Unicuritiba

Cálculo do tamanho da amostra nas pesquisas em Administração

Dados Internacionais de Catalogação na Publicação (CIP)
Biblioteca do Centro Universitário Curitiba - UNICURITIBA

L936 Luchesa, Cláudio J.
Cálculo do tamanho da amostra nas pesquisas em administração. / Cláudio J. Luchesa. Curitiba: Edição do autor, 2011.

27 p.

Apresentado como produção científica dos autores.

1. Pesquisa. 2. Tamanho de amostra. 3. Estatística. 4. Grau de confiança
I. Chaves Neto, Anselmo. II. Título.

CDD (21. ed.) – 519.5

CÁLCULO DO TAMANHO DA AMOSTRA – SÍNTESE

Depois de definidos os objetivos e a metodologia, o que se precisa saber numa pesquisa em Administração é quantos elementos ou quantas observações da variável de interesse deve-se tomar da população ou universo amostrado. este número chama-se tamanho da amostra e é representado, muitas vezes, por n. assim, com essas observações estima-se por ponto a média μ ou por intervalo de confiança de nível 1 - α. a construção do intervalo de confiança é feita usando-se o pivô z que tem distribuição normal com média 0 e variância 1, ou seja, z ~n(0, 1). o valor de z define a grandeza de 1 - α. Por exemplo: podemos estar interessados em pesquisar quantos alunos do Centro Universitário Curitiba deveremos entrevistar para saber quantos deles poderiam estar interessados em comprar determinado produto e a que preço, com um grau de confiança de 99 %. Ou poderíamos estar interessados também em conhecer o rendimento médio dos egressos do Curso de Administração, com um grau de confiança de 95 %.

Uma população é considerada finita quando o número de elementos que a compõe é determinado e conhecido. Já uma população é considerada infinita quando o número de elementos que a compõe é muito grande e se desconhece esse número. Um exemplo de população finita é o número de eleitores de Curitiba. Embora esse número seja grande, ele é determinado e conhecido. Mas, quando se tem como população correspondente ao número de pinos que estão sendo continuamente fabricados e evidentemente se desconhece esse número, essa população pode ser considerada infinita.

Os procedimentos descritos a seguir são genéricos, isto é: prestam-se a qualquer pesquisa. Os exemplos, todavia, foram escolhidos de modo a adequar-se, da melhor maneira possível, às pesquisas em monografias e trabalhos de conclusão de curso do Curso de Administração deste Centro Universitário Curitiba.

Duas linhas de raciocínio e cálculo podem ser empregadas: estimativa de valores ou estimativa de proporções. Além disso, o cálculo do tamanho das amostras, nas pesquisas em Administração, apresenta diversas possibilidades. As mais comuns são sintetizadas pela tabela abaixo:

Cálculo do Tamanho de Amostra	Estimativa do parâmetro populacional μ ($\bar{x}$ é a estimativa pontual de μ)	População Infinita População Finita
	Estimativa do parâmetro p (a proporção amostral $\hat{p}$ é a estimativa pontual de p)	População Infinita População Finita

1. QUADRO DOS SÍMBOLOS UTILIZADOS

- **μ** - (pronuncia-se mi) média de todos os valores de uma população qualquer que se esteja pesquisando.
- **x** - valor individual de uma variável aleatória, ou valor de uma única observação da população.
- **$\bar{x}$** - (pronuncia-se xis barra) média dos valores de uma amostra de tamanho n, é o estimador natural do parâmetro verdadeiro μ.
- **e** - margem de erro do valor que se estima para um parâmetro populacional, média, variância, desvio padrão, etc..... É também conhecido como erro de estimativa.
- **N** - número de valores que compõem uma população finita.
- **n** - número de observações que se fez de uma população para compor uma amostra, ou seja, o tamanho da amostra tomada da população que se está pesquisando.
- **z** - representa a variável aleatória normal padrão, ou seja, z ~N(0, 1), que por não depender de parâmetro desconhecido

facilita os cálculos.

σ - desvio padrão populacional.

s - desvio padrão amostral.

$\hat{p}$ - proporção amostral, que estima a verdadeira proporção populacional p.

$\hat{q}$ - complemento da proporção de uma amostra: $\hat{q} = 1 - \hat{p}$

2. CÁLCULO A PARTIR DE VALORES

2.1. Estimativa do parâmetro populacional μ

a) **Fórmula**: $\mu = \bar{x}$ (estimativa pontual)

b) **O que é**? Quando queremos estimar a média de uma população, μ, tomamos uma amostra de tamanho n dessa população e calculamos a média da amostra $\bar{x}$. Tomamos, então, essa média $\bar{x}$ como se fosse a média da população μ, .

c) Exemplo: Queremos saber, por exemplo, qual é o salário médio dos egressos dos cursos de Administração do país. Porém, não temos condições de fazer contato com todos os egressos desses cursos. Então, tomamos uma amostra de n egressos, calculamos a média dessa amostra e assumimos que ela representa o salário médio de todos os administradores em funções executivas. Temos então:
n = número de egressos para os quais perguntamos o seu salário,
μ = salário médio de todos os egressos do curso e
$\bar{x}$ = salário médio dos egressos que nos responderam o valor do seu salário,

2.2. Cálculo do erro de estimativa

a) **Fórmula**: $e = \bar{x} - \mu$ $\quad e = z.\frac{\sigma}{\sqrt{n}}$ (*1)

b) **O que é**? Quando se toma uma amostra de tamanho n de uma população qualquer e se deseja estimar a sua média populacional μ, mediante o cálculo da média amostral, $\bar{x}$, incorre-se sempre em certo erro e. O erro da estimativa é definido pela diferença entre a média amostral $\bar{x}$ e a verdadeira média populacional μ. Assim, podemos escrever que: $e = \bar{x} - \mu$, conforme escrito acima em (*1). Porém, como não sabemos o verdadeiro valor da média da população, μ, estimamos que ela seja $\bar{x}$, média da amostra, mais ou menos um valor de erro e, a um grau de confiança $1 - \alpha$, com correspondente escore normal padrão dado por z, ou seja, z ~N(0; 1); a expressão (*1), acima, permite estimar este erro. O desvio padrão populacional, σ, dividido pela raiz quadrada do tamanho da amostra, $\sqrt{n}$, é chamado de erro padrão e, na verdade, é o desvio padrão da estatística que está estimando o parâmetro. Multiplicando o erro padrão pelo escore padronizado z, que corresponde ao grau de confiança que desejamos, obtém-se o valor do erro de estimativa da média.

Observações: esta fórmula (*1) se aplica: ❶ a populações normais, ❷ com desvios padrão populacionais, σ, conhecidos, ❸ para amostras maiores do que 30, n > 30, e ❹ para populações infinitas.

c) **Exemplo**: Deseja-se fazer uma pesquisa de salários entre executivos de empresas, formados em Administração. De pesquisas anteriores dos Conselhos Regionais de Administração no país, sabe-se que o desvio padrão dos salários de todos os executivos enquadrados nesta categoria é de \$ 6.250,00; foram entrevistados 600 executivos, cujo salário médio foi de \$ 9.600,00. Admite-se que esta população é infinita e deseja-se um grau de confiança de 95 %.

