

Agência Nacional de Vigilância Sanitária



GUIA PARA TRATAMENTO ESTATÍSTICO DA VALIDAÇÃO ANALÍTICA

VIGENTE A PARTIR DE 12/09/2017

Início do período de contribuições: 13/09/2017

Fim do período de contribuições: 12/03/2018



GUIA PARA TRATAMENTO ESTATÍSTICO DA VALIDAÇÃO ANALÍTICA

Este Guia expressa o entendimento da Anvisa sobre as melhores práticas com relação a procedimentos, rotinas e métodos considerados adequados ao cumprimento de requisitos técnicos ou administrativos exigidos pela legislação. Não confere ou cria novas obrigações, devendo ser utilizado por agentes públicos e privados como referência para cumprimento legislativo.

Abordagens alternativas são possíveis, de modo que sua inobservância não caracteriza infração sanitária, nem constitui motivo para indeferimento de petições, desde que atendidos os requisitos exigidos pela legislação, ainda que por meio diverso daquele previsto nesta recomendação.

As recomendações contidas neste Guia produzem efeitos a partir da data de sua publicação no Portal da Anvisa e ficam sujeitas ao recebimento de sugestões da sociedade por meio de formulário eletrônico, disponível em http://formsus.datasus.gov.br/site/formulario.php?id_aplicacao=34184.

As contribuições* recebidas serão avaliadas e poderão subsidiar a revisão do Guia e a consequente publicação de uma nova versão. Independente da decisão da área, será publicada análise geral das contribuições e racional que justifique a revisão ou não do Guia.

*A fim de garantir maior transparência ao processo de elaboração dos instrumentos regulatórios editados pela Anvisa, esclarecemos que os nomes dos responsáveis pelas contribuições (pessoas físicas e jurídicas) são considerados informações públicas e serão disponibilizados de forma irrestrita nos relatórios e outros documentos gerados a partir dos resultados deste Guia. Já o e-mail e o CPF dos participantes, considerados informações sigilosas, terão seu acesso restrito aos agentes públicos legalmente autorizados e às pessoas a que se referem tais informações, conforme preconiza o artigo 31, §1º, inciso I da Lei nº 12.527, de 18 de novembro de 2011. Outras informações que venham a ser consideradas sigilosas pelos participantes poderão ser apensadas em campo específico no formulário eletrônico.



SUMÁRIO

I. ESCOPO.....	4
II. INTRODUÇÃO.....	4
III. BASE LEGAL.....	4
IV. AVALIAÇÃO ESTATÍSTICA NA LINEARIDADE.....	4
A. Considerações iniciais	4
B. Análise visual	5
C. Do modelo.....	6
1. Decisão entre MMQO ou MMQP	6
2. Utilização do MMQO	8
3. Utilização do MMQP.....	9
4. Avaliação dos coeficientes	9
D. Análise de variância (Anova)	10
1. Anova em MMQO	10
2. Anova em MMQP.....	12
E. Avaliação da correlação linear.....	13
F. Análise de resíduos	13
G. Resumo	15
V. AVALIAÇÃO ESTATÍSTICA DO EFEITO MATRIZ.....	16
VI. GLOSSÁRIO	19
VII. REFERÊNCIAS.....	19



I. ESCOPO

O objetivo deste guia é recomendar modelos e procedimentos de análise estatística de alguns parâmetros da validação, para cumprimento da Resolução RDC nº 166/2017(doravante chamada simplesmente RESOLUÇÃO), que dispõe sobre a validação de métodos analíticos.

Este guia traz recomendações da Anvisa sobre o que se entende como a melhor maneira de cumprir a RESOLUÇÃO. Procedimentos diferentes podem ser aceitos, desde que técnica e cientificamente embasados e que cumpram o disposto na RESOLUÇÃO.

II. INTRODUÇÃO

A validação analítica tem como objetivo fornecer evidências documentais de que o método é adequado para o uso pretendido. Para isso, é necessário testar alguns parâmetros do método analítico, conforme estabelecido na RESOLUÇÃO.

Após testar os parâmetros e obter os resultados, é imprescindível uma avaliação adequada, para que se possa chegar a uma conclusão correta a partir dos resultados obtidos. Para alguns parâmetros, esta avaliação envolve análise estatística.

