



Departamento de Estatística
UNIVERSIDADE FEDERAL DE JUIZ DE FORA



Inferência Estatística

Departamento de Estatística

Parâmetros, Estimadores e Estimativas

▶ Parâmetro:

- ▶ As quantidades da população, em geral desconhecidas, sobre as quais temos interesse, são denominadas parâmetros e, usualmente representadas por letras gregas tais como θ , μ e σ , entre outras.

▶ Estimador e estimativa:

- ▶ À combinação dos elementos da amostra, construída com a finalidade de representar, ou estimar, um parâmetro de interesse na população, denominamos de *estimador*. Em geral, denotamos os estimadores por símbolos com o acento circunflexo: $\hat{\theta}$, $\hat{\mu}$, $\hat{\sigma}$, etc.
- ▶ Aos valores numéricos assumidos pelos estimadores denominamos *estimativas*.



Departamento de Estatística

UNIVERSIDADE FEDERAL DE JUIZ DE FORA



Inferência Estatística

- ▶ Vamos considerar o conjunto de dados relativos aos pesos obtido em entrevistas com 898 alunos da UFJF;
- ▶ Com o objetivo puramente ilustrativo, considere que esses 898 alunos componham nossa população alvo;
- ▶ Nesse caso conhecemos a média e variância da população (parâmetros de interesse):
 - ▶ $\mu = 63,42$ kg
 - ▶ $\sigma^2 = 156,03$ kg²
 - ▶ $\sigma = 12,49$ kg
- ▶ Vamos retirar 5 amostras aleatórias de tamanho 30 dessa população e anotar os resultados obtidos para a variável peso.



Departamento de Estatística

UNIVERSIDADE FEDERAL DE JUIZ DE FORA



Exemplo

- ▶ Vamos representar uma amostra de tamanho n , a ser retirada de uma população, por (X_1, X_2, \dots, X_n) .
- ▶ Amostra 1:
 - ▶ (75; 46,5; 62; 56; 42; 45; 64; 54; 60; 80 ...)
- ▶ Amostra 2:
 - ▶ (84; 56; 75; 78; 49; 58; 72; 68; 59; 54; ...)
- ▶ Amostra 3:
 - ▶ (60; 62; 78; 95; 55; 73; 76; 72; 50; 56; ...)
- ▶ Amostra 4:
 - ▶ (55; 69; 55; 56; 58; 95; 46; 69,5; 60; 54; ...)
- ▶ Amostra 5:
 - ▶ (62; 53; 53; 80; 75; 62; 55; 50; 52; 74; ...)



Exemplo

- ▶ Poderíamos utilizar, como estimativa da média populacional ($\mu = 63,42$ kg), os valores da média amostral ou da mediana amostral, por exemplo:
- ▶ **Amostra 1:**
 - ▶ $\hat{\mu}_1 = \bar{x}_1 = 62,15$ kg; ou
 - ▶ $\hat{\mu}_2 = md_1 = 62,5$ kg
- ▶ **Amostra 2:**
 - ▶ $\hat{\mu}_3 = \bar{x}_2 = 64,7$ kg; ou
 - ▶ $\hat{\mu}_4 = md_2 = 63,5$ kg
- ▶ **Amostra 3:**
 - ▶ $\hat{\mu}_5 = \bar{x}_3 = 63,23$ kg; ou
 - ▶ $\hat{\mu}_6 = md_3 = 60$ kg

Amostra 4:

$$\hat{\mu}_7 = \bar{x}_4 = 61,48 \text{ kg; ou}$$

$$\hat{\mu}_8 = md_4 = 59 \text{ kg}$$

Amostra 5:

$$\hat{\mu}_9 = \bar{x}_5 = 62,9 \text{ kg; ou}$$

$$\hat{\mu}_{10} = md_5 = 62 \text{ kg}$$



Inferência Estatística

- ▶ Como escolher qual estimador utilizar?
- ▶ Deve-se estudar as propriedades de um estimador.



Departamento de Estatística

UNIVERSIDADE FEDERAL DE JUIZ DE FORA



Propriedades dos Estimadores

▶ Vício

- ▶ Um estimador $\hat{\theta}$ é *não viciado* ou *não viesado* para um parâmetro θ se $E(\hat{\theta}) = \theta$. Ou seja, um estimador é não viciado se o seu valor esperado coincide com o parâmetro de interesse.

▶ Consistência

- ▶ Um estimador $\hat{\theta}$ é *consistente*, se, à medida que o tamanho da amostra aumenta, seu valor esperado converge para o parâmetro de interesse e sua variância converge para zero:

- ▶ $\lim_{n \rightarrow \infty} E(\hat{\theta}) = \theta$;
- ▶ $\lim_{n \rightarrow \infty} Var(\hat{\theta}) = 0$.

