PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE TRANSPORTES E GESTÃO TERRITORIAL – PPGTG
DEPARTAMENTO DE ENGENHARIA CIVIL – ECV

DISCIPLINA: TGT410026 – FUNDAMENTOS DE ESTATÍSTICA

8ª AULA: ESTIMAÇÃO POR INTERVALO

Vimos que a estimação por ponto usa uma estatística da amostra para estimar um parâmetro da população.

Entretanto, não se pode esperar que uma estimação por ponto produza o valor exato do parâmetro populacional. Mesmo que não seja possível obter o valor exato de um parâmetro populacional a partir dos dados de uma amostra, vimos que a distribuição amostral de uma estatística pode nos ajudar a verificar quão próximo do valor do parâmetro amostral está o valor da estatística da amostra.

A partir de agora, vamos adicionar um valor ao estimador por ponto, denominado ***margem de erro***, para produzir uma estimação por intervalo do parâmetro populacional de interesse. Esse intervalo mostrará o grau de confiança na nossa estimação além de evidenciar a aproximação entre a estatística amostral e o parâmetro populacional.

De maneira geral a estimação por intervalo possui a seguinte estrutura:

Estimação por ponto ± Margem de erro

Assim temos:

$\bar{x}$ ± Margem de erro, para estimar a média populacional μ

$\bar{p}$ ± Margem de erro, para estimar a proporção da população p

O leitor já pode antever o papel fundamental que desempenham as distribuições amostrais de $\bar{x}$ e $\bar{p}$.

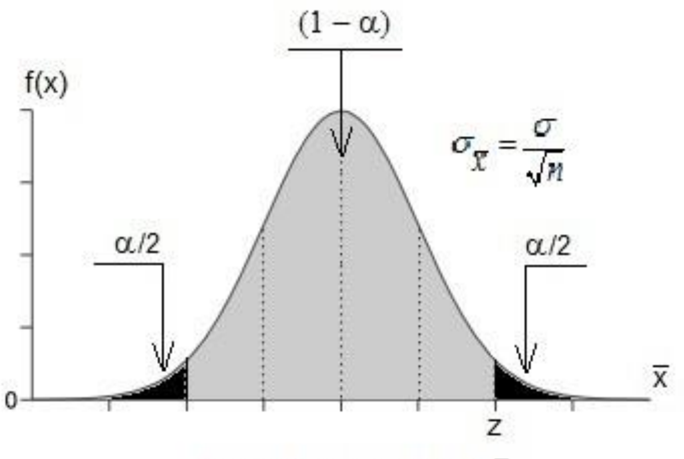
Do ponto de vista conceitual já possuímos todas as informações necessárias para desenvolver estimações por intervalo.

MÉDIA DA POPULAÇÃO: σ CONHECIDO

A estimação por intervalo de uma média populacional μ com desvio padrão ***σ conhecido*** é estabelecida usando a seguinte expressão:

$$\mu \Rightarrow \bar{x} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

em que ($1 - \alpha$) é o coeficiente de confiança (ou nível de confiança ou grau de confiança) adotado na estimação; σ é o desvio padrão da população; n é o tamanho da amostra; $z_{\alpha/2}$ é o valor de z que produz uma área de $\alpha/2$ na cauda superior da distribuição normal padrão de probabilidade.



Distribuição amostral de $\bar{x}$

IMPORTANTE:
A distribuição amostral mostra como os valores de $\bar{x}$ estão distribuídos nas proximidades da média populacional μ; consequentemente a distribuição amostral de $\bar{x}$ fornece informações sobre as possíveis diferenças entre $\bar{x}$ e μ.

Vamos usar a tabela de áreas da distribuição normal padrão para verificar como determinar a margem de erro de nossa estimação. Vamos imaginar que nosso nível de confiança ($1 - \alpha$) seja de 95%. Vamos pesquisar na tabela um valor de z, tal que o intervalo $-z$ a $+z$ produza uma área sob a curva normal padrão de 0,95. Ou seja, a área de (-z a 0) somada com a área de (0 a +z) seja igual a 0,95. Esse valor é z = 1,96 que produz a área de 0,4750. Por simetria, para que a área sob a curva normal seja igual a 0,95, temos que considerar o intervalo [-z a +z], ou seja, 0,4750 + 0,4750 = 0,950.

