CURSO DE GRADUAÇÃO EM AGRONOMIA
ESTATISTICA BÁSICA
SEMESTRE: 2020-1
Prof. José Lemos
AULAS EM EAD

Tentem conseguir pela Internet, além das referencias que já lhes mandei antes este livro

INTRODUÇÃO À ESTATISTICA

Rafael BISQUERRA; Jorge Castelar SARRIERA; Francesc MARTINEZ

Editora: Artmed

Aulas dos dias 23 – 24 de julho de 2020

Estatística é disciplina de apoio a todas as outras disciplinas. A ciencia não avança se o cientista não tiver uma boa formação em Ciências Estatísticas ou se ao menos não tiver como buscar apoio de profissionais que dominem esta Ciência fascinante.

Nesta Pandemia causada pelo “Virus Chinês”, ficou evidente que o desconhecimento de Estatistica por quem toma decisões, presta informações pode provocar estragos enormes.

Voces estão acompanhando as bobagens que são faladas acerca dos acontecimentos que envolvem a doença causada pelo virus.

Em Estatística, trabalhamos com variáveis que, por Definição, são:

Elementos, objetos, situações, fenômenos, que podem assumir valores ou categorias diferentes. Essas diferenças podem ser no espaço ou no tempo

Exemplos de variáveis que apresentam variações ESPACIAIS.

V1: Nomes de vocês. São 22 nomes diferentes

V.2: Sexo. Tem moças e rapazes

V3: Idade, cada um de vocês tem idades diferentes medidas em meses

V4: Alturas. Voces tem alturas diferentes que são medidas em centimetros

V5: Mas na frente cada um de voces terá uma nota em Estatística.

EXERCICIO 1:

CRIEM MAIS 10 VARIÁVEIS QUE APRESENTEM VARIAÇÕES ESPACIAIS

Exemplos de Variáveis com mudanças **TEMPORAIS**:
V1: As alturas dos vegetais perenes hoje são diferentes daquelas observadas no ano passado;
V2: As chuvas de 2020, são diferentes das que aconteceram em 2019, 2018,2017
V3: O conhecimento que voces tem hoje é diferente daqueles que tinham em 2019, 2018, 2017...
V4: A produção agrícola do Ceará varia anualmente
V5: A quntidade de automóveis que circulam nas ruas de Fortaleza é diferente daquela que acontecia em 2019, 2018,...
EXERCICIO 2:
APRESENTE DEZ (10) VARIÁVEIS COM AFERIÇÕES TEMPORAIS

Variáveis quanto à forma em que são mensuradas:

Qualitativas e Quantitativas

Variáveis qualitativas não assumem grandezas, mas categorias.

Exemplos:

Sexo

Categorias: Macho; Fêmea

Tabagismo:

Categorias: fumantes; não fumantes

Escolaridade dos agricultores

Categorias: analfabetos; fundamental incompleto; fundamental completo; nível médio...

Nesses casos, pode-se atribuir valores numéricos para as categorias, mas esses valores não tem interpretação na sua magnitude

Por exemplo, se queremos estudar se o comportamento das mulheres é diferente em relação ao dos homens no que se refere ao vicio do tabagismo:

Podemos atribuir o Peso ou escore 1 para a categoria sexo feminino e Peso ou escore 0 (zero) para a categoria sexo masculino. Neste caso não se pode dizer que 1 > 0. Apenas podemos dizer que se tratam de categorias diferentes.

Mas se estivermos tratando de escolaridade e atribuirmos os escores:

0 = Analfabetos; 1 = fundamental incompleto; 2 = fundamental completo; 3 = nivel médio completo; 4 = nível superior completo.

Neste caso as categorias podem ser ORDENADAS. Podemos escrever que

4 > 3 > 2 > 1> 0

as categorias associadas à variável podem ser ORDENADAS. Isso não é possivel fazer quando se trata de sexo, por exemplo.

EXISTE UM ARCABOUÇO TEÓRICO TODO ELE DESENHADO PARA ESTUDAR VARIÁVEIS COMO ESSAS, QUE APENAS PODEM SER COMPARADAS. CHAMA-SE ESSE ARCABOUÇO DE

ESTATISTICA NÃO PARAMÉTRICA

EXERCICIO 3:

APRESENTE 10 VARIÁVEIS QUALITATIVAS QUE PODEM SER UTILIZADAS NA AGRICULTURA.

VARIÁVEIS QUANTITATIVAS

Nestes casos as observações das variáveis são quantificáveis, podem ser ordenadas e guardam escalas de razão entre elas.

Exemplos:

V1: Idade: Uma pessoa de 40 anos de idade tem o dobro da idade de quem tem apenas 20 anos. Podemos dizer que em termos de idade

40 anos > 20 anos.

40anos = 2.(20 anos)

V2: Um agricultor que possua 50 hectares de terras terá 1/3 da área de outro que possuir 150 hectares

50 ha < 150 ha.

50ha = (1/3)150ha.

V3: Uma pessoa que tenha renda mensal de R10.000,00 terá renda que valerá dez (10) vezes alguem que tenha renda de apenas R$1000,00

R$10.000,00 > R$1000,00 e

R$1000,00 = (1/10)R$10.000,00

A NOSSA DISCIPLINA TRABALHARÁ COM VARIÁVEIS QUANTITATIVAS. ESSE ARCABOUÇO TEÓRICO É CHAMADO DE

ESTATISTICA PARAMÉTRICA.

EXERCICIO 4: ELENQUE DEZ (10) VARIÁVEIS QUANTITATIVAS APLICADAS À AGRICULTURA

Variáveis podem ser observadas de formas **DISCRETAS** e **CONTÍNUAS**

Variáveis **DISCRETAS** são observadas de forma pontual

Variáveis **CONTÍNUAS** são observadas em intervalos de classes.

