

A SENSIBILIDADE DOS TESTES PARAMÉTRICOS VS ROBUSTEZ DOS NÃO PARAMÉTRICOS E AS DIFERENÇAS NA COMPARAÇÃO DE PARÂMETROS DISCRETOS DO MOVIMENTO DO CG EM DIFERENTES CAE'S

C. M.B. Rodrigues *, M.V. Correia *, J. M.C.S. Abrantes**, J. Nadal*** e M.A.B. Rodrigues****

* INESC TEC (anteriormente INESC Porto) e
Faculdade de Engenharia da Universidade do Porto, Portugal

** MovLab - Laboratório de Tecnologias e Interfaces /

Universidade Lusófona de Humanidades e Tecnologias, Lisboa, Portugal

*** Programa de Engenharia Biomédica - COPPE/UFRJ, Rio de Janeiro, Brasil

**** Departamento de Eletrônica e Sistemas/Universidade Federal de Pernambuco, Recife, Brasil

e-mail: c.rodrigues@fe.up.pt

Resumo: A comparação dos resultados sob diferentes condições experimentais sempre requer escolher o teste estatístico mais adequado. A verificação das condições tais como o tamanho da amostra, normalidade, independência, emparelhamento e homogeneidade de variância é essencial para a escolha do método estatístico mais adequado, bem como os objetivos a serem alcançados com a aplicação destes testes. Para avaliar o efeito da aplicação de testes paramétricos (P) e não paramétricos (NP) em condições hostis, foi realizada uma análise com base em uma amostra pequena ($n = 6$) de parâmetros discretos relativos ao movimento do CG em diferentes contramovimentos dos membros inferiores avaliados no salto vertical máximo padronizado. Os métodos P e NP foram aplicados com e sem satisfazer as condições de aplicação e os resultados foram comparados em um amplo conjunto de variáveis dinâmicas e cinemáticas, avaliando-se os casos e as condições em que os testes P e NP mostraram resultados semelhantes ou diferentes. A conclusão aponta para a utilidade da aplicação de teste P, mesmo quando as condições de aplicação não são verificadas, caso em que se recomenda a aplicação de testes NP para confirmação do resultado do teste P.

Palavras-chave: Testes estatísticos paramétricos, testes estatísticos não-paramétricos, robustez, sensibilidade.

Abstract: *The comparison of results under different experimental conditions always requires choosing the most suitable statistical test. The verification of conditions as the sample size, normality, independence, pairing, homogeneity of variances, among others are essential for choosing the best suited statistical method, but also the goals to be achieved with the application of these tests. To assess the best route of application of parametric (P) and non-parametric (NP) tests in hostile conditions, we performed an analysis based on a small sample ($n = 6$) of discrete parameters relating to the movement of the CG in different countermovement of the lower limbs evaluated in standardized maximum*

vertical jump. We applied P and NP methods in case the conditions of application were checked and where these conditions were not checked and compared the results in a comprehensive set of dynamic and kinematic variables, assessing the cases and conditions in which P and NP tests showed similar or different results. Our finding point towards the interest and utility of P test application even in cases where the application conditions are not checked, in which case it is recommended to apply NP testing for confirmation of P tests results.

Key-words: *Non-parametric statistical tests, parametric statistical tests, robustness, sensitivity.*

Introdução

Se por um lado a lei dos grandes números e a aproximação da média amostral pela distribuição normal são válidas de acordo com o teorema do limite central para amostras suficiente grandes [1], nem sempre se encontram disponíveis para estudo amostras suficientemente grandes de modo a garantir as condições de aproximação pela distribuição normal. Por outro lado ainda, a análise estatísticas de amostras de maior dimensão, nem sempre homogêneas, pode mascarar por via da compensação estatística das médias resultante da agregação, efeitos associados a estratégias individuais de movimento que se pretendem analisar.

Além disso, a análise estatística de grupos é normalmente utilizada para comparação entre grupos de sujeitos ou condições experimentais em que a média de cada classe agrupada é utilizada para representar cada classe e os resultados são em seguida generalizados a toda a classe, confundindo desta forma resultados gerais que seriam válidos para todos e cada um dos membros de uma classe, com resultados agregados que seriam válidos para a classe considerada de forma agregada [2]. A questão coloca-se desta forma entre as opções de obtenção do comportamento médio de uma classe e a

variação de cada sujeito em relação a este comportamento médio, ou a extração de características do comportamento individual de cada sujeito e a inferência de características comuns entre os sujeitos pertencentes à classe considerada. Para amostras homogêneas os resultados esperados através das opções anteriores seriam semelhantes, no entanto o movimento humano está permanentemente sujeito a um conjunto de restrições e variabilidade do sistema nervoso central no controle do movimento, pelo que o mesmo sujeito não é capaz de realizar o mesmo movimento sob as mesmas condições em duas tentativas consecutivas. A opção consiste por isso na utilização de estratégias alternativas complementares com vista à captação das características pretendidas no estudo em causa e a minimização do enviesamento de outros fatores.

