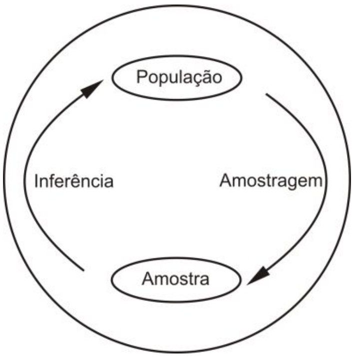
# AMOSTRAGEM

## 1)Conceitos em amostragem

Inferência estatística – é o processo de obter informação sobre uma população a partir de resultados observados na amostra.

Amostragem – é o processo de retirada dos “n” elementos amostrais, na qual deve seguir um método adequado (tipos de amostragem).




## 2) Plano de amostragem

1°) Definir claramente os objetivos da pesquisa;

2°) Definição da população;

3°) Definição da unidade amostral;

4°) Forma de seleção dos elementos da população;

5°) Tamanho da amostra.

Exemplo:

**População Alvo:** Moradores de uma cidade.

**Objetivo:** Tipo de residência: - Própria; alugada; emprestada.

- Um piso; dois pisos; três ou mais pisos.

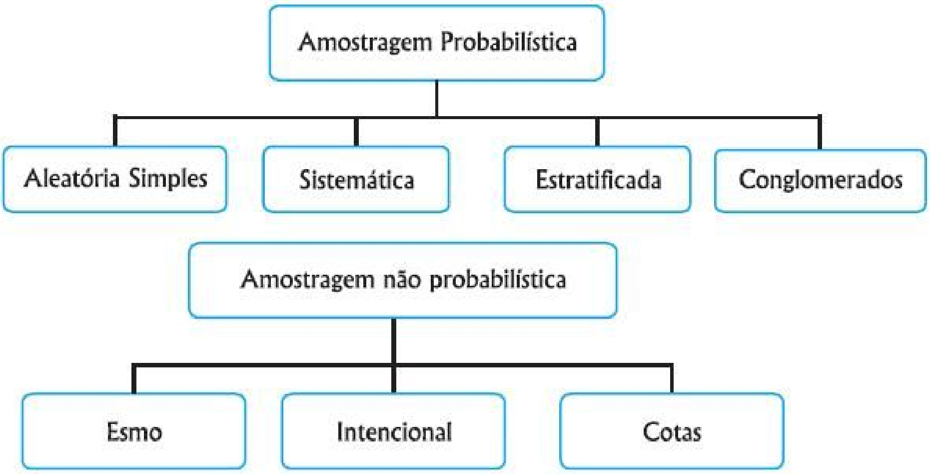
**Unidade Amostral:** Domicílios (residências).

**Elementos da população:** Família por domicílio.

**Tipo de amostragem:** por exemplo, aleatória simples, sistemática, estratificada.

## 3) Tipos de amostragem

Podemos classificar a amostragem em não-probabilística e probabilística (mais utilizada). Dentro da amostragem não probabilística temos a amostragem a esmo, intencional e cotas, para a amostragem probabilística existe a amostragem simples ou ocasional, sistemática, estratificada e por conglomerados.




Quadro resumo amostragem

Vejamos alguns tipos principais de amostragem probabilística:

a) Amostragem simples ou ocasional

É o processo mais elementar e freqüentemente utilizado. Todos os elementos da população têm igual probabilidade de serem escolhidos. Para uma população finita o processo deve ser sem reposição. Todos os elementos da população devem ser numerados. Para realizar o sorteio dos elementos devemos usar a **Tabela de Números Aleatórios**.

b) Amostragem Sistemática

Trata-se de uma variação da amostragem aleatória simples, conveniente quando a população esta naturalmente ordenada, como fichas em um fichário, lista telefônica, etc.

Sendo $N$ o tamanho da população e $n$ o tamanho da amostra desejado, define-se a quantidade $\frac{N}{n} = K$ , chamado intervalo de amostragem.

Faz-se um sorteio entre os números $1, 2, 3, ..., k$, e se obtém o valor $i$, onde será o meu primeiro elemento, os demais elementos poderão ser calculados pelo termo geral de uma progressão aritmética.

$$a_n = i + (n-1).k$$

A amostra sistemática é freqüentemente utilizada em pesquisas que obrigam que a seleção seja feita durante a etapa de coleta de dados, por pessoas que não estão familiarizadas com tabelas de números aleatórios ou com uso de software.

Exemplo de amostragem simples e sistemática.

Suponha que uma empresa de telefonia fixa deseja saber o grau de satisfação de seus usuários com serviços prestados. O número de assinantes é da ordem de 50.000 e nos desejamos selecionar uma amostra aleatória de 1.000 assinantes com o intuito de obter a avaliação sobre os serviços.

No caso a amostragem simples.

Primeiramente devemos ter os assinantes numerados seqüencialmente de 1 a 50.000 e somente após seriam selecionados os 1.000 assinantes. A seleção seria feita com o uso de uma tabela de números aleatórios ou de software que gere estes números.

No caso da amostragem sistemática

$$N = 50.000$$

$$n = 1.000$$

$$\frac{N}{n} = \frac{50.000}{1.000} = 50$$

Portanto devemos selecionar 1 assinante entre os primeiros 50. Fazendo-se um sorteio entre 1, 2, ..., 50, digamos que foi sorteado o número 10.

