Introdução a Estatística Aplicada I - Teoria de Probabilidades

Material adaptado de "Laws of probability, Bayes´Theorem, and the central Limit Theorem" de G. Jogesh Babu

Prof^{a} Doutora Raquel Nicolette

Sumário

Introdução

Formalização matemática
- Espaço amostral, experimento e evento
- Variáveis aleatórias

Probabilidade
- Definições de probabilidade
- Regra da Adição
- Probabilidade condicional
- Regra da multiplicação
- Independência de eventos
- Teorema de Bayes

Distribuições de Probabilidade
- Distribuições discretas
- Distribuições contínuas

Existem fenômenos aleatórios na natureza?

Jogar uma moeda é aleatório?

Não necessariamente. Mas modelando as saídas como aleatórias tem-se uma representação parciomoniosa da realidade.

De que maneira um *elétron* irá girar?

Isso é aleatório?

Não podemos excluir a possibilidade de uma nova teoria ser inventada algum dia que poderia explicar a rotação, mas modelá-lo como aleatório é bom o suficiente.

Aleatoriedade não é uma imprevisibilidade total; mas podemos quantificar o que é imprevisível.

Se a **matemática** e a **teoria de probabilidades** fosse bem entendidas diversos séculos atrás, como são hoje, talvez as pessoas tivessem modelado a ocorrência de um eclipse solar como um evento aleatório, uma probabilidade atribuída baseada na ocorrência empírica.

Subsequentemente, alguém poderia ter revisado o modelo, observando que o eclipse solar ocorre somente no dia de lua nova.

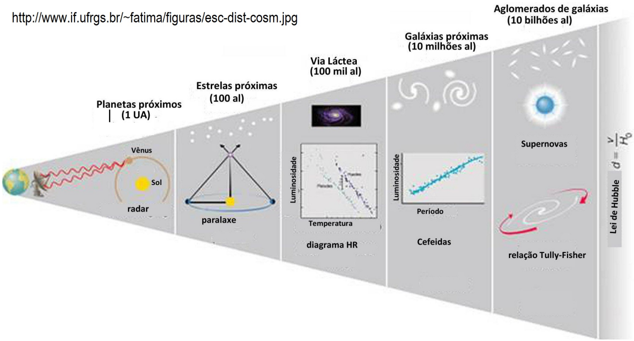
Após algum tempo, o fenômeno poderia ser completamente entendido e o modelo ter mudado de estocástico ou aleatório, para um modelo determinístico.

Assim muitas vezes encontramos eventos cujo resultado é incerto. A incerteza poderia ser causa de:

- nossa incapacidade de observar com precisão todas as entradas necessárias para calcular o resultado
- custo excessivo para observar todos os dados de entrada
- falta de entendimento sobre o fenômeno
- dependência das escolhas a serem feitas no futuro, como o resultado de uma eleição

Escada de distância cósmica

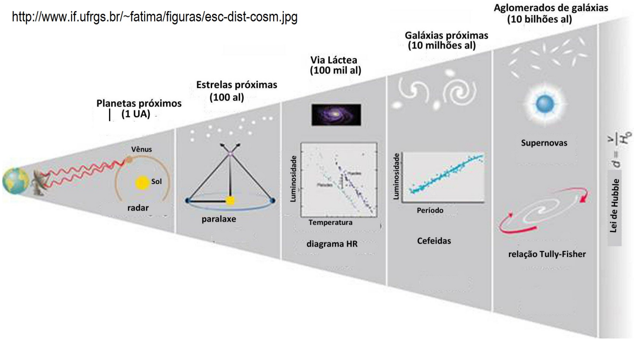
http://www.if.ufrgs.br/~fatima/figuras/esc-dist-cosm.jpg



Muitas objetos no sistema solar foram medidos com bastante precisão pelos gregos antigos e babilônicos usando métodos geométricos e trigonométricos.

Escada de distância cósmica

http://www.if.ufrgs.br/~fatima/figuras/esc-dist-cosm.jpg

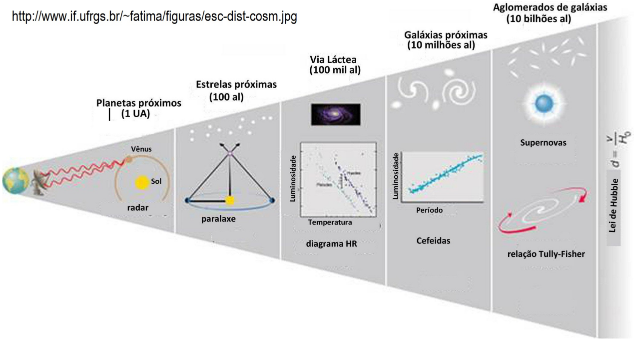


A distâncias das estrelas do segundo degrau são encontradas por ideias de aralaxe, calculando o desvio angular ao longo de 6 meses.