Os modelos estatísticos partem de alguns pressupostos que precisam ser investigados antes de sua utilização. Caso essa investigação não ocorra ou não seja completa, os resultados obtidos pelos modelos estatísticos podem não ser confiáveis.

III. BASE LEGAL

Resolução – RDC nº 166 de 24 de julho de 2017, publicada no DOU de 25 de julho de 2017.

IV. AVALIAÇÃO ESTATÍSTICA NA LINEARIDADE

A. CONSIDERAÇÕES INICIAIS

Conforme definido no art. 23 da RESOLUÇÃO, linearidade é a capacidade de um método de gerar respostas analíticas diretamente proporcionais à concentração do analito. A resposta analítica em questão é dependente do método, podendo ser, por exemplo,



área do pico, absorvância, volume usado de agente titulante, entre outros. É importante notar que, se uma dada resposta analítica não varia conforme a concentração do analito, ela não é adequada.

É pressuposto dos modelos estatísticos utilizados para o tratamento dos dados de validação, conforme abordagem proposta nesse Guia, que as medições são realizadas em indivíduos independentes. Por isso, existe a necessidade de que as amostras sejam preparadas de maneira independente, isto é, a partir de diferentes pesagens da Substância Química de Referência (SQR). No contexto da linearidade, quando não for viável o preparo de amostras independentes desde o início do procedimento (isto é, desde a pesagem), poderá ser aceito o preparo de soluções diluídas separadamente a partir de uma mesma solução estoque, sendo esta preparada a partir da pesagem da SQR. A injeção ou análise repetida da mesma solução não é adequada para aplicação dos modelos estatísticos aqui mencionados.

Para avaliar se a resposta analítica é, de fato, proporcional à concentração do analito, primeiramente se plotam os dados de concentração e de resposta analítica em um plano cartesiano, com a concentração (variável independente) no eixo das abscissas (x) e a resposta analítica (variável dependente) no eixo das ordenadas (y). Assim, cada resultado de concentração e resposta analítica gerará um ponto neste plano. Esses pontos são referidos como (x_1, y_1) , (x_2, y_2) , (x_3, y_3) . . . (x_i, y_i) . . . (x_n, y_n) , onde n representa o número total de pontos.

O método descrito posteriormente para definição do modelo depende da definição do centroide. O centroide é a posição $(\bar{x}; \bar{y})$, onde \bar{x} é a média aritmética dos valores de x e \bar{y} é a média aritmética dos valores de y .

B. ANÁLISE VISUAL

Antes de qualquer avaliação estatística propriamente dita, é importante fazer uma análise visual para observar se os pontos, quando dispostos no gráfico, têm relação linear aparente. Embora essa análise visual não seja comprobatória, é recomendável que ela preceda o uso de qualquer modelo estatístico, porque, neste caso, nenhum modelo poderá ser usado se a relação for visivelmente não linear – o que pode demonstrar que a resposta analítica não é adequada ou que necessita de transformação matemática.



C. DO MODELO

O “modelo” descrito nesta seção é uma equação que pretende prever os resultados de y a partir de qualquer valor de x . Considerando que a resposta analítica deve ser diretamente proporcional à concentração, para cumprimento da RESOLUÇÃO, esse modelo necessariamente será uma equação de primeiro grau, representada conforme abaixo:

$$y = a + bx + \varepsilon$$

Onde a é definido como coeficiente linear ou intercepto e b é definido como coeficiente angular ou inclinação da reta e ε como o erro do modelo.

O método recomendado para a estimativa da equação é o método dos mínimos quadrados (MMQ). Este método pode ter um fator de ponderação, sendo então denominado método dos mínimos quadrados ponderados (MMQP) ou pode ser aplicado diretamente, sem fator de ponderação, sendo então denominado método dos mínimos quadrados ordinários (MMQO). O MMQO também pode ser entendido como um caso particular do MMQP no qual o fator de ponderação é igual a 1.

A partir da definição do modelo, obtém-se a equação da reta que será utilizada para cálculo dos valores esperados da regressão, que serão necessários para a Anova.

1. Decisão entre MMQO ou MMQP

Para escolha do método a ser utilizado, deve-se considerar a variância dos resultados de y para cada valor de x . Se a variância de y for constante (situação denominada homocedasticidade), pode-se utilizar o MMQO. Caso sejam observadas alterações de variância de y (situação denominada heterocedasticidade), é recomendável que se utilize o MMQP, pois a utilização do MMQO poderá gerar resultados tendenciosos.

