

USO DA ESTATÍSTICA COMO FERRAMENTA PARA DETERMINAR A PERIODICIDADE DAS CALIBRAÇÕES: EXEMPLOS.

Valter Quilici¹

¹ CDTN, Belo Horizonte, Brasil. quilici@cdtn.br

Resumo: Os testes de hipótese são ferramentas bastante conhecidas e aplicáveis a praticamente todas as áreas da Ciência.

Todo metrologista deve executar seu ensaio com repetitividade ou reprodutibilidade. Preocupação essa expressa nos itens 5.2.1 e 5.2.2 da ABNT NBR ISO / IEC 17025:2005. [Metas referentes à competência do metrologista]

Todo equipamento-chave para uma medição deve ter confiabilidade metrológica. Essa preocupação está expressa nos itens 5.5.2; 5.5.10; 5.6.1 e 5.6.6.3 da ABNT NBR ISO / IEC 17025:2005 [Metas referentes a equipamento: exatidão, calibração, confiança no desempenho (verificações intermediárias), programa de calibração].

Há um método simples e de baixo custo que serve tanto para garantir a confiabilidade metrológica do desempenho de pessoas, equipamentos ou MRCs quanto estabelecer a periodicidade das calibrações: ele é o emprego dos testes de hipótese [Cochran & Snedecor] associado a uma carta de controle.

Neste trabalho busca-se, de forma didática e por meio de exemplos genéricos, mostrar esse procedimento.

Palavras chave: Confiabilidade metrológica, ensaio, instrumentação e estatística aplicada.

1. ESTATÍSTICA E METROLOGIA

Instituições com renomada tradição nos aspectos de rastreabilidade metrológica e exatidão de medida empregam desde longa data *carta de controle* como ferramenta para demonstrar a confiabilidade metrológica. Por exemplo, o NIST (antigo NBS) dos Estados Unidos faz uso dessa carta para pilhas padrão de Weston desde 1850 [6].

Também o setor industrial produtivo usa essas cartas para manter sua produção dentro dos limites de confiabilidade. Num mundo globalizado onde a produção deve ter como base o mais baixo custo aliado à qualidade, as cartas de controle assumem vital importância. À medida que a metrologia garante, por amostragem, que um parafuso está dentro dos limites de tolerância de projeto, isto significa que se esse parafuso for produzido, por exemplo, na China ele terá grande probabilidade de rosquear numa porca especificada para ele e produzida no Paraguai. Muitos anos de uso garantem a qualidade da produção industrial.

Portanto a estatística é companheira de longa data da metrologia, em alguns setores. No entanto, os metrologistas, por vários motivos, continuam perdendo muito do que a estatística pode proporcionar como ferramenta básica num laboratório. Por exemplo, um metrologista com alguma

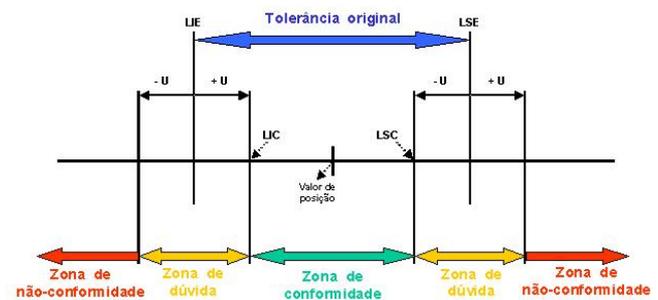
habilidade em estatística pode usar os testes de hipóteses para definir o “intervalo entre calibrações” sucessivas para cada um dos equipamentos do seu laboratório. Este artigo objetiva fornecer alguns exemplos dessa utilização.

2. ALGUMAS FERRAMENTAS ESTATÍSTICAS

A seguir são apresentadas algumas ferramentas estatísticas que são fundamentais para este artigo.

2.1. Carta de controle

Quando realizamos uma medição de uma grandeza repetidas vezes ou quando fazemos o projeto de um equipamento destinado à medição, para cada *valor da grandeza* $[\mu]$, é definido ou especificado um *intervalo de tolerância*, dado por: $(\mu \pm ku)$, onde ku é um valor determinado tal que, nesse intervalo, existe uma probabilidade de $P\%$ de que o *valor verdadeiro da grandeza medida* esteja lá. Portanto, *tolerância* é o intervalo de variabilidade de aceitação do valor de uma grandeza. Esse é o princípio básico para estabelecimento de uma carta de controle [Figura 1].