A distribuição amostral de $\bar{x}$ segue uma distribuição normal com desvio padrão da média (erro padrão) igual a $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$. Assim, fica estabelecido o denominado ***erro amostral***, pela expressão:

$$\pm 1{,}96.\sigma_{\bar{x}} = \pm 1{,}96.\frac{\sigma}{\sqrt{n}}$$

Genericamente, a margem de erro de uma estimação da média populacional para o caso de σ conhecido é:

$$\pm z_{\alpha/2}.\sigma_{\bar{x}} = \pm z_{\alpha/2}.\frac{\sigma}{\sqrt{n}}$$

Com a margem de erro determinada pela expressão acima, a forma geral de uma *estimação da média populacional para o caso de* ***σ*** *conhecido* é a seguinte:

$$\mu \Rightarrow \bar{x} \pm z_{\alpha/2}.\frac{\sigma}{\sqrt{n}}$$

Alguns valores para o nível de confiança normalmente usados nos processos de estimação:

Nível de confiança	**α**	**α/2**	$\mathbf{Z_{\alpha/2}}$
90%	0,10	0,05	1,645
95%	0,05	0,025	1,960
99%	0,01	0,005	2,576

Exemplo:

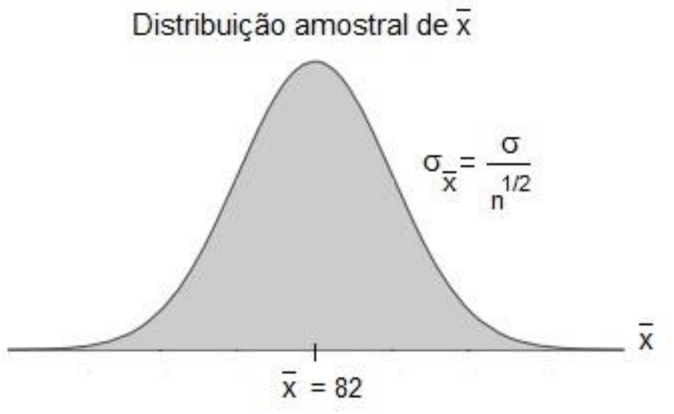
A *Lloyd´s Store* seleciona uma amostra aleatória simples de cem clientes para saber qual quantia eles gastam em cada ida às compras. Considerando a variável de interesse *x = a quantia gasta em cada ida às compras*, a média amostral $\bar{x}$ fornece uma estimação por ponto de μ, que é a quantia média gasta em cada ida às compras pela população de todos os clientes da empresa. A empresa usa essa pesquisa semanal há vários anos.

Baseando-se nos dados históricos, a empresa assume um valor conhecido de σ = US$ 20 para o desvio padrão da população; além disso, os dados históricos indicam claramente que a população (!) segue uma distribuição normal.

Na última semana a *Lloyd´s* pesquisou 100 clientes (*n = 100*) e obteve a média $\bar{x}$ da amostra = US$ 82.

Estabelecer uma estimação por intervalo da quantia média gasta pelos clientes em cada ida às lojas.

SOLUÇÃO:
Trata-se de uma variável com distribuição amostral que segue uma distribuição normal; além disso, o desvio padrão da população é conhecido.



Resta então estabelecer o nível de confiança a adotar e determinar a margem de erro da estimação. Supor (1 – α) = 95%.

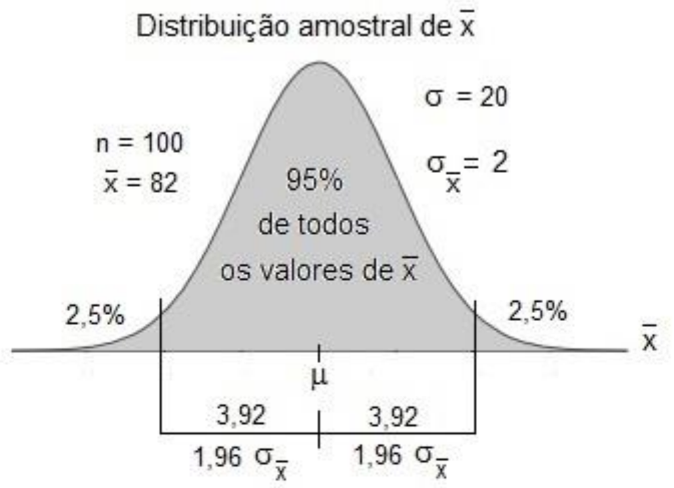
$$\mu \Rightarrow 82 \pm z_{\alpha/2} \cdot \frac{20}{\sqrt{100}}$$