A variância dos valores de y para cada um dos níveis de concentração é calculada da seguinte maneira:

$$S_{y_i}^2 = \frac{1}{m-1} \sum_{j=1}^m (y_{ij} - \bar{y}_i)^2$$

Em que: j representa o j -ésimo resultado para o conjunto de réplicas i ; m representa o número de medidas (réplicas) para a resposta analítica y em cada ponto.



Para avaliar se os dados são homocedásticos ou não, é recomendado aplicar o teste de Cochran. As hipóteses são:

$$\begin{cases} H_0: \sigma_{y_1}^2 = \sigma_{y_2}^2 = \sigma_{y_i}^2 \dots = \sigma_{y_n}^2 \\ H_1: \text{pelo menos um dos } \sigma_{y_i}^2 \text{ s diferente,} \end{cases} \quad i = 1, 2, \dots, n$$

Em que $\sigma_{y_i}^2$ representa a variância dos valores de y para um determinado nível de concentração

A hipótese nula (H_0) é que as variâncias sejam todas iguais e a hipótese alternativa (H_1) é que pelo menos uma das variâncias seja diferente das outras.

A estatística a ser testada é:

$$C = \frac{S_{y_{max}}^2}{\sum_i^n S_{y_i}^2} = \frac{\text{maior var\~ancia em y}}{\text{soma de todas as vari\~ancias em y}}$$

Em que o valor de C é a razão entre a maior variância observada para os conjuntos de dados de y e a soma das variâncias de y observadas para todos os níveis de concentração.

O valor de C calculado deve ser comparado ao valor crítico ao nível de significância de 5%. A conclusão do teste será:

Se $C < C_{crítico}$, aceita-se a hipótese nula (dados homocedásticos)

Se $C \geq C_{crítico}$, rejeita-se a hipótese nula (dados heterocedásticos)

Os valores de C crítico ao nível de significância de 5% estão descritos na Tabela 1.


Tabela 1 - Valores de C crítico com nível de significância de 5%.

Número de Pontos	N° de medidas (réplicas) para a variável y			
	2	3	4	5
5	0,841	0,684	0,598	0,544
6	0,781	0,616	0,532	0,48
7	0,727	0,561	0,48	0,431
8	0,68	0,516	0,438	0,391
9	0,638	0,478	0,403	0,358
10	0,602	0,445	0,373	0,331
11	0,57	0,417	0,348	0,308
12	0,541	0,392	0,326	0,288
13	0,515	0,371	0,307	0,271
14	0,492	0,352	0,291	0,255
15	0,471	0,335	0,276	0,242
16	0,452	0,319	0,262	0,23
17	0,434	0,305	0,25	0,219
18	0,418	0,293	0,24	0,209
19	0,403	0,281	0,23	0,2
20	0,389	0,27	0,22	0,192

2. Utilização do MMQO

Se os dados forem homocedásticos, poderá ser utilizado o MMQO. Neste caso, os coeficientes são calculados conforme segue:

$$b = \frac{\sum_{i=1}^n [(x_i - \bar{x})(y_i - \bar{y})]}{[\sum_{i=1}^n (x_i - \bar{x})^2]}$$

$$a = \bar{y} - b\bar{x}$$

Em que b representa a inclinação da reta e a representa o intercepto.

É importante lembrar que cada um dos pontos está sujeito a erro experimental. Diante dos referidos erros, assume-se que estes estão na resposta analítica (y). O erro experimental também estará presente nos coeficientes.

A reta de regressão linear deve sempre passar pelo centroide.



3. Utilização do MMQP

Ao aplicar o método dos mínimos quadrados ponderados, os coeficientes da reta de regressão devem ser estimados utilizando-se um fator de ponderação calculado a partir do inverso da variância de y , s_i^2 , levando em consideração a média harmônica de todas as variâncias de y .

