INTRODUÇÃO

A estratégia de avaliação dos riscos ambientais constitui grande desafio para os profissionais de higiene ocupacional, vez que não há normalização consolidada definindo o número de medições a serem realizadas em cada GHE (Grupo Homogêneo de Exposição), frequência de avaliação e tratamento estatístico dos dados no Brasil. A NR-9, que trata do PPRA (Programa de Prevenção dos Riscos Ambientais), determina que o monitoramento dos riscos deve ser feito de maneira sistemática e repetitiva, no entanto, não define a estratégia a ser seguida. Essa norma determina também que pelo menos uma vez por ano deve ser feita análise global do PPRA visando promover ajustes e estabelecer novas metas. Todavia, essa regra, na maioria das vezes, é interpretada equivocadamente, pois os profissionais da área entendem que deve ser adotada a estratégia de avaliar os riscos ambientais sempre a cada ano, na data que expirou o cronograma das ações do PPRA. Ora, realizar avaliação dos riscos ambientais somente em determinada época do ano não tem nenhum critério científico, pois, como será visto posteriormente, as variações da intensidade e concentrações variam em função de diversos fatores.

A NR-22 adotou para avaliação das poeiras minerais, com base no NIOSH (National Institute for Occupational Safety and Health), critério do tamanho da amostra em função do número de trabalhadores do GHE. No entanto, não definiu a estratégia de avaliação e tratamento estatístico dos dados conforme definido na norma do NIOSH.

A Instrução Normativa n. 1 de 20.12.95, que trata da avaliação ocupacional do benzeno, estabelece estratégia de avaliação e tratamento estatístico dos dados, podendo, desse modo, ser tomada como referência na avaliação dos outros agentes ambientais.

Em 2018, a FUNDACENTRO publicou o *Guia Técnico sobre estratégia de amostragem e interpretação de resultados de avaliações quantitativas de agentes químicos em ambientes de trabalho*. O guia detalha os procedimentos técnicos e define estratégia de amostragem para a determinação de concentração de agentes químicos no ar dos ambientes de trabalho e também orienta a análise estatística dos dados obtidos e a comparação destes dados com os valores de referência estabelecidos pela legislação brasileira. Este guia pode ser adotado pelos higienistas brasileiros, facilitando a compreensão da matéria e solucionando dúvidas do dia-a-dia dos profissionais da área de saúde e segurança do trabalhador.

A ABHO (Associação Brasileira de Higienistas Ocupacionais) em breve disponibilizará a versão em português do livro *A Strategy for Assessing and*

Managing Occupational Exposures – Estratégia para avaliar e gerenciar exposições ocupacionais da AIHA (American Industrial Hygiene Association).

A definição da estratégia de avaliação e o tratamento estatístico dos agentes ambientais torna-se necessário devido às variações de concentrações que podem ocorrer durante a jornada de trabalho, na semana, no mês ou até mesmo nas estações do ano, bem como as alterações nos processos produtivos, máquinas e equipamentos em manutenção ou sem manutenção; funcionários que faltam, erros de amostragens dentre outros fatores e que podem gerar resultados apresentando maior ou menor concentração; assim, busca-se a estimativa da exposição real dos trabalhadores aos agentes durante o trabalho de forma mais confiável possível com o auxílio da estatística.

De acordo com o NIOSH, alguns fatores que podem alterar os dados ambientais levantados em uma medição podem ser:

— erro no método analítico;

— erro de amostragem;

— variação de concentração.

Desse modo, tomar decisão sobre a exposição ocupacional de determinado agente, com base somente em uma medição, não é aconselhável.

Neste manual, será abordada a estratégia de avaliação dos riscos ambientais, tomando como referências as fontes bibliográficas sobre a matéria, especialmente as relacionadas a seguir:

— NIOSH — Occupational Exposure Sampling Strategy Manual — EUA, 1977

— AIHA — A strategy for assessing and managing occupational exposures — EUA, 2006

— INSH — Método de avaliação à exposição a agentes químicos por inalação — Espanha

— IN 01, 20.12.95 — Avaliação da concentração de benzeno em ambientes de trabalho — Brasil

— NTP 270 — Evaluación de la exposición al ruido. Determinación de niveles representativos. Instituto Nacional de Seguridad e Higiene em el Trabajo — Espanha

— NTP 950 — Estratégia de medição e avaliação à exposição ao ruído

Com a finalidade de facilitar o entendimento, este manual está dividido em quatro capítulos, sendo que no primeiro serão apresentados conceitos básicos de estatísticas; no segundo as definições dos parâmetros utilizados nas avaliações ambientais; no terceiro capítulo serão tratadas as estratégias de avaliação de agentes químicos e no quarto capítulo as estratégias de avaliação de agentes físicos.

CAPÍTULO I

NOÇÕES BÁSICAS DE ESTATÍSTICA

A estratégia de avaliação dos riscos ambientais é ligada ao tratamento estatístico dos dados, sendo assim, nesse capítulo, serão abordados os conceitos e noções básicas de estatística visando proporcionar aos leitores melhor entendimento do conteúdo deste livro. Este capítulo tem como objetivo apenas fornecer definições básicas de estatística, que serão utilizadas no tratamento dos dados das avaliações dos riscos ambientais, abordados posteriormente. Todavia, é importante que os leitores consultem outras bibliografias sobre estatística visando esclarecer possíveis dúvidas sobre a matéria. A pretensão não é ensinar estatística, apenas fornecer elementos básicos para a compreensão da estratégia de avaliação dos riscos ambientais, bem como tratar estatisticamente os dados, conforme abordados nos capítulos seguintes.

A estatística é parte da matemática e sua finalidade é extrair informações de dados para obter uma melhor compreensão dos fenômenos, por meio de coleta, organização, descrição, análise e interpretação dos dados. A estatística basicamente se divide em estatística descritiva e inferencial ou indutiva.

A estatística descritiva apenas descreve os dados observados em determinada amostra, sem previsão sobre os parâmetros do universo. Já a estatística inferencial ou indutiva analisa e interpreta os dados pelos parâmetros do universo.

No estudo estatístico temos as medidas de tendência central (média, mediana e moda) e as medidas de dispersão ou variabilidade (amplitude total, desvio padrão e coeficiente de variação).

