II Workshop de Acreditação de Produtores de Materiais de Referência e de Provedores de Ensaios de Proficiência

Aspectos técnicos relacionados à norma ISO 13528:2005

Roberto Gonçalves Junqueira

Professor Associado
ALM/FAFAR/UFMG

Setembro, 2011

Conteúdo

- Visão geral dos principais procedimentos estatísticos
- Determinação de valor designado e sua incerteza
- Determinação do desvio-padrão da proficiência
- Cálculo do índice de desempenho
- Verificação da homogeneidade e estabilidade

Estatística de Ensaios de Proficiência

- ISO 13528:2005. Statistical Methods for use in proficiency testing by interlaboratory comparisons.
- IUPAC 2006. Thompson, M; Ellison, S.L.R.; Wood, R. The international harmonized protocol for proficiency testing of analytical chemistry Laboratories.
- ISO/IEC 17043:2010. Conformity Assessment – General requirements for proficiency testing.

Recomendações para limitar a incerteza do valor designado

$$X \pm u_X$$

X é o valor designado de uma quantidade particular e

u_X é sua incerteza padrão, que depende:

- do método pela esta qual foi estimada;
- do número de laboratórios, quando se usam os dados dos participantes do ensaio de proficiência em sua estimativa;
- de outros fatores não controlados.

Recomendações para limitar a incerteza do valor designado

O desvio-padrão para avaliação da proficiência σ_p é utilizado para avaliar a estimativa da tendência do laboratório.

$$u_X \leq 0,3\sigma_P$$

Se essa condição não se verificar, a incerteza do valor designado deverá ser fornecida aos participantes, pois esta poderá afetar seu desempenho na proficiência.

Avaliação da incerteza do valor designado

Exemplo

Suponha que o valor designado X é determinado como a média dos resultados de 11 laboratórios e que o desvio padrão da proficiência é o desvio-padrão s desses mesmos 11 resultados, assim $\sigma_p = s$.

Uma primeira aproximação para u_X pode ser obtida por

$$u_X = s/\sqrt{11} = 0,3s$$

que satisfaz-se a condição estabelecida.

Mas, se o numero de laboratórios é $p < 11$, $u_X > s/\sqrt{(11)}$ e a condição $u_X \leq 0,3\sigma_p$ não pode ser satisfeita.

Recomendações para a escolha do número replicatas de medição

A variação da **repetitividade** contribui para a variação entre as tendências dos laboratórios em um EP.

Para minimizar essa influência, o **número de replicatas** n definido por cada laboratório deve satisfazer a condição:

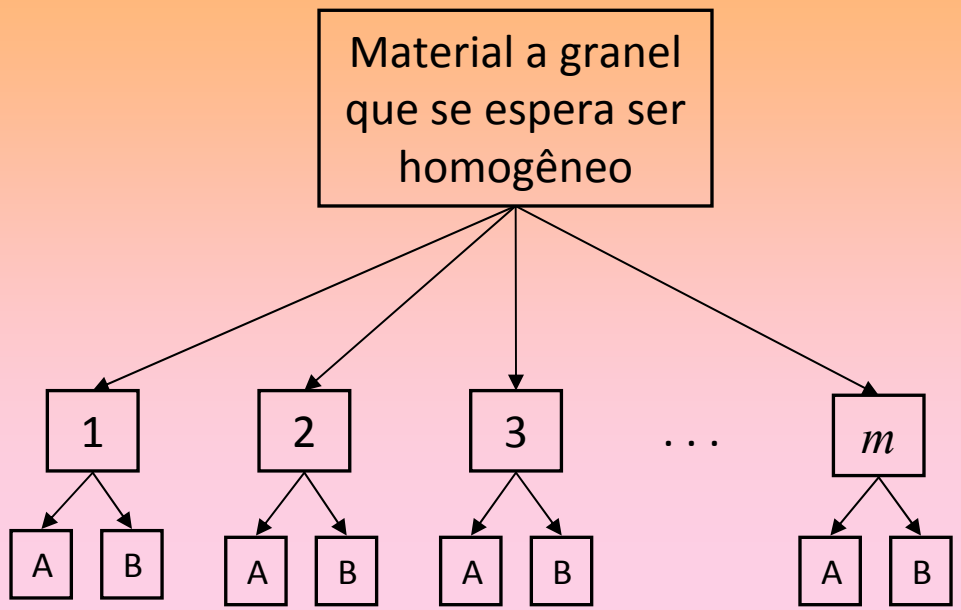
$$\sigma_r / \sqrt{n} \leq 0,3\sigma_p$$

Quando isto se verifica, o desvio-padrão (prévio) de repetitividade σ_r contribui com menos de 10% do desvio-padrão da proficiência σ_p .

Homogeneidade e estabilidade das amostras

- FEARN, T. & THOMPSON, M. 2001. *A new test for sufficient homogeneity*. Analyst, v. 126, p.1414-1417.
- RSC 2004. *Test for 'sufficient homogeneity' in a reference material*. In: Analytical Methods Committee Technical Brief of The Royal Society of Chemistry.
- ISO 13528:2005. *Annex B. Homogeneity and stability checks of samples*. In: Statistical Methods for use in proficiency testing by interlaboratory comparisons.
- IUPAC 2006. *3.11 Testing for sufficient homogeneity and stability*. In: Thompson, M; Ellison, S.L.R.; Wood, R. The international harmonized protocol for proficiency testing of analytical chemistry Laboratories.
- ILAC 2008. *Discussion Paper on Homogeneity and Stability Testing.* (PTCG_HandS_April0808).
- ISO/IEC 17043:2010. *B.5 Demonstration of proficiency test item homogeneity and stability.* In: Conformity Assessment – General requirements for proficiency testing.