Exemplos de variáveis **DISCRETAS**:

V1: Idades dos estudantes dessa turma. Vocês tem as idades medidas em anos

V2: O numero de carteiras que existem em cada uma das salas de aula do Campus do Pici;

V3: O numero de vacas leiteiras que um criador possui em um ano

V4: O total de chuvas que caiu no município de Quixeramobim entre janeiro e julho de 2020

Exemplos de
Variáveis CONTINUAS que são observadas em intervalos de classes.

V1 – População brasileira atual na faixa etária de 15 a 20 anos;

V2 – Total de agricultores cearenses que possuem áreas entre 5 e 20 hectares segundo o último Censo Agropecuário do Ceará

V3: População de fortaleza que tem renda mensal variando entre um (1) e três (3) salários mínimos em julho de 2020;

V4: população que foi a óbito no Brasil devido ao ataque do “virus chinês” que tinha idade acima de 70 anos.

EXERCICIO:

ELENCAR MAIS DEZ VARIÁVEIS CONTINUAS E DEZ VARIÁVEIS DISCRETAS. DÊEM PREFERENCIA ÀS VARIÁVEIS LIGADAS AO SETOR RURAL.

VARIÁVEIS ALEATÓRIAS

A ciência Estatística trabalha com variáveis aleatórias. Variáveis que apresentam várias possibilidades de ocorrencias e que não temos condições de definir o que acontecerá com ela num determinado experimento

O que caracteriza uma VARIÁVEL ALEATÓRIA é o fato de não sabermos o que acontecerá se buscarmos fazer uma experimentação com essa variável.

Exemplo 1:

Neste momento eu, o PROFESSOR de vocês na disciplina de Estatística Básica não sei o que acontecerá com as notas que voces terão na disciplina. Eu apenas sei que elas, as notas, variarão entre zero e dez. Eu só posso afirmar antecipadamente que os que se dedicarem mais a estudar terão CHANCE MAIOR ou PROBABILIDADE MAIOR de se saírem bem

Variáveis Aleatórias

Exemplo 2:

Quando uma mulher engravida ninguém saberá com certeza qual será o sexo da criança que ela dará à luz em nove meses. Mas todos sabem que a criança será do Sexo Masculino ou Feminino.

Exemplo 3:

Quando vamos fazer pesquisa com agricultores, se estivermos interessados em saber se a escolaridade dele pode influenciar na assimilação da tecnologia que queremos introduzir num determinado local, por exemplo, para elevar a produção por hectare de uma dada cultura, nós escolhemos um determinado grupo. Mas não sabemos, antes de conversar com o agricultor, qual o seu nivel de escolaridade. Mas sabemos antecipadamente que ele somente pode ser ALFABETIZADO ou ANALFABETO. Sendo ALFABETIZADO podemos estar interessados em saber quantos anos estudou, se completou o nivel fundamental, médio...

VARIÁVEL ALETATÓRIA, PORTANTO, é aquela que não sabemos antecipadamente o que acontecerá se fizermos um experimento ou uma pesquisa. Caso saibamos o que irá acontecer com certeza no nosso experimento, estaremos diante de variáveis DETERMINSITCAS ou NÃO ALEATÓRIAS

EXERCÍCIO

CRIE CINCO (5) VARIÁVEIS ALEATÓRIAS ASSOCIADAS À AGRICULTURA

SEGUNDO COMBO DE AULAS GRAVADA NO DIA 30/07 AULAS DOS DIAS 30 E 31 DE JULHO DE 2020

ESTATÍSTICA DESCRITIVA UNIVARIADA
BASEADA NO CAPITULO 3 DO LIVRO DE BISQUERRA et al.
SERÁ ENVIADA AMANHA PARA VOCES.

INTRODUÇÃO

Uma vez sabendo que em Estatística trabalhamos com variáveis discretas, continuas e, no nosso caso, variáveis quantitativas,
esse capitulo da disciplina consistirá em DESCREVER os comportamentos dessas variáveis quantitativas observadas de forma discreta ou contínua

Por estas razões a Seção é
Chamada de Estatísticas
Descritivas.
Voces podem encontrar no
Capitulo 3 do Livro do Bisquerra,
na pagina 39

Na vida real existem muitas variáveis que interagem:

Peso e Altura de uma pessoa, por exemplo.

Disponibilidade macronutrientes (N,P,K) no Solo e desenvolvimento das lavouras;

Produção de leite qualidade do rebanho.

Na realidade existe um emaranhado de variáveis que podem ser responsáveis pela ocorrência ou pelo desenvolvimento de um fenômeno.

Para um agricultor produzir um bem agrícola ele precisa ter vigor físico, tem intuição, ter área em tamanho adequado, ter acesso a sementes e / ou mudas de qualidade; ter acesso a outros insumos; ter disponibilidade de água....

Ou seja, há um complexo de variáveis que são responsáveis pelo sucesso ou pelo insucesso da sua produção.

Temos portanto o que chamamos em Ciência Estatística, Análise Multivariada.

O estudo dessas interações mais complexas entre variáveis aprendemos com o avanço do conhecimento estatístico e também com o acesso a uma base matemática melhor, que envolve estudo com Vetores, Matrizes...

Vamos começar a nossa jornada estudando

ANALISE UNIVARIADA

Com apenas uma variável

por exemplo, variáveis como:

Idade

peso

produção de feijão de um município

Área colhida

Precipitação de chuvas...

Chamaremos genericamente uma Variável de X, Y, Z

ou da forma que avaliarmos conveniente.

Vamos começar com o que chamamos de Distribuição de Frequências.

