampliação de sua capacidade agora maximizará seu lucro esperado?

Solução

|  |  |  |
| --- | --- | --- |
| *Tab.:2* | *Ampliar* | *Retardar Ampliação* |
| *Condições favoráveis*  p(cf) =1/3 | *R$ 328.000,00* | *R$ 160.000,00* |
| *Recessão*  p(rec)=2/3 | *-R$ 80.000,00* | *R$ 16.000,00* |

Dado: Cond. Favoráveis (P1) = 1/3

Recessão (P2) = 2/3

1º Caso: Ampliar agora:





2º Caso: Retardar ampliação:





Conclusão:

Como o resultado financeiro do 2º caso excede o 1º caso. Decorre que o retardamento da ampliação maximizará os lucros esperados pelo fabricante.

**9.1 Distribuição Binomial**

Em muitos problemas aplicados, o que nos interesse é reconhecer a probabilidade de um evento ocorrer x vezes n tentativas.

Por exemplo:

* Probabilidade de nascer 2 filhas se o casal pretende ter 3 filhos
* Probabilidade de ocorrer 3 falhas de transmissão de dados em 10 tentativas
* Probabilidade de que 60 espectadores dentre 200, lembrem-se de quais produtos foram anunciados.

O tipo de probabilidade que estamos procurando, refere-se a obter x sucessos em n tentativas faremos as seguintes

Hipóteses 1 : N provas independentes e do mesmo tipo serão realizadas

: Cada prova (tentativa) possui dois resultados mutuamente exclusivos, que chamaremos sucesso e fracasso.

: A probabilidade de sucesso, indicada por p, será constante em cada tentativa.

Fórmula Binominal



Onde x: nº de sucessos.

P: Probabilidade de sucesso.

1-p: probabilidade de fracasso.



Logo:

Para essa distribuição temos:

* **Média** 
* **Desvio Padrão** 

Exemplos:

* Suponha que haja uma probabilidade de 0,60 de um carro ser roubado em certa cidade e ser recuperado. Encontre a probabilidade de que:  
    
  A) No máximo três dentre 10 carros furtados sejam recuperados.

B) No mínimo sete dentre 10 carros furtados sejam recuperados.

Resolução:

1º Passo:

Dados

* Prob. 

Prob. 











(b) Prob. 







Exercícios

* Com uso de computador, faça um programa para calcular a probabilidade do exercício anterior, considerando p = 0,63.
* Os registros de uma loja de vendas de computadores indicam que 70% de todos os compradores de computadores novos, exigem modem interno. Determine as probabilidades de que, entre 10 compradores de computadores novos, 0,1,2,3 ou 4 exijam o modem interno.

**9.2 Distribuição de Poisson**

A distribuição de Poisson pode ser usada para encontrar um número x de sucessos, dentro de um “continuum” de tempo ou espaço.

Exemplos:

* A chegada de certo número de chamadas telefônicas em um intervalo de tempo.
* O número de veículos que possam em uma esquina em certo tempo.

Neste caso, Supõe-se que os eventos são independentes, e o processo estacionário.



Onde:

P(x,t) é a probabilidade de obtermos x sucessos num intervalo de tempo (t).  
T tempo  
x Nº de sucessos  
 Nº médio de sucessos por unidade de tempo  
e Nº de Euler (2,7183...)

Para essa distribuição:

* Média 
* Desvio-Padrão 

Exemplos?

* Uma fábrica de pneus verificou que ao testar seus pneus nas pistas, havia em média um estouro de pneu a cada 5.000 km.
* Qual a probabilidade que um teste de 3.000 km haja no máximo um pneu estourado?
* Qual a probabilidade de que um carro ande 8.000 sem estourar nenhum pneu?

Resolução:

1º Passo:

Dados:



X = nº de estouros (“sucesso”)



* 









* 







* Se um banco recebe em média 6 cheques sem fundo por dia, qual é a probabilidade de receber 4 cheques sem cobertura em um dia qualquer?

Dados:

λ = 6 cheques/dia

t = 1 dia

x = 4 cheques

Fórmula





* O número de falhas mensais de um computador, é uma variável aleatória que segue a distribuição de Poisson, com λ = 1,8. Determine as probabilidades de que este computador funciona durante 1 mês:
* Sem qualquer falha
* Com apenas uma falha

Resolução:

* 
* 