CE 002 - Estatística I

Agronomia - Turma B

Professor Walmes Marques Zeviani

Laboratório de Estatística e Geoinformação
Departamento de Estatística
Universidade Federal do Paraná

1º semestre de 2012

Sumário

Laboratório de Estatística e Geoinformação

Estrutura do curso

O curso é divido em 3 partes:

Probabilidade noções de cálculo de probabilidade, variáveis aleatórias e propriedades, modelos para distruição de variáveis aleatórias;

Estatística descritiva métodos para amostragem, organização, tratamento, análise, apresentação e interpretação de dados. Emprego de estatísticas descritivas e representações gráficas.

Inferência estatística ferramentas para fazer inferências baseado em dados amostrais. Inclui métodos de estimação pontual e intervalar, testes de hipótese e predição.

Laboratório de Estatística e Geoinformação

O que é estatística?

O que é estatística?

A estatística utiliza-se das **teorias probabilísticas** para explicar a frequência da ocorrência de eventos, tanto em estudos observacionais quanto em experimento **modelar** a aleatoriedade e a incerteza de forma a estimar ou possibilitar a **previsão de fenômenos** futuros, conforme o caso.

A estatística é uma ciência que se dedica à **coleta**, **análise** e **interpretação** de dados. Preocupa-se com os métodos de recolha, organização, resumo, apresentação e interpretação dos dados, assim como **tirar conclusões** sobre as características das fontes donde estes foram retirados, para melhor **compreender** as situações.

fonte: Wikipedia (`http://pt.wikipedia.org/wiki/Estatística`)

Laboratório de Estatística e Geoinformação

A estatística moderna é uma tecnologia quantitativa para a ciência experimental e observacional que permite avaliar e estudar as **incertezas** e os seus efeitos no planejamento e interpretação de experiências e de observações de **fenômenos** da natureza e da sociedade.

Segundo Rao (1999), a estatística é uma ciência que estuda e pesquisa sobre: o levantamento de dados com a máxima quantidade de **informação** possível para um dado custo; o processamento de dados para a quantificação da quantidade de incerteza existente na resposta para um determinado problema; a **tomada de decisões** sob condições de incerteza, sob o menor **risco** possível. Finalmente, a estatística tem sido utilizada na pesquisa científica, para a **otimização** de recursos econômicos, para o aumento da qualidade e produtividade, na otimização em análise de decisões, em questões judiciais, previsões e em muitas outras áreas.

fonte: `http://vsites.unb.br/ie/est/complementar/estatistica.htm`

Laboratório de Estatística e Geoinformação

Por que estudar estatística?

- A disponibilidade de parelhos modernos, muitos dos quais acoplados a computadores, permitem a quantificação de muitos fenômenos. A massa de dados gerada precisa ser analisada adequadamente.
- Na Ciência, são realizados estudos experimentais ou observacionais, em que o interesse é comparar grupos/tratamentos ou ainda determinar fatores prognósticos/risco importantes.
- O material biológico estudado é sempre uma amostra e o objetivo final é tirar conclusões sobre toda a população de interesse com base na amostra.

fonte: `http://leg.ufpr.br/~silvia/CE055/node4.html`

Laboratório de Estatística e Geoinformação

Por que estudar estatística?

A Estatística está presente em todas as áreas da ciência que envolvam o planejamento do experimento, a construção de modelos, a coleta, o processamento e a análise de dados e sua consequente transformação em informação, para **postular**, refutar ou validar **hipóteses científicas** sobre um fenômeno observável. Desta forma, a Estatística pode ser pensada como a **ciência de aprendizagem a partir de dados**. No mundo moderno, a alta competitividade na busca de tecnologias e de mercados tem provocado uma constante corrida pela informação. Essa é uma tendência crescente e irreversível. O aprendizado a partir de dados é um dos desafios mais relevantes da era da informação em que vivemos. Em linhas gerais, podemos dizer que a Estatística fornece técnicas e métodos de análise de dados que auxiliam o processo de **tomada de decisão** nos mais variados problemas **onde existe incerteza**.

fonte: `http://www.famat.ufu.br/node/170`

Laboratório de Estatística e Geoinformação

Probabilidades - Objetivos do estudo

- Entender e descrever espaços amostrais e eventos para experimentos aleatórios;
- Interpretar probabilidades e usar probabilidades de resultados para obter probabilidades de eventos;
- Calcular probabilidades e eventos conjuntos, como união e intersecção de eventos individuais;
- Calcular e interpretar probabilidades condicionais;
- Determinar a independência de eventos;
- Usar o teorema de Bayes;
- Entender variáveis aleatórias.

Laboratório de Estatística e Geoinformação

Definições

Experimento aleatório (ou fenômeno aleatório) é a situação na qual observamos um sistema que produz resultados que não poder sem previstos, muito embora seja repetido toda vez da mesma maneira. Exemplo: lançar um dado/moeda, pesar um fruto a.a., etc.

Espaço amostral (Ω) Conjunto de todos os resultados possíveis de um experimento aleatório. O Ω pode conter um número finito ou infinito de pontos. Exempo: {cara,coroa}, {1,2,3,4,5,6}, $\mathbb{N}$, $\mathbb{R}$.

Ponto amostral (ω) são elementos de um espaço amotral. Exemplo: $\omega_1 =$ cara, $\omega_2 =$ coroa.

Evento é um subconjunto de pontos do espaço amostral de um experimentos aleatório. Exemplo: A = "sair face par", B = "sair face menor que 3".

Exemplos

Experimento lançar o dado e observar o resultado da face.