1. MÉDIA

1.1. Média Aritmética

A média aritmética é o parâmetro básico na determinação do valor dos dados levantados. Considerando os valores x_1; x_2, x_3.....x_n a média aritmética é igual a:

$$\overline{x} = \frac{x_1 + x_2 + x_3 \ldots\ldots x_n}{n}$$

Exemplo: as notas de matemática de 4 alunos foram 9, 10, 6 e 7. A média aritmética $\overline{x}$ é igual a:

$$\overline{x} = \frac{9,0 + 10,0 + 6,0 + 7,0}{4} = 8,0$$

1.2. Média Ponderada

A média ponderada é derivada da média aritmética que leva em conta os pesos. **Exemplo**: em um concurso a prova de matemática têm peso 2 e a de português peso 3. Considerando que determinado concorrente obteve notas 6,0 em matemática e 7,0 em português, a média ponderada é igual a:

$$\overline{x} = \frac{6,0 \times 2 + 7,0 \times 3}{5} = 6,6$$

1.3. Média Geométrica

A média geométrica é a raiz do produto de n valores, conforme equação:

$$\overline{x} = \sqrt[n]{x_1 \,.\, x_2 \,.\, x \ldots x_n}$$

Exemplo: A média de geométrica de 2, 5 e 10 é igual a:

$$\overline{x} = \sqrt[3]{2 \,.\, 5 \,.\, 10} = 4,64$$

1.4. Moda

É o valor que mais se repete num conjunto de dados. **Exemplo**: as notas na prova de português de seis alunos foram 5, 6, 6, 7, 7 e 7. A moda é igual a 7. Quando não há nenhum valor que foi registrado com maior frequência, não há moda nos dados coletados.

1.5. Mediana

A mediana é o valor central exato de determinada amostra.

Exemplo: As notas de geografia de cinco alunos foram 4, 5, 7, 4 e 7. Primeiramente deve-se ordenar os valores na ordem crescente: 4, 4, 5, 7 e 7. Nesse caso, a mediana é igual a 5. Quando a quantidade de dados da amostra for par, a identificação da mediana não é direta. Exemplo: As notas de história de seis alunos foram 4, 4, 5, 6, 7 e 7. Na amostra há dois valores centrais 5 e 6. Nesse caso, a mediana é a média de 5 e 6 que é igual a 5,5.

1.6. Amplitude Total

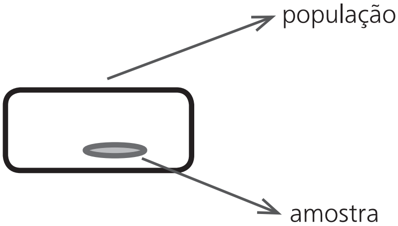
A amplitude mede o intervalo mínimo e máximo dos dados da amostra. **Exemplo**: a quantidade de gols assinalados por determinado time em 5 partidas foram: 5, 3, 1, 2 e 2. A amplitude total é igual a: $X_{max} - X_{min}$, ou seja, 5 - 1 = 4. O uso da amplitude total apresenta desvantagem quando a variação dos dados é grande. Sendo assim, o uso da variância e do desvio padrão são mais adequados na análise de dispersão dos dados. Para o estudo da dispersão dos dados, inicialmente, é necessário definir população e amostra.

1.7. População ou universo

Elementos ou indivíduos com pelo menos uma característica comum ao objeto em estudo. Exemplos: população de uma cidade, empregados de determinada empresa, entre outros.

1.8. Amostra

É o subconjunto ou parte da população. Numa pesquisa com todos os habitantes de uma cidade selecionam-se, por exemplo, uma amostra de 200 pessoas. Normalmente, os parâmetros da população é amostral.



A tabela a seguir mostra a nomenclatura dos principais parâmetros usados na estatística para população e amostra.

Parâmetro	População	Amostra
Média	μ	$\bar{x}$
Desvio padrão	σ	s
Tamanho	N	n

1.9. Variância

A variância é a dispersão dos valores da amostra em relação à média. Para a população a variância é definida pela equação:

$$\sigma^2 = \frac{\Sigma (x_i - \mu)^2}{N}$$

Onde:

σ^2 — variância

x_i — variável

μ — média

N — tamanho

Para a amostra a variância é expressa pela seguinte equação:

$$s^2 = \frac{\Sigma (x_i - \bar{x})^2}{n - 1}$$

Onde:

s^2 — variância

x_i — variável

$\bar{x}$ — média

n — tamanho

Exemplo: As notas, na prova de história, de cinco alunos (amostra) foram 6, 5, 7, 8 e 10. Para o cálculo da variância, inicialmente, deve-se calcular a média aritmética:

$$\bar{x} = \frac{6 + 5 + 7 + 8 + 10}{5} = 7{,}2$$

Para facilitar o entendimento os dados e os cálculos foram agrupados na tabela a seguir:

X_i (notas)	**$\bar{x}$ (média)**	**$(X_i - \bar{x})$**	**$(X_i - \bar{x})^2$**
6,0	7,2	-1,2	1,44
5,0	7,2	-2,2	4,84
7,0	7,2	-0,2	0,04
8,0	7,2	0,8	0,64
10,0	7,2	2,8	7,84
Soma			14,8

A média da soma deve ser dividida por n-1. Sendo assim, o valor da variância ou dispersão é igual a:

$$s^2 = \frac{14{,}8}{5 - 1} = 3{,}7$$

1.10. Desvio Padrão

Ao elevar ao quadrado as diferenças entre os dados e a média, a variância não fica na mesma unidade. Sendo assim, para a adequação das unidades calcula-se o desvio padrão (S), que para a amostra, é a raiz quadrada da variância:

$$s = \sqrt{s^2}$$

Para o exemplo anterior temos:

$$s = \sqrt{3{,}7} = 1{,}92$$

O desvio padrão pode ser calculado diretamente por meio das seguintes equações:

a) Desvio padrão da população

$$\sigma = \sqrt{\frac{\Sigma\,(x_i - \mu)^2}{N}}$$

Onde:

σ — desvio padrão

x_i — variável

μ — média

N — tamanho

b) Desvio padrão da amostra

$$s = \sqrt{\frac{\Sigma\,(x_i - \bar{x})^2}{n - 1}}$$

Onde:

s — desvio padrão

x_i — variável

$\bar{x}$ — média

n — tamanho

1.11. Coeficiente de Variação

O coeficiente de variação para a amostra é a razão do desvio padrão sobre a média multiplicado por 100, conforme equação:

$$CV = \frac{s}{\bar{x}} \cdot 100$$

Quanto maior o coeficiente de variação maior é a dispersão dos dados.