$$\mu \Rightarrow 82 \pm z_{\alpha/2} .2$$

Para (1 – α) = 95% temos α = 0,05 e α/2 = 0,025. Assim, para uma área na aba superior da curva normal igual a 0,025, ou para uma área entre 0 e $z_{\alpha/2}$ igual a 0,5 – 0,025 = 0,4750, verificamos na tabela que esse valor corresponde a z = 1,96. Por simetria, esse valor de área corresponde a z = -1,96. Assim, a margem de erro da estimação por intervalo com nível de confiança de 95% é:

$$\mu \Rightarrow 82 \pm 1{,}96 x 2$$

$$\mu \Rightarrow 82 \pm 3{,}92$$



O intervalo [78,08 a 85,92] é dito intervalo de confiança e dizemos que temos 95% de confiança em que o intervalo [78,08 a 85,92] inclui a média populacional μ.

Refazer o exemplo usando níveis de confiança de 90% e 99%. Combinando os três resultados (margens de erro) correspondentes aos três níveis de confiança, o que dizer a respeito da margem de erro da estimação e da amplitude dos intervalos de confiança.

MÉDIA DA POPULAÇÃO: σ DESCONHECIDO

Quando desejamos fazer estimação por intervalo da média populacional e não temos uma boa estimativa do desvio padrão da população, precisamos lançar mão da amostra para estimar μ e σ. Essa situação representa o caso de ***σ desconhecido***.

Quando o desvio padrão da amostra *s* é usado para estimar *σ* a margem de erro e a estimação da média populacional baseiam-se em uma distribuição de probabilidade conhecida como ***distribuição t***.

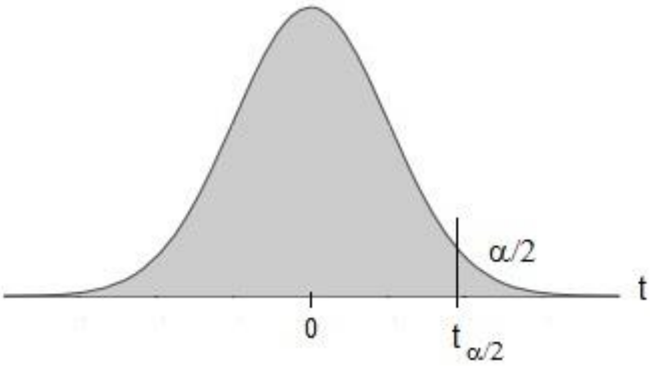
A *distribuição t* é utilizada em inúmeras aplicações, mesmo naquelas situações em que as populações se desviam da distribuição normal. É a distribuição de probabilidade indicada quando nos deparamos com amostras de tamanho pequeno.

Matematicamente a distribuição t foi desenvolvida baseando-se na suposição de uma distribuição normal para a população da qual é extraída a amostra. A distribuição t possui como parâmetro o número de graus de liberdade, dada pela relação $gl = (n - 1)$, sendo n o tamanho da amostra. A medida que gl aumenta, a diferença entre a distribuição t e a distribuição normal padrão torna-se cada vez menor.

A figura a seguir ilustra o caso de distribuições t com 10 e 20 graus de liberdade. Note que, aumentando o grau de liberdade menor é a variabilidade em relação à distribuição normal; note também que a média da distribuição t é igual a zero.

Inserir aqui a figura 8.4

Do ponto de vista prático, o uso da *distribuição t* nos casos de desvio padrão populacional desconhecido é idêntica ao uso da distribuição normal padrão. No caso de estimação com desvio padrão conhecido usamos a notação $z_{\alpha/2}$ para indicar o valor de z que produz uma área de α/2 (0,025) na cauda superior de uma distribuição normal padrão; no caso em tela, em que o desvio padrão é desconhecido vamos usar a notação $t_{\alpha/2}$ para indicar o valor t que produz uma área α/2 (0,025) na cauda superior da *distribuição t*.