▶ Eficiência

- ▶ Dados dois estimadores $\hat{\theta}_1$ e $\hat{\theta}_2$, não viciados para um parâmetro θ , dizemos que $\hat{\theta}_1$ é mais eficiente do que $\hat{\theta}_2$ se $Var(\hat{\theta}_1) < Var(\hat{\theta}_2)$.



Estimadores para média, proporção e variância

Parâmetro	Estimador	Propriedades
μ	$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$	Não viciado e consistente
π	$\hat{\pi} = \frac{\text{frequência amostral com a característica}}{n}$	Não viciado e consistente
σ^2	$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$	Não viciado e consistente



Departamento de Estatística

UNIVERSIDADE FEDERAL DE JUIZ DE FORA



Distribuições Amostras

- ▶ Estimadores são funções de variáveis aleatórias, sendo assim, eles também são variáveis aleatórias;
- ▶ Vamos estudar a distribuição de probabilidade de um dos estimadores mais utilizados:
 - ▶ A média amostral



Departamento de Estatística

UNIVERSIDADE FEDERAL DE JUIZ DE FORA



Distribuição da Média Amostral

- ▶ Considere X como a variável aleatória que representa o peso (em kg) dos 898 alunos da UFJF que consideramos como nossa população alvo na aula passada;
- ▶ Sabemos que X tem média igual a 63,42 kg e variância igual a 156,03 kg²;
- ▶ Se forem retiradas 50 amostras dessa população, e calculadas as médias para cada amostra, teremos 50 valores, possivelmente, distintos que poderiam ser considerados como estimativas da média da população;
- ▶ $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_{50}$;
- ▶ Podemos considerar as médias obtidas como uma nova variável aleatória \bar{X} .



Departamento de Estatística

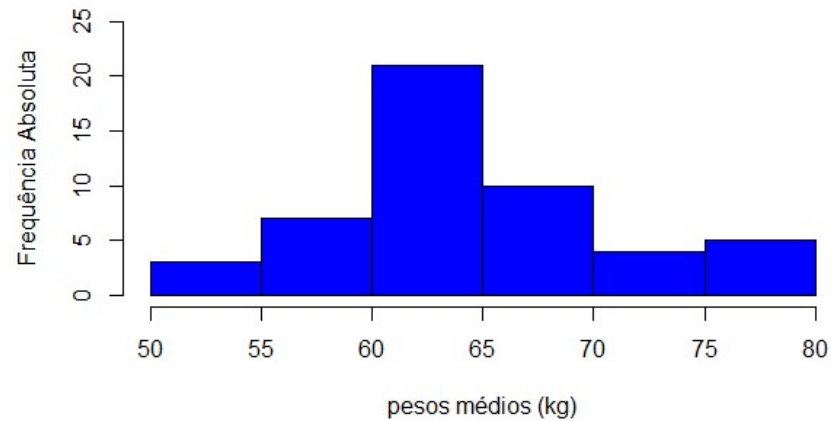
UNIVERSIDADE FEDERAL DE JUIZ DE FORA



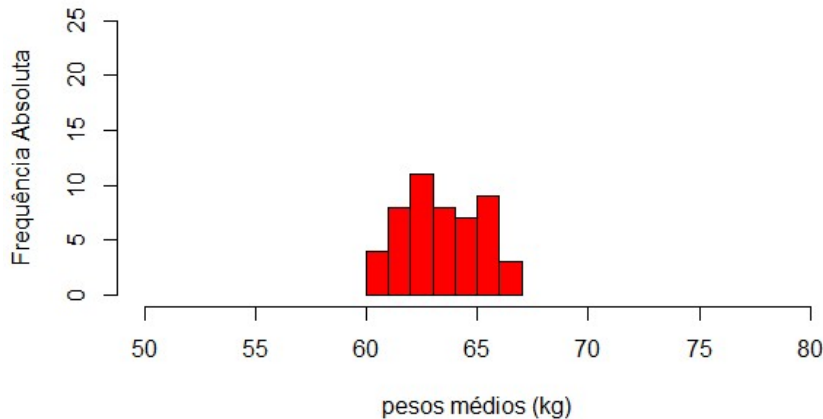
Distribuição da Média Amostral

- ▶ Histogramas das médias das idades obtidas para 50 amostras aleatórias retiradas da população de 898 alunos ($\mu = 63,42$ kg). Considerando tamanhos amostrais de:

