**Disciplina: Estatísticas Educacionais Tarefa 2**

O empresário gostou da análise, mas pretende obter um produto mais completo. Você diz que vai fazer uma outra análise baseada na famosa Teoria da Resposta ao Item (TRI). Pensando na bolada maior ainda você recorre a um especialista em TRI, que lhe aconselha os seguintes passos:

**Etapas da Análise TRI:**

1. Crie uma base dicotômica com os códigos 1 (Acerto) e 0 (Erro) no item. Você terá uma matriz de 48 colunas e cerca de 73 mil linhas.
2. Uma medida alternativa de discriminação de um item é dada pelo Coeficiente de Correlação entre o Item (0 e 1) e o Escore, chamado também de *Correlação Ponto-Bisserial*. No entanto, exclua o tal item do Escore, de forma que só poderá chegar a 47.
3. Crie um arquivo de texto chamado Base01.dat [Lembre-se que se os dados estiverem no Excel você deve usar *Salvar Como* e escolher o tipo *Texto formatado (separado por espaços) (\*.prn).* Pode-se usar a extensão PRN, ou alterar para DAT ou TXT. Verifique o arquivo gerado contando quantas colunas há exatamente no DAT, que equivale à LARGURA da coluna. Um exemplo de sintaxe SPS está Anexa.
4. Prepare uma sintaxe extensão BLM do BILOG-MG para rodar os seus dados dicotômicos. Rode e veja os Gráficos (Run / Plot) de ajuste dos itens.
5. Monte uma macro no Excel (VBA) para obter todas as tabelas e gráficos.

**Prazo de Entrega**: Classroom (como todo cliente, há urgência). A Multa Contratual é de 10% por dia de atraso.

**SÍNTESE DE SINTAXE DO IBM-SPSS**

A melhor forma de iniciar as sintaxes no SPSS é fazer a operação via menu e na etapa final clicar em **PASTE** ou **COLAR** ao invés de concluir. Seque algumas sintaxes criadas desta forma.

**EXEMPLOS DE LEITURA TIPO SPSS:**

GET

 FILE=' C:\TRI\Dados01.sav'.

**EXEMPLOS DE LEITURA TIPO TEXTO (DAT, TXT ETC.):**

GET DATA /TYPE = TXT

 /FILE = ' C:\TRI\Dados.TXT'

 /FIXCASE = 1

 /ARRANGEMENT = FIXED

 /FIRSTCASE = 1

 /IMPORTCASE = ALL

 /VARIABLES =

 /1

INSCRICAO 0-7 F8

SEXO 10-10 A1

UF 11-18 F8

NOTA 20-27 COMMA8.2

RESPOSTAS 837-1036 A200 .

CACHE.

EXECUTE.

**OBS.** No geral, os números têm formato F8.2, onde 8 representa o número total de dígitos e 2 a quantidade de casas decimais; para inteiros podemos usar apenas F8; quando o separador for vírgula, usa-se COMMA8.2; para *strings* de tamanho 8 usamos A8.

**EXEMPLOS DE SALVAR DADOS SPSS:**

SAVE OUTFILE=' C:\TRI\Dados01.sav'

 /COMPRESSED.

**EXEMPLOS DE SALVAR DADOS EM TIPO TEXTO (DAT, TXT ETC.):**

Sintaxe SPSS para salvar três variáveis em um arquivo DAT (texto). A primeira variável (ID) será a Identificação do indivíduo, com 12 caracteres (A12), sendo acrescido mais um caracter em branco (1X) no arquivo DAT. Depois serão colocadas as respostas (0 ou 1, F1.0, ou seja, numérico com nenhuma casa decimal) no DAT, sem espaço separador.

WRITE OUTFILE='C:\TRI\Dados01.dat'

 TABLE

/ID (A12,1X)

Item1 (F1.0)

Item2 (F1.0)

EXECUTE.

***Se tivermos várias variáveis com o mesmo formato poderemos adotar:***

WRITE OUTFILE='C:\TRI\Dados01.dat'

 TABLE

/ID (A12,1X)

Item1 TO Item10 (10(F1.0))

EXECUTE.

**EXEMPLOS DE CODIFICAÇÃO/RÓTULO:**

VALUE LABELS TP\_SEXO

 M 'Masculino'

 F 'Feminino'.

\* --------------------------------------------------------------------------------------------------------------------

\* Atribuindo Value Labels aos Estados

\* --------------------------------------------------------------------------------------------------------------------.

VALUE LABELS CO\_UF\_INSC

 11 'Rondonia'

 12 'Acre'

 13 'Amazonas'

 14 'Roraima'

 15 'Para'

 16 'Amapa'

 17 'Tocantins'

 21 'Maranhao'

 22 'Piaui'

 23 'Ceara'

 24 'Rio Grande do Norte'

 25 'Paraiba'

 26 'Pernambuco'

 27 'Alagoas'

 28 'Sergipe'

 29 'Bahia'

 31 'Minas Gerais'

 32 'Espirito Santo'

 33 'Rio de Janeiro'

 35 'Sao Paulo'

 41 'Parana'

 42 'Santa Catarina'

 43 'Rio Grande do Sul'

 50 'Mato Grosso do Sul'

 51 'Mato Grosso'

 52 'Goias'

 53 'Distrito Federal' .