Procedimento para teste de homogeneidade



$$\sigma_{sam} < 0,3\sigma_p$$

$$\sigma_r < 0,5\sigma_p$$

Condições ideais

Amostra controle dos laboratórios

$m \geq 10$

Duplicatas de análise

Análise em ordem aleatória

$$D_i = y_{iA} - y_{iB}$$

$$S_{DD} = \sum_{i=1}^{m} D_i^2$$

$$C = \frac{D_{\max}^2}{S_{DD}}$$

Teste C de Cochran

Tratamento de valores discrepantes

$$S_i = y_{iA} + y_{iB}$$

$$s_{sam}^2 = \left(s_S^2/2 - s_r^2\right)/2$$

$$\bar{S} = (1/m)\sum_i S_i$$

$$s_S^2 = \sum_i^m \left(S_i - \bar{S}\right)^2 \Big/ (m-1)$$

$$s_{sam}^2 \leq \frac{\chi_{m-1}^2 0,3^2 \sigma_p^2}{m-1} + \frac{\left(F_{m-1,m} - 1\right)s_r^2}{2}$$

Homogeneidade suficiente

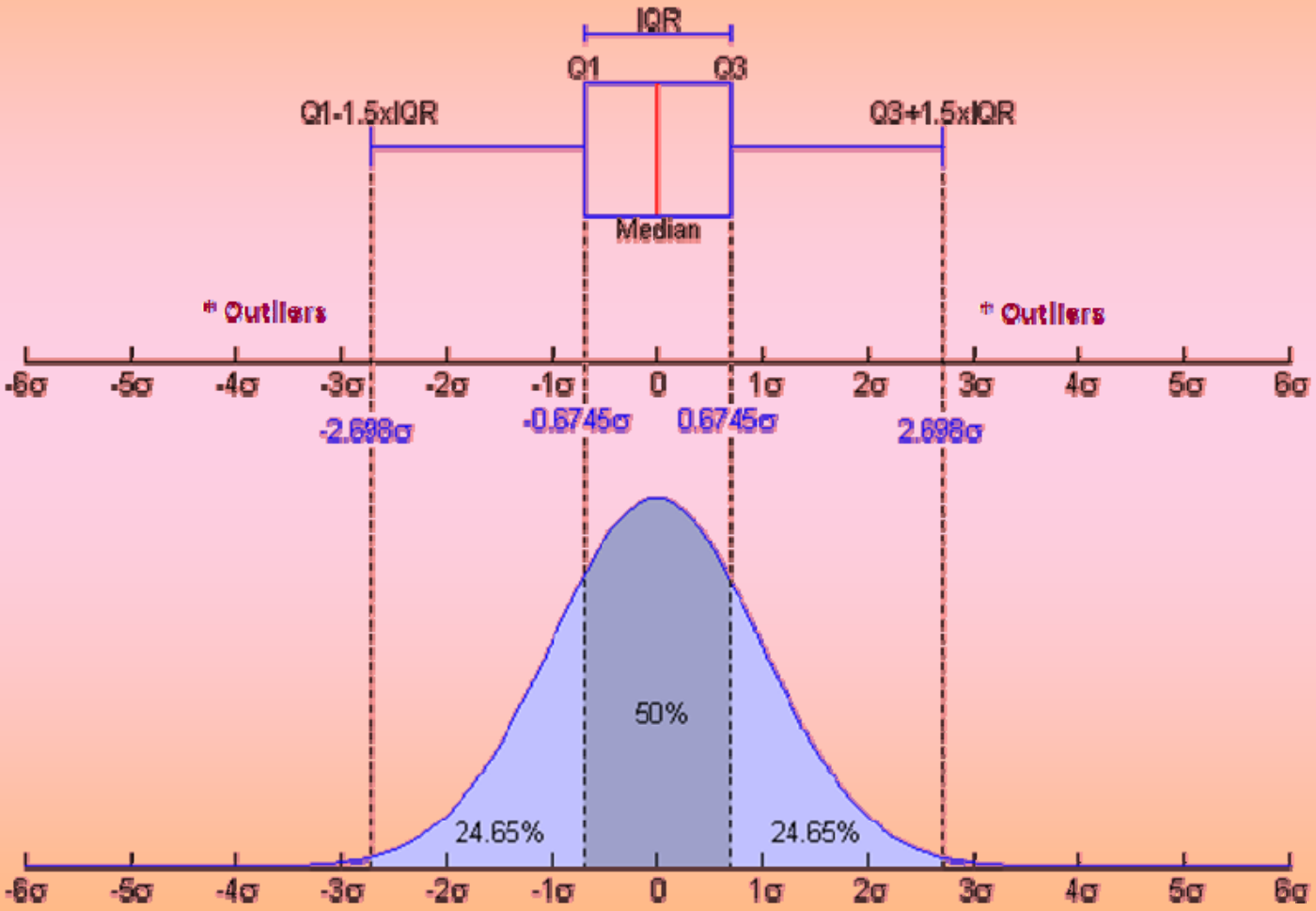
s_{sam}: estimativa da variação na composição entre os recipientes (amostras)

s_r : estimativa do desvio padrão de repetitividade (estima a variabilidade analítica)

σ_p : desvio-padrão da proficiência (desvio-padrão alvo)

(IUPAC, 2006)

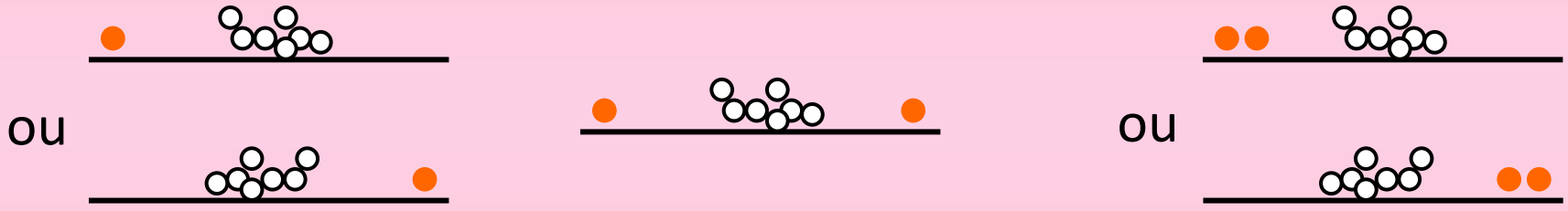
Tratamento de valores discrepantes pelo gráfico de Box-Whiskers



(TUKEY, 1977)

Tratamento de valores discrepantes pelo método de Grubbs

Fazer o tratamento em ciclos até que valores discrepantes não sejam mais detectados ou até que seja retirado 22,2% dos dados originais.



$$G_1 = \frac{\left|\bar{X} - X_i\right|}{s} \qquad G_2 = \frac{X_n - X_1}{s} \qquad G_3 = 1 - \left(\frac{(n-3)s_{n-2}^2}{(n-1)s^2} \right)$$

Para cada ciclo, será considerado valor discrepante o resultado que fornecer G1, G2 ou G3 superior ao valor crítico.