A margem de erro e a estimação por intervalo de uma média populacional no caso de σ desconhecido é:

$$\mu \Rightarrow \bar{x} \pm t_{\alpha/2} . \frac{s}{\sqrt{n}}$$

Em que s é o desvio padrão da amostra, n é o tamanho da amostra, $t_{\alpha/2}$ é o valor t que produz uma área igual a $\alpha/2$ na cauda superior da distribuição t e $gl = (n – 1)$ graus de liberdade.

Para se trabalhar com a *distribuição t* existem tabelas que fornecem as áreas desejadas em função do número de graus de liberdade $gl = (n – 1)$ e do $\alpha/2$ adotado.

Exemplo:

Foi idealizado um estudo para estimar a média dos débitos de cartões de crédito da população de famílias de um país. Uma amostra de $n = 85$ famílias forneceu os saldos de cartões de crédito mostrados na tabela a seguir.

Nenhuma informação a respeito de estimativas do desvio padrão da população está disponível.

Saldos de cartões de crédito de uma amostra de 85 famílias					
9619	5994	3344	7888	7581	9980
5364	4652	13627	3091	12545	8718
8348	5376	968	943	7959	8452
7348	5998	4714	8762	2563	4935
381	7530	4334	1407	6787	5938

2998	3678	4911	6644	5071	5266
1686	3581	1920	7644	9536	10658
1962	5625	3780	11169	4459	3910
4920	5619	3478	7979	8047	7503
5047	9032	6185	3258	8083	1582
6921	13236	1141	8660	2153	
5759	4447	7577	7511	8003	
8047	609	4667	14442	6795	
3924	414	5219	4447	5915	
3470	7636	6416	6550	7164	

Realizar a estimação por intervalo da média dos débitos de cartões de crédito da população, adotando um nível de confiança de 95%.

Resolução:
Nesse caso, deveremos utilizar os dados da tabela (dados amostrais) para estimar a média e o desvio padrão da população. Para o cálculo da média utilizar a expressão:

$\bar{x} = \frac{\sum x_i}{n}$ e para o desvio padrão: $s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$;

os valores encontrados foram:

$\bar{x} = 5.900$

$s = 3.058$

Com 95% de confiança e *n – 1 = 84* graus de liberdade, (1 – α) = 0,95, α = 0,05 e α/2 = 0,025, a tabela de *distribuição t* fornece $t_{0,025} = 1,989$.

Podemos agora realizar a estimação por intervalo da média populacional:

$$\mu \Rightarrow 5.900 \pm 1,989.\frac{3.058}{\sqrt{85}}$$

$$\mu \Rightarrow 5.900 \pm 660$$

A estimação por ponto da média populacional é 5.900; a margem de erro é 660, e o intervalo de confiança é:
5.900 – 660 = 5.240 a 5.900 + 660 = 6.560.

Desse modo os resultados mostram que temos 95% de confiança que a média dos saldos de cartões de crédito da população de todas as famílias está ente 5.240 e 6.560.

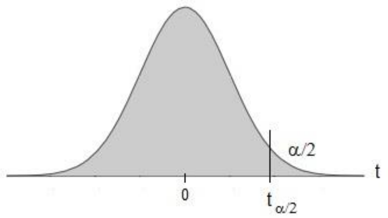


Tabela da distribuição t (um pequeno trecho...)						
Graus de	Área na cauda superior					
Liberdade	0,20	0,10	0,05	0,025	0,01	0,005
1	1,376	3,078	6,314	12,706	31,821	63,656
5	0,920	1,476	2,015	2,571	3,365	4,032
9	0,883	1,383	1,833	2,262	2,821	3,250
12	0,873	1,356	1,782	2,179	2,681	3,055
16	0,865	1,337	1,746	2,120	2,583	2,921
19	0,861	1,328	1,729	2,093	2,539	2,861
25	0,856	1,316	1,708	2,060	2,485	2,787
29	0,854	1,311	1,699	2,045	2,462	2,756
30	0,854	1,310	1,697	2,042	2,457	2,750
40	0,851	1,303	1,684	2,021	2,423	2,704
45	0,850	1,301	1,679	2,014	2,412	2,690
50	0,849	1,299	1,676	2,009	2,403	2,678
53	0,848	1,298	1,674	2,006	2,399	2,672
80	0,846	1,292	1,664	1,990	2,374	2,639
84	0,846	1,292	1,663	1,989	2,372	2,636

Como trabalhar com amostras pequenas?

