

1. Introdução;
2. Principais tabelas e problemas típicos;
3. Conclusões e sumário.

Edmundo Eboli Bonini\*

## PRINCIPAIS TABELAS ESTATÍSTICAS APLICADAS À CONTABILIDADE E À AUDITORIA

### 1. INTRODUÇÃO

O Comitê de Amostragem Estatística do American Institute of Certified Public Accountants — AICPA — publicou, em fevereiro de 1962, um relatório especial intitulado *A Amostragem estatística e o auditor independente*. Esse relatório tratava da natureza geral da amostragem estatística e de sua aplicabilidade à auditoria, concluindo com o seguinte parágrafo:

“São recomendáveis um maior conhecimento de amostragem estatística e pesquisa adicional sobre sua aplicabilidade, por parte da profissão. Seguindo esta conclusão, o comitê deu atenção, posteriormente, à relação entre amostragem estatística e os padrões de auditoria geralmente aceitos, acreditando que a publicação de seus pontos de vista sobre o assunto possa servir a uma finalidade bastante útil.”<sup>1</sup>

As leis da estatística, como todas as leis científicas, nasceram da observação de um modo de comportamento uniforme, de um conjunto de dados estatísticos. Os profissionais desta ciência têm a preocupação de observar o maior número possível de dados que parece repetir-se na futura lei e, só depois disso, considerar a lei como exata, sujeita, no entanto, a alterações quando negar a homogeneidade que se supõe existir em todo o conjunto. Mas, quantos desses dados estatísticos devem ser estudados, antes de a lei possa ser formulada? Haverá possibilidade de estudar todos os que existem? Será isso necessário, para que a lei tenha aplicação em todos os casos que porventura forem aparecendo?

Em qualquer pesquisa, de qualquer natureza, é importante levantarmos uma amostra, para inferirmos sobre o comportamento do todo. Nesta inferência, cometemos erros que são medidos por meio do cálculo de probabilidade, e o mais importante modelo de probabilidade é o modelo normal. Em qualquer levantamento estatístico, admitimos que todos os elementos que deverão fazer parte da amostra são escolhidos com igual probabilidade de dela fazerem parte.

Iremos analisar as tabelas relativas a mecanismo da amostragem, dimensionamento, intervalos de confiança e tipos de amostragem.

### 2. PRINCIPAIS TABELAS E PROBLEMAS TÍPICOS

#### 2.1. *Números aleatórios ou randômicos ou dígitos aleatórios ou números ao acaso*

Uma amostra é aleatória se todo elemento a ser escolhido de um coletivo tiver probabilidade constante. O melhor método consiste em numerar cada um dos dados do coletivo  $N$  com um número da série de 1 a  $N$ , e escolher de uma tábua de dígitos aleatórios, uma lista telefônica,  $n$  números que correspondam a  $n$  unidades. Para a construção de tabelas de números aleatórios, foram ideados diversos métodos,<sup>2</sup> que procuram evitar o aparecimento de alguns números ou grupo de números ou ordem de colocação e que foram aplicados a várias provas de hipóteses, entre elas as mencionadas por Kendall.<sup>3</sup> As principais tabelas são:

\* Professor do Departamento de Contabilidade de Atuária da Faculdade de Economia e Administração da Universidade de São Paulo.

a) Tippett, L.H.C. *Random sampling numbers*. 1. ed. London, 1927; 2. ed. London, 1952. (Tract. of Computers, n. 15).

b) Kendall, M.G. e Babington, S.B. *Tables of random sampling numbers*. London, 1939. (Tract. of Computers, n. 24).

c) Fischer, R. A. & Yates, F. *Statistical tables*. Table n. 33.

d) Horton, H. B. & Smith, R. T. *Tables of 105.000 random decimal digits*. Washington, 1949. Interstate Commerce Commission.

e) *A Million random digits with 100.000 normal deviates*. The Rand Corporation, 1955.

O uso das tábuas para obter amostras aleatórias, no caso de escolher  $n$  dados, dentro de um coletivo de  $N$ , poderá ser feito da seguinte forma:

a) ordenar e numerar os  $N$  dados do coletivo;

b) começar por qualquer bloco da tábua e percorrê-la em qualquer direção, escolhendo os  $n$  números menores do que  $N$  e os dados correspondentes. Se aparecerem números maiores do que  $N$ , são desprezados.

Temos a seguir um extrato:

90752	52210	83974	29992	65831	38857	50490
64364	67412	33339	31926	14883	24413	59744
08962	00358	31662	25388	61642	34072	81249
95012	68379	93526	70765	10592	04542	76463
15664	10493	20492	38391	91132	21999	59516
16408	81899	04153	53381	79401	21438	83035
18629	81953	05520	91962	04739	13092	97662
73115	35101	47498	87637	99016	71060	88824
57491	16703	23167	49323	45021	33132	12544
30405	83946	23792	14422	15059	45799	22716

## 2.2 Letras do alfabeto aleatórias

Os documentos contábeis poderão ter uma composição alfanumérica e, assim, poderemos compor as amostras aleatórias com as duas tabelas — letras e números aleatórios — considerando que os dois primeiros números representem as letras do alfabeto. Uma tabela de letras aleatórias poderá ser observada no livro de Herbert Arkin, *Handbook of sampling for auditing and accounting*, editado pela McGraw Hill em 1963. Um extrato desta tabela é dado a seguir:

JHLKI	LOWDS	RKDBJ	QEQZT	NPOHN	XPSIB	ECNFO	QNZKP	ZZATA	SMHIB
HZVGI	FZVVM	LPELT	EBBSB	VZGRUH	BNGLD	TORLE	ZMIRAY	EIHJ	PGTNI
JLZQY	TSLTL	KWRWN	YGEWT	EOQYV	FYLTK	WIIHS	LUGJF	JLUSN	MLEJL
PTCCP	BFQFO	PLRRT	KOAOG	FWRMS	NGJMT	WXYKI	RSJPK	MXPYD	MIRGO
EMGKD	QALGO	LITIF	XDIAZ	GTDOH	WEWQH	QLOZL	EMJUF	IFDOB	GKYFR
CDPDE	IHNPF	COHAT	RRFSF	HZDOQ	XCSKY	KTOYD	NEGIJ	GPHSA	OGZAY
PDIJD	GNGNR	EMOPE	FLTEB	XBOIR	RYVOP	UHNNO	TPTET	SKWVF	TRSWH
FXNIKQ	JAKKA	DOARL	EGNGV	TNOZY	RKNSV	PIOSC	YSNBJ	RNPTB	ONSOB
RFSIJ	CDKVG	SOKQH	MJYFS	FLOLL	ENBNC	KZPGJ	WONKR	LMFIC	NQGLY
WNVWB	PPWYV	QYHXX	NKUAM	UXILA	GKFEF	GBAFI	ZQJUK	NZOLY	

## 2.3 Meses aleatórios

Os documentos contábeis poderão ter uma composição mensal e numérica, daí a utilização de duas tabelas: meses aleatórios e dígitos aleatórios. Temos a seguir um extrato:

MAR	SEP	MAR	AUG	SEP	MAR	DEC	AUG	MAY	JAN
MAY	OCT	SEP	SEP	JUN	APR	MAY	MAY	OCT	APR
DEC	FEB	DEC	SEP	DEC	FEB	JUN	MAY	SEP	DEC
NOV	JUL	MAY	NOV	DEC	FEB	DEC	AUG	JUN	APR
DEC	FEB	JUN	JUL	MAY	OCT	NOV	JAN	NOV	NOV

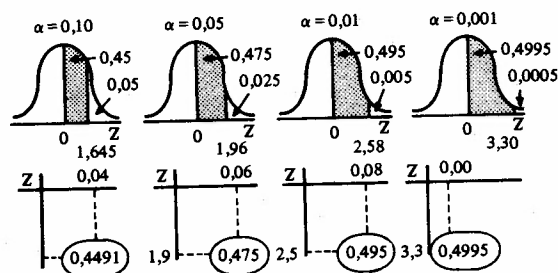
2.4. Tamanho da amostra  $n$  para amostragem por atributos (taxa de ocorrência) — população finita  $N$

$$n = \frac{Z_{\alpha/2}^2 p' q' N}{e^2(N-1) + Z_{\alpha/2}^2 p' q'}$$

onde  $n$  é o tamanho da amostra,  $N$  é o tamanho do universo,  $p'$  é a taxa de ocorrência decimal na amostra,  $q' = 1 - p'$  é a proporção de insucesso na amostra,  $\alpha$  é o nível de significância,  $1 - \alpha$  é o nível de confiança (este é, normalmente, de 90%, 95%, 99% e 99,9%),  $Z_{\alpha/2}$  é o valor da variável normal reduzida, que depende de frequência  $\alpha$ .

$\alpha$	: 0,10	0,05	0,01	0,001
$\alpha/2$	: 0,05	0,025	0,005	0,0005
$Z_{\alpha/2}$	: 1,645	1,96	2,58	3,30

Os valores de  $Z_{\alpha/2}$  são obtidos através da curva normal, sendo que  $Z_{\alpha/2}$  são as abscissas que correspondem às áreas de  $\frac{1 - \alpha}{2}$ . Assim, teremos:



Para utilizarmos a fórmula de  $n$ , deveremos especificar:

a) nível de significância ( $\alpha$ ) ou nível de confiança ( $1 - \alpha$ )

b) valor da taxa de ocorrência  $p'$ ,  $q' = 1 - p'$

c) tamanho da população  $N$

d) margem de confiança (erro)  $e$

Vejam os exemplos: Desejo fazer uma circularização de clientes. O meu cardex apresenta 4.500 clientes. Por meio de dados passados, ou de uma amostra piloto ( $n = 30$  clientes), sei que a percentagem de violações é de 2% (0,02). Qual o número mínimo de clientes a serem analisados, admitindo-se um erro (precisão) de 1% (0,01) e um nível de confiança de 95%. Substituindo os dados:

$N = 4.500; e = 0,01; p' = 0,02; q' = 0,98; 1 - \alpha = 0,95$   $Z_{\alpha/2} = Z_{0,025} = 1,96$ , na fórmula de  $n$ , teremos:

$$n = \frac{1,96^2 \cdot 0,02 \cdot 0,98 \cdot 4500}{0,01^2 \cdot 4499 + 1,96^2 \cdot 0,02 \cdot 0,98} \cong 645$$