Se os pontos individuais são assinalados por (x_1, y_1) , (x_2, y_2) , (x_3, y_3) ... (x_i, y_i) ... (x_n, y_n) , os desvios padrão correspondentes são s_1, s_2, s_3 ... s_i ... s_n . Pode-se, em seguida, definir os pesos individuais, w_1, w_2, w_3 ... w_i ... w_n , como sendo:

$$w_i = \frac{s_i^{-2}}{\sum_{i=1}^n s_i^{-2}/n}$$

A partir dos pesos individuais, tem-se:

$$b_w = \frac{\sum_{i=1}^n w_i x_i y_i - n \bar{x}_w \bar{y}_w}{\sum_{i=1}^n w_i x_i^2 - n \bar{x}_w^2}$$

$$a_w = \bar{y}_w - b \bar{x}_w$$

Em que: b_w representa a inclinação ponderada da reta e \bar{x}_w e \bar{y}_w representam as coordenadas do centroide ponderado e a_w representa o intercepto ponderado.

A linha de regressão ponderada deve passar através do centroide ponderado, cujas coordenadas são dadas por:

$$\bar{x}_w = \sum_{i=1}^n w_i x_i / n \quad \text{e} \quad \bar{y}_w = \sum_{i=1}^n w_i y_i / n$$

4. Avaliação dos coeficientes

Conforme disposto no inciso V do art. 26 da RESOLUÇÃO, a comprovação da significância do coeficiente angular é fundamental para que seja demonstrada a linearidade do método. Tal avaliação é realizada por meio do teste F da Anova, conforme descrito a seguir.

A avaliação do coeficiente linear pode trazer informação importante, dependendo do método que está sendo validado. Recomenda-se que seja verificado, por meio de teste t, se o coeficiente linear é ou não estatisticamente diferente de zero.

É esperado que o coeficiente linear não seja estatisticamente diferente de zero. Caso ele seja, é importante fazer uma avaliação do impacto deste resultado na linearidade do



método. Recomenda-se considerar, entre outros fatores, o valor do coeficiente linear em relação às concentrações testadas e a possível interferência do diluente ou de outras substâncias no sinal analítico que está sendo avaliado.

Em geral, a linearidade do método precisará ser reavaliada quando o valor do coeficiente linear for estatisticamente diferente de zero e tiver magnitude significativa para o sinal analítico na concentração de trabalho. Além disso, caso seja encontrado intercepto estatisticamente diferente de zero e com magnitude significativa frente às respostas analíticas, é recomendável que se utilize uma curva de calibração ao invés de um ponto único para padronização na rotina de análise.

D. ANÁLISE DE VARIÂNCIA (ANOVA)

A análise de variância (Anova) permite avaliar se uma determinada reta de regressão estimada explica satisfatoriamente a relação existente entre a concentração e o sinal.

A Anova tem fatores diferentes para o cálculo, dependendo do método utilizado (MMQO ou MMQP).

A Anova é realizada por meio da avaliação de todas as fontes de variação do sistema:

- Regressão: variação das respostas analíticas em relação ao centroide (amplitude do modelo)
- Resíduos: variação das respostas analíticas reais em relação ao modelo (erros)
- Total: soma das variações da regressão e dos resíduos.

Para que seja comprovado que y realmente varia em função de x (comprovação fundamental para a linearidade), é necessário demonstrar que a variação devida à regressão (modelo) é suficientemente maior que a devida ao erro (resíduos).

1. Anova em MMQO

As somas dos quadrados estão relacionadas pela fórmula abaixo:

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^n (y_i - \hat{y}_i)^2$$



Em que:

\bar{y} é a média aritmética dos valores de y no centroide

y_i é um valor individual de y obtido experimentalmente em um determinado ponto (i -ésimo ponto)

\hat{y}_i é um valor individual de y calculado pelo modelo (equação) em um determinado ponto (i -ésimo ponto)

n é o número total de pontos

$\sum_{i=1}^n (y_i - \bar{y})^2$ é a soma dos quadrados total, também conhecida como SQTot

$\sum_{i=1}^n (\hat{y}_i - \bar{y})^2$ é a soma de quadrados da regressão, também conhecida como SQReg

$\sum_{i=1}^n (y_i - \hat{y}_i)^2$ é a soma de quadrados devidos aos erros (ou resíduos), também conhecida como SQRes

Essa fórmula também é conhecida como:

$$\mathbf{SQTot = SQReg + SQRes}$$

Após obtenção das somas dos quadrados, podem ser obtidos os quadrados médios, dividindo-se as somas pelo número de graus de liberdade. Tem-se então:

$$\text{Quadrado Médio da Regressão} = \text{QMReg} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{p-1}$$

$$\text{Quadrado Médio dos Resíduos} = \text{QMRes} = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n-p}$$

$$\text{Quadrado Médio Total} = \text{QMTot} = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n-1}$$

p é o número de parâmetros do modelo. Neste caso, $p=2$.

