MPEIDO

Probabilidades

Aleatório

Uma experiência em que o **resultado não é previsível**. Não se pode prever nada más há probabilidades associadas com eventos regulares e tudo isso.

Probabilidade

Medida do grau de certeza associado a um resultado proveniente de um fenómeno de acaso.

A uma experiência aleatória são associados:

- Espaço de amostragem – conjunto de resultados possíveis
  - Representado por **S**
  - Resultados têm de ser mutuamente exclusivos e não divisíveis
  - **Discreto se for contável** (número finito de elementos, ou infinito em que se pode estabelecer uma correspondência biunívoca com o conjunto dos inteiros
  - **Continuo se não for contável**
- Conjunto de acontecimentos
  - **Subconjunto** de espaço de amostragem
- Lei de probabilidade
  - **Atribui probabilidade** aos vários eventos (0 a 1)

Como é que atribuímos probabilidades a eventos então?

Há diferentes teorias, temos a **teoria clássica** (de Laplace), **frequencista** e por fim, **matemática**.

Teoria clássica

Primeiro reduzir o fenómeno a um conjunto de resultados elementares, “casos”, igualmente prováveis**. P(evento) = ºcasos favoráveis / ºcasos possíveis.**

Por exemplo, com o lançamento de 1 dado, todos os eventos têm a mesma probabilidade. A probabilidade de obter 5 é dividir 1 (favoráveis) / 6 (possíveis)

Regras básicas

OU: P(sair face maior que 4) = P(sair face 5 ou 6) = P(5,6) = 2/6 = P(5) + P(6)

P(A U B) = P(A) + P(B) – P(A ∩ B) (se não forem mutuamente exclusivos)

Complemento: P(~A) = 1 – P(A)

E: P(face par e face menor ou igual a 4) = P(face par) x P(face menor ou igual a 4), ou seja ½ x 2/3 = 1/3

Se os acontecimentos forem independentes, P (A ∩ B) = P(A) x P(B)

Abordagem frequencista

Repete-se a experiência muitas vezes (N).
Seja k o número de vezes que ocorre o acontecimento que nos interessa (sair face 5 num dado).
F (**frequência relativa** de ocorrência) = k/n. Esta frequência é uma **medida empírica da probabilidade**.
Se a frequência relativa convergir para um certo valor quando N aumenta, então o **limite para +infinito da frequência relativa é a probabilidade do acontecimento**.
Numa experiência, a **frequência relativa de todos os resultados** em soma dá 1.
adsadsdas

Por exemplo a probabilidade de sair 2 caras em 3 lançamentos.
Em Matlab, faz-se uma coluna de 3 números random de 0 a 1. Se sair menos que 0.5 digamos que é cara. Nesse caso, o elemento da coluna fica com o valor 1, senão, 0.
No fim, soma-se a coluna. Nesta abordagem frequencista, é necessário realizar a experiência muitas vezes. Então faz-se mais colunas no início (N). Somam-se os valores booleanos de cada coluna e ficamos com um vetor de 1 linha e N colunas. Os elementos deste vetor podem ter valores entre 0 a 3 (ou seja, podem sair 0 caras até 3). Nós queremos a probabilidade de saírem exatamente 2 caras. Então se cada elemento desse vetor for 2, substituímos o seu valor por o valor 1, senão, 0. Por fim, faz-se a soma desse novo vetor e divide-se o resultado por N. **Ou seja, fizeram-se N experiências e em cada experiência, se o valor obtido foi o desejado, essa experiência conta como 1. Faz-se a soma de todos estes valores e divide-se pelo número de experiências. Quanto maior N, o resultado tende a ser mais fiável**.

```
% contar num ocorrências de “2 caras”

% contar num caras (1s) em cada experiência

% (que se encontra numa coluna da matriz lancamentos)

numCarasNaExperiencia= sum (lancamentos);

% contar vezes em que esse número de caras é % 2

numOcorrencias = sum (numCarasNaExperiencia ==2)

% calcular freq relativa

fr = numOcorrencias / N

% usar como estimativa da probabilidade

 pA= fr
```

Esta abordagem tem algumas desvantagens:

- Gasta bastantes recursos na simulação
- Experiências têm de poder ser repetidas em condições idênticas
- Se espaço amostral for infinito temos problemas.
- Quantos ensaios se tem de efetuar para termos medidas fiáveis?

Teoria Axiomática de probabilidade

Axiomas são leis boring que sabemos por dedução e.g S tem prob 1.
P(A) = 1 – P(~A)
P(A U B) = P(A) + P(B) – P (AB) (se não forem mutuamente exclusivos)
…

Probabilidades Condicionais

A ocorrência de um acontecimento depende da ocorrência do outro.