Se a população fosse infinita, teríamos:

$$n = \frac{1,96^2 \cdot 0,02 \cdot 0,98}{0,01^2} = 753$$

O valor de  $p'$  que torna máximo  $p'q' = p'(1 - p') = p' - p'^2$ , seja:

$w = p' - p'^2; w' = 0 \Rightarrow 1 - 2p' = 0 \Rightarrow p' = 1/2$  e o valor máximo de  $p'q'$  é  $1/4$ . Então, podemos escrever:  $n \leq 0,25 \cdot \frac{Z_{\alpha/2}^2}{e^2}$ . Se  $\alpha = 0,05$  e  $e = 0,04$ , teremos  $n = 600$ .

Em função da tabela 1, poderemos tirar as seguintes conclusões:

- a) à medida que aumentamos a tolerância (margem de confiança, erro), diminui o número de elementos a serem observados;
- b) para uma mesma tolerância, se o universo aumenta, isto não implica aumentar na mesma proporção o número de elementos da amostra;
- c) se dobrarmos a tolerância, para um mesmo universo, não implica em reduzir pela metade o número de elementos a serem observados.

Tabela 1

Taxa de ocorrência  $p' = 0,02; q' = 0,98$   
Nível de significância 95% =  $Z_{0,025} = 1,96$   
Margem de confiança (erro)

N	$\pm 0,5\%$	$\pm 0,75\%$	$\pm 1\%$	$\pm 1,25\%$	$\pm 1,5\%$	$\pm 2\%$
4,400	1789	1027	643	435	311	181
4,500	1805	1032	645	436	312	181
4,600	1821	1037	648	437	312	181
4,700	1836	1042	649	438	313	181
4,800	1851	1047	651	438	313	182
4,900	1866	1052	653	439	314	182

2.5 Tabela de  $d_2$  — fatores para estimar o valor do desvio-padrão populacional  $\delta$  em função da amplitude média  $R$

O valor de  $\delta$  é dado por:  $\delta = \frac{\bar{R}}{d_2}$ , onde  $d_2$  é a tabela

que depende do número de elementos que compõe cada grupo. Assim, teremos:

Tabela 2

Nº de elementos de cada grupo	2	3	4	5	6	7	8	9	10
$d_2$	1,128	1,693	2,059	2,326	2,534	2,704	2,847	2,970	3,078

Exemplo: levantamos um grupo de 48 documentos reunindo-os em 8 grupos, sendo cada grupo composto de 6 documentos. Em cada grupo tivemos os seguintes maiores e menores valores:

Grupos	: 1	2	3	4	5	6	7	8
Maior valor	: 21	19	32	41	45	51	50	26
Menor valor	: 12	9	10	30	20	30	20	10
Amplitude (R)	: 9	10	22	11	25	21	20	16

Temos que  $(\bar{R}) = \frac{\sum \bar{R}}{8} = \frac{144}{8} = 18; d_{2;6} = 2,534$

Então, o valor de  $\delta$  será:  $\delta = \frac{\bar{R}}{d_{2;6}} = \frac{18}{2,534} = 7,103$

2.6 Tamanho da amostra  $n$  para amostragem por variáveis (valor monetário) — população finita  $N$

$$n = \frac{N}{\left(\frac{e}{\sigma}\right)^2 \frac{1}{Z_{\alpha/2}^2} (N - 1) + 1}$$

onde:  $n$  é o tamanho da amostra;  $N$  = tamanho do universo;  $\delta$  = desvio-padrão populacional;  $e$  = precisão ou margem de confiança;  $\alpha$  = nível de significância;  $1 - \alpha$  = nível de confiança;  $Z_{\alpha/2}$  é o valor da variável normal reduzida que depende do nível de significância.

Observando os saldos do exemplo acima, qual o número mínimo de documentos a ser analisado, sendo o erro da amostragem (precisão) de 1;  $1 - \alpha = 0,95$ ;  $N = 5.000$ ?

Solução:  $n = \frac{5.000}{\left(\frac{1}{7,103}\right)^2 \frac{1}{1,96^2} 4999 + 1} \cong 189$ . Se  $N \rightarrow \infty$

$$n = \frac{1}{\left(\frac{e}{\delta}\right)^2 \frac{1}{Z_{\alpha/2}^2}} = \left(\frac{Z_{\alpha/2} \cdot \delta}{e}\right)^2 = \left(\frac{1,96 \cdot 7,103}{1}\right)^2 \cong 194$$

Em função da tabela abaixo, poderemos tirar as seguintes conclusões:

- a) à medida que aumenta a margem de confiança  $e$ , diminui o número de elementos;
- b) mantendo-se a margem de confiança  $e$  constante em relação ao desvio-padrão populacional e aumentando-se o nível de confiança, o número de elementos a serem observados aumenta;
- c) aumentando-se  $\frac{e}{\delta}$ , não se reduz na mesma proporção o número de elementos a serem observados, evidentemente mantendo-se constante o nível de confiança.