(BURKE, 2001)

Verificação da estabilidade

$$\left|\mu_H - \mu_E\right| \leq 0,3\sigma_p$$ (ISO 13528:2005)

Se este critério não for cumprido, deve-se examinar a preparação das amostras e os procedimentos de armazenamento para verificar se será possível introduzir melhorias.

$$\left|\mu_0 - \mu_1\right| < 0,1\sigma_p$$ (IUPAC, 2006)

Uma distinção conceitual deve ser feita entre instabilidade significativa e instabilidade consequencial. Pode haver uma variação significativa nos resultados, mas esta variação pode ter efeito desprezível no cálculo do escore-z.

$$\left|b_1\right| < t_{0,975;n-2}s\left(b_1\right)$$ (ISO GUIDE 35:2006)

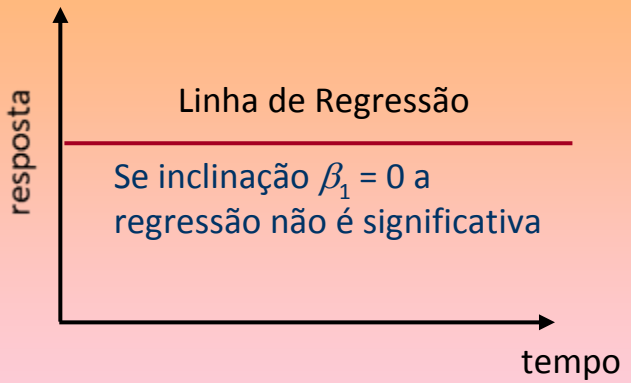
Se não se conhece o mecanismo físico-químico que descreve a degradação do material, uma reta de inclinação b_1 e erro-padrão $\mathrm{s}(b_1)$ pode ser usada para avaliar a tendência de alteração.

Teste de Hipótese para a Inclinação

Hipóteses

$H_0 : \beta_1 = 0$ (a regressão linear não é significativa)

$H_1 : \beta_1 \neq 0$ (a regressão linear é significativa)

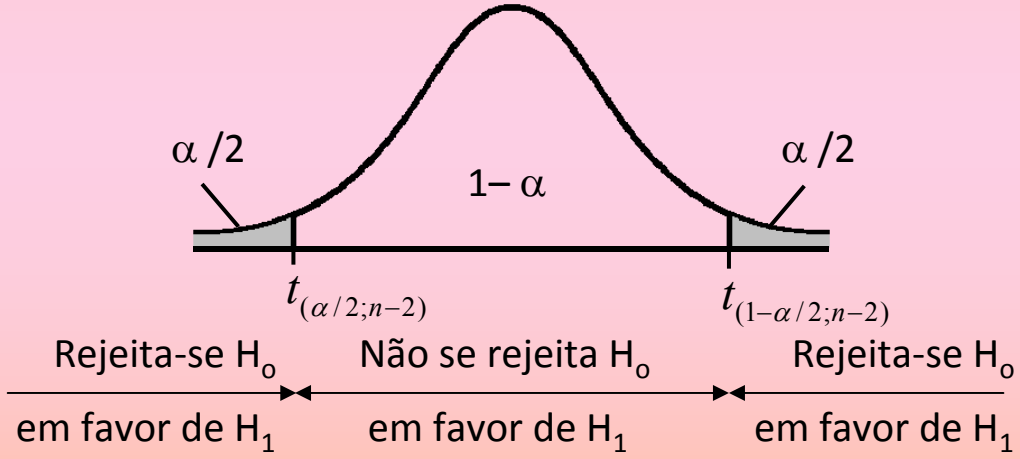


Estatística do teste

$$t = \frac{b_1 - 0}{s(b_1)} \qquad b = S_{XY} / S_{XX} \qquad s(b_1) = \sqrt{s^2_{res} / S_{XX}}$$

$$s^2_{res} = \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2 \Big/ (n-2)$$

$$t = \frac{b_1}{\sqrt{s^2_{res} \Big/ \sum_{1=1}^{n} \left(X_i - \bar{X}\right)^2}}$$



Regra de decisão

Se $\left| t \right| > t_{1-\alpha/2;n-2gl} \Rightarrow 2p < \alpha$ (Rejeita -se H$_o$ em favor de H$_1$)

Intervalos de confiança

$$P\left\{ b_1 \pm t_{1-\alpha/2;n-2gl} \sqrt{s^2_{res} / S_{XX}} \right\} = 100(1-\alpha)\%$$

(DANIEL, 2006)

Apresentação dos resultados

- recomenda-se que os resultados individuais não sejam arredondados mais do que $\sigma_r/2$;
- os participantes devem fornecer os valores reais de suas medições, mesmo que valores negativos sejam registrados;
- os resultados não devem ser truncados, não sendo aceitáveis “< 0,1” ou “menor que o limite de detecção”.

Determinação do valor designado

Escolha do método

- Formulação
- Materiais de Referência Certificados
- Materiais de Referência
- Valor de consenso de laboratórios especialistas
- Valor de consenso de laboratórios participantes

Análise robusta x eliminação de dispersos

Determinação do valor designado

Formulação

- Mistura de constituintes

  deriva-se pela massa empregada

- Amostras brancas adicionadas do analito

  - podem não representar situações reais

  - podem produzir amostras não homogêneas

Determinação do valor designado

Materiais de Referência Certificados

- utiliza-se o valor certificado
- a incerteza padrão do valor designado

  é derivada a partir da incerteza do certificado
- as limitações de uso referem-se a custo e disponibilidade

Exemplo

X_{CRM} = 21,62 unidades LA $u_{X;\mathrm{CRM}}$ = 0,26 unidades LA

Determinação do valor designado

Materiais de Referência (RM)

- Preparados antes da distribuição
- Testados contra material de referência certificado (CRM)

método da média das diferenças

$$D_i = \bar{X}_{i;RM} - \bar{X}_{i;CRM}$$

$$\bar{D} = \frac{1}{m}\sum_{i=1}^{m} D_i$$

$$X_{RM} = X_{CRM} + \bar{D}$$

$$u_{X;RM} = \sqrt{u^2_{X;CRM} + u^2_D}$$

Determinação do valor designado

Valor de consenso de laboratórios especialistas

- Os laboratórios especialistas podem ser participantes

  - análise robusta

  - outro método estatístico referenciado

$$u_X = \frac{1,25}{p}\sqrt{u_{X;CRM}^2 + u_D^2} \qquad p > 10$$

dados com distribuição normal

Determinação do valor designado

Valor de consenso de laboratórios participantes

- Determinado por análise robusta (Algoritmo A)
- Outros métodos referenciados

$$X = x^* = mediana(X_i) \qquad (i = 1, 2, ..., p)$$

$$u_X = 1,25\, s^* / \sqrt{p} \qquad p > 10$$

- Pode não haver um consenso real
- O consenso pode ser tendencioso

Determinação do valor designado

Comparação do valor designado com o consenso

- Quando X é obtido com materiais de referência

  - estima-se a mediana de consenso (x^*)