Feito pela primeira vez pelo matemático Friedrich Bessel com precisão de até 100anos-luz, o err oaqui é maior que no intervalo anterior.

Escada de distância cósmica

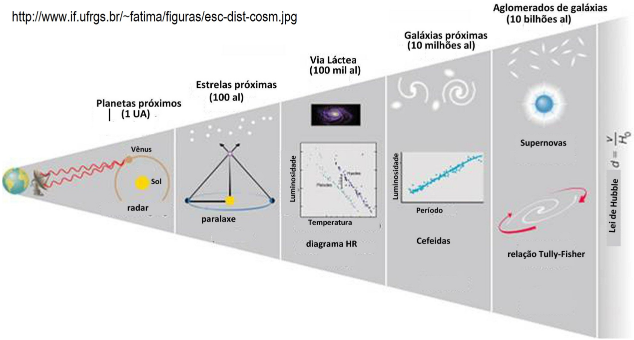
http://www.if.ufrgs.br/~fatima/figuras/esc-dist-cosm.jpg



Distâncias de estrelas moderadamente distantes podem ser obtidas através de uma combinação de brilho aparente e distância para as estrelas mais próximas usando o diagrama de Hertsprung-Russel. Desta forma o método funciona para estrelas até 300.000 anos-luz e o erro é significativamente maior.

Escada de distância cósmica

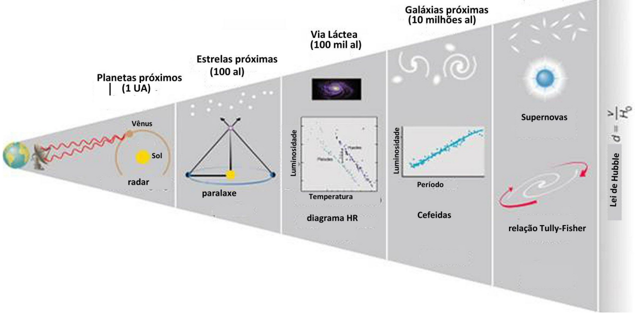
http://www.if.ufrgs.br/~fatima/figuras/esc-dist-cosm.jpg



A distância ao próximo e último leote de estrelas é obtida através da alocação das oscilações de seus brilhos. Este método funciona até estrelas há 13.000.000 anos-luz.

Escada de distância cósmica

http://www.if.ufrgs.br/~fatima/figuras/esc-dist-cosm.jpg

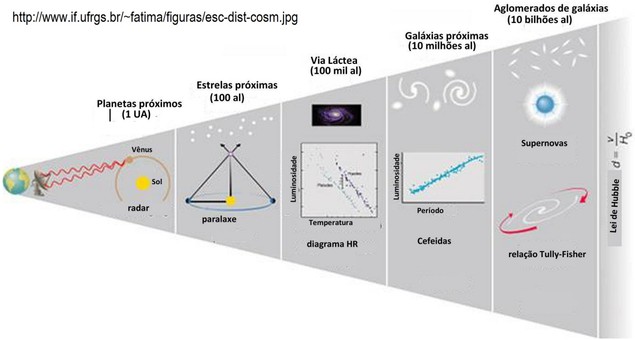


Em cada degrau da escada erros e incertezas são acumulados.

Cada passo herda todos os problemas anteriores, e também os erros intrínsecos a cada passo tendem a aumentar para objetos mais distantes.

Escada de distância cósmica

http://www.if.ufrgs.br/~fatima/figuras/esc-dist-cosm.jpg



Então precisamos entender a incerteza. E a única maneira de compreender essa noção cientificamente é fornecer uma estrutura para ela. Essa estrutura deve ser suficientemente rica para se prestar a quantificação.

Moedas, dados e etc

- A estrutura que precisa ser entendida ao jogar uma moeda é intuitiva. Nós assumimos a probabilidade $\frac{1}{2}$ para cair **cara** e a probabilidade $\frac{1}{2}$ para cair **coroa.**

- Similarmente paraa saída 1,2,3,4,5 e 6 do lançamento de um dado assume probabilidade $\frac{1}{6}$

- Seguindo o mesmo pensamento para o resultado de 000001 a 999999 de um bilhete de loteria ser o vencedor é de $\frac{1}{99999}$.