Aplicando na fórmula do termo geral de uma progressão aritmética, obtemos:

$$a_n = i + (n-1).k$$

$$a_1 = 10 + (1-1).50 = 10$$

$$a_2 = 10 + (2-1).50 = 60$$

$$a_3 = 10 + (3-1).50 = 110$$

E assim sucessivamente, logo temos então os números 10, 60, 110, ...

c) Amostragem estratificada

É um processo de amostragem usado quando nos deparamos com populações heterogêneas, no qual pode-se distinguir subpopulações mais ou menos homogêneas, denominados estratos.

Após a determinação dos estratos, seleciona-se uma amostra aleatória de cada uma subpopulação (estrato).

As diversas subamostras retiradas das subpopulações devem ser proporcionais aos respectivos números de elementos dos estratos e guardarem a proporcionalidade em relação a variabilidade de cada estrato, obtendo-se uma estratificação ótima.

Exemplo: Vamos obter uma amostra estratificada de 10% da população para a pesquisa da estatura de 90 alunos de uma escola sendo que destes 54 sejam meninos e 36 sejam meninas.

São, portanto dois estratos (gênero feminino e gênero masculino) e queremos uma amostra de 10% da população.

| Gênero | População | 10% | Amostra |
|---|---|---|---|
| M | 54 | $\frac{10x54}{100}=5{,}4$ | 5 |
| F | 36 | $\frac{10x36}{100}=3{,}6$ | 4 |
| Total | 90 | $\frac{10x90}{100}=9{,}0$ | 9 |

Numeramos os alunos de 1 a 90, sendo que de 01 a 54 correspondem meninos e de 55 a 90, meninas.

Tomando na tabela de números aleatórios a primeira e a segunda coluna da esquerda, de cima para baixo, obtemos os seguintes números:

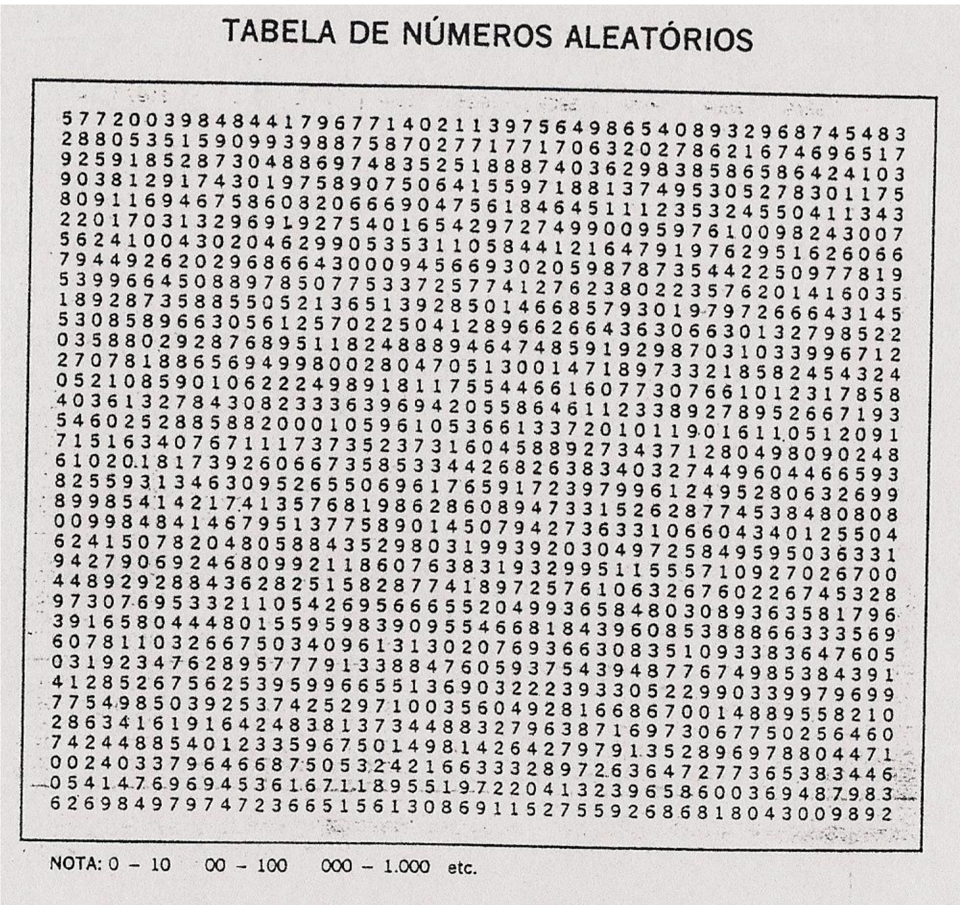


TABELA DE NÚMEROS ALEATÓRIOS

57720039848441796771402113975649865408932968745483
28805351590993988758702771771706320278621674696517
92591852873048869748352518887403629838586586424103
90381291743019758907506415597188137495305278301175
80911694675860820666904756184645111235324550411343
22017031329691927540165429727499009597610098243007
56241004302046299053531105844121647919762951626066
79449262029686643000945669302059878735442250977819
53996645088978507753372577412762380223576201416035
18928735885505213651392850146685793019797266643145
53085896630561257022504128966266436306630132798522
03588029287689511824888946474859192987031033996712
27078188656949980028047051300147189733218582454324
05210859010622249891811755446616077307661012317858
40361327843082333639694205586461123389278952667193
54602528858820001059610536613372010119016110512091
71516340767111737352373160458892734371280498090248
61020181739260667358533442682638340327449604466593
82559313463095265506961765917239799612495280632699
89985414217413576819862860894733152628774538480808
00998484146795137758901450794273633106604340125504
62415078204805884352980319939203049725849595036331
94279069246809921186076383193299511555710927026700
44892928843628251582877418972576106326760226745328
97307695332110542695666552049936584803089363581796
39165804448015595983909554668184396085388866333569
60781103266750340961313020769366308351093383647605
03192347628957779133884760593754394877674985384391
41285267562539599665513690322239330522990339979699
77549850392537425297100356049281668670014889558210
28634161916424838137344883279638716973067750256460
74244885401233596750149814264279791352896978804471
00240337964668750532421663332897263647277365383446
05414769694536167118955197220413239658600369487983
62698497974723665156130869115275592686818043009892

NOTA: 0 – 10    00 – 100    000 – 1.000 etc.

