ESTATÍSTICA II

Michelle Hanne Soares de Andrade

michellehanne@cefetmg.br

1º. SEMESTRE 2018

Revisão Geral Estatística I

Resumo – População e Amostra

- **População e Amostra:** Ao examinar um grupo qualquer, considerando todos os seus elementos, estamos tratando da **população** ou **universo**. Nem sempre isso e possível. Nesse caso, examinamos uma pequena parte chamada **amostra**.
- Uma **população** pode ser finita ou infinita. Por exemplo:
  - a população dos alunos de sua escola é finita e a população constituída de todos os resultados (cara ou coroa) em sucessivos lances de uma moeda e infinita.
- Se uma **amostra** é representativa de uma população, podemos obter conclusões importantes sobre a população.

Medidas de Posição

- Indicam alguma tendência central ou comportamento esperado com respeito extraídos de uma amostra qualquer.
- Dentre as principais medidas, destacam-se: **media, mediana e moda.**

Medidas de Posição

- Sejam $x_1, x_2, \dots x_n$ n observações de um fenômeno aleatório qualquer. Denominamos media aritmética da amostra.

$$\overline{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i,$$

Média Aritmética

- Pode ser interpretada como o centro de massa (baricentro das observações).

- Caso os dados estejam organizados em uma tabela de frequências, isto é, se para cada xi foi verificado um número fi relativo a sua frequência, então

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i f_i}{\sum_{i=1}^{n} f_i}.$$

Exemplo

Uma pesquisa avaliou a idade dos 25 alunos de uma turma de engenharia. Os dados foram organizados na seguinte tabela de frequências:

Idade	**freq.**
19	3
20	5
21	1
22	8
23	4
24	1
25	0
26	3
Total	25

A média aritmética pode ser obtida diretamente do dispositivo (1):

$$\overline{x} = \frac{\sum_{i=1}^{8} x_i f_i}{\sum_{i=1}^{8} f_i} = \frac{548}{25} = 21,92.$$

Mediana - dados discretos ou não agrupados

- Caso os dados de interesse sejam discretos ou, sejam contínuos e ainda não tenham passado por um agrupamento, então o calculo da mediana é mais simples.
- Basta observar a paridade do número n de observações:
  - **se n e ímpar, então a mediana e exatamente o valor central das observações.**

$$x_{\left(\frac{n+1}{2}\right)};$$

  - caso contrario, a mediana e dada pela media aritmética dos dois valores centrais

$$\frac{x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)}}{2}.$$

Mediana - dados discretos ou não agrupados

- Exemplos:
  - No exemplo dos salários, temos a seguinte ordenação: R$1000,00, R$1800,00, R$2500,00, R$2500,00, R$20000. Como o número de observações e ímpar, temos que a mediana e dada pelo elemento **central x(3) = 2500.**
  - No exemplo das classes sociais, temos a seguinte ordenação: B, B, B, B, B, M, M, M, M, M, M, M, M, M, M, A, A, A, A, A, A, A, A, A, A,A, A, A, A, A. Como o numero de observações é par, segue que a mediana é dada pelo ponto médio de x(15) e x(16). A mediana então ficaria entre os valores M e A.

Moda (Mo)

- Moda de dados agrupados sem intervalo de classe:

Nº DE MENINOS	f_i
0	2
1	6
2	10
3	12
4	4
	$\Sigma = 34$

Mo = 3,

(pois 3 tem frequência 12)

Moda (Mo)

- Moda de dados agrupados com intervalo de classe:
- A classe de maior frequência é chamada classe modal. A moda será, então, o ponto médio desta classe.

$$\boxed{Mo = \frac{\ell^* + L^*}{2}}$$

i	ESTATURAS (cm)	f_i
1	150 ⊢ 154	4
2	154 ⊢ 158	9
3	158 ⊢ 162	11 ←
4	162 ⊢ 166	8
5	166 ⊢ 170	5
6	170 ⊢ 174	3
		$\Sigma = 40$

$$Mo = \frac{\ell^* + L^*}{2}$$

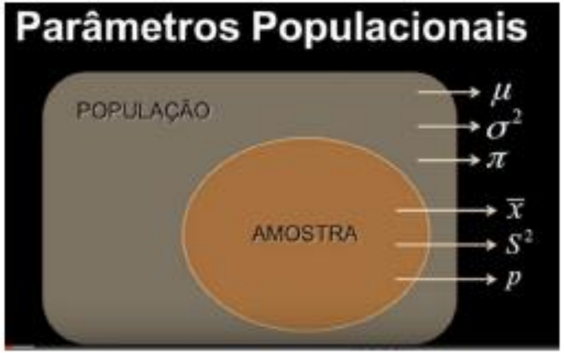
$$Mo = \frac{158 + 162}{2} = \frac{320}{2} = 160$$

