

# Uma alternativa para avaliar discordância entre duas medidas via modelo de regressão linear simples sem intercepto

## An alternative to evaluate disagreement between two measures via simple linear regression model without intercept

Maria Clara Vieira Borba<sup>1</sup>; Eduardo Yoshio Nakano<sup>2</sup>

### Resumo

Este artigo propõe um método estatístico para avaliar a discordância entre duas medidas quantitativas, quando as mesmas seguem uma distribuição normal. Este método consiste em realizar um teste de hipóteses no coeficiente de um modelo de regressão simples sem intercepto. Para ilustrar esse método e comparar com outros métodos que são erroneamente utilizados (tais como o teste t-pareado ou correlação de Pearson), foram utilizados dados artificiais da literatura e dados simulados. Cabe ressaltar que, enquanto os Coeficientes de Concordância de Kendall e o Coeficiente de Correlação Intraclasse medem a concordância entre duas medidas quantitativas, este novo método é uma alternativa para situações cujo interesse é comprovar que duas medidas são discordantes. Este trabalho também apresenta a implementação computacional do método no software livre R. Por último, foi mostrado que o método proposto cumpriu o seu papel, mesmo quando os outros procedimentos (erroneamente utilizados) falharam.

**Palavras-chave:** Concordância. Discordância. Teste-reteste. Confiabilidade.

### Abstract

This paper proposes a statistical method to assess the disagreement between two normally distributed measures. This method consists to test the coefficient of a simple linear regression model without intercept. Artificial dataset from literature and simulated data were used to illustrate this method and compare it to other methods that are wrongly used (such as the paired t-test or the Pearson correlation coefficient). It is noteworthy that, whereas Kendall's coefficient of concordance and the Intraclass correlation coefficient assess agreement between two quantitative measurements, this new method is an alternative to test the disagreement between these two measurements. This work also presents a computational implementation of the method in the free software R. Finally, this work showed that the proposed method fulfilled its role, even when other procedures (wrongly used) have failed.

**Keywords:** Concordance. Disagreement. Test-retest. Reliability.

<sup>1</sup> Mestranda em Estatística, Universidade de Brasília; mariaclara.vborba@gmail.com

<sup>2</sup> Professor Doutor do Departamento de Estatística da Universidade de Brasília; nakano@unb.br

## Introdução

Em muitas situações, deseja-se avaliar a concordância entre dois ou mais métodos que deveriam medir a mesma quantidade. Isto pode ser feito através de testes estatísticos de concordância. Em pesquisas biomédicas, por exemplo, no qual existem diferentes formas de avaliar uma determinada medida clínica, buscam-se sempre novos métodos menos invasivos, com menores efeitos colaterais e que sejam mais simples ou mais baratos do que os métodos já consagrados.

Tem-se como objetivo, então, verificar se esses novos métodos são equivalentes a ponto de se poder substituir um pelo outro. Para que isso ocorra, os dois métodos em comparação devem apresentar um erro que não tenha impacto significativo para que sejam permutáveis. Os testes de concordância também podem ser aplicados em estudos onde os níveis de confiabilidade teste-reteste são avaliados, em que uma mesma medida é realizada duas ou mais vezes com um intervalo entre as aplicações. Existem diversos métodos de se avaliar a concordância entre duas medidas. Pode-se citar, por exemplo, o Coeficiente de Concordância de Kendall (CONOVER, 1999) e o Coeficiente de Correlação Intraclasse (HAGGARD, 1958) no caso de medida quantitativas e os testes de Kendall Tau-a, Goodman Kruskal – Gamma e o D de Somer (CIRILLO; FERREIRA; LIMA, 2003) quando as medidas são dicotômicas.

Por outro lado, existem situações em que o interesse é verificar se duas medidas quantitativas são discordantes. Nessa situação, o Coeficiente de Concordância de Kendall e o Coeficiente de Correlação Intraclasse não são capazes de apresentar um resultado satisfatório. Isso ocorre porque ambos os testes têm como hipótese nula a discordância das observações e, em um teste estatístico de significância, a não rejeição da hipótese nula não implica a sua aceitação (MAGALHÃES; LIMA, 2015).

Altman e Bland (1983) mostraram casos em que técnicas como: o teste de comparação de médias, o coeficiente de correlação e a técnica de regressão

linear simples eram usadas de forma inadequada para a verificação da concordância. De fato, o método estatístico correto para a resolução desse problema não é óbvio. Segundo Bland e Altman (1986), muitos estudos realizam erroneamente um teste t pareado de comparação de médias para verificar a discordância entre duas medidas. No entanto, duas medidas que apresentam viés médio (vício ou diferença média) igual a zero não serão necessariamente concordantes, visto que os mesmos podem apresentar erros grandes.