57 28 ~~92~~ 90 80 22 56 79 53 18 ~~53~~ 03 27 05 40

Temos então:

28 22 53 18 03 – para os meninos;

57 90 80 56 – para as meninas;

A amostragem pode ser com ou sem reposição dos elementos.

Consideremos $N$ o número de elementos de uma população, e seja $n$ o número de elementos de uma amostra, então:

Se o processo de retirada dos elementos for **com** reposição o número de amostras possíveis será:

$$\text{Número de amostras} = N^n$$

Se o processo de retirada dos elementos for **sem** reposição o número de amostras possíveis será:

$$\text{Número de amostras} = C_N^n = \frac{N!}{n!(N-n)!}$$

## 4) Dimensionamento da amostra

Procedimentos:

1º) Analise o questionário, ou roteiro da entrevista e escolha uma variável que julgue mais importante para o estudo. Se possível, escolha mais do que uma variável.

2º) Verifique o nível de mensuração da variável: se nominal, ordinal ou intervalar.

3º) Considere o tamanho da população: finita ou infinita.

4º) Se a variável escolhida for intervalar e a população considerada infinita, você poderá determinar o tamanho da amostra pela fórmula:

$$n = \left(\frac{Z.\sigma}{d}\right)^2$$

onde: $Z$ = abscissa da distribuição normal padrão, fixado um nível de $(1-\alpha)\%$ de confiança para construção do intervalo de confiança para a média.

Se o nível for de 90%, $Z = 1{,}645$.

Se o nível for de 95,5%, $Z = 2$.

Se o nível for de 95%, $Z = 1{,}96$.

Se o nível for de 99%, $Z = 2{,}57$.

Geralmente, utiliza-se $Z = 2$, isto é, admite-se $(1-\alpha)\% = 95{,}5\%$.

$\sigma$ = desvio padrão da população, expressão na unidade variável. Você poderá avaliá-lo de, pelo menos, uma das três maneiras:

- Especificações técnicas;
- Resgatar o valor de estudos semelhantes;
- Fazer conjecturas sobre possíveis valores.

**Obs.:** Se o desvio padrão não for conhecido devemos utilizar um valor preliminar obtido por processos como fazer uma aproximação $\sigma \cong \dfrac{amplitude}{4}$ ou realizar um estudo piloto, iniciando o processo de amostragem. Com base na primeira coleção de pelo menos 31 valores amostrais selecionados aleatoriamente, calcular o desvio-padrão amostral e utilizá-lo em lugar do desvio padrão populacional. Este valor pode ser refinado com a obtenção de mais dados amostrais.

$d$ = erro amostral, expresso na unidade da variável. O erro amostral é a máxima diferença que o investigador admite suportar entre $\mu$ e $\overline{X}$, isto é: $\left|\mu - \overline{X}\right| < d$, onde $\mu$ é a verdadeira média populacional, que ele não conhece, e $\overline{X}$ será a média amostral a ser calculada a partir da amostra.

5º) Se a variável escolhida for intervalar e a população finita, tem-se:

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

onde: $Z$ = abscissa da distribuição normal padrão (veja comentaria do item 4º);

$\sigma$ = desvio padrão da população (veja comentaria do item 4º);

$N$ = tamanho da população;

$d$ = erro amostral (veja comentaria do item 4º).

6º) Se a variável escolhida for nominal ou ordinal, e a população considerada infinita, você poderá determinar o tamanho da amostra pela fórmula:

$$n = \frac{Z^2.\hat{p}.\hat{q}}{d^2}$$

onde: $Z$ = abscissa da distribuição normal padrão (veja comentaria do item 4º);

$\hat{p}$ = estimativa da verdadeira proporção de um dos níveis da variável escolhida. Por exemplo, se a variável escolhida for porte da empresa, $\hat{p}$ podará ser a estimativa da verdadeira proporção de grandes empresas do setor que está sendo estudado. Será expresso em decimais. Assim, se $\hat{p} = 30\%$, teremos: $\hat{p} = 0{,}30$.

$\hat{q} = 1 - \hat{p}$

**Obs.:** Se $\hat{p}$ e $\hat{q}$ forem desconhecidos, substituímos $\hat{p}$ e $\hat{q}$ por 0,5, obtendo a seguinte estimativa: $n = \dfrac{Z^2.0{,}25}{d^2}$

$d$ = erro amostral, expresso em decimais. O erro amostral neste caso será a máxima diferença que o investigador admite suportar entre $p$ e $\hat{p}$, isto é: $|p - \hat{p}| < d$, em que $p$ é a verdadeira proporção, que ele não conhece, e $\hat{p}$ será a proporção (freqüência relativa) do evento a ser calculado a partir da amostra.

7º) Se a variável escolhida for nominal ou ordinal e a população finita, tem-se:

$$n = \frac{Z^2.\hat{p}.\hat{q}.N}{d^2(N-1) + Z^2.\hat{p}.\hat{q}}$$

onde: $Z$ = abscissa da distribuição normal padrão (veja comentaria do item 4º);

$\hat{p}$ = estimativa da proporção (veja comentaria do item 6º);

$\hat{q} = 1 - \hat{p}$ (veja comentaria do item 6º);

$d$ = erro amostral (veja comentaria do item 6º);

$N$ = tamanho da amostra.