Solução:
Temos: $\sigma = 6.250$, $n = 600$ e $z = 1,96$ - para um grau de confiança de 95 %, z terá o valor de 1,96; basta consultar a tabela normal padronizada – ver anexos. Aplicamos, então, a fórmula (*1):

$$e = z.\frac{\sigma}{\sqrt{n}}; \qquad e = 1,96.\frac{6.250}{\sqrt{600}} = 500$$

2.3. Intervalo de Confiança

a) **Fórmula**: $P[(\bar{x} - e) < \mu < (\bar{x} + e)] = 1 - \alpha$ ou, substituindo e por seu valor na fórmula (*1):

$$\left(\bar{x} - z.\frac{\sigma}{\sqrt{n}}\right) < \mu < \left(\bar{x} + z.\frac{\sigma}{\sqrt{n}}\right)$$

b) **O que é**? O intervalo de confiança para a média populacional estimada pela média de uma amostra $\bar{x}$ é um intervalo de valores, acima e abaixo dessa média, dentro do qual se acredita que possa estar a verdadeira média da população μ, com um grau de confiança de tantos pontos percentuais quantos queiramos, 95 % por exemplo, determinado por z.
Este grau de confiança significa que, se tomássemos muitas amostras da mesma população e construíssemos intervalos de confiança para as suas médias, poderíamos estar seguros de que 95 % destes intervalos conteriam a verdadeira média populacional μ.

c) **Exemplo**: Na pesquisa sobre os salários dos executivos de empresas, formados em Administração, encontrou-se uma média amostral, $\bar{x}$, de 9.600,00 e, no item 2.2. acima estimamos um erro de 500,00 para esta média. Por conseguinte, o intervalo de confiança para este salário médio, com um grau de confiança de 95 %, será de:

$(\bar{x} - e) < \mu < (\bar{x} + e)$ $\qquad$ $(9.600 - 500) < \mu < (9.600 + 500)$ $\qquad$ ou:
$9.100 < \mu < 10.100$

Podemos dizer, então, que, se tomássemos muitas amostras de 600 salários de executivos formados em Administração (lembrar que foi tomada uma amostra de 600 executivos) poderíamos estar seguros de que 95 % dos intervalos de confiança construídos com aquelas médias conteriam a verdadeira média populacional.

2.4. Tamanho da Amostra – População Infinita

a) **Fórmula**: de (*1): $e = z.\frac{\sigma}{\sqrt{n}}$ deduz-se que: $n = \left(z.\frac{\sigma}{e}\right)^2$ (*2)

OBS.: Muitas vezes o desvio padrão populacional σ é desconhecido, então o que se faz é trabalhar com o erro e como múltiplo ou fração do desvio padrão, p.ex., e = 2σ ou e = σ/4. Dessa forma na expressão (*2) o parâmetro σ é eliminado da expressão.

b) **O que é**? Quando queremos calcular quantos elementos precisamos tomar na nossa amostra aplicamos a fórmula (*2) acima. Cabe lembrar, como se observa na fórmula, que o tamanho da amostra irá depender: 1º) do desvio padrão da população, σ, do tamanho do erro que estamos dispostos a aceitar, e, e do grau de confiança que desejamos; 90 %, 95 %, 99 %, etc....., representado pelo escore z.
Para o uso desta fórmula (*2) aplicam-se as mesmas restrições já referenciadas em 2.2.: ela se aplica ❶ à população normal, ❷ com desvio padrão populacional, σ, conhecido, ❸ para amostras maiores do que 30, n > 30, e ❹ para populações infinitas.

c) **Exemplo**:

Uma universidade está pesquisando o tempo médio diário que alunos das faculdades do país passam assistindo televisão. A pesquisa determina um grau de confiança de 96 % e deseja-se a estimativa com um erro de 0,25 horas, ou quinze minutos. Um estudo piloto, realizado especificamente para esta pesquisa mostrou um desvio padrão de 1,87 horas. Neste caso o desvio padrão populacional σ foi estimado pelo desvio padrão amostral s = 1,87 horas em um trabalho anterior. Porém, se não existisse esse trabalho prévio, o que geralmente é a realidade, poderia se supor o erro como 20% de σ, ou seja, e = 0,20σ. Precisa-se calcular quantas entrevistas serão necessárias para esta pesquisa.

<u>Solução</u>:
Basta aplicar a fórmula (*2), na qual:
z = 2,06 basta consultar a tabela normal padronizada – ver anexos
σ = 1,87
e = 0,25

$$n = \left(z.\frac{\sigma}{e} \right)^2 \qquad n = \left(2{,}06.\frac{1{,}87}{0{,}25} \right)^2 = 237{,}4$$

Resposta: 238 entrevistas.

3. CÁLCULO A PARTIR DE VALORES QUANDO AS CONDIÇÕES ❶ , ❷ , ❸ OU ❹ NÃO ESTÃO SATISFEITAS

3.1. E se o desvio padrão da população for desconhecido? Viola-se a restrição ❷

a) **Condição**: para aplicar as fórmulas (*1) e (*2), precisamos conhecer o desvio padrão da população. Mas, como poderemos calcular erro, intervalo de confiança e média, se não conhecemos os desvio padrão da população, como acontece na grande maioria das pesquisas? Se $\sigma = ?$

b) **O que fazer**? Existem duas possibilidades, estatisticamente válidas, para contornar esta dificuldade:

1ª) Podemos estimar o desvio padrão populacional empregando uma regra prática, decorrente de experiências empíricas:

$$\sigma = \frac{\text{amplitude}}{4} \qquad (*3)$$

2ª) Ou então, poderemos usar o desvio padrão amostral, s, desde que seja tomada uma amostra igual ou maior do que 31 elementos; desde que $n \geq 31$. O Teorema do Limite Central garante que amostras maiores do que 30 são normais, independentemente do formato da distribuição de probabilidade da população da qual está sendo retirada a amostra.
3ª.) Ou então suporíamos, como descrito anteriormente, o erro como um múltiplo ou uma fração do desvio padrão σ.

c) Exemplo 1: A cadeia de lanchonetes Marreco Donaldo, MD, está interessada em conhecer o gasto médio por pessoa, dos clientes de uma cadeia de lanchonetes concorrente. Os executivos da MD admitem que a despesa de um cliente possa variar entre R$ 3,00 e R$ 15,00. Eles desejam um grau de confiança de 98 % e admitem uma margem de erro de R$ 0,25. Pede-se calcular o tamanho da amostra para esta pesquisa.

<u>Solução</u>:
Basta aplicar a fórmula (*2) mas, antes, é preciso calcular o desvio padrão. O desvio padrão poderia ser calculado 1) mediante uma pesquisa inicial, com 31 amostras ou mais, ou 2) mediante o emprego da fórmula empírica (*3), já que se sabe de antemão, por medições anteriores, na própria cadeia de lanchonetes MD, que as despesas individuais variam entre R$ 3,00 e R$ 15,00

1º Cálculo do desvio padrão pela fórmula (*3):

$$\sigma = \frac{\text{amplitude}}{4} \qquad s = \frac{15-3}{4} = 3,0$$

2º Calculo do tamanho da amostra – sabemos da tabela normal padronizada que, para um grau de confiança de 98 % z será igual a 2,33:

$$n = \left(z.\frac{\sigma}{e}\right)^2 \qquad n = \left(2,33.\frac{3}{0,25}\right)^2 = 781,76 \qquad \text{ou seja: 782 amostras.}$$

d) **<u>Exemplo 2</u>**: Deseja-se conduzir uma pesquisa de salários dos executivos formados em Administração no Brasil; admite-se que esta população seja infinita. Precisa-se calcular quantas entrevistas serão necessárias nesta pesquisa. Deseja-se um grau de confiança de 95 % e admite-se um erro de 500,00 a mais ou a menos, no salário médio.

<u>Solução</u>:
Como não conhecemos o desvio padrão da população, realizamos uma pesquisa amostral entre 31 executivos formados em Administração, escolhidos randomicamente no país.
A pesquisa apresentou o resultado da Tabela I abaixo:

TABELA I – SALÁRIOS DE 31 ADMINISTRADORES

26.300	20.000	22.800	19.700	19.300
18.200	14.600	21.000	25.000	22.400
4.200	19.800	19.800	23.500	5.300
9.800	8.500	15.400	16.000	12.500
18.200	7.600	16.400	12.500	24.200
13.400	6.800	26.000	21.300	5.900

O cálculo mostrou um desvio padrão de $ 6.538,00 que poderá ser empregado em substituição do desvio padrão populacional. O erro admitido é de 500,00 e o escore z para o grau de confiança desejado é de 1,96. Então, o tamanho da amostra será dado por:

$$n = \left(1{,}96.\frac{6.538}{500}\right)^2 = 656{,}8 \qquad \text{ou:} \qquad 657 \text{ entrevistas}$$

3.2. E se a população for finita, o que acontece com o erro? Viola-se a restrição ❹

a) **<u>Condição</u>**: Na maior parte das pesquisas em Administração a população tem um número limitado de elementos. Contudo, a fórmula (*1) pressupõe que a população seja infinita. Assim, se a aplicássemos a populações finitas estaríamos violando a restrição ❹. Procuramos, por exemplo, conhecer certos aspectos de segmentos específicos de mercado da Região Metropolitana de Curitiba e, tanto a população desta região como um todo, quanto qualquer um dos seus segmentos, possuem um número definido, limitado, de pessoas.