Temos três tipos de Frequências

Absolutas;

Relativas; e

Relativas acumuladas.

Vou mostrar para vocês cada um desses conceitos com uma tabela. Tabelas e Gráficos são instrumentos muito úteis em Estatística.

Vamos supor que a turma de voces que tem 22 estudantes matriculados tenham a distribuição de idades mostrada na primeira coluna da tabela a seguir.

Na segunda coluna, eu mostro o total de voces que me disseram as respectivas idades. São as

FREQUENCIAS ABSOLUTAS DA VARIÁVEL IDADE.

Na terceira coluna eu calculo a participação relativa das faixas etárias, no total. São as

FREQUÊNCIAS RELATIVAS DAS IDADES DE VOCÊS

Na Quarta Coluna eu **ACUMULO** as frequências absolutas, somando a frequencia absoluta da célula atual com o que aconteceu na célua anterior. De tal sorte que o ultimo acumulado **SEMPRE SERÁ IGUAL A 1,00**

Vamos ver a Tabela

Idade em Anos	Frequência absoluta	Frequência Relativa	Frequência Relativa Acumulada
17	4	(4/22) = 0,18	0,18
18	6	(6/22) = 0,27	0,18 + 0,27 = 0,45
19	5	(5/22) = 0,23	0,45 + 0,23 = 0,68
20	4	(4/22) = 0,18	0,68+0,18 = 0,86
22	3	(3/22) = 0,14	0,86 + 0,14 = 1,00
Total	22	(22/22) = 1,00	

Em geral o cálculo das frequências é o primeiro tratamento que damos a uma variável.

Essas frequencias, sobretudo a RELATIVA, já nos dá uma idéia do grau de homogeneidade ou de heterogeneidade da variável

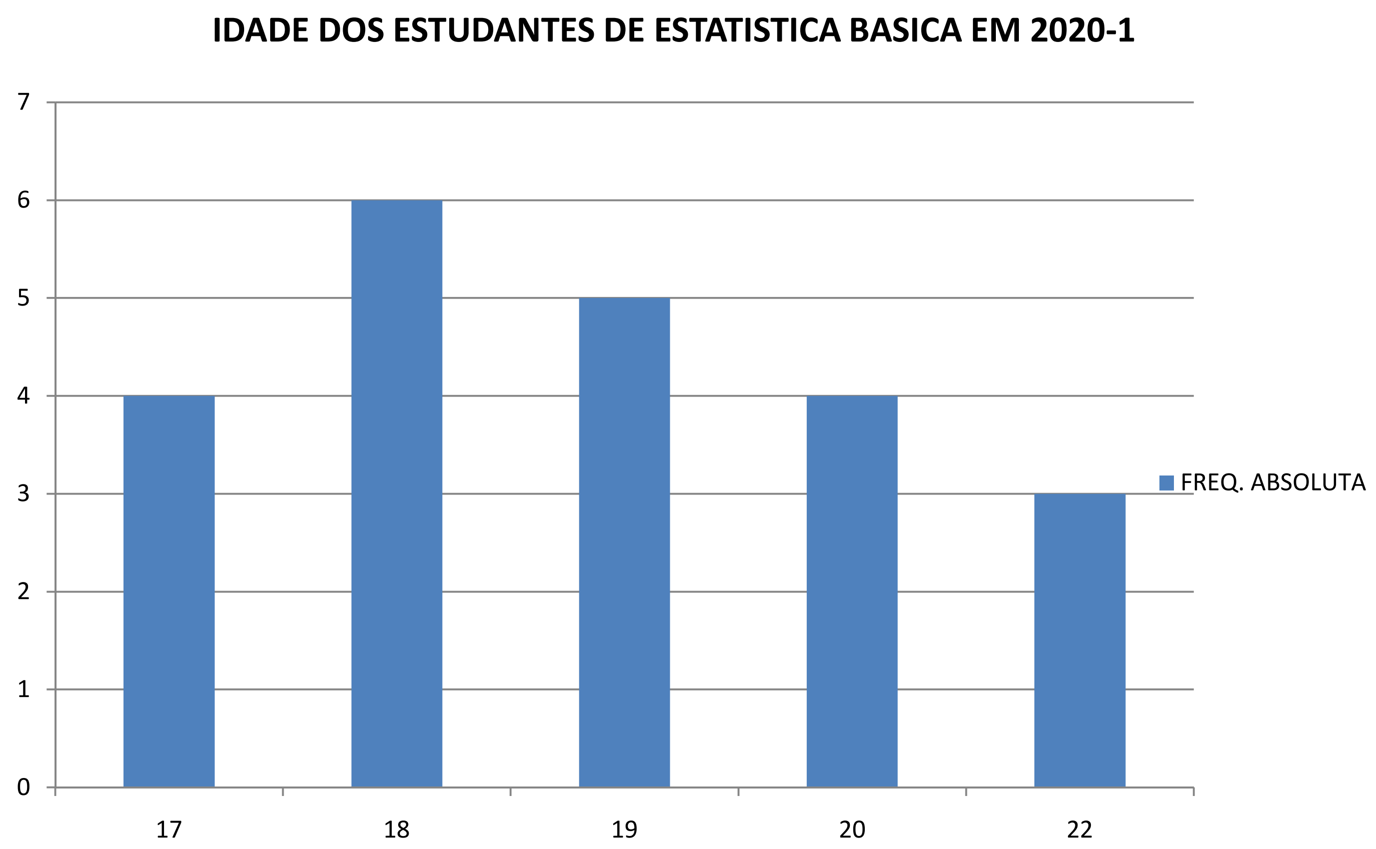
As frequências Relativas e Relativas Acumuladas podem ser multiplicadas por 100. Neste caso a leitura é em Percentagem.