Neste contexto, este trabalho apresenta um procedimento alternativo, a partir de um modelo de regressão linear simples sem intercepto, que permitirá avaliar a discordância entre duas medidas quantitativas. As principais diferenças da técnica proposta em relação aos testes usados erroneamente são apresentadas a seguir. Além disso, a mesma será ilustrada com dados fictícios obtidos de Hirakata e Camey (2009) e dados simulados. Toda simulação e análise serão realizadas através do software livre R (R CORE TEAM, 2015).

## Metodologia

Considere dois métodos que deveriam medir uma mesma quantidade. Sem perda de generalidade, considere que  $X$  é o valor observado pelo método padrão e  $Y$  a medida feita por outro método alternativo. A proposta deste trabalho é verificar a discordância entre duas medidas,  $X$  e  $Y$ , através de um modelo de regressão linear simples sem intercepto. A técnica de regressão linear simples avalia uma possível relação linear entre duas variáveis, isto é, se a relação entre essas duas variáveis pode ser representada por uma linha reta.

A equação de um modelo de regressão simples é dada por (NETER, 1996):

$$Y = \beta_0 + \beta_1 X + \varepsilon \quad (1)$$

em que  $\beta_0$  e  $\beta_1$  são, respectivamente, o intercepto e o coeficiente angular da reta e  $\varepsilon$  é o erro aleatório com distribuição Normal.

As hipóteses de interesse são:

$H_0$ : Os métodos concordam entre si;

$H_a$ : Os métodos discordam entre si.

$$EP(\hat{\beta}_1) = \sqrt{\frac{\sum_{i=1}^n [(y_i - M) - \hat{\beta}_1(x_i - M)]^2}{(n-1)\sum_{i=1}^n (x_i - M)^2}} \quad (3)$$

Diz-se que os dois métodos são equivalentes (ou concordantes) se apresentarem valores similares, isto é, quando  $X=Y$  ou, de forma equivalente, quando  $D=X-Y=0$ . Note que, sob a hipótese de concordância e supondo que ambas as médias de  $X$  e  $Y$  são iguais à zero, a reta de regressão resultante passa pela origem e tem coeficiente angular igual a 1. Assim, a concordância dessas duas medidas é representada pela equação (1) quando  $\beta_0=0$  e  $\beta_1=1$ . Neste contexto, uma forma alternativa de verificar a discordância entre as medidas  $X$  e  $Y$  é ajustar um modelo de regressão linear simples sem intercepto ( $\beta_0=0$ ) e testar a hipótese do coeficiente angular da reta ser diferente de um ( $\beta_1 \neq 1$ ). Assim, as hipóteses de interesse são descritas, em termos estatísticos, por:

$$H_0: \beta_1=1 \quad \text{vs.} \quad H_a: \beta_1 \neq 1$$

em que  $\beta_1$  é o coeficiente angular de uma reta de regressão sem intercepto.

Situações em que os valores de  $X$  e  $Y$  não apresentam média zero podem ser facilmente adaptadas. Segundo Altman e Bland (1983), como não se sabe o verdadeiro valor da quantidade de interesse, então a melhor estimativa que se tem é a média dos valores de  $X$  e  $Y$ . Assim, o procedimento se resume em subtrair os valores de  $X$  e  $Y$  por  $M$ , que representa a média de todos os valores de  $(X+Y)/2$ . Dessa forma, o coeficiente  $\beta_1$  do modelo de regressão sem intercepto e seu erro padrão podem ser estimados, respectivamente, por

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - M)(y_i - M)}{\sum_{i=1}^n (x_i - M)^2} \quad (2)$$

e

em que  $M = \frac{1}{n} \sum_{i=1}^n \frac{x_i + y_i}{2}$ ,  $x_i$  é a  $i$ -ésima observação segundo o método padrão,  $y_i$  é a  $i$ -ésima observação segundo o método alternativo e  $n$  é o tamanho da amostra.