Sabemos que na maioria dos casos um tamanho da amostra $n \geq 30$ produz resultados adequados para estimação por intervalo de uma média populacional. Entretanto se a distribuição for altamente inclinada ou se contiver pontos fora da curva recomenda-se o uso de tamanho de amostra maiores. Se a distribuição não está normalmente distribuída, mas é mais ou menos simétrica, pode-se esperar que tamanhos de amostra em torno de 15 elementos produzam bons *intervalos de confiança aproximados*.

De qualquer maneira, é sensato elaborar um histograma quando estamos diante de amostras de tamanho pequeno. O histograma pode auxiliar a análise da distribuição da população. Vejamos um exemplo prático:

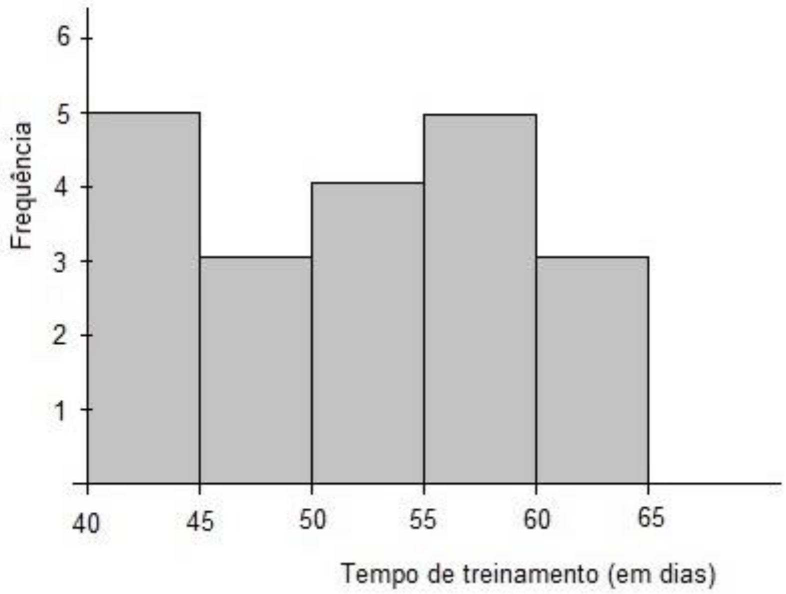
A MotorAço Industrial está interessada em desenvolver um novo programa auxiliado por computador para treinar os empregados do setor de manutenção a fazer reparos nas máquinas. A fim de avaliar plenamente o programa, o diretor do departamento de manufaturas solicitou uma estimativa do tempo médio populacional necessário

para que os empregados do setor de manutenção concluam o treinamento auxiliado por computador.
Uma amostra de 20 empregados é selecionada, tendo cada empregado da amostra concluído o programa de treinamento. Os dados coletados estão na tabela a seguir:

Tempo de treinamento, em dias, correspondente à amostra de 20 empregados da MotorAço

52	59	54	42
44	50	42	48
55	54	60	55
44	62	62	57
45	46	43	56

Na dúvida, o analista da empresa elaborou um histograma para facilitar a análise quanto à distribuição populacional de tempo de treinamento dos empregados do setor de manutenção da empresa MotorAço. O histograma é mostrado a seguir:



Após a análise dos dados e do histograma, efetuar a estimação do tempo médio populacional utilizando um nível de confiança de 95%.
Resultados:

$$\bar{x} = 51{,}5$$

$$s = 6{,}84$$

$$gl = (n-1) = 19$$

$$(1-\alpha) = 0{,}95$$

$$\alpha = 0{,}05$$

$$\alpha / 2 = 0{,}025$$

$$t_{\alpha/2} = t_{0{,}025} = 2{,}093$$

A estimação por ponto é $\bar{x} = 51,5$, a margem de erro $\pm 2,093.\frac{6,84}{\sqrt{20}} = \pm 3,2$, e o intervalo de confiança é: [48,3 a 54,7] dias.