**EXEMPLOS DE RECODIFICAÇÃO:**

\* --------------------------------------------------------------------------------------------------------------------

\* Criando as Regiões com labels

\* --------------------------------------------------------------------------------------------------------------------.

RECODE

 CODUF

 (11 thru 17=1) (18 thru 29=2) (31 thru 35=3) (41 thru 43=4) (50 thru 53=5) INTO Regiao .

 VALUE LABELS Regiao 1 'Norte' 2 'Nordeste' 3 'Sul' 4 'Sudeste' 5 'Centro-Oeste' .

FORMAT Regiao (F1.0).

EXECUTE .

**EXEMPLOS DE TRANSFORMAÇÃO:**

COMPUTE TIPO = 1 .

COMPUTE ORDEM=$CASENUM .

COMPUTE NOTA100 = (NOTA-M)/S \*100+500.

COMPUTE Tamanho=LENGTH(RTRIM(LTRIM(RESPOSTAS,'.'))).

EXECUTE.

**EXEMPLOS DE FILTRO:**

USE ALL.

SELECT IF (ID\_PROVA=2).

EXECUTE.

**EXEMPLO: DELETANDO VARIÁVEIS:**

DELETE VARIABLES VAR1 VAR2.

DELETE VARIABLES VAR3 TO VAR10.

EXECUTE.

**EXEMPLO: ORDENANDO:**

SORT CASES BY

 INSCRICAO (A) .

**EXEMPLO: SEPARANDO UM STRING DE RESPOSTAS (A50) EM 50 VARIÁVEIS Q1,Q2, ..., Q50 COM AS RESPOSTAS INDIVIDUAIS**

DO REPEAT Q = Q1 TO Q50 / #J = 1 TO 50 .

 STRING Q (A1).

 COMPUTE Q= SUBSTR(RESPOSTAS,#J,1) .

END REPEAT .

FORMAT Q1 TO Q50 (A1) .

EXECUTE .

**EXEMPLO: FAZENDO A CORREÇÃO PARA 0-1 (DICOTÔMICA) DE UMA PROVA COM APENAS UM CADERNO (GABARITO) E COLOCANDO EM QC1, QC2,..., QC50.**

STRING #GAB (A50).

COMPUTE #GAB='DBEDEBADBDBBAADABBCEBDECDECABCCBCCDEBDAAAEECDEABEC' .

DO REPEAT Q = QC1 TO QC50 / R = Q1 TO Q50 / #J = 1 TO 50 .

 Q= (R=SUBSTR(#GAB,#J,1)) .

END REPEAT .

FORMAT QC1 TO QC50 (F1.0) .

EXECUTE .

**EXEMPLO: FAZENDO A CORREÇÃO PARA 0-1 (DICOTÔMICA) DE UMA PROVA COM VÁRIOS (8) CADERNOS (GABARITOS)**

STRING #Gab1 #Gab2 #Gab3 #Gab4 #Gab5 #Gab6 #Gab7 #Gab8 (A63).

COMPUTE #GAB1='DBEDEBADBDBBAADABBCEBDECDECABCCBCCDEBDAAAEECDEABECACDCDBCEAAECE' .

COMPUTE #GAB2='DBEEDBADBDBABADBABCBEDEDCCEABCCBCCDEDBAAAEECDEAEBCACDDBCCEAAECE' .

COMPUTE #GAB3='DBEBEDADBDBABABADBCEBEDDCCEBACCBCDECDBAAAEECDEABEACCDBDCCAEAECE' .

COMPUTE #GAB4='DBEBDEABDDBBAABDACBEBDCEDBACECCBCCDEDBAAAEECDEABEACCDCBDCEAEACE' .

COMPUTE #GAB5='DAEACBEAABAADADEEDBEDBCEEDAEBDEECDBCBCCCDBCBDCEADACBBBEDBCADACC' .

COMPUTE #GAB6='DCEACAEAEBBADBDEDCBEEBCEEDAEADEEDDBCBCCCDBCADCEADACBEBEDBDADACD' .

COMPUTE #GAB7='DCEACAEAEBBADBDEDCBEEBCEEDAEADEEDEBCCCCCDCEBDCEADACBEAEDAAADBCD' .

COMPUTE #GAB8='ECEACABAEBBAEBDEEBBDEBAEEDAEADEADABCECDCDBEBDCEEDEECEEEDAEBDCCD' .

DO REPEAT Q = QC1 TO QC63 / R = Q1 TO Q63 / #J = 1 TO 63 .

 IF (ID\_PROVA=1) Q= (R=SUBSTR(#GAB1,#J,1)) .

 IF (ID\_PROVA=2) Q= (R=SUBSTR(#GAB2,#J,1)) .

 IF (ID\_PROVA=3) Q= (R=SUBSTR(#GAB3,#J,1)) .

