



SME0122

Introdução à Inferência Estatística

2º semestre de 2011

Prof. Cibele Russo

cibele@icmc.usp.br

<http://www.icmc.usp.br/~cibele>

Sala 3-162, ramal 6618

Aulas: Quartas e sextas-feiras das 8h10 às 9h50 na sala 4-003

Horário de atendimento: Terças-feiras das 16h às 17h ou mediante agendamento por e-mail.

Coteia Wiki: <http://wiki.icmc.usp.br/index.php/SME-122>

Objetivos gerais da disciplina

- Introduzir noções básicas da Inferência Estatística
- Discutir e desenvolver problemas de estimação (pontual e intervalar) de parâmetros da população com base em elementos da amostra
- Discutir e desenvolver testes de hipóteses a respeito de parâmetros da população com base em evidências fornecidas pela amostra

Programa - Parte 1

- Revisão de Probabilidades. Variáveis discretas e contínuas
- Estatística descritiva
- Inferência estatística. População e amostra. Princípios de estimação, função de verossimilhança, momentos amostrais
- Amostras Aleatórias, Estatísticas e Distribuições Amostrais
- Estimador de Mínimos Quadrados
- Estimador de Máxima Verossimilhança
- Noções sobre Estatística Bayesiana: O estimador de Bayes
- Estimação por intervalo. Aplicações

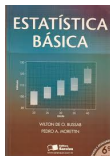
Programa - Parte 2

- Testes de Hipóteses: Testes paramétricos
- Teste para a média com variância conhecida
- Teste para a média com variância desconhecida
- Nível descritivo
- Testes Qui-Quadrado
- Comparação de duas médias
- Análise de Variância
- Análise de Regressão. Aplicações.

Bibliografia principal



Magalhães, Marcos Nascimento e Lima, Antonio Carlos Pedroso de Lima: Noções de probabilidade e estatística. São Paulo EDUSP 2005.



Bussab, W. O., Morettin, P. A. – Estatística Básica, Saraiva (preferencialmente a 6a Edição)

Bibliografia complementar

- DeGroot, M. H. and Schervish, M. J., "Probability and Statistics", 3th. Edition, Addison-Wesley, 2001.
- Box, G. E., Hunter, W. and Hunter, J.S. Statistics for Experimenters, Wiley, 1978.
- Mood, A. M., Graybill, F. A.; Boes, D. C. Introduction to the Theory of Statistics, 3rd edition, McGraw-Hill, Singapore, 1974.
- Walpole, R. E., Myers, R.H.; Myers, S. L., Ye, K. - Probability and Statistics, 7th. Ed. Prentice Hall, 2004.

Critério de Avaliação

- $MF = (2 P1 + 3 P2) / 5$

P1: Prova 1; P2: Prova 2

(**Sub:** somente com justificativa oficial, por exemplo atestado médico reconhecido pela UBAS)

- Critério de Recuperação (Rec)

Se $3 \leq MF < 5$: O aluno poderá fazer a prova Rec

Nota da Rec: NR

Nova média final: NMF

$NMF = MF + (NR/2,5)$, se $NR \geq 7,5$; ou $\text{Max } MF, NR$, se

$NR \leq 5,0$; ou $5,0$, se $5,0 \leq NR < 7,5$

Datas importantes

- Datas importantes

1ª prova: 07/10/2011 das 8h às 10h.

2ª prova: 07/12/2011 das 8h às 10h.

Prova de recuperação: 20/01/2012 das 8h às 10h.