Tabela 3

$\frac{e}{d}$	Tamanho da amostra — valor monetário amostras aleatórias Nível de confiança			
	90%	95%	99%	99,9%
	Tamanho do universo N = 6.000			
0,10	259	362	600	922
0,14	135	190	322	509
0,18	83	117	119	319
0,20	67	95	162	261

### 2.7 Intervalo de confiança para a proporção ou frequência ou taxa de ocorrência $p'$ ou confiabilidade da amostra para frequência relativa — população finita N

$$p' \pm Z_{\alpha/2} \sqrt{\frac{p'q'}{n}} \sqrt{1 - \frac{n}{N}}$$

onde:  $p'$  é a taxa de ocorrência na amostra;  $q' = 1 - p'$ ;  $n$  é o tamanho da amostra;  $1 - \alpha$  é o nível de confiança;  $\alpha =$  nível de significância;  $Z_{\alpha/2}$  é o valor da variável normal reduzida que depende de  $\alpha$ ;  $N$  é o tamanho da população.

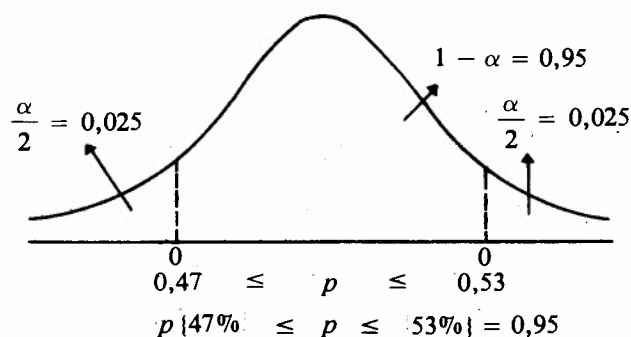
82

Exemplo: Sendo  $p' = 0,5$ ;  $q' = 0,5$ ;  $1 - \alpha = 0,95$ ;  $n = 1.000$ ;  $N = 10.000$ , qual o intervalo de confiança?

$$0,5 \pm 1,96 \sqrt{\frac{0,5 \cdot 0,5}{1.000}} \sqrt{1 - \frac{1.000}{10.000}} =$$

$$= 0,5 \pm 1,96 \cdot 0,0158114 \cdot 0,9487 = 0,5 \pm 0,0293793$$

ou 0,4706 e 0,529 ou 47,06% e 53%



### 2.8 Confiabilidade da amostra para o cálculo do intervalo de confiança de média — população finita N — grande amostra $n \geq 30$

$$\pm \frac{Z_{\alpha/2}}{\sqrt{n}} \sqrt{1 - \frac{n}{N}}$$

onde:  $n$  é o tamanho da amostra;  $N$  é o tamanho da população;  $Z_{\alpha/2}$  é o valor da variável normal reduzida.

Damos a seguir um extrato desta tabela.

Tabela 4

Tamanho da amostra	Nível de confiança			
	90%	95%	99%	99,9%
	Tamanho do universo 500			
40	$\pm 0,2494$	$\pm 0,2972$	$\pm 0,3912$	$\pm 0,5005$
50	0,2207	0,2630	0,3461	0,4414
60	0,1992	0,2373	0,3125	0,3996
70	0,1824	0,2173	0,2860	0,3658
	Tamanho do universo 1.000			
40	$\pm 0,2548$	$\pm 0,3036$	$\pm 0,3997$	$\pm 0,5113$
50	0,2268	0,2702	0,3556	0,4535
60	0,2059	0,2453	0,3229	0,4130
70	0,1897	0,2260	0,2982	0,3804

Em função do extrato acima, verificamos que:

- mantendo-se constante  $n$  e  $N$ , à medida que aumenta o nível de confiança, aumenta o erro da amostragem, como múltiplo do desvio-padrão;
- para um mesmo universo  $N$ , o nível de confiança  $1 - \alpha$ , à medida que aumenta o tamanho da amostra  $n$ , diminui o erro da amostragem, como múltiplo do desvio-padrão.

Seja um inventário composto de 4.000 itens de valor total de Cr\$ 1.289.640,00; a média por item será: Cr\$ 1.289.640 ÷ 4.000 = Cr\$ 322,41. Levantamos uma amostra-piloto de 120 itens e obtivemos uma média para as amplitudes de Cr\$ 148,49 (dividimos os 120 itens em 20 grupos de 6 itens cada, e obtivemos 20 amplitudes cuja soma foi de Cr\$ 2.969,72).

$$\text{O valor de } \delta \text{ será: } \delta = \frac{\bar{R}}{d_2} = \frac{148,49}{2,534} = 58,60$$

O erro da amostragem como múltiplo do desvio-padrão é:

$$\pm \frac{Z_{\alpha/2}}{\sqrt{n}} \sqrt{1 - \frac{n}{N}} = \pm \frac{2,58}{120} \sqrt{1 - \frac{120}{4000}} = \pm 0,2319$$

Neste caso consideramos:  $n = 120$ ;  $N = 4.000$ ;  $1 - \alpha = 0,99$ .

