

Instruções para o trabalho

Este documento contém todas as instruções para o desenvolvimento do segundo trabalho prático. Siga os requisitos apresentados.

1. Resumo:

O trabalho consiste em desenvolver um script em bash – possivelmente com o auxílio dos filtros nativos do linux como grep, awk e sed – para executar um projeto de experimento (PDE) especificado num arquivo texto de configuração. Este documento apresenta mais detalhes sobre o trabalho e foi estruturado da seguinte forma:

1. Resumo – Apresenta essa breve descrição.
2. Projeto de experimentos – Explica o que se entende por projeto de experimentos (PDE).
3. Arquivo de configuração – Descreve o formato utilizado no arquivo de configuração do PDE, aquele que será fornecido como entrada para o script.
4. Exemplo – Apresenta um exemplo e explica o que o script deverá fazer.

O trabalho pode ser executado em grupo de até dois alunos.

2. Projeto de experimentos

O projeto de experimentos é uma técnica bastante utilizada por pesquisadores interessados em avaliar como certos fatores observados afetam o comportamento de algum procedimento sendo estudado. Você facilmente encontrará muito material disponível na internet descrevendo como desenvolver projetos de experimentos e diversos livros de estatística especializados nesse assunto. O nosso objetivo aqui, entretanto, não é aprofundar nesse tema, mas apenas apresentar algumas informações básicas para esclarecer o que será cobrado no trabalho.

Considere, por exemplo, que você deseja avaliar o comportamento de algum programa PROG que recebe como entrada os valores das variáveis A, B, C. Para estudar esse comportamento, você planejou uma bateria de testes variando as entradas das variáveis A, B, C com um conjunto de valores predefinidos. Isso basicamente descreve um projeto de experimentos. Formalizando então e definindo melhor os termos utilizados usaremos esse exemplo para ilustrar:

- Evento estudado: é o comportamento do programa PROG;
- Fatores: são as variáveis estudadas que afetam o evento, no caso as variáveis A, B e C;
- Níveis dos fatores: são os conjuntos de valores que cada fator vai assumir no PDE;
- Ensaios: são as instâncias de experimentos realizados no PDE.

3. Arquivo de configuração

O simples exemplo abaixo descreve uma configuração de PDE. Daremos um nome para esse arquivo, por exemplo, config.txt.

```

1  ## Projeto de experimentos para o programa PROG
2
3  FATORES:
4  A = (5, 20, 60)
5  B = (1.5, 10, 20, 40)
6  C = (1, 0)
7
8  COMANDO:
9  prog $A -f $B -d $C
10
11 ENSAIOS:
12 1 = 5, 10, *
13 2 = 20, 40, 0
14 3 = *, 1.5, 1
15
16 ## fim do exemplo

```

4. Exemplo

A chamada do script, chamado `projexp`, será feita via linha de comando, acompanhada de dois parâmetros que são o nome do arquivo de configuração e o nome do arquivo de saída (que vai armazenar um log dos resultados):

`projexp config.txt saidas.txt`

Segue uma breve descrição de cada parte do arquivo:

- As linhas que seguem a descritor `FATORES` indicam os fatores estudados seguidos dos respectivos níveis. Para simplificar considere que os níveis vão tratar números, inteiros ou flutuantes.
- A linha que segue o descritor `COMANDO` indica qual linha de código que será disparada pelo script para realizar cada ensaio. Note que ela pode incluir variáveis (denotadas por dólar+fator, no exemplo, `$A`, `$B` e `$C`) cujos valores serão indicados nos ensaios que seguem.
- As linhas que seguem o descritor `ENSAIOS` indicam que ensaios serão realizados no PDE. Cada ensaio é apresentado numa linha indicando os níveis assumidos pelos fatores, seguindo a ordem em que esses aparecem na linha indicada por `COMANDO`. Um asterisco no fator indica que todos os níveis serão testados.
- Considerando que um ensaio pode descrever uma série de experimentos, podem ocorrer sobreposições e algum experimento pode ser repetido mais de uma vez, o script não vai checar nem controlar isso. Para efeito de exemplo, um projeto fatorial completo (testar todas as combinações de níveis dos fatores) pode ser descrito facilmente pelo ensaio `*, *, *`.

Seguindo a descrição da configuração no exemplo, o script `projexp` fará executar os experimentos:

```

prog 5 -f 10 -d 1
prog 5 -f 10 -d 0
prog 20 -f 40 -d 0
prog 5 -f 1.5 -d 1
prog 20 -f 1.5 -d 1
prog 60 -f 1.5 -d 1

```

As saídas impressas pelo programa `prog` nos experimentos serão gravadas dentro do arquivo texto `saidas.txt`. Por exemplo, se a saída do programa `prog` for uma linha de texto como “media=5.5 max=15.3”, um possível arquivo de log teria a seguinte aparência:

```
1  ## Projeto de experimentos para o programa PROG
2
3  EXPERIMENTO 5, 10, 1 — DURAÇÃO 00:00:00:013 s
4  media=5.5 max=15.3
5
6  EXPERIMENTO 5, 10, 0 — DURAÇÃO 00:00:01:027 s
7  media=6.5 max=11.3
8
9  EXPERIMENTO 20, 40, 0 — DURAÇÃO 00:00:00:103 s
10 media=9.1 max=22.7
11
12  etc ...
13
14  ## fim do exemplo
```

Um projeto de experimentos completo envolveria analisar os dados coletados, mas isso não será contemplado pelo script objetivo deste trabalho. Bom estudo.