Quanto menor o valor do coeficiente de variação, menor é a dispersão da amostra, ou seja, os dados são mais homogêneos.

1.12. Erro Padrão da Amostra

Calcula-se o erro padrão da amostra devido a aleatoriedade das mesmas. O erro padrão ou margem de erro é calculado pela seguinte equação:

$$E_p = \frac{s}{\sqrt{n}}$$

Onde:

s — desvio padrão

n — é o tamanho da amostra

2. PROBABILIDADES

O estudo da probabilidade auxilia o higienista ambiental a sair da perspectiva da incerteza para a perspectiva do risco. A probabilidade é um número entre 0 (zero) e 1 (um); ou de 0 a 100%. Desse modo, a probabilidade pode ser expressa da seguinte forma: $0 \leq P \leq 1$ ou $0 \leq P \leq 100\%$;

2.1. Distribuição Normal

A distribuição normal é uma das mais importantes da estatística. Essa distribuição é definida pela média e pelo desvio padrão. Os dados coletados se distribuem numa curva em forma de um sino, conforme mostra a figura 1:

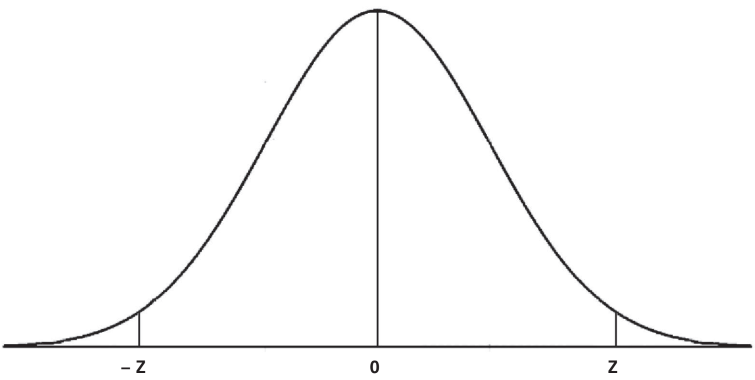


Figura 1 — Curva da distribuição normal

A curva normal padronizada apresenta a média (μ) igual a zero e desvio padrão (σ) igual a 1,0. A probabilidade é calculada pela área abaixo da curva. A área total sob a curva é igual a 1,0. Para obter a probabilidade, calcula-se a média e o desvio padrão. A área sob a curva é obtida por meio de tabelas as quais são encontradas com a probabilidade acumulada ou bilateral. A tabela 1 (página 31) mostra a tabela de distribuição normal bilateral, enquanto a tabela 2 (página 32) mostra a distribuição acumulada ou unilateral.

No caso de teste da significância para ambos os lados deve-se utilizar a tabela 1; no teste somente de um dos lados utiliza-se a tabela 2. O Z é utilizado para estimar coeficientes. A escolha da tabela a ser utilizada deve ser feita por meio do teste de hipótese.

a) Testes de hipóteses[1]

As tabelas estatísticas apresentam valores em limites bilaterais e unilaterais, ou seja, para nível de significância para dois lados ou só um lado. Assim, por exemplo, para confiança de 95%, os valores de Z estão compreendidos entre -1,96 e + 1,96, para limite bilateral. A área sob a curva nesse intervalo é igual a 0,95 ou 95% (ver tabela 1). No caso do limite unilateral o valor de Z igual 1,645 sendo a área sob a curva também igual a 0,95 (Ver tabela 2). Para a escolha da tabela usada é necessário efetuar os testes das hipóteses:

— Hipótese nula (H_0): é a hipótese que assume como verdade para a construção do teste. É chamada hipótese nula ou verdadeira.

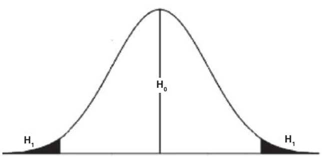
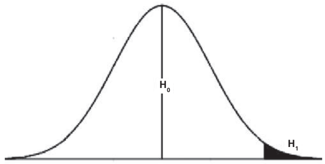
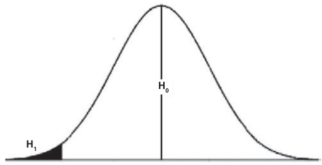
— Hipótese Alternativa (H_1): será uma hipótese complementar de H_0, fornecerá alternativa à hipótese nula.

Assim, o teste consiste em aceitar ou rejeitar a Hipótese Nula (H_0). Exemplo: a altura média dos homens de determinada cidade é de 1,70 metros. Qual a probabilidade com nível de significância de 5% ou 95% de confiança da existência de homens dessa cidade com altura superior a essa média? A hipótese nula ou verdadeira H_0=1,70 metros. A hipótese alternativa H_1 será a verdadeira, se houver homens com altura superior a 1,70 metros.

Se a formulação da hipótese alternativa indicar que o parâmetro é maior ou menor, o valor de teste (valor verdadeiro), o teste será Unilateral. Se a formulação da hipótese alternativa indicar que o parâmetro é diferente do valor de teste (verdadeiro) será Bilateral[2]. De maneira reduzida pode-se identificar as hipóteses de acordo com a regra a seguir:

---

(1) GRINGS, José Fernando. *Testes de hipóteses*. Disponível em: <http://www.omatematico.com>. Acesso em: 15.8.15.
(2) INE 7002 — Inferência Estatística — Testes de Hipóteses. Capítulo 10. Disponível em: <http://www.inf.ufsc.br/~marcelo/Cap10.pdf>. Acesso em: 15.8. 2015.