A exatidão da amostragem ou erro por item será:

$$\delta \frac{Z_{\alpha/2}}{\sqrt{n}} \sqrt{1 - \frac{n}{N}} = 58,60 \cdot 0,2319 = 13,59$$

A avaliação da exatidão da amostragem para o total do universo é: 4.000 · Cr\$ 13,59 = Cr\$ 54.360,00

Podemos concluir que existem 99 chances em 100 de que o inventário de Cr\$ 1.289.640,00 apresente uma flutuação de Cr\$ 54.360,00 para mais ou para menos. Os limites de controle são:

Limite superior de controle:

$$\text{Cr\$ } 1.289.640,00 + \text{Cr\$ } 54.360,00 = \text{Cr\$ } 1.344.000,00$$

Limite inferior de controle:

$$\text{Cr\$ } 1.289.640,00 - \text{Cr\$ } 54.360,00 = \text{Cr\$ } 1.235.280,00$$

## 2.9 Confiabilidade da amostra para o cálculo do intervalo de confiança da média — população infinita $N \rightarrow \infty$ , grande amostra $n \geq 30$

$$\pm \frac{Z_{\alpha/2}}{\sqrt{n}} \sqrt{1 - \frac{n}{N}}, \text{ quando } N \text{ tende para infinito,}$$

$\sqrt{1 - \frac{n}{N}}$  tende para 1 e, portanto, a expressão inicial

ficará  $\pm \frac{Z_{\alpha/2}}{\sqrt{n}}$  onde  $n$  é o tamanho da amostra e  $Z_{\alpha/2}$

o valor da variável normal reduzida.

Qual a confiabilidade da amostra, ou seja, o fator que deverei multiplicar pelo desvio-padrão, a fim de obter a exatidão da amostragem e o intervalo de confiança para a média? Considerar uma amostra de 2.500 itens e o nível de confiança de 95%.

Solução:

$$\pm \frac{Z_{\alpha/2}}{\sqrt{n}} = \pm \frac{1,96}{\sqrt{2.500}} = \pm \frac{1,96}{50} = \pm 0,0392$$

Em função da tabela abaixo, poderemos concluir o seguinte:

- matendo-se constante o tamanho da amostra  $n$ , à medida que aumenta o nível de confiança, aumenta o erro da amostragem como múltiplo do desvio-padrão;
- para um mesmo nível de confiança, à medida que aumenta o tamanho da amostra, diminui o erro da amostragem como múltiplo do desvio-padrão.

Tabela 5

Tamanho da amostra	Nível de confiança			
	90%	95%	99%	99,9%
Erro da amostragem como múltiplo do desvio-padrão				
30	±0,3003	±0,3578	±0,4710	±0,6025
31	0,2954	0,3520	0,4634	0,5927
32	0,2908	0,3465	0,4561	0,5834
33	0,2864	0,3412	0,4491	0,5745

## 2.10 Fator de correção para a população finita

$$\sqrt{\frac{N-n}{N-1}} \equiv \sqrt{1 - \frac{n}{N}}$$

onde:  $n$  e  $N$  são tamanhos de amostra e população respectivamente. Assim, poderemos ter a seguinte tabela:

Tabela 6

Proporção da amostra na população	Fator de correção	Proporção da amostra na população	Fator de correção
$\frac{n}{N}$ %	Finita	$\frac{n}{N}$ %	Finita
10,0	0,9487	14,7	0,9236
10,1	0,9482	14,8	0,9230
10,2	0,9476	14,9	0,9225
10,3	0,9471	15,0	0,9220

O valor 0,9487 foi obtido com:  $n = 100$ ;  $N = 1.000$

$$e \sqrt{1 - \frac{100}{1.000}} = \sqrt{1 - 0,10} = \sqrt{0,9} = 0,9487. \text{ Podemos observar que, à medida que aumenta a relação } \frac{n}{N},$$

diminui o fator de correção, uma vez que diminui  $\sqrt{1 - \frac{n}{N}}$ .

## 2.11 Amostragem por descoberta — probabilidade de incluir pelo menos uma ocorrência na amostra

$$P(E; e \geq 1; N; n) = 1 - P(E; 0; N; n) = 1 - \frac{\binom{E}{0} \binom{N-E}{n}}{\binom{N}{n}}$$

Adota-se o modelo *hipergeométrico* (sem reposição) quando, em uma população de  $N$  documentos, há  $E$  errados. Escolhendo-se  $n$  documentos aleatoriamente, qual a probabilidade de encontrarmos  $e$  errados?

Quando  $\frac{n}{N} \leq 0,10$ , adota-se o modelo binomial como sendo uma boa aproximação de modelo hipergeométrico.