A partir daí, pode-se obter o valor de F , dividindo-se o quadrado médio da regressão (modelo) pelo quadrado médio dos resíduos (erros):

$$F = \frac{\frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{p-1}}{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n-p}} = \frac{\text{QMReg}}{\text{QMRes}}$$

Para ilustração, estão apresentados no Quadro 1 os valores considerados para Anova no MMQO.


Quadro 1 - Resumo da Anova nos MMQO

Fonte de Variação	Graus de liberdade	Soma de Quadrados	Quadrado Médio	F
Regressão	$p - 1$	$\sum_{i=1}^n (x_i - \bar{x})^2$	$\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{p - 1}$	$\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{p - 1} \div \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n - p}$
Resíduos	$n - p$	$\sum_{i=1}^n (y_i - \bar{y})^2$	$\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n - p}$	
Total	$n - 1$	$\sum_{i=1}^n (y_i - \bar{y})^2$	$\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n - 1}$	

As hipóteses do teste são:

$$\begin{cases} H_0: b = 0 \\ H_1: b \neq 0 \end{cases}$$

Comparando-se o valor de F com um valor crítico, toma-se a decisão de aceitar ou rejeitar a hipótese nula de que $b=0$.

Se $F_{calculado} \geq F_{(\alpha,1,n-2)} \Rightarrow$ rejeita-se a hipótese nula de que $b=0$ e assume-se que y efetivamente varia em função de x e que o método pode ser considerado linear.

$F_{(0,05;1;n-2)}$ tem distribuição de probabilidade F de Snedecor com 1 grau de liberdade no numerador e $n-2$ graus de liberdade no denominador ao nível de significância de 5%.

2. Anova em MMQP

Quando se utiliza o MMQP, os mesmos fatores de ponderação devem ser utilizados na Anova. Por isso, as fórmulas são diferentes da Anova em MMQO, embora o racional aplicado seja o mesmo.

Para ilustração, estão apresentados no Quadro 2 os valores considerados para Anova no MMQP.

Quadro 2 - Resumo da Anova em MMQP

Fonte de Variação	Graus de Liberdade	Soma de Quadrados	Quadrado Médio	F
Regressão	$p - 1$	$\sum_{i=1}^n w_i (x_i - \bar{x})^2$	$\frac{\sum_{i=1}^n w_i (x_i - \bar{x})^2}{p - 1}$	$\frac{\sum_{i=1}^n w_i (x_i - \bar{x})^2}{p - 1} \div \frac{\sum_{i=1}^n w_i (y_i - \bar{y})^2}{n - p}$
Resíduos	$n - p$	$\sum_{i=1}^n w_i (y_i - \bar{y})^2$	$\frac{\sum_{i=1}^n w_i (y_i - \bar{y})^2}{n - p}$	
Total	$n - 1$	$\sum_{i=1}^n w_i (y_i - \bar{y})^2$	$\frac{\sum_{i=1}^n w_i (y_i - \bar{y})^2}{n - 1}$	



Na Anova para MMQP, o F crítico e a regra de decisão para a hipótese nula são iguais ao Anova para MMQO.

E. AVALIAÇÃO DA CORRELAÇÃO LINEAR

Após determinado o modelo, pode ser calculado o coeficiente de correlação de Pearson (r), que representa o grau de associação linear entre as variáveis.

Caso seja usado o MMQO, o r é calculado da seguinte maneira:

$$r = \frac{\sum_{i=1}^n [(x_i - \bar{x})(y_i - \bar{y})]}{\sqrt{[\sum_{i=1}^n (x_i - \bar{x})^2][\sum_{i=1}^n (y_i - \bar{y})^2]}}$$

Em que: r representa o coeficiente de correlação de Pearson.

Caso seja usado o MMQP, o r é calculado da seguinte maneira:

$$r_w = \frac{\sum_i^n w_i \cdot \sum_i^n w_i x_i y_i - \sum_i^n w_i x_i \cdot \sum_i^n w_i y_i}{\sqrt{[\sum_i^n w_i \cdot \sum_i^n w_i x_i^2 - (\sum_i^n w_i x_i)^2][\sum_i^n w_i \cdot \sum_i^n w_i y_i^2 - (\sum_i^n w_i y_i)^2]}}$$

Em que: r_w representa o coeficiente de correlação de Pearson ponderado.

