

BIOESTATÍSTICA

AULA 2

Anderson Castro Soares de Oliveira
Jose Nilton da Cruz

Departamento de Estatística/ICET/UFMT

Estatística Descritiva

ESTATÍSTICA DESCRITIVA

- Um dos objetivos da Estatística é sintetizar os valores que uma ou mais variáveis podem assumir
- Estatística descritiva é parte da estatística que lida com a organização, resumo e apresentação de dados. Esta é feita por meio de:
 - Tabelas;
 - Gráficos;
 - Medidas Descritivas (média, variância, entre outras).
- A estatística descritiva a ser utilizada depende do tipo de variável:
 - Variável qualitativa- gráficos e tabelas
 - Variável quantitativa- gráficos, tabelas e medidas descritivas

Tabelas

TABELAS

- Todas as variáveis podem ser resumidas por meio de uma tabela, mas a construção é diferenciada dependendo do tipo de variável.
- Toda tabela deve ser, clara, objetiva e auto-explicativa.
- Os elementos fundamentais da tabela são:
 - título - aponta o fenômeno, época e local de ocorrência;
 - cabeçalho - especifica o conteúdo das colunas;
 - coluna indicadora - detalha as linhas;
 - corpo - mostra os dados;
 - linhas - contém uma sequência horizontal de informações;
 - colunas - contém uma sequência vertical de informações.

- Complementarmente, tem-se:
 - fonte - cita o informante (caracterizando a confiabilidade dos dados);
 - notas - esclarecem o conteúdo e indicam a metodologia adotada na obtenção ou elaboração da informação;
 - chamadas - clarificam pontos específicos da tabela.

Variáveis Qualitativas

- O primeiro passo da descrição de dados qualitativos é criar uma tabela de distribuição de frequências.
- Frequência - medida que quantifica a ocorrência dos valores de uma variável de um conjunto de dados. As frequências podem ser:
 - Absoluta (fa) - contagem das observações de uma variável;
 - Relativa (fr) - divisão da frequência absoluta pelo total de observações, assim representa a proporção de ocorrência dos elementos de uma variável.

$$fr = \frac{fa}{n}$$

- Percentual (fp) - é a frequência relativa multiplicada por 100, assim representa a porcentagem de ocorrência dos elementos de uma variável.

$$fp = 100 \times fr$$

Variáveis Qualitativas

- Para verificar as matrizes de inseminação em bovinos, uma empresa obteve 18 amostras de diferentes raças e verificou sua qualidade de acordo com parâmetros específicos, obtendo os seguintes resultados

Tabela 1: Variáveis observadas de 18 bovinos de uma empresa.

amostra	Raça	Qualidade Sêmen	amostra	Raça	Qualidade Semen
1	Gir	Bom	10	Holandesa	Ótimo
2	Gir	Bom	11	Holandesa	Ótimo
3	Gir	Bom	12	Holandesa	Regular
4	Gir	Bom	13	Holandesa	Regular
5	Gir	Ótimo	14	Nelore	Bom
6	Gir	Regular	15	Nelore	Bom
7	Gir	Regular	16	Nelore	Ótimo
8	Holandesa	Bom	17	Nelore	Ótimo
9	Holandesa	Bom	18	Nelore	Regular

Variáveis Qualitativas

Tabela 2: Distribuição de frequência da raça de 18 bovinos de uma empresa

Raça	Frequência Absoluta (fa)	Frequência Relativa (fr)	Frequência Percentual (fp)
Gir	7		
Holandesa	6		
Nelore	5		
Total	18		

Variáveis Qualitativas

Tabela 2: Distribuição de frequência da raça de 18 bovinos de uma empresa

Raça	Frequencia Absoluta (fa)	Frequencia Relativa (fr)	Frequencia Percentual (fp)
Gir	7	0,39	
Holandesa	6	0,33	
Nelore	5	0,28	
Total	18	1,00	

Variáveis Qualitativas

Tabela 2: Distribuição de frequência da raça de 18 bovinos de uma empresa

Raça	Frequência Absoluta (fa)	Frequência Relativa (fr)	Frequência Percentual (fp)
Gir	7	0,39	39%
Holandesa	6	0,33	33%
Nelore	5	0,28	28%
Total	18	1,00	100%

Variáveis Qualitativas

Tabela 3: Distribuição de frequência da qualidade de sêmen de 18 bovinos de uma empresa.