Experimentos e espaço amostral

Um experimento, que ao ser realizado sob as mesmas condiçoes não produz os mesmos resultados, é denominado um **experimento aleatório**.

Exemplo:

- Lançar uma moeda
- Jogar um dado
- Ganhar na loteria

O conjunto de todos os possíeis resultados de um experimento aleatório é chamado de **espaçoo amostral** (Ω). Pode conter um número finito ou infinito de pontos.

Um espaço amostral finito ou infinito enumerável é chamado de **espaço amostral discreto**.

Exemplo:

- Lançar uma moeda: $\Omega = \{cara, coroa\}$
- Jogar um dado : $\Omega = \{1, 2, 3, 4, 5, 6\}$

Eventos

Todo resultado ou subconjunto de resultados de um experimento aleatório, é chamado de **evento**.

Assim, todo subconjunto $\mathbb{E} \subseteq \Omega$ é chamado de evento.

Em termos de probabilidade rigorosos um espaço de probabilidade consiste em:

- um espaço amostral Ω
- um conjunto $\mathbb{F}$ de subconjuntos de Ω que são chamados de eventos
- uma função $\mathbb{P}$ que atribui uma probabilidade a cada evento que obedecer alguns axiomas.

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 = $\{\omega : \omega \leq 5\}$.

- Experimento: calcular a distância de uma estrela
- Espaço amostral: $\Omega = \mathbb{R}^+$.
- Pontos amostrais: espaço amostral é infinito.
- Eventos: A = "Distância menor que 4,5 anos-luz", B = $\{x : x \geq 6\text{anos-luz}\}$.

Variáveis aleatórias

De vota aos dados. Supondo que estamos apostando e temos uma situação como esta:

Resultado	1	2	3	4	5	6
Reais líquidos obtidos	-8	2	0	4	-2	4

O nosso interesse no resultado é apenas através da sua ssociação com o valor monetário. Então estamos interessados em uma função do espaço de resultados de Ω para os números reais $\mathbb{R}$.

Essa função é chamada de variável aleatória.

Seja X o montante de dinheiro ganho em um lance de dados. Estamos interessados em $\{X = x\} \subset \Omega$

{X=0}	para	**x=0**	|	evento	**{3}**
{X=2}	para	**x=2**	|	evento	**{2}**
{X=4}	para	**x=4**	|	evento	**{4,6}**
{X=-2}	para	**x=-2**	|	evento	**{5}**
{X=-8}	para	**x=-8**	|	evento	**{1}**

Notação: letras maiúsculas representam a variável aleatória enquanto as letras minúsculas os valores que a variável assume.

Seja X o montante de dinheiro ganho em um lance de dados. Estamos interessados em $\{X = x\} \subset \Omega$

{X=0}	para	**x=0**	|	evento	**{3}**
{X=2}	para	**x=2**	|	evento	**{2}**
{X=4}	para	**x=4**	|	evento	**{4,6}**
{X=-2}	para	**x=-2**	|	evento	**{5}**
{X=-8}	para	**x=-8**	|	evento	**{1}**

Exemplo: pra encontrar a probabilidade que você ganhou 4 reais, isto é $P(X = 4)$ você quer encontrar a probabilidade atribuída ao evento $\{4, 6\}$, portanto

$$P(\omega \in \Omega : X(\omega) = 4) = P(4, 6) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

$$\{X = x\} \text{ é uma abreviatura de } \{\omega \in \Omega : X(\omega) = x\}$$

Se resumirmos os possíveis valores diferentes de zero de $P(\{X = x\})$, obtemos uma função chamada de função de massa de probabilidade de x, denotada por $f(x)$ ou $p(x)$ ou $f_X(X)$.

$$f_X(X) = P(\{X = x\} = \begin{cases} 1/6 & para \quad x = 0 \\ 1/6 & para \quad x = 2 \\ 2/6 & para \quad x = 4 \\ 1/6 & para \quad x = -2 \\ 1/6 & para \quad x = -8 \\ 0 & para \quad qualquer \quad outro \quad valor \quad de \quad x \end{cases}$$