Assim, a estatística do teste  $H_0: \beta_1=1$  é dada por (NETER, 1996):

$$t = \frac{\hat{\beta}_1 - 1}{EP(\hat{\beta}_1)} \quad (4)$$

que segue uma distribuição  $t$ -Student com  $(n-1)$  graus de liberdade. Assim, o valor-p do teste é dado por:

$$\text{valor-p} = P(t_{(n-1)} > |t|). \quad (5)$$

Os valores de  $\hat{\beta}_1$ ,  $EP(\hat{\beta}_1)$ , da estatística do teste  $t$  e do valor-p do teste de concordância, expressos, respectivamente, por (2), (3), (4) e (5), podem ser facilmente obtidos através do comando *lm* do software livre R (R CORE TEAM, 2015).

Note que quando  $X$  e  $Y$  possuem distribuição Normal, o estimador (2) é equivalente ao Estimador de Mínimos Quadrados de  $\beta_1$  (NETER, 1996). Assim, a reta de regressão estimada é aquela que passa pela origem e minimiza o quadrado das distâncias verticais de cada ponto a ela. Note que a distância horizontal de cada ponto com a reta de regressão também implica em erro de medição entre as medidas  $X$  e  $Y$ . Neste contexto, as distâncias horizontais dos

pontos à reta de regressão (que, sob a hipótese de concordância, é a reta  $Y = X$ ) também devem ser consideradas em um teste de concordância.

A proposta deste trabalho para resolver esse problema é inverter as variáveis  $X$  e  $Y$  (isto é, considerar  $X$  como resposta e  $Y$  como variável explicativa), ajustar o modelo de regressão sem intercepto e calcular o valor-p dado por (5). A técnica de regressão linear pode ser aplicada normalmente com as variáveis invertidas, visto que nos problemas de um teste de concordância ambas as medidas ( $X$  e  $Y$ ) são aleatórias. Note que esse procedimento é análogo à realização de uma regressão inversa sem intercepto (IYER, 2003).

Neste contexto, a concordância será rejeitada se pelo menos uma das regressões rejeitarem a hipótese  $H_0: \beta_1=1$ . Assim, o valor-p do teste de concordância será o valor mínimo dos dois valores-p obtidos.

O Quadro 1 apresenta uma sintaxe de uma função do R (R CORE TEAM, 2015) para a realização do teste de concordância proposto neste trabalho.

**Quadro 1** - Script no R para realizar o teste de concordância.

```
disagree <- function(x,y) {
  disagreeel <- function(x,y) {
    n <- length(x) #tamanho da amostra
    M <- mean((x+y)/2)
    X <- x-M
    Y <- y-M
    est.beta <- sum(X*Y)/sum(X^2)
    sd.est.beta <- sqrt(
      sum((Y-est.beta*X)^2)/(n-1)/sum(X^2)
    )
    t <- (est.beta-1)/sd.est.beta
    df <- n-1
    pvalue <- 2*pt(-abs(t),df=n-1)
    return(round(c(Beta.l=est.beta,EP=sd.est.beta,
      t=t,df=df,Valor.p=pvalue),5))
  }
  result1 <- disagreeel(x,y)
  result2 <- disagreeel(y,x)
  Valor.p <- min(result1[5],result2[5])
  plot(x,y,xlim=c(min(x,y),max(x,y)),ylim=c(min(x,y),
    max(x,y)),pch=16)
  abline(coef=c(0,1),lty=2)
  output <- cbind(rbind(result1[1:4],
    result2[1:4]),Valor.p)
  rownames(output) <- c("(X,Y)","(Y,X)")
  return(output)
}

### Exemplo
x<-rnorm(30,1,2)
y<-x+rnorm(30,0,1)
disagree(y,x)
```

### *Diferença da metodologia proposta com os testes de concordância tradicionais*

Os testes de concordância existentes, como o Coeficiente de Concordância de Kendall (CONOVER, 1999) e o Coeficiente de Correção Intraclasse (HAGGARD, 1958) definem como hipótese nula a discordância entre duas medidas. Já o teste proposto neste trabalho, que é baseado em uma regressão linear simples sem intercepto, considera como hipótese nula a concordância entre os dois métodos avaliados. Desta forma, a principal diferença entre o método proposto em relação àqueles já existentes está no tipo de erro que se deseja controlar. Note que um teste de hipóteses estatístico controla apenas o erro Tipo I (rejeitar a hipótese nula quando ela é verdadeira), cuja probabilidade é chamada de Nível de Significância do teste (MAGALHÃES e LIMA, 2015). A Tabela 1 apresenta os possíveis resultados ao se realizar um teste para verificar a concordância entre duas medidas,  $X$  e  $Y$ .

