

CENTRO DE ESTATÍSTICA APLICADA – CEA

RELATÓRIO DE CONSULTA

TÍTULO DO PROJETO: Determinação de critérios de desempenho a serem utilizados no processo de detecção, promoção e seleção de talentos esportivos no treinamento a longo prazo.

PESQUISADORA: Maria Tereza Silveira Böhme

INSTITUIÇÃO: Escola de Educação Física e Esporte – USP

FINALIDADE: Publicação

PARTICIPANTES DA ENTREVISTA: Maria Tereza Silveira Böhme
Rinaldo Artes
Carmen Saldiva de André
Frederico Poletto
Flávio Aparecido Moraes
Luiz Fernando Freitas Pereira

DATA: 19/04/2002

FINALIDADE DA CONSULTA: Críticas à análise estatística

RELATÓRIO ELABORADO POR: Flávio Aparecido Moraes
Luiz Fernando Freitas Pereira

1. INTRODUÇÃO

A detecção e seleção de talentos esportivos precoces são de suma importância para potencializar o desenvolvimento esportivo desses atletas. Entretanto, até agora, há poucos estudos que estabelecem métodos para esta detecção.

Nesta primeira parte do projeto pretende-se determinar as características (biológicas e de desempenho físico) que diferenciam os indivíduos treinados dos que não receberam treinamento. Assim, numa etapa futura, com o auxílio de profissionais que trabalham como técnicos esportivos no início do processo de treinamento, será possível determinar quais características devem ser consideradas nesta detecção.

2. DESCRIÇÃO DO ESTUDO

A amostra desta pesquisa é composta por adolescentes de ambos os sexos, na faixa etária de 11 a 14 anos, divididos em dois grupos – treinados e não treinados. Os treinados eram participantes do projeto “Esporte Talento”, do Centro de Práticas Esportivas da Universidade de São Paulo e Fundação Ayrton Senna (CEPEUSP/AYRTON SENNA), das modalidades esportivas basquetebol, handebol, canoagem e futebol; os não treinados eram alunos da Escola de Aplicação da Faculdade de Educação da Universidade de São Paulo (EAFEUSP). A amostra é composta por 111 adolescentes (37 do sexo masculino e 64 do sexo feminino) que foram treinados e 85 adolescentes (34 do sexo masculino e 51 do sexo feminino) que não foram treinados.

Cada adolescente foi submetido a testes para determinação do crescimento físico, adiposidade corporal, maturação sexual e aptidão física, ao longo de três semestres consecutivos (abril/98, setembro/98 e abril/99); dessa forma, seriam obtidas três medidas das variáveis de interesse para cada adolescente. Não foi possível a determinação completa das três medidas para cada indivíduo em todas as variáveis devido ao abandono do estudo por alguns adolescentes.

3. DESCRIÇÃO DAS VARIÁVEIS

Os participantes do estudo foram classificados segundo as seguintes variáveis :

- Treinamento: treinado ou não treinado;
- Sexo: masculino ou feminino;
- Tempo de medição: primeira, segunda ou terceira medição;
- Faixa etária: de 10 a 13 anos ou maior que 13 anos;

Foram também observadas as variáveis:

- Indicadoras do crescimento físico: estatura (cm), peso (kg), envergadura (cm), altura tronco-cefálica (cm), comprimento de membros inferiores (cm), diâmetros de úmero e trocânter (cm), perímetros de braço contraído (cm) e perna (cm);
- Indicadoras da adiposidade corporal (mm): dobras cutâneas tricipital, subscapular, supra espinal, abdominal e perna;
- Indicadoras da maturação sexual:
 - estágios de desenvolvimento da pilosidade pubiana: de P1 (infantil) a P5 (adulto) para ambos os sexos;
 - estágios de desenvolvimento de mamas: de M1 (infantil) a M5 (adulto) para o sexo feminino;
 - estágios de desenvolvimento de genitais: de G1 (infantil) a G5 (adulto) para o sexo masculino;
- Indicadoras de aptidão física:
 - Teste de impulsão horizontal (cm): relacionado à força explosiva de membros inferiores;
 - Teste de arremesso de “medicine ball” (cm): relacionado à força explosiva de membros superiores;
 - Teste de flexão de tronco (número de vezes): relacionado à força da musculatura do tronco;
 - Teste de agilidade (segundos): indicadora da capacidade de mudar o corpo de direção o mais rápido possível;