NOTA IMPORTANTE:

Quando σ é conhecido, a margem de erro $z_{\alpha/2}(\sigma / \sqrt{n})$ É FIXA E É A MESMA para todas as amostras de tamanho *n*.
Quando σ é desconhecido, a margem de erro $t_{\alpha/2}(s / \sqrt{n})$ VARIA DE AMOSTRA PARA AMOSTRA; isso ocorre porque o desvio padrão varia, dependendo da amostra selecionada.

Aplicação:

1) Para uma distribuição t com 16 graus de liberdade, encontre a área ou probabilidade, em cada região apresentada a seguir:
a) À direita de 2,120
b) À esquerda de 1,337
c) À esquerda de -1,746
d) À direita de 2,583
e) Entre -2,120 e 2,120
f) Entre -1,746 e 1,746

2) Encontre os valores t em cada um dos seguintes casos:
a) Área da cauda superior igual a 0,025, com 12 graus de liberdade;
b) Área da cauda inferior igual a 0,05, com 50 graus de liberdade;
c) Área da cauda superior igual a 0,01, com 30 graus de liberdade;
d) Em que 90% da área se situa entre esses dois valores t com 25 graus de liberdade;
e) Em que 95% da área se situa entre esses dois valores t com 45 graus de liberdade.

3) Uma amostra aleatória simples com *n = 54* produziu uma média amostral igual a 22,5 e um desvio padrão da amostra igual a 4,4.
a) Desenvolva um intervalo de confiança de 90% para a média populacional;
b) Estabeleça um intervalo de confiança de 95% para a média populacional;
c) Estabeleça um intervalo de confiança de 99% para a média populacional;
d) O que acontece à margem de erro e ao intervalo de confiança, quando o grau de confiança é aumentado?

4) O número médio de horas de vôo dos pilotos da Continental Airlines equivale a 49 horas por mês. Suponha que essa média tenha se baseado em tempos de vôo reais de uma amostra de 110 pilotos da Continental Airlines e que o desvio padrão da amostra tenha sido de 8,5 horas.
a) Com 95% de confiança, qual é a margem de erro?
b) Qual é a estimação por intervalo, com 95% de confiança, do tempo de vôo médio da população de pilotos?
c) O número médio de horas de vôo dos pilotos da United Airlines equivale a 36 horas por mês. Use os resultados que obteve no item (b) para discutir as diferenças entre os tempos de vôo dos pilotos das duas empresas aéreas. Um relatório da Infraero mostrou que a United Airlines tem o custo de mão-de-obra mais elevado entre todas as empresas aéreas. A informação contida nesse exercício oferece subsídios para compreendermos por que a United Airlines poderia esperar custos de mão-de-obra mais elevados?

COMO DETERMINAR O TAMANHO DA AMOSTRA

Sabemos que a estimação por intervalo de uma média populacional é dada pela expressão:

$$\mu \Rightarrow \bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

Chamando de E = a margem de erro, resolver a expressão para $\sqrt{n}$;

$$E = z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \therefore \sqrt{n} = \frac{z_{\alpha/2}.\sigma}{E}$$

$$n = \frac{(z_{\alpha/2})^2.\sigma^2}{E^2}$$

OBSERVAÇÃO:
Note que o tamanho da amostra está dependendo do valor do desvio padrão σ.
Se σ não é conhecido, deve-se providenciar um valor planejado para o desvio padrão; na prática, pode-se adotar um dos seguintes procedimentos:
a) usar uma estimativa do desvio padrão da população, calculada a partir de dados de estudos anteriores;
b) usar um estudo piloto para selecionar uma amostra preliminar; o desvio padrão amostral poderá ser usado como valor planejado;
c) usar como aproximação tosca do desvio padrão, a diferença entre os maiores e os menores valores de dados da população. Muitas

vezes a amplitude dividida por 4 é sugerida como um valor planejado de σ.