Estas fórmulas são básicas para qualquer tipo de composição da amostra; todavia, existem fórmulas específicas segundo o critério de composição da amostra.

Se o investigador escolhe mais de uma variável, deve optar pelo maior "n" obtido.

Exemplos:

a)Suponha que a variável escolhida em estudo seja o peso de certa peça e que a população seja infinita. Pelas especificações do produto, o desvio padrão é de 10kg. Logo, admitindo um nível de confiança de 95,5% e m erro amostral de 1,5kg, temos:

$\sigma = 10kg$ $\quad$ $d = 1{,}5kg$ $\quad$ $(1-\alpha)\% = 95{,}5\%$, ou seja: $Z = 2$

$$n = \left(\frac{Z.\sigma}{d}\right)^2 = \left(\frac{2.10}{1{,}5}\right)^2 = 177{,}77 \cong 178$$

Ou seja: com uma amostra aleatória simples de 178 peças, há um erro máximo de 1,5kg para construir um intervalo de confiança para o peso médio, com nível de confiança de 95,5%.

b) Admitimos os mesmos dados do exemplo anterior e uma população finita de 600 peças.

Logo:

$\sigma = 10kg$ $d = 1,5kg$ $(1-\alpha)\% = 95,5\%$, ou seja: $Z = 2$ $N = 600$

$$n = \frac{Z^2.\sigma^2.N}{d^2(N-1)+Z^2\sigma^2} = \frac{2^2.10^2.600}{1,5^2(600-1)+2^2.10^2} = 137,1 \cong 138$$

c) Suponha que a variável escolhida em um estudo seja a proporção de eleitores favoráveis ao candidato X e que o investigador tenha elementos para suspeitar que essa porcentagem seja de 30%. Admitir a população infinita, um nível de confiança de 99% e um erro amostral de 2% (ou seja, que a diferença entre a verdadeira proporção de eleitores do candidato X e a estimativa a ser calculada na amostra seja no máximo de 2%).

Assim:

$(1-\alpha)\% = 99\%$, ou seja: $Z = 2,57$ $\hat{p} = 30\% = 0,30$ $\hat{q} = 1-0,30 = 0,70$

$d = 2\% = 0,02$

$$n = \frac{Z^2.\hat{p}.\hat{q}}{d^2} = \frac{(2,57)^2.(0,30)(0,70)}{(0,02)^2} = 3.467,57 \cong 3.468$$

Ou seja, consultando, aleatoriamente, 3.468 eleitores, poderemos fazer inferência da verdadeira proporção de eleitores do candidato X, com erro máximo de 2%.

d) Admitir os mesmo dados do exemplo anterior e que a população de eleitores seja finita de 20.0000 eleitores. Logo:

$$n = \frac{Z^2.\hat{p}.\hat{q}.N}{d^2(N-1)+Z^2.\hat{p}.\hat{q}} = \frac{(2,57)^2.(0,30)(0,70)(20.000)}{(0,02)^2(20.000-1)+(2,57)^2(0,30)(0,70)} = 2.955,33 \cong 2.956$$

**Obs.:** Como determinar o valor de Z:

- Se $Z = 90\%$ então $Z_{\frac{\alpha}{2}} = Z_{\frac{90\%}{2}} = 45\% = 0,450$, ir na tabela da distribuição normal e ver em qual linha e coluna está o valor 0,450, localizado vemos que o valor na linha é 1,6 e na coluna é 0,04, o que corresponde a 1,6+0,04=1,64.

 Se $Z = 98\%$ então $Z_{\frac{\alpha}{2}} = Z_{\frac{98\%}{2}} = 49\% = 0{,}490$, ir na tabela da distribuição normal e ver em qual linha e coluna está o valor 0,490, localizado vemos que o valor na linha é 2,3 e na coluna é 0,03, o que corresponde a 2,3+0,03=2,33.

**Exercícios**

1) Um estatístico deseja estimar a renda média para o primeiro ano de trabalho de um tecnólogo em Grãos. Quantos valores de renda devem ser tomados, se o estatístico deseja ter 95% de confiança em que a média amostral esteja a menos de R$ 500,00 da verdadeira média populacional? Suponha que saibamos, por um estudo prévio, que para tais rendas, $\sigma$ = R$ 6250,00.

2) Baseado nos dados do exercícios 1, utilize uma margem de erro maior, como R$1.000,00 e determine qual seria o tamanho da amostra necessária nesta situação.

3) Uma pesquisa é planejada para determinar as despesas médicas anuais das famílias dos empregados de uma grande empresa. A gerência da empresa deseja ter 95% de confiança de que a média da amostra está no máximo com uma margem de erro de  $50 da média real das despesas médicas familiares. Um estudo-piloto indica que o desvio-padrão pode ser calculado como sendo igual a $400.

a) Qual o tamanho de amostra necessário?

b) Se a gerência deseja estar certa em uma margem de erro de  $25, que tamanho de amostra será necessário?

4) O teste de QI padrão é planejado de modo que a média seja 100 e o desvio padrão para adultos normais seja 15. Ache o tamanho da amostra necessária para estimar o QI médio dos instrutores de estatística. Queremos ter 99% de confiança em que nossa média amostral esteja a menos de 1,5 pontos de QI da verdadeira média. A média para esta população é obviamente superior a 100, e o desvio-padrão é provavelmente inferior a 15, porque se trata de um grupo com menor variação do que um grupo selecionado aleatoriamente da população geral; portanto, se tomamos $\sigma = 15$, estaremos sendo conservadores, por utilizarmos um valor que dará um tamanho de amostra no mínimo tão grande quanto necessário. Suponha $\sigma = 15$ e determine o tamanho da amostra necessário.