No caso específico da Tabela diriamos que

18% tinham idade de 17 anos;

27% tinham idade de 18 anos e assim por diante.

A leitura da frequência acumulada seria a seguinte. A partir da segunda célula da quarta coluna:
45% dos estudantes tem idade entre 17 e 18 anos;
68% tem idade entre 17 e 19 anos...


IDADE DOS ESTUDANTES DE ESTATISTICA BASICA EM 2020-1
7
6
5
4
3
2
1
0
17
18
19
20
22
FREQ. ABSOLUTA

ESTATISTICAS DESCRITIVAS
MEDIDAS DE TENDENCIA CENTRAL
DE UMA VARIÁVEL ALEATÓRIA X

1 – MEDIANA

2 – MODA

3 – MÉDIA ARITIMÉTICA

4 – MÉDIA PONDERADA.

MEDIANA DE X

DEFINIÇÃO: Valor central (que divide ao meio) as observadas de X, depois da seqüências de observações tiver sido hierarquizada. A hierarquia dos valores da variável tanto pode ser em ordem crescente como decrescente.

Há duas situações na definição da MEDIANA.

1 – Quando o numero de observações é impar, faz-se a hierarquia e busca-se o valor que divide a amostra em DUAS PARTES IGUAIS

2 – Quando o numero de observações é par, fazemos a hierarquia da variável e identificamos os dois valores centrais. Somamos esses valores e dividimos por 2. Esse será o valor da mediana.

DESDOBRAMENTOS DA MEDIANA.

Em muitos casos, em vez de dividir a amostra em duas partes iguais, depois da variável ter sido hierarquizada, podemos dividir em QUATRO (4) partes iguais. Neste caso temos

QUARTIS.

Podemos dividir em CINCO (5) partes iguais. Neste caso teremos

QUINTIS.

Podemos dividir em dez (10) partes iguais e ai teremos

DECIS

Podemos ainda dividir a amostra com a variável devidamente hierarquizada em 100 partes iguais. Neste caso teremos

PERCENTIS.

Em muitas pesquisas que fazemos com agricultores utilizamos mais dessas formas de organização dos dados:
Quartis, Quintis, Decis, Pecentis.
Por exemplo, suponhamos que queremos saber o comportamento da produção na Bacia Leiteira do Ceará.
Tomamos os Produtores de leite e hierarquizamos em ordem crescente da produção diária de leite.
Depois dividimos a amostra em Cinco partes iguais. QUINTIS por exemplo. E ai observamos os 20% que tem a menor produção de leite (Quintil inferior) e comparamos, por exemplo com os 20% de agricultores que apresentam as maiores produções de leite (Quintil Superior). Ai observamos o que os agricultores que estão no Quintil Superior fazem que permitem eles estarem nesta situação e comparamos como as tecnicas de produção de leite dos produtores situados no Quintil inferior.

MODA

Por definição a Moda de uma variável aleatória é o valor ou a categoria de valores (no caso de variável continua) que surge com maior frequência na amostra.

Na aula expositiva eu falarei da utilidade dessa Medida de tendencia central e mostrarei que ela faz parte do nosso cotidiano e não nos demos conta até aqui.

AULA DO DIA 31/07/2020
Valor Esperado de Uma Variável Aleatória
Esperança Matemática
Média

Valor esperado de uma Variável Aleatória
Esperança Matemática ou Média

Em geral calculamos a media aritimética de uma variável aleatória X cujos valores observados são

$x1, x2, ..., xn$

Da seguinte forma

$$X_{média} = (x1 + x2 + ... + xn) / N$$

N é o número de observações.

Média Ponderada

Vimos na aula passada os conceitos de FREQUENCIAS ABSOLUTAS e FREQUENCIAS RELATIVAS.

Vimos aquele exemplo da sala de aula hipotética que vou pedir permissão para voces para trazer para cá

Idade em Anos Xi	Frequência absoluta	Frequência Relativa	Frequência Relativa Acumulada
17	4	(4/22) = 0,18	0,18
18	6	(6/22) = 0,27	0,18 + 0,27 = 0,45
19	5	(5/22) = 0,23	0,45 + 0,23 = 0,68
20	4	(4/22) = 0,18	0,68+0,18 = 0,86
22	3	(3/22) = 0,14	0,86 + 0,14 = 1,00
Total	22	(22/22) = 1,00	

Observem que cada idade acontece com determinada frequencia absoluta que é transformada em frequencia relativa

Essas frequencias são os PESOS associados a cada uma das idades.

Vamos calcular a média da idade dessa turma. Este é um caso típico de média ponderada

Em geral a média ponderada será calculada pela equação a seguir:

$$X_{\text{média}} = (P1X1 + P2X2 + \ldots + PNXN) / N$$

Neste caso P1, P2, ..., PN são os pesos, ou as frequencias associados a cada uma das observações de Xi

Aplicando essa equação no nosso exemplo, obteremos

$$X_{média} = (4.17 + 6.18 + 5.19 + 4.20 + 3.22)\ /22 = 19 \text{ ANOS}$$

Os pesos utilizados também poderiam ser as frequencias relativas de cada uma das idades. Neste caso não precisa dividir por N. O Cálculo da média ponderada seria feito da seguinte forma:

$$X_{média} = 0{,}18.\ 17 + 0{,}27.18 + 0{,}23.19 + 0{,}18.20 + 0{,}14.22 = 19 \text{ ANOS}$$

Mais na frente aprenderemos um conceito mais geral e mais elegante de média, entendida como valor esperado de uma variável aleatória.

Algumas propriedades associadas à média:

Média de (X + Y) = média de X + média de Y

Media de (X.Y) = média de X. média de Y

EXERCICIOS DE APLICAÇÃO

Na Tabela a seguir estão mostradas as precipitações médias de chuvas do Ceará entre os anos de 1985 e 2018 de acordo com a FUNCEME.