  - compara-se ao valor designado (X)

$$u_{x^*-X} = \sqrt{\frac{\left(1,25 s^*\right)}{p} + u_X^2} \qquad p > 10$$

$$\left| x^* - X \right| < 2 \cdot u_{x^*-X}$$

Caso contrário as razões para a variabilidade devem ser investigas

Cálculo da média e desvio-padrão por análise robusta

Algoritmo A para média e desvio-padrão

- Inicia-se com $x^* = mediana$

$$s^* = 1{,}483 \text{ x } \text{mediana } |x_i - x^*|$$

- Limita-se os dados em $x^* + 1{,}5s^*$ e $x^* - 1{,}5s^*$

- Como opção pode-se usar a primeira estimativa de x^* e s^* e não fazer iterações

Cálculo da média e desvio-padrão por análise robusta

Algoritmo A para média e desvio-padrão

- Calcula-se novas estimativas:

$$x^* = (\Sigma x_i)/p$$

$$s^* = 1{,}134\sqrt{\Sigma(x_i^* - x^*)^2/(p-1)}$$

- Limita-se os dados novamente em $1{,}5s^*$

  Recalcula-se novos x^* e s^*

- Repete-se até convergência

  (estabilização dos valores de x^* e s^*)

Valores faltantes

$n_{\min} \geq 0{,}59n$ → Considerar as medições do participante no cálculo das estatísticas

$n_{\min} < 0{,}59n$ → Não considerar as medições do participante no cálculo das estatísticas

$s_{\bar{X}} = \sigma_r / \sqrt{n}$

$\downarrow n \quad \uparrow s_{\bar{X}}$ → Quando n se reduz para 0,59n, o desvio-padrão aumento de um fator de 1,3.

Determinação do desvio-padrão da proficiência

Escolha do método

- Por valor prescrito (decisão regulatória)
- Por percepção (expectativas baseadas na experiência)
- Por modelo estatístico geral (Horwitz/Thompson)
- Por meio de estudos de precisão
- Pelos dados da rodada

Análise robusta x eliminação de dispersos

Determinação do desvio-padrão da proficiência

Valor prescrito

- Uma adequação a um propósito de uso, conforme:
  - determinado pelo julgamento de especialista;
  - decisão regulatória.

- Pode não ser um valor realista em relação à reprodutibilidade do método de medição.

Determinação do desvio-padrão da proficiência

Por percepção

- Uma estimativa de rodadas anteriores do ensaio de proficiência ou expectativas baseadas na experiência
- Pode corresponder ao nível de desempenho desejável ou requerido.
- Pode não ser um valor realista

$$\sigma_L = \sqrt{\sigma_R^2 - \sigma_r^2} \qquad \sigma_p = \sqrt{\left(\phi \times \sigma_L\right)^2 + \left(\sigma_r^2 / n\right)} \qquad \phi \geq 0,5 \quad aceitável$$

σ_r: desvio padrão de repetitividade

σ_R: desvio padrão de reprodutibilidade

Determinação do desvio-padrão da proficiência

Modelo estatístico geral

- Modelo de Horwitz/Thompson

$$\sigma_R = \begin{cases} 0{,}22c & \text{se } c < 1{,}2 \times 10^{7} \quad \text{(THOMPSON, 2000)} \\ 0{,}02c^{0{,}8495} & \text{se } 1{,}2\times10^{7} \leq c \leq 0{,}138 \quad \text{(HORWITZ, 1982)} \\ 0{,}01c^{0{,}5} & \text{se } c > 0{,}138 \quad \text{(THOMPSON, 2000)} \end{cases}$$

σ_R : desvio padrão de reprodutibilidade

c : fração de maça (adimensional, para 1 ppm = 1 mg/kg, $c = 10^{-6}$)

Determinação do desvio-padrão da proficiência

Resultados de estudos de precisão

- Estudos de precisão intermediária (s_R parcial)

$$\sigma_L = \sqrt{\sigma_R^2 - \sigma_r^2} \qquad \sigma_p = \sqrt{\sigma_L^2 + \left(\sigma_r^2 / n\right)}$$

σ_r: desvio padrão de repetitividade

σ_R: desvio padrão de reprodutibilidade

(ISO 5725-2:1994)

Determinação do desvio-padrão da proficiência

Pelos dados da rodada

- O desvio-padrão usado para avaliar a proficiência dos participantes em uma rodada é derivado dos dados dos próprios participantes da mesma rodada.
- O valor do desvio padrão da proficiência pode variar substancialmente de rodada para rodada.

Análise robusta x eliminação de dispersos

Estatísticas de desempenho

Estimativa da tendência do laboratório

$$D = x - X$$

em que:

x: é o resultado do participante

X: é o valor designado

99,7%: $-3{,}0\sigma_p < D < 3{,}0\,\sigma_p$ **A**

95,0%: $-2{,}0\sigma_p < D < 2{,}0\,\sigma_p$ **W**

- A diferença simples entre o resultado do participante e o valor designado pode ser adequada para determinar o desempenho e é facilmente compreendida pelos participantes.

Estatísticas de desempenho

Diferença percentual

$$D_{\%} = 100\left(x - X\right)/X$$

em que:

x: é o resultado do participante

X: é o valor designado

99,7%: $-300\sigma_p < \mathrm{D} < 300\ \sigma_p$ **A**

95,0%: $-200\sigma_p < \mathrm{D} < 200\ \sigma_p$ **W**

- É interpretada de maneira semelhante à diferença simples

  A: sinal de ação

  W: sinal de alerta

Estatísticas de desempenho

Escores e escores percentuais

- Os dados são ordenados e ranqueados

$$i = 1, 2, \ldots, p$$

- O escore percentual pode ser definido por

$$100(i - 0{,}5)/p$$

  - Não se assume qualquer distribuição
  - Não é dependente de x* e s*

Estatísticas de desempenho

O escore z

$$z = (x - X)/\sigma_p$$

99,7%: $-3{,}0\sigma_p < z < 3{,}0\ \sigma_p$ **A**

95,0%: $-2{,}0\sigma_p < z < 2{,}0\ \sigma_p$ **W**

- Gráficos de probabilidade normal (p > 100) auxiliam na interpretação.
- Para p < 10 participantes não devem ser dados sinais de alerta baseados em uma rodada única.