b) **<u>O que fazer</u>**? Basta multiplicar o erro padrão da fórmula (*1) por um fator de correção para populações finitas, conhecido vulgarmente como

CPF, ou seja, coeficiente de população finita: $\sqrt{\frac{N-n}{N-1}}$. A fórmula para o cálculo do erro fica sendo então:

$$e = z.\frac{\sigma}{\sqrt{n}}.\sqrt{\frac{N-n}{N-1}} \qquad (*4)$$

Onde:
N = tamanho da população
n = tamanho da amostra
Existe, neste caso uma condição a ser satisfeita: o número de amostras a serem colhidas, n, deverá ser igual a pelo menos 5 % do tamanho da população N: $n \geq 0{,}05.N$

c) **<u>Exemplo</u>**: O município de Arapiraca das Dornas deseja fazer uma pesquisa do peso de papéis descartados mensalmente pelas residências da cidade, para efeito do planejamento da coleta de lixo. O peso médio do papel descartado em um mês, por uma amostra de 62 residências, foi de 9,4281 kg. e o desvio padrão dessa amostra foi de 4,1681 kg. Como a cidade possui 2.637 residências, deseja-se conhecer o intervalo de confiança para esta média, com grau de confiança de 95 %.

<u>Solução</u>:
Conhecemos o desvio padrão amostral, s, que, para 62 indivíduos, substitui o desvio padrão populacional. Trata-se de uma amostra finita, na qual:
n = 62
N = 2.637
$\bar{x}$ = 9,4281
s = 4,1681
z para o grau de confiança de 95 % = 1,96
Podemos, então empregar a fórmula (*4):

$$e = z.\frac{s}{\sqrt{n}}.\sqrt{\frac{N-n}{N-1}} \qquad e = 1{,}96.\frac{4{,}1681}{\sqrt{62}}.\sqrt{\frac{2.637-62}{2.637-1}} = 1{,}025 \text{ kg}$$

Donde, o intervalo de confiança é definido por:

$(\bar{x} - e) < \mu < (\bar{x} + e) \qquad (9{,}4281 - 1{,}025) < \mu < (9{,}4281 + 1{,}025)$ ou: $8{,}4031 < \mu < 10{,}4531$

3.3. Como fica o tamanho da amostra se a população for finita? Viola-se ❹

a) **<u>Condição</u>**: Da mesma forma que acontece com a fórmula para o cálculo do erro, (*1), também a fórmula de cálculo do tamanho da amostra, (*2), torna-se inadequada se estivermos tratando com uma população finita, como no caso da maior parte das pesquisas em Administração, em que a população tem um número limitado de elementos. Neste caso estaríamos violando a condição ❹ se aplicássemos a fórmula (*2), uma vez que ela permite calcular o tamanho da amostra em pesquisas com populações infinitas.

b) **<u>O que fazer</u>**? Basta multiplicar o erro padrão da fórmula (*2) pelo mesmo fator de correção para populações finitas. A partir da fórmula (*4), para o cálculo do erro em populações finitas, podemos facilmente deduzir uma fórmula para o cálculo da amostra em populações finitas.

Assim, de (*4): $e = z.\frac{\sigma}{\sqrt{n}}.\sqrt{\frac{N-n}{N-1}}$ tiramos que:

$$n = \frac{z^2.\sigma^2.N}{e^2.(N-1) + z^2.\sigma^2} \qquad (*5)$$

c) **<u>Exemplo</u>**: Seja o mesmo caso do exemplo de 3.2. O município de Arapiraca das Dornas deseja fazer uma pesquisa do peso de papéis descartados mensalmente pelas residências da cidade, para efeito de planejamento da coleta de lixo. O peso médio de papel descartado, por uma amostra teste de 31 residências, foi de 9,4281 kg. e o desvio padrão dessa amostra foi de 4,1681 kg. Como a cidade possui 2.637 residências, deseja-se calcular o tamanho da amostra a ser coletada, com grau de confiança de 95 %.

Solução:
Com os mesmos dados, já havíamos calculado o erro. Basta, agora, aplicarmos a fórmula (*5), na qual:
N = 2.637
$\bar{x}$ = 9,4281
s = 4,1681
z para o grau de confiança de 95 % = 1,96
e = 1,025
n = ?

$$n = \frac{z^2.\sigma^2.N}{e^2.(N-1)+z^2.\sigma^2}$$

$$n = \frac{1{,}96^2.4{,}1681^2.2637}{1{,}025^2.(2637-1)+1{,}96^2.4{,}1681^2} = 62$$

3.4. E se a amostra for menor do que 30 e a distribuição da população é desconhecida? Violam-se simultaneamente as restrições ❷ e ❸

a) **Condição**: Pode ocorrer de não ser possível colher uma amostra maior do que 30 e/ou, além disso, desconhecermos o desvio padrão da população (embora este possa ser estimado, conforme foi mostrado em 3.1.). Se aplicássemos, agora, o escore z, tirado da tabela normal padronizada, estaríamos violando as restrições ❷ e ❸ .

b) **O que fazer**? Neste caso emprega-se a distribuição t de Student, com n – 1 graus de liberdade, no lugar do escore z. As fórmulas para o cálculo do erro são as mesmas, apenas substituindo-se z por t. Observa-se que, mesmo assim é preciso que a população tenha distribuição normal. As fórmulas para o cálculo do erro passam, então, a ser:

1º Para populações infinitas:

$$e = t.\frac{\sigma}{\sqrt{n}} \qquad (*6)$$

E, aqui ocorre um problema. Para obtermos o valor de t há necessidade do número de graus de liberdade ν = n – 1, que depende do tamanho da amostra n. Então, o que se faz é tomar uma amostra piloto de tamanho n_0, estimar o desvio padrão por s_0, obter t com n_0 -1 graus de liberdade e dimensionar o tamanho da amostra por n' (fixado o erro de estimativa). Se o tamanho da amostra obtido n' foi maior que n_0 deve-se tomar mais n' – n_0 outras observações; se não recalcula-se o erro e pronto.

2º Para populações finitas:

$$e = t.\frac{\sigma}{\sqrt{n}}.\sqrt{\frac{N-n}{N-1}} \qquad (*7)$$

c) **Exemplo**: Uma pesquisa sobre uso do tempo constatou que 20 Administradores, selecionados aleatoriamente, gastam uma média de 2,40 horas por dia, desvio padrão de 1,30 horas, com trabalho meramente burocrático. Os dados aparentam uma distribuição normal. Pede-se construir um intervalo de confiança de 95 % para esta média amostral, considerando todos os Administradores.

Solução:

Como temos uma amostra menor do que 30 e uma distribuição normal, em vez do escore z, aplicamos a tabela t. Além disso, como se tratam de todos os Administradores, podemos considerar uma população infinita. Assim, aplicamos a fórmula (*6) com os dados:

$\bar{x} = 2{,}4$

$n = 20$

$\sigma = 1{,}3$

t – para 95 % e 19 = 20 – 1 graus de liberdade = 2,093 (ver tabela)

$e = t.\frac{\sigma}{\sqrt{n}}$ $\quad$ $e = 2{,}093.\frac{1{,}3}{\sqrt{20}} = 0{,}6$ $\quad$ Então:

$(\bar{x} - e) < \mu < (\bar{x} + e)$ $\quad$ $(2{,}4 - 0{,}6) < \mu < (2{,}4 + 0{,}6)$ ou: $1{,}8 < \mu < 3{,}0$

Observação: como se tratam de amostras pequenas, não há sentido em calcular o tamanho da amostra.

4. CÁLCULO A PARTIR DE PROPORÇÕES

Em diversos tipos de pesquisas o interesse maior está na proporção que determinada parcela da população representa, em relação a uma população total, cuja característica está-se examinando. Podemos estar interessados, por exemplo, em saber qual é a proporção de estudantes universitários que possuem computador portátil, o chamado lap-top. Ou podemos pretender saber qual é a porcentagem de residências que estão com a televisão sintonizada em determinado canal, em determinado horário.
Para este tipo de pesquisa empregam-se procedimentos de cálculo mediante proporções. Eles são largamente empregados nas pesquisas de opinião, especialmente nas pesquisas eleitorais.