P(A | B) = P(AB) / P(B) (indefinida se P(B) = 0)
Regra de Bayes: P(causa | efeito) = P(efeito | causa) * P(causa) / P(efeito)
$P(A \cap B) = P(B) \times P(A \mid B) = P\ A \times P(B|A)$

Independência

2 acontecimentos são independentes se P(AB) = P(A)P(B), ou P(A) = 0 e implica P(A|B) = P(A)
Em geral, dois acontecimentos que sejam mutuamente exclusivos e tenham probabilidade não nula não podem ser independentes
A experiência de Bernoulli consiste em realizar uma experiência e registar de um dado acontecimento se verifica ou não.
p = probabilidade de sucesso
n = ensaios

- k sucessos em n experiências podem ocorrer de C_k^n maneiras

- Então a probabilidade pedida é:

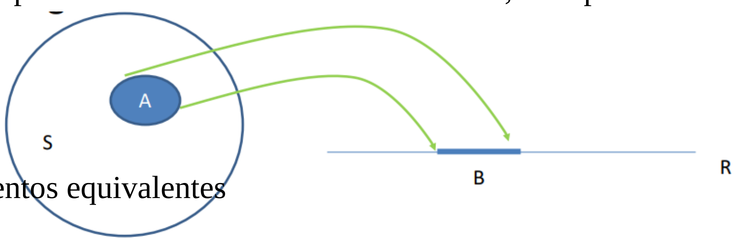
$$P_n(k) = C_k^n \; p^k \; (1-p)^{n-k}$$

Lei Binomial

Variáveis aleatórias

Uma função que mapeia o espaço de amostragem na reta real é designada de variável aleatória. É o resultado numérico das nossas experiências aleatórias
Cada acontecimento A tem uma certa probabilidade X.

Caso continuo
Se os conjuntos que representam os eventos forem contínuos, o mapeamento é para um segmento da reta real



A e B são acontecimentos equivalentes

Tipos de variáveis aleatórias

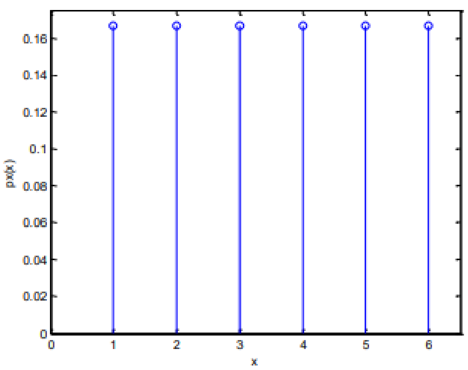
- **Discreta:** se os valores que a variável aleatória pode assumir forem finitos ou infinitos, mas contáveis.
- **Continua:** se os valores que pode assumir formarem um ou vários intervalos disjuntos
- **Mista:** onde se verificam os atributos que definem os 2 tipos anteriores

As variáveis aleatórias são caracterizáveis pelo conjunto de valores que podem assumir e as probabilidades associadas, ou seja pela **distribuição de probabilidades**

Por exemplo, uma variável aleatória discreta X é especificada por:

1. **Conjunto de valores que pode assumir**: $x_i = 1,2\ldots$
2. **Probabilidade associada a cada um desses valores**: $p_X(x_i)$ – **Função massa** de probabilidade/ função de probabilidade

Implica que o somatório de todas as probabilidades para cada acontecimento seja 1.
Por exemplo, o lançamento de um dado:
$x_i = \{1,2,3,4,5,6\}$ (**espaço de amostragem**)
$p_X(x_i) = 1/6$ (**probabilidade para cada acontecimento**)

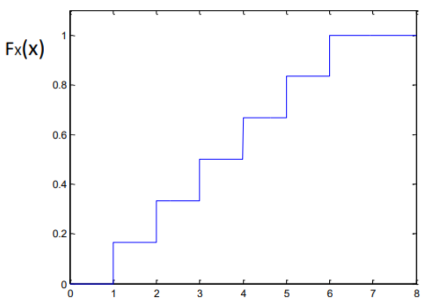


Em Matlab:
xi = 1:6; (array de 1 a 6[espaço de amostragem])

p = ones (1,6) /6; (array de 6 elementos de valor 1/6 [probabilidade de cada acontecimento no espaço de amostragem])

stem (xi, p), xlabel('x'), ylabel('px(x)');