Uma população é composta de 200 documentos. De experiências passadas, sabemos que a percentagem de erros é de 1% ( $E = 0,01 \cdot 200 = 2$ ). Levantando-se uma amostra de 10 documentos ( $n = 10$ ), qual a probabilidade de termos pelo menos um documento errado na amostra?

$$P(200; 2; 10; e \geq 1) = 1 - \frac{\binom{2}{0} \binom{200-2}{10}}{\binom{200}{10}}$$

$$= 1 - \frac{2! \cdot 198!}{0!2! \cdot 10!188!} = 1 - \frac{198!}{10!188!} \cdot \frac{10!190!}{200!}$$

$$= 1 - \frac{190 \cdot 189}{200 \cdot 199} = 1 - 0,90226 = 0,0977 \text{ ou } 9,8\%$$

Em função da tabela seguinte, poderemos tirar as seguintes conclusões:

a) mantendo-se constante  $n$  e  $N$ , à medida que aumentamos a taxa de ocorrência (erros) na população, aumenta a probabilidade de incluir pelo menos uma ocorrência na amostra;

b) para uma mesma população e taxa de ocorrência na população, a probabilidade de incluir pelo menos uma ocorrência na amostra aumenta à medida que aumentamos o tamanho da amostra  $n$ .

Tabela 7

Quando a taxa de ocorrência é					
Tamanho da Amostra	0,5%	1%	1,5%	2,5%	3%
Probabilidade de encontrar pelo menos uma ocorrência é					
Tamanho do universo é 200					
10	5%	9,8%	14,3%	18,7%	22,8%
25	12,5%	23,5%	33,2	41,6	49,1
40	20,0	36,1	49,0	59,4	67,6
Quando a taxa de ocorrência é					
Probabilidade de encontrar pelo menos uma ocorrência é					
Tamanho do universo é 200					
	4%	5%	7,5%	10%	12,5%
10	34,2%	40,9%	55,0%	66,0%	74,6%
20	65,6	74,6	87,5	94,0	97,2

## 2.12 Plano de amostragem para aceitação

O plano de amostragem para aceitação inclui o número de elementos da população  $N$ , o tamanho da amostra  $n$  e um número de aceitação  $c$ . O número de aceitação é o número máximo de erros ou outro evento que é tolerado na amostra, para o universo ser considerado satisfatório. Admitindo-se o modelo hipergeométrico:

$$P_H(N; E; n; e) = \frac{\binom{E}{e} \binom{N-E}{n-e}}{\binom{N}{n}}$$

O modelo binomial é:  $P_b(n; p; x) = \binom{n}{x} p^x q^{n-x}$ ,  $p = 1 - q$ ,  $p = \frac{x}{n}$ ;  $\mu = np$  e  $\delta^2 = npq$ . O modelo de Poisson apresenta a forma:

$$P_p(\mu; x) = \frac{e^{-\mu} \mu^x}{x!}, \mu = np = \delta^2$$

Em geral, na construção de planos de amostragem simples, para que uma partida ou lote seja aceitável na inspeção, fixa-se o número máximo de erros que se permite, na amostra de tamanho  $n$ ; tal número máximo de erros,  $c$ , denomina-se *número de aceitação*. Ocorrendo na amostra um ou mais erros além de  $c$ , a partida será rejeitada; define-se assim, em consequência, o *número de rejeição*  $r = c + 1$ . Naturalmente, devemos ter:  $0 \leq c \leq (n - 1)$  e  $1 \leq r \leq n$ .

Vejamos dois exemplos do modelo hipergeométrico (sem reposição):

1. seja uma partida de  $N = 50$  documentos com número de erros  $E = 2$ . Qual a probabilidade de aceitação da partida, inspecionando-se uma amostra de  $n = 10$ , com um número de aceitação  $c = 1$ ?

No caso,  $f = \frac{n}{N} = 0,20 > 0,10$  e, por isso, deveremos utilizar a distribuição hipergeométrica. Se fosse  $f < 0,10$ , deveríamos utilizar o modelo binomial.

$$P_H(N; E; n; e) = \frac{\binom{E}{e} \binom{N-E}{n-e}}{\binom{N}{n}} \text{ para } e = 0; e = 1, \text{ teremos:}$$

$$P_H(50; 2; 10; 0) = \frac{\binom{2}{0} \binom{48}{10}}{\binom{50}{10}} = 0,637; P_H(50; 2; 10; 1) =$$

$$= \frac{\binom{2}{1} \binom{48}{9}}{\binom{50}{10}} = 0,326$$

As probabilidades de aceitação e rejeição são respectivamente:

$$F(c) = \sum_{e=0}^c \frac{\binom{E}{e} \binom{N-E}{n-e}}{\binom{N}{n}} e + 1 - F(c)$$

Para  $c = 1$ , será:  $F(1) = P(0 \leq e \leq 1) = 0,637 + 0,326 = 0,963$ , e a probabilidade de rejeição será:  $1 - F(1) = 0,037$

2. Se levantarmos uma amostra de  $n = 40$  documentos, de uma população composta de  $N = 200$  documentos, sabendo-se que a percentagem ou taxa de erros é de 1%, ou seja,  $E = 2$ , qual a probabilidade de aceitação para o número de aceitação (número máximo de erros ou outro evento que é tolerado na amostra para o universo ser considerado satisfatório) 1,  $c = 1$ ?