Estatísticas de desempenho

Número E_n

$$E_n = \frac{x - X}{\sqrt{U_{lab}^2 + U_{ref}^2}}$$

X: valor designado em um laboratório referência

U_{ref}: incerteza expandida de X

U_{lab}: incerteza expandida de um resultado particular x

95,0%: $-1 < E_n < 1$

Estatísticas de desempenho

Escore z’

$$z' = (x - X) \Big/ \sqrt{\sigma_p^2 + u_X^2}$$

Sendo u_X a incerteza padrão do valor designado X

Pode ser usado se o valor designado não foi obtido por consenso

99,7%: $-3{,}0\sigma_p < z < 3{,}0\ \sigma_p$ **A**

95,0%: $-2{,}0\sigma_p < z < 2{,}0\ \sigma_p$ **W**

Estatísticas de desempenho

Escore ζ

$$\zeta = (x - X) \Big/ \sqrt{u_x^2 + u_X^2}$$

Só deve ser usado se o valor designado

não foi obtido por consenso

99,7%: $-3{,}0\sigma_p < z < 3{,}0\ \sigma_p$ **A**

95,0%: $-2{,}0\sigma_p < z < 2{,}0\ \sigma_p$ **W**

Estatísticas de desempenho

Escore E_z

$$E_{z-} = \frac{x - (X - U_X)}{U_x} \qquad E_{z+} = \frac{x - (X + U_X)}{U_x}$$

Resultado satisfatório: $-1{,}0 < E_{z-}$ e $E_{z+} < 1{,}0$

Quando E_{z-} ou E_{z+} está fora do intervalo: desempenho questionável

Resultado insatisfatório: E_{z-} e E_{z+} fora do intervalo -1,0 ;+1,0

Métodos gráficos para combinação de índices de desempenho

Aplicação

- Facilitam a apresentação e interpretação dos dados
- Podem ser distribuídos aos participantes
  - os resultados são codificados
  - permitem aos participantes compararem seus dados com os dos demais participantes
- São úteis para o coordenador avaliar o ensaio

Métodos gráficos para combinação de índices de desempenho

Histogramas de índices de desempenho

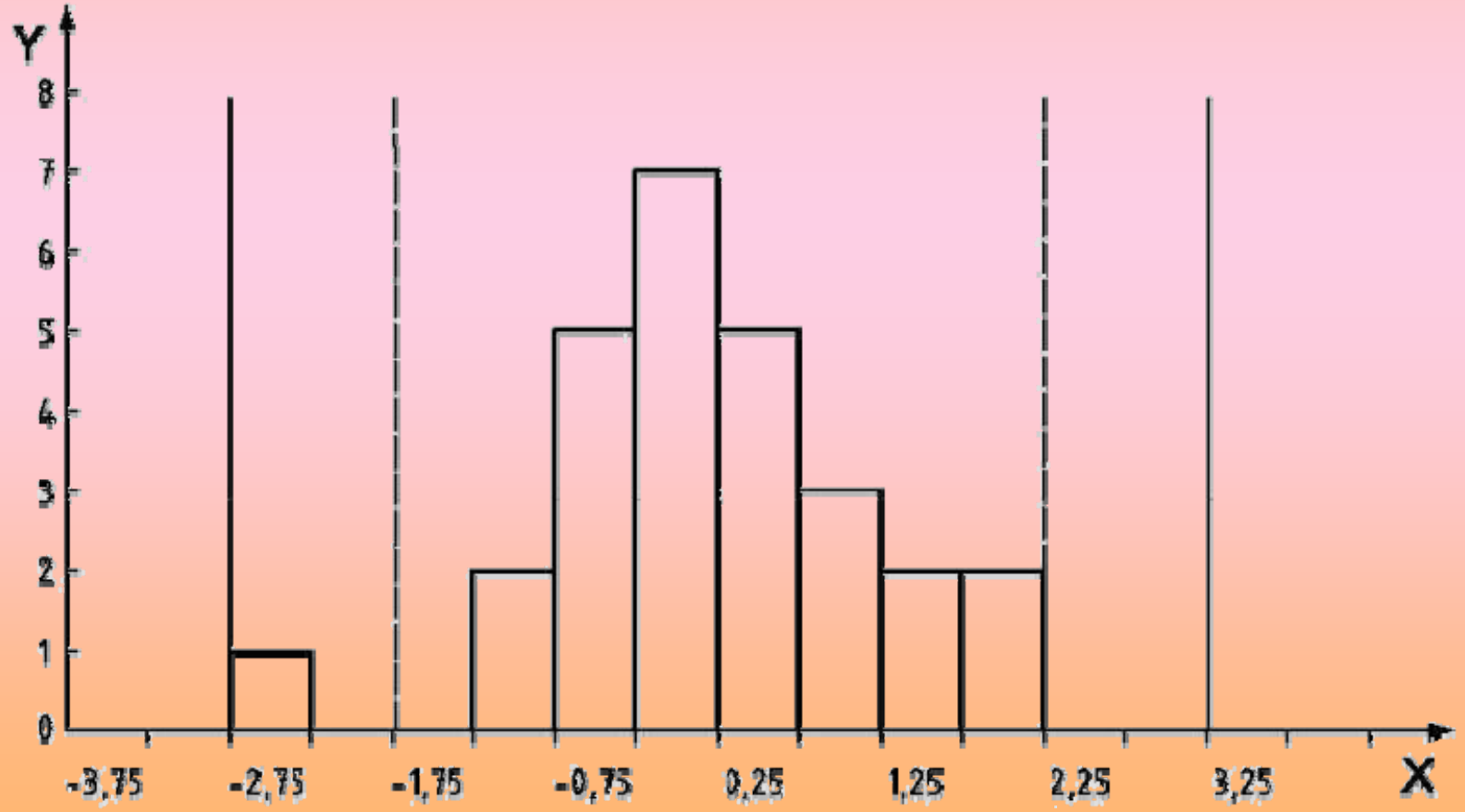
- São úteis quando o número de participantes é baixo.
- Auxiliam na interpretação do z-escore, da estimativa da tendência do laboratório ou da diferença percentual.

