



Universidade Federal do Ceará

Faculdade de Economia

Métodos Quantitativos

Vicente Lima Crisóstomo

Fortaleza, 2015

Amostragem

■ Amostragem

- Técnica de pesquisa na qual um sistema preestabelecido de amostras é considerado idôneo para representar o universo pesquisado, com margem de erro aceitável.

■ Inferência Estatística

- Envolve formulação de juízo sobre um todo a partir do exame de um subconjunto, ou amostra, deste todo.

Sumário

- Introdução
- Estatística Descritiva
- Probabilidade
- Distribuições de Probabilidades
- Amostragem e Distribuições Amostrais
- Estimação
- Testes de Significância
- Análise de Variância
- Teste de Significância para Proporções
- Testes Não Paramétricos
- Correlação e Regressão

Amostragem

■ Exemplos de Amostragem do “cotidiano” e outros

- Provar um pouco da comida
- Ouvir uma fração de uma música
- Assistir uma fração/parte de um filme
- Ler uma parcela de um livro
- Provar um pouco de certo vinho
- Fazer um teste de um carro
- Hospedar-se alguns dias numa cidade
- Conhecer alguns moradores da cidade para conhecer os costumes do local
- Conhecer algumas empresas para saber sobre o conjunto delas

Amostragem

- População
 - Universo completo de estudo
- Censo
 - Exame de todos os elementos do universo, ou população
- Amostragem
 - Exame de uma parcela, ou fração, da população

Amostragem

- Finalidade
 - Permitir fazer-se generalizações sobre uma população sem necessitar realizar-se um censo.
- Importante
 - Obter a amostra que melhor represente a população
- É importante definir-se claramente
 - População e amostra em cada estudo
- População pode ser, quanto ao tamanho:
 - Limitada ou finita
 - Ilimitada ou infinita

Amostragem

- População Limitada ou finita
 - Produtos disponíveis em um estabelecimento comercial
 - Alunos do curso T da universidade Y
 - Automóveis da cidade de Goiânia
 - Eleitores da cidade de Manaus
 - Eleitores do Brasil

Amostragem

- População ilimitada ou infinita
 - Resultados do lance de um dado
 - Resultados do lance de uma moeda
 - Produção de uma indústria ou de um setor
 - Nascimentos de indivíduos de certa espécie
 - Peixes no oceano Atlântico

Amostragem

■ Amostragem X Censo

- Censo seria o ideal mas nem sempre possível
- Censo viável quando
 - População muito pequena
 - Um grupo de 20 pessoas reunidas
 - Grande tamanho da amostra em relação à população
 - Uma amostra com grande desvio padrão com tamanho pouco inferior ao da população
 - Precisão completa exigida
 - Levantamento do faturamento de produtos para revenda em uma loja
 - Situações nas quais se dispõe de toda a informação

Amostragem

■ Amostragem X Censo

- Censo seria o ideal mas nem sempre possível
 - Censo inviável para **população “infinita”**
 - Habitantes de um país
 - Animais de certa espécie na floresta
 - Censo incompatível com **limitação temporal**
 - Coleta de dados individuais implicará muito tempo para grande amostra
 - Censo inviável para **testes destrutivos**
 - Teste de produtos industriais: lâmpadas, alimentos
 - Censo para grande população pode ter **custo** proibitivo
 - Pesagem e mensuração de animais
 - Censo pode ter menor **precisão** na coleta de dados
 - Grande quantidade de indivíduos observados pode depreciar precisão dos dados

Amostragem

■ Plano de Amostragem

- Estratégia para obtenção da amostra ideal, ou, a amostra que melhor represente a população objeto de estudo

■ Amostragem Aleatória

- Cada indivíduo da população tem igual probabilidade de compor a amostra

Amostragem

■ Amostragem Aleatória

- Cada indivíduo da população tem igual probabilidade de compor a amostra
- Populações com variável discreta
 - Cada valor tem a mesma chance de compor a amostra
- Populações com variável contínua
 - A probabilidade de incluir na amostra qualquer intervalo de valores é igual à percentagem da população que se encontra naquele intervalo