$$P_H(N; E; n; c = 1) = P_H(N; E; n; e = 0) + P_H(N; E; n; e = 1)$$

$$P_H(200; 2; 40; c = 1) = P_H(200; 2; 40; 0) + P_H(200; 2; 40; 1) =$$

$$= P_H(200; 2; 40; c=1) = \frac{\binom{2}{0} \binom{200-2}{40-0}}{\binom{200}{40}} + \frac{\binom{2}{1} \binom{200-2}{40-1}}{\binom{200}{40}} =$$

$$\frac{198!}{40!158!} + \frac{40!60!}{200!} + 2 \cdot \frac{198!}{39!159!} + \frac{40!60!}{200!} =$$

$$P_H(200; 2; 40; c = 1) = \frac{159.160}{200.199} + \frac{2.40.160}{200.199} =$$

$$= 0,639 + 0,3216 = 0,9606 \text{ ou } 96,1\%$$

Se  $c = 2$ , teremos:

$$P_H(200; 2; 40; c = 2) = \frac{\binom{2}{0} \binom{200-2}{40-0}}{\binom{200}{40}} + \frac{\binom{2}{1} \binom{200-2}{40-1}}{\binom{200}{40}} +$$

$$+ \frac{\binom{2}{2} \binom{200-2}{40-2}}{\binom{200}{40}} = 0,639 + 0,3216 + 0,0391$$

$P_H(200; 2; 40; c = 2) = 0,9997$  ou  $99,97\%$ , aproximadamente  $100\%$

Os cálculos acima poderão ser observados na tabela 8:

Tabela 8

		Tamanho do universo 200 Taxa de erros contidos no universo de:				
Tamanho da amostra	Número de aceitação	0,5%	1%	1,5%	2%	2,5%
A probabilidade de aceitação é:						
(n)	(c)					
40	0	80,0%	63,9%	51,0%	40,7%	32,4%
40	1	100,00	96,1%	89,8	82,1	73,8
40	2	100,00	100,0%	99,3	97,4	94,4
40	3	100,00	100,0	100,0	99,9	99,4
50	0	75,0	56,2	42,0	31,3	23,3
50	1	100,0	93,8	84,5	73,9	63,3
50	2	100,0	100,0	98,5	95,1	89,9
50	3	100,0	100,0	100,0	99,7	98,6

Em função da tabela 8, temos as seguintes observações:

a) mantendo-se  $N$ ,  $n$  e  $c$ , à medida que *aumenta* a taxa de erros na população  $E$ , *diminui* a probabilidade de aceitação;

b) mantendo-se  $N$ ,  $n$  e  $E$  (taxa de erros na população), à medida que *aumenta* o número de aceitação  $c$ , *aumenta* a probabilidade de aceitação;

c) mantendo-se  $N$ ,  $c$  (número de aceitação, número máximo de erros na amostra, para aceitarmos o lote ou universo ou partida) e  $E$  (taxa de erros ou número

de erros no universo), à medida que *aumenta* o tamanho da amostra  $n$ , *diminui* a probabilidade de aceitação;

d) mantendo-se  $n$ ,  $c$ ,  $E$ , à medida que *aumenta* o tamanho do universo  $N$ , *aumenta* a probabilidade de aceitação;

e) uma amostra com um número de erros igual ao da população, apresenta a probabilidade de  $100\%$  de aceitação. Por exemplo:

$$P_H(200; 2; 40; c = 2) = P_H(200; 2; 40; 0) + P_H(200; 2; 40; 1) + P_H(200; 2; 40; 2) = 0,639 + 0,3216 + 0,0391 = 0,997 \text{ ou } 100\%.$$

### 2.13 Tabela de amostragem por conglomerados — erro da amostragem como múltiplo do desvio-padrão populacional $\delta$

$$\frac{t_{\alpha/2; n-1}}{\sqrt{n-1}}$$

Para o cálculo do intervalo de confiança para a média, temos que:

$$P\left\{x - t_{\alpha/2; n-1} \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{\alpha/2; n-1} \frac{s}{\sqrt{n}}\right\} = 1 - \alpha$$

considerando-se uma população infinita. O erro da amostragem ou margem de confiança é dado por:

$$t_{\alpha/2; n-1} \frac{s}{\sqrt{n}} = t_{\alpha/2; n-1} \sqrt{\frac{\sum(x_i - \bar{x})^2}{n-1}} = t_{\alpha/2; n}$$

$$= t_{\alpha/2; n-1} \frac{1}{\sqrt{n-1}} \sqrt{\frac{\sum(x_i - \bar{x})^2}{n}}$$

$$= \frac{t_{\alpha/2; n-1}}{\sqrt{n-1}} \delta, \text{ onde } \delta = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n}}$$

Apresentamos a seguir um extrato desta tabela:

Tabela 9

Número de conglomerados	Nível de confiança		
	90%	95%	99%
	Erro de amostragem como múltiplo do desvio-padrão		
10	0,6100	0,7533	1,0833
11	0,5724	0,7052	1,0024
12	0,5427	0,6633	0,9377
13	0,5138	0,6293	0,8833
14	0,4909	0,5991	0,8348
15	0,4704	0,5719	0,7964

Vejamos dois exemplos:

a) Qual o erro de amostragem para  $n = 15$ ,  $1 - \alpha = 0,90$  ou 90%?

