Universidade de Fortaleza

PPGIA – Introdução a estatística aplicada a ciência dos dados

Elaborado por Prof. Elizabeth S. Furtado em colaboração prof. Furlan Duarte

Individual: Ponto de participação

Postar a resposta no Unifor on-line

Objetivos:

* Conhecer os cuidados que se deve ter com a coleta de dados, para um experimento com usuário;
* Saber formalizar o problema para um modelo de pesquisa a ser investigado, definindo as hipóteses.
* Conhecer a amostra para um modelo de pesquisa de teste das preferências de usuários por um sistema interativo;
* Analisar os dados a partir de gráficos: *Boxplot*, histograma, gráficos de pontos e de Pareto.

Considere um processo (Figura 1), composto por um conjunto de etapas a serem realizadas pelos organizadores da investigação (testadores). Considere também o processo ilustrado na imagem abaixo, que é aplicado para a realização de uma sessão de teste de usabilidade de um sistema de software. Os testadores do sistema iniciam com a seleção dos indivíduos que farão parte de uma sessão individual de teste. Tais indivíduos, chamados usuários de um sistema, recebem explicações sobre o teste, realizam as atividades da investigação, e por fim, respondem um questionário.

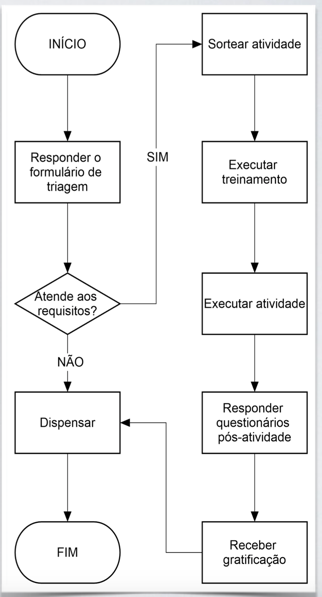


Figura 1. Processo de etapas idealizado pelo Fran de oliveira

Considere ainda que esse processo foi aplicado para analisar as preferências de 60 usuários entre dois sistemas, que eles não conheciam ainda. Os dados gerados pelo experimento servem para saber qual o melhor sistema na percepção dos usuários participantes do experimento.

Agora, para algumas etapas do processo mencionadas a seguir, responda as seguintes perguntas, para esse experimento realizado.

1ª. Questão. Considere que testadores analisaram na etapa “atende aos requisitos?” se um participante teria a idade requerida (entre 18 e 24 anos). Explique o que significaria, em um gráfico XY, se ter no eixo Y idade e no eixo X o perfil da amostra (número de indivíduos selecionados)?.

2ª. Questão. Considere que todos os indivíduos selecionados irão realizar as mesmas atividades usando cada um dos sistemas.

Pede-se:

1. Justifique a necessidade de “sortear atividade” para o exemplo em estudo.
2. Dê exemplos de variáveis que podem ser riscos para os resultados senão houvesse o sorteamento das atividades.
3. Considere que, é necessário que os sistemas sejam analisados com relação à 8 funcionalidades. Tecnicamente, isto significaria que cada indivíduo deveria fazer 8 atividades usando sistema. Mas, mesmo sorteando a ordem das atividades, os indivíduos se cansariam em ter que realizar todas e o experimento poderia também trazer viés. Como eles resolveriam este problema? Quais as implicações da(s) decisão(ões)?

3ª. Questão. Sobre a etapa “receber treinamento”, considere que ela se refere à forma como um experimento é esperado ser conduzido pelos investigadores. Por exemplo, todos os participantes devem receber o mesmo treinamento sobre cada sistema. O treinamento consistiria no testador mostrar a utilidade e estrutura do sistema, antes do usuário usar o sistema. Depois o testador diria que passaria X atividades para o usuário realizar usando o sistema. Para o exemplo em estudo, dê exemplos de variáveis que podem ser riscos para os resultados se os indivíduos não receberem este treinamento.

4ª. Questão.

Para cada um dos 4 problemas dados a seguir, pede-se:

a) defina a hipótese para um modelo de pesquisa da preferência de usuários por um sistema interativo;

b) coloque a amostra coletada para o problema no software de sua escolha;

c) Gere um ou mais gráficos mencionados na aula e ilustre seu gráfico. Analise a utilidade dos gráficos que você gerou.

1- Problema:

analisar as preferências de usuários entre dois sistemas

Amostra é prefsAB.cvs

2- Problema:

analisar as preferências de usuários entre três sistemas

Amostra é prefsABC.cvs

3- Problema:

analisar as preferências de usuários, por sexo, entre dois sistemas

Amostra é prefsABsex.cvs

4- Problema:

analisar as preferências de usuários, por sexo, entre dois sistemas

Amostra é prefsABCsex.cvs

5ª. Questão. Considere o arquivo, tempoporsite.cvs, em que 600 participantes, usaram um dos dois sites A ou B, e o tempo de uso de cada um deles foi computado.

Este arquivo descreve um teste A / B do site em que o tempo no site dos participantes foi medido em duas variações do mesmo site.

Metade foram expostos ao site A e metade a variação de A, que é o site B. Nenhum deles tinha usado o site antes.

A questão de investigação foi a seguinte: Qual site causa no usuário o interesse em ficar o maior tempo possível.

Pede-se:

a) Analise os conceitos apresentados em aula (desvio padrão, mediana).

Quantos sujeitos estavam neste teste A / B do site?

Quantos sujeitos foram expostos a cada variação do site A e B?

Qual foi o tempo médio para o site "B"?

Qual foi o desvio padrão de tempo para o site "A"?

b) Gere um ou mais gráficos mencionados na aula e ilustre seu gráfico. Analise a utilidade dos gráficos que você gerou.

Faça os histogramas para cada um dos sites

Faça um boxplot comparando os sites.