Amostragem

■ Amostragem Aleatória

- O processo de escolha de indivíduos é aleatório
- Assim a amostra será aleatória
 - Cada indivíduo tem a mesma possibilidade de figurar na amostra a cada prova (teste): $(1 / n)$
 - Grupos de indivíduos têm a mesma chance que outros do mesmo tamanho de figurar na amostra
 - Grupos maiores têm maior probabilidade de compor a amostra

Amostragem

■ Amostragem Aleatória

- Há processos naturalmente aleatórios
 - Chegada de clientes a uma loja
 - Chamadas telefônicas a um help-desk
 - Número de mensagens eletrônicas que chegam por dia
 - Compra de produtos em uma loja virtual
- Outros casos
 - Conhece-se o universo de pesquisa
 - Pode-se enumerar os indivíduos
 - Escolha aleatória de componentes da amostra

Amostragem

■ Amostragem Aleatória

- Probabilística
- Sempre mais recomendável

■ Amostragem Não Probabilística

- Subjetiva
- Por julgamento
 - Julga-se *a priori*
 - os elementos mais representativos
 - indivíduos que devem ser representados na amostra
 - Conhece-se bem a população e quer-se ter elementos representativos de cada sub grupo
 - Populações e/ou amostras pequenas

Amostragem

■ Amostragem Não Probabilística

- Por julgamento
 - Exemplos
 - Matriz de uma empresa pesquisando filiais
 - Pesquisa sobre determinada terapia médica
 - Pesquisa de um novo medicamento
 - Pesquisa em um determinado grupo de pessoas
 - Pesquisa em um determinado grupo de trabalhadores

Amostragem

- Amostragem Probabilística (Aleatória)
 - Sistemática
 - Sistematização da escolha aleatória
 - Estratificada
 - Segmentação da população em grupos homogêneos
 - Por conglomerado
 - Formação de grupos heterogêneos

Amostragem

- Amostragem Probabilística (Aleatória)
 - Sistemática
 - Sistematiza-se a escolha aleatória
 - Aleatoriza-se os indivíduos
 - Escolhe-se cada k-ésimo elemento
 - $K = \text{tamanho população} / \text{tamanho amostra}$
 - Exemplo
 - Pesquisa sobre funcionários da empresa X

Amostragem

- Amostragem Probabilística (Aleatória)
 - Estratificada
 - Divide-se a população em subgrupos (estratos)
 - Grupos de características similares
 - Pouca variabilidade em amostras do mesmo grupo
 - Exemplos
 - Tempo de lazer de pessoas de distintas faixas de renda
 - Duração de férias de pessoas de distintas faixas de renda
 - Hábitos de lazer por faixa etária
 - Pode-se compor a amostra com alguns elementos de cada subgrupo?

Amostragem

- Amostragem Probabilística (Aleatória)
 - Por conglomerado
 - Formação de grupos heterogêneos
 - Ideal é que cada grupo represente bem a população
 - Seria uma minipopulação
 - Exemplo
 - Pesquisa de intenção de voto
 - Coloca-se na amostra um grupo de indivíduos heterogêneos que represente a população
 - Pesquisa sobre municípios brasileiros
 - A amostra deve conter municípios com características distintas de modo a bem representar o conjunto
 - Pesquisa sobre satisfação de funcionários
 - Deve-se ter funcionários de todas as categorias?

Amostragem

- Amostragem tem como finalidade
 - Obter uma indicação de valores de parâmetros da população
 - Média
 - Mediana
 - Desvio padrão (variabilidade)
 - Proporção de elementos com certa característica
 - Haverá divergência entre valores
 - Amostral e populacional?
 - De distintas amostras da mesma população?
 - Qual a proximidade entre parâmetro populacional e amostral?

Amostragem

- Qual a proximidade entre parâmetro populacional e amostral?
 - Fatores que influem na resposta
 - Tamanho da amostra
 - Menor variabilidade entre estatísticas de grandes amostras
 - Variabilidade da população em si
 - População com alta variabilidade leva a estatísticas amostrais com elevada variabilidade
 - A estatística sendo considerada
 - Variabilidade de distintas estatísticas são descritas por distintas distribuições de probabilidades
 - Populações com muita variabilidade geram estatísticas amostrais com maior variabilidade

Distribuições Amostrais

- Uma distribuição amostral é uma distribuição de probabilidades que indica até que ponto uma estatística amostral tende a variar devido a variações casuais na amostragem aleatória