Algumas regras

- A USP exige no mínimo 70% de presença nas aulas
- A docente poderá fazer a chamada em qualquer instante do horário da aula, mesmo se os alunos já tiverem assinado a lista de presença, e eventuais ausências não serão abonadas
- Questionamentos a respeito da matéria poderão ser feitos a qualquer momento
- Conversas paralelas sobre outros assuntos não serão toleradas - os outros alunos têm direito a assistir a aula em silêncio

Algumas regras

- Os alunos devem desligar seus telefones celulares durante a aula
- Não é permitido o uso de notebook durante as aulas, exceto em possíveis aulas práticas
- Ao enviar e-mail para a docente, identificar o código da disciplina no assunto SME0122
- Eventuais dúvidas, consultar Normas da USP
<http://www.usp.br/leginf/>

Algumas regras - Sobre provas

- Será permitido o uso de calculadora **própria**
- Fraude em prova implicará a reprovação direta do(s) envolvido(s)
- Falsidade ideológica na realização da prova (ou assinatura de presença) implicará abertura de processo disciplinar

Apoio computacional



<http://www.r-project.org/>

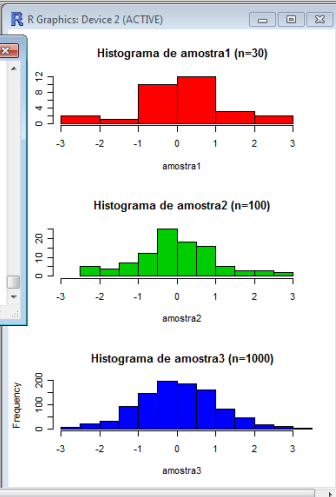


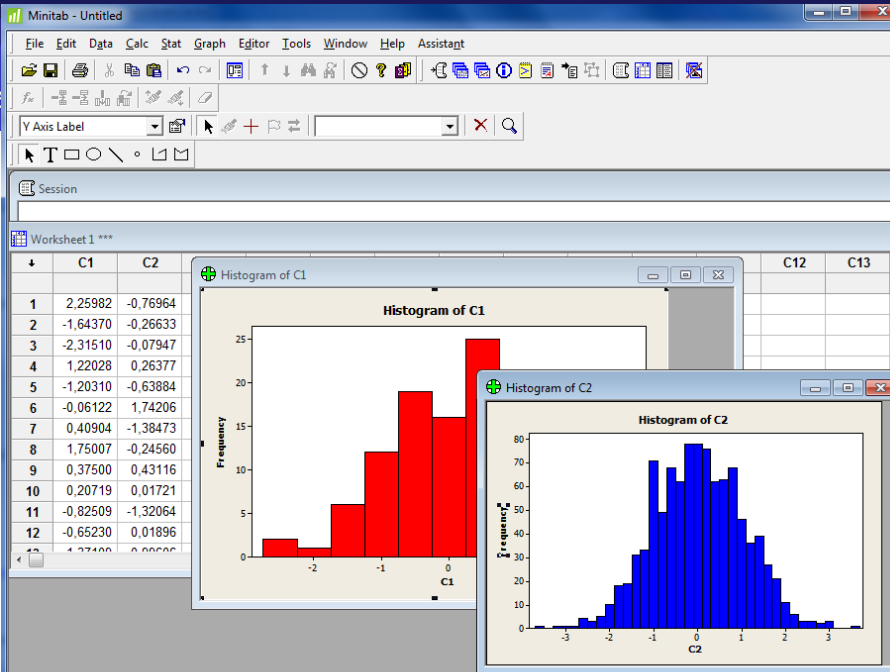
<http://www.minitab.com>



<http://www.doornik.com/download.html>

```
RGui
Arquivo  Editar  Visualizar  Misc  Pacotes  Janelas  Ajuda
[Icons]
R Console
> amostra1 <- rnorm(30,0,1)
> amostra2 <- rnorm(100,0,1)
> amostra3 <- rnorm(1000,0,1)
>
> faixax <- range(amostra1, amostra2, amostra3)
>
> par(mfrow=c(3,1))
>
> hist(amostra1, xlim=faixax, col=2, main="Histograma de amostra1 (n=30)")
> hist(amostra2, xlim=faixax, col=3, main="Histograma de amostra2 (n=100)")
> hist(amostra3, xlim=faixax, col=4, main="Histograma de amostra3 (n=1000)")
>
>
>
>
> |
```