$$Mo = 160 \text{ cm}$$

Resumo

Parâmetro (População)	
Valor médio	μ
Desvio padrão	σ
Proporção	p
Correlação	ρ

Estatística (Amostra)	
Média	$\bar{x}$
Desvio padrão	s
Proporção	$\hat{p}$
Correlação	r



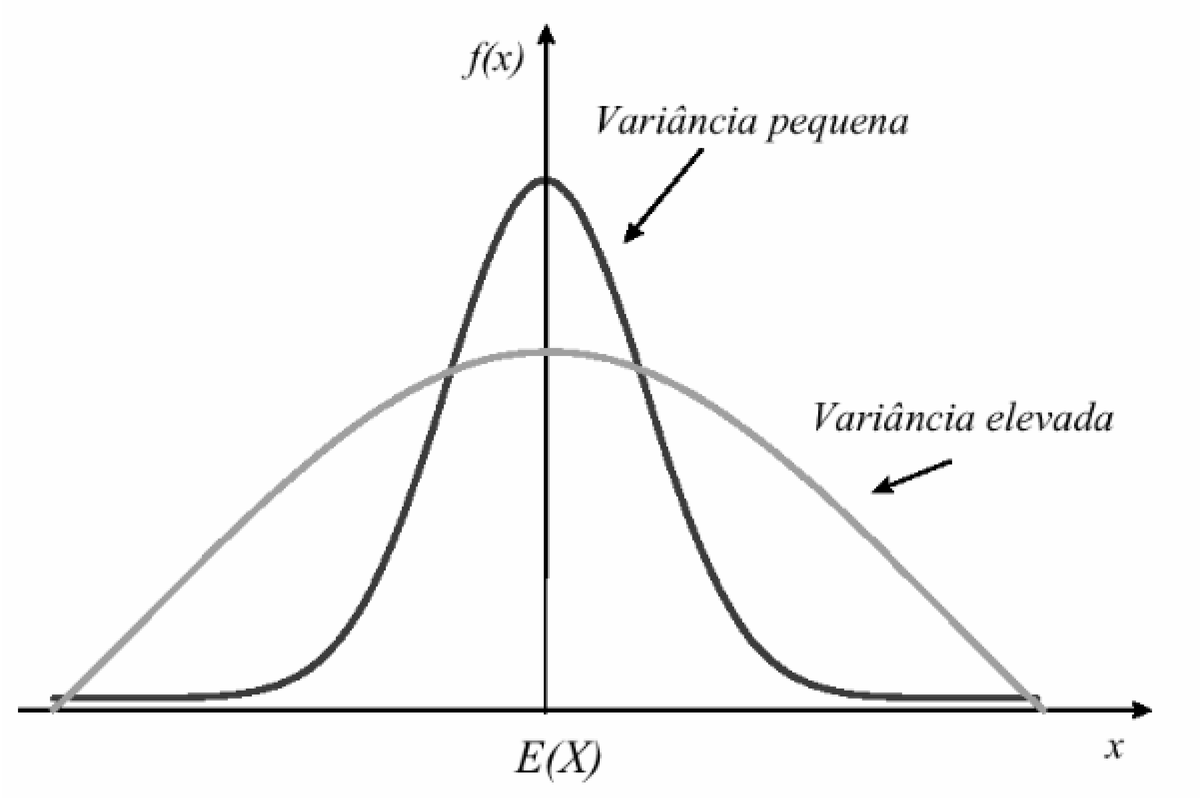
Resumo - Desvio Padrão (S) x Variância (S^2)

- **Desvio Padrão (S) x Variância (S^2)**
  - O desvio padrão é a medida mais usada na comparação de diferenças entre conjuntos de dados, por ter grande precisão. O desvio padrão determina a dispersão dos valores em relação à média e é calculado por meio da raiz quadrada da variância.