4.1. Estimativa de uma Proporção Populacional

a) **Fórmula**: $\hat{p} = \frac{x}{n}$ (*8) onde: x = sucessos e n = amostra de tamanho n.

b) **O que é**? Trata-se uma simples proporção, como o número de eleitores que votariam em determinado candidato, em relação ao número total de eleitores entrevistados. Tais proporções são calculadas sempre na base 1 e, assim, existem três possibilidades: 1) proporção, 2) probabilidade e 3) porcentagem.

c) **Exemplo 1** - proporção: 527 residências estão com a TV ligada em determinado canal de televisão, em determinado horário, das 1628 residências entrevistadas.

Solução: Temos:
n = população amostrada = 1.628
x = sucessos = televisões ligadas no determinado canal = 527

$\hat{p} = \frac{527}{1.628} = 0{,}3237 \text{ na base 1, ou} : 32{,}37\,\%$ das residências estão sintonizando o canal.

d) **<u>Exemplo 2</u>** - <u>probabilidade</u>: 3 casos prováveis em 15 casos possíveis.

<u>Solução</u>: Temos:
n = população amostrada = 15
x = sucessos = casos prováveis = 3

$\hat{p} = \frac{3}{15} = 0{,}2 \text{ ou} : 20{,}00\,\%$ ou, como costumamos dizer: uma probabilidade de 20 %.

e) **<u>Exemplo 3</u>** - <u>porcentagem</u>: 32 % das residências estão sintonizando determinado canal de televisão.

<u>Solução</u>: Temos:
n = população amostrada (como se trata de porcentagem) = 100
x = sucessos = 32

$\hat{p} = \frac{32}{100} = 0{,}32 \text{ ou} : 32{,}00\,\%$

4.2. Cálculo do erro de estimativa de uma proporção e intervalo de confiança

a) **<u>Fórmula</u>**: $e = z_{\alpha/2}\sqrt{\frac{\hat{p}.\hat{q}}{n}}$ (*9)

b) onde: $\hat{p} = \frac{x}{n}$ e $\hat{q} = 1 - \hat{p}$

c) **<u>O que é</u>**? O erro **de estimativa** de uma proporção é a diferença, para mais ou para menos, que aceitamos em nossa pesquisa, em função do nível de confiança desejado e representado pelo escore z. Da mesma forma que no caso da média temos o escore padronizado z como o coeficiente do erro padrão para obter o erro de estimativa. O valor de $\hat{p}$ já conhecemos e o valor de $\overline{q}$ é o complemento de $\hat{p}$ para 1.

Ou seja: $\hat{q} = 1 - \hat{p}$. Das 527 residências que estão com a televisão sintonizada no canal cuja audiência se está pesquisando, das 1628 entrevistadas, por exemplo, temos:

$$\hat{p} = \frac{527}{1628} = 0{,}3237 \qquad \text{e} \qquad \overline{q} = \frac{1628 - 527}{1628} = 0{,}6763 \qquad \text{donde:}$$

$$\hat{p} + \hat{q} = 0{,}3237 + 0{,}6763 = 1 \qquad \text{donde generalizando:} \qquad \hat{p} + \hat{q} = 1$$

Quando $\hat{p}$ for desconhecido faz-se $\hat{p} \times \hat{q} = 0{,}25$, que é o maior valor que pode ser obtido pelo produto $\hat{p} \times \hat{q}$:

TABELA II – PRODUTOS p x q

$\hat{p}$	$\hat{q}$	$\hat{p} \times \hat{q}$
0,4	0,6	0,24
0,5	0,5	0,25
0,6	0,4	0,24

d) **<u>Exemplo 1</u>**: Uma empresa de seguros deseja estimar a porcentagem de motoristas que usam o celular ao volante. Um amostra de 850 motoristas acusou 544 que usam celular enquanto dirigem. Pede-se determinar o intervalo de confiança com um grau de confiança de 90 % para esta porcentagem.

Solução: Calculamos a proporção e depois o erro, aplicando a fórmula (*9):

$\hat{p} = \frac{544}{850} = 0{,}64$ Donde: $\hat{q} = 0{,}36$ *GC* = 90 % Donde: z = 1,645

Aplicando a fórmula (*9):

$$e = z_{\alpha/2}\sqrt{\frac{\hat{p}.\hat{q}}{n}} \qquad e = 1{,}645\sqrt{\frac{0{,}64 \times 0{,}36}{850}} = 0{,}027$$

Intervalo de confiança = $(0{,}64 - 0{,}027) < \hat{p} < (0{,}64 + 0{,}027) =$ $0{,}613 < \hat{p} < 0{,}667$

e) **Exemplo 2**: Como ficaria o intervalo de confiança se não fosse possível calcular o valor de $\hat{p}$?

Solução: Basta fazer o produto $\hat{p} \times \hat{q} = 0{,}25$

GC = 90 % Donde: z = 1,645

Aplicando a fórmula (*9): $e = z_{\alpha/2}\sqrt{\frac{\hat{p}.\hat{q}}{n}}$ $\qquad e = 1{,}645\sqrt{\frac{0{,}25}{850}} = 0{,}028$

Intervalo de confiança = $(0{,}64 - 0{,}028) < \hat{p} < (0{,}64 + 0{,}028) =$ $0{,}612 < \hat{p} < 0{,}668$

4.3. Tamanho da amostra em população infinita

a) **Fórmula**: Da fórmula (*9) deduzimos que $n = z^2_{\alpha/2} . \frac{\hat{p}.\hat{q}}{e^2}$ (*10)

Quando $\hat{p}$ é desconhecido, fazer: $\hat{p}.\hat{q} = 0,25$; que corresponde ao pior caso para o erro padrão, ou seja, se p = 0,5 o erro padrão é máximo.

b) **<u>O que é</u>**? Esta fórmula permite estabelecer o número de amostras para calcular uma proporção $\hat{p}$ desejada, a um grau de confiança estabelecido por z, aceitando-se o erro percentual representado por e.
Geralmente esta é a fórmula empregada pelos institutos de pesquisa eleitoral; quando uma pesquisa informa que o resultado da pesquisa poderá ser 2,5 % maior ou menor, é ao valor de e que estão se referindo. Note-se que nem o tamanho da população e nem o da amostra constam da fórmula; tais valores já estão implícitos no valor de $\hat{p}$.
Note-se, novamente, e além disso, que, para o mesmo valor de z e de e, n será máximo quando $\hat{p}.\hat{q} = 0,25$.

c) **<u>Exemplo</u>**: Um instituto de pesquisas deseja avaliar o desempenho de um político em campanha. Para isso, precisa estabelecer o tamanho da amostra de eleitores a serem entrevistados. Deseja-se uma margem de erro de 3 pontos percentuais, para mais ou para menos, e um nível de confiança de 95 %.

<u>1ª Hipótese</u>: estudos anteriores mostravam que o político detinha 18 % das preferências do eleitorado.
<u>Solução</u>: Basta aplicar a fórmula (*10), onde z = 1,96, e = 0,03, $\hat{p} = 0,18$ e $\hat{q} = 0,82$

$$n = z^2_{\alpha/2}.\frac{\hat{p}.\hat{q}}{e^2} \qquad n = 1,96^2.\frac{0,18 \times 0,82}{0,03^2} = 630,01 = 631 \quad \text{entrevistas}$$

<u>2ª Hipótese</u>: não se tem nenhuma informação anterior sobre o desempenho do político.
<u>Solução</u>: Basta aplicar a fórmula (*10), onde z = 1,96, e = 0,03, e considerando o produto $\hat{p}.\hat{q} = 0,25$:

$n = z^2_{\alpha/2} . \frac{\hat{p}.\hat{q}}{e^2}$ $\quad$ $n = 1{,}96^2 . \frac{0{,}25}{0{,}03^2} = 1067{,}11 = 1068$ entrevistas

4.4. Tamanho da amostra em população finita

a) **Fórmulas**: $e = z_{\alpha/2}\sqrt{\frac{\hat{p}.\hat{q}}{n}}.\sqrt{\frac{N-n}{N-1}}$ (*11) $\quad$ $n = \frac{N.\hat{p}.\hat{q}.z^2_{\alpha/2}}{\hat{p}.\hat{q}.z^2_{\alpha/2} + (N-1).e^2}$

(*12)

b) **O que é**? Trata-se do mesmo procedimento empregado para o cálculo do tamanho de amostra em população infinita. Só que, neste caso, o valor encontrado deverá ser multiplicado pelo mesmo fator de correção para populações finitas, já apresentado anteriormente em 3.2. A fórmula (*9) transforma-se na fórmula (*11) e a fórmula (*10) transforma-se na fórmula (*12).

c) **Exemplo 1**: Um instituto de pesquisas foi contratado para avaliar determinado programa de televisão em uma pequena cidade do interior, cuja população soma 5.000 habitantes. Deseja-se um grau de confiança de 97 % e admite-se uma margem de erro de 2 pontos percentuais.