- Teste de velocidade de 30 m (segundos): relacionado à velocidade de deslocamento;
- “Teste de sentar e alcançar” (cm): indicadora da capacidade de flexão da articulação do quadril;
- Teste de corrida de 9 minutos (m): indicadora da capacidade de resistência geral aeróbica;

4. SITUAÇÃO DO PROJETO

A pesquisadora já havia realizado uma análise dos dados. Foi utilizado o aplicativo “SPSS for Windows” versão 8.0. O principal interesse da pesquisadora era eliminar algumas dúvidas sobre as técnicas por ela utilizadas e verificar possíveis erros conceituais em sua análise.

Numa primeira etapa da análise, a pesquisadora aplicou um procedimento de imputação de dados. Utilizou-se um método baseado em modelos de regressão (ver apêndice A). A pesquisadora informou que menos de 20% dos dados da análise foram imputados.

Para cada variável de aptidão física, foi ajustado um modelo de análise de variância com dois fatores fixos sendo eles: treinamento e tempo de medição (com medidas repetidas neste fator) (Neter et. al, 1996).

A pesquisadora realizou análises de variância separadas para cada sexo, pois sabe-se que em todas as variáveis de aptidão física observadas há grande diferença entre os sexos, principalmente na faixa etária em que se encontravam os adolescentes do estudo.

Também foram realizadas análises discriminantes (Johnson et. al, 1999) separadas por tempo de medição, para determinar quais variáveis de aptidão física mais diferenciam os grupos de adolescentes treinados e não treinados. Nesta parte do estudo, como forma de avaliação da análise discriminante, a pesquisadora apresentou as porcentagens de classificação correta.

5. SUGESTÕES DO CEA

A técnica de imputação de dados realizada é adequada (ver apêndice A). Porém quando se utilizam dados com imputações nas análises de variância é necessário corrigir o número de graus de liberdade da soma de quadrados residual e consequentemente da estatística de teste F. Foi sugerido pelo CEA que a pesquisadora realizasse novamente todas as análises sem a imputação de dados e verificasse possíveis diferenças entre os resultados obtidos.

Para a análise de variância com medidas repetidas é necessário que a matriz de variância e covariância das observações dos indivíduos satisfaça à condição de esfericidade (Winer et. al, 1992). O SPSS v.8 realiza um teste para avaliar se a condição de esfericidade é satisfeita. No apêndice B há um exemplo de como interpretar este teste e proceder se esta suposição for satisfeita ou não.

Para a inclusão de outros fatores (como maturação sexual) nas análises de variâncias com medidas repetidas foi sugerido a uso de variáveis do tipo “dummy”. Estas variáveis assumem os valores 0 ou 1 e servem para indicar se as observações pertencem ou não a um determinado nível de um fator. Para a implementação desta análise o uso do software SAS v.8 foi recomendado, pois assim não é necessária a imputação dos dados. Há um exemplo de como construir as variáveis do tipo “dummy” e como utilizar o SAS para realizar a análise no apêndice C.

Para a análise discriminante, foi sugerido que a pesquisadora utilizasse o método “Leave-One-Out”, também denominado “Jackknife” (Johnson et. al, 1999) como melhor forma de validação desta análise, pois já está disponível no aplicativo utilizado (SPSS for Windows). Esse método classifica cada elemento da amostra segundo a função discriminante calculada com todos os elementos, exceto o que será classificado. Assim, elimina-se o vício de utilizar a mesma amostra para determinar a função discriminante e a validação.

Por último, foi sugerido que este projeto seja encaminhado para triagem do CEA para ser analisado durante o 2º semestre de 2.002.