Resumo - Variância (S^2)

- **Variância (S^2)**
  - Sendo a variância calculada a partir dos quadrados dos desvios, ela é um número em unidade quadrada em relação a variável em questão, o que, sob o ponto de vista prático é um inconveniente; por isso, tem pouca utilidade na estatística descritiva, **mas é extremamente importante na inferência estatística e em combinações de amostras.**

Resumo – Variância (S^2)



Desvio Padrão

- O conceito de variância é bastante rico, contudo, deve ser utilizado com cautela já que trata do problema original em escala quadrática.
- O desvio padrão surge como uma alternativa para corrigir este detalhe e assim facilitar a análise dos resultados.
- Tal medida e dada pela raiz quadrada da variância amostral:

$$S = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}x_i^2 - \overline{x}^2}.$$

Desvio Padrão

Em geral, a variância amostral possui propriedades matemáticas melhores, enquanto o desvio padrão oferece interpretações mais razoáveis.

Desvio Padrão

Desvio Padrão (S)

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}}$$

x_i são os valores da variável

média

Exemplo:

Calcular o desvio padrão

da população

representada por:

{ - 4 , -3 , -2 , 3 , 5}

Xi	$\bar{X}$	(Xi - $\bar{X}$)	(Xi - $\bar{X}$)^2
- 4	- 0,2	- 3,8	14,44
- 3	- 0,2	- 2,8	7,84
- 2	- 0,2	- 1,8	3,24
3	- 0,2	3,2	10,24
5	- 0,2	5,2	27,04
		Σ =	62,8

Sabemos que n = 5 e 62,8 / 5 = 12,56.

A raiz quadrada de 12,56 é o desvio padrão = **3,54**

Resumo – Desvio Padrão

- **O desvio padrão é uma medida que só pode assumir valores não negativos e quanto maior for, maior será a dispersão dos dados.**
- Algumas propriedades do desvio padrão, que resultam imediatamente da definição, são:
  - O desvio padrão é sempre não negativo e será tanto maior, quanta maior a variabilidade entre os dados.
- **Se S = 0, então não existe variabilidade, isto é, os dados são todos iguais.**

Resumo – Exercício

- Tendo por base uma amostra da altura de uma parcela da população apresentada na Tabela 5.2, determinar:

  a) A variância das alturas;

  b) O desvio-padrão das alturas.

Tabela 5.2 – Estatura de uma amostra de uma população A

Altura (cm)	Nº de pessoas
150 |— 158	5
158 |— 166	18
166 |— 174	42
174 |— 182	27
182 |— 190	8
$\sum$	100

Fonte: Dados fictícios, apenas para fins ilustrativos.

Resumo – Exercício - Solução

Solução

a) A variância das alturas

Usando a fórmula 5.13 obtém-se o seguinte resultado:

$$s^2 = \frac{\sum X_i^2 f_i - \frac{\left(\sum X_i f_i\right)^2}{n}}{n-1}$$

Para calcular a variância, necessita-se conhecer as informações a seguir, cujos valores estão calculados na Tabela 5.3:

$$\sum X_i f_i$$

$$\sum X_i^2 f_i$$

Resumo – Exercício - Solução

Solução

a) A variância das alturas

Usando a fórmula 5.13 obtém-se o seguinte resultado:

$$s^2 = \frac{\sum X_i^2 f_i - \frac{\left(\sum X_i f_i\right)^2}{n}}{n-1}$$

Para calcular a variância, necessita-se conhecer as informações a seguir, cujos valores estão calculados na Tabela 5.3:

$\sum X_i f_i$

$\sum X_i^2 f_i$

Tabela 5.3 – Tabela auxiliar da Tabela 5.2

Altura (cm)	Nº de pessoas	X_i	$X_i f_i$	$X_i^2 f_i$
150 |— 158	5	154	770	118.580
158 |— 166	18	162	2916	472.392
166 |— 174	42	170	7140	1.213.800
174 |— 182	27	178	4806	855.468
182 |— 190	8	186	1.488	276.768
$\sum$	100		17.120	2.937.008

Fonte: Dados fictícios, apenas para fins ilustrativos.

Resumo – Exercício - Solução

Substituindo os valores na fórmula 5.13 obtém-se os seguinte resultados:

$$s^2 = \frac{2937008 - \frac{17120^2}{100}}{100 - 1}$$

$$s^2 = \frac{2937008 - 2930944}{100 - 1} = \frac{6064}{99}$$

$s^2 = 61,25$

Resposta:

A variância das alturas é de 61,25 cm²

b) O desvio-padrão das alturas

Solução

Extrai-se a raiz quadrada da variância.