Solução:

$$\frac{t_{\alpha/2; n-1}}{\sqrt{n-1}} = \frac{1,7613}{\sqrt{15-1}} = 0,4707; \quad t_{\alpha/2; n-1} = t_{0,10; 14} = 1,7613$$

(Consultar *Estatística — teoria e exercícios* de E. E. Bonini e S. E. Bonini, p. 434)

b) Qual o erro da amostragem para  $n = 20$ ;  $1 - \alpha = 0,95$  ou 95%?

Solução:

$$\frac{t_{\alpha/2; n-1}}{\sqrt{n-1}} = \frac{2,0930}{\sqrt{20-1}} = 0,4795; \quad t_{\alpha/2; n-1} = t_{0,05; 19} = 2,0930$$

### 3. CONCLUSÃO E SUMÁRIO

Em qualquer pesquisa de cunho científico, é fundamental a aplicação de métodos quantitativos.

Raramente é possível levantar todos os elementos de uma população devido a restrições de tempo, presteza, especialização, economia, etc. Em função das restrições levantadas, observamos uma amostra com o objetivo de inferir sobre o coletivo.

O Comitê de Amostragem Estatística do Instituto Americano de Contadores Públicos Certificados publicou, em fevereiro de 1962, um relatório especial intitulado *A Amostragem Estatística e o Auditor Independente*. Este relatório tratava da natureza geral da amostragem estatística e de sua aplicabilidade à auditoria, observando, entre outros, os seguintes detalhes:

— São recomendáveis um maior conhecimento de amostragem estatística e pesquisa adicional sobre sua aplicabilidade, por parte da profissão.

— O Comitê é de opinião que o uso da amostragem estatística é permitido, de acordo com os padrões de auditoria geralmente aceitos.

— As amostras estatísticas são avaliadas em termos de *precisão*, que se expressa como uma escala de valores positivos e negativos, em torno do resultado da amostra, e *confiança*, que se expressa como a proporção dessas escalas, com base em todas as possíveis amostras semelhantes do mesmo tamanho e que incluíram o valor da população real.

— Embora a amostragem estatística forneça ao auditor uma medida de precisão e confiança, as técnicas estatísticas não definem para ele os valores de cada uma, necessários à segurança no trabalho de auditoria.

— A especificação da precisão e confiança necessárias em um determinado teste é uma função da auditoria e deve basear-se em julgamento, da mesma maneira que uma decisão quanto à segurança necessária em auditoria, quando não se usa a amostragem estatística.

— Ao determinar a extensão de um dado teste de auditoria e do método para selecionar itens a serem examinados, o auditor pode considerar a possibilidade de usar as técnicas de amostragem estatística comprovadamente vantajosas em certas circunstâncias. O uso de amostragem estatística não reduz a necessidade do julgamento do auditor, mas fornece certas medidas estatísticas quanto aos resultados dos testes de auditoria, as quais, de outra forma, não se podem obter!<sup>4</sup>

O presente artigo tem por objetivo eliminar a subjetividade dos levantamentos contábeis por amostragem, ou seja, responder às perguntas de como, onde, quanto e por que dos elementos comprobatórios contábeis.

<sup>1</sup> Instituto de Auditores Independentes do Brasil. 5. Seção Regional. *Exposição de normas de auditoria*. São Paulo/Paraná, 1977. n.1, p. 25.

<sup>2</sup> Horton, H. B. A Method of obtaining random numbers. *Ann. of Mat. Statist.*, p. 81-5, 1948.

<sup>3</sup> Kendall, M. G. & Babington, S. B. Randomness and random sampling numbers. *Journ. of Statistic. Soc.*, 1:147 seg., 1938.

<sup>4</sup> *Exposição de normas de auditoria*. cit.

### REFERÊNCIAS BIBLIOGRÁFICAS

Arkin, H. *Handbook of sampling in auditing and accounting*. New York McGraw Hill, 1962.

Barnes, R. M. *Work sampling*. 2.ed. John Wiley & Sons, 1964.

Bonini, E. E. & Bonini, S. E. *Estatística — teoria e exercícios*. São Paulo, Loyola, 1976.

Cohran, G. William. *Técnicas de amostragem*. Fundo de Cultura, 1965.

Grant, E.L. *Statistical Quality Control*. McGraw Hill.

Hill, Henry P. Roth; Joseph, L & Arkin, H. *Sampling in auditing*. New York, The Ronald Press, 1962.

Instituto de Auditores Independentes do Brasil. 5. Seção Regional. *Exposição de Normas de Auditoria*. São Paulo/Paraná, 1977. n.1.

Lourenço, Ruy de C.B. Filho. *Controle estatístico de qualidade*. Rio de Janeiro, Ao Livro Técnico, 1964.

Madow, William G. *Teoria dos levantamentos por amostragem*. IBGE.

Rodrigues, Milton da Silva. *Vocabulário brasileiro de estatística*, IBGE, 1956.

Tagliacarne, Guglielmo. *Pesquisa de mercado, técnica e prática*, Atlas, 1976.

Trueblood, Robert M. & Cyest, R. M. *Sampling techniques in accounting*, Englewood Cliffs, New Jersey, Prentice Hall, 1957.

Vance, Lawrence L. & Neter, John. *Statistical sampling for auditors and accountants*. New York, John Wiley & Sons, 1956.