Espaço amostral $\Omega = \{1,2,3,4,5,6\}$.

Pontos amostrais $\omega_1 = 1$, $\omega_2 = 2$, $\ldots$, $\omega_6 = 6$.

Eventos A = "sair face par", B = $\{\omega : \omega \leq 4\}$.

Experimento retirar uma carta de um baralho de 54 cartas.

Espaço amostral $\Omega = \{\clubsuit A, \clubsuit 2, \ldots, \heartsuit A, \ldots, \spadesuit A, \ldots, \diamondsuit J, \diamondsuit Q, \diamondsuit K\}$.

Pontos amostrais $\omega_1 = \clubsuit A$, $\omega_2 = \clubsuit 2$, $\ldots$, $\omega_{54} = \diamondsuit K$.

Eventos A = "sair um ais", B = "sair uma letra", C = "sair carta de $\clubsuit$".

Experimento pesar um fruto a.a.

Espaço amostral $\Omega = \mathbb{R}^+$.

Pontos amostrais espaço amostral é infinito.

Eventos A = "peso menor que 50g", B = $\{x : x \geq 100\text{g}\}$.

Operações com eventos

União é o evento que consiste da união de **todos** os pontos amostrais dos eventos que a compõem. Denotamos a união do evento A com B por $A \cup B$.

Interseção é o evento composto pelos pontos amostrais **comuns** aos eventos que a compõem. Denotamos a interseção de A com B por $A \cap B$.

Complemento é o conjunto de pontos do espaço amostral que não estão no evento. Denotamos o complemento do evento A por A^c.

Laboratório de Estatística e Geoinformação

Exemplo

Considere o lancamento de um dado e os eventos $A = \{1,2,3,4\}$, $B = \{\omega : \omega \leq 3\}$, $C =$ "face par", $D =$ "face primo".

Uniões

$A \cup B = \{1,2,3,4\}$, $A \cup C = \{1,2,3,4,6\}$, $A \cup D = \{1,2,3,4,5\}$.

Intersecções

$A \cap B = \{1,2,3\}$, $A \cup C = \{2,4\}$, $A \cup D = \{1,2,3\}$.

Complementos

$A^c = \{5,6\}$, $B^c = \{\omega : \omega > 3\}$, $D^c = \{4,6\}$.

Laboratório de Estatística e Geoinformação

Tipos de eventos

Disjuntos (mutuamente exclusivos) são eventos que possuem intersecção nula, ou seja, $A \cap B = \{\varnothing\}$. compõem. Denotamos a união do evento A com B por $A \cup B$.

Complementares são eventos que a união é o espaço amostral, ou seja, $A \cup B = \Omega$. Denotamos a interseção de A com B por $A \cap B$.

Exaustivos (disjuntos e complementares) são eventos que atendem ambas propriedades.

Laboratório de Estatística e Geoinformação

Probabilidades

Probabilidade de um evento

quando o espaço amostral é discreto, a probabilidade de um evento A, denotada por $P(A)$, é a soma das probabilidades do elementos do espaço amostral (ω) que compõem A.

Axiomas de probabilidade

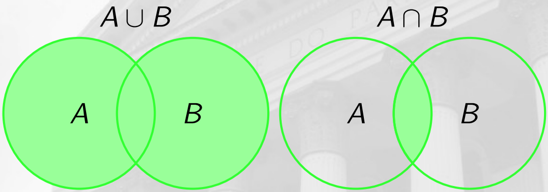
- $P(\Omega) = 1$.
- $0 \leq P(A) \leq 1$.

Laboratório de Estatística e Geoinformação

Regra da adição

Probabilidade de uma União

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \tag{1}$$



Probabilidade condicional

Permite avaliar a probabilidade quando informações adicionais se tornam disponíveis ou quando certas condições são assumidas. A probabilidade de um evento B dado a realização do evento A é representado por

$$P(B|A). \tag{2}$$

Qual a probabilidade de sair um número par ao lançar um dado? Qual a probabilidade de sair um número par sabendo que o resultado é (ou condicional ao resultado ser) menor ou igual a 3?

Fórmula da probabilidade condicional

$$P(B|A) = \frac{P(A \cap B)}{P(A)}, \text{ assumindo que } P(B) \neq 0. \tag{3}$$

Exemplo

Um lote 400 de produtos eletrônicos fabricados é classificado quando à apresentar defeito da superfície e defeito de funcionamento. A classificação é resumida na seguinte tabela

	Falha na superfície		
Falha no funcionamento	sim (evento S)	não	total
sim (evento F)	10	18	28
não	30	342	372
total	40	360	400

Qual a probabilidade de retirar um elemento com defeito funcional dado que apresenta defeito na superfície?

$$P(F|S) = 10/40 = 0.25. \tag{4}$$

Aplicando a fórmula da probabilidade condicional verificamos que

$$P(F|S) = \frac{P(F \cap S)}{P(S)} = \frac{10/400}{40/400} = 10/40 = 0.25. \tag{5}$$

Árvore de probabilidades

É um diagrama para visulizar probabilidades condicionais. O diagrama inicia da esquerda e cada ramo representa um possível resultado para o primeiro evento que é não condicional, ou marginal. Em seguida, cada ramo é dividido para representar os resultados do segundo evento dado os possíveis resultados do primeiro evento, ou seja, uma probabilidade condicional. Seguindo essa regra, o diagrama pode representar probabilidades condicionais para muitos eventos. Em cada nó do diagrama a soma das probabilidades é 1.

Laboratório de Estatística e Geoinformação

Exemplo

Árvore de probabilidades