 IF (ID\_PROVA=4) Q= (R=SUBSTR(#GAB4,#J,1)) .

 IF (ID\_PROVA=5) Q= (R=SUBSTR(#GAB5,#J,1)) .

 IF (ID\_PROVA=6) Q= (R=SUBSTR(#GAB6,#J,1)) .

 IF (ID\_PROVA=7) Q= (R=SUBSTR(#GAB7,#J,1)) .

 IF (ID\_PROVA=8) Q= (R=SUBSTR(#GAB8,#J,1)) .

END REPAEAT .

FORMAT QC1 TO QC63 (F1.0) .

EXECUTE .

**EXEMPLOS DE CONSTRUÇÃO DE ESCORES:**

COMPUTE Escore = SUM(QC1 TO QC63) .

FORMAT Escore (F2.0) .

EXECUTE .

**OBTENDO FREQUÊNCIAS:**

FREQUENCIES

 VARIABLES=Q1 TO Q63 Escore

 /ORDER= ANALYSIS

**SINTAXE DO BILOG-MG**

Curso TRI: Heliton Tavares, UFPA

Situação 1: Calibração dos dados da População 1 com ~ N(0,1)

>COMMENTS

(1) As duas primeiras linhas deste arquivo são reservadas para um título geral. Caso

 não haja, devem ser deixadas em brando.

(2) Neste exemplo há apenas um grupo em estudo. Será feira a calibração dos itens e scoring (habilidades).

(3) Os dados são simulados considerando 50 itens de múltipla escolha, cujas respostas estão no arquivo DADOS1.DAT.

(4) Os parâmetros dos itens serão estimados na Escala Normal (0,1) [D=1.702, default].

 Para usar a Logística (D=1) deve-se acrescentar a opção LOGistic no comando GLOBAL. Será adotado o Modelo

 Logístico de 3 parâmetros (NPARM=3).

(5) O SAVE indica que algumas estimativas serão salvas; parâmetros dos itens em CursoTRI.PAR e habilidades em CursoTRI.SCO

(6) No arquivo de dados Dados1.dat, os 9 primeiros caracteres estão reservados para IDentificação

 do respondente. Este número aparece sempre em dois lugares: NIDCHAR=9 e (9A1,50A1)

(7) Podemos dar nomes gerais ou apropriados aos itens: (item01,Item02, ...,items50) cujo step é (1)

(8) No processos de Calibração serão adotados 31 pontos de quadratura (NQPT), 50 ciclos EM (CYCLES), com critério de

 convergência de 0.001 (CRIT). Serão montados os gráficos para todosos itens (PLOT=1), com até 20 pontos de ajuste

 Qui-Quadrado (CHI=20). O DIAgnostic=1,ou 2 ou 3 ou 4 apresenta vários níveis de Detalhamento o processo.

(9) As proficiências serão estimadas por MV (METhod=1). Posteriormente serão obtidas

 pelo método EAP (METhod=2) e MAP (METhod=3). Melhor não mostrar as habilidades na tela (NOPRINT), nem

 transformar/padronizar as estimativas (RSCTYPE=0).

(10) O BILOG-MG lê apenas os três primeiros caracteres de cada opção de comando, de forma

 que poderíamos escrever apenas MET=1 (ao invés de METHOD=1), LOG (ao invés de LOGistic) etc.

(11) O BILOGMG só lê até a coluna 80 desta sintaxe, ignorando os caracteres posteriores.

(12) Tudo o que escrevermos antes de GOLBAL (ver abaixo) vira comentário. Os comandos vão começar agora.

>GLOBAL DFNAME='Dados1.dat', NPARM=3, SAVE;

>SAVE PARM='CursoTRI.PAR', SCORE='CursoTRI.SCO';

>LENGTH NITEMS=50;

>INPUT NTOTAL=50, NIDCHAR=9;

>ITEMS INAMES=(Item01(1)Item50);

>TEST TNAME=UFPA, INUMBER=(1(1)50);

(9A1,50A1)

>CALIB NQPT=31, CYCLES=50, NEWTON=0,CRIT=0.001, PLOT=1, CHI=20, DIAgnostic=4;

>SCORE NOPRINT, METHOD=1, RSCTYPE=0;

**SINTAXE DO SPSS PARA IMPORTAR PROFICIÊNCIAS DO BILOG-MG**

GET DATA  /TYPE = TXT

 /FILE = 'C:\CursoTRI.SCO'

 /FIXCASE = 2

 /ARRANGEMENT = FIXED

 /FIRSTCASE = 3

 /IMPORTCASE = ALL

 /VARIABLES =

 /1  Grupo 0-2 F3.0

 ID 5-16 A30

 /2 Peso 0-5 COMMA6.2

 Calib 6-6 1X

 Teste 7-14 A8

 Tentados 15-19 F5.0

 Acertados 20-24 F5.0

 PAcerto 25-34 COMMA10.2

 Prof01 35-46 COMMA12.2

 EP 47-58 COMMA12.2  .

CACHE.

EXECUTE.