*OxEdit - F:\Disciplinas\Inferência Estatística 2011 02\1Motivação.ox

File Edit Search View Run Tools Window Help



rel 1Motivação.ox +Ox Output

```

1
2 #include <oxstd.h>
3
4 semente();
5
6 main()
7 {
8     ranseed(semente());
9
10    decl amostra1, amostra2, amostra3, amostra4;
11
12    amostra1 = rann(10,1); //Gera amostra (n=30) da normal
13    amostra2 = rann(100,1); //Gera amostra (n=100) da normal
14    amostra3 = rann(1000,1); //Gera amostra (n=1000) da normal
15    amostra4 = rann(10000,1); //Gera amostra (n=10000) da normal padrão
16
17    print("  Amostra 1   Amostra 2   Amostra 3   Amostra 4  \n");
18    print("          n=10      n=100      n=1000     n=10000  \n");
19    print("Medias", meanc(amostra1) ~ meanc(amostra2) ~ meanc(amostra3) ~ meanc(amostra4)); //Imprime médias amostrais
20    print("Variancias", variance(amostra1) ~ variance(amostra2) ~ variance(amostra3) ~ variance(amostra4)); //Imprime variâncias amostrais
21
22 }
23
24 /* Função semente, que gera uma semente de acordo com a hora do relógio */
25 semente()
26 {
27     decl h, m, s, stim = time();
28     sscanf(stim[0:1], "%d", &h);
29     sscanf(stim[3:4], "%d", &m);
30     sscanf(stim[6:7], "%d", &s);
31     return h * 3600 + m * 60 + s;
32 }

```

*OxEdit - Ox Output

File Edit Search View Run Tools Window Help



rel 1Motivação.ox +Ox Output

```

----- Ox at 11:44:42 on 02-Aug-2011 -----

Ox Console version 5.10 (Windows) (C) J.A. Doornik, 1994-2008
This version may be used for academic research and teaching only

  Amostra 1   Amostra 2   Amostra 3   Amostra 4
          n=10      n=100      n=1000     n=10000
Medias
-0.10071    0.045101   -0.012733   -0.0064810
Variancias
  0.65457    0.83841   0.98516     1.0031

```

Motivação - Alguns exemplos

O tempo de vida de chips de computadores de uma determinada marca são normalmente distribuídos com parâmetros média $\mu = 1,4 \times 10^6$ horas e desvio-padrão $\sigma = 3 \times 10^5$ horas. Qual a probabilidade aproximada de, num lote com 100 chips, pelo menos 20 terem tempo de vida menor que $1,8 \times 10^6$ horas?

Motivação - Alguns exemplos

O tempo de vida de um tipo de componente eletrônico tem distribuição exponencial com parâmetro λ desconhecido. Uma amostra aleatória de n desses componentes foi testada durante T horas e observou-se o número X de componentes que falharam. Como obter um bom estimador de λ baseado em X ?

Motivação - Alguns exemplos

Um aluno faz um teste de múltipla escolha com 10 questões, cada uma com 5 alternativas (somente uma alternativa correta). O aluno acerta 4 questões. É possível deduzir (estatisticamente) que este aluno sabe a matéria?

O que é estatística?

“Estatística é um conjunto de técnicas que permite, de forma sistemática, **organizar, descrever, analisar e interpretar** dados oriundos de estudos ou experimentos, realizados em qualquer área do conhecimento”

(Magalhães e Lima. Noções de Probabilidade e Estatística, Edusp, 2002).

O que é probabilidade?

“Probabilidade é a teoria matemática utilizada para estudar a **incerteza** oriunda de fenômenos de caráter *aleatório*.”

(Magalhães e Lima. Noções de Probabilidade e Estatística, Edusp, 2002).

O que é inferência estatística?

“Inferência estatística é o estudo de técnicas que possibilitam a **extrapolação**, a um grande conjunto de dados, das informações e conclusões obtidas a partir de subconjuntos de valores, usualmente de dimensão muito menor.”

(Magalhães e Lima. Noções de Probabilidade e Estatística, Edusp, 2002).