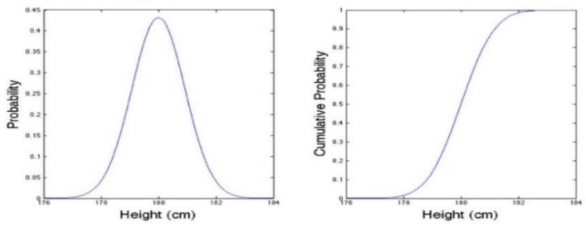
Uma variável aleatória discreta pode ser também especificada pela sua **função distribuída acumulada.** É uma função cujo limite para +infinito é 1 e -infinito é 0, ou seja o somatório da função dá 1. Por exemplo no gráfico de cima, fica assim:



Em x = 1, temos a probabilidade de 1, em x = 2, temos a probabilidade de 2 + a probabilidade acumulada de trás, and so on. É uma **função em escada**.

Variáveis aleatórias continuas

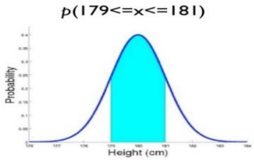
- Exemplo de par de funções de densidade e de distribuição



Skip nesta parte 😴

Probabilidades e função de densidade

- $P(a < X \leq b) = \int_a^b f_X(x)dx$



- A probabilidade é a área debaixo da curva
- Área total da curva =1

Caracterização das variáveis aleatórias

Consideremos N lançamentos de um dado
Média = (numDe1s x 1 + numDe2s x 2 …) /N = numDe1s/N x 1 + NumDe2s/N x 2 + …
Valor esperado: Valor médio de X ao repetirmos as experiências indefinidamente:

$$\lim_{n\to\infty} \frac{X_1 + X_2 + \cdots + X_n}{n}$$

Só existe valor esperado se os resultados forem finitos.
É um **termo enganador** e na realidade não é algo que devemos esperar que ocorra.
Por exemplo, o **valor médio do lançamento do dado**: (1 + 2 + 3 + 4 + 5 + 6) / 6 = **3,5**.
(Somatório espaço amostragem) / len (espaço amostragem).
É impossível sair 3,5.
O valor esperado de uma variável designa-se por **E[X].**

E[aX] = a * E[X]
E[X+Y] = E[X] + E[Y]
E[X+c] = E[X] + c

- No caso discreto: $E[X] = \sum_i x_i p(x_i)$
- No caso contínuo: $E[X] = \int_{-\infty}^{+\infty} x\, f_X(x)dx$

Exemplo de cálculo de E[X]

x_i	$p_X(x_i)$	$x_i\, p_X(x_i)$
-1	.1	-.1
0	.2	.0
1	.4	.4
2	.2	.4
3	.1	.3
		1.0

E[X]= 1.0

A média pode não ser suficiente. Podemos ter a mesma média em turmas muito diferentes. Pode haver outliers a tirarem 20's e 2's e noutra turma toda a gente ter 11's

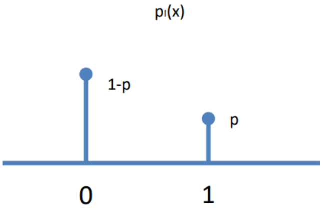
Essa dispersão é dada pela **Variância**
Ideia base: Usar a diferença dos valores da variável para a media (valor esperado) e fazer a sua média.
Para evitar o cancelamento de diferenças negativas e positivas, em vez de usar diretamente o valor da diferença, utilizar o seu valor quadrático.

Então, Var(X) = E [(X - E(X))^2]
Var(X) = E[X^2] – E^2[X]
A **raiz quadrada da variância** é o **desvio padrão**
var(X+c) = var(X)
var (c X) = C^2 var(X)
Exemplo disto tudo:

$S_I = \{0,1\}$
p= Pr(A)

$p_I(1)$ =p
$p_I(0)$ =1-p

Valor esperado E[I] ?
Var(I) = ?



$$E[I] = \sum_i x_i p(x_i)$$
$$= 0 \times (1-p) + 1 \times p$$
$$= \boldsymbol{p}$$
$$Var(I) = E[I^2] - (E[I])^2$$

$$E[I^2] = 0^2 \times (1-p) + 1^2 \times p = \text{p}$$

$$Var(I) = p - p^2 = \boldsymbol{p\,(1-p)}$$

E[X] pode ser interpretado como valor medio de X, centro de gravidade da função massa de probabilidade (caso discreto) ou função de densidade de probabilidade.

Desvio padrão / variância dá uma **medida da dispersão da variável aleatória**. **Pequenos valores indicam var aleatória muito concentrada em torno da média**. Se for **zero**, não temos var aleatória e **todos os valores são iguais à média**.