Solução: Basta aplicar a fórmula (*12). Temos que:
Para o nível de confiança de 97 % teremos z = 2,17
e = 0,02
Como não temos nenhuma informação anterior, faremos $\hat{p}.\hat{q} = 0{,}25$
N = 5000

$$n = \frac{N.\hat{p}.\hat{q}.z^2_{\alpha/2}}{\hat{p}.\hat{q}.z^2_{\alpha/2} + (N-1).e^2}$$

$$n = \frac{5000 \times 0{,}25 \times 2{,}17^2}{0{,}25 \times 2{,}17^2 + (5000-1) \times 0{,}02^2} = 1852{,}83 = 1853$$

d) **Exemplo 2**: Com os dados já calculados no exemplo 1, vamos confirmar o valor da margem de erro.

Solução: Basta aplicar a fórmula (*11). Temos que:
Para o nível de confiança de 97 % teremos z = 2,17
e = ? não conhecemos.
Como não temos nenhuma informação anterior, faremos $\hat{p}.\hat{q} = 0{,}25$
N = 5000
n = 1852,83

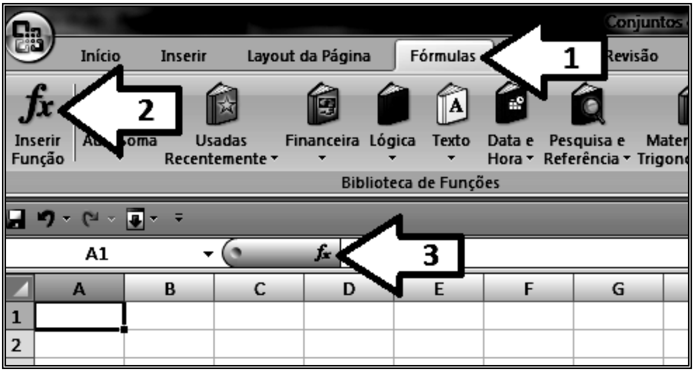
$$e = z_{\alpha/2}\sqrt{\frac{\hat{p}.\hat{q}}{n}}.\sqrt{\frac{N-n}{N-1}} \qquad e = 2{,}17\sqrt{\frac{0{,}25}{1852{,}83}}.\sqrt{\frac{3.147{,}17}{4999}} = 0{,}02$$

5. USO DA PLANILHA EXCEL

Os imensos recursos operacionais oferecidos pela planilha EXCEL facilitam sobremaneira os cálculos estatísticos de modo geral e do tamanho da amostra em particular. Os procedimentos descritos a seguir são baseados na planilha EXCEL do Office 2007. Existem duas maneiras para executar as funções estatísticas:
1) a partir da seqüência necessária para inserir funções estatísticas ou
2) a partir da sintaxe da função desejada.

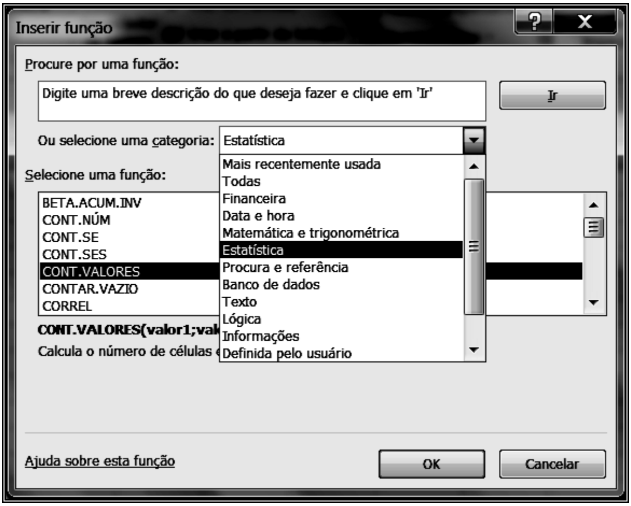
5.1. Inicio da seqüência necessária para inserir funções estatísticas

FIGURA I – DUAS POSSIBILIDADES PARA INSERIR FUNÇÕES



A seqüência necessária para inserir funções estatísticas pode ser iniciada a partir de dois comandos (Figura I acima):
a) clicar no menu Fórmulas (1) e, depois, clicar em ***fx*** (2) ou (3) ou :
b) clicar diretamente no comando ***fx*** (3); este comando está disponível em todos os menus da planilha: Início, Inserir, Layout da Página etc.....
Depois que se clica em ***fx*** , em (2) ou (3), abre-se o sub-menu da Figura II abaixo:

FIGURA II – COMANDOS PARA INSERIR FUNÇÕES ESTATÍSTICAS



Neste sub-menu selecionar a categoria Estatística e, depois, a função desejada. A função será executada na célula em que se encontrar o cursor, na planilha. Cada função exigirá o preenchimento de um sub-menu específico, para que se possa inserir os argumentos específicos de cada função. Por exemplo, para inserir a função desvio padrão; seguem-se os passos:

1º passo: clique em: ***fx***

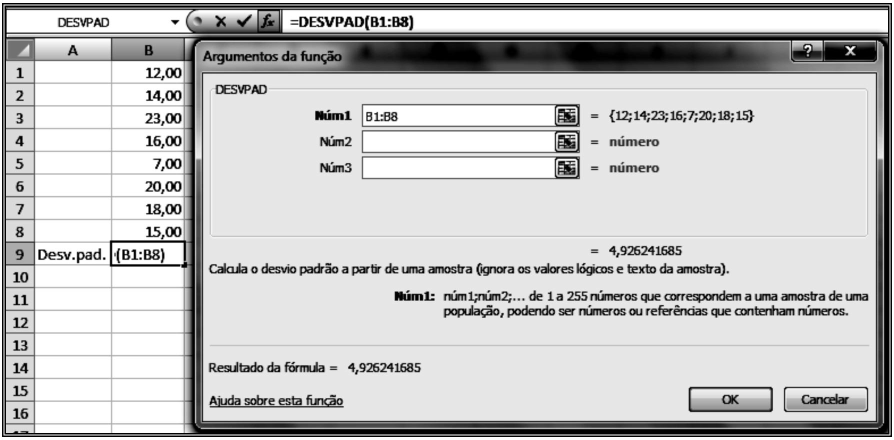
2º passo: selecione a categoria: Estatística

3º passo: dentro da categoria Estatística selecione: DESVPAD

4º passo: clique em: OK

Abrirá o seguinte sub-menu:

FIGURA III – SUB-MENU PARA INSERIR OS ARGUMENTOS DO DESVIO PADRÃO

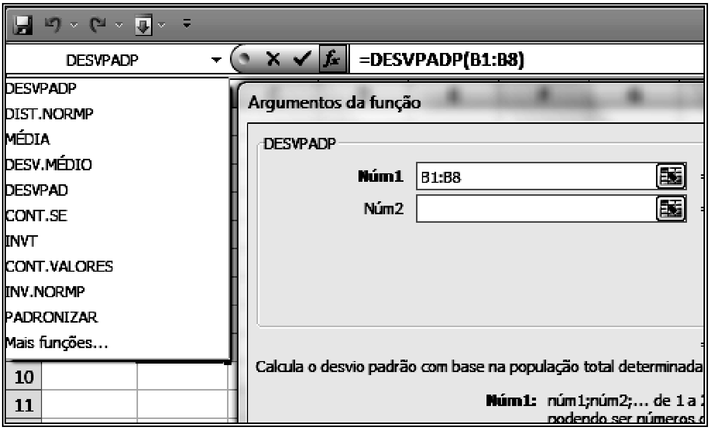


Neste sub-menu, a planilha pede que sejam inseridos os argumentos da função; no caso do desvio padrão os argumentos são apenas os valores sobre os quais se quer calcular o desvio padrão.