Qualidade Semen	Frequência Absoluta (fa)	Frequência Relativa (fr)	Frequência Percentual (fp)	Frequência Acumulada (FA)	Frequência Percentual Acumulada (FP)
Regular	5	0,28	28%		
Bom	8	0,44	44%		
Ótimo	5	0,28	28%		
Total	18	1,00	100%		

Variáveis Qualitativas

Tabela 3: Distribuição de frequência da qualidade de sêmen de 18 bovinos de uma empresa.

Qualidade Semen	Frequência Absoluta (fa)	Frequência Relativa (fr)	Frequência Percentual (fp)	Frequência Acumulada (FA)	Frequência Percentual Acumulada (FP)
Regular	5	0,28	28%	5	28%
Bom	8	0,44	44%	13	72%
Ótimo	5	0,28	28%	18	100%
Total	18	1,00	100%		

Variáveis Quantitativas

- A tabela de distribuição de frequências de uma variável discreta é, em geral bastante semelhante à das variáveis qualitativas, pois os valores inteiros que a variável assume podem ser considerados como "categorias"
- Quando uma variável discreta tem muitas categorias, o procedimento para construir a tabela de distribuição de frequências, deve ser o mesmo que variáveis contínuas.

Variáveis Quantitativas

- A construção de tabelas de distribuição de frequências para variáveis quantitativas contínuas é feita agrupando os dados em classes e obtendo as frequências observadas em cada classe.
- É importante notar que ao resumir dados referentes a uma variável contínua sempre se perde alguma informação já que não temos ideia de como se distribuem as observações dentro de cada classe.

Variáveis Quantitativas

- Para construir tabelas de distribuição de frequências para variáveis contínuas, envolve os seguintes passos
 - Decidir sobre o número de classes k , entre 5 e 20. O número de classes $k \cong \sqrt{n}$ ou $k = 1 + 3,322\log_{10}(n)$
 - Determinar a amplitude dos dados: $A = \text{Max} - \text{Min}$.
 - Determinar a amplitude de classe h :

$$h = \frac{A}{k - 1}$$

Variáveis Quantitativas

- Determinar o limite inferior da primeira classe LI_1 :

$$LI_1 = Min - \frac{h}{2}$$

- Determinar o limite superior da primeira classe $LS_1 = LI_1 + h$ sendo que o limite inferior da segunda classe $LI_2 = LS_1$, e assim

$$LS_2 = LI_2 + h$$

e assim, sucessivamente

Variáveis Quantitativas

- Dados ordenados, relativos a média diária de produção de leite de 30 vacas da raça holandesa.

6,94	7,27	7,46	7,97	8,03	8,37
8,56	8,66	8,88	8,95	9,30	9,33
9,55	9,76	9,80	9,82	9,98	9,99
10,14	10,19	10,42	10,44	10,66	10,88
10,88	11,16	11,80	11,88	12,25	12,34

Variáveis Quantitativas

- $k = \sqrt{30} = 5,47 \approx 5$
- $A = Max - Min = 12,34 - 6,94 = 5,40$
- $h = \frac{A}{k-1} = \frac{5,40}{4} = 1,35$
- $LI_1 = Min - \frac{h}{2} = 6,94 - \frac{1,35}{2} = 6,94 - 0,67 = 6,27$

Variáveis Quantitativas

Tabela 4: Distribuição de frequência da produção média diária de leite de 30 vacas da raça holandesa.

Classes		Frequência Absoluta (fa)	Frequência Relativa (fr)	Frequência Percentual (fp)
6,27	└	7,62		
7,62	└			
	└			
	└			
	└			

Variáveis Quantitativas

Tabela 4: Distribuição de frequência da produção média diária de leite de 30 vacas da raça holandesa.

Classes		Frequência Absoluta (fa)	Frequência Relativa (fr)	Frequência Percentual (fp)
6,27	├	7,62		
7,62	├	8,97		
8,97	├	10,32		
10,32	├	11,67		
11,67	├	13,02		

Variáveis Quantitativas

Tabela 4: Distribuição de frequência da produção média diária de leite de 30 vacas da raça holandesa.

