PROBABILIDADES E DISTRIBUIÇÕES DE PROBABILIDADES

- Probabilidades

$$P(\text{resultado favorável}) = \frac{\text{número de resultados favoráveis}}{\text{nº de resultados igualmente possíveis}}$$

A probabilidade é uma medida da incerteza dos fenômenos.

Traduz-se por um número real compreendido de 0 (zero) e 1 (um).

Probabilidade a priori ou matemática: cálculo a partir de modelo estatístico e sem experimentação, determinando as probabilidades de acontecimentos futuros;

Probabilidade a posteriori: estimativa por meio de dados experimentais, da verdadeira probabilidade ou valor mais provável.

Espaço amostral (S): todos os possíveis resultados do experimento.

Evento: qualquer subconjunto do espaço amostral.

A probabilidade da união de dois eventos:

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

Evento Mutuamente Exclusivos (Disjuntos): $P(A \cap B) = 0$

Eventos independentes: são aqueles que não exercem ação entre si, isto é, cada evento comportando-se da maneira que lhe é própria.

condição necessária e suficiente para que dois eventos sejam independentes:

$$P(A \cap B) = P(A)\ x\ P(B)$$

Probabilidade Condicionada: Se A e B são dois eventos, a probabilidade de B ocorrer, depois de A ter acontecido, é definida por P(B/A) e é denominada probabilidade condicional de ocorrer B, sabendo da ocorrência de A.

Se a ocorrência ou não de A não afetar a probabilidade da ocorrência de B, então P(B/A) = P(B). Neste caso, A e B são independentes.

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

Exemplo 1. Imagens de duas épocas distintas foram classificadas em 3 classes: floresta (*F*), capoeira (*C*) e área agrícola (*A*). A fim de comparar as mudanças entre as épocas, fez-se a tabulação cruzada entre as imagens classificadas.

i) Os eventos A2 e F1 são independentes?

ii) Selecionado um ponto aleatoriamente:

a) ser floresta na época 1;
b) ser floresta em ambas as épocas;
c) ser capoeira em qualquer época;
d) não ter mudado de classe entre as épocas analisadas;
e) ser capoeira na época 2, tendo sido área agrícola na época 1; e
f) ser capoeira na época 2, não tendo sido área agrícola na época 1.

		Época 1		
		Floresta (F1)	Capoeira (C1)	Área Agrícola (A1)
Época 2	Floresta (F2)	100	0	0
	Capoeira(C2)	0	150	50
	Área Agrícola (A2)	20	30	100

Distribuições de Probabilidades

Funções estatísticas que permitem o cálculo de probabilidade sem a necessidade de realização do experimento.

A distribuição adequada para representar a variável é obtida em função de características dessa variável.

Distribuições discretas: Uniforme, Bernoulli, Binomial, Poisson, Multinomial, Geométrica, ...

Distribuições Contínuas: Uniforme, Exponêncial, Normal, Gamma, ...

DISTRIBUIÇÃO BINOMIAL

É apropriada nas experiência onde ocorre somente duas situações (sucesso ou fracasso)

Exemplos: lançamento de moedas; fecundação de óvulos; defeito de equipamentos; teste de doença; etc..

$$P(X = x) = C_n^x p^x q^{n-x}$$

n é o número de experiências; p é a probabilidade do sucesso;
q = 1 - p é a probabilidade do fracasso

$$C_n^x = \frac{n!}{(n-x)!\,x!}$$

Um experimento se enquadra como um experimento binomial se as seguintes condições são satisfeitas:

- A variável é discreta
- Em cada experimento ocorre apenas o sucesso (p) ou fracasso (q)
- Os experimentos repetidos são independentes
- A probabilidade do sucesso (p) permanece constante de experimento para experimento
- Um número fixo de n experiências são realizadas

Exemplo 1: Estudos indicam que 70% das empresas de certo setor da economia atendem as normas ambientais. Se 10 empresas desse setor, qual a probabilidade de:

a) Todos atenderem as normas?

b) No máximo 6 atenderem as normas?

DISTRIBUIÇÃO NORMAL

Mais importante dos modelos para variáveis contínuas.

Aplicação em todos as áreas da ciência.

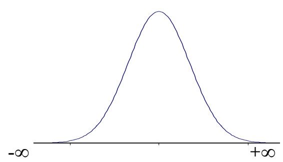
Uma variável X tem distribuição normal se sua função densidade de probabilidade é dada por:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}.e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

μ = média da variável X ($\mu \in \Re$)
σ = desvio padrão de X ($\sigma \in \Re$+)
e = 2,718
π= 3,141516
$x \in \Re$

O gráfico da normal tem a forma de sino

A distribuição é simétrica em relação a média (μ)



Cálculo de probabilidade

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

A distribuição normal padronizada

Utiliza-se das propriedades da média e do desvio padrão, para definir a variável normal padronizada Z.

Livros apresentam as tabelas de Z

$$Z = \frac{x-\mu}{\sigma}$$

Tabela 1. Área sob a curva normal padronizada compreendida entre os valores 0 e Z

Z	0	1	2	3	4	5	6	7	8	9
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990
3.1	0.4990	0.4991	0.4991	0.4991	0.4992	0.4992	0.4992	0.4992	0.4993	0.4993
3.2	0.4993	0.4993	0.4994	0.4994	0.4994	0.4994	0.4994	0.4995	0.4995	0.4995
3.3	0.4995	0.4995	0.4995	0.4996	0.4996	0.4996	0.4996	0.4996	0.4996	0.4997
3.4	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4998
3.5	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998
3.6	0.4998	0.4998	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.7	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.8	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.9	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000

Exercício de aplicação:

1) Sabe-se que a variável X que representa o tempo gasto para se realizar uma determinada análise de laboratório, tem distribuição normal com média de 40 minutos e desvio padrão de 10 minutos. Uma análise será realizada:

a) Qual a probabilidade que o tempo para a realização da análise fique entre 30 e 50 minutos?

b) Qual a probabilidade que o tempo seja superior a 60 minutos?

c) Qual a probabilidade do tempo estar entre 40 e 50 minutos.

d) Acima de que valor de tempo espera-se realizar 90% das análises?

2) Suponha que a emissão de CO_2 por veículo no Brasil tenha distribuição normal com média de 2 ton e desvio padrão de 0,6 ton.

a) Ao selecionar um veículo aleatoriamente qual a probabilidade que a emissão: i) fique entre 1,4 e 2,6 ton? ii) seja superior a 2,5 ton? iii) seja inferior a 1 ton? iv) fique entre 2,2 e 2,6 ton?

b) Se a emissão for classificado em Baixa, Média e Alta, de acordo com o seguinte critério: Baixa → 10% das menores emissões; Alta→ 15% das maiores e Média→ os demais 75%, quais serão os limites estimados para a classificação?

c) Se selecionarmos 5 veículos, qual a probabilidade de que em 3 tenhamos emissão superior a 2,9 ton?