Distribuições Amostrais

- Pode usar-se parâmetros amostrais para inferir-se sobre parâmetros populacionais?
- Inferência Estatística (raciocínio indutivo)
 - Envolve formulação de juízo sobre a população a partir do exame de uma amostra extraída desta população
- Uma Distribuição Amostral é
 - uma distribuição de probabilidades que indica até que ponto uma estatística amostral tende a variar devido a variações casuais na amostragem aleatória

Distribuições Amostrais

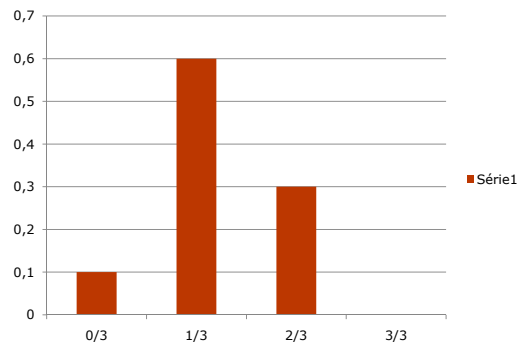
- Suponha uma população sobre a qual quer-se saber algo
 - Comprimento e peso de peixes de um lago
 - Peso dos animais de um rebanho
 - Altura dos alunos de determinada faixa etária
 - Salário dos trabalhadores de determinada categoria no país
 - Rentabilidade das empresas de um setor
 - Nível de recolhimento de impostos de empresas de certo setor

Distribuições Amostrais

- Se é viável fazer-se um censo, ok
- Se não é viável o censo
 - Tem-se que usar amostras
 - Peso dos animais de um rebanho de 1.000 cabeças
 - Pode-se escolher amostras de 2, 3, 4... 1.000 animais
 - Exemplo: população de 5 animais com pesos de 200, 203, 206, 209, 212
 - Amostras de 3 indivíduos
 - Qual a **Proporção de animais com peso inferior a 205?**
 - Analisar todas as amostras possíveis

- Amostras de 3 indivíduos => 10 combinações

Animal	Peso	Amostras	Pesos			Média amostral			proporção abaixo de 205			
A	200	ABC	200	203	206	203	1	1	0	2/3	0,1	0,3
B	203	ABD	200	203	209	204,5	1	1	0	2/3	0,1	
C	206	ABE	200	203	212	206	1	1	0	2/3	0,1	
D	209	ACD	200	206	209	204,5	1	0	0	1/3	0,1	0,6
E	212	ACE	200	206	212	206	1	0	0	1/3	0,1	
Média:	206	ADE	200	209	212	206	1	0	0	1/3	0,1	
		BCD	203	206	209	206	1	0	0	1/3	0,1	
		BCE	203	206	212	207,5	1	0	0	1/3	0,1	
		BDE	203	209	212	207,5	1	0	0	1/3	0,1	
		CDE	206	209	212	209	0	0	0	0/3	0,1	0,1