O foco do presente trabalho consiste no estudo do ciclo de alongamento e encurtamento (CAE) muscular, o qual é um fenómeno essencial nos membros inferiores no ser humano que se manifesta quer na corrida e nos saltos, quer na marcha. Este ciclo se caracteriza por uma ação reversível de alongamento a que se segue um encurtamento muscular, normalmente associado a um contra-movimento, conduzindo a uma ação concêntrica muscular mais intensa após a ação excêntrica do que seria conseguida apenas em ação concêntrica. Na corrida e nos saltos o CAE tem normalmente como objetivo a maximização do desempenho, enquanto na marcha e em actividades sub-máximas o CAE tem como objetivo a antecipação de carga e a sincronização de ação muscular com vista à proteção da integridade estrutural relativamente ao impacto no contato [3]. Enquanto tarefa de controle de contato, a locomoção humana quer seja na corrida, saltos ou marcha depende em grande parte da força de reação do solo (FRS) durante a fase de contato, que atua conjuntamente com a ação da gravidade para produzir o movimento pretendido do centro de gravidade (CG). Assim, o sistema nervoso central (SNC) produz em cada instante a ativação muscular necessária com recurso nomeadamente ao CAE para em conjunto com as demais estruturas presentes, produzir as ações internas que em conjunto com as ações externas vão produzir a aceleração pretendida do CG. Assim, o SNC é responsável pelo controle da FRS, sendo esta considerada como um sinal de controle pelo SNC relativamente ao movimento pretendido do CG [4].

Nesta perspectiva, analisou-se a componente vertical da FRS durante a fase de contato em diferentes situações de contra-movimento (CM) dos membros inferiores, avaliados de acordo com Asmussen e Bonde-Petersen [5], em saltos de impulsão máxima vertical *Squat Jump* (SJ) sem CM, *Counter Movement Jump* (CMJ) com CM longo e *Drop Jump* (DJ) com CM curto, com vista à comparação dos testes estatísticos emparelhados, paramétricos e não-paramétricos e da avaliação da sensibilidade e robustez dos testes em análise de parâmetros discretos com amostras de dimensão reduzida relativas à força de reação solo enquanto ação responsável pelo movimento e da

aceleração produzida no CG enquanto consequência desta ação em diferentes CM's e CAE associados.

Materiais e métodos

Os dados analisados são relativos a uma amostra de seis sujeitos, estudantes universitários de educação física e desporto com aptidões semelhantes e sem treino prévio específico. Os sujeitos foram informados sobre o protocolo experimental tendo cada sujeito assinado uma declaração escrita de consentimento informado de acordo com a *World Medical Association Declaration of Helsinki*. Os sujeitos foram pesados ($76,7 \pm 6,7$ kg) e medida a sua estatura ($1,789 \pm 0,049$ m), tendo cada sujeito realizado um conjunto de três ensaios de acordo com o protocolo SJ, CMJ e DJ [5], dos quais foi selecionado o melhor ensaio para cada sujeito e tipo de salto com base no método do tempo de voo por ausência de contato com a plataforma de força. Durante os ensaios foram registados os dados das forças e momentos de força de contato com o solo através da plataforma de força modelo BP2416-4000CE (AMTI, EUA) operando à frequência de amostragem 1000Hz com amplificador MiniAmp MSA-6 da mesma marca. De acordo com a componente vertical GRF_z da força vertical de reação do solo durante a fase de suporte em que esta se mantém aproximadamente constante, foi determinada a massa de cada sujeito $m = Fg / g$, com $g = 9,81$ m/s² a aceleração da gravidade. A partir de GRF_z foi obtida a força vertical resultante $FR_z = GRF_z - Fg$ e a partir desta a aceleração vertical $az = FR_z / m$ do CG durante a fase de impulsão em cada tipo de salto. A partir dos perfis temporais de GRF_z e az foram extraídos um conjunto de parâmetros discretos nomeadamente o valor mínimo *min imp*, máximo *max imp* e médio *mean imp* de força e aceleração durante a fase de impulsão, os valores médio de força na fase descendente *mean down* e ascendente *mean up* de impulsão, bem como o nível de força e aceleração no instante de inversão *start up* da fase descendente para a fase ascendente de impulsão.