Classes			Frequência Absoluta (fa)	Frequência Relativa (fr)	Frequência Percentual (fp)	Frequência Acumulada (FA)	Frequência Percentual Acumulada (FP)
6,27	├	7,62	3	0,10	10%	3	10%
7,62	├	8,97	7	0,23	23%	10	33%
8,97	├	10,32	10	0,33	33%	20	67%
10,32	├	11,67	6	0,20	20%	26	87%
11,67	├	13,02	4	0,13	13%	30	100%
			30	1,00	100%		

Gráficos

GRÁFICOS

- Os gráficos são uma forma de apresentação visual dos dados.
- Gráficos fornecem menor grau de detalhes que as tabelas
- Uma representação gráfica coloca em evidência as tendências, as ocorrências ocasionais, os valores mínimos e máximos e também as ordens de grandezas dos fenômenos que estão sendo observados.

Variáveis Qualitativas

- Os gráficos podem ser apresentados por meio de frequências absolutas, relativas ou percentuais.
- Os gráficos mais utilizados são barras, colunas e setores.

Gráficos de barras

- Gráfico de barras - formado por retângulos horizontais.
- O objetivo deste gráfico é de comparar grandezas e é recomendável para variáveis cujas categorias tenham designações extensas.
- Cada barra representa uma categoria ou um atributo
- O tamanho da barra representa a intensidade de uma categoria.

GRÁFICOS PARA VARIÁVEIS QUALITATIVAS

Gráficos de barras

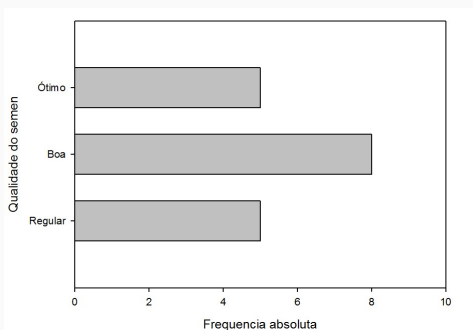


Figura 1: Distribuição de frequência da qualidade de sêmen de 18 bovinos de uma empresa.

Gráficos de colunas

- Gráfico de colunas - Difere do gráfico de barras por serem seus retângulos dispostos verticalmente ao eixo das abscissas sendo mais indicado quando as designações das categorias são breves.

GRÁFICOS PARA VARIÁVEIS QUALITATIVAS

Gráficos de colunas

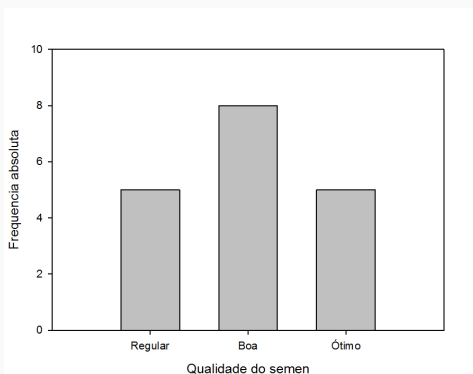


Figura 2: Distribuição de frequência da qualidade de sêmen de 18 bovinos de uma empresa.

Gráficos de setores

- Gráfico de setores - a variável em estudo é projetada num círculo, de raio arbitrário, dividido em setores com áreas proporcionais às frequências das suas categorias.
- Este tipo de gráfico deve ser utilizado apenas quando tem poucas categorias.
- Cada fatia do gráfico é proporcional à porcentagem correspondente àquela categoria. Por exemplo, se uma determinada categoria representa 40% do total o ângulo da fatia será de $0,4 \times 360^\circ = 144^\circ$.

GRÁFICOS PARA VARIÁVEIS QUALITATIVAS

Gráficos de setores

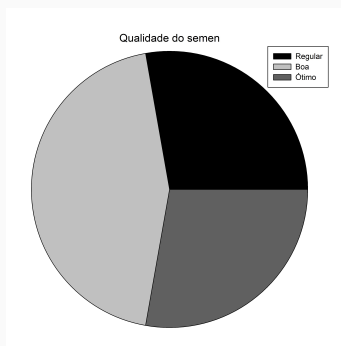


Figura 3: Distribuição de frequência da qualidade de sêmen de 18 bovinos de uma empresa.

Gráficos temporais

- Muitas variáveis são medidas em intervalos ao longo do tempo.
- Neste caso podemos colocar o tempo no eixo x e os valores da variável no eixo y .
- Este gráfico tem o objetivo de avaliar as mudanças da variável ao longo do tempo.