Variável aleatória discreta

1. Uma função definida no espaço amostral Ω e com valores num conjunto enumerável de pontos é chamada de variável aleatória discreta.
   X assume o valor x_i e podemos calcular a probabilidade de X = x_i ou seja $P(X = x_i)$ e denotamos por:

$$P(x_i) = P(X = x_i)$$

2. Valor esperado

$$E(X) = \sum_{i=1}^{n} x_i P(X = xi) = \sum_{i=1}^{n} x_i pi$$

   se a v.a. assume valores $x_1, x_2, ..., x_n$

3. Variância: $\sigma^2 = Var(X) = E(X - \mu)^2$

$$Var(X) = E(X^2) - (E(X))^2$$

4. Uma função definida no espaço amostral Ω e assumindo valores reais é chamada de variável aleatória contínua.

Neste caso temos uma função densidade de probabilidade f.d.p que tem as propriedades:

1. $f(x) \geq 0$
2. $P(a \leq X \leq b) = \int_a^b f(x)dx$
3. $\int_{-\infty}^{\infty} f(x)dx$

Valor esperado: $E(X) = \int_{-\infty}^{\infty} xf(x)dx$

Variância:

$\sigma^2 = Var(X) = E(X - \mu)^2$

$$Var(X) = E(X^2) - (E(X))^2$$

Variável aleatória contínua

Uma v.a. é classificada como contínua se assume valores em qualquer intervalo dos números reais, ou seja, um conjunto de valores não enumerável. Dessa forma, não é possível atribuir probabilidades para um ponto específico, apenas para intervalos da reta.

Atribuímos probabilidades à intervalos de valores,através de uma **função**. Portanto, as probabilidades são representadas por áreas.

A **função densidade de probabilidade** (fdp) atribui probabilidades à intervalos de valores do tipo $[a, b]$, e é definida como

$$P[a < x < b] = \int_a^b f(x)dx$$

com as seguintes propriedades

i) é uma funçãonão negativa

$$f(x) \geq 0$$

ii) a área total sob a curva deve ser igual a 1

$$\int_{-\infty}^{+\infty} f(x)dx = 1$$

Variável aleatoria contínua

Esperança

A esperança de uma v.a. contínua tem o mesmo sentido e interpretação que tem-se em uma v.a. discreta: é a **média** ou **valor esperado** da V.A.

$$(X) = \int_{-\infty}^{+\infty} x \cdot f(x)dx$$

Variança

É o grau de dispersão dos valores de uma v.a. em relação a sua média (X). A forma geral para o cálculo em V.A.s contínuas é

$$\begin{aligned}(X) &= [(X - (X))^2] \\ &= \int_{-\infty}^{+\infty} [x - (X)]^2 \cdot f(x)dx\end{aligned}$$

Definições de probabilidade

Definição Clássica

Consideramos um espaço amostral Ω com $n(\Omega)$ eventos simples, supondo que sejam **igualmente prováveis**. Seja A um evento de Ω, composto de $n(A)$ eventos simples. A probabilidade de A, $P(A)$ será

$$P(A) = \frac{n(A)}{n(\Omega)}$$

Definições de probabilidade

Definições de probabilidade

Exemplo: Se um dado fosse lançado **10** vezes, e contássemos quantas vezes saiu a face 4, qual seria a probabilidade desse evento?

```
Console ~/
> ##Tamanho da amostra##
> n<-10
>
> ## Objeto para guardar o resultado##
> x<-numeric(n)
>
> ## Repetição##
> for(i in 1:n){
+
+     ##Amostra de tamanho 1 dos elementos de 1 a 6##
+
+     x[i]<-sample(1:6,size=1)}
>
> ##Total de elementos com valor = 4##
> sum(x==4)
[1] 2
>
> ##Proporção de elementos com valor = 4##
> sum(x==4)/length(x)
[1] 0.2
```

Definições de probabilidade

Exemplo: Repetindo diversas vezes o mesmo experimento anterior: se um dado fosse lançado **100** vezes,**1.000** vezes, **10.000** vezes, **100.000** vezes e **1.000.000** vezes e contássemos quantas vezes saiu a face 4, qual seria a probabilidade desse evento?

n	proporção
10	0,2
100	0,19
1000	0,163
10000	0,1659
100000	0,1684
1000000	0,1672
10000000	0,1668

Definições de probabilidade

Então,

$$\lim_{n \to \infty} \frac{n(A)}{n} = P(A) \approx 0,1668$$

As probabilidades calculadas a partir de frequências relativas, são **estimativas** da verdadeira probabilidade

Lei dos Grandes Números

A Lei dos Grandes Números nos diz que as estimativas dadas pelas frequências relativas tendem a ficar melhores com mais observações.