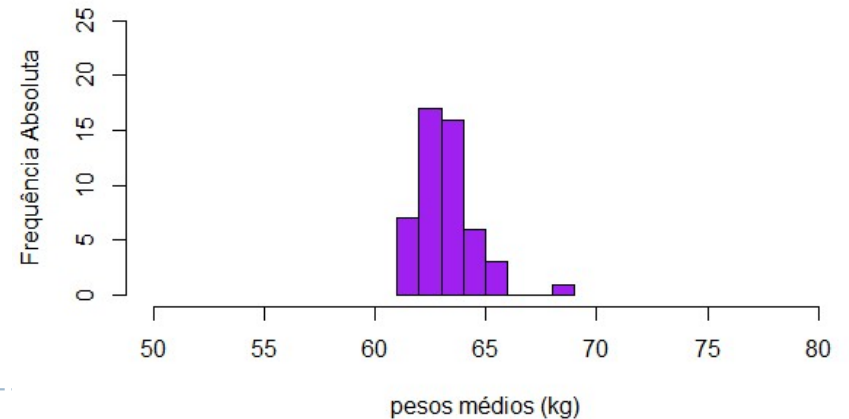
n = 5



n = 50



n = 100



Teorema Central do Limite

- ▶ Suponha uma amostra aleatória simples de tamanho n retirada de uma população com média μ e variância σ^2 (note que a distribuição de probabilidade da variável aleatória não é especificada). Representando tal amostra por n variáveis aleatórias independentes (X_1, X_2, \dots, X_n) e, denotando sua média por \bar{X} , temos que:

$$\frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \xrightarrow{n \rightarrow \infty} Z$$

- ▶ com $Z \sim N(0,1)$.



Departamento de Estatística

UNIVERSIDADE FEDERAL DE JUIZ DE FORA



Teorema Central do Limite

- ▶ Garante que para n grande a distribuição da média amostral, devidamente padronizada, segue uma distribuição Normal Padrão;
- ▶ Esse teorema permite que utilizemos a distribuição Normal para estudar \bar{X} probabilisticamente;
- ▶ Estudos, envolvendo simulações, mostram que, em muitos casos, valores de n ao redor de 30 fornecem aproximações bastante boas para aplicações práticas.



Departamento de Estatística

UNIVERSIDADE FEDERAL DE JUIZ DE FORA



Exercício 1

- ▶ O fabricante de uma lâmpada afirma que seu produto tem vida média de 1600 horas, com desvio padrão de 250 horas. O dono de uma empresa compra 100 lâmpadas desse fabricante. Qual é a probabilidade de que a vida média dessas lâmpadas ultrapasse 1650 horas?



Exercício 2

- ▶ Sabe-se que a altura dos alunos da UFJF tem média igual a 1,60 m e variância de $0,02 \text{ m}^2$. Assumindo que os dados segue a distribuição normal, que foi selecionado uma amostras de tamanho 5 para o estudo da distribuição amostral da média. Determine:
 - ▶ A distribuição de probabilidade da altura dos alunos da UFJF.
 - ▶ A distribuição de probabilidade amostral da média das alturas dos alunos da UFJF.
 - ▶ A probabilidade da altura de um aluno selecionado aleatoriamente seja maior que 1,50.
 - ▶ A probabilidade da média das altura dos alunos da UFJF ser menor que 1,89.
-



Teorema Central do Limite - Aplicação

- ▶ Suponha que:
- ▶ π represente a proporção de indivíduos com determinada característica em uma população (valor desconhecido)
- ▶ $\hat{\pi}$ represente o estimador dessa proporção, considerando uma amostra de tamanho n , dado por:

$$\hat{\pi} = \frac{\text{número de indiv. na amostra com dada característica}}{n}$$

- ▶ Tem-se que:
- ▶ $E(\hat{\pi}) = \pi$ e $Var(\hat{\pi}) = \frac{\pi(1-\pi)}{n}$
- ▶ Considerando o Teorema Central do Limite, tem-se que para n suficientemente grande:

$$\frac{\hat{\pi} - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}} \xrightarrow{n \rightarrow \infty} N(0,1)$$



Exercício 3

- ▶ Uma grande empresa, que possui 2.500 funcionários, sorteará 100 funcionários para participarem de um processo de capacitação. Sabe-se que na empresa há 1800 homens e 700 mulheres. Qual a probabilidade da proporção de mulheres na amostra estar entre 30% e 40%?