Exemplos:

1) Um estudo sobre o custo de aluguel de automóveis revelou que o custo médio para alugar um carro de porte médio era de aproximadamente R$ 55 por dia com desvio padrão de R$ 9,65. Ao projetar um novo estudo para saber a evolução do custo diário de aluguel de automóveis de tamanho médio, foi projetada uma margem de erro da média populacional do custo de aluguel de R$ 2 e um grau de confiança de 95%.
Qual o tamanho da amostra a ser selecionada para realizar tal estudo?

$$n = \frac{(z_{\alpha/2})^2 . \sigma^2}{E^2} = \frac{(1,96)^2 (9,65)^2}{2^2} = 89,43$$

$$n = 90$$

2) Estima-se que a amplitude de um conjunto de dados seja 36. Determinar:
a) qual o valor planejado vc usaria para o desvio padrão da população? (σ = 9)
b) com 95% de confiança, qual seria o tamanho da amostra que forneceria uma margem de erro igual a 3? (n = 35)
c) com 95% de confiança, qual seria o tamanho da amostra que forneceria uma margem de erro igual a 2? (n = 78)

3) A Santur fornece informações sobre o custo de pernoites em quartos de hotel em todo o território catarinense. Use R$ 2 como margem de erro desejada e R$ 22,50 como valor planejado para o desvio padrão da população para encontrar o tamanho de amostra recomendado a seguir:
a) para uma estimação por intervalo de confiança de 90% do custo médio populacional dos quartos de hotel. (n = 343)
b) para uma estimação por intervalo de confiança de 95% do custo médio populacional dos quartos de hotel. (n = 487)
c) para uma estimação por intervalo de confiança de 99% do custo médio populacional dos quartos de hotel. (n = 840)
d) quando a margem de erro é fixada, o que acontece ao tamanho da amostra quando o grau de confiança aumenta? Você recomendaria

que a Santur utilizasse um grau de confiança de 99%? Discuta. (Não recomendaria!)

ESTIMAÇÃO POR INTERVALO DA PROPORÇÃO DA POPULAÇÃO

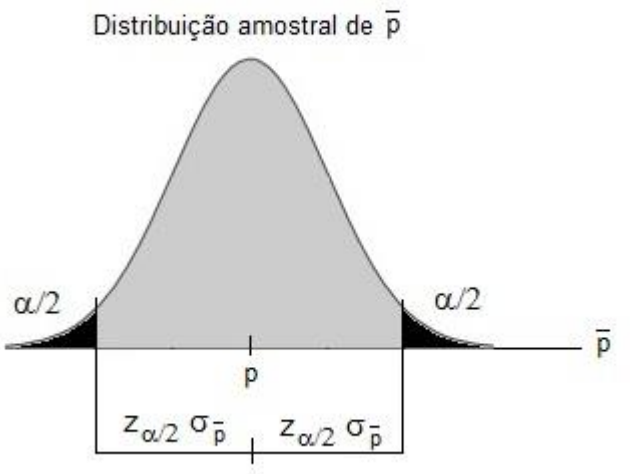
Vimos que a forma geral da estimação por intervalo de uma proporção populacional é:

$$p \pm margem\ de\ erro$$

Vimos também que a distribuição amostral de $\bar{p}$ pode ser aproximada por meio de uma distribuição normal quando $np \geq 5$ e $n(1-p) \geq 5$.

A figura a seguir ilustra a distribuição amostral de $\bar{p}$. A distribuição amostral de $\bar{p}$ é a proporção p da população e o desvio padrão de $\bar{p}$ é dado por:

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



Uma vez que a distribuição amostral de $\bar{p}$ está normalmente distribuída, a estimação por intervalo da proporção populacional possui uma margem de erro dada por:

$$Margem\ de\ erro = z_{\alpha/2}\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

A expressão geral da estimação da proporção populacional é:

$$p \Rightarrow \bar{p} \pm z_{\alpha/2}\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