H_0 ; $X = \mu$	H_0 ; $X = \mu$	H_0 ; $X = \mu$
$H_1 : X \neq \mu$	$H_1 : X > \mu$	$H_1 : X < \mu$
Bilateral	Unilateral à direita	Unilateral à esquerda
		

(GRINGS, 2015)

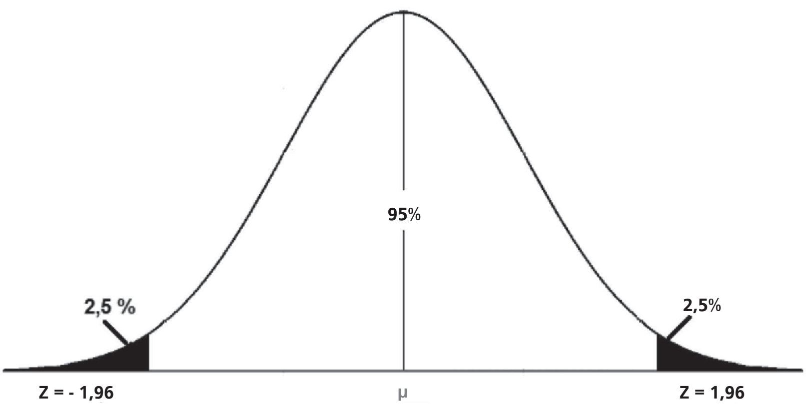
O valor de x é calculado e μ é a média da população, podendo ser aplicada a regra também para amostras.

Exemplo: Considerando que o peso médio da população de uma cidade é igual a 75 quilos. Qual o percentual com 95% de certeza da existência de pessoas com peso superior a 75 quilos? Nesse caso, o teste é unilateral à direita.

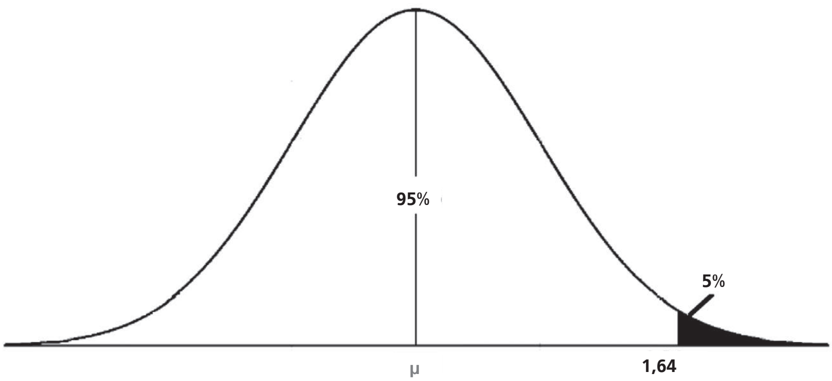
Se o objetivo fosse conhecer com 95 % de certeza as pessoas com peso inferior a 75 quilos, o teste seria unilateral à esquerda.

Se o objetivo fosse definir o limite e confiança de 95% para o peso de 75 quilos o teste seria bilateral. Nesse teste, o peso pode ser para mais quanto para menos.

A figura abaixo mostra a distribuição normal bilateral. Os valores de Z, nesse tipo de distribuição, estão na tabela 1. Nessa tabela, o valor de confiança 95% corresponde a z = 1,96 (ver tabela 1, valor assinalado no final deste capítulo). Nessa tabela, o valor da área para z = 1,96 é igual a 0,475. Esse valor multiplicado por dois (ambos os lados) será igual 0,95 ou 95%.



A figura a seguir mostra a distribuição unilateral. Os dados de z, nesse tipo de distribuição, estão na tabela 2. Nessa tabela, o valor de confiança 95% corresponde a z = 1,65 (ver tabela 2 no final deste capítulo).



Na tabela t da distribuição Student aplica-se a mesma regra. A tabela 3 mostra os valores para os testes unilateral e bilateral. Assim, para o grau de liberdade (n - 1) igual a 10 e nível de significância 5% (95% de confiança), na distribuição unilateral, o valor de t é de 1,8125. Já na bilateral, o valor de t é 2,2281 (ver valores assinalados na tabela 3, no final deste capítulo). Portanto, no uso das tabelas, é importante realizar teste de hipóteses. Como será visto posteriormente, as normas de tratamento estatístico dos dados determinam a hipótese adotada e, consequentemente, a tabela a ser usada unilateral ou bilateral.

b) Determinação de Z

A distribuição normal é utilizada em tratamento estatístico para grandes amostras, geralmente com n > 30. Os valores das áreas sob a curvas são em função de z, que é obtido pela seguinte equação:

Para a população:

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

Onde:

x — é a variável

μ — é a média

σ — Desvio padrão

No caso de amostras o valor de z deve ser substituído por t (distribuição student). Conforme equação à seguir:

Para amostras

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

Onde:

$\bar{x}$ — Média da amostra

s — Desvio padrão da amostra

n — número de amostras

μ — média da população

Exemplo: Numa cidade de 2000 habitantes a média do peso da população é de 70 quilos e o desvio padrão é igual a 5 quilos. Escolhendo um indivíduo aletoriamente, com peso de 76 quilos, qual a probabilidade com 95% de certeza de sortear uma pessoa com peso superior a 76 quilos?

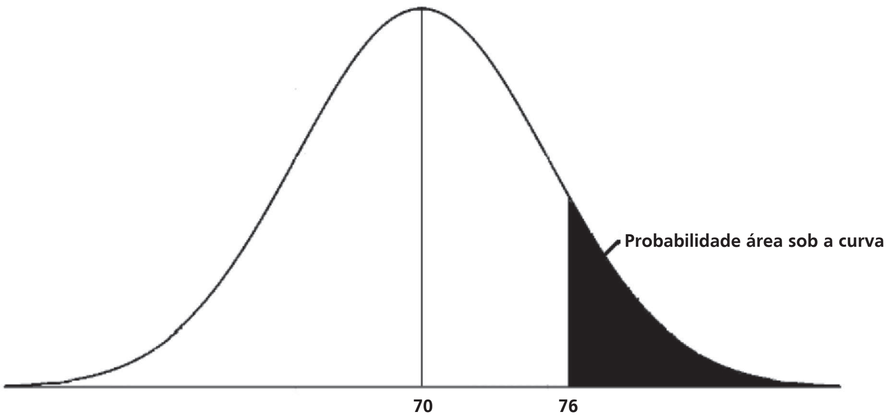
De acordo com o exposto temos:

— Média dos pesos da população (μ) é igual a 70 quilos;

— Desvio padrão (σ) é igual a 5 quilos;

— Peso do indivíduo selecionado (x) = 76 quilos.

De acordo com o teste de hipótese a distribuição é unilateral, pois a pesquisa visa verificar a probabilidade de certa pessoa com peso maior que a média (70 quilos).