Exemplo

$X = \left\{ \text{concentração em } \frac{\text{unidades}}{\text{ml}} \text{ de determinada vitamina no sangue} \right\}$

$$X_S \sim N(14, 6^2)$$

$$X_D \sim N(10, 6^2)$$



$$X_E \sim N(18, 6^2)$$

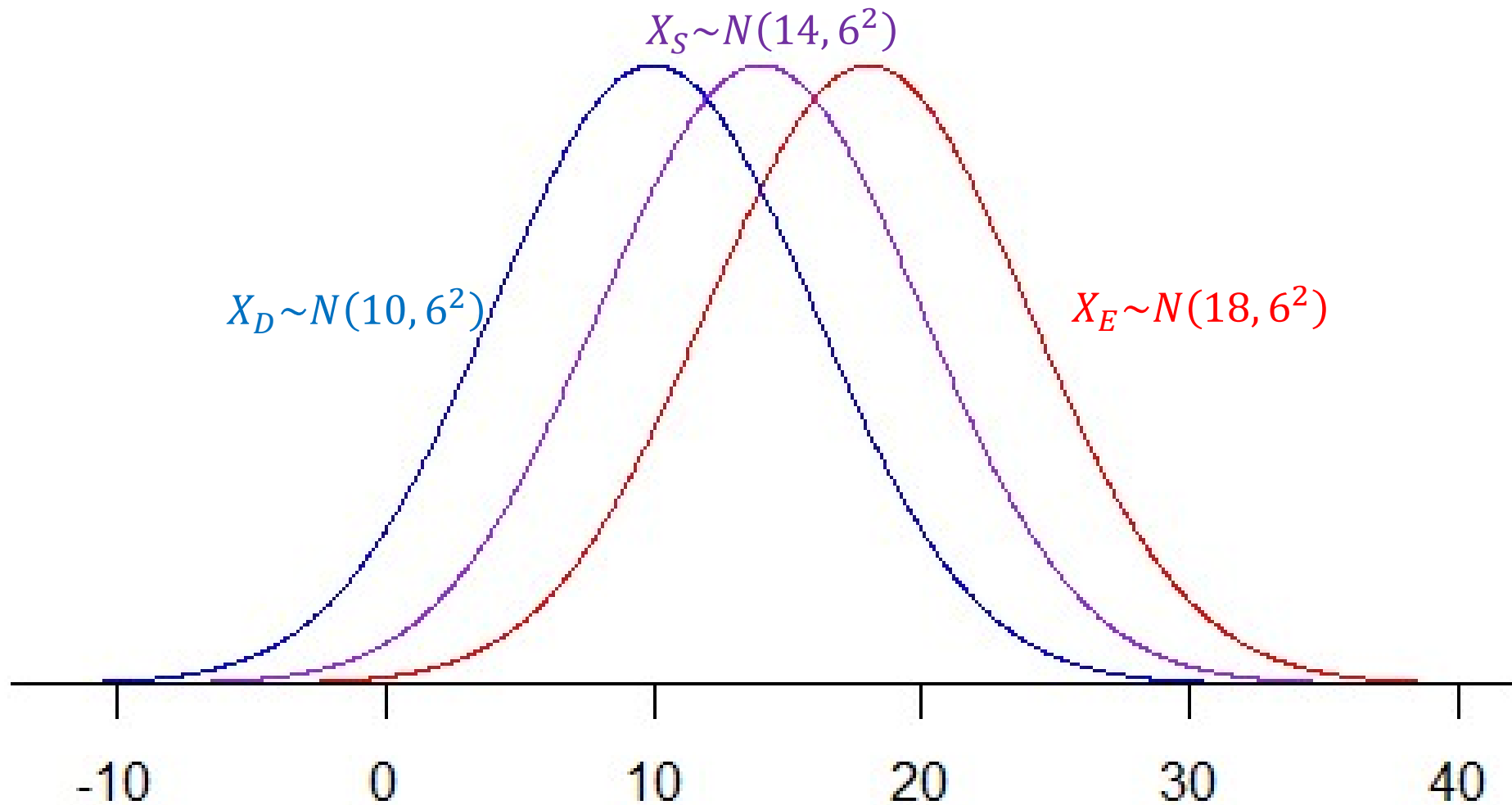


Departamento de Estatística

UNIVERSIDADE FEDERAL DE JUIZ DE FORA



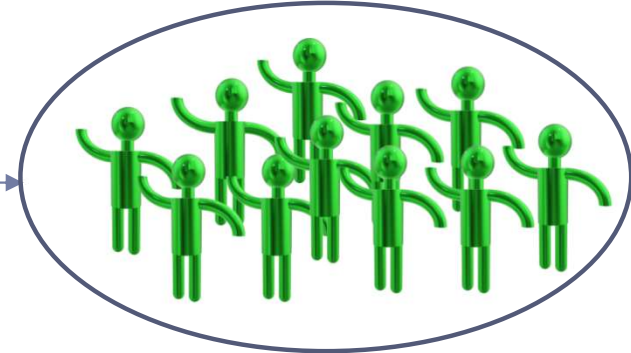
Exemplo



Exemplo



$n = 30$



$(X_1, X_2, \dots, X_{30})$

\bar{x}_{obs}

$\mu = 10$

$\mu = 14$

$\mu = 18$

A média amostral pode ser utilizada para fazer inferência sobre o valor μ da população