6. REFERÊNCIAS BIBLIOGRÁFICAS

JOHNSON, R. A. and WICHERN, D. W. (1999). **Applied multivariate statistical analysis**. 4th.ed. Upper Saddle River, N.J. : Prentice Hall. 816p.

NETER, J., WASSERMAN, W. KUTNER, M. H. and NACHTSHEIM, C. J. (1996). **Applied linear statistical models**. 4.ed. New York: McGraw-Hill. 1408p.

WINER, B. J., BROWN, D. R. and MICHELS, K. M. (1991). **Statistical principles in experimental design**. 3ed. New York: McGraw-Hill. 1057p.

APÊNDICE A

O método de imputação por regressão do software SPSS v.8 consiste em:

1. Estimar uma equação que relacione a variável que contém dados ausentes com outras variáveis do banco de dados. Esta equação é obtida através da Regressão Linear Múltipla;

2. Utilizando a equação encontrada, prever o valor perdido para a variável de interesse;

3. Adicionar a esta previsão uma componente aleatória.

Exemplo:

Vamos supor que temos os seguintes dados:

Indivíduo	Altura (m)	Peso (kg)	X
1	1,79	75	124,00
2	1,73	70	125,00
3	1,65	68	119,00
4	1,70	74	130,00
5	1,56	49	102,50
6	1,88	90	*
7	1,69	65	118,00
8	1,83	87	140,00

O método de imputação por regressão irá calcular, para este caso, a seguinte equação, considerando apenas os dados completos nas variáveis altura, peso e X simultaneamente:

Regressão Linear Múltipla

$$X = 43,2 * \text{Altura (m)} + 0,701 * \text{Peso (kg)} + E,$$

onde E é uma componente aleatória

$$\text{QMR} = 12$$

Com esta equação, o valor a ser imputado para a variável X do sexto indivíduo será:

$X = 43,2 * 1,88 + 0,701 * 65 + E$, onde E pode ser obtido de duas maneiras diferentes.

A primeira maneira, definida como padrão pelo SPSS v.8 é obter aleatoriamente um E de algum indivíduo, por exemplo, do primeiro indivíduo. Assim temos que:

$$E = 124,00 - 43,2 * 1,79 - 0,701 * 75 = - 5,903$$

A segunda maneira é supor que E tem distribuição Normal com média 0 e variância igual ao quadrado médio residual (QMR) da regressão linear múltipla. No nosso caso a variância é 12. Desta forma, podemos calcular aleatoriamente um E.

Escolhendo a primeira maneira de obter E, o valor a ser imputado para a variável X do sexto indivíduo será:

$$X = 43,2 * 1,88 + 0,701 * 65 - 5,903 = 120,88$$

APÊNDICE B

Uma importante suposição do modelo da análise de variância para medidas repetidas é a suposição de esfericidade da matriz de variância e covariância das observações. Abaixo há um exemplo de saída do software SPSS de uma análise de variância com medidas repetidas em 1 fator e suas devidas interpretações. Neste exemplo, o interesse do estudo é comparar o efeito de 4 tipos de drogas sobre a realização de uma determinada tarefa. Cinco indivíduos foram observados sob o efeito de cada droga.

Fonte:

<http://www.ssc.uwo.ca/sscl/statsexamples/spss/anova/rmnogroupresults.html>

Mauchly's Test of Sphericity(b)
Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon(a)		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
DRUG	.186	4.572	5	.495	.605	1.000	.333
Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.							
a May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the layers (by default) of the Tests of Within Subjects Effects table.							
b Design: Intercept Within Subjects Design: DRUG							

Na tabela acima, podemos verificar que podemos aceitar a suposição de esfericidade, pois o Sig (p-valor) é maior que 5%. Caso contrário deveríamos analisar um dos testes abaixo que ajustam os p-valores à falta da suposição de esfericidade, como o Greenhouse-Geisser, Huynh-Feldt e Lower-bound, mostrados a seguir.