Como determinar o tamanho da amostra:
A exemplo da determinação do tamanho da amostra na estimação da média populacional, aqui o raciocínio é o mesmo. Denotando por *E* a margem de erro e desenvolvendo a expressão em função de *n*, teremos:

$$n = \frac{(z_{\alpha/2})^2 . \bar{p}(1-\bar{p})}{E^2}$$

OBSERVE QUE não poderemos usar a fórmula supra diretamente porque $\bar{p}$ somente será conhecido depois de selecionarmos a amostra. Precisamos, então, de um valor planejado para $\bar{p}$ que denotaremos por *p**. Com isso usaremos a seguinte expressão para o cálculo do tamanho da amostra:

$$n = \frac{(z_{\alpha/2})^2 . p^*(1-p^*)}{E^2}$$

com as seguintes recomendações:
a) use a proporção amostral de uma amostra anterior;
b) use um estudo piloto para selecionar uma amostra preliminar; a proporção dessa amostra preliminar pode ser utilizada como *p**;
c) use o julgamento para atribuir o valor de *p**;
d) usar *p*=0,50*. Justificativa:
Alguns valores possíveis para o termo *p*(1 – p*)* presente no numerador da expressão acima são:

p*	*p*(1 – p*)*	
0,10	0,10(0,90)=0,09	
0,20	0,20(0,80)=0,16	
0,30	0,30(0,70)=0,21	
0,40	0,40(0,60)=0,24	
0,50	0,50(0,50)=0,25	— maior valor para *p*(1 – p*)*
0,60	0,60(0,40)=0,24	
0,70	0,70(0,30)=0,21	
0,80	0,80(0,20)=0,16	
0,90	0,90(0,10)=0,09	

Isso mostra que ao adotarmos p* = 0,50 como valor planejado, estaremos cientes que esse valor recomenda um tamanho de amostra maior. De fato, ficamos tranquilos, pois se a proporção amostral vier a ser um valor diferente de 0,50, a margem de erro será menor que o valor previsto. ASSIM, AO ADOTARMOS $p^* = 0{,}50$, GARANTIMOS QUE O TAMANHO DA AMOSTRA SERÁ SUFICIENTE PARA OBTERMOS A MARGEM DE ERRO DESEJADA.

NOTA:
A margem de erro desejada para estimar uma proporção populacional é, em geral, 0,10 ou menos. Em pesquisas de opinião pública realizadas por organizações como o Instituto Gallup, Harris e Ibope, uma margem de erro entre 0,02 e 0,03 é comum. Com essas margens de erro a expressão estudada fornece sempre um tamanho n da amostra que satisfaz a condição essencial $np \geq 5$ e $n(1-p) \geq 5$ para que se possa usar a distribuição normal como uma aproximação à distribuição amostral de $\bar{x}$.

APLICAÇÕES:
1) Qual o tamanho da amostra para efetuar uma estimativa da proporção populacional considerando uma confiança de 95% e margem de erro igual a 0,03. Suponha que não haja dados históricos disponíveis para que se possa planejar um valor para $\bar{p}$.

2) Uma empresa realizou uma pesquisa entre 346 pessoas que procuravam emprego para saber porque as pessoas trocam de emprego tão frequentemente. A resposta mais escolhida (152 vezes) foi "melhor remuneração em outro lugar".
a) qual é a estimação por ponto da proporção de pessoas que procuram emprego que escolheriam "melhor remuneração em outro lugar" como a razão para trocar de emprego? (0,4393)
b) qual é a estimação por intervalo de confiança de 95% da proporção populacional? (0,3870 a 0,4916)

3) Uma pesquisa realizada conjuntamente pelo jornal USA Today, CNN e Instituto Gallup para a presidência da república tomou como amostra 491 eleitores em potencial em junho/2000. Uma das principais finalidades da pesquisa era obter uma estimativa da proporção dos eleitores em potencial que eram favoráveis a cada candidato. Suponha que, nesta ocasião, o valor planejado de $\bar{p}$ foi 0,50 e um grau de confiança de 95%.

a) em junho qual foi a margem de erro planejada para a pesquisa?

b) quanto mais se aproximam as eleições de novembro, maior precisão e menores margens de erro são desejadas. Suponha as seguintes margens de erro sejam solicitadas para as pesquisas realizadas durante a campanha à Presidência da República. Calcule o tamanho da amostra recomendada para cada pesquisa:

Pesquisa (Época)	**Margem de erro**	**Tamanho da amostra – *n***
Junho	0,0442	491
Setembro	0,04	601
Outubro	0,03	1068
Início de Novembro	0,02	2401
Véspera das eleições	0,01	9604