Gráficos temporais

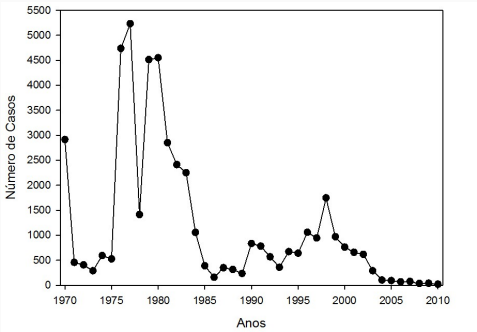


Figura 4: Número de casos de raiva canina por ano no Brasil no período 1970 a 2010.

Histograma

- Uma forma de representar graficamente a distribuição de frequência das variáveis contínuas é por meio do histograma.
- O histograma é semelhante ao gráfico de barras verticais, no eixo vertical pode-se utilizar as frequências ou densidades de frequências e no eixo horizontal as classes.
- É uma versão gráfica de uma tabela de frequências

Histograma

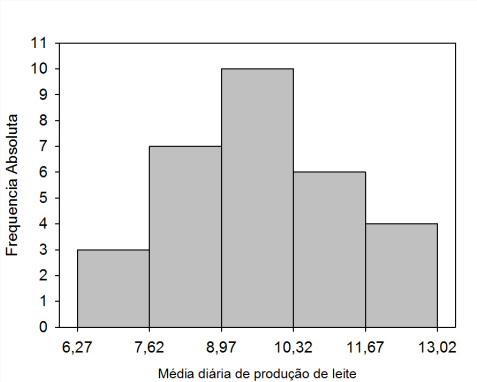


Figura 5: Distribuição de frequência da produção média diária de leite de 30 vacas da raça holandesa.

Medidas Descriptivas

MEDIDAS DESCRITIVAS

- As medidas descritivas são resumos numéricos de variáveis quantitativas e podem ser divididas em:
 - Medidas de Posição - São medidas de tendência central, ou seja, representativas do valor central, ao redor do qual se agrupam a maioria dos valores.
 - Média, Mediana, Moda
 - Medidas de dispersão - são medidas para medir a variabilidade dos dados.
 - Variância, Desvio Padrão, Coeficiente de Variação

MEDIDAS DE POSIÇÃO

Média

- A média de uma população ou amostra é a soma de todos os elementos da população (amostra) dividida pelo número de elementos. Esta medida apresenta a mesma unidade dos dados.
- Para a população a média é representada por

$$\mu = \frac{\sum_{i=1}^N x_i}{N}$$

em que N é o tamanho da população

- Para a amostra a média é representada por

$$\bar{X} = \frac{\sum_{i=1}^n x_i}{n}$$

em que n é o tamanho da amostra.

MEDIDAS DE POSIÇÃO

Média

- Considere o conjunto de dados 612, 983, 623, 883, 666 , 970.
- A média é dado por:

$$\begin{aligned}\bar{X} &= \frac{\sum_{i=1}^n x_i}{n} = \frac{612 + 983 + 623 + 883 + 666 + 970}{6} \\ &= \frac{4737}{6} = 789,5\end{aligned}$$

Mediana

- Num conjunto de dados ordenados, a mediana (M_d) é o valor que deixa metade da frequência abaixo dele. A mediana, como a média, possui a mesma unidade de cada observação.
- Depois de ordenar os dados, a mediana pode ser obtida por meio da expressão,

$$M_d = \begin{cases} X_{\frac{n+1}{2}}, & \text{se } n \text{ for ímpar;} \\ \frac{X_{\frac{n}{2}} + X_{\frac{n+2}{2}}}{2}, & \text{se } n \text{ for par} \end{cases}$$

Mediana

- Considere o conjunto de dados: 5, 2, 6, 13, 9, 15, 10.
- Primeiro é necessário ordenar os dados: 2, 5, 6, 9, 10, 13, 15. Como se de uma conjunto com $n = 7$ (ímpar), então:

$$Md = X_{\frac{n+1}{2}} = X_{\frac{7+1}{2}} = X_4$$

- Logo a Mediana é igual ao elemento que está na quarta posição do conjunto de dados, assim

$$Md = 9$$

Mediana

- Considere o conjunto de dados ordenados: 1, 2, 3, 4, 6, 8.
Como se de uma conjunto com $n = 6$ (par), então

$$Md = \frac{X_{\frac{n}{2}} + X_{\frac{n+2}{2}}}{2} = \frac{X_{\frac{6}{2}} + X_{\frac{6+2}{2}}}{2} = \frac{X_3 + X_4}{2}$$