Calcule:

1 - Mediana

2 – Média aritmética

ANO	CHUVA
1985	1888,4
1986	1314,5
1987	738,4
1988	1148,5
1989	1257,7
1990	623,9
1991	771,2
1992	671,3
1993	418,5
1994	1156,1
1995	1067,1
1996	1064,1
1997	702
1998	529,4
1999	871,9
2000	1062,9
2001	700
2002	951,5
2003	965
2004	1168,5
2005	673,5
2006	830,1
2007	986,7
2008	969,1
2009	1341,7
2010	594,3
2011	1034,5
2012	388,8
2013	1300
2014	551,2
2015	532,7
2016	554,6
2017	698,2
2018	802,2

Na Tabela a seguir, extraída do IBGE temos as populações do Ceará em 2015
Responda à seguintes questões:
1 - Qual a Faixa etária que tinha o maior contingente populacional
em 2015?
2 – como é designada a frequência associada ao maior contingente populacional?

Grupos de idade	População residente (1 000 pessoas)								
	Total	Homens	Mulheres	Urbana			Rural		
				Total	Homens	Mulheres	Total	Homens	Mulheres
Total	**8 924**	**4 377**	**4 546**	**6 474**	**3 097**	**3 377**	**2 450**	**1 281**	**1 169**
0 a 4 anos	**565**	298	267	413	208	205	152	90	62
Menos de 1 ano	**116**	59	57	83	43	40	33	16	17
1 a 4 anos	**449**	239	210	330	165	166	119	74	45
5 a 9 anos	**683**	351	332	468	248	220	215	103	112
10 a 14 anos	**716**	389	327	499	271	228	217	118	99
15 a 19 anos	**819**	423	396	576	292	284	243	132	112
15 a 17 anos	**489**	261	229	341	179	162	149	82	67
18 ou 19 anos	**330**	163	167	235	113	122	95	50	45
20 a 24 anos	**750**	373	377	574	277	297	177	96	80
25 a 29 anos	**722**	354	368	538	266	273	184	88	95
30 a 34 anos	**716**	349	368	544	258	286	173	91	82
35 a 39 anos	**640**	311	328	475	226	249	165	85	80
40 a 44 anos	**556**	260	296	432	194	238	125	66	58
45 a 49 anos	**569**	273	296	409	192	217	161	81	79
50 a 54 anos	**480**	216	264	360	157	202	120	58	61
55 a 59 anos	**375**	172	203	273	115	158	102	57	45
60 a 64 anos	**381**	167	213	263	110	153	117	57	60
65 a 69 anos	**331**	149	182	230	100	129	102	49	53
70 anos ou mais	**619**	292	327	421	184	237	198	108	90

Todos os exercícios tem que ser enviados para o e-mail somel0225@gmail.com

Até quarta-feira próxima

Dia 5 /08/ 2020

O envio dos exercícios constará como frequencias nas aulas e farão parte da composição da nota final.

AULAS DOS DIAS 6 E 7 DE AGOSTO DE 2020

ESTATÍSTICAS DESCRITIVAS MEDIDAS DE VARIABILIDADE.

1 – AMPLITUDE DA VARIÁVEL ALEATORIA X (APX)

Amplitude da distribuição de uma variável aleatória é a diferença entre o maior valor e o menor valor.

Para se calcular a amplitude, hierarquiza-se o rol de observações da variável em ordem crescente ou decrescente. Identificam-se os valores extremos da variável aleatoria X: Maximo (Xmx) e mínimo (Xmn). Assim:

APX = Xmx - Xmn

EXEMPLO.

A FUNCEME MOSTRA QUE ENTRE 2010 E 2019 AS CHUVAS MEDIAS ANUAIS DO CEARÁ FORAM AS MOSTRADAS NA TABELA DO PROXIMO SLIDE

ANO	CHUVAS EM MM
2010	537
2011	996,6
2012	362,1
2013	546,9
2014	546,1
2015	523,1
2016	554,6
2017	698,2
2018	802,8
2019	846,2

NESTE CASO COMO SO TEMOS DEZ OBSERVAÇÕES É FACIL IDENTIFICAR QUE A CHUVA MÁXIMA ACONTECEU EM 2011 E A MINIMA OCORREU EM 2015

Xmx = 996,6 mm; Xmn = 523,1mm

PORTANTO:

APx = 996,6mm – 523,1mm = 473,5mm

A amplitude não é de muita utilidade na maioria dos nossos problemas agronômicos.

A não ser quando estamos interessados apenas em saber os valores extremos de uma variável. A sua utilidade é principalmente para esses casos.

Segunda medida de dispersão

DESVIO EM TORNO DA MÉDIA

(ξi)

Vamos chamar a Média de uma variável aleatória Xi (i = 1, 2, ..., n) de (Xmd)

O desvio em torno da média será dado por

$$\xi_i = X_i - X_{md}$$

Os valores do desvio serão positivos,

quando Xi > Xmd,

ou negativos,

quando Xi < Xmd

VAMOS DAR UM EXEMPLO NO PROXIMO SLIDE MOSTRANDO AS CHUVAS QUE ACONTECERAM ENTRE 2000 E 2019 NO CEARÁ SEGUNDO A FUNCEME.

ANO	CHUVAS (mm)
2000	888,6
2001	566,5
2002	804
2003	825,5
2004	1028,1
2005	580,9
2006	759
2007	663,4
2008	919,7
2009	1211,3
2010	537
2011	996,6
2012	362,1
2013	546,9
2014	546,1
2015	523,1
2016	554,6
2017	698,2
2018	802,8
2019	846,2

Na Tabela mostrada no Slide anterior apresentamos as chuvas anuais do Ceará entre 2000 e 2019 A média de chuvas ESTIMADA para o período é de 733,03 mm.