Foram aplicados com recurso ao *Statistical Package for Social Sciences* IBM SPSS 9.5.0.0 (SPSS, EUA) os testes de normalidade de Kolmogorov-Smirnov-Lillefors e de Shapiro-Wilk a todos os parâmetros associados a GRF_z e az nos vários tipos de salto considerados para avaliação da normalidade da distribuição, bem como os testes de Levene para avaliação da homogeneidade das variâncias dos dados amostrais ao nível de significância 0,05. Após avaliação dos pressupostos de normalidade e homogeneidade das variâncias, foram aplicados os testes paramétricos *paired samples* T-Test para amostras emparelhadas e os testes não paramétricos de Wilcoxon para avaliação da amplitude das diferenças emparelhadas.

Resultados

Apresentam-se na Tabela 1 os resultados de aplicação dos testes Kolmogorov-Smirnov para avaliação da normalidade e de Levene para avaliação da homogeneidade das variâncias de *GRFz* e *az* e na Tabela

2 os resultados da comparação estatística das médias emparelhadas com os testes paramétricos Paired samples T-Test e não paramétricos Wilcoxon Signed Ranks Test para as três condições experimentais, SJ, CMJ e DJ para os parâmetros discretos de *GRFz* e *az*.

Tabela 1: Estatísticas de teste e significâncias *($p < 0,05$) dos testes Kolmogorov-Smirnov para avaliação da normalidade e de Levene para avaliação da homogeneidade das variâncias de *GRFz* e *az* em SJ (S), CMJ (C) e DJ (D)

<i>GRFz</i> (N)	Kolmogorov-Smirnov						Levene's Test					
	SJ	p	CMJ	p	DJ	p	S-C	p	C-D	p	S-D	p
<i>min imp</i>	0,240	0,200*	0,194	0,200*	0,203	0,200*	3,185	0,105*	3,049	0,111*	7,613	0,020
<i>max imp</i>	0,206	0,200*	0,254	0,200*	0,154	0,200*	0,370	0,556*	5,944	0,035	6,989	0,025
<i>mean imp</i>	0,165	0,200*	0,163	0,200*	0,160	0,200*	1,281	0,284*	7,210	0,023	4,063	0,071*
<i>mean down</i>	0,312	0,068	0,304	0,088	0,188	0,200*	0,005	0,947*	13,104	0,005	13,022	0,005
<i>mean up</i>	0,224	0,200*	0,193	0,200*	0,177	0,200*	0,108	0,749*	0,009	0,925*	0,102	0,756*
<i>start up</i>	0,170	0,200*	0,203	0,200*	0,258	0,200*	0,469	0,509*	0,006	0,940*	0,209	0,657*
<i>take-off</i>	0,492	0,000	0,362	0,014	0,264	0,200*	24,80	0,001	0,704	0,421*	6,117	0,033
<i>max land</i>	0,197	0,200*	0,222	0,200*	0,225	0,200*	0,331	0,578*	1,295	0,282*	4,056	0,072*
<i>az</i> (m/s ²)	SJ	p	CMJ	p	DJ	p	S-C	p	C-D	p	S-D	p
<i>min imp</i>	0,216	0,200*	0,188	0,200*	0,223	0,200*	9,729	0,011	5,889	0,036	17,324	0,002
<i>max imp</i>	0,191	0,200*	0,237	0,200*	0,170	0,200*	3,292	0,100*	2,884	0,120*	5,313	0,044
<i>mean imp</i>	0,266	0,200*	0,287	0,134	0,169	0,200*	4,363	0,063*	8,859	0,014	4,983	0,050
<i>mean down</i>	0,464	0,000	0,202	0,200*	0,195	0,200*	0,887	0,369*	8,207	0,017	8,153	0,017
<i>mean up</i>	0,127	0,200*	0,316	0,062	0,201	0,200*	0,147	0,710*	8,291	0,016	15,268	0,003
<i>start up</i>	0,235	0,200*	0,181	0,200*	0,310	0,073	0,378	0,552*	9,270	0,012	11,540	0,007
<i>take-off</i>	0,331	0,039	0,266	0,200*	0,319	0,057	7,653	0,020	0,008	0,931*	8,570	0,015
<i>max land</i>	0,194	0,200*	0,178	0,200*	0,158	0,200*	0,398	0,542*	0,695	0,424*	2,547	0,142*

Tabela 2: Estatísticas de teste e significâncias *($p < 0,05$) dos testes paramétricos Paired samples T-Test e não paramétricos Wilcoxon Signed Ranks Test para os valores médios de amostras emparelhadas de *GRFz* e *az*.