Assim,

$$s = \sqrt{s^2}$$

$$s = \sqrt{61,25}$$

s = 7,8263

Resposta: As estaturas das pessoas estão dispersas em média 7,83 cm em relação à média da distribuição.

Teoria das Probabilidades

Experimentos Aleatórios Estocásticos

- Experimentos que ao serem repetidos nas mesmas condições não produzem o mesmo resultado são **denominados experimentos aleatórios (estocásticos)**.

Exemplos:

- Quanto se retira um lote de peças num processo de produção, observa-se que o número de peças defeituosas varia de lote para lote;
- Uma lâmpada é escolhida no processo de fabricação e se observa o seu tempo de duração, verifica-se que esse tempo varia de lâmpada para lâmpada.

Experimentos Determinísticos

- Os experimentos que ao serem repetidos nas mesmas condições conduzem ao mesmo resultado são **denominados determinísticos**.
- Exemplos:
- Se uma pedra cai de uma certa altura, pode-se determinar sua posição e velocidade para qualquer instante de tempo posterior a queda;
- Se agua é aquecida a 100 graus centigrados ao nível do mar, ela entrara em ebulição.

Experimentos

- Lance uma moeda ate que ocorra uma cara e, então conte o n° de vezes que a moeda foi lançada. O espaço amostral deste experimento é

$$E = \{1, 2, 3, \ldots, \infty\}$$

- O infinito refere-se ao caso em que nunca ocorre cara e, assim a moeda é lançada um n° infinito de vezes.
- Este exemplo mostra um **espaço amostral infinito enumerável.**

Espaços Amostrais e Eventos

- Chamamos de espaço amostral ao conjunto de todos os resultados possíveis de um fenômeno aleatório.
- Usualmente tal conjunto é representado por Ω.
- Os subconjuntos de Ω são denominados eventos e são representados por meio de letras maiúsculas:

$$A,\ B,\ C,\ \text{etc.}$$

- O conjunto vazio sempre pertence a Ω e é representado por $\emptyset$.

Espaços Amostrais e Eventos - Exemplo

- Considere o fenômeno aleatório caracterizado pelo lançamento de um dado.
- Neste caso,

$$\Omega = \{1, 2, 3, 4, 5, 6\}.$$

- Ω é dito um espaço amostral finito.
- Alguns exemplos de eventos:
  1. $E_1 =$ "o número obtido é par" $= \{2, 4, 6\} \subset \Omega$
  2. $E_2 =$ "o número obtido é ímpar" $= \{1, 3, 5\} \subset \Omega$
  3. $E_3 =$ "o número obtido é menor que 5" $= \{1, 2, 3, 4\} \subset \Omega$
  4. $E_4 =$ "o número obtido é maior que 6" $= \emptyset \subset \Omega$

Probabilidade Empírica

Repetindo-se um experimento E um grande número de vezes e calculando-se a frequência relativa do evento A, obtém-se um número “p” que pode ser tomado como a probabilidade da ocorrência de A, que nesse caso, poderia ser tomada como:

$$P(A) = p = \lim_{n \to \infty} \frac{f(A)}{n}$$

$$P(A) = \frac{numero\ de\ ocorrencias\ de\ A}{numero\ de\ repeticao\ do\ experimento}$$

Probabilidade Empírica - Exemplo

- Uma escola particular pretende oferecer um treinamento esportista aos seus alunos. Dos 300 alunos entrevistados, 142 optaram pelo voleibol, 123 indicaram o basquete e 35 indicaram o futebol.
- Selecionado aleatoriamente um desses alunos, qual a probabilidade de obter alguém que prefere o voleibol?

$$P(V)=\frac{142}{300}=0.47333.$$

- **A medida que o número de observações aumenta, as aproximações tendem a ficar cada vez mais próximas da probabilidade efetiva.**

Exemplos – Ocorrência de um Evento

Exemplo 1

- No lançamento de uma moeda não viciada, qual a probabilidade de cair coroa?
  - numero de eventos desejados = 1 (só ha uma face coroa)
  - numero de eventos possíveis = 2 (ha 2 faces na moeda).
  - Portanto, P = 1/2
- Nesse caso os eventos são chamados de equiprováveis.
- Caso idêntico a esse é o sexo de criança esperada apos uma gravidez:
  - P(menina) = 1/2 e P(menino) = 1/2.