- Logo para obter a mediana é necessário obter os elementos que estão na terceira e quarta posição do conjunto de dados, assim:

$$Md = \frac{3 + 4}{2} = 3,5$$

Moda

- A moda M_o de um conjunto de dados é o valor mais frequente
- Para obter a moda basta observar qual o dado que mais se repete.
- No conjunto de dados 3 , 5 , 8 , 10 , 12 não apresenta moda. O conjunto é amodal
- No conjunto de dados 7 , 8 , 9 , 10 , 10 , 10 , 11 , 12 a moda é igual a 10, pois é único que se repete. Conjunto unimodal.
- No conjunto de dados 2 , 3 , 4 , 4 , 4 , 5 , 6 , 7 , 7 , 7 , 8 , 9 temos duas modas: 4 e 7. O conjunto é bimodal.

Comparação entre Média, Mediana e Moda

- Média
 - Definição: Soma de todos os valores dividido pelo total de elementos do conjunto.
 - Vantagens: Reflete cada valor; possui boas propriedades matemáticas.
 - Limitações: É influenciada por valores externos.
 - Quando usar:
 1. Deseja-se obter a medida de posição que possui a maior estabilidade;
 2. Houver necessidade de um tratamento algébrico posterior.

Comparação entre Média, Mediana e Moda

- Mediana
 - Definição: Valor que divide o conjunto em duas partes iguais.
 - Vantagens: Menos sensível a valores extremos que a média.
 - Limitações: Difícil de determinar para grande quantidade de dados
 - Quando usar:
 1. Deseja-se obter o ponto que divide o conjunto em partes iguais;
 2. Há valores extremos que afetam de maneira acentuada a média;

Comparação entre Média, Mediana e Moda

- Moda
 - Definição: Valor mais frequente.
 - Vantagens: Valor "típico"; maior quantidade de valores concentrados neste ponto
 - Limitações: Não se presta a análise matemática; Pode não haver moda para certos conjuntos de dados
 - Quando usar:
 1. Deseja-se obter uma medida rápida e aproximada da posição;
 2. A medida de posição deve ser o valor mais típico da distribuição.

Simetria

- A determinação das medidas de posição permite discutir sobre a simetria da distribuição dos dados.
 - Distribuição simétrica: $\bar{X} = M_d = M_o$
 - Distribuição assimétrica: ocorrem diferenças entre os valores da média, mediana e moda. A assimetria pode ser:
 - à direita: $\bar{X} > M_d > M_o$
 - à esquerda: $\bar{X} < M_d < M_o$

MEDIDAS DE DISPERSÃO

- As medidas de posição são importantes para caracterizar um conjunto de dados, mas não são suficientes para caracterizar completamente a distribuição dos dados.
- Considere as amostras as duas amostras abaixo
 - Grupo A: 185 185 185
 - Grupo B: 187 183 185
- A média para os dois grupos é a mesma $\bar{X}_A = 185$ e $\bar{X}_B = 185$.

MEDIDAS DE DISPERSÃO

- Os 2 conjuntos não diferem entre si e consideramos somente a média. Entretanto o grupo A tem todas as observações iguais a média. Já no grupo B ocorre uma certa dispersão nos dados.
- As medidas de variabilidade ou dispersão possibilitam que façamos distinção entre os conjuntos quanto à sua homogeneidade, isto é, o grau de concentração em torno de uma medida de tendência central.

MEDIDAS DE DISPERSÃO

- O conceito de variabilidade é essencial a Estatística
- Variabilidade significa que alguma coisa tem propensão para variar ou mudar.
- Qualquer que seja o fenômeno que se tenha interesse em analisar terá variabilidade.

MEDIDAS DE DISPERSÃO

- Variabilidade natural - é inerente ao indivíduo. Refere-se as diferenças encontradas entre vários indivíduos ou entre o mesmo indivíduo ao longo do tempo
- Variabilidade sistemática - é ocasionado por causas atribuíveis, como por exemplo, máquinas desajustadas, erros do operador, matéria prima com defeito.