Tests of Within-Subjects Effects
Measure: MEASURE_1
Sphericity Assumed

Tabela a ser analisada

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power(a)
DRUG	698.200	3	232.733	24.759	.000 ¹	74.277	1.000
Error(DRUG)	112.800	12	9.400				

a. Computed using alpha = .05

Tests of Within-Subjects Effects
Measure: MEASURE_1
Greenhouse-Geisser

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power(a)
DRUG	698.200	1.815	384.763	24.759	.001	44.928	.999
Error(DRUG)	112.800	7.258	15.540				

a. Computed using alpha = .05

Tests of Within-Subjects Effects
Measure: MEASURE_1
Huynh-Feldt

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power(a)
DRUG	698.200	3.000	232.733	24.759	.000	74.277	1.000
Error(DRUG)	112.800	12.000	9.400				

a. Computed using alpha = .05

Tests of Within-Subjects Effects
Measure: MEASURE_1
Lower-bound

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power(a)
DRUG	698.200	1.000	698.200	24.759	.008	24.759	.953
Error(DRUG)	112.800	4.000	28.200				

a. Computed using alpha = .05

Como a suposição de esfericidade é satisfeita para o nosso exemplo, devemos observar agora o Sig (p-valor) da tabela 'Sphericity Assumed'. Como o p-valor é menor que 5%, chegamos a conclusão que há o efeito do fator droga, ou seja, há pelo menos

uma média referente a cada droga que difere das demais. Se a suposição não fosse satisfeita, deveríamos observar o p-valor de algum dos três testes ajustados.

APÊNDICE C

Abaixo segue um exemplo de implementação de uma análise de variância no programa SAS e da codificação de variáveis qualitativas. Este estudo teve por objetivo avaliar 27 indivíduos segundo uma característica y em três períodos distintos (aos 8, 10 e 12 anos de idade), segundo treinamento (sim ou não).

Suponha que a variável maturação sexual (matura) assuma os valores 1 ou 2 indicando o nível de maturação (“imaturo” ou “maduro”). Podemos criar uma variável “dummy” (ms1) que tem a seguinte correspondência com a variável original:

Valor original	Valor da “dummy”
Maturação	Ms1
1	0
2	1

Dessa forma o parâmetro associado a ms1 indica o acréscimo na variável resposta quando se passa do nível “imaturo” para o nível “maduro” de maturação sexual.

Imagine agora que a variável maturação sexual (matura) assuma os valores 1, 2 ou 3 indicando o nível de maturação (“imaturo”, “moderado” e “maduro”, respectivamente). Podemos criar duas variáveis “dummy” (ms1 e ms2) que têm a seguinte correspondência com a variável original:

Valor original	Valor da “dummy”	
Maturação	Ms1	Ms2
1	0	0
2	1	0
3	0	1

Dessa forma o parâmetro associado a ms1 indica o acréscimo esperado na variável resposta quando se passa do nível “imaturo” para o nível “moderado” de

maturação sexual e o associado a ms2 indica acréscimo esperado na variável resposta quando se passa do nível “imaturo” para o nível “maduro” de maturação sexual.

Há uma maneira alternativa de se criar variáveis “dummy” se a variável em questão for ordinal. No nosso exemplo, a variável matura é ordinal. Esta maneira alternativa está detalhada abaixo e implementada no programa SAS:

Valor original	Valor da “dummy”	
Maturação	Ms1	Ms2
1	0	0
2	1	0
3	1	1

Assim o parâmetro associado a ms1 indica o acréscimo na variável resposta quando se atinge o nível “moderado” de maturação sexual e o associado a ms2 indica o acréscimo na variável resposta quando se atinge o nível “maduro” de maturação sexual.

A implementação dos outros dois exemplos de variável dummy é semelhante ao apresentado a seguir.