Probabilidades

Axiomas de probabilidade

Vamos considerar **probabilidade** como sendo uma função $P(\cdot)$ que associa valores numéricos é um evento A do espaço amostral, e que satisfaz as seguintes condições

i) $P(\Omega) = 1$; $P(\varnothing) = 0$

ii) $0 \leq P(A) \leq 1$

iii) $P(A \cup B) = P(A) + P(B)$ se, e seomente se $A \cap B = \varnothing$

Os axiomas asseguram que as probabilidades podem ser interpretadas como **frequências relativas**.

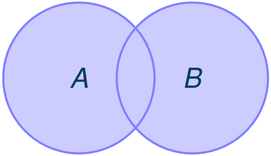
Regra da adição

Regra da adição

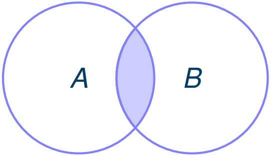
A probabilidade da união entre dois eventos quaisquer, A e B, é definida pela **regra da adição de probabilidades**

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$A \cup B$



$A \cap B$

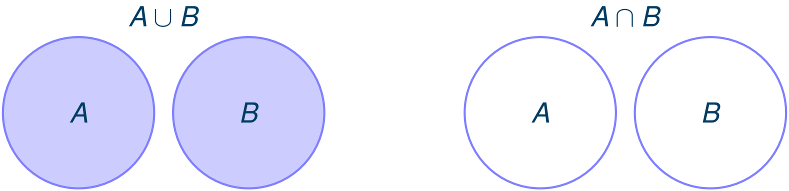


Regra da Adição

A regra da adição pode ser simplificada, **se e somente se** os eventos A e B forem **disjuntos** (ou mutuamente exclusivos)

$$P(A \cup B) = P(A) + P(B)$$

pois, neste caso, $A \cap B = \varnothing \quad \Rightarrow \quad P(A \cap B) = P(\varnothing) = 0$



Regra da Adição

Regra do complementar

Como consequência da regra da adição, obtemos que, para qualquer evento A,

$$P(A) = 1 - P(A^c)$$

Verifique através de $P(A \cup A^c) = P(A) + P(A^c) - P(A \cap A^c)$

Probabilidade condicional

Em diversos casos, um o fenômeno aleatório pode ser separado em etapas.

A informação do que ocorreu em uma determinada etapa pode influenciar nas probabilidades de ocorrências das etapas sucessivas.

Assim admitimos que ganhamos informação, e podemos *recalcular* as probabilidades de interesse.

Estas probabilidades *recalculadas* recebem o nome de **probabilidade condicional**.

Probabilidade condicional

Para entender a ideia de probabilidade condicional, considere o seguinte exemplo:

- Um dado foi lançado, cada um dos seus resultados é igualmente provável. Mas suponha que eu dê uma olhada e diga que X é par. Qual é a probabilidade do resultado é menor que 4?
- Qualé é a probabilidade do resultado pertencer a $\{1, 2, 3\}$ com essa "nova" informação?

$$\Omega = \{1, 2, 3, 4, 5, 6\},\ n(\Omega) = 6$$

$$A = \text{é par} = \{2, 4, 6\},\ n(A) = 3 \quad \Rightarrow \quad P(A) = \frac{n(A)}{n(\Omega)} = \frac{3}{6}$$

$$B = \text{é menor que } 4 = \{1, 2, 3\},\ n(B) = 3 \quad \Rightarrow \quad P(B) = \frac{n(B)}{n(\Omega)} = \frac{3}{6}$$

$$C = \text{é par dado que é menor que } 4 = \{2\},\ n(C) = 1 \quad \Rightarrow \quad P(C) = \frac{n(C)}{n(B)} = \frac{1}{3}$$

Probabilidade condicional

Definição

Para dois eventos A e B de um mesmo espaço amostral, $P(A|B)$ denota a probabilidade de A ocorrer, dado que B ocorreu, e é definido como

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

da mesma forma que a probabilidade de B ocorrer, dado que A ocorreu é dada por

$$P(B|A) = \frac{P(B \cap A)}{P(A)}$$

Probabilidade condicional

Voltando ao exemplo e aplicando a definição de probabilidade condicional:

$$P(A \cap B) = \frac{n(A \cap B)}{n(\Omega)} = \frac{1}{6}$$

$$P(B) = \frac{n(B)}{n(\Omega)} = \frac{3}{6}$$

$$\begin{aligned} P(A|B) &= \frac{P(A \cap B)}{P(B)} \\ &= \frac{1/6}{3/6} \\ &= \frac{1}{3} \end{aligned}$$