MEDIDAS DE DISPERSÃO

Variância e Desvio Padrão

- A variância é baseada pelo quadrado dos desvios dos dados em relação à média.
- Para a população a variância é representada por

$$\sigma^2 = \frac{\sum_{i=1}^N (x_i - \mu)^2}{N}$$

em que N é o tamanho da população

- Para a amostra a variância é representada por

$$S^2 = \frac{\sum_{i=1}^n (x_i - \bar{X})^2}{n - 1}$$

em que n é o tamanho da população

Variância e Desvio Padrão

- O desvio padrão é a raiz quadrada positiva da variância. Esta medida é expressa na mesma unidade dos dados.
- Para a população o desvio padrão é representada por

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

- Para a amostra o desvio padrão é representada por

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

em que n é o tamanho da população

MEDIDAS DE DISPERSÃO

Variância e Desvio Padrão

- Considere o conjunto de dados 612, 983, 623, 883, 666 , 970, temos que $\bar{X} = 789,5$
- Como se trata de uma amostra,temos:

$$\begin{aligned} S^2 &= \frac{\sum_{i=1}^n (x_i - \bar{X})^2}{n - 1} \\ &= \frac{(612 - 789,5)^2 + (983 - 789,5)^2 + (623 - 789,5)^2 + (883 - 789,5)^2 + (666 - 789,5)^2 + (970 - 789,5)^2}{6 - 1} \\ &= \frac{(-177,5)^2 + (193,5)^2 + (-166,5)^2 + (93,5)^2 + (-123,5)^2 + (180,5)^2}{5} \\ &= \frac{31506,25 + 37442,25 + 27722,25 + 8742,25 + 15252,25 + 32580,25}{5} \\ &= \frac{153245,5}{5} = 30649,1 \end{aligned}$$

MEDIDAS DE DISPERSÃO

Coeficiente de Variação

- O coeficiente de variação (CV) é uma medida de dispersão que expressa o desvio padrão em termos da média de forma percentual

$$CV = 100 \frac{S}{\bar{X}}$$

- Se as amostras tiverem unidade diferentes ou médias diferentes o CV pode ser utilizado para comparar a variabilidade entre duas amostras.
- Supondo uma amostra com $\bar{X} = 20$ e $S^2 = 225$, qual o coeficiente de variação:

$$CV = 100 \frac{S}{\bar{X}} = 100 \frac{\sqrt{225}}{20} = 75\%$$

Boxplot

BOXPLOT

- O gráfico Boxplot (ou desenho esquemático) é uma análise gráfica que oferece a idéia da posição, dispersão, assimetria, caudas e dados discrepantes.
- O boxplot é um gráfico que possibilita representar a distribuição de um conjunto de dados com base em alguns de seus parâmetros descritivos:
 - A mediana (Md ou Q_2),
 - O primeiro quartil (Q_1) - valor que deixa 25% dos dados abaixo dele
 - O terceiro quartil (Q_3) - valor que deixa 75% dos dados abaixo dele
- A principal função do boxplot é identificar a presença de outliers (valores discrepantes) no conjunto de dados.

BOXPLOT

- Para construir um boxplot desenhamos uma "caixa" com o nível superior dado pelo terceiro quartil (Q_3) e o nível inferior pelo primeiro quartil (Q_1). A mediana (Q_2) é representada por um traço no interior da caixa e segmentos de reta são colocados da caixa até dos limites inferior (LI) e superior (LS), dados por

$$LI = Q_1 - 1.5dq$$

$$LS = Q_3 + 1.5dq$$

em que $dq = Q_3 - Q_1$ denominando diferença quartílica.

BOXPLOT

- Para traçarmos o boxplot utilizamos as seguintes etapas:
 - Construir um retângulo de tal maneira que suas bases têm alturas correspondentes aos primeiro e terceiro quartis da distribuição.
 - Cortar o retângulo por um segmento paralelo às bases, na altura correspondente à mediana;
 - Traçar um segmento paralelo ao eixo, partindo do ponto médio da base superior do retângulo até o maior valor observado que NÃO supere LS;
 - Traçar um segmento paralelo ao eixo, partindo do ponto médio da base inferior do retângulo, até o menor valor que NÃO é menor LI;
 - Caso tenha valores que superior a LS ou inferior a LI, marcar os pontos, estes valores são considerados observações discrepantes.
 - Podemos opcionalmente marcar o valor da média;