**Tabela 1.** Possíveis resultados ao realizar um teste de concordância.

		Decisão	
		X e Y concordantes	X e Y discordantes
Real	X e Y concordantes	Decisão correta	Erro B
	X e Y discordantes	Erro A	Decisão correta

Assim, em qualquer teste de concordância, dois possíveis erros podem ser cometidos: concluir que os métodos são concordantes quando, na realidade, eles são discordantes (Erro A) ou concluir que os métodos são discordantes quando, na realidade, eles são concordantes (Erro B). Se em um determinado

problema deseja-se controlar o Erro A (situação em que o Erro A é o mais grave) então o Coeficiente de Concordância de Kendall ou o Coeficiente de Correlação Intraclasse pode ser utilizado. No entanto, se o desejo é controlar o Erro B (situação em que o Erro B é o mais grave) então o teste proposto neste trabalho é o mais indicado.

### Resultados e discussões

A metodologia apresentada neste trabalho foi ilustrada por um conjunto de dados hipotéticos de Hirakata e Camey (2009) que estimou o tempo (em semanas) de gestação de 50 mulheres através da data da última menstruação (DUM) e pelo ultrassom. A Tabela 2 apresenta os valores observados com as seguintes variáveis:

**id:** identificação da mulher;

**DUM:** número de semanas de gestação a partir da DUM;

**U1:** número de semanas de gestação calculado a partir do ultrassom obtido após a 20ª semana de gestação;

**U2:** número de semanas de gestação calculado a partir do ultrassom obtido antes da 20ª semana de gestação;

**X1:** número de semanas de gestação calculado a partir do ultrassom obtido após a 20ª semana de gestação (U1) com erro não sistemático que cresce de acordo com os valores de U1.

**X2:** valores transformados de U1 para diminuir o erro em 80% entre U1 e DUM;

**X3:** valores transformados de U1 para aumentar o erro em 100% entre U1 e DUM;

**X4:** valor médio de DUM somado com um erro aleatório.

**Tabela 2.** Dados artificiais do tempo de gestação (em semanas) de 50 mulheres.

id	DUM	U1	U2	X1*	X2*	X3*	X4*
1	36,7	37,0	37,8	35,8	36,8	37,3	38,8
2	37,6	37,8	38,6	37,6	37,6	38,0	38,6
3	39,1	39,5	40,9	40,8	39,2	39,9	39,2
4	39,8	39,4	41,1	40,6	39,7	39,0	36,8
5	38,3	38,4	40,3	38,9	38,3	38,5	38,2
6	37,0	36,6	37,5	35,7	36,9	36,2	36,7
7	39,0	38,7	39,5	39,5	38,9	38,4	37,0
8	39,1	39,3	40,2	40,4	39,1	39,5	38,5
9	38,6	38,6	39,2	39,3	38,6	38,6	37,9
10	37,5	37,5	38,7	37,0	37,5	37,5	37,9
11	39,3	39,4	41,1	40,6	39,3	39,5	38,2
12	38,2	37,8	39,7	37,6	38,1	37,4	36,6
13	36,3	36,1	38,1	35,0	36,3	35,9	37,2
14	38,8	38,5	39,1	39,1	38,7	38,2	37,0
15	38,1	38,5	39,5	39,2	38,2	38,9	39,0
16	38,0	38,0	39,6	38,0	38,0	38,0	37,8
17	38,0	38,1	38,1	38,3	38,0	38,2	38,1
18	37,8	37,6	38,9	37,2	37,8	37,4	37,2
19	36,6	36,4	37,6	35,3	36,6	36,2	37,2
20	36,4	36,8	36,7	36,0	36,5	37,2	39,1
21	37,9	37,6	38,0	37,2	37,8	37,3	37,0
22	39,4	39,2	40,4	40,2	39,4	39,0	37,3
23	37,3	37,5	38,8	37,0	37,3	37,7	38,5
24	37,0	36,7	37,7	35,9	36,9	36,4	37,0
25	38,1	38,2	38,5	38,5	38,1	38,3	38,2
26	36,4	36,5	37,9	35,5	36,4	36,6	38,2
27	38,0	38,3	38,4	38,6	38,1	38,6	38,9
28	38,8	38,6	39,9	39,4	38,8	38,4	37,4
29	37,2	37,1	37,7	36,4	37,2	37,0	37,6
30	37,9	38,3	39,7	38,7	38,0	38,7	39,2
31	37,7	37,4	38,9	36,8	37,6	37,1	37,1
32	38,2	38,0	38,4	38,1	38,2	37,8	37,4
33	36,1	36,4	37,4	35,4	36,2	36,7	38,9
34	39,0	39,3	40,9	40,4	39,1	39,6	39,0
35	37,6	37,9	38,4	37,8	37,7	38,2	38,9
36	38,1	38,0	39,5	38,1	38,1	37,9	37,7
37	38,9	39,1	40,7	40,0	38,9	39,3	38,6
38	38,1	38,3	39,0	38,7	38,1	38,5	38,6
39	38,0	37,8	38,0	37,7	38,0	37,6	37,4
40	38,2	37,9	38,5	37,9	38,1	37,6	37,1
41	38,6	38,7	39,7	39,6	38,6	38,8	38,2
42	38,9	39,1	40,1	40,1	38,9	39,3	38,5
43	37,8	37,7	38,0	37,4	37,8	37,6	37,4
44	38,4	38,3	38,8	38,8	38,4	38,2	37,4
45	37,2	36,9	37,5	36,2	37,1	36,6	36,7
46	37,6	37,3	38,0	36,7	37,5	37,0	36,8
47	38,1	38,1	38,2	38,3	38,1	38,1	37,7
48	36,7	36,5	36,9	35,6	36,7	36,3	37,0
49	37,8	37,6	39,5	37,3	37,8	37,4	37,4
50	38,2	38,4	39,2	39,0	38,2	38,6	38,8