5) Uma assistente social deseja saber o tamanho da amostra (n) necessário para determinar a proporção da população atendida por uma Unidade de Saúde, que pertence ao município de Alegrete. Não foi feito um levantamento prévio da proporção amostral e, portanto, seu valor é desconhecido. Ela quer ter 90% de confiança que sua o erro máximo de estimativa (d) seja de 5% (ou 0,05). Quantas pessoas necessitam ser entrevistadas?

6) Baseado nos dados do exercício 5, utilize uma margem de erro maior, como 0,20 (20%) e determine qual seria o tamanho da amostra necessário quando o nível de confiança é 90% e quando é 95%.

7) Dada a seguinte população: (rendas em R$1.000,00)

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 29 | 06 | 34 | 12 | 15 | 31 | 34 | 20 | 08 | 30 |
| 08 | 15 | 24 | 22 | 35 | 31 | 25 | 26 | 20 | 10 |
| 30 | 04 | 16 | 21 | 14 | 21 | 16 | 18 | 20 | 12 |
| 31 | 20 | 12 | 18 | 12 | 25 | 26 | 13 | 10 | 05 |
| 13 | 19 | 30 | 17 | 25 | 29 | 25 | 28 | 32 | 15 |
| 10 | 21 | 18 | 07 | 16 | 14 | 11 | 22 | 21 | 36 |
| 32 | 17 | 15 | 13 | 08 | 12 | 23 | 25 | 13 | 21 |
| 05 | 12 | 32 | 21 | 10 | 30 | 30 | 10 | 14 | 17 |
| 34 | 22 | 30 | 48 | 19 | 12 | 08 | 07 | 15 | 20 |
| 26 | 25 | 22 | 30 | 33 | 14 | 17 | 13 | 10 | 09 |

a) Calcular o tamanho da amostra para estimar a média, sendo $d = R\$2.000,00$, $\sigma = R\$7.000,00$ e $(1-\alpha)\% = 95,5\%$.

b) Retirar uma amostra aleatória simples, considerando o tamanho amostral obtido no item anterior.

c) Calcular sua média amostral.

d) Calcular o desvio padrão amostral.

e) Calcular a média da população e verifique se $\left|\mu - \overline{X}\right| \leq d$.

8) Sendo $\hat{p} = \hat{q} = 0,5$, população infinita, $d = 0,05$ e $(1-\alpha)\% = 95,5\%$, determinar o tamanho amostral.

9) Sendo $\hat{p} = \hat{q} = 0,5$, população de 200.000, $d = 0,05$ e $(1-\alpha)\% = 95,5\%$, determinar o tamanho amostral. Comparar com o resultados obtido no exercício 8.

10) Qual é o tamanho da amostra que o Departamento de Trânsito de uma grande cidade deve tomar para estimar a porcentagem de semáforos defeituosos, se o objetivo é ter 95,5% de confiança em não errar mais de 3%.

11) Estudos anteriores mostram que o desvio padrão da altura dos homens que cursam o Instituto Federal Farroupilha é de 10cm. Querendo estimar a atura média de todos os homens dessa universidade, com tolerância de 3cm e probabilidade de 0,955, quantas observações deverão ser utilizadas?

12) Qual é o tamanho necessário da amostragem que um alfaiate deve usar se deseja estimar o tempo médio que os fregueses levam ao trocador? Ele acha que o desvio padrão é de 3 minutos, com base em amostras anteriores, e deseja estimar a média a menos de 1 minuto, usando um nível d 95,5%.

13) Em Alegrete, há 10.000 árvores. Qual deve ser o tamanho da amostra que o Departamento de Jardins precisa tomar para estimar a porcentagem de plantas que merecem podas, se o objetivo é ter 99% de confiança de não errar mais de 3%?

14) Segundo dados de uma pesquisa anterior, 40% dos alunos de certa escola são “gremistas”. Admita um erro de 2,5%, $(1-\alpha)\% = 95,5\%$, para dimensionar tamanho de amostra de tricolores. Sabe-se que a escola tem 5.000 alunos matriculados.

15) Um engenheiro encarregado do controle de qualidade deseja estimar a fração de artigos defeituosos de um grande lote de lâmpadas. Com base em sua experiência, ele sabe que a fração efetiva de lâmpadas defeituosas deve estar próxima de 0,2. Que tamanho deve ter uma amostra, se ele desejar estimar a verdadeira fração de lâmpadas defeituosas com tolerância de 0,01, usando um nível de confiança de 98%

**Gabarito**

1) Um estatístico deseja estimar a renda média para o primeiro ano de trabalho de um tecnólogo em Grãos. Quantos valores de renda devem ser tomados, se o estatístico deseja ter 95% de confiança em que a média amostral esteja a menos de R$ 500,00 da verdadeira média populacional? Suponha que saibamos, por um estudo prévio, que para tais rendas, $\sigma$ = R$ 6250,00.