BOXPLOT

- Para um conjunto de dados com as seguintes medidas:

$$Md = 9,81$$

$$Q_1 = 8,71$$

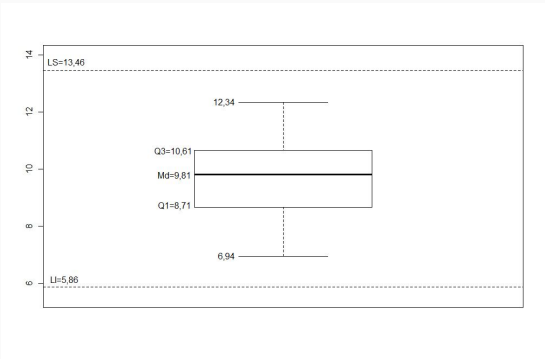
$$Q_3 = 10,61$$

$$dq = 10,61 - 8,71 = 1,9$$

$$LI = 8,71 - 1,5 \times 1,9 = 5,86$$

$$LS = 10,61 + 1,5 \times 1,9 = 13,46$$

BOXPLOT



R Commander

Estatística Descritiva

- Para fazer a estatística descrita no R, usaremos os dados do arquivo **anemia** que contém os dados de um estudo a respeito anemia infecciosa em 1100 equinos
 - Fazenda - localidade do animal
 - Função - 0 apenas serviço e 1 serviço e exposição
 - Sexo - 0 Feminino e 1 Masculino
 - Idade - idade dos animais em meses
 - Carrapato - tem carrapatos 0 Não e 1 Sim
 - IDGA - diagnóstico para anemia infecciosa sendo 0 negativo e 1 positivo
 - Doente - situação real do animal para anemia infecciosa sendo 0 não doente e 1 doente

Estatística Descritiva

- É possível obter uma análise preliminar dos dados utilizando menu Estatísticas
 - Estatísticas $\Rightarrow$ Resumos $\Rightarrow$
Conjunto de dados ativo
- Este menu irá retornar
 - Variáveis qualitativas - frequências absolutas;
 - Variáveis quantitativas - menor valor, primeiro quartil, mediana, média, terceiro quartil e maior valor.

Tabelas

- Para construir tabelas de variáveis qualitativas utiliza-se o menu Estatísticas
 - Estatísticas ⇒ Resumos ⇒
Distribuição de Frequência

Tabelas

- Para construir tabelas de variáveis quantitativas é necessário utilizar o comando
- Variáveis Quantitativas Contínuas - `fdt` pacote `fdth`
`fdt(x, k, start, end, h, breaks=c("Sturges", "Scott", "FD"))` em que:
 - `x` - são os dados que devem agrupados;
 - `k` - numero de intervalos de classe;
 - `start` - limite inferior da primeira classe;
 - `end` - limite superior da ultima classe;
 - `h` - amplitude da classe;
 - `breaks` - caso não seja definido `k`, `start`, `end` e `h`, pode-se utilizar um método para isso.

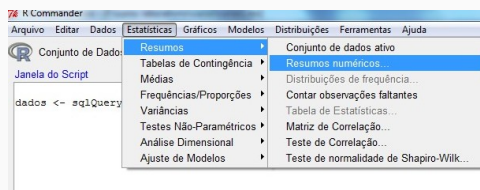
Gráficos

- Os gráfico podem ser feitos menu Gráficos.
- Os gráficos podem ser alterados utilizando alguns comandos
 - `main` para adicionar titulo;
 - `xlab` titulo para o eixo x;
 - `ylab` titulo para o eixo y;
 - `xlim` delimita os valores de x;
 - `ylim` delimita os valores de y;
 - `col` para definir cor;
 - `horiz` para definir se as barras são horizontais ou verticais.

Resumo estatístico

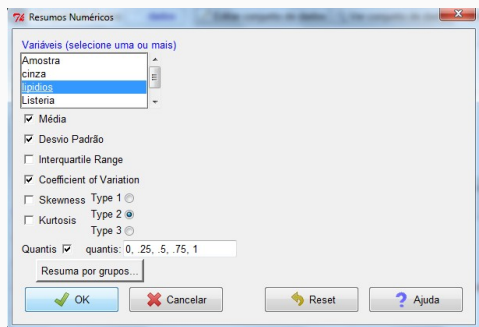
- O resumo estatístico apresenta as principais medidas de posição e dispersão, para isso utiliza-se o menu Estatísticas

- Estatísticas ⇒ Resumos ⇒ Resumos Numéricos



Resumo estatístico

- Na janela que se abre seleciona a variável e pode-se escolher as medidas a ser calculadas.



Resumo estatístico

- Para se calcular a moda deve-se utilizar o pacote `modeest`.
 - Variável discreta - função `mfv`
 - Variável contínua por diferentes métodos - função `mlv`