Legenda:

LIE ≡ limite inferior da especificação [projeto]

LSE ≡ limite superior da especificação [projeto]

LIC ≡ limite inferior de controle [medição]

LSC ≡ limite superior de controle [medição]

U ≡ incerteza do processo de medição usado para verificar a conformidade [ku]

Figura 1: Exemplo de carta de controle básica.

2.2. O teste de Cochran (Homogeneidade de variâncias a uma certa probabilidade)



Este é um teste de hipóteses desenvolvido por William Gemmill Cochran (1909 – 1980) e baseia-se na distribuição F [desenvolvida por Sir Ronald Aylmer Fischer (1890 – 1962)].

Fundamento:

A idéia principal é comparar a maior variância de uma população de dados com a soma de todas as variâncias dessa população, para uma certa probabilidade. A aplicabilidade é para uma população maior ou igual a 3 conjuntos.

O teste estatístico é:

$$C = \frac{\text{maior variância}}{\text{soma de todas as variâncias}} = \frac{s_{\max}^2}{\sum_{i=1}^q s_i^2} \quad [\text{equação 1}]$$

onde q = número de conjuntos de uma população = número de conjuntos.

Para efetuá-lo devem ser realizados os seguintes passos:

Passo 1: Se cada um dos q conjuntos tem n elementos com v graus de liberdade, usando a equação acima você pode calcular o valor de C . O número de graus de liberdade (v) deve ser o mesmo.

Passo 2: Através da tabela de Cochran (Tabela de valores de C) verifique qual é o valor de C tabelado para v graus de liberdade. A tabela de valores mais comumente usada é a para a probabilidade $p = 5\%$.

Passo 3: Constituir as hipóteses. Por exemplo:

Se $C(\text{calculado}) > C(\text{tabelado}) \Rightarrow$ Não há homogeneidade.

Se $C(\text{calculado}) < C(\text{tabelado}) \Rightarrow$ Há homogeneidade.

Neste teste todos os conjuntos de dados devem ter o mesmo “grau de liberdade”.

2.3. O teste de Snedecor (Homogeneidade entre 2 variâncias a uma certa probabilidade)

Fundamentos:



Este é um teste de hipóteses desenvolvido por George Waddell Snedecor (1881 – 1974) e baseia-se na distribuição F [desenvolvida por Sir Ronald Aylmer Fischer (1890 – 1962)].

A idéia principal é comparar a variância de uma população de dados, duas a duas, para uma certa probabilidade. Isto é feito verificando se a variância de 2 conjuntos da população são iguais para um desejado nível de confiança ($95\% \approx t$ student).

O teste estatístico é:

$$F_{\text{Calc}} = \frac{s_1^2}{s_2^2} \quad \text{sendo } s_1^2 > s_2^2 \quad [\text{equação 2}]$$

Onde o *denominador* é o conjunto da população considerado como o *padrão*.

Passo 1: Se o conjunto q_1 tem n_1 elementos com v_1 graus de liberdade e o conjunto q_2 tem n_2 elementos com v_2 graus de liberdade, através da fórmula acima calcula-se o valor de F .

Passo 2: Através da tabela de Snedecor (Tabela de valores de F) verifique qual é o valor de F tabelado para v_1 graus de liberdade do numerador e v_2 graus de liberdade do denominador. A tabela de valores mais comumente usada é a para a probabilidade $p = 5\%$.

Passo 3: Constituir as hipóteses. Por exemplo:

Se $F(\text{tabelado}) > F(\text{calculado}) \Rightarrow$ Há homogeneidade.

Se $F(\text{tabelado}) < F(\text{calculado}) \Rightarrow$ Não há homogeneidade.

É importante observar que, ao contrário do teste de Cochran, no teste de Snedecor não é necessário que os dois conjuntos tenham o mesmo “grau de liberdade”.

3. ALGUMAS DEFINIÇÕES

Além das definições mencionadas na bibliografia para efeito deste artigo emprega-se:

Instrumentos de medição iguais: são os de mesmo fabricante, mesmo modelo, podendo ser de mesma série ou não.