Queremos determinar o tamanho n da amostra, dado que $\alpha = 0{,}05$ (95% de confiança). Desejamos que a média amostral seja a menos de R$ 500 da média populacional, de forma que $d = 500$. Supondo $\sigma = R\$6250{,}00$, aplicamos a equação abaixo, obtemos:

$$n = \left(\frac{Z.\sigma}{d}\right)^2 = \left(\frac{1{,}96.6250}{500}\right)^2 = 600{,}25 \cong 601$$

Devemos, portanto, obter uma amostra de ao menos 601 rendas de primeiro ano, selecionadas aleatoriamente, de tecnólogos da faculdade que tenham feito um curso Tecnólogo em Grãos. Com tal amostra teremos 95% de confiança em que a média amostral difira em menos de R$500,00 da verdadeira média populacional.

2) Baseado nos dados do exercícios 1, utilize uma margem de erro maior, como R$1.000,00 e determine qual seria o tamanho da amostra necessária nesta situação.

Dados do problema:

$\alpha = 0{,}05$ (95% de confiança) $Z = 1{,}96$

$d = 1000$.

$\sigma = R\$6250{,}00$, aplicando a equação abaixo, obtemos:

$$n = \left(\frac{Z.\sigma}{d}\right)^2 = \left(\frac{1{,}96.6250}{1000}\right)^2 = 150{,}063 \cong 151$$

Devemos, portanto, obter uma amostra de ao menos 151 rendas de primeiro ano, selecionadas aleatoriamente, de tecnólogos da faculdade que tenham feito um curso Tecnólogo em Grãos. Com tal amostra teremos 95% de confiança em que a média amostral difira em menos de R$1.000,00 da verdadeira média populacional.

3) Uma pesquisa é planejada para determinar as despesas médicas anuais das famílias dos empregados de uma grande empresa. A gerência da empresa deseja ter 95% de confiança de que a média da amostra está no máximo com uma margem de erro de $50 da média real das despesas médicas familiares. Um estudo-piloto indica que o desvio-padrão pode ser calculado como sendo igual a $400.

a) Qual o tamanho de amostra necessário?

b) Se a gerência deseja estar certa em uma margem de erro de $25, que tamanho de amostra será necessário?

a) Dados do problema:

$\alpha = 0{,}05$ (95% de confiança) $Z = 1{,}96$

$d = 50$.

$\sigma = 400$ , aplicando a equação abaixo, obtemos:

$$n = \left( \frac{Z.\sigma}{d} \right)^2 = \left( \frac{1{,}96.400}{50} \right)^2 = 245{,}862 \cong 246$$

Devemos, portanto, obter uma amostra de ao menos 246 despesas médicas, selecionadas aleatoriamente, de famílias dos empregados. Com tal amostra teremos 95% de confiança em que a média amostral difira em menos de $50 da verdadeira média populacional.

b) Dado do problema:

$\alpha = 0{,}05$ (95% de confiança) $Z = 1{,}96$

$d = 25$.

$\sigma = 400$ , aplicando a equação abaixo, obtemos:

$$n = \left( \frac{Z.\sigma}{d} \right)^2 = \left( \frac{1{,}96.400}{25} \right)^2 = 983{,}449 \cong 984$$

Devemos, portanto, obter uma amostra de ao menos 984 despesas médicas, selecionadas aleatoriamente, de famílias dos empregados. Com tal amostra teremos 95% de confiança em que a média amostral difira em menos de $25 da verdadeira média populacional.

4) O teste de QI padrão é planejado de modo que a média seja 100 e o desvio padrão para adultos normais seja 15. Ache o tamanho da amostra necessária para estimar o QI médio dos instrutores de estatística. Queremos ter 99% de confiança em que nossa média amostral esteja a menos de 1,5 pontos de QI da verdadeira média. A média para esta população é obviamente superior a 100, e o desvio-padrão é provavelmente inferior a 15, porque se trata de um grupo com menor variação do que um grupo selecionado aleatoriamente da população geral; portanto, se tomamos $\sigma = 15$, estaremos sendo conservadores, por utilizarmos um valor que dará um tamanho de amostra no mínimo tão grande quanto necessário. Suponha $\sigma = 15$ e determine o tamanho da amostra necessário.

Dado do problema:

$\alpha = 0{,}01$ (99% de confiança) $Z = 2{,}57$

$d = 1{,}5$.

$\sigma = 15$, aplicando a equação abaixo, obtemos:

$$n = \left(\frac{Z.\sigma}{d}\right)^2 = \left(\frac{2{,}57.15}{1{,}5}\right)^2 = 660{,}49 \cong 661$$

Devemos, portanto, obter uma amostra de ao menos 661 instrutores de estatística, selecionadas aleatoriamente. Com tal amostra teremos 99% de confiança em que a média amostral difira em menos de 1,5 da verdadeira média populacional.

5) Uma assistente social deseja saber o tamanho da amostra (n) necessário para determinar a proporção da população atendida por uma Unidade de Saúde, que pertence ao município de Alegrete. Não foi feito um levantamento prévio da proporção amostral e, portanto, seu valor é desconhecido. Ela quer ter 90% de confiança que sua o erro máximo de estimativa (d) seja de 5% (ou 0,05). Quantas pessoas necessitam ser entrevistadas?

Dados do problema:

$\alpha = 0{,}1$ (90% de confiança) $Z = 1{,}645$

$d = 5\% = 0{,}05$, aplicando a equação abaixo, obtemos:

$$n = \frac{Z^2.\hat{p}.\hat{q}}{d^2} = \frac{(1{,}645)^2.0{,}25}{(0{,}05)^2} = 270{,}602 \cong 271$$

Devemos, portanto, obter uma amostra de 271 pessoas para determinar a proporção da população atendida na Unidade de Saúde, que se origina do município de Alegrete.

6) Baseado nos dados do exercício 5, utilize uma margem de erro maior, como 0,20 (20%) e determine qual seria o tamanho da amostra necessário quando o nível de confiança é 90% e quando é 95%.