NO PRÓXIMO SLIDE MOSTRAMOS
A VARIÁVEL CHUVAS, A SUA MÉDIA
E OS DESVIOS EM TORNO DELA.

CHUVA MÉDIA(mm) = 733

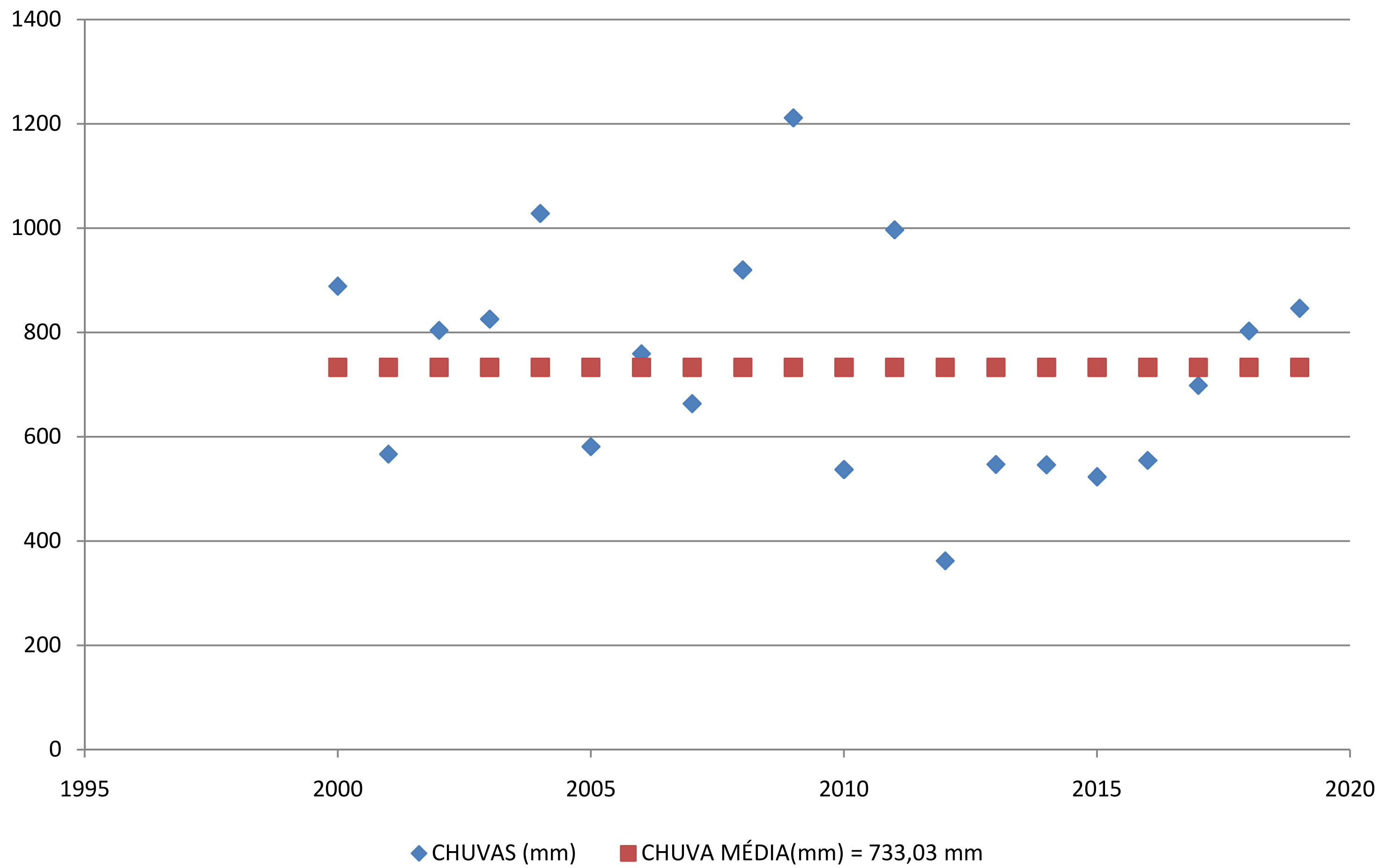
ANO	CHUVAS (mm)	mm	DESVIO = ξi
2000	888,6	733,03	155,57
2001	566,5	733,03	-166,53
2002	804	733,03	70,97
2003	825,5	733,03	92,47
2004	1028,1	733,03	295,07
2005	580,9	733,03	-152,13
2006	759	733,03	25,97
2007	663,4	733,03	-69,63
2008	919,7	733,03	186,67
2009	1211,3	733,03	478,27
2010	537	733,03	-196,03
2011	996,6	733,03	263,57
2012	362,1	733,03	-370,93
2013	546,9	733,03	-186,13
2014	546,1	733,03	-186,93
2015	523,1	733,03	-209,93
2016	554,6	733,03	-178,43
2017	698,2	733,03	-34,83
2018	802,8	733,03	69,77
2019	846,2	733,03	113,17
SOMA DOS DESVIOS			**ZER0**

NO GRÁFICO A SEGUIR OS PONTOS ABAIXO E ACIMA DA MÉDIA (LINHA VERMELHA), SÃO AS OBSERVAÇOES ANUAIS DE CHUVAS.

AS DISTANCIAS ENTRE CADA PONTO E A RETA VERMELHA SÃO OS DESVIOS EM TORNO DA MÉDIA DE CHUVAS (733,03MM).

AS OBSERVAÇÕES QUE ESTÃO ACIMA TERÃO DESVIOS POSITIVOS EM RELAÇÃO À MÉDIA.