Quanto maior o valor de r , mais forte é a correlação entre as duas variáveis e mais realista é o modelo (equação) proposto. O valor mínimo aceito para r ou r_w é de 0,990, conforme disposto no § 3º do art. 27 da RESOLUÇÃO. Os valores de r e r_w devem ser arredondados considerando pelo menos 3 casas decimais.

F. ANÁLISE DE RESÍDUOS

Resíduo é a diferença entre o valor encontrado de y e o valor estimado pelo modelo (equação) em um determinado ponto, $(y_i - \hat{y}_i)$. O resíduo quantifica a distância entre o valor real e o estimado.

Quando o erro do modelo é devido apenas às variações comuns de análise, espera-se que os resíduos sejam independentes entre si e tenham distribuição normal, ou seja:

$$\varepsilon_i \sim^{iid} N(0, \sigma^2)$$

É importante que a normalidade, a independência e a homocedasticidade sejam investigadas.



Para avaliação da normalidade dos resíduos, recomenda-se:

- Avaliação visual do gráfico
- Testes estatísticos de normalidade, tais como: Shapiro Wilk, Lilliefors, gráficos de probabilidade ou histograma.

Os gráficos para avaliação visual devem incluir a distribuição de todos os resíduos em função da concentração e da resposta, preferencialmente com uma linha reta no valor de média e outras duas linhas delimitando os valores esperados. Para estimar os valores esperados, pode-se utilizar testes como o de Grubbs, D de Cook, entre outros. Valores fora da faixa esperada são considerados *outliers* (valores aberrantes). Neste caso, é importante investigar se existe alguma causa analítica para esses valores aberrantes e se, mesmo com esses valores, os dados ainda podem ser utilizados.

Exemplo de gráfico para avaliação dos resíduos está descrito na figura.

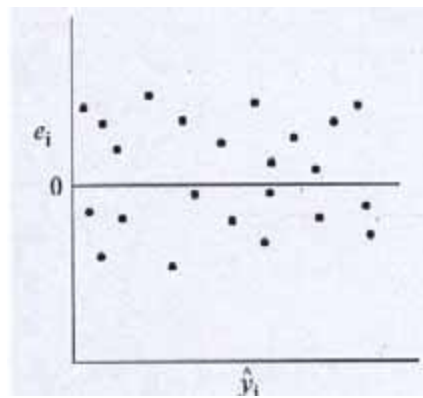


Figura 1 - Gráfico de distribuição dos resíduos em função da resposta analítica. Neste caso, a variância dos resíduos parece ser constante (homocedasticidade).



G. RESUMO

Para ilustração, apresenta-se fluxo resumido para análise estatística dos dados referentes ao parâmetro linearidade