Dados do problema para 90% de nível de confiança:

$\alpha = 0{,}1$ (90% de confiança) $Z = 1{,}645$

$d = 20\% = 0{,}20$, aplicando a equação abaixo, obtemos:

$$n = \frac{Z^2.\hat{p}.\hat{q}}{d^2} = \frac{(1{,}645)^2.0{,}25}{(0{,}20)^2} = 16{,}913 \cong 17$$

Devemos, portanto, obter uma amostra de 17 pessoas para determinar a proporção da população atendida na Unidade de Saúde, que se origina do município de Alegrete.

Dados do problema para 90% de nível de confiança:

$\alpha = 0{,}05$ (95% de confiança) $Z = 1{,}96$

$d = 20\% = 0{,}20$, aplicando a equação abaixo, obtemos:

$$n = \frac{Z^2.\hat{p}.\hat{q}}{d^2} = \frac{(1{,}96)^2.0{,}25}{(0{,}20)^2} = 24{,}01 \cong 25$$

Devemos, portanto, obter uma amostra de 25 pessoas para determinar a proporção da população atendida na Unidade de Saúde, que se origina do município de Alegrete.

7) Dada a seguinte população: (rendas em R$1.000,00)

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **29** | 06 | **34** | 12 | **15** | 31 | 34 | 20 | 08 | 30 |
| 08 | 15 | 24 | 22 | 35 | 31 | 25 | **26** | 20 | 10 |
| 30 | **04** | 16 | 21 | 14 | 21 | **16** | **18** | 20 | 12 |
| 31 | 20 | 12 | 18 | 12 | 25 | 26 | **13** | **10** | **05** |
| **13** | 19 | 30 | **17** | 25 | 29 | 25 | 28 | 32 | 15 |
| 10 | 21 | **18** | **07** | 16 | **14** | **11** | 22 | **21** | **36** |
| **32** | **17** | 15 | 13 | 08 | 12 | 23 | 25 | 13 | 21 |
| **05** | **12** | 32 | **21** | 10 | 30 | **30** | 10 | **14** | **17** |
| 34 | **22** | 30 | 48 | 19 | 12 | 08 | 07 | **15** | **20** |
| **26** | **25** | 22 | **30** | 33 | 14 | **17** | 13 | 10 | **09** |

a) Calcular o tamanho da amostra para estimar a média, sendo $d = R\$2.000{,}00$, $\sigma = R\$7.000{,}00$ e $(1-\alpha)\% = 95{,}5\%$.

Dados do problema:

$(1-\alpha)\% = 95{,}5\% \quad Z = 2$

$d = R\$2.000{,}00$

$\sigma = R\$7.000{,}00$ aplicando a equação abaixo, obtemos:

$$n = \frac{Z^2.\sigma^2.N}{d^2(N-1)+Z^2\sigma^2} = \frac{(2^2).(7000^2).100}{(2000^2).(100-1)+(2^2).(7000^2)} = 33{,}1081 \cong 34$$

b) Retirar uma amostra aleatória simples, considerando o tamanho amostral obtido no item anterior.

Utilizando uma tabela de números aleatórios obtemos os seguintes números para uma seleção de 34 números:

57, 28, 92, 90, 80, 22, 56, 73, 53, 18, 53, 03, 27, 05, 40, 54, 71, 61, 82, 89, 00 62, 94, 44, 97, 39, 60, 03, 41, 77, 28, 74, 00, 05, 62, 72, 80, 59, 38, 91, 01, 24, 44, 99, 92, 08, 58.

Selecionamos os primeiro 34 números

29, 34, 15, 26, 04, 16 18, 13, 10, 05, 13, 17, 18, 07, 14, 11, 21 36, 32, 17, 05, 12, 21, 30, 14, 17, 22, 15, 20, 26, 25, 30, 17, 09.

c) Calcular sua média amostral.

$$\overline{X} = \frac{\sum x_i}{n} = \frac{29+34+15+...+30+17+09}{34} = \frac{619}{34} = 18{,}206$$

d) Calcular o desvio padrão amostral.

$$S^2 = \frac{1}{n-1}\left[\sum x_i^2 - \frac{\left(\sum x_i\right)^2}{n}\right] = \frac{1}{34-1}\left(13661 - \frac{(619)^2}{34}\right) = \frac{1}{33}(13661 - 11269{,}44) = 72{,}471$$

$$S = \sqrt{S^2} = \sqrt{72{,}471} = 8{,}53$$

e) Calcular a média da população e verifique se $\left|\mu - \overline{X}\right| \le d$ .

$$\mu = \frac{\sum x_i}{N} = \frac{29+06+34+...+13+10+09}{100} = \frac{1962}{100} = 19{,}62$$

$$\left|\mu - \overline{X}\right| \le d \Rightarrow |19{,}62 - 18{,}206| = 1{,}414 \le 2000$$

8) Sendo $\hat{p} = \hat{q} = 0,5$, população infinita, $d = 0,05$ e $(1-\alpha)\% = 95,5\%$, determinar o tamanho amostral.

Dados do problema:

$(1-\alpha)\% = 95,5\%$ e $Z = 2$

$\hat{p} = \hat{q} = 0,5$

$d = 0,05$, aplicando a equação abaixo, obtemos:

$$n = \frac{Z^2.\hat{p}.\hat{q}}{d^2} = \frac{(2)^2.0,25}{(0,05)^2} = 400$$

9) Sendo $\hat{p} = \hat{q} = 0,5$, população de 200.000, $d = 0,05$ e $(1-\alpha)\% = 95,5\%$, determinar o tamanho amostral. Comparar com o resultados obtido no exercício 8.