Probabilidade condicional

Verifica-se assim que tem-se duas maneiras de calcular a probabilidade condicional $P(A|B)$:

1. Diretamente, pela considerando a probabilidade de A em relação ao espaço amostral reduzido B
2. Empregando a definiçãoo acima, onde $P(A \cap B)$ e $P(B)$ são calculadas em relaçãoo ao espaço amostral original Ω

Regra da multiplicação

A regra da multiplicação deriva do conceito de probabilidade condicional. Dado que

$P(A|B) = \dfrac{P(A \cap B)}{P(B)}$ temos que

$P(A \cap B) = P(B) \cdot P(A|B)$

Desta forma obtemos a probabilidade de uma interseção pelo produto de uma probabilidade marginal com uma probabilidade condicional.

Regra da multiplicação

$$\begin{aligned} P(A \cap B) &= P(B) \cdot P(A|B) \\ &= P(A) \cdot P(B|A) \end{aligned}$$

Independência de eventos

Quando saber que o evento B ocorreu, não altera a probabilidade de ocorrência de A dizemos que os aventos A e B são **independentes**

Os eventos A e B são **eventos independentes** se a ocorrência de B não altera a probabilidade de ocorrência de A, ou seja, eventos A e B sâo independentes se

$$P(A|B) = P(A) \quad \text{e também que} \quad P(B|A) = P(B)$$

Com isso, e a regra da multiplicação, temos que

$$P(A \cap B) = P(B) \cdot P(A|B) = P(B) \cdot P(A)$$
$$P(A \cap B) = P(A) \cdot P(B|A) = P(A) \cdot P(B)$$

Isso significa que se dois eventos são independentes, a probabilidade de ocorrência simultânea $P(A \cap B)$ é o produto das probabilidades marginais, $P(A)$ e $P(B)$.

Teorema de Bayes

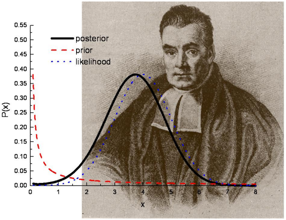
- Sejam $B_1, B_2, ...B_n$ eventos que formam uma partição do espaço amostral Ω
- Para quantificar nossas suspeitas que e seja B_i foi a causa da ocorrência de A, gostaríamos de obter $P(B_i|A)$
- Assim assumimos que encontrar $P(A|B_i)$ é direto para i
- Além disso, assumios que temos conhecimento prévio de $P(B_i)$ para cada i
- O Teorema da Bayes é a receita para obter a quantidade $P(B_i|A)$
- É a base da Inferência Bayesiana.

Simplificando, nosso objetivo em encontrar $P(B_i|A)$ é determinar como nossa observação A modifica nossas probabilidade B

A álgebra direta quevela que:

$$P(B_i|A) = \frac{P(B_i).P(A|B_i)}{P(A)} = \frac{P(B_i).P(A|B_i)}{\sum_{i=1}^{n} P(B_j).P(A|B_j)}$$

A identidade acima é o que chamamos de Teorema de Bayes, o qual é nomeado à ***Thomas Bayes***, um britânico do século XVIII, matematico e ministro presbiteriano.



Teorema de Bayes

$$P(B_i|A) = \frac{P(B_i).P(A|B_i)}{P(A)} = \frac{P(B_i).P(A|B_i)}{\sum_{i=1}^{n} P(B_j).P(A|B_j)}$$

Observando que o denominador não depende de i, podemos resumir o teorema de Bayes a sua essência

$$P(B_i|A) \propto P(B_i).P(A|B_i)$$

Frequentemente chamamos o lado esquerdo da equação acima como a probabilidade *a posteriori* de B_i, assim o teorema de Bayes expressado sucintamente, afirma que a probabilidade *a posterior* é proporcional a *verossimilhança* vezes a *priori*.

Teorema de Bayes

$$P(B_i|A) = \frac{P(B_i).P(A|B_i)}{P(A)} = \frac{P(B_i).P(A|B_i)}{\sum_{i=1}^{n} P(B_j).P(A|B_j)}$$

Existem muitas controvérsias e parentes paradoxos associados a probabilidades condiconais. A raiz dessa causa as vezes e uma especificação incompleta das condiçoes na formulação dos problemas, embora também existam alguns "paradoxos" que exploram a aparente incapacidade inerente das pessoas de modificar as probilidades prioris corretamente quando confrontadas com novas informações.