4. APLICAÇÕES

4.1. Caso 1: o laboratório só tem um instrumento de medição.

Considerando a Figura 1, fazendo nela uma rotação de 90° à esquerda obtém-se como eixo das ordenadas o eixo original. Colocando um eixo das abcissas para a grandeza “Data” obtém-se a “carta de controle” mais comum usada nos nossos laboratórios [Figura 2]

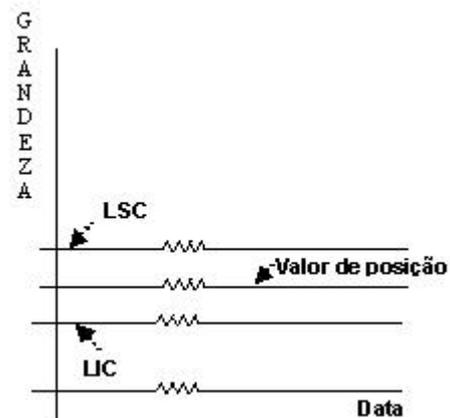


Figura 2: Carta de controle tipo Grandeza x Data

Aproveitando a idéia de “verificação intermediária” apresentado em 5.5.10 da ABNT NBR ISO IEC 17025:2005 [5] pode-se colocar na carta de controle os valores da grandeza medidos quando dessa verificação. Enquanto esses valores se mantiverem no intervalo entre LIC e LSC [$LIC < G < LSC$] significa que o equipamento se mantém dentro da faixa de operação especificada pelo fabricante no manual desse equipamento. Quando na verificação o valor estiver fora, para cima ou para baixo, chegou a hora de fazer a calibração desse equipamento.

Quando o equipamento voltar da calibração, no mesmo gráfico, torna-se a colocar os valores da nova verificação intermediária. Quando o valor estiver fora outra vez, terá chegado a hora de mandar calibrar novamente esse equipamento.

Com no mínimo 3 rodadas de carta de controle pode-se estabelecer com segurança a periodicidade da calibração do equipamento.

Por analogia pode-se usar raciocínio idêntico para material de referência certificado [MRC] quando o laboratório só tem um.

4.2. Caso 2: o laboratório tem dois instrumentos de medição “iguais”.

Neste caso há um melhor uso da estatística uma vez que o metrologista pode utilizar o teste de Snedecor. Por meio do teste de Snedecor esse metrologista vai descobrir se os dois equipamentos são compatíveis, são homogêneos. Para isso é preciso seguir os passos apresentados em 2.3.

A aplicação típica do teste de Snedecor será mostrada no exemplo a seguir.

Exemplo 1: O mesmo metrologista realizou uma série de medidas de uma certa grandeza utilizando o equipamento A e o equipamento B. Deve-se salientar que isso é feito em condições de reprodutibilidade. Como resultado esse metrologista apresentou os dados da tabela 1.

Tabela 1: Dados para o teste de Snedecor

| Número da Medida | Equipamento | |
|------------------|--------------|--------------|
| | A | B |
| 01 | 51 | 50 |
| 02 | 52 | 56 |
| 03 | 52 | 55 |
| 04 | 51 | 51 |
| 05 | 51 | ☆ |
| 06 | 52 | ☆ |
| Média | 52 [51,500]* | 53 [53,000]* |
| Desvio Padrão | 0,548 | 2,944 |
| Variância | 0,300 | 8,667 |

Legenda: ☆ ≡ nenhuma medida
* o valor entre colchetes é o resultado antes do arredondamento

Usando os fundamentos de 2.3 e a equação 2, produz –se o F_{Calc} :

$$F_{Calc} = 28,89$$

Quando o extrato da tabela de Snedecor [Tabela 2] é consultado, para este exemplo, pode-se concluir que:

$$F_{Tab} = 5,41$$

Das hipótese estabelecidas em 2.3 pode-se concluir que, como $F_{Tab} < F_{Calc}$, então não há homogeneidade entre os resultados dos dois equipamentos para uma mesma grandeza medida em condições de reprodutibilidade.

Um dos dois equipamentos não está conforme. Note que o equipamento com menor variância é considerado o

equipamento que mede correto. É o equipamento em que o metrologista deposita credibilidade.