Limites de advertência e de ação para índices de desempenho

Estatística de desempenho	Intervalo de classe	Limites de advertência	Limites de ação	Amplitude do histograma
Tendência	$0,3\sigma_p$ a $0,5\sigma_p$	$\pm 2,0\sigma_p$	$\pm 3,0\sigma_p$	$6,0\sigma_p$
%Diferença	$30\sigma_p/X$ a $50\sigma_p/X$	$\pm 200\sigma_p/X$	$\pm 300\sigma_p/X$	$600\sigma_p$
Z-escore	0,3 a 0,5	± 2,0	± 3,0	6,0

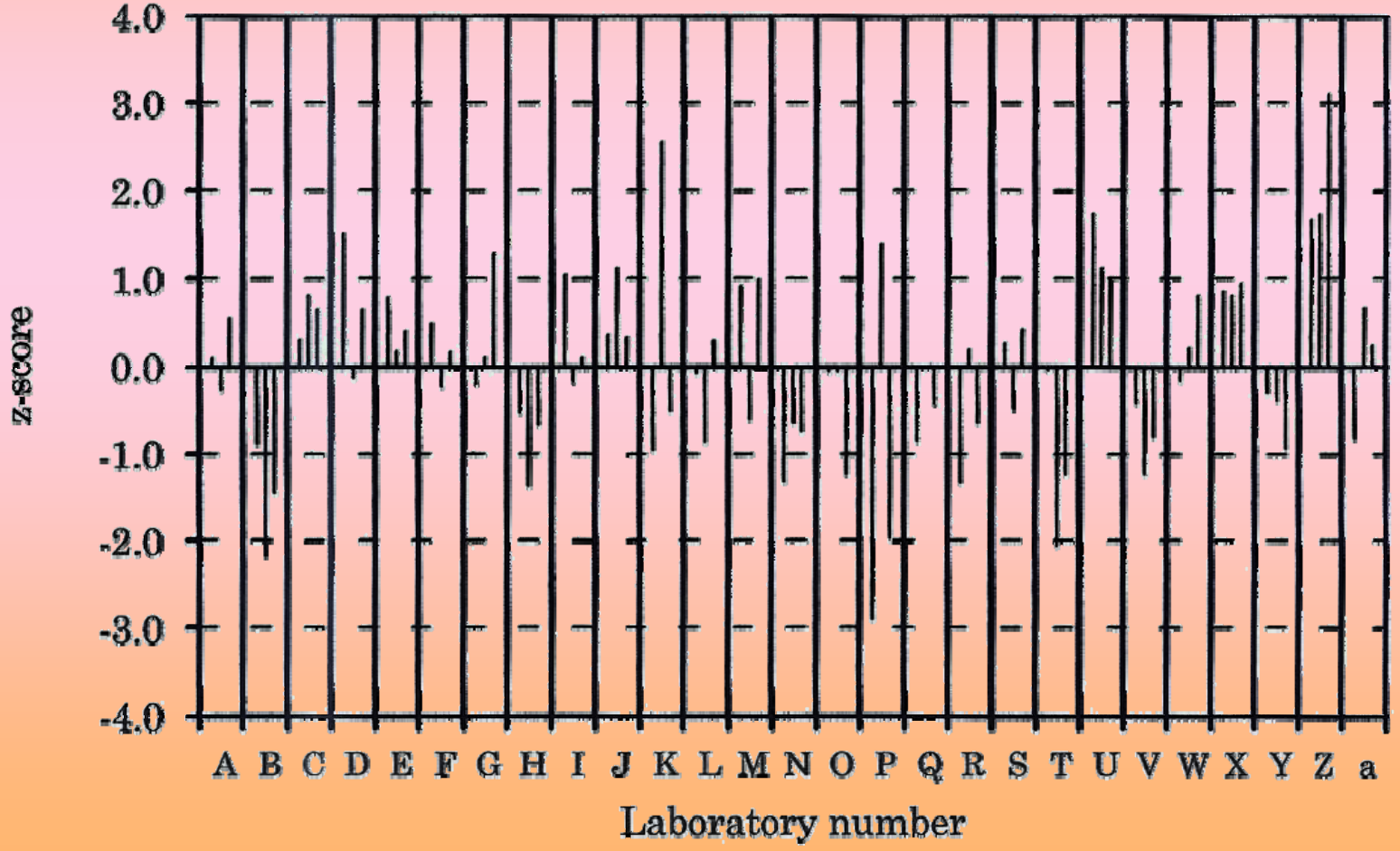
Métodos gráficos para combinação de índices de desempenho

Histogramas de índices de desempenho



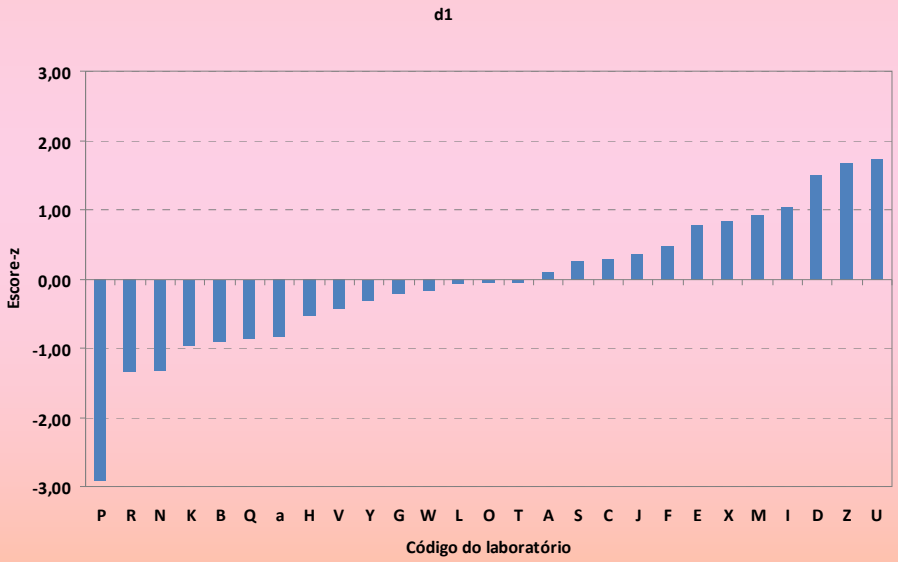
Métodos gráficos para combinação de índices de desempenho