Exemplos – Ocorrência de um Evento

Exemplo 2

- No lançamento de um dado não viciado, qual a probabilidade de cair face 4?
  - numero de eventos desejados = 1 (só ha uma face 4)
  - numero de eventos possíveis = 6 (ha 6 faces no dado).
  - Portanto, P = 1/6

Ocorrência de dois Evento - Exemplos

Exemplo 1- Qual a probabilidade de cair face 3 ou face 6 em um único lançamento de dado?

- P(3) = 1/6
- P(6) = 1/6
- **P(3 ou 6) = 1/6 + 1/6 = 2/6 = 1/3**

Exemplo 2- Qual a probabilidade de se retirar uma dama de um baralho previamente embaralhado?

- P (dama copas) = P (dama ouro) = P (dama espadas) = P (dama paus) = 1/52
- **Portanto, P (dama) = 1/52 + 1/52 + 1/52 + 1/52 = 4/52 = 1/13**

Ocorrência de dois Evento - Exemplos

Exemplo 3- Consideremos o experimento do lançamento de duas moedas, onde o espaço amostral é $\Omega = \{ KK, KC, CK, CC\}$.

- Sejam os eventos A = “ocorrência de exatamente 1 cara” e B = “duas faces iguais”.
  - Então $A = \{KC, CK\}$ e $B = \{CC, KK\}$; logo, $A \cup B = \Omega$ e $A \cap B = \emptyset$.
  - Seja C o evento “pelo menos uma cara”; então $C = \{KC, CK, KK\}$ e $B \cup C = \Omega$ e $B \cap C = \{KK\}$.

Probabilidade Condicional - Exemplo

Exemplo 1- De um baralho de 52 cartas (13 de ouro, 13 de espadas, 13 de copas e 13 de paus) qual é a probabilidade de, **ao ser retirada uma carta, ser uma dama de copas?**

- P = 1/52

Exemplo 2- De um baralho de 52 cartas (13 de ouro, 13 de espadas, 13 de copas e 13 de paus) qual é a probabilidade de, **ao ser retirada uma carta, ser uma dama, sabendo-se que a carta retirada e de copas.**

- Como há 4 tipos (se não soubesse que naipe é): **P = 1/52 X 4**
- Como já se sabe que a carta é de copas, temos apenas 1 dama em um total de 13 cartas. A probabilidade é então: **P(Q, copas) = 1/13.**

Exercícios

- **Solução:** Sabemos que a probabilidade da mulher engravidar em um mês e de 20%, que na forma decimal e igual a 0,2. A probabilidade dela não conseguir engravidar e igual a 1 - 0,2, ou seja, e igual a 0,8.
- Eventos consecutivos e independentes (pelo menos enquanto ela não engravida). Como a mulher só deve engravidar no quarto mês, então a probabilidade dos três meses anteriores deve ser igual a probabilidade dela não engravidar no mês, logo:

$$P(E) = 0.8 \, x \, 0.8 \, x \, 0.8 \, x \, 0.2 \rightarrow P(E) = 0{,}1024$$

- **0,1024 multiplicado por 100% é igual a 10,24%.**
- **Então: A probabilidade de a mulher vir a engravidar somente no quarto mês e de 10,24%.**

Exercícios

1) Alguns amigos estão em uma lanchonete. Sobre a mesa há duas travessas.

- Em uma delas há 3 pasteis e 5 coxinhas.
- Na outra ha 2 coxinhas e 4 pastéis.

- Se ao acaso alguém escolher uma destas travessas e também ao acaso pegar um dos salgados, qual a probabilidade de se ter pegado um pastel?

Exercícios

- **Solução**: A probabilidade de escolhermos 1 dentre 2 travessas e igual 1/2.
- A probabilidade de escolhermos um pastel na primeira travessa e 3 em 8, ou seja, e 3/8 e como a probabilidade de escolhermos a primeira travessa e 1/2, temos:

$$P(A)=\frac{1}{2}x\frac{3}{8}\rightarrow P(A)=\frac{3}{16}$$

- A probabilidade de escolhermos um pastel na segunda travessa e 4 em 6, isto e 4/6 e como a probabilidade de escolhermos a segunda travessa e igual a 1/2, temos:

$$P(B)=\frac{1}{2}x\frac{4}{6}\rightarrow P(B)=\frac{1}{3}$$

- Então a probabilidade de escolhermos um pastel e igual a: A probabilidade de se ter pegado um pastel é 25/48

$$P(E)=P(A)+P(B)\rightarrow P(E)=\frac{3}{16}+\frac{1}{3}=\frac{25}{48}$$