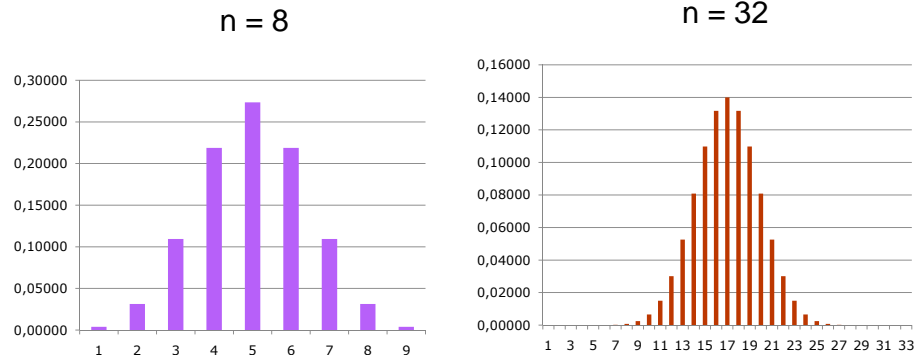


Distribuições Amostrais

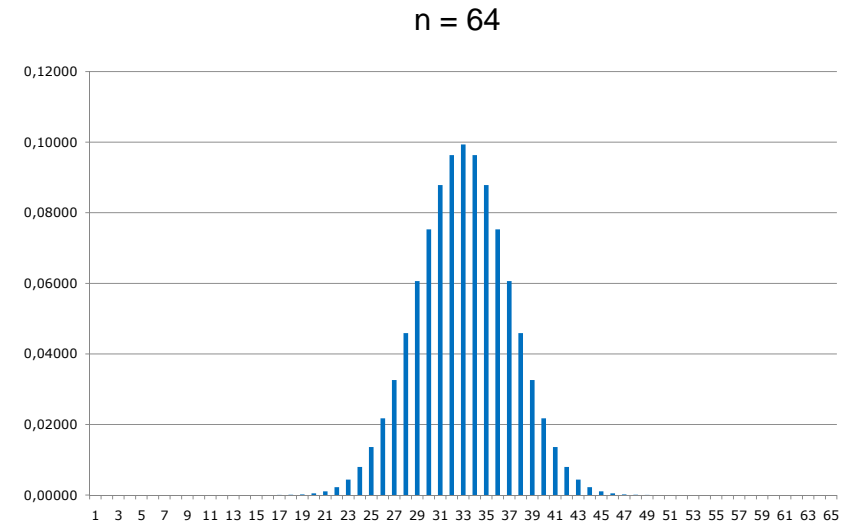
- Efeito do tamanho da amostra sobre uma distribuição amostral
 - Parâmetro amostral X populacional
 - Amostragem aleatória
 - Alta propensão a que a estatística amostral se aproxime do parâmetro populacional
 - Quanto maior a amostra
 - Mais próximo Parâmetro amostral do populacional

■ Efeito do tamanho da amostra sobre uma Distribuição Amostral

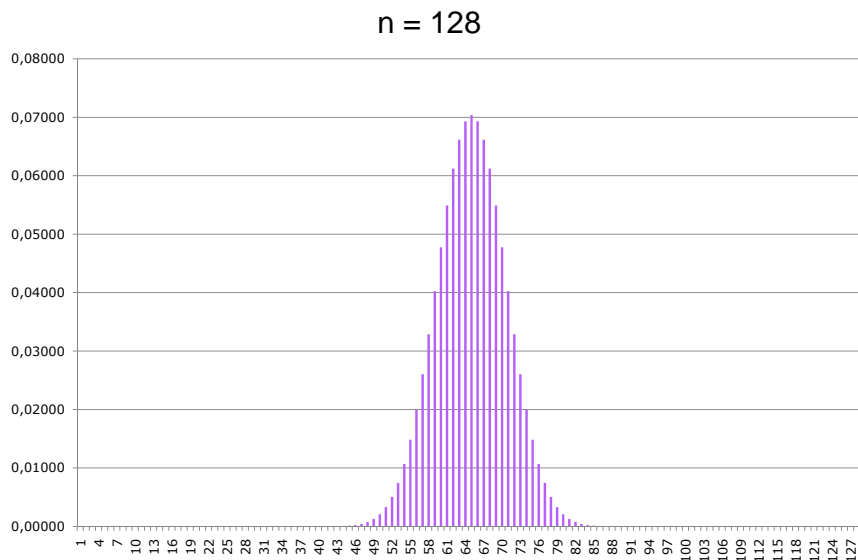
- ***Maior amostra => Mais proximidade da Normal***
- Exemplo: Distribuição Binomial com $p = 0,5$ e crescente n