Gráfico de barras

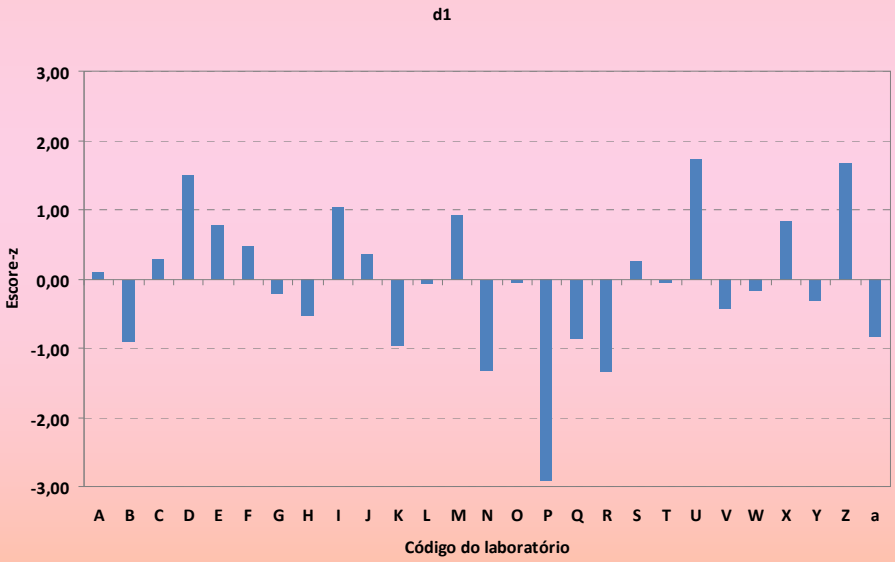


Métodos gráficos para combinação de índices de desempenho

Gráfico de barras



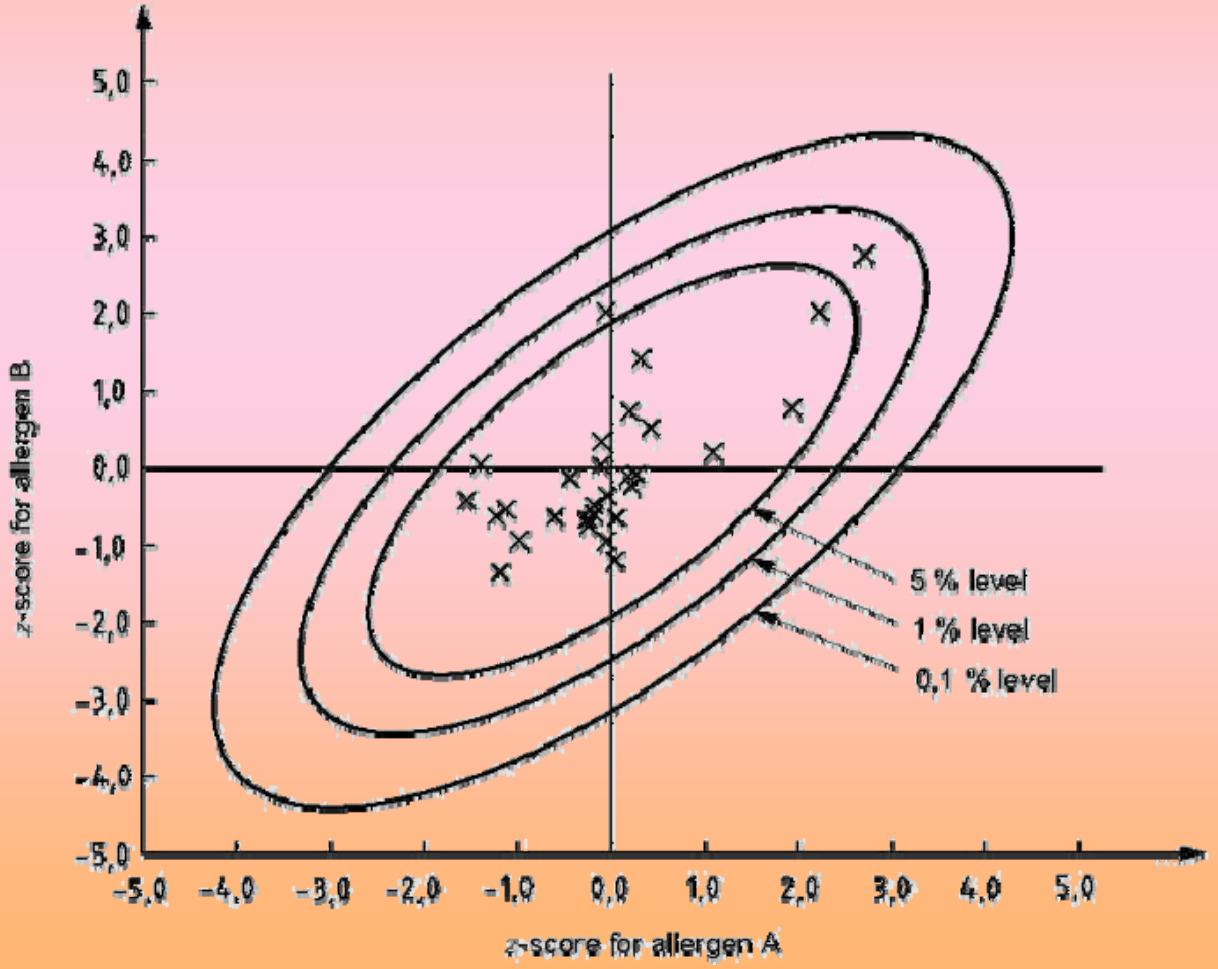
Valores de z ordenados



Ordem aleatória dos laboratórios

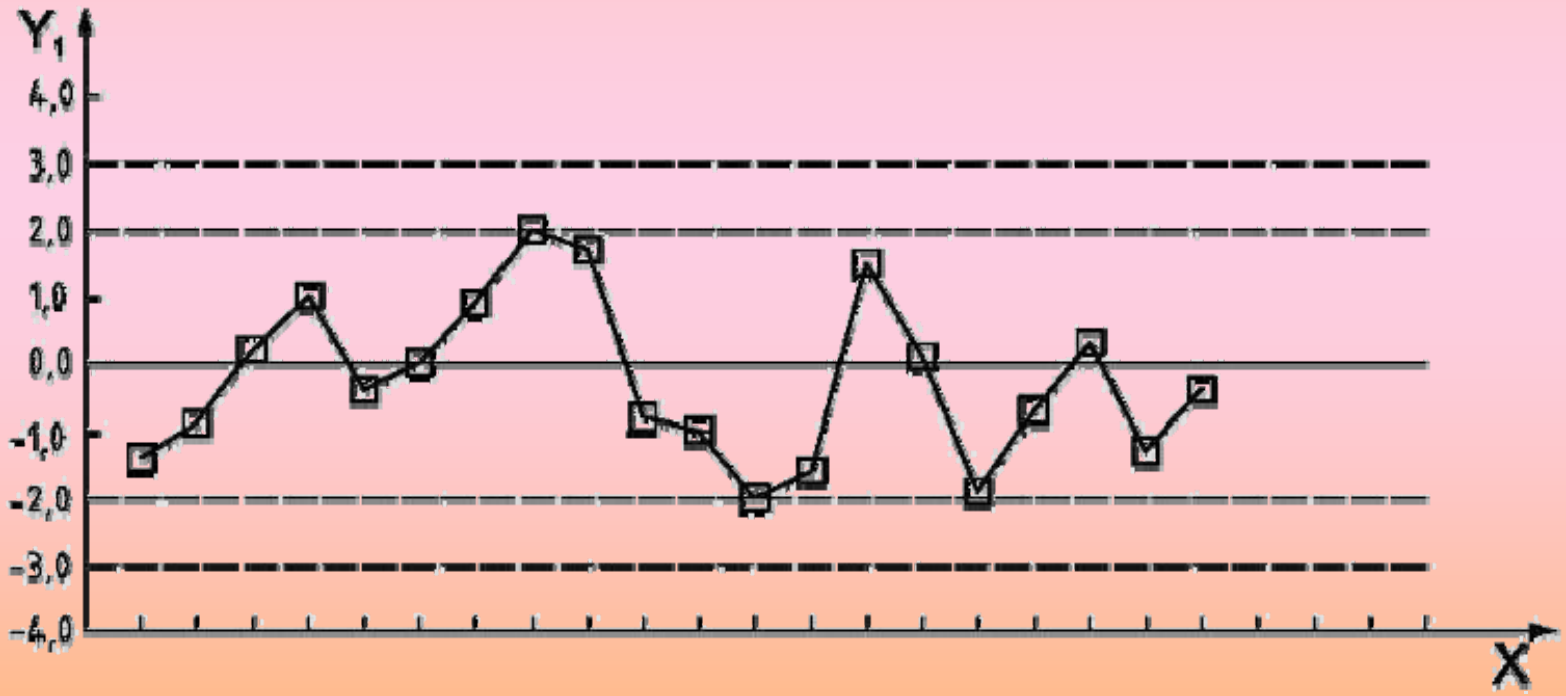