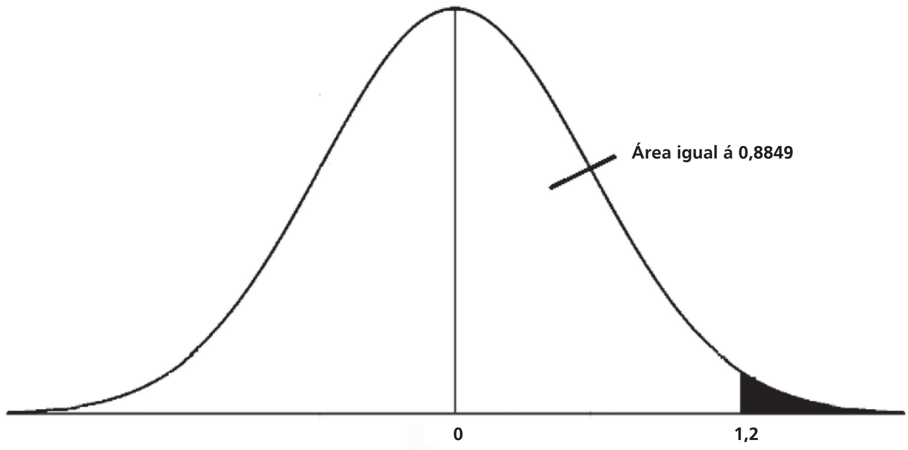


Para encontrar a probabilidade temos que calcular o Z. Conforme equação abaixo:

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

$$z = \frac{76 - 70}{5} = 1,2$$

Na tabela 2, no final deste capítulo, deve-se entrar com o valor de Z calculado e obter a área sob a curva. Nesse caso, para o valor de Z igual a 1,2 a área sobre a curva é igual a 0,8849 conforme mostra a curva abaixo:



Considerando que a área total sob a curva é igual a 1, o valor da probabilidade desejada será igual a: 1,0 - 0,8849 = 0,1151. Esse valor multiplicado por 100 é igual a 11,51%. Assim a probabilidade de selecionar uma pessoa com peso superior a 76 quilos é de 11,51%.

No caso de amostras, a determinação da probabilidade é obtida de maneira similar. Assim, considerando o exemplo anterior, ou seja, cidade de 2000 habitantes a média do peso da população é de 70 quilos. Em uma amostragem de 40 pessoas o peso médio foi de 71 quilos e desvio padrão da amostra igual a 4,95. Assim, qual a probabilidade do peso superar 70 quilos?

Temos os seguintes dados:

— Média dos pesos da população (μ) é igual a 70 quilos;

— tamanho da amostra = 40 pessoas

— média da amostra $\bar{x}$ = 71 quilos

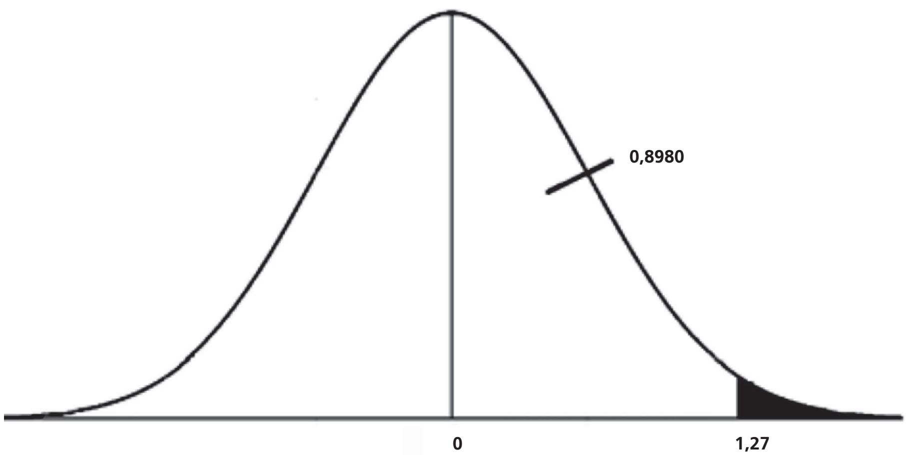
Para utilização do Z em dados amostrais, substitui-se o desvio padrão da população pelo erro padrão da amostra E_p que é igual a $\frac{s}{\sqrt{n}}$, na equação teremos:

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

Nesse caso, o valor de Z é igual a:

$$z = \frac{71 - 70}{\frac{4,95}{\sqrt{40}}} = 1,27$$

Entrando na tabela 2 com valor de Z igual a 1,27, a área sob a curva é igual 0,8980 conforme mostra a curva a seguir:



Considerando que a área total sob a curva é igual a 1 o valor da probabilidade desejada será igual a: 1,0 - 0,8980 = 0,1020. Esse valor multiplicado por 100 é igual a 10,20%. Assim a probabilidade da média dos pesos superar 71 quilos é de 10,20%.

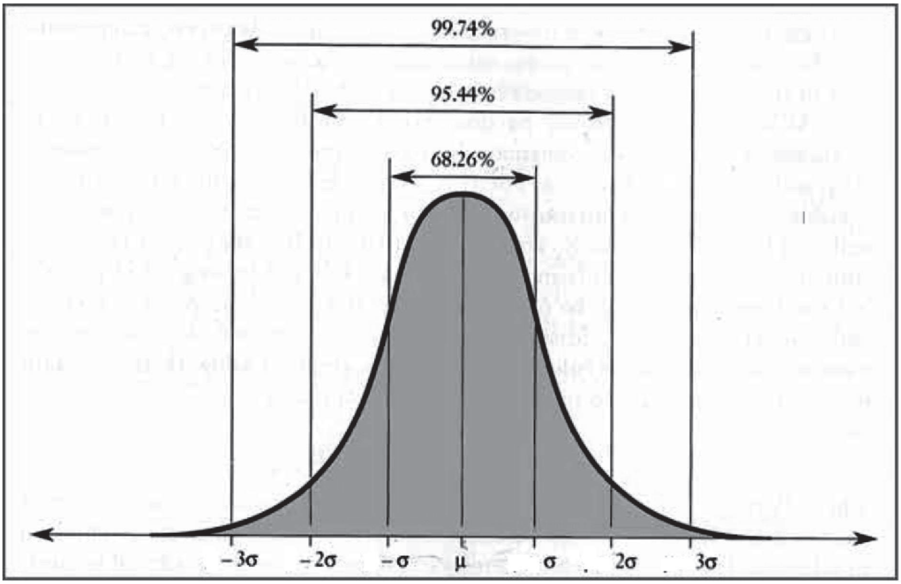
c) considerações sobre a curva normal

Considerando a média μ e o desvio padrão σ a figura abaixo mostra as probabilidades:

— Desvio padrão igual 1σ a probabilidade é aproximadamente 68,26%

— Desvio padrão igual 2σ a probabilidade é aproximadamente 95,44%

— Desvio padrão igual 3σ a probabilidade é aproximadamente igual 99,74%



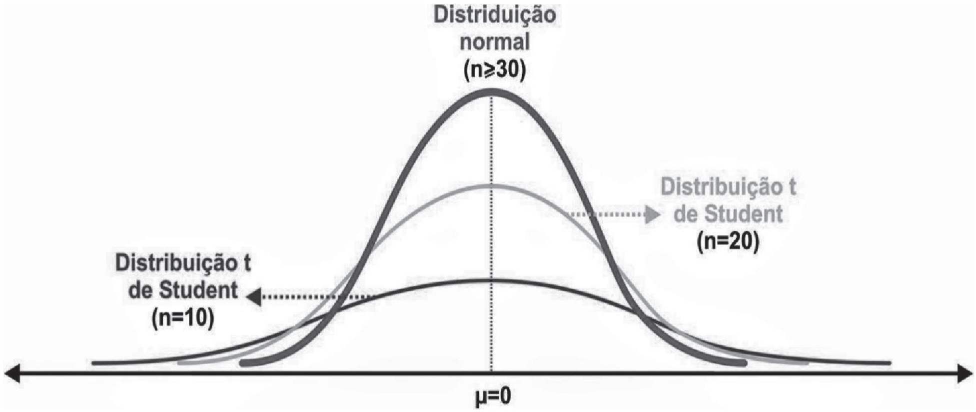
Exemplo:

Suponhamos que a média de pessoas que acessam diariamente determinado prédio comercial é igual a 200 e o desvio padrão igual a 20, a probabilidade é igual a 68,26%. Assim, existe 68,26% de chance de 180 a 220 pessoas acessarem o prédio por dia.

Considerando o dobro do desvio padrão (2 x 20 = 40), a probabilidade será de 95,44%. Assim, existe 95,44% de chance de 160 a 240 pessoas acessarem o prédio por dia.

2.2. Distribuição Student

A distribuição Student é similar à normal. Os dados coletados também se distribuem em forma de um sino, porém a sua cauda mais é longa. A figura 5 mostra a comparação das curvas da distribuição Student em relação à normal. Quanto maior o tamanho da amostra, mais a distribuição Student se aproxima da normal.



A distribuição Student ou t é utilizada nas seguintes condições:

— não se conhece o desvio padrão da população (σ) e sua distribuição é normal.

— o tamanho da amostra é pequeno, ou seja, inferior a 30 ($n \leq 30$);

Quanto maior n (número de amostras), maior a aproximação em relação à distribuição normal. Para n > 30 podemos utilizar distribuição normal com valores críticos 'z'.

Nessa distribuição, o z é substituído por t que é determinado em função do grau de liberdade. Na avaliação dos riscos ambientais, na maioria das vezes, o tamanho da amostra é pequeno ($n \leq 30$), por essa razão a distribuição de Student é muito utilizada.

O valor é calculado por meio da seguinte equação:

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

Onde:

$\bar{x}$ — média da amostra

μ — média da população

s — desvio padrão da amostra

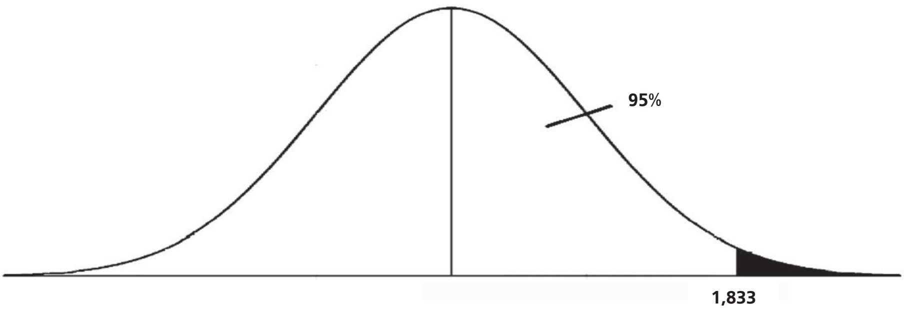
n — tamanho da amostra

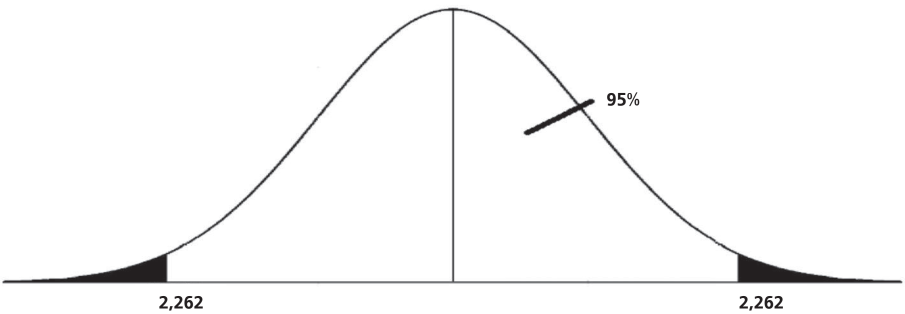
t — é obtido na tabela 3 no final deste capítulo, devendo entrar nessa tabela considerando o grau de liberdade.

GL — grau de liberdade igual a: GL = n - 1

Com o grau de liberdade e a confiança desejada utiliza-se a tabela 3 no final deste capítulo, para a determinação do valor da área sobre a curva. A tabela 3 fornece os valores para teste bilateral e unilateral.

Exemplo: considerando o tamanho da amostra igual a 10, os valores de t que delimitam a área sob a curva com 95% de confiança serão: com o grau de liberdade (10 - 1), para nível de confiança de 95% (5% na tabela 3) encontramos o valor de t igual a 1,8331 para teste unilateral, e 2,2622 para teste bilateral, conforme mostram as curvas a seguir.