■ Maior amostra => Mais proximidade da Normal



■ Maior amostra => Mais proximidade da Normal



Distribuições Amostrais

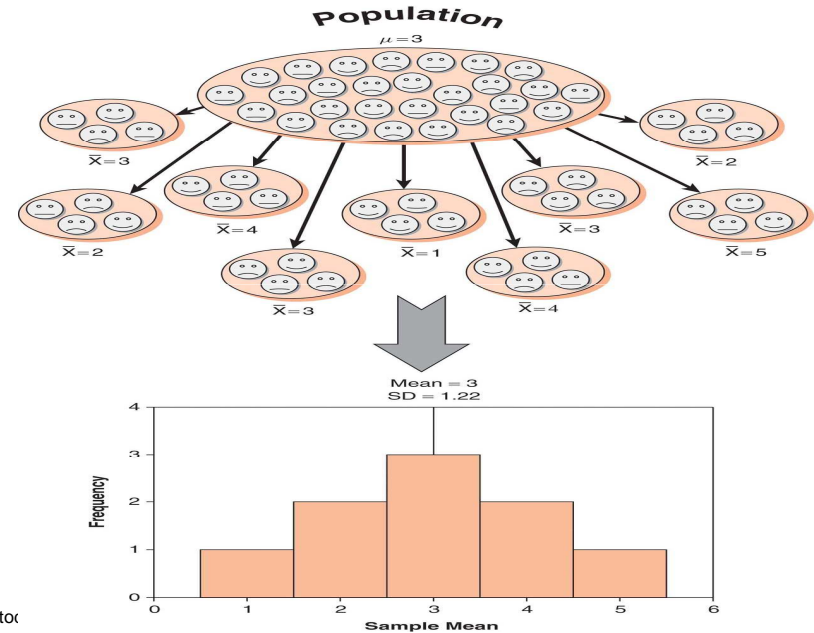
- Efeito do tamanho da amostra sobre uma distribuição amostral
 - Aumento no tamanho da amostra
 - Distribuição tende para a Normalidade
 - População mais simétrica em relação à média
 - Mais acelerada normalidade
 - Menos variabilidade entre proporções amostrais
 - Erro do parâmetro amostral em relação ao populacional decresce
 - Distribuição Normal
 - Adequada para descrever distribuições de Médias Amostrais

Distribuições de Médias Amostrais

■ Uma Distribuição Amostral de Médias

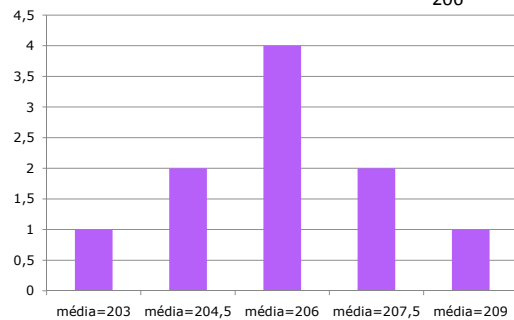
- Indica probabilidade de diversas médias amostrais
- Distribuição de Frequências de médias amostrais
- Tende à Normalidade
- Média da Distribuição Amostral de Médias é igual à média populacional
- Aumentando-se o número de observações por amostra
 - Há uma redução do desvio padrão da distribuição amostral

Amostra representa bem a população?

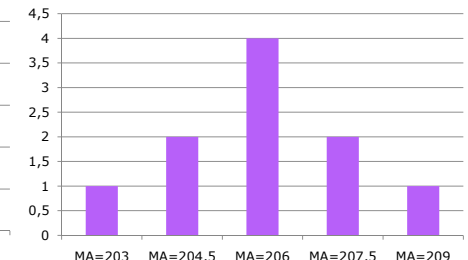
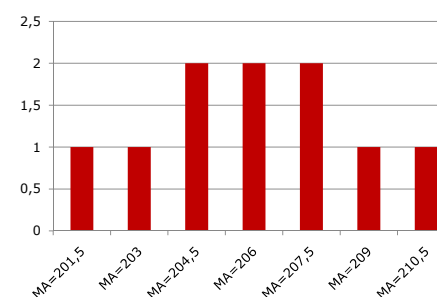


- População de 5 animais com pesos de 200, 203, 206, 209, 212
 - Amostras de 3 indivíduos => 10 combinações

Animal	Peso	Amostras	Pesos			Média amostral	Frequência	MA*Freq
A	200	ABC	200	203	206	203	1	203
B	203	ABD	200	203	209	204,5	2	409
C	206	ABE	200	203	212	206	4	824
D	209	ACD	200	206	209	204,5		
E	212	ACE	200	206	212	206		
Média:	206	ADE	200	209	212	206		
		BCD	203	206	209	206		
		BCE	203	206	212	207,5	2	415
		BDE	203	209	212	207,5		
		CDE	206	209	212	209	1	209
					206	10	2060	