Observações:

1) Note-se que, no exemplo acima, a própria planilha já sugeriu a série sobre a qual calcular o desvio padrão: células B1 a B8. Se a planilha não sugerir a série de valores desejada, basta selecionar com o mouse o conjunto desejado. Clicando em OK, o resultado aparecerá na célula B9.
2) Note-se, também, que no topo da planilha, ao lado de ***fx*** aparecerá a sintaxe da função: =DESVPAD(B1:B8)
3) Note-se, também, que, à esquerda de ***fx*** , aparecerá parte da sintaxe: DESVPAD. Clicando na seta que aparece ao lado, abre-se outro sub-menu com as demais funções estatísticas, como na figura IV adiante. Clicando em qualquer uma dessas funções, reabrirá o sub-menu dos argumentos da função sobre a qual se clicou, conforme a Figura III.

FIGURA IV – UM SEGUNDO SUB-MENU PARA INSERIR FUNÇÕES ESTATÍSTICAS



5.2. Inserindo as funções a partir de *fx*

Todas as funções estatísticas, necessárias ao cálculo do tamanho da amostra, podem ser executadas a partir deste procedimento; os passos 1º e 2º acima: clicar em *fx* e selecionar a categoria Estatística – são comuns a todas. As operações necessárias ao cálculo do tamanho de amostras, e suas respectivas funções na planilha, são os seguintes:

1) Contagem do número de valores da Série: **CONT.VALORES**
A planilha conta o número de células não vazias, no intervalo de células que deverá ser selecionado logo a seguir, depois de clicar em OK, como argumento da função.
2) Soma da Série: **SOMA**
A planilha calcula a soma de todos os números no intervalo de células que deverá ser selecionado logo a seguir, depois de clicar em OK, como argumento da função.
Observação:

A função soma encontra-se na categoria Matemática e Trigonométrica e não na categoria Estatística.

3) Média da Série: **MÉDIA**

A planilha calcula a média aritmética de todos os números no intervalo de células que deverá ser selecionado logo a seguir, depois de clicar em OK, como argumento da função.

4) Variância Amostral: **VAR**

A planilha calcula a variância amostral – divide a soma dos quadrados da diferença entre cada valor e a média de todos os valores por n-1 - de todos os números no intervalo de células que deverá ser selecionado logo a seguir, depois de clicar em OK, como argumento da função.

5) Variância Populacional: **VARP**

A planilha calcula a variância populacional – divide a soma dos quadrados da diferença entre cada valor e a média de todos os valores por n - de todos os números no intervalo de células que deverá ser selecionado logo a seguir, depois de clicar em OK, como argumento da função.

6) Desvio Padrão Amostral **DESVPAD**

A planilha calcula o desvio padrão amostral – extrai a raiz quadrada da variância amostral - de todos os números no intervalo de células que deverá ser selecionado logo a seguir, depois de clicar em OK, como argumento da função.

7) Desvio Padrão Populacional **DESVPADP**

A planilha calcula o desvio padrão populacional – extrai a raiz quadrada da variância populacional - de todos os números no intervalo de células que deverá ser selecionado logo a seguir, depois de clicar em OK, como argumento da função.

8) Escore z padronizado: **INV.NORMP**

A planilha calcula o escore z padronizado, para uma probabilidade que precisará ser fornecida como argumento, logo a seguir, depois de clicar em OK. Observa-se que este escore z corresponde a uma distribuição normal padronizada, cuja média é 0 e desvio padrão 1 e encontra-se na tabela ao final.

9) Valor da Estatística t: **INVT**

A planilha calcula o valor da tabela t de Student, para uma probabilidade e um número de graus de liberdade, que precisarão ser fornecidos como argumento, logo a seguir, depois de clicar em OK.

10) Maior Valor da Série: **MÁXIMO**

A planilha seleciona o maior valor da série.

11) Menor Valor da Série: **MÍNIMO**

A planilha seleciona o menor valor da série.

12) Mediana da Série: **MED**

A planilha calcula a mediana da série. A mediana é o valor que separa os valores da série em duas partes. É definida também como o valor central, que separa o segundo do terceiro quartis.

13) Moda da Série: **MODO**

A planilha calcula a moda da série. Moda é o valor que mais ocorre na série.

14) Erro Padrão da Série: **INT.CONFIANÇA**

A planilha calcula o erro padrão da série, para uma população infinita, segundo a fórmula de 2.2. Após clicar em OK, um sub-menu pedirá para que seja fornecido o valor de alfa, ou seja: o percentual de erro aceito: 5 %, 2,5 % etc..... Ele deverá ser digitado na forma centesimal: 0,05 ou 0,025 e assim por diante. O sub-menu pedirá, também, o valor do desvio padrão e o tamanho da amostra. É preciso notar que este valor, que a planilha chama de Int.Confiança (intervalo de confiança) é, na verdade, o erro de estimativa. O intervalo de confiança para a média é um erro de estimativa abaixo da média e outro acima dela.

EXEMPLO 1 – ESTATÍSTICA DESCRITIVA

Seja calcular as estatísticas da série de valores da Tabela III abaixo, empregando as funções estatísticas a partir de ***fx*** . Os números da Tabela III estão digitados na planilha, coluna D linhas 1 a 50.

Admitamos tratar-se do resultado de um teste do espaço necessário em metros, para frenagem de um carro em determinada velocidade. São 50 observações, medidas em 50 testes de frenagem.

Na Tabela IV encontram-se as estatísticas dos espaços de frenagem da Tabela III: contagem, soma, média, variância, desvio padrão, escore z

para um grau de confiança de 95 %, maior e menor valores da série, amplitude, mediana e moda.
Na primeira coluna está a estatística, na do meio o está o valor da estatística. Finalmente, na coluna da direita está a sintaxe do comando.

TABELA III – ESPAÇOS DE FRENAGEM

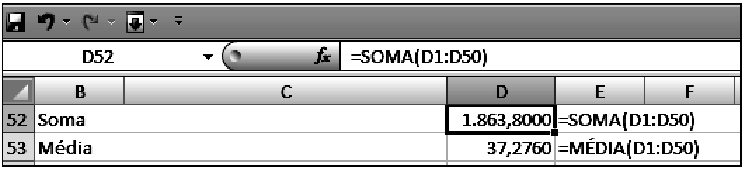
33,2	47,7	33,2	43,4	49,2	21,0	33,3	33,1	38,2
28,1	37,0	32,8	37,3	39,7	43,7	31,4	29,9	
38,3	38,8	32,6	40,7	24,3	43,7	34,6	43,1	
36,7	46,5	47,8	41,5	32,2	34,3	39,3	44,3	
41,7	38,6	43,3	36,8	35,6	36,6	38,2	29,6	
48,5	38,0	23,7	36,3	29,5	44,4	33,6	34,9	

TABELA IV – ESTATÍSTICAS DO ESPAÇO DE FRENAGEM

Contagem do número de valores:	50,0000	=CONT.VALORES(D1:D50)
Soma:	1.863,8000	=SOMA(D1:D50)
Média:	37,2760	=MÉDIA(D1:D50)
Variância amostral:	41,9129	=VAR(D1:D50)
Variância populacional:	41,0746	=VARP(D1:D50)
Desvio padrão amostral:	6,4740	=DESVPAD(D1:D50)
Desvio padrão populacional:	6,4089	=DESVPADP(D1:D50)
Escore z padronizado - gc 95 %:	1,6449	=INV.NORMP(0,95)
Maior valor da série:	49,2000	=MÁXIMO(D1:D50)
Menor valor da série:	21,0000	=MÍNIMO(D1:D50)
Amplitude:	28,2000	Obs. 1
Mediana:	37,1500	=MED(D1:D50)
Moda:	33,2000	=MODO(D1:D50)

A função soma aparecerá ao lado de ***fx*** , quando se executa a função, conforme Figura V abaixo .