Dados do problema:

$(1-\alpha)\% = 95,5\%$ e $Z = 2$

$N = 200.000$

$\hat{p} = \hat{q} = 0,5$

$d = 0,05$, aplicando a equação abaixo, obtemos:

$$n = \frac{Z^2.\hat{p}.\hat{q}.N}{d^2(N-1) + Z^2.\hat{p}.\hat{q}} = \frac{(2^2).(0,25).(200000)}{(0,05^2).(200000-1) + (2^2).(0,25)} = \frac{200000}{500,9975} = 399,204 \cong 400$$

Comparando os resultados, verifica-se que o cálculo do tamanho amostral para uma população de 200.000 dá, aproximadamente, o mesmo resultado, se considerarmos a população infinita.

10) Qual é o tamanho da amostra que o Departamento de Trânsito de uma grande cidade deve tomar para estimar a porcentagem de semáforos defeituosos, se o objetivo é ter 95,5% de confiança em não errar mais de 3%.

Dados do problema:

$(1-\alpha)\% = 95,5\%$ e $Z = 2$

$\hat{p} = \hat{q} = 0,5$

$d = 3\% = 0,03$, aplicando a equação abaixo, obtemos:

$$n = \frac{Z^2.\hat{p}.\hat{q}}{d^2} = \frac{(2)^2.0,25}{(0,03)^2} = \frac{1}{0,0009} = 1111,111 \cong 1112$$

11) Estudos anteriores mostram que o desvio padrão da altura dos homens que cursam o Instituto Federal Farroupilha é de 10cm. Querendo estimar a atura média de todos os homens dessa universidade, com tolerância de 3cm e probabilidade de 0,955, quantas observações deverão ser utilizadas?

Dados do problema:

$(1-\alpha)\% = 95{,}5\%$ e $Z = 2$

$\sigma = 10cm$

$d = 3$, aplicando a equação abaixo, obtemos:

$$n = \left(\frac{Z.\sigma}{d}\right)^2 = \left(\frac{2.10}{3}\right)^2 = 44{,}4 \cong 45$$

12) Qual é o tamanho necessário da amostragem que um alfaiate deve usar se deseja estimar o tempo médio que os fregueses levam ao trocador? Ele acha que o desvio padrão é de 3 minutos, com base em amostras anteriores, e deseja estimar a média a menos de 1 minuto, usando um nível d 95,5%.

Dados do problema:

$(1-\alpha)\% = 95{,}5\%$ e $Z = 2$

$\sigma = 3cm$

$d = 1$, aplicando a equação abaixo, obtemos:

$$n = \left(\frac{Z.\sigma}{d}\right)^2 = \left(\frac{2.3}{1}\right)^2 = 36$$

13) Em Alegrete, há 10.000 árvores. Qual deve ser o tamanho da amostra que o Departamento de Jardins precisa tomar para estimar a porcentagem de plantas que merecem podas, se o objetivo é ter 99% de confiança de não errar mais de 3%?

Dados do problema:

$N = 10.000$

$(1-\alpha)\% = 99\%$ e $Z = 2{,}57$

$\hat{p} = \hat{q} = 0{,}50$

$d = 3\% = 0{,}03$, aplicando a equação abaixo, obtemos:

$$n = \frac{Z^2.\hat{p}.\hat{q}.N}{d^2(N-1) + Z^2.\hat{p}.\hat{q}} = \frac{(2{,}57^2).(0{,}25).(10000)}{(0{,}03^2).(10000-1) + (2{,}57^2).(0{,}25)}$$

$$n = \frac{16512{,}25}{10{,}650325} = 1550{,}398 \cong 1551$$

14) Segundo dados de uma pesquisa anterior, 40% dos alunos de certa escola são "gremistas". Admita um erro de 2,5%, $(1-\alpha)\% = 95,5\%$, para dimensionar tamanho de amostra de tricolores. Sabe-se que a escola tem 5.000 alunos matriculados.

Dados do problema:

$N = 5.000$

$(1-\alpha)\% = 95,5\%$ e $Z = 2$

$\hat{p} = 40\% = 0,40$

$\hat{q} = (1-0,4) = 0,60$

$d = 2,5\% = 0,025$, aplicando a equação abaixo, obtemos:

$$n = \frac{Z^2.\hat{p}.\hat{q}.N}{d^2(N-1)+Z^2.\hat{p}.\hat{q}} = \frac{(2^2).(0,24).(5000)}{(0,025^2).(5000-1)+(2^2).(0,24)}$$

$$n = \frac{4800}{4,084375} = 1175,21 \cong 1176$$

15) Um engenheiro encarregado do controle de qualidade deseja estimar a fração de artigos defeituosos de um grande lote de lâmpadas. Com base em sua experiência, ele sabe que a fração efetiva de lâmpadas defeituosas deve estar próxima de 0,2. Que tamanho deve ter uma amostra, se ele desejar estimar a verdadeira fração de lâmpadas defeituosas com tolerância de 0,01, usando um nível de confiança de 98%

Dados do problema:

$(1-\alpha)\% = 98\%$ e $Z = 2,33$

$\hat{p} = 0,2$

$\hat{q} = (1-0,2) = 0,8$

$d = 0,01$, aplicando a equação abaixo, obtemos:

$$n = \frac{Z^2.\hat{p}.\hat{q}}{d^2} = \frac{(2,33^2).(0,16)}{(0,01^2)} = \frac{0,868624}{(0,01^2)} = 8686,24 \cong 8687$$