\*As variáveis X1 a X4 foram geradas pelos autores.  
Fonte: Hirakata e Camey (2009).

A variável número de semanas de gestação estimada pela última data da menstruação (DUM) foi confrontada com as variáveis U1, U2 e X1 a X4 pelos testes t pareado, correlação de Pearson e o teste de concordância proposto. Os resultados são apresentados na Tabela 3.

**Tabela 3.** Resultados dos testes de comparação da variável DUM com as variáveis U1, U2, X1, X2, X3 e X4.

DUM vs.	Diferença Média* (DP)	Valor p t- pareado	Correl. Pearson (valor p)	Valor p Teste Proposto
U1	0,014 (0,248)	0,692	0,961 (0,000)	0,069
U2	0,948 (0,588)	0,000	0,854 (0,000)	0,000
X1	0,036 (0,827)	0,760	0,958 (0,000)	0,000
X2	0,004 (0,064)	0,659	0,997 (0,000)	0,353
X3	0,028 (0,497)	0,692	0,866 (0,000)	0,000
X4	0,090 (1,161)	0,586	0,022 (0,878)	0,000

\* Diferença Média ou Vício.

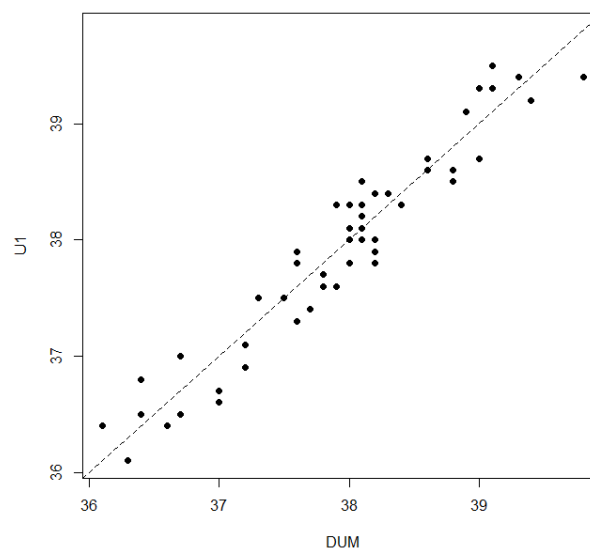
Fonte: Os autores.

três últimas colunas apresentam os resultados dos valores de p, para a verificação do vício (pelo teste t-pareado), correlação (pelo teste da correlação de Pearson) e discordância (pelo novo método proposto), respectivamente. Quando o valor p para o teste do vício for maior ou igual a 0,05, o vício entre as duas medidas será chamado de “vício zero” ou “vício nulo” e quando o valor p deste mesmo teste for menor que 0,05, o vício entre as duas medidas será chamado de “vício constante”. Quando o valor p para o teste de correlação for maior ou igual a 0,05, denomina-se por “medidas não correlacionadas” e quando o valor p deste mesmo teste for menor que 0,05, denomina-se por “medidas correlacionadas”. Quando o valor p para o teste proposto for maior ou igual a 0,05, denomina-se por “não discordantes” ou “erro moderado” e quando o valor p deste mesmo teste for menor que 0,05, denomina-se “discordantes”.