FIGURA V – APARÊNCIA DA FUNÇÃO SOMA

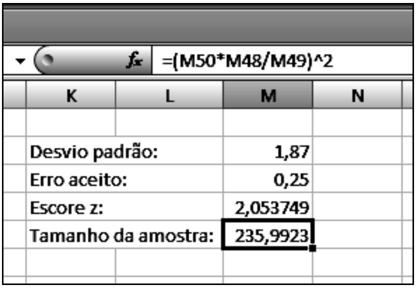


EXEMPLO 2 – CÁLCULO DO TAMANHO DA AMOSTRA

2.a. População infinita, desvio padrão conhecido
A planilha EXCEL não dispõe de uma função específica para o cálculo do tamanho da amostra. Assim, para este cálculo será necessário escrever a função. Seja, por exemplo, resolver o exemplo c, de 2.4.:
Uma universidade está pesquisando o tempo médio diário que alunos das faculdades do país passam assistindo televisão. A pesquisa determina um grau de confiança de 96 % e deseja-se a estimativa com um erro de 0,25 horas, ou quinze minutos. Um estudo piloto, realizado especificamente para esta pesquisa mostrou um desvio padrão de 1,87 horas. Precisa-se calcular quantas entrevistas serão necessárias para esta pesquisa.

Solução:
Conhecemos:
$\sigma = 1{,}87$
$e = 0{,}25$
Digitaram-se o desvio padrão e o erro aceito nas células M2 e M3. Empregando a função =INV.NORMP(0,98) obteve-se o escore z para o grau de confiança de 96 % na célula M4. Finalmente, escreveu-se a fórmula para o cálculo do tamanho da amostra na célula M5, conforme demonstrado na Figura VI abaixo,

FIGURA VI – POPULAÇÃO INFINITA, DESVIO PADRÃO CONHECIDO


=(M50*M48/M49)^2

K	L	M	N
Desvio padrão:		1,87	
Erro aceito:		0,25	
Escore z:		2,053749	
Tamanho da amostra:		235,9923	

Observar que, em 2.4. obteve-se, para este mesmo problema, o resultado de 237,4 entrevistas. A diferença decorre da precisão no cálculo do escore z. Em 2.4. o valor do escore z foi tomado da tabela normal padronizada e aqui, foi calculado pela planilha.

2.b. <u>População infinita, desvio padrão desconhecido</u>
<u>Solução</u>:
Neste caso, basta empregar o desvio padrão estimado, conforme referenciado em 3.1. e empregar o mesmo procedimento acima, do exemplo 2.a.

2.c. <u>População finita, desvio padrão conhecido</u>
Seja, por exemplo, resolver o exemplo c, de 3.3.: O município de Arapiraca das Dornas deseja fazer uma pesquisa do peso de papéis descartados mensalmente pelas residências da cidade, para efeito de planejamento da coleta de lixo. O peso médio de papel descartado, por uma amostra teste de 31 residências, foi de 9,4281 kg. e o desvio padrão dessa amostra foi de 4,1681 kg. Como a cidade possui 2.637 residências, deseja-se calcular o tamanho da amostra a ser coletada, com grau de confiança de 95 % e é aceito um erro de 1,025 kg.

<u>Solução</u>:
Conhecemos:

N = 2.637
n = ?
$\bar{x}$ = 9,4281
σ = 4,1681
e = 1,025

Digitaram-se o tamanho da população e a média amostral nas células T1 e T2, respectivamente. Digitou-se o desvio padrão amostral na célula T3. Na célula T4, empregando a função =INV.NORMP(0,975) obteve-se o escore z para o grau de confiança de 95 % e, na célula T5 digitou-se o erro aceitável. Finalmente, escreveu-se a fórmula para o cálculo do tamanho da amostra na célula T6, conforme demonstrado na Figura VII abaixo,

FIGURA VI – POPULAÇÃO FINITA, DESVIO PADRÃO CONHECIDO

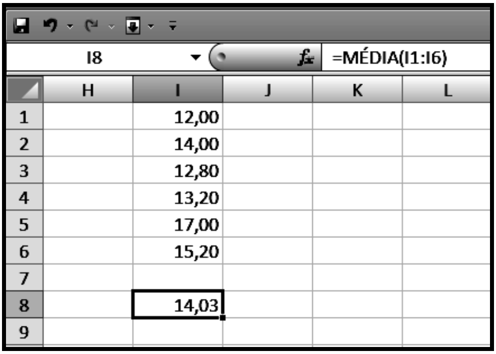
fx =((T51^2)*(T50^2)*T48)/((T52^2)*(T48-1)+(T51^2)*(T50^2))

Q	R	S	T	U	V
Tamanho da população (N):			2637		
Média amostral (xis barra):			9,4281		
Desvio padrão amostral (s):			4,1681		
Escore z para GC de 95 %:			1,959964		
Erro aceitável:			1,025		
Tamanho da amostra:			62,05088		

5.2. Executar as Funções a partir da Sintaxe das Funções

Para executar as funções diretamente na planilha, basta colocar o cursor da planilha na célula desejada e digitar a sintaxe do comando. Desejamos, por exemplo, calcular a média de 12,0, 14,0, 12,8, 13,2, 17,0 e 15,2; estes valores encontram-se digitados nas células I1 a I6, conforme Figura V abaixo. Coloca-se, então, o cursor na célula I8 e digita-se a sintaxe do comando que calcula a média: =MÉDIA(I1:I6)

FIGURA V – DIGITAÇÃO DA SINTAXE



Veja-se que ao lado de ***fx*** , canto superior direito da figura acima, aparece a sintaxe digitada e, na célula I8 o resultado do cálculo. As sintaxes paras as funções necessárias ao cálculo das estatísticas descritivas e para o cálculo do tamanho da amostra são as mesmas já apresentadas acima, em 5.2. No entanto, observa-se que as funções necessárias ao cálculo da amostra têm as seguintes sintaxes:

1) Soma: =SOMA(I1:I6)

A planilha executará a soma de todos os valores existentes entre as células I1 e I6; lembrar que as células devem estar separadas por dois pontos.

2) Média: =MÉDIA(A1:A6)

A planilha calculará a média aritmética de todos os valores existentes entre as células A1 e A6.

3) Contar o número de valores: =CONT.NÚM(B1:B197)

A planilha mostrará o número de valores existentes entre as células B1 e B197.

4) Variância: =VAR(N8:N35)

A planilha calculará a variância de todos os valores existentes entre as células N8 e N35.

5) Desvio Padrão: =DESVPAD(C1:C35)

A planilha calculará o desvio padrão dos valores existentes entre as células C1 e C35. Lembrar que o desvio padrão é igual à raiz quadrada da variância.

6) Valor de z na tabela normal padronizada: =INV.NORMP(x)

A planilha mostrará o valor de z para um determinado grau de confiança, que deverá ser digitado no lugar de x.

Exemplo: deseja-se o valor de z para um grau de confiança de 95 %. A planilha entende o grau de confiança como sendo unicaudal. Então, o grau de confiança a ser digitado deverá ser: 1 – 0,95 = 0,05 / 2 = 0,025 donde: 1,0 – 0,025 = 0,975 ou 97,5

=INV.NORMP(0,975)

O resultado apresentado será 1,959964 – confirmar na tabela.

7) Coeficiente t de Student: =INVT(x1;x2)

A planilha mostrará o valor da tabela t de Student, para um grau de confiança igual ao valor de x1 e a tantos graus de liberdade quanto x2.

No espaço x1 digitar o grau de confiança que se deseja; no espaço x2 o número de graus de liberdade. Notar que, também para esta sintaxe, os valores devem ser separados uns dos outros por ponto e vírgula.

Exemplo: =INVT(0,05;15)

Deseja-se o valor da tabela t para o grau de confiança de 5 % (0,05) e com quinze graus de liberdade. O resultado será 2,1314 – confirmar na tabela.

8) Escore z: =PADRONIZAR(x1;x2;x3)

A planilha calculará o escore z de um determinado valor, dados a média e o desvio padrão da série à qual pertence o valor determinado. No espaço x1 digitar o valor que se quer padronizar; no espaço x2 a média da série e, no espaço x3 digitar o desvio padrão da série. Notar que, para esta sintaxe, os valores devem ser separados uns dos outros por ponto e vírgula.

Exemplo: =PADRONIZAR(0,909;0,9128;0,0395)

Deseja-se o escore z de 0,909; a média da série é 0,9128 e seu desvio padrão é 0,0395. O resultado será -0,0962.