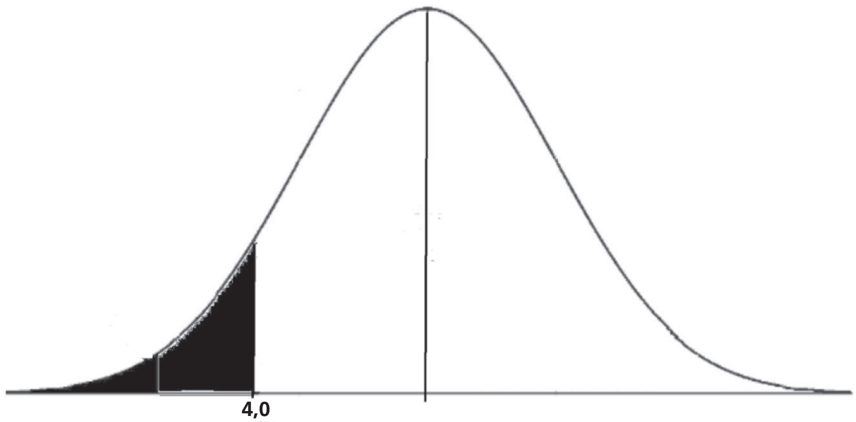
Exemplo:

Foram coletadas 6 (seis) amostras de dióxido de enxofre obtendo-se as seguintes concentrações:

Amostra	Concentração (ppm)
1	2,5
2	4,0
3	5,0
4	3,0
5	4,0
6	2,9

Considerando o limite de tolerância dessa substância igual 4,0 ppm (NR-15, anexo 11), calcular a probabilidade da concentração ser inferior a esse limite.

Nesse caso, graficamente teremos:



A probabilidade da concentração ser inferior a 4 ppm corresponde a área hachurada. Para encontrar essa área é necessário calcular o t e recorrer a tabela da distribuição student.

a) Concentração média do dióxido de enxofre é igual a:

$$\bar{x} = \frac{2,5 + 4,0 + 5,0 + 3,0 + 4,0 + 2,9}{6} = 3,6\text{ppm}$$

b) O desvio padrão da amostra é igual a:

$$s = \sqrt{\frac{(2,5 - 3,6)^2 + (4,0 - 3,6)^2 + (5,0 - 3,6)^4 + (3,0 - 3,6)^2 + (4,0 - 3,6)^2 + (2,9 - 3,6)^2}{6 - 1}} = 0,85 \text{ ppm}$$

c) Cálculo de t

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} \quad \text{substituindo os valores teremos:} \quad t = \frac{3,6 - 4,0}{\frac{0,85}{\sqrt{6}}} = -1,14$$

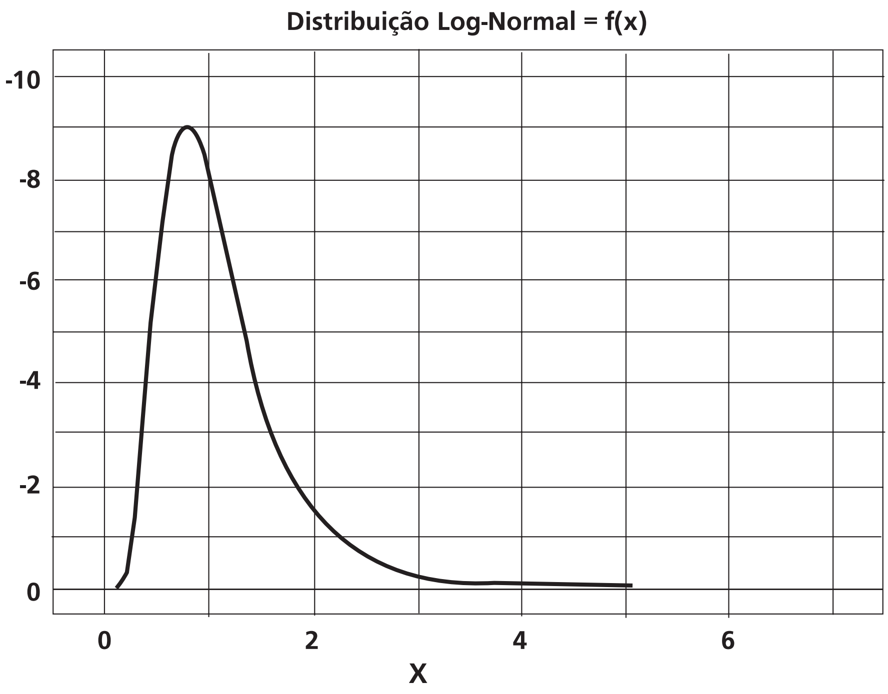
Na tabela t-Student, com o grau de liberdade igual a 5 (n-1), o valor de t que mais de aproxima do valor encontrado é 1,15. Sendo assim, a probabilidade que corresponde a esse valor é aproximadamente 15% (teste unilateral). Portanto, a probabilidade do limite ser superado é de 15%. Nas tabelas da distribuição Student, os valores possuem grandes intervalos. Sendo assim, é importante determinar a probabilidade em calculadoras eletrônicas ou no Excel.

gl	Teste Unilateral								
	15%	**10%**	**5%**	**2,5%**	**2%**	**1%**	**0,5%**	**0,1%**	**0,05%**
	Teste Bilateral								
	30%	**20%**	**10%**	**5%**	**4%**	**2%**	**1%**	**0,2%**	**0,1%**
1	1,9626	3,0777	6,3137	12,7062	15,8945	31,8210	63,6559	318,2888	636,5776
2	1,3862	1,8856	2,9200	4,3027	4,8487	6,9645	9,9250	22,3285	31,5998
3	1,2498	1,6377	2,3534	3,1824	3,4819	4,5407	5,8408	10,2143	12,9244
4	1,1896	1,5332	2,1318	2,7765	2,9985	3,7469	4,6041	7,1729	8,6101
5	1,1558	1,4759	**2,0150**	2,5706	2,7565	3,3649	4,0321	5,8935	6,8685

2.3. Distribuição Logaritmo Normal

Na distribuição Log-Normal é aplicado na variável o logaritmo decimal ou neperiano. Assim, a variável x tem distribuição log-normal quando y = logx tem distribuição normal. A figura a seguir mostra a distribuição log-normal.