Distribuições discretas de probabilidade

Os modelos de probabilidade são utilizados para descrever vários fenômenos ou situações que encontramos na natureza, ou experimentos por nós construídos.

Tais modelos são expressos por uma famíliaa de **distribuições de probabilidade** que dependem de um ou mais **parâmetros**.

Uma V.A. fica completamente caracterizada pela sua **função de probabilidade** e seus parâmetros.

Distribuição de Bernoulli

Um experimento que assume somente dois valores é um experimento de *Bernoulli* e sua v.a. X assume dois valores 0 ou 1, com probabilidade p de sucesso e $(1-p)$ de fracasso.

1. $P(X=0)=1-p$
2. $P(X=1)=p$
3. $E(X)=p$
4. $Var(X)=p-p^2=p(1-p)$

Distribuição Binomial

1. n ensaios de Bernoulli
2. ensaios independente
3. probabilidade de sucesso é sempre igual a p, $0 < p < 1$

Neste caso dizemos que a v.a. X tem distribuição Binomial com parâmentros n e p e representamos por $X \; Bin(n; p)$

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

$$P(X = k) = p^k (1 - p)^{n-k}$$

$$E(X) = np$$

$$Var(X) = np(1 - p)$$

Distribuição de Poisson

Definição: Seja um experimento realizado nas seguintes condições:

i) As ocorrências são independentes
ii) As ocorrências são aleatórias
iii) A variável aleatória *X* é o número de ocorrências de um evento **ao longo de algum intervalo** (de tempo ou espaço)

Distribuição de Poisson

Vamos associar a v.a. X o número de ocorrências em um intervalo.

X pode assumir os valores $0, 1, \ldots,$ representando o número de sucessos em um intervalo t.

$$P(X = k) = \frac{e^{-\lambda t}(\lambda t)^k}{k!}$$

$$P(X = k) = \frac{e^{-\mu}(\mu)^k}{k!}$$

$E(X) = \lambda$

$Var(X) = \lambda$

X	X Counts	p(x)	Values of X	E(x)	V(x)
Discrete uniform	Outcomes that are equally likely (finite)	$\frac{1}{b-a+1}$	$a \le x \le b$	$\frac{b+a}{2}$	$\frac{(b-a+2)(b-a)}{12}$
Binomial	Number of sucesses in n fixed trials	$\binom{n}{x} p^x (1-p)^{n-x}$	$x = 0,1,\ldots,n$	np	$np(1-p)$
Poisson	Number of arrivals in a fixed time period	$\frac{e^{-\lambda}\lambda^x}{x!}$	$x = 0,1,2,\ldots$	λ	λ
Geometric	Number of trials up through 1st success	$(1-p)^{x-1}p$	$x = 1,2,3,\ldots$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Negative Binomial	Number of trials up through kth success	$\binom{x-1}{k-1}(1-p)^{x-k}p^k$	$x = k,\ k+1,\ldots$	$\frac{k}{p}$	$\frac{k(1-p)}{p^2}$
Hyper - geometric	Number of marked individuals in sample taken without replacement	$\frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$	$\max(0, M+n-N) \le x \le \min(M,n)$	$n * \frac{M}{N}$	$\frac{nM(N-M)(N-n)}{N^2(N-1)}$

Distribuições contínuas

Existem diversos modelos coninúos de probabilidade. Alguns deles:

- Uniforme
- Exponencial
- Gama

Um dos modelos mais importantes, tanto do ponto de vista teórico quanto prático, é o **modelo normal**.

Conhecido por volta de 1733 pelo matemático francês Abraham De Moivre, e serve para explicar inúmeros fenômenos naturais, físicos, psicológicos, sociológicos, . . .

A distribuição normal é extremamente importante pois serve de fundamento para muitas técnicas de **inferência** e **aproximações**.

Distribuição Normal

Definição: Dizemos que uma v.a. X tem distribuição ormal se sua fdp é

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right], \quad -\infty < x < \infty$$

onde $\mu \in \mathbb{R}$ é a média da população, $\sigma \in \mathbb{R}^+$ é o desvio-padrão populacional.