#### *Medidas correlacionadas, com vício zero e não discordantes*

A Figura 1 apresenta o gráfico de dispersão para as variáveis DUM e U1. Pode-se verificar que não há evidência de discordância entre os valores. Pela Tabela 3, é possível ver que o vício não é significativo ( $p=0,692$ ), a correlação é significativa ( $p=0,000$ ) e o teste proposto não rejeita a concordância entre as medidas ( $p=0,069$ ).

**Figura 1** - Gráfico de dispersão para as variáveis DUM e U1.

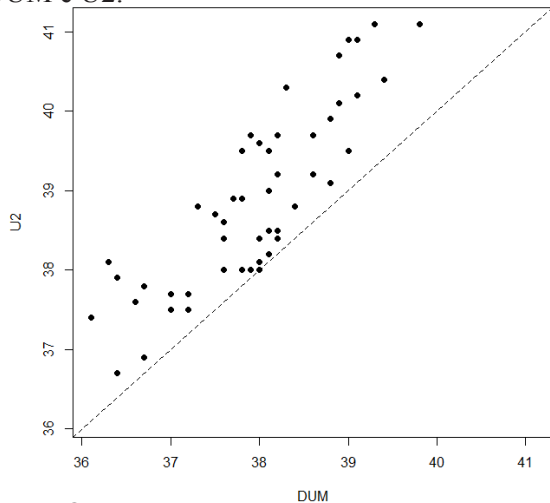


Fonte: Os autores.

#### *Medidas correlacionadas, com vício constante e discordantes*

A Figura 2 apresenta o gráfico de dispersão para as variáveis DUM e U2. Pode-se verificar que há discordância entre os valores. Pela Tabela 3, é possível ver que o vício é significativo ( $p=0,000$ ), a correlação é significativa ( $p=0,000$ ) e o teste proposto rejeita a concordância entre as medidas ( $p=0,000$ ). Veja que este é um exemplo que mostra que a presença de correlação não implica concordância. Como esperado, medidas com vício constante implica discordância e o teste proposto foi capaz de detectar essa discordância.

**Figura 2** - Gráfico de dispersão para as variáveis DUM e U2.

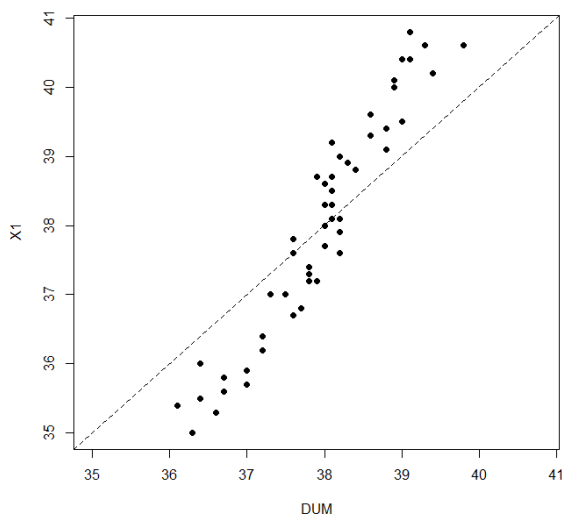


**Fonte:** Os autores.

*Medidas correlacionadas, com vício nulo e discordantes*

A Figura 3 apresenta o gráfico de dispersão para as variáveis DUM e X1. Pode-se verificar que há discordância entre os valores. Pela Tabela 3, é possível ver que o vício não é significativo ( $p=0,760$ ), a correlação é significativa ( $p=0,000$ ) e o teste proposto rejeita a concordância entre as medidas ( $p=0,000$ ). Veja que este é um exemplo que mostra que a presença de correlação, mesmo com vício nulo ainda não garante que as medidas são concordantes.