indiv	peso	pesos			Média amostral		Fr	Amostras				Pesos		Média amostral		Fr
A	200	AB	200	203	201,5	MA=201,5	1	ABC	200	203	206	203	MA=203	1		
B	203	AC	200	206	203	MA=203	1	ABD	200	203	209	204,5	MA=204,5	2		
C	206	AD	200	209	204,5	MA=204,5	2	ABE	200	203	212	206	MA=206	4		
D	209	AE	200	212	206	MA=206	2	ACD	200	206	209	204,5	MA=207,5	2		
E	212	BC	203	206	204,5	MA=207,5	2	ACE	200	206	212	206	MA=209	1		
Méd	206	BD	203	209	206	MA=209	1	ADE	200	209	212	206				
DP	4,243	BE	203	212	207,5	MA=210,5	1	BCD	203	206	209	206				
		CD	206	209	207,5			BCE	203	206	212	207,5				
		CE	206	212	209			BDE	203	209	212	207,5				
		DE	209	212	210,5			CDE	206	209	212	209				



Distribuições Amostrais

- Média da Distribuição Amostral é igual à Média Populacional

$$\mu_{\bar{x}} = \mu_x$$

- Para população muito grande, desvio padrão amostral será:

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$$

- Onde:

- Desvio Padrão da distribuição amostral ($\sigma_{\bar{x}}$)
- Desvio Padrão da distribuição populacional (σ_x)

Distribuições Amostrais

- Teorema do Limite Central

- Não é necessário conhecer a distribuição de uma população para fazer-se inferências sobre ela a partir de amostras
- Pode-se fazer inferência a partir de amostras “grandes”
 - 30 (trinta) ou mais observações

Distribuições Amostrais

- **Teorema do Limite Central**

- 1- Se População em estudo tem Distribuição Normal, a Distribuição de Médias Amostrais será Normal para todos os tamanhos de amostra
- 2- Se População em estudo não é Normal, a Distribuição de Médias Amostrais será aproximadamente Normal para grandes amostras

Distribuições Amostrais

- Fórmulas para cálculos com Distribuições Amostrais de Médias

- Média da Distribuição Amostral = Média da População

$$\mu_{\bar{x}} = \mu_x$$

- Desvio Padrão da Distribuição Amostral (DA) (população grande)

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$$

Distribuições Amostrais

■ Aplicação

■ Grande população com:

- Média = 20; Desvio Padrão = 1,4
- Amostra de 49 observações

$$\mu_{\bar{x}} = \mu_x = 20$$

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}} = \frac{1,4}{\sqrt{49}} = \frac{1,4}{7} = 0,2$$

- Variabilidade das médias amostrais com relação à média populacional

Distribuições Amostrais

■ Aplicação

■ Grande população com:

- Média = 20; Desvio Padrão = 1,4
- Amostra de 49 observações
- $N > 30 \Rightarrow$ Distribuição Amostral Normal

- Porcentagem de possíveis médias amostrais que diferem por mais 0,2 da média populacional

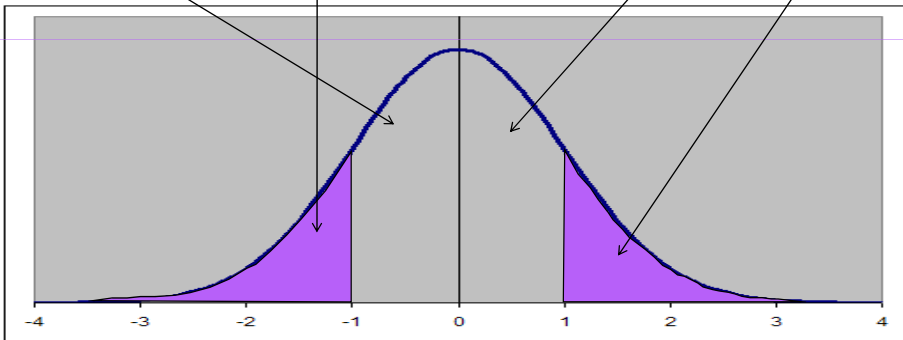
- % de médias acima de 20,2, mais o % de médias abaixo de 19,8
Equivale a
- % de médias acima de +1DPA (Desvio Padrão Amostral) ($x=20,2$)
mais % de médias abaixo de +1DPA ($x=19,8$)