<i>GRFz</i> (N)	Paired samples T-Test						Wilcoxon Signed Ranks Test					
	S-C	p	C-D	p	S-D	p	S-C	p	C-D	p	S-D	p
<i>min imp</i>	7,968	0,001*	-7,610	0,001*	-3,578	0,016*	2,201	0,028*	-2,201	0,028*	-2,201	0,028*
<i>max imp</i>	-2,587	0,049*	-11,126	0,000*	-12,870	0,000*	-2,023	0,043*	-2,201	0,028*	-2,201	0,028*
<i>mean imp</i>	2,005	0,101	-14,523	0,000*	-13,797	0,000*	1,992	0,046*	-2,201	0,028*	-2,201	0,028*
<i>mean down</i>	4,034	0,010*	-12,107	0,000*	-11,969	0,000*	1,992	0,046*	-2,201	0,028*	-2,201	0,028*
<i>mean up</i>	-1,854	0,123	-14,375	0,000*	-17,001	0,000*	-1,363	0,173	-2,201	0,028*	-2,201	0,028*
<i>start up</i>	-14,271	0,000*	-13,030	0,000*	-16,451	0,000*	-2,201	0,028*	-2,201	0,028*	-2,201	0,028*
<i>take-off</i>	1,625	0,165	-1,835	0,126	-0,524	0,623	1,604	0,109	-1,483	0,138	-0,135	0,893
<i>max land</i>	-0,196	0,852	1,006	0,361	0,754	0,485	0,314	0,753	0,943	0,345	0,524	0,600
<i>az</i> (m/s ²)	S-C	p	C-D	p	S-D	p	S-C	p	C-D	p	S-D	p
<i>min imp</i>	7,427	0,001*	-6,730	0,001*	-3,510	0,017*	2,201	0,028*	-2,201	0,028*	-2,201	0,028*
<i>max imp</i>	-2,504	0,054	-12,067	0,000*	-14,863	0,000*	-2,023	0,043*	-2,201	0,028*	-2,201	0,028*
<i>mean imp</i>	1,558	0,180	-14,288	0,000*	-13,221	0,000*	1,363	0,173	-2,201	0,028*	-2,201	0,028*
<i>mean down</i>	0,208	0,844	-12,923	0,000*	-12,873	0,000*	-0,105	0,917	-2,201	0,028*	-2,201	0,028*
<i>mean up</i>	-2,333	0,067	-13,059	0,000*	-12,344	0,000*	-1,992	0,046*	-2,201	0,028*	-2,201	0,028*
<i>start up</i>	-8,593	0,000*	-9,562	0,000*	-10,533	0,000*	-2,201	0,028*	-2,201	0,028*	-2,201	0,028*
<i>take-off</i>	1,552	0,181	-1,838	0,125	-0,676	0,529	1,214	0,225	-1,483	0,138	-0,105	0,917
<i>max land</i>	-0,247	0,814	1,169	0,295	0,821	0,449	0,314	0,753	1,153	0,249	0,734	0,463

Discussão

Os níveis de significância obtidos nos testes de normalidade de Kolmogorov-Smirnov-Lillefors e de Shapiro-Wilk (Tabela 1) foram suficientemente elevados, superiores a 0,05 na generalidade dos parâmetros associados a GRF_z e az em todos os tipos de salto pelo que é razoável assumir a normalidade da distribuição na população de proveniência dos dados amostrais. O reduzido tamanho da amostra ($n=6$) confere pouco significado aos resultados anteriores dos testes de normalidade. No entanto, a generalidade dos autores refere que a maior parte dos testes estatísticos são robustos relativamente à violação dos testes de normalidade, Stergiou (2004). Um outro pressuposto para aplicação dos testes paramétricos das médias em amostras emparelhadas foi avaliado através do teste de Levene para verificação da homogeneidade das variâncias. A maior parte dos parâmetros em GRF_z e az apresentaram significâncias superiores a 0,05 na comparação SJ-CMJ bem como em um número significativo de parâmetros nas comparações CMJ-DJ e SJ-DJ. Ainda assim, este requisito perde relevância, em virtude do número de sujeitos ser o mesmo em todas as situações, SJ, CMJ e DJ, já que se trata da mesma amostra de sujeitos em diferentes condições experimentais.