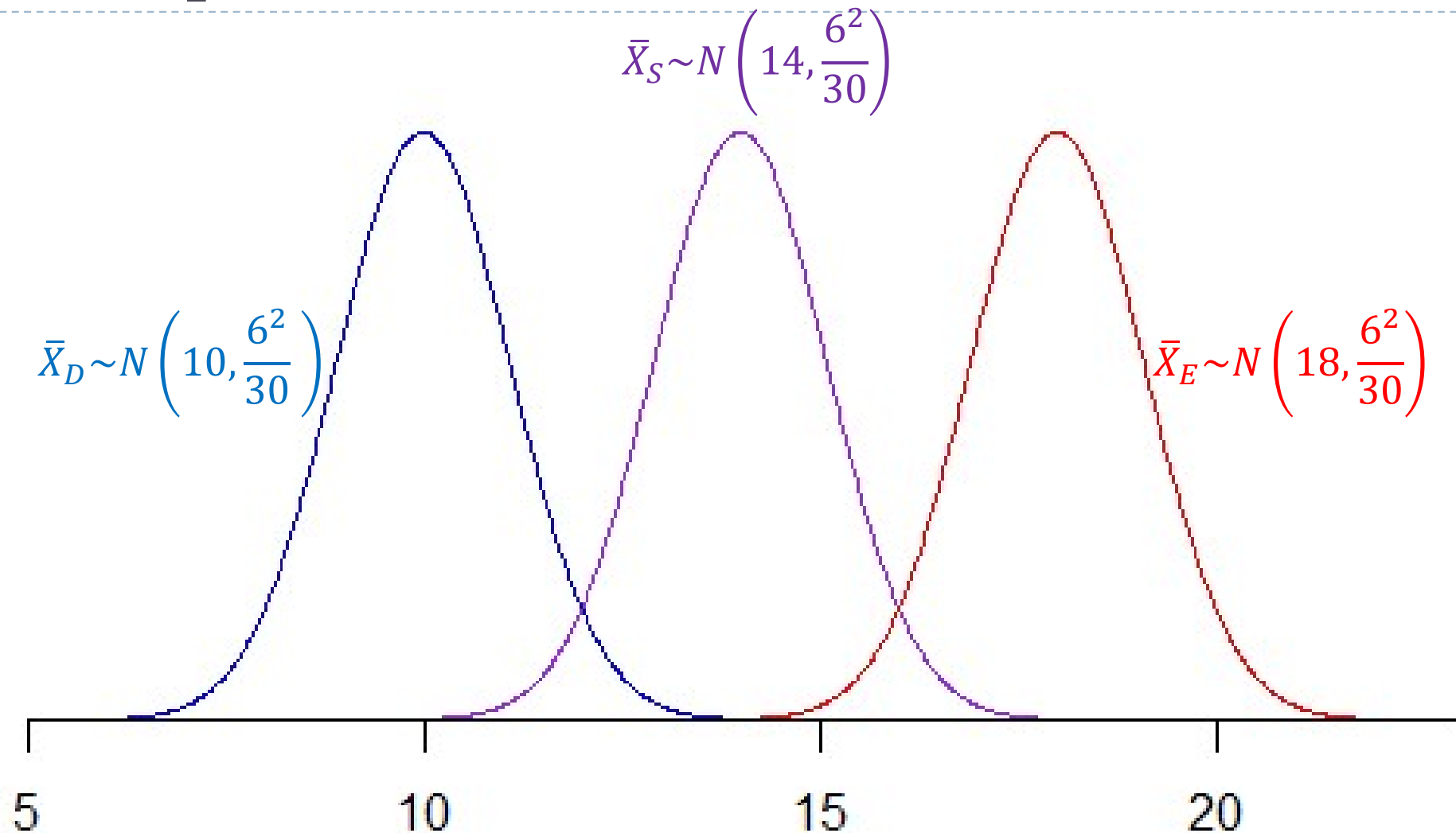


Departamento de Estatística

UNIVERSIDADE FEDERAL DE JUIZ DE FORA



Exemplo



Departamento de Estatística

UNIVERSIDADE FEDERAL DE JUIZ DE FORA