$$\frac{19,8 - 20}{0,2} = -1\sigma_{\bar{x}}$$

$$\frac{20,2 - 20}{0,2} = +1\sigma_{\bar{x}}$$

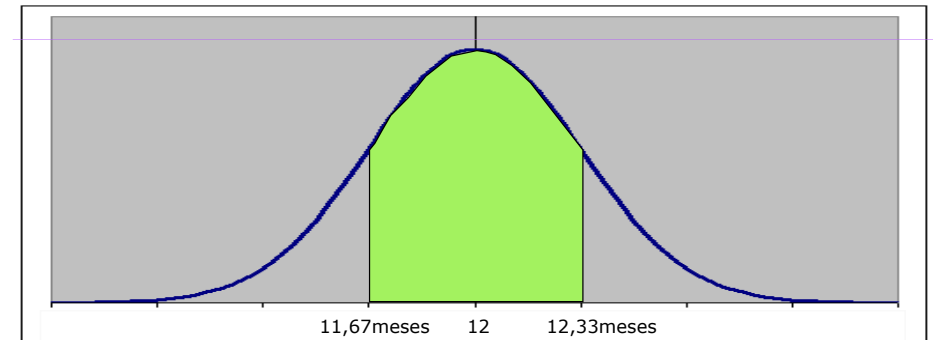
- $P(x < 19,8 \text{ ou } x > 20,2) = P(x < 19,8) + P(x > 20,2) =$
 $P(z < -1) + P(z > +1) = 0,1587 + 0,1587 = \mathbf{0,3174}$

- $P(x < 19,8) = P(z < -1) = 0,5 - 0,3413 = 0,1587$
- $P(x > 20,2) = P(z > +1) = 0,5 - 0,3413 = 0,1587$



- Um fabricante declara que sua pilha tem vida esperada (média) de 12 meses. Sabe-se que o desvio padrão (populacional) correspondente é de 2 meses. Que porcentagem de amostras de pilhas de N observações terá vida média de 1/3 de mês em torno da média?

- $N = 36; N = 64; \dots$



Distribuições Amostrais

- Média Populacional = 12 meses
- Desvio Padrão Populacional = 2 meses
- % de amostras de pilhas com vida média variando 1/3 de mês (0,3333 mês) em torno da média?
- N = 32; N = 64; N = 128; N = 256; N = 512.

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$$

$$z = \frac{x - \bar{x}}{\frac{\sigma_x}{\sqrt{n}}}$$

N	Média DA	DP DA	MDA inferior	MDA superior	z inferior	z superior	P(inf < MDA < sup) = P(inf < z < sup)		
32	12	0,354	11,67	12,33	-0,93	0,93	0,3238	0,3238	<u>0,6476</u>

(11,67-12)/0,354 = (12,33-12)/0,354
= -0,93 = 0,93

Distribuições Amostrais

- Média Populacional = 12 meses
- Desvio Padrão Populacional = 2 meses
- % de amostras de pilhas com vida média variando 1/3 de mês (0,3333 mês) em torno da média?
- N = 32; N = 64; N = 128; N = 256; N = 512.

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$$

$$z = \frac{x - \bar{x}}{\frac{\sigma_x}{\sqrt{n}}}$$

N	Média DA	DP DA	MDA inferior	MDA superior	z inferior	z superior	P(inf < MDA < sup) = P(inf < z < sup)		
32	12	0,354	11,67	12,33	-0,93	0,93	0,3238	0,3238	<u>0,6476</u>
64	12	0,250	11,67	12,33	-1,32	1,32	0,4049	0,4049	<u>0,8098</u>
128	12	0,177	11,67	12,33	-1,87	1,87	0,4693	0,4693	<u>0,9386</u>
256	12	0,125	11,67	12,33	-2,64	2,64	0,4959	0,4959	<u>0,9918</u>
512	12	0,088	11,67	12,33	-3,73	3,73	0,4999	0,4999	<u>0,9998</u>

Distribuições Amostrais

- Distribuições de Proporções Amostrais
 - Uma DPA indica quão provável é determinado conjunto de proporções amostrais, conhecidos o Tamanho da Amostra e a Proporção Populacional
 - Média (proporção ou porcentagem média) da DA é igual à proporção populacional

$$\bar{p} = p$$

- Para grande população, fórmula do DP Amostral

$$\sigma_p = \sqrt{\frac{p(1-p)}{n}}$$