**Figura 3** - Gráfico de dispersão para as variáveis DUM e X1.

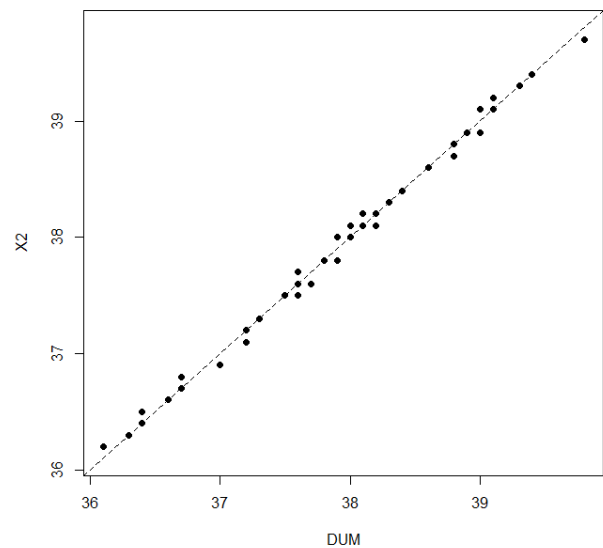


**Fonte:** Os autores.

*Medidas correlacionadas, com vício nulo e erro pequeno (não discordantes)*

A Figura 4 apresenta o gráfico de dispersão para as variáveis DUM e X2. Pode-se verificar que não há discordância entre os valores. Pela Tabela 3, é possível ver que o vício não é significativo ( $p=0,659$ ), a correlação é significativa ( $p=0,000$ ) e o teste proposto não rejeita a concordância entre as medidas ( $p=0,353$ ).

**Figura 4** - Gráfico de dispersão para as variáveis DUM e X2.

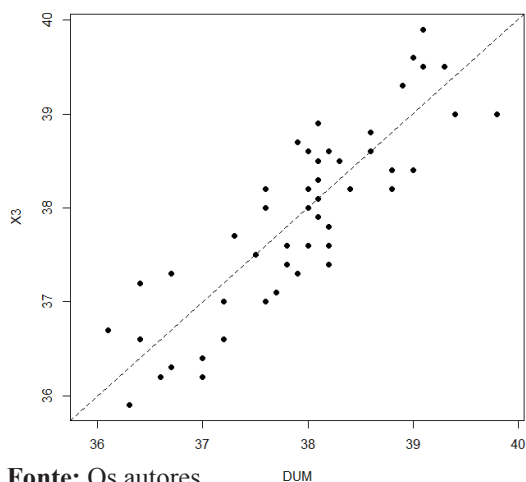


**Fonte:** Os autores.

*Medidas correlacionadas, com vício nulo e erro moderado (discordantes)*

A Figura 5 apresenta o gráfico de dispersão para as variáveis DUM e X3. Pode-se verificar que há discordância entre os valores (erro significativo). Pela Tabela 3, é possível ver que o vício não é significativo ( $p=0,692$ ), a correlação é significativa ( $p=0,000$ ) e o teste proposto rejeita a concordância entre as medidas ( $p=0,000$ ). Veja que este é outro exemplo que mostra a presença de correlação, mesmo com a ausência de vício não implica concordância. Nesta situação, o teste proposto detectou as diferenças (discordâncias) entre as duas medidas, algo que o teste t e o coeficiente de correlação não foram capazes de detectar.

**Figura 5** - Gráfico de dispersão para as variáveis DUM e X3.

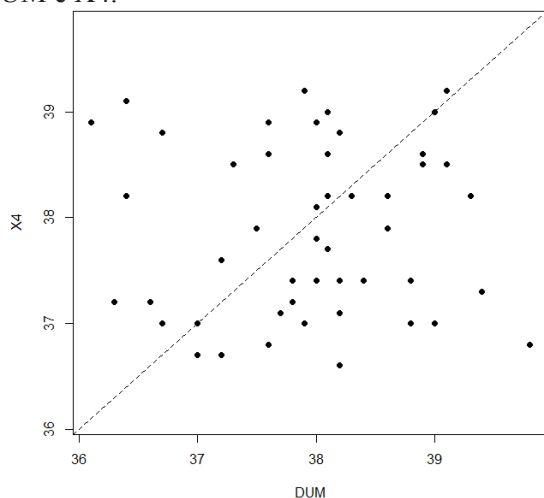


Fonte: Os autores.

*Medidas não correlacionadas, com vício nulo e erro grande (discordantes)*

A Figura 6 apresenta o gráfico de dispersão para as variáveis DUM e X4. Pode-se verificar que há discordância entre os valores (erro significativo). Pela Tabela 3, é possível ver que o vício não é significativo ( $p=0,586$ ), a correlação também não é significativa ( $p=0,878$ ) e o teste proposto rejeita a concordância entre as medidas ( $p=0,000$ ). Como esperado, a ausência de correlação implica discordância e o teste proposto foi capaz de detectar essa discordância.

**Figura 6** - Gráfico de dispersão para as variáveis DUM e X4.



Fonte: Os autores.