Entrada de dados e especificação do modelo

Na tela do editor de programas do SAS digite:

```
data dados;
```

```
input Indiv Treina $ y1 y2 y3 matura1 matura2 matura3;
```

```
y=y1; ldade=8; matura=matura1; output;
```

```
y=y2; ldade=10; matura=matura2; output;
```

```
y=y3; ldade=12; matura=matura3; output;
```

```
drop y1-y3 matura1-matura3;
```

```
datalines;
```

```
1 n 21.0 20.0 21.5 1 1 2
2 n 21.0 21.5 24.0 2 2 3
3 n 20.5 24.0 24.5 1 2 3
4 n 23.5 24.5 25.0 1 1 2
```

5	n	21.5	23.0	22.5	2	2	2
6	n	20.0	21.0	21.0	2	3	3
7	n	21.5	22.5	23.0	1	1	1
8	n	23.0	23.0	23.5	1	2	3
9	n	20.0	21.0	22.0	1	1	1
10	n	16.5	19.0	19.0	2	3	3
11	n	24.5	25.0	28.0	1	3	3
12	s	26.0	25.0	29.0	1	2	2
13	s	21.5	22.5	23.0	3	3	3
14	s	23.0	22.5	24.0	2	3	3
15	s	25.5	27.5	26.5	1	1	2
16	s	20.0	23.5	22.5	1	2	3
17	s	24.5	25.5	27.0	2	2	3
18	s	22.0	22.0	24.5	1	2	3
19	s	24.0	21.5	24.5	1	1	2
20	s	23.0	20.5	31.0	3	3	3
21	s	27.5	28.0	31.0	2	3	3
22	s	23.0	23.0	23.5	1	1	2
23	s	21.5	23.5	24.0	2	3	3
24	s	17.0	24.5	26.0	1	2	3
25	s	22.5	25.5	25.5	2	3	3
26	s	23.0	24.5	26.0	1	2	2
27	s	22.0	21.5	23.5	2	2	3

;

```
data dados1; set dados;
    if matura=1 then do; ms1=0; ms2=0; end;
    if matura=2 then do; ms1=1; ms2=0; end;
    if matura=3 then do; ms1=1; ms2=1; end;
run;
```

Criação das variáveis do tipo
Dummy, a partir da variável
matura

```

data dados2; set dados1;
  idadex=idade-8;
run;

```

A variável idade deve ser decrescida do menor valor

```

proc mixed data=dados2 method=ml covtest;
  class indiv treina;
  model y = treina idadex treina*idadex ms1 ms2 /s;
  repeated / type=un subject=indiv r;
run;

```

Especificação do modelo

Em seguida, executar o comando Submit na barra de ferramentas do programa SAS. As saídas deverão aparecer na tela denominada Output. As principais tabelas a serem analisadas são:

Tabela 1: Estimativas dos parâmetros do modelo

Solution for Fixed Effects						
Effect	Treina	Estimate	Standard Error	DF	t Value	Pr > t
Intercept		22.5299	0.5763	25	39.09	<.0001
Treina	n	-1.3403	0.7976	25	-1.68	0.1053
Treina	s	0
idadex		0.6571	0.1529	25	4.30	0.0002
idadex*Treina	n	-0.2157	0.1871	25	-1.15	0.2599
idadex*Treina	s	0
ms1		0.1089	0.5068	25	0.21	0.8316
ms2		0.1893	0.5047	25	0.38	0.7108

O intercepto refere-se à média da variável y dos indivíduos que treinam, têm 8 anos de idade e estão no estágio inicial de maturação sexual. No nosso exemplo, não há efeito de maturação sexual ao nível de significância de 5%.

Tabela 2: Análise de variância

Type 3 Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
Treina	1	25	2.82	0.1053
idadex	1	25	17.36	0.0003
idadex*Treina	1	25	1.33	0.2599
ms1	1	25	0.05	0.8316
ms2	1	25	0.14	0.7108

A um nível de 5% de significância, podemos dizer que não existe efeito de interação entre idade e sexo.

Neste exemplo, não é necessário testar a hipótese de esfericidade, pois não assumimos nenhuma estrutura na matriz de variância e covariância das observações na especificação do modelo através do comando “type=un”.