Tabela 2: Extrato dos valores da tabela de Snedecor [Pontos percentuais da Distribuição F para $p = 5\%$]

| | | v_1 ≡ “graus de liberdade” do numerador | | | | | | | |
|---|---|---|------|------|------|------|------|------|------|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| v_2 ≡ “graus de liberdade” do denominador | 1 | 161 | 200 | 216 | 225 | 230 | 234 | 237 | 239 |
| | 2 | 18,5 | 19,0 | 19,2 | 19,2 | 19,3 | 19,3 | 19,4 | 19,4 |
| | 3 | 10,1 | 9,55 | 9,28 | 9,12 | 9,01 | 8,94 | 8,89 | 8,85 |
| | 4 | 7,71 | 6,94 | 6,59 | 6,39 | 6,26 | 6,16 | 6,09 | 6,04 |
| | 5 | 6,61 | 5,79 | 5,41 | 5,19 | 5,05 | 4,95 | 4,88 | 4,82 |
| | 6 | 5,99 | 5,14 | 4,76 | 4,53 | 4,39 | 4,28 | 4,21 | 4,15 |
| | 7 | 5,59 | 4,74 | 4,35 | 4,12 | 3,97 | 3,87 | 3,79 | 3,73 |
| | 8 | 5,32 | 4,46 | 4,07 | 3,84 | 3,69 | 3,58 | 3,50 | 3,44 |

Fonte: **Biometrika**, V 33 (1943), p 73 – 88.

Pode-se, também, olhar outros aspectos.

Quando o mensurando é o valor de um padrão fica fácil ver “qual deles é o equipamento que mede correto”.

Quando não se dispõe de padrão usa-se a “carta de controle” para chegar a conclusão de “qual deles é o equipamento que mede correto”.

É claro que o exemplo é didático para ficar mais fácil o entendimento do uso do teste de hipóteses.

4.3. Caso 3: o laboratório tem mais de dois instrumento de medição “iguais”.

Neste caso há um melhor uso da estatística uma vez que o metrologista pode utilizar o teste de Cochran. Por meio do teste de Cochran esse metrologista vai descobrir se os três (ou mais) equipamentos são compatíveis, são homogêneos.

A aplicação típica do teste de Cochran será mostrada no exemplo 2, à seguir.

Exemplo 2: O mesmo metrologista realizou uma série de medidas de uma certa grandeza utilizando os equipamentos A, B, C e D. Os resultado obtidos por esse metrologista são os da tabela 3.

Usando os fundamentos de 2.2 e a equação 1, produz –se o C_{Calc} :

$$C_{Calc} = 0,712$$

Quando o extrato da tabela de Cochran [Tabela 4] é consultado, para este exemplo [$q = 4$ e $v = 5$], pode-se concluir que:

$$C_{Tab} = 0,690$$

Das hipótese estabelecidas em 2.2 pode-se concluir que, como $C_{Calc} > C_{Tab}$, então, para uma mesma grandeza medida

em condições de reprodutibilidade, não há homogeneidade entre os resultados do equipamento que apresentou a maior variância e os outros.

Isso pode ser interpretado como uma indicação de que está na hora de “calibrar” esse equipamento.

Tabela 3: Dados para o teste de Cochran

| Número da Medida | Equipamento | | | |
|------------------|--------------|--------------|--------------|--------------|
| | A | B | C | D |
| 01 | 71 | 72 | 70 | 71 |
| 02 | 72 | 73 | 74 | 72 |
| 03 | 72 | 74 | 75 | 72 |
| 04 | 71 | 72 | 71 | 73 |
| 05 | 72 | 73 | 76 | 72 |
| 06 | 71 | 74 | 73 | 74 |
| Média | 72 [71,500]* | 73 [73,000]* | 73 [73,167]* | 72 [72,333]* |
| Desvio Padrão | 0,548 | 0,894 | 2,317 | 1,033 |
| Variância | 0,300 | 0,800 | 5,367 | 1,067 |

* o valor entre colchetes é o resultado antes do arredondamento

Tabela 4: Extrato dos valores da tabela de Cochran [Pontos percentuais da Distribuição F para p = 5%]

| Número de conjuntos q | v ≡ “graus de liberdade” de cada conjunto | | | | |
|-----------------------|---|-------|-------|-------|-------|
| | 1 | 2 | 3 | 4 | 5 |
| 2 | - | 0,975 | 0,939 | 0,906 | 0,877 |
| 3 | 0,967 | 0,871 | 0,798 | 0,746 | 0,707 |
| 4 | 0,906 | 0,768 | 0,684 | 0,629 | 0,690 |
| 5 | 0,841 | 0,684 | 0,598 | 0,544 | 0,506 |
| 6 | 0,781 | 0,616 | 0,532 | 0,480 | 0,445 |
| 7 | 0,727 | 0,561 | 0,480 | 0,431 | 0,397 |
| 8 | 0,680 | 0,516 | 0,438 | 0,391 | 0,300 |

Fonte: *Annals of Eugenics* (1941). p50

Ainda continuando este raciocínio, resta a dúvida: - Como está o desempenho dos outros equipamentos (o A, o B e o D)?