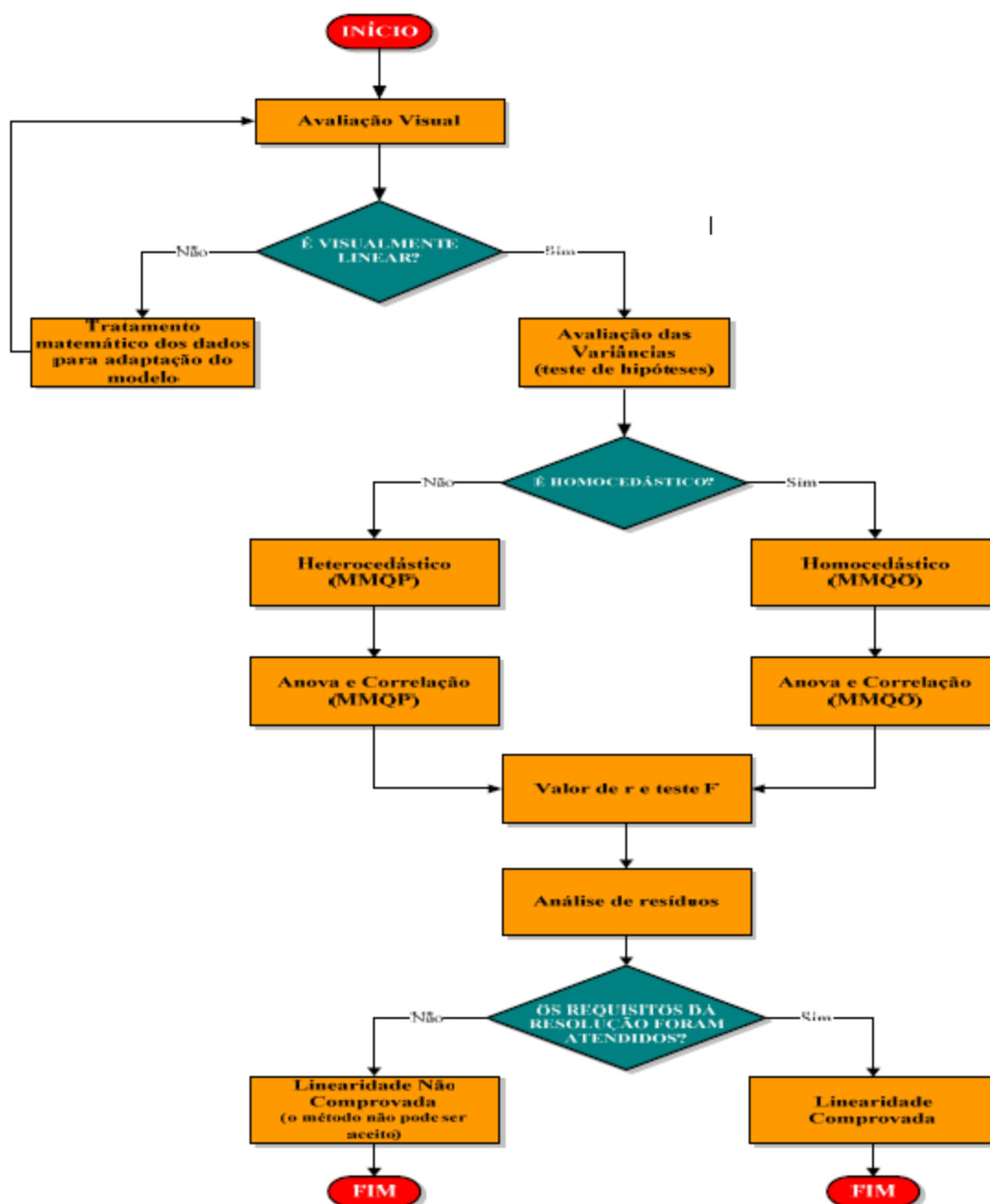


Figura 2 - Fluxograma para avaliação de linearidade.



V. AVALIAÇÃO ESTATÍSTICA DO EFEITO MATRIZ

Para comprovação da ausência de efeito matriz, é necessário demonstrar que as retas de concentração *versus* resposta analítica da solução de SQR e da solução amostra fortificada com SQR são paralelas, conforme disposto nos artigos 29 e 30 da RESOLUÇÃO. As retas poderão ser consideradas paralelas quando seus coeficientes angulares ou inclinações não forem estatisticamente diferentes.

Existem algumas maneiras para avaliar se as inclinações não são estatisticamente diferentes. Para qualquer uma delas, é fundamental que os pressupostos estatísticos sejam atendidos e, se necessário, testados.

Uma das maneiras possíveis de se avaliar o efeito matriz é o teste t. É importante notar que o uso adequado do teste t só é possível diante de um conjunto de dados que permita determinar adequadamente a variância dos valores que serão comparados, por isso existe a necessidade de que a comparação envolva um número adequado de curvas conforme a variância; quanto maior a variância, maior o número de curvas necessárias.

Para correta aplicação do teste t, também é fundamental que se comparem as variâncias de cada um dos grupos de amostras que se pretende testar. No caso, as variâncias da inclinação da curva de solvente contaminado com SQR e da curva de amostra contaminada com SQR precisam ser comparadas, pois o método de cálculo do teste t vai depender da igualdade ou não dessas variâncias. Assim, antes da aplicação do teste t, é importante que as variâncias sejam comparadas por meio do teste F de Snedecor.