Note que as Figuras 1, 4, 5 e 6 são situações em que as medidas apresentam vício nulo, variando apenas o erro entre as duas medidas consideradas. Considerando um nível de significância de 5%, a concordância é rejeitada nas situações descritas pelas Figuras 5 e 6, que apresentam os maiores erros (pontos mais distantes da reta  $X=Y$ ). Note ainda que, apesar de não rejeitar a concordância nas situações descritas pelas Figuras 1 e 4, o teste proposto apresenta boa sensibilidade, considerando os erros entre as medidas, o que pode ser observado através do valor-p do teste. Note que o teste t pareado e o coeficiente de correlação não apresentam essa sensibilidade, apresentando resultados praticamente iguais nas situações descritas pelas Figuras 1 e 4, ignorando o maior erro de medida apresentado na situação descrita pela Figura 1.

## Conclusões

Quando o objetivo é analisar a concordância de duas medidas quantitativas, Altman e Bland (1983) descreveram exemplos mostrando que a comparação de médias através de um teste t pareado, a correlação de Pearson ou a regressão linear simples são procedimentos inadequados e sugerem o uso do gráfico denominado Bland-Altman.

Este trabalho propôs um procedimento alternativo para a verificação da discordância de duas medidas, quando as mesmas seguem uma distribuição Normal, através de um teste de hipóteses do coeficiente angular de uma regressão linear simples sem intercepto.

Os resultados obtidos mostraram que o procedimento proposto é útil para verificar se duas medidas são discordantes. Além disso, esse teste não compete com os testes de concordância usualmente utilizados, como o Coeficiente de Concordância de Kendall (CONOVER, 1999) e o Coeficiente de Correlação Intraclasse (HAGGARD, 1958), pois o mesmo tem um propósito diferente. A diferença entre esses testes está no tipo de erro que se deseja controlar. Se o objetivo é confirmar que duas medidas



são concordantes, então o Coeficiente de Correlação Intraclasse ou o teste de concordância de Kendall pode ser utilizado. No entanto, situações cujo interesse é comprovar que duas medidas são discordantes, o teste proposto neste trabalho deve ser utilizado. Neste contexto, o teste proposto neste trabalho pode ser visto como um Teste de Discordância.

Os resultados obtidos nas ilustrações mostraram que o teste de discordância proposto cumpre o seu papel mesmo quando outros procedimentos (erroneamente utilizados) falham. Além disso, com o algoritmo disponibilizado neste trabalho em um software livre, qualquer pesquisador poderá realizar esse novo teste de discordância sem custo algum.

### Agradecimentos

Os autores agradecem aos revisores anônimos, cujas sugestões contribuíram muito para melhora deste artigo.

### Referências

ALTMAN, D. G.; BLAND, J. M. Measurement in medicine: the analysis of method comparison studies. *Journal of the Royal Statistical Society. Series D (The Statistician)*, v. 32, n. 3, p. 307-317, 1983.

BLAND, J. M.; ALTMAN, D. G. Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*, n. 1, p. 307-310, 1986.

CIRILLO, M. A.; FERREIRA, D. F.; LIMA, P. S. G. Uma alternativa para avaliar a eficiência de meios de cultivo in vitro via modelo de regressão logística. *Semina: Ciências Exatas e Tecnológicas*, v. 24, n. 1, p. 101-110, 2013.

CONOVER, W. J. *Practical nonparametric statistics*. 3. ed. New York: J. Wiley, 1999.

HAGGARD, E. A. *Intraclass correlation and the analysis of variance*. New York: Dryden Press, 1958.

HIRAKATA, V. N.; CAMEY, S. A. Análise de concordância entre métodos de Bland-Altman. *Rev. HCPA*, v. 29, n. 3, p. 261-268, 2009.

IYER, H. Statistical calibration and measurements. In: R. KHATTREE, R.; RAO, C.R. (Ed.). *Handbook of statistics 22*. Amsterdam: Elsevier Science B.V., 2003. Chapter 20.

MAGALHÃES, M. N.; LIMA, A. C. P. *Noções de probabilidade e estatística*. 7. ed. São Paulo: EDUSP, 2015.

NETER, J.; KUTNER, M. H.; NACHTSHEIM, C. J.; WASSERMAN, W. *Applied linear statistical models*. 4. ed. Boston: McGraw-Hill/Irwin, 1996.

R CORE TEAM R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2015. Disponível em: <<http://www.R-project.org>>. Acesso em: 13 jun.2016.

Recebido em 20 Novembro, 2015 - Received on November 20, 2015.

Aceito em 4 Abril, 2016 - Accepted on April 4, 2016.