O primeiro pressuposto necessário para aplicação dos testes estatísticos paramétricos é de que os dados experimentais sejam numéricos de modo a permitir a realização de operações aritméticas. Os testes não paramétricos podem ser aplicados entre outras, em variáveis ordinais, mas a maior parte destes testes atende apenas à ordem, maior, menor ou igual, e não à amplitude da diferença entre pares de valores. O teste de Wilcoxon faz utilização da informação sobre a amplitude da diferença entre os pares de valores do mesmo parâmetro, sendo por isso mais sensível na detecção de diferenças quando estas existem. Ainda assim, o teste de Wilcoxon tem como pressuposto que as diferenças entre pares de valores associados sejam uma amostra com distribuição simétrica, o que pode ser avaliado qualitativamente pela forma do histograma da variável diferença de pares e quantitativamente por um coeficiente de *skewness* de Pearson para avaliação de assimetria da distribuição de probabilidade amostral. Na ausência de verificação das condições anteriores, o teste de Wilcoxon permite ainda assim comparar em qual das condições experimentais a população possui valores superiores.

Os resultados das comparações dos valores esperados nas diferentes condições experimentais SJ, CMJ e DJ são consistentes entre os testes paramétricos *Paired samples T-Test* e os testes não paramétricos *Wilcoxon Signed Ranks Test* com exceção apenas dos parâmetros GRF_z *mean imp*, az *max imp* e az *mean up* na comparação SJ-CMJ, para os quais foram detetadas diferenças nos testes NP ao contrário dos testes P que não detetaram diferenças para estes parâmetros ao nível de significância ($p<0,05$). Este fato poderá estar

associado às significâncias destes parâmetros muito próximas por excesso em P e por defeito em NP do nível de significância 0,05 definida como estatisticamente significativa (ES). Desta forma uma pequena variação do nível p em torno de 0,05 conduz a uma classificação como ES se a variação for por defeito ou como não ES se a variação for por excesso.

Conclusão

Perante condições mais hostis, relativas a amostras de dimensão reduzida ($n\leq 30$), ausência de normalidade, não homogeneidade de variâncias ou assimetria de variáveis emparelhadas, os testes não paramétricos revelam-se uma alternativa viável aos testes paramétricos, apesar de possuírem à partida menor sensibilidade. Apesar de os testes paramétricos exigirem condições mais rigorosas na verificação de pressupostos para aplicação, eles são mais sensíveis, permitindo detectar mais facilmente as diferenças quando elas existem. Desta forma, os testes paramétricos são frequentemente utilizados, mesmo quando nem todos os pressupostos de aplicação são verificados, desde que utilizados os testes não paramétricos correspondentes para verificação dos resultados.

Agradecimentos

Ao Prof. Carlos Carvalho que disponibilizou o LMH/ISMAI e supervisionou os ensaios bem como aos colegas que contribuíram no processo de recolha de dados. Aos Prof.s Velhote Correia e João Abrantes que orientaram e co-orientaram o meu Doutoramento. Ao Prof. Jurandir Nadal que supervisionou a mobilidade no PEB da COPPE/UFRJ. Ao Prof. Marco Benedetti que viabilizou a mobilidade Pós-doc no Dep. de Eletrônica e Sistemas da UFPE. À Univ. do Porto e à Università degli Studi di Roma "La Sapienza" que viabilizaram as mobilidades EBWII e BE MUNDUS.

Referências

- [1] Bendat, JA, Piersol, A.G. Random Data Analysis and Measurement Procedures. 3rd ed. New York: John Wiley and Sons; 2000.
- [2] Stergiou, N. Innovative Analyses of Human Movement. IL: Human Kinetics; 2004.
- [3] Komi, PV, Ishikawa, M & Linnamo, V. Identification of Stretch-Shortening Cycles in different sports. In: Proceedings of the 29th Conference of the International Society Biomechanical in Sports 2011 Jun 27- Jul 1; Porto, Portugal. 2011. 11(2): 31-33.
- [4] Winter, DA. Biomechanics and Motor Control of Human Movement. 4^a ed. New Jersey: John Wiley & Sons; 2009.
- [5] Asmussen E, Bonde-Petersen F. Storage of elastic energy in skeletal muscles in man. Acta Physiol. Scand. 1974; 91: 385-92.