Para cada reta (SQR em solvente e amostra fortificada com SQR) será determinada uma equação de regressão da mesma maneira que na linearidade:

$$y_1 = a_1 + b_1x_1 + \varepsilon$$

$$y_2 = a_2 + b_2x_2 + \varepsilon$$

Em que:

y_1 corresponde ao sinal analítico para a curva de calibração obtida com SQR em solvente;

y_2 corresponde ao sinal analítico para a curva de calibração obtida com a amostra fortificada com SQR;

x_1 corresponde à concentração da SQR em solvente (grupo 1);

x_2 corresponde à concentração do analito nas amostras fortificadas com SQR (grupo 2);



a_1 corresponde ao intercepto da curva de calibração obtida com o SQR em solvente;

a_2 corresponde ao intercepto da curva de calibração obtida com a amostra fortificada com SQR;

b_1 corresponde à inclinação da curva de calibração obtida com o SQR em solvente;

b_2 corresponde à inclinação da curva de calibração obtida com a amostra fortificada com SQR;

ε corresponde ao erro do modelo.

Os erros ε são independentes e identicamente distribuídos como uma distribuição normal com média 0 e variância σ^2 , $N(0, \sigma^2)$.

O teste t compara as inclinações das retas ajustadas segundo as equações definidas acima, testando a hipótese nula de que as inclinações são iguais (retas paralelas). Temos então:

$$\begin{cases} H_0: b_1 = b_2 \\ H_1: b_1 \neq b_2 \end{cases}$$

Quando o teste F indica variâncias semelhantes, o valor de T é calculado conforme abaixo:

$$T = \frac{\widehat{b}_1 - \widehat{b}_2}{S_{\widehat{b}_1 - \widehat{b}_2}}$$

Em que:

\widehat{b}_1 é a estimativa de b_1 , usando as n_1 observações do grupo 1;

\widehat{b}_2 é a estimativa de b_2 , usando as n_2 observações do grupo 2;

$S_{\widehat{b}_1 - \widehat{b}_2}$ é a estimativa do desvio-padrão da diferença entre os coeficientes angulares dos modelos. Este desvio-padrão é igual à raiz quadrada da seguinte variância:

$$S_{\widehat{b}_1 - \widehat{b}_2}^2 = S_{P,y/x}^2 \left[\frac{1}{(n_1-1)S_{x_1}^2} + \frac{1}{(n_2-1)S_{x_2}^2} \right],$$

Em que:

$$S_{P,y/x}^2 = \frac{(n_1 - 2)S_{y/x_1}^2 + (n_2 - 2)S_{y/x_2}^2}{n_1 + n_2 - 4},$$



Em que:

S_{y/x_1}^2 é o quadrado médio do resíduo para o grupo 1;

S_{y/x_2}^2 é o quadrado médio do resíduo para o grupo 2;

$S_{x_1}^2$ é a variância dos X para o grupo 1;

$S_{x_2}^2$ é a variância dos X para o grupo 2.

O valor obtido de T deve ser comparado com o valor crítico, considerando $n_1 + n_2 - 4$ graus de liberdade e nível de significância de 5%.

Se $T_{calculado} \geq T_{crítico} \Rightarrow$ rejeita-se a hipótese nula e considera-se que existe efeito matriz significativo e o método não é suficientemente seletivo. Para que este método seja utilizado, devem ser propostas ações corretivas que minimizem a distorção no resultado causada pelo efeito matriz.

Se a hipótese nula não for rejeitada, considera-se que não há efeito matriz significativo. Neste caso, as retas de regressão podem ser consideradas paralelas e a mesma inclinação pode ser aplicada nos dois modelos. Teremos então:

$$y_1 = a_1 + b_3x + \varepsilon \text{ e}$$

$$y_2 = a_2 + b_3x + \varepsilon$$



VI. GLOSSÁRIO

Anova – Análise de Variância

ANVISA- Agência Nacional de Vigilância Sanitária

MMQ - Método dos Mínimos Quadrados

MMQP - Método dos Mínimos Quadrados Ponderados

MMQO - Método dos Mínimos Quadrados Ordinários

R - Coeficiente de correlação de Pearson

RDC – Resolução de Diretoria Colegiada

SQR - Substância Química de Referência

VII. REFERÊNCIAS

Miller, J. N; Miller, J. C. Statistics and Chemometrics for Analytical Chemistry. 6. Ed. England: Pearson, 2010.

Ministério da Agricultura, Pecuária e Abastecimento (MAPA), Guia: Validação e Controle de Qualidade Analítica (Fármacos em Produtos para Alimentação Animal e Medicamentos Veterinários), Brasília, 2011.