AS OBSERVAÇÕES ABAIXO DA MÉDIA TERÃO DESVIOS NEGATIVOS EM RELAÇÃO À MÉDIA



1400
1200
1000
800
600
400
200
0
1995
2000
2005
2010
2015
2020
CHUVAS (mm)
CHUVA MÉDIA(mm) = 733,03 mm

PROPRIEDADE IMPORTANTE DA MÉDIA:

A SOMA DOS DESVIOS EM TORNO DA MÉDIA SERÁ **SEMPRE** IGUAL A ZERO

OUTRA MEDIDA DE DISPERSÃO

VARIÂNCIA (s^2)

Por definição a variância é a soma dos quadrados dos desvios (SQD) dividido pelo numero de observações (N) menos um

$$s^2 = ((SQD / (N – 1))$$

EM SEGUIDA MOSTRA-SE A SEQUENCIA AINDA UTILIZANDO A SÉRIE DE PRECIPITAÇÃO DE CHUVAS NO CEARÁ DE 2000 A 2019.

ANO	CHUVAS (mm)	CHUVA MÉDIA(mm) = 733,03 mm	DESVIO = ξi	DESVIO AO QUADRADO
2000	888,6	733,03	155,57	24202,02
2001	566,5	733,03	-166,53	27732,24
2002	804	733,03	70,97	5036,74
2003	825,5	733,03	92,47	8550,7
2004	1028,1	733,03	295,07	87066,3
2005	580,9	733,03	-152,13	23143,54
2006	759	733,03	25,97	674,44
2007	663,4	733,03	-69,63	4848,34
2008	919,7	733,03	186,67	34845,69
2009	1211,3	733,03	478,27	228742,2
2010	537	733,03	-196,03	38427,76
2011	996,6	733,03	263,57	69469,14
2012	362,1	733,03	-370,93	137589,1
2013	546,9	733,03	-186,13	34644,38
2014	546,1	733,03	-186,93	34942,82
2015	523,1	733,03	-209,93	44070,6
2016	554,6	733,03	-178,43	31837,26
2017	698,2	733,03	-34,83	1213,13
2018	802,8	733,03	69,77	4867,85
2019	846,2	733,03	113,17	12807,45
SOMA			**0**	**854711,682**

OBSERVA-SE QUE

SQD = 854711,682

N – 1 = 20 – 1 = 19

PORTANTO,

s^2 = 854711,682 / 19 = 44984,83

DESVIO PADRÃO (s).

Por Definição o Desvio Padrão associado a uma variável aleatória é a raiz quadrada da variância,

ou a variância elevada à potencia

0,5 ou ½

ASSIM, escreve-se

$$\mathbf{s = (s^2)^{0,5}}$$

O desvio padrão afere a dispersão média em torno da média. É aferido na mesma unidade de medida da média.

No caso estudado, as chuvas do Ceará entre 2000 e 2019 tiveram média de 733,03 mm, com desvio padrão entorno de 212,10mm

A variância, tal como mostrada até aqui, assim como o desvio padrão,são medidas absolutas de variabilidade.

EXERCICIOS DE APLICAÇÃO DAS AULAS DOS DIAS 6 E 7 DE AGOSTO

NA TABELA MOSTRADA NO PRÓXIMO SLIDE TEMOS AS PRECIPITAÇÕES DE CHUVAS NO CEARÁ ENTRE OS ANOS DE 1990 E 2019 SEGUNDO A FUNCEME

1 - Identifique o ano que teve a maior chuva e o que teve a menor chuva.

2 – calcule a amplitude das chuvas no período.

3 – calcule a precipitação média no período.

4 – complete a tabela mostrando os desvios em torno da média e os quadrados dos desvios

5 – façam o gráfico mostrando a média com a dispersão em torno dela.

6 – Calcule a soma dos desvios em torno da média

7 – calcule a variância das chuvas no período.

8 – calcule o desvio padrão das chuvas no período.

Caso se considere ANOS DE ESTIAGEM aqueles que chove menos de 600 mm, pergunta-se:

9 – No período analisado, em quantos anos aconteceram estiagens

10 – calcule a amplitude, a média e o desvio padrão das chuvas acontecidas nos períodos de estiagem.

DATA LIMITE DE ENVIO DOS EXERCIDIOS: 12/08/2020

ANO	CHUVAS ANUAIS (MM)
1990	567,8
1991	703,6
1992	607,9
1993	376,9
1994	1022
1995	947,1
1996	956,5
1997	667,2
1998	432,6
1999	771,4
2000	888,6
2001	566,5
2002	804
2003	825,5
2004	1028,1
2005	580,9
2006	759
2007	663,4
2008	919,7
2009	1211,3
2010	537
2011	996,6
2012	362,1
2013	546,9
2014	546,1
2015	523,1
2016	554,6
2017	698,2
2018	802,8
2019	846,2

AULA DO DIA 13 DE AGOSTO CONCLUSÃO DA UNIDADE DE ESTATISTICAS DESCRITIVAS.

O último tópico que iremos discutir se refere as Medidas de Variabilidade, é o

Coeficiente de Variação (CV)

ANTES DE APRESENTAR O CV, VAMOS FAZER UM EXERCICIO.

EU NÃO MOSTREI AINDA PRA VOCES COMO SE CALCULA A MÉDIA DE UMA VARIÁVEL ALEATÓRIA CONTINUA.

VAMOS FAZER ISSO AGORA A PARTIR DA TABELA A SEGUIR QUE FOI RETIRADA DO IBGE E QUE JÁ PASSEI PRA VOCES. NELA EU COLOQUEI APENAS AS DUAS PRIMEIRAS COLUNAS

POPULAÇÃO DO CEARÁ EM 2017 POR ESTRATO DE IDADE.