Notação: $X \sim (\mu, \sigma^2)$

Esperançaa e variância: $(X) = \mu$ e $(X) = \sigma^2$

Distribuição normal

Para calcularmos a probabilidade da distribuição normal, deve-se calcular a área entre os pontos *a* e *b*, ou seja,

$$P[a < X < b] = \int_a^b \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] dx$$

No entanto, a função distribuição normal não possui forma fechada, portanto o cálculo de probabilidades não pode ser feito diretamente pela integral, apenas por aproximações numéricas.

Para contornar esse problema, os valores de probabilidade são obtidos para uma distribuição normal padrão (Z) com $\mu = 0$ e $\sigma^2 = 1$,

$$Z = \frac{X-\mu}{\sigma} \sim \mathrm{N}(0,1)$$

A vantagem é que podemos fazer uma única tabela com as integrais aproximadas de Z, ao invés de uma tabela para cada par (μ, σ^2).

Modelo normal

Se $Z \sim (0, 1)$, entãoo sua fdp é

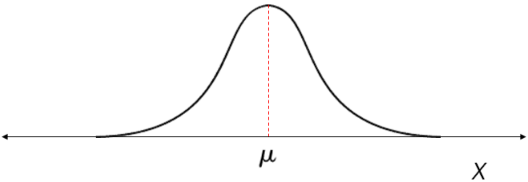
$$f(z) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}(z)^2\right]$$

Para se obter a probabilidade de Z estar entre a e b,

$$P[a < Z < b] = \int_a^b \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}(z)^2\right] dz$$

As integrais (áreas) para valores de Z entre 0,00 e 3,99 estão na tabela. Portanto, para qualquer valor de X entre a e b, podemos calcular a probabilidade correspondente através da transformação,

$$P[a < X < b] = P\left[\frac{a-\mu}{\sigma} < Z < \frac{b-\mu}{\sigma}\right]$$

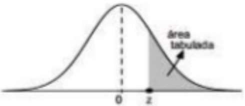


- **Suas média, mediana e moda são iguais.**

- **Tem forma de sino e é simétrica em torno da média.**

- **A área total sob a curva é de 100%.**

Distribuição normal padrão.



z	segunda decimal de z									
	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,4960	0,4920	0,4880	0,4840	0,4801	0,4761	0,4721	0,4681	0,4641
0,1	0,4602	0,4562	0,4522	0,4483	0,4443	0,4404	0,4364	0,4325	0,4286	0,4247
0,2	0,4207	0,4168	0,4129	0,4090	0,4052	0,4013	0,3974	0,3936	0,3897	0,3859
0,3	0,3821	0,3783	0,3745	0,3707	0,3669	0,3632	0,3594	0,3557	0,3520	0,3483
0,4	0,3446	0,3409	0,3372	0,3336	0,3300	0,3264	0,3228	0,3192	0,3156	0,3121
0,5	0,3085	0 ,3050	0 ,3015	0 ,2981	0 ,2946	0 ,2912	0 ,2877	0 ,2842	0 ,2810	0 ,2776
0,6	0,2743	0 ,2709	0 ,2676	0 ,2643	0 ,2611	0 ,2578	0 ,2546	0 ,2514	0 ,2483	0 ,2451
0,7	0,2420	0 ,2389	0 ,2358	0 ,2327	0 ,2296	0 ,2266	0 ,2236	0 ,2206	0 ,2177	0 ,2148
0,8	0,2119	0 ,2090	0 ,2061	0 ,2033	0 ,2005	0 ,1977	0 ,1949	0 ,1922	0 ,1894	0 ,1867
0,9	0,1841	0 ,1814	0 ,1788	0 ,1762	0 ,1736	0 ,1711	0 ,1685	0 ,1660	0 ,1635	0 ,1611
1,0	0,1587	0 ,1562	0 ,1539	0 ,1515	0 ,1492	0 ,1469	0 ,1446	0 ,1423	0 ,1401	0 ,1379
1,1	0,1357	0 ,1335	0 ,1314	0 ,1292	0 ,1271	0 ,1251	0 ,1230	0 ,1210	0 ,1190	0 ,1170
1,2	0 ,1151	0 ,1131	0 ,1112	0 ,1093	0 ,1075	0 ,1056	0 ,1038	0 ,1020	0 ,1003	0 ,0985
1,3	0 ,0968	0 ,0951	0 ,0934	0 ,0918	0 ,0901	0 ,0885	0 ,0869	0 ,0853	0 ,0838	0 ,0823
1,4	0 ,0808	0 ,0793	0 ,0778	0 ,0764	0 ,0749	0 ,0735	0 ,0722	0 ,0708	0 ,0694	0 ,0681

Obrigada pela atenção!