Como o nome sugere o logaritmo de uma variável com distribuição Log-Normal com parâmetros μ e σ tem uma distribuição Normal com média μ e desvio padrão σ. Esta relação significa que dados provenientes de uma distribuição log-Normal podem ser analisados segundo uma distribuição normal se trabalharmos com o logaritmo dos dados ao invés dos valores originais.



a) Média da amostra

$$\bar{y} = \frac{\Sigma \log x}{n}$$

Onde:

x — variável

n — tamanho

b) Desvio padrão da amostra

$$s = \sqrt{\frac{(x - \bar{y})^2}{n - 1}}$$

Onde:

x — variável

$\bar{y}$ — média da amostra em log-normal

n — tamanho

c) Probabilidade

Para o cálculo da probabilidade determina-se o valor de z e na tabela 2 encontra-se a área sob a curva.

d) Limite de confiança

Com os valores da média, do desvio padrão, e do tamanho da amostra determina-se o Limite de confiança utilizando a tabela normal ou Student. Na estratégia de avaliação dos riscos ambientais, a distribuição log-normal é muito utilizada como será visto posteriormente.

3. LIMITE DE CONFIANÇA OU INTERVALO DE CONFIANÇA

A média da amostra dificilmente é igual à média real ou média populacional. Por esse motivo, na estatística dos dados é calculada a margem de erro ou erro padrão em relação à média e, consequentemente, o Limite de Confiança (LC). O Erro Padrão ou margem de erro é dado pela seguinte equação:

$$EP = z \cdot \frac{\sigma}{\sqrt{n}}$$

Onde:

EP — erro padrão ou margem de erro

σ — desvio padrão da população

z — valor da tabela normal

n — tamanho da amostra

Quando o desvio padrão da população (σ) é desconhecido e a amostra é pequena ($n \leq 30$) utiliza-se a distribuição Student. Nesse caso, o Erro Padrão é calculado pela equação:

$$EP = t \cdot \frac{s}{\sqrt{n}}$$

Onde:

t — o valor de t crítico é obtido na tabela 3 (distribuição Student)

s — desvio padrão da amostra

n — tamanho da amostra

Como visto anteriormente, na tabela distribuição Student temos que usar o grau de liberdade (n-1). De acordo com o NIOSH (National Institute for Occupational Safety and Health, 1977), em higiene ocupacional utiliza-se o limite de confiança unilateral com nível de confiança de 95%, sendo t = 1,645.

Com os dados do erro padrão ou margem de erro, determina-se os intervalos de confiança da média amostral, conforme as equações:

$$LIC = \bar{x} - EP \text{ ou } LIC = \bar{x} - t\ \frac{s}{\sqrt{n}}$$

$$LSC = \bar{x} + EP \text{ ou } LSC = \bar{x} + t\ \frac{s}{\sqrt{n}}$$

LIC — Limite Inferior de Confiança

LSC — Limite Superior de Confiança

$\bar{x}$ — média da amostra

t — obtido na tabela 3 para nível de significância de 5% ou confiança 95%. Lembrando que deve-se entrar nessa tabela com o grau de liberdade (n - 1)

Exemplo:

Determinar o Limite de Confiança de 95% da População para a média e população desconhecida, sabendo que a distribuição dos dados se aproxima da curva normal de acordo com os seguintes dados:

$\bar{x}$ (média amostral) = 20

S (desvio padrão da amostra) = 3,0

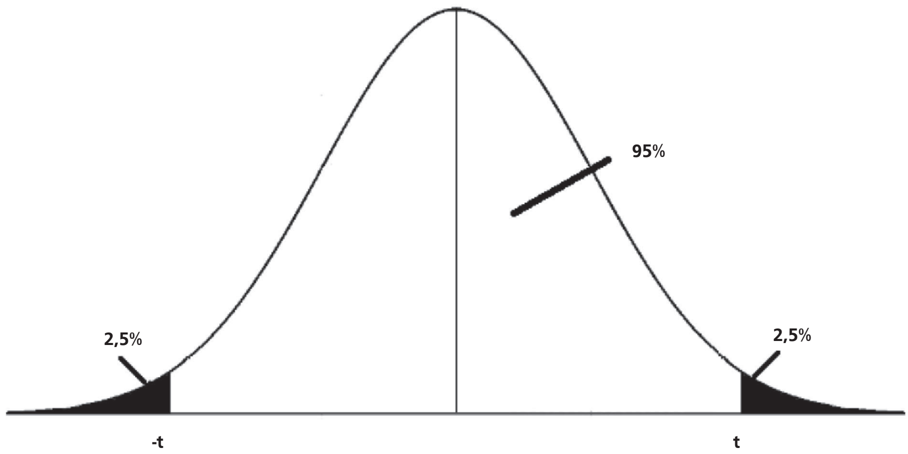
n = 10

Nesse caso, vamos utilizar a distribuição Student, pois o desvio padrão da população é desconhecido e o tamanho da amostra n ≤ 30. Sendo assim:

$$LC = \bar{x} \pm t \frac{s}{\sqrt{n}}$$

Primeiramente, determina-se t na tabela 3. Nessa tabela, deve entrar com o valor do grau de liberdade (n - 1) e o valor da significância para a confiança de 95%. Ou seja, 100 - 95 = 5%. Como a curva é simétrica temos 2,5% para cada lado na bilateral.

Nota: Como visto anteriormente, o NIOSH utiliza o limite de confiança unilateral.



Assim, com o valor de (10 – 1 = 9) e significância 5% bilateral ou 2,5% unilateral, na tabela 3, o valor de t é igual a 2,26 para distribuição. Com esses dados os Limites de confiança são:

$$LIC = 20 - 2{,}26 \frac{3}{\sqrt{10}} = 17{,}86$$

$$LSC = 20 + 2{,}26 \frac{3}{\sqrt{10}} = 22{,}14$$

Portanto, a verdadeira média está compreendida no intervalo de 17,86 a 22,14 com 95% de certeza.

Na avaliação dos riscos ambientais, a tomada de decisão é baseada na teoria estatística e testes de hipóteses, que estão intimamente ligados ao conceito dos intervalos de confiança (isto é, o cálculo do intervalo de confiança que se espera conter a exposição média real). Como será visto mais adiante algumas normas consideram os limites de confiança bilateral, enquanto outros determinam os limites na distribuição unilateral.