Para avaliar esse aspecto basta eliminar na Tabela 3 os dados do equipamento C. Aplica-se novamente o teste de Cochran. Dessa aplicação resulta o valor:

$$C_{Calc} = 0,492$$

Quando o extrato da tabela de Cochran [Tabela 4] é consultado, para este caso [q = 3 e v = 5], pode-se concluir que:

$$C_{Tab} = 0,707$$

Das hipótese estabelecidas em 2.2 pode-se concluir que, como $C_{Calc} < C_{Tab}$, então, para uma mesma grandeza medida em condições de reprodutibilidade, há homogeneidade entre os resultados do equipamento que apresentou a maior variância e os outros. Assim sendo os equipamentos A, B e D apresentam resultados coerentes e não precisam ser “calibrados”.

Outra vez cabe lembrar que o exemplo tem propósito didático. No dia a dia de um laboratório que busca excelência, os valores dos desvios padrão são muito próximos. O teste de Cochran, mesmo para esse laboratório, mantém sua eficácia.

5. CONCLUSÃO

O uso da estatística baseado em carta de controle e nos testes de Snedecor e Cochran oferece ferramentas que vão tornar mais fácil o dia a dia dos laboratórios que buscam uma forma científica de estabelecer planos de calibração de equipamentos. As ferramentas apresentadas permitem estabelecer a periodicidade da calibração dos equipamentos independente da grandeza que esse equipamento mede. Lembre-se que todos os exemplos apresentados neste artigo são genéricos e não fazem referência à grandeza objeto da medição do equipamento [mensurando].

AGRADECIMENTOS

Agradeço ao Prof. Dr. João Mário Andrade Pinto, amigo e colega de trabalho a disponibilidade e a boa vontade demonstrados na discussão de “temas” em estatística. À Rosane Rodrigues Fraga e Paulo Gomide pela revisão gramatical do texto.

À CNEN e ao CDTN agradeço o apoio, o incentivo e a oportunidade de crescimento profissional que são proporcionados aos profissionais que integram seus recursos humanos.

REFERÊNCIAS

- [1] ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **ABNT NBR ISO 9000**: sistemas de gestão da qualidade - Fundamentos e vocabulário. Rio de Janeiro: 2005.
- [2] ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **ABNT ISO/IEC Guia 2**: termos gerais e suas definições referentes à normalização e atividades correlatas. Rio de Janeiro: 1993.
- [3] INMETRO. **Vocabulário internacional de termos fundamentais e gerais de metrologia**. 3ª ed. Brasília: 2003.
- [4] ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **ABNT NBR ISO 9001**: sistemas de gestão da qualidade - Requisitos. Rio de Janeiro: 2000.
- [5] ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **ABNT NBR ISO / IEC 17025**: requisitos gerais para a competência de laboratórios de ensaio e calibração. Rio de Janeiro: 2005.
- [6] FLUKE. **Calibration: philosophy in practice**. 2nd edition. Everett - Washington: Fluke, 1994. 1 V.
- [7] HOLMAN, J. P., **Experimental methods for engineers**. 7th ed. New York: McGrawHill Book, July 2000. 720p. ISBN 0 - 07 - 366055 - 8.
- [8] COSTA, S.F. **Introdução ilustrada à estatística**. 4ª ed. São Paulo: Harbra, 2005. 399p
- [9] LEVIN, J. **Estatística aplicada a ciências humanas**. 2ª ed. São Paulo: Harbra, 1987. 392p.
- [10] MERRINGTON, M.; THOMPSON, C.M.; PEARSON, E.S. Tables of percentage points of the inverted beta(f) distribution. **Biometrika**, v.33 p 73 – 88, 1943.

[11] COCHRAN, W.G. The distribution of the largest of a set estimated variances as a fraction of their total. **Annals of Eugenics**, v11 p 47 – 51, 1941.