ANEXO 1 – O ESCORE z

O chamado escore z, de um valor qualquer da variável aleatória, é o número de desvios padrão que aquele valor dista da média. É o resultado do expoente da fórmula (*1) acima:

$$z = \frac{x_i - \mu}{\sigma}$$

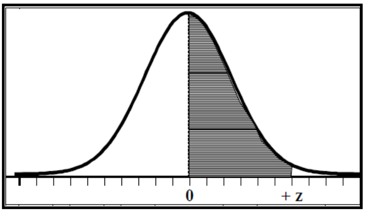
Por exemplo: as alturas de um grupo de mulheres pesquisadas tem distribuição normal com média 63,6 e desvio padrão 2,5. Calcular a quantos desvios padrão da média encontra-se a mulher cuja altura é de 68,6.

$$z = \frac{68{,}6 - 63{,}6}{2{,}5} = 2{,}00$$

Com efeito: 63,6 + 2 x 2,5 = 68,6

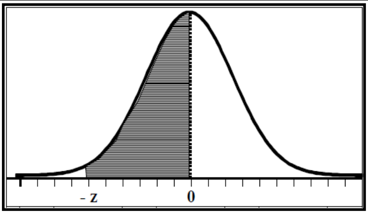
POSSIBILIDADES

Existem nove possibilidades de localização de áreas na curva normal padronizada. É preciso atenção a isto, para que se possa verificar a coerência do resultado. São as seguintes:



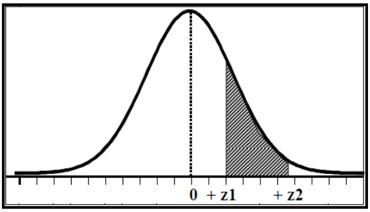
Área de 0 até + z

Leitura direta na tabela



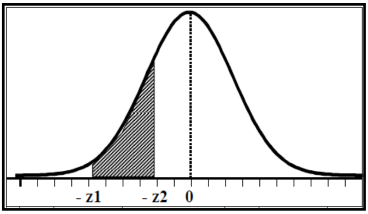
Área de 0 até - z

Leitura direta na tabela; considerar que, pelo fato da curva ser simétrica ao redor de 0, a área de 0 a - z é igual à área de 0 a + z



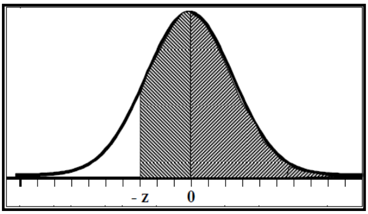
Área entre z1 e z2

Calcular a área de 0 a z2 menos a área de 0 a z1



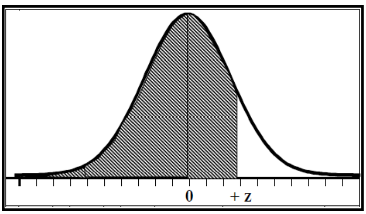
Área entre - z1 e - z2

Calcular a área de 0 a – z1 menos a área de 0 a – z2



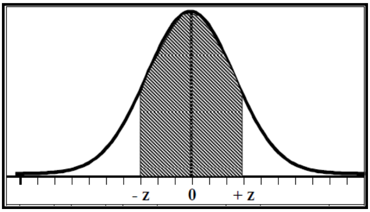
Área total à direita de - z1

Ler diretamente na tabela a área de 0 a z1 e somar 0,5



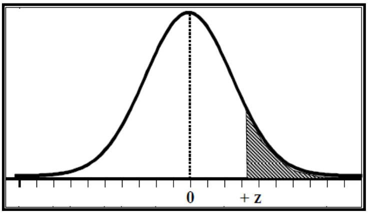
Área total à esquerda de + z1

Ler diretamente na tabela a área de 0 a z1 e somar 0,5



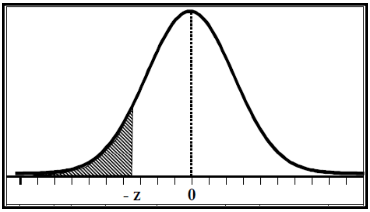
Área entre - z e + z

Somar a área de 0 a – z com a área de 0 a +z



Área à direita de + z

Ler diretamente na tabela a área de 0 a + z e diminuí-la de 0,5



Área à esquerda de - z

Ler diretamente na tabela a área de 0 a + z e diminuí-la de 0,5

ANEXO 2 – TABELA DA DISTRIBUIÇÃO NORMAL PADRONIZADA

Tabela G Áreas na Cauda Direita sob a Distribuição Normal Padronizada

Cada valor da tabela indica a proporção da área total sob a curva normal contida no segmento delimitado por uma perpendicular levantada na média e uma perpendicular levantada à distância de z desvios padrões unitários.

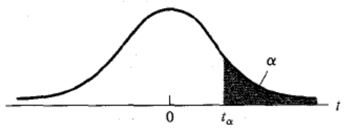
Média z

Ilustrando: 43,57% da área sob uma curva normal estão entre a ordenada máxima e um ponto 1,52 desvios padrões adiante.

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2518	0.2549
0.7	0.2580	0.2612	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4986	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990
3.1	0.4990	0.4991	0.4991	0.4991	0.4992	0.4992	0.4992	0.4992	0.4993	0.4993
3.2	0.4993	0.4993	0.4994	0.4994	0.4994	0.4994	0.4994	0.4995	0.4995	0.4995
3.3	0.4995	0.4995	0.4995	0.4996	0.4996	0.4996	0.4996	0.4996	0.4996	0.4997
3.4	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4998	0.4998
3.5	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998
3.6	0.4998	0.4998	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.7	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.8	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.5000	0.5000	0.5000
3.9	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000

ANEXO 3 – TABELA DA DISTRIBUIÇÃO t DE STUDENT

Tabela H Distribuições t



Esta tabela dá os valores de t_α que correspondem a uma área α na cauda direita (superior) específico de graus de liberdade.

Graus de liberdade	Na cauda superior área α							
	.1	.05	.025	.01	.005	.0025	.001	.0005
1	3.078	6.314	12.706	31.821	63.657	127.32	318.31	636.62
2	1.886	2.920	4.303	6.965	9.925	14.089	22.327	31.598
3	1.638	2.353	3.182	4.541	5.841	7.453	10.214	12.924
4	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.029	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	3.833	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10	1.371	1.812	2.228	2.764	3.169	3.581	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.372	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073
16	1.337	1.746	2.120	2.583	2.921	3.252	3.686	4.015
17	1.333	1.740	2.110	2.567	2.898	3.222	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.197	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.850
21	1.323	1.721	2.080	2.518	2.831	3.135	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792
23	1.319	1.714	2.069	2.500	2.807	3.104	3.485	3.767
24	1.318	1.711	2.064	2.492	2.797	3.091	3.467	3.745
25	1.316	1.708	2.060	2.485	2.787	3.078	3.450	3.725
26	1.315	1.706	2.056	2.479	2.779	3.067	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.057	3.421	3.690
28	1.313	1.701	2.048	2.467	2.763	3.047	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.038	3.396	3.659
30	1.310	1.697	2.042	2.457	2.750	3.030	3.385	3.646
40	1.303	1.684	2.021	2.423	2.704	2.971	3.307	3.551
60	1.296	1.671	2.000	2.390	2.660	2.915	3.232	3.460
120	1.289	1.658	1.980	2.358	2.617	2.860	3.160	3.373
∞	1.282	1.645	1.960	2.326	2.576	2.807	3.090	3.291

Referências Bibliográficas

TRIOLA, M. F. Introdução à Estatística. 7. ed. Rio de Janeiro: LTC, 1999

STEVENSON, W. J. Estatística Aplicada à Administração. 1. ed. São Paulo: Harbra, 2001

COSTA NETO, P. L. O. Estatística. 1. ed. 18ª reimpressão. São Paulo: Editora Edgard Blücher Ltda, 2000

NEUFELD, J. L. Estatística Aplicada à Administração usando EXCEL. 1. ed. 5ª reimpressão. São Paulo: Pearson, 2009

Tabelas:

STEVENSON, W. J. Estatística Aplicada à Administração. 1. ed. São Paulo: Harbra, 2001
Distribuição Normal Padronizada – pg. 461
Distribuição t de Student – pg. 462