Métodos gráficos para combinação de índices de desempenho

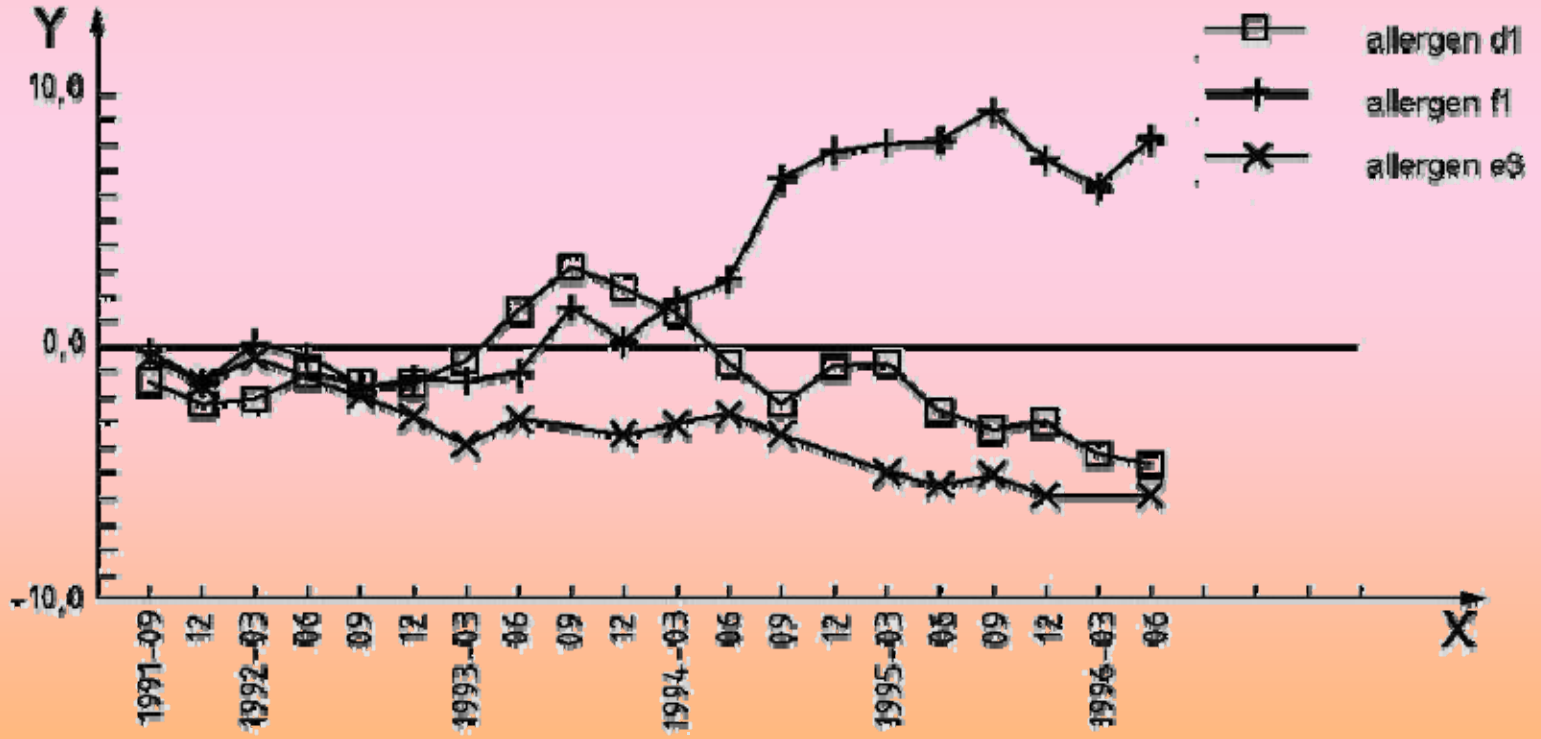
Gráfico de Youden



Métodos gráficos para combinação de índices de desempenho

Carta controle de Shewhart para escore z



Métodos gráficos para combinação de índices de desempenho

Carta controle de soma cumulativa para escore z