Estrato de Idades	Ponto Médio do Estrato (PM)	Freq. Absoluta (FA)	Freq. Relativa (FR)	PMxFR
0 a 4 anos	2	565	0,06	0,13
5 a 9 anos	7	683	0,08	0,54
10 a 14 anos	12	716	0,08	0,96
15 a 19 anos	17	819	0,09	1,56
20 a 24 anos	22	750	0,08	1,85
25 a 29 anos	27	722	0,08	2,18
30 a 34 anos	32	716	0,08	2,57
35 a 39 anos	37	640	0,07	2,65
40 a 44 anos	42	556	0,06	2,62
45 a 49 anos	47	569	0,06	3
50 a 54 anos	52	480	0,05	2,8
55 a 59 anos	57	375	0,04	2,4
60 a 64 anos	62	381	0,04	2,65
65 a 69 anos	67	331	0,04	2,49
70 a 100 anos	85	619	0,07	5,9
TOTAIS		**8922**	**1**	**34,3**

Quando a variável é colocada de forma continua com estratos com diferentes **FREQUENCIAS ABSOLUTAS,** procede-se da forma que está na tabela.

CALCULAM-SE OS PONTOS MÉDIOS DOS ESTRATOS. ISSO É FEITO SOMANDO OS VALORES EXTREMOS E DIVINDO POR DOIS. O PRIMEIRO ESTRATO, POR EXEMPLO VARIA DE

0 – 4 ANOS

A MÉDIA DO ESTRATO SERÁ

$(0 + 4)/2 = 2$

CALCULAM-SE AS FREQUENCIAS RELATIVAS (FR) DOS ESTRATOS QUE SÃO DADAS PELA RELAÇÃO

FR = (FA / TOTAL)

A POPULAÇÃO TOTAL DO CEARÁ EM 2017 ERA DE 8922 PESSOAS

(O DADO QUE ESTÁ NO IBGE NÃO ESTÁ CERTO, A SOMA DELES DEU 8924)

A FR ASSOCIADA AO PRIMEIRO ESTRATO SERÁ:

FR1 = 565/8922 = 0,06

A SEGUIR APLICA-SE O CONCEITO DE VALOR ESPERADO OU DE MEDIA DE UMA VARIÁVEL ALEATORIA X QUE É DADO POR

$$E(X) = FR1.X1 + FR2X2 + FRnXN$$

$$E(X) = 0,06.2 + 0,08.7 + ... + 0,07.85$$

$$E(X) = 34,3$$

ASSIM, A IDADE MÉDIA DOS CEARENSES **EM 2017 ERA DE 34,3 ANOS**

COEFICIENTA DE VARIAÇÃO (CV)

DEFINIÇÃO:
Por Definição o CV de uma variável aleatória X é a relação percentual entre o desvio padrão (DPx) e a Média da Variável (MEx)

CV = (DPx / Mex).100

O CV AFERE A HOMOGENEIDADE/HETEROGENEIDADE DA DISTRIBUIÇÃO DA VARIÁVEL X

Quanto maior o CV, mas heterogênea a distribuição.

CV, ao contrario do Desvio Padrão e da Amplitude, que são aferidos na mesma unidade de medida de X, é um indicador adimencional
Por isso permite a comparação entre variáveis distintas que estão medidas na mesma unidade de medida ou não.

AULA DO DIA 14/08/2020
TEORIA DAS PROBABILIDADES

AULA DO DIA 14/08/2020
TEORIA DAS PROBABILIDADES

EVENTOS ALEATORIOS
O QUE É ALEATORIEDADE?
O QUE SÃO VARIÁVEIS ALEATORIAS?
EVENTOS ALEATORIOS E EVENTOS DETERMINISTICOS

ESPAÇO AMOSTRAL

Define-se como espaço amostral, ou espaço das possibilidades, ao conjunto “S” de todos os resultados possíveis de ocorrer em um experimento sujeito às leis do acaso.

Probabilidade e as suas Propriedades

A Probabilidade é um número associado a um evento, destinado a medir a sua possibilidade de ocorrência

Definições

N = número de casos possíveis de ocorrência do evento E

k = número de casos favoráveis à ocorrência do evento E

Define-se a probabilidade de ocorrencia do evento E pela seguinte equação:

$$P(E) = k / N$$

Assim, a probabilidade de ocorrência do evento E se obtêm dividindo-se dentro de um espaço amostral o numero de casos favoráveis à ocorrencia de E naquele espaço amostral (S), pelo total de todas as possibilidades de casos no mesmo espaço S.

São as seguintes as propriedades associadas à probabilidade de ocorrência do evento E:

a) $0 \leq P(E) \leq 1$;

b) $P(S) = 1$

c) se A, B, C, ..., são eventos mutuamente excludentes:

$$P(A \cup B \cup C \ldots) = P(A) + P(B) + P(C) + \ldots$$

d) $P(A \cup B) = P(A) + P(B) + P(A \cap B)$

e) $\Sigma[P(E)] = 1$

Exercícios para a próxima semana:
1 – Aquela planilha que eu já mandei pra vocês com as popualações do Ceará em 2017, que vou repetir em seguida:
a) calcule a média de idade dos homens e das mulheres que moravam no Ceará naquele ano (lembrando que tem que calcular os pontos médios);
b) Calculem a probabilidade de uma pessoa sorteada (de qualqer sexo) ter idade entre 19 e 19 anos;
e) calcular a probabilidade de uma Cearense mulher ter mais de 50 anos e de um Cearense homem ter mais de 50 anos.

NA PRATICA A PROBABILIDADE É APROXIMADA PELA FREQUENCIA RELATIVA