Variáveis aleatórias: Distribuições

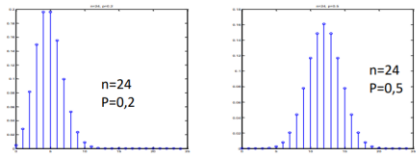
Distribuições discretas

Distribuição de Bernoulli

Diretamente relacionada com as experiências de Bernoulli. → 1 se acontecer o que queremos, senão 0.

Distribuição binomial

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



$$F_X(x) = \sum_{k=0}^{\lfloor x \rfloor} \binom{n}{k} p^k (1-p)^{n-k}$$

According to the U.S. Census Bureau, approximately 6% of all workers in Jackson, Mississippi, are unemployed.

In conducting a random telephone survey in Jackson, what is the probability of getting two or fewer unemployed workers in a sample of 20?

6% desempregado => $p = 0{,}06$

Tamanho da amostra é 20 => $n = 20$

94% têm emprego => $1 - p = 0{,}94$

x é o número de sucessos que se pretende

Resolução

$$n = 20$$
$$p = 0{,}06$$
$$q = 1 - p = 0{,}94$$
$$P(X \le 2) = P(X=0) + P(X=1) + P(X=2)$$
$$= 0{,}2901 + 0{,}3703 + 0.2246 = 0{,}8850$$

$$P(X=0) = \frac{20!}{0!(20-0)!}(0{,}06)^0(0{,}94)^{20-0} = (1)(1)(0{,}2901) = 0{,}2901$$

$P(X=1) = \ldots$

$P(X=2) = \ldots.$

Distribuição Geométrica

- Seja X o número de vezes que é necessário repetir uma experiência de Bernoulli até obter um sucesso
  - Prob. Sucesso: p   prob. Falha = 1-p
- $p_X(k) = p(1-p)^{k-1}, k = 1,2,3, \ldots$

  Porque teremos k-1 insucessos e depois sucesso

Considere o serviço de atendimento via telefone do Helpdesk da UA.

Supondo que a probabilidade de se conseguir contactar o suporte é p=0,1 (só ao fim de 10 tentativas ☹).

Determine a probabilidade de necessitar de menos de 3 chamadas até conseguir expor o seu problema ?

Solução:

$\Pr(nº\ chamadas < 3) =$

$\Pr(1\ chamada\ OU\ 2\ chamadas)$

→ $= p(1-p)^{1-1} + p(1-p)^{2-1} = p(2-p) = 0{,}19$

E[X] = 1/p

Var(X) = (1 – p) / p^2

Distribuição de Poisson

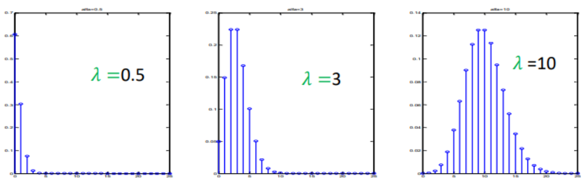
Considere-se que temos uma variavel binomial, n (amostras) cresce e p decresce por forma a np = lambda

Distribuição para vários valores do parâmetro λ

Função de probabilidade:

$$p_X(k) = \Pr(X = k) = \frac{\lambda^k}{k!}e^{-\lambda}, k = 0,1,2,3\ \ldots$$

Tem apenas um parâmetro, o lambda



E[X] = lambda

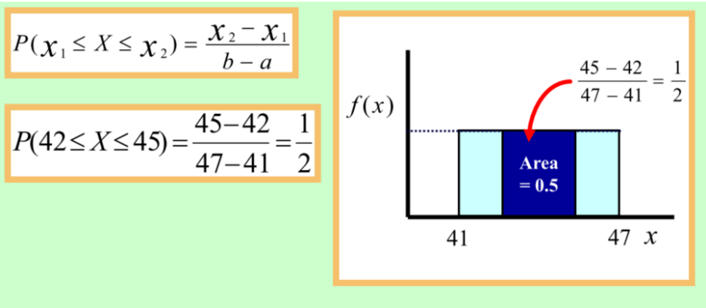
Var(X) = lambda

A distribuição de Poisson foca-se apenas no numero de ocorrencias (discreto) num intervalo de tempo contínuo.
Problemas envolvendo a distribuiçao binomial em que n é grande e p é pequeno, geram eventos raros. Entao estes problemas são candidatos à utilização da distribuição de Poisson.

Rule of thumb: Se n > 20 e np <= 7, usa-se poisson em vez da binomial.

Distribuiçoes continuas

- P(42<=X<=45)   com U(41,47)



Skip nisto

Os slides agora é so letras eu não vou ler isto