

# Estatística e Probabilidade

#### Bacharelado em Sistemas de Informação

Aula 7: Distribuições de Probabilidade \_\_\_\_\_
Prof. Dr. Samuel Sanches



#### **VARIÁVEIS ALEATÓRIAS**

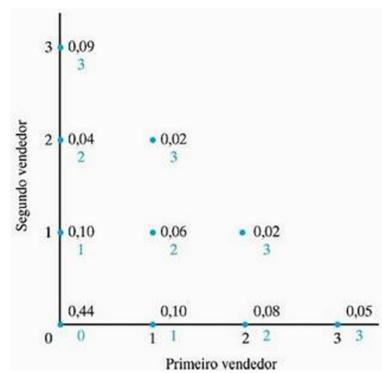
★ Resultados dos experimentos => variáveis aleatórias

★ (não são variáveis e nem aleatórias, como o pé de cabra, não é nem pé e nem cabra)



#### **VARIÁVEIS ALEATÓRIAS**

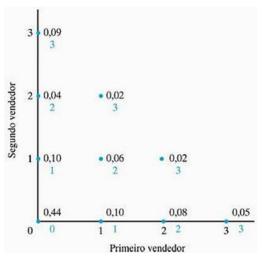
- ★ Exemplo da venda dos caminhões, vamos analisar quantos os dois vendedores vão vender juntos na semana, essa será nossa variável aleatória (azul no gráfico), veja que são valores finitos (não temos infinitas possibilidades), ou seja, **discretos**.
- ★ Na matemática quando associamos números com pontos, temos uma função, então a nossa variável aleatória está mais para uma função no nosso problema, aqui entram as contínuas.





★ Probabilidade => variável aleatória => valor particular => somamos as probabilidades do espaço amostral em que aquele valor é assumido.

Número de caminhões vendidos	Probabilidade	
0	0,44	
1	0,20	
2	0,18	
3	0,18	



★ A tabela ilustra a distribuição, onde assumimos probabilidades para os valores particulares assumidos.



- ★ Para um dado equilibrado, possibilidades são 1, 2, 3, 4, 5, 6, a probabilidade é 1/6:
- ★ Costume escrever como fórmula: f(x) = 1/6para x = 1, 2, 3, 4, 5 e 6

onde f(1) => probabilidade de obter 1

Número de pontos na jogada de um dado	Probabilidade		
1	16		
2	16		
3	16		
4	16		
5	16		
6	16		

★ Por serem probabilidades, estarão no intervalo de 0 a 1 e a soma de todos os valores de uma distribuição de probabilidade deve ser igual a 1.



- ★ Lançamento 4x de uma moeda (16 resultados): KKKK, KKKC, KKCK, KCKK, CKKK, KKKK, KKCC, KCKC, KCKK, CKKC, CKCK, CCKK, KCCC, CKCC, CCKC, CCCK e CCCC.
- ★ Distribuição para o resultado que temos cara:

Número de formas	Probabilidade
0	1/16
1	4/16
2	6 16
3	4 16
4	1/16



★ **Exemplo**: Verifique que a distribuição de probabilidade do número de caras obtidas em quatro lançamentos de uma moeda equilibrada é dada por:

$$f(x) = \frac{\binom{4}{x}}{16}$$
 para  $x = 0, 1, 2, 3 e 4$ 

Substituindo: 
$$\binom{4}{0} = 1$$
,  $\binom{4}{1} = 4$ ,  $\binom{4}{2} = 6$ ,  $\binom{4}{3} = 4$  e  $\binom{4}{4} = 1$ 

Obtendo: f(0) = 1/16, f(1) = 4/16, f(2) = 6/16, f(3) = 4/16 e f(4) = 1/16



★ Exemplo: Verifique se a correspondência dada pela expressão pode ser a distribuição de probabilidade de alguma variável aleatória.

Substituindo os valores de x em f(x):

$$f(1) = \frac{4}{15}$$
,  $f(2) = \frac{5}{15}$  e  $f(3) = \frac{6}{15}$ 

Nenhum valor negativo ou > 1 e a soma:

$$\frac{4}{15} + \frac{5}{15} + \frac{6}{15} = 1$$

Pode ser a distribuição de alguma variável aleatória!



- $\star$  x sucessos e **n** x fracassos em **n** tentativas.
- ★ Número fixo de provas.
- ★ A probabilidade de sucesso é a mesma em cada prova.
- ★ As provas são todas independentes.
- ★ A probabilidade de se obter **x** sucessos em **n** provas independentes é:

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}$$
 para  $x = 0, 1, 2, ..., ou n$ 

onde **p** é a probabilidade constante de sucesso em cada prova.



**★ Exemplo**: A expressão para o número de caras foi apresentada, utilizando a distribuição binomial, chegue na expressão, onde n = 4 e p = 1/2.

$$f(x) = {4 \choose x} \left(\frac{1}{2}\right)^x \left(1 - \frac{1}{2}\right)^{4 - x} = {4 \choose x} \left(\frac{1}{2}\right)^4 = \frac{{4 \choose x}}{16}$$

para x = 0, 1, 2, 3 e 4.



**★ Exemplo**: Se a probabilidade de um eleitor qualquer (escolhido aleatoriamente) votar na eleição é 0,7, qual é a probabilidade de 2 entre 5 votarem na eleição?

$$x = 2, n = 5, p = 0.70 \text{ e} {5 \choose 2} = 10$$

$$f(2) = {5 \choose 2} (0,70)^2 (1 - 0,70)^{5-2} = 10(0,70)^2 (0,30)^3 = 0,132$$



★ Exemplo: A probabilidade de uma pessoa fazendo compras em um mercado que aproveite uma promoção é de 0,3. Determine a probabilidade de que dentre 6 pessoas, haja 0, 1, 2, 3, 4, 5 e 6 aproveitando a promoção. Esboce um histograma da distribuição.

n = 6 e p = 0,3, com x = 0, 1, 2, 3, 4, 5, 6  

$$f(0) = {6 \choose 0} (0,30)^{0} (0,70)^{6} = 0,118$$

$$f(1) = {6 \choose 1} (0,30)^{1} (0,70)^{5} = 0,303$$

$$f(2) = {6 \choose 2} (0,30)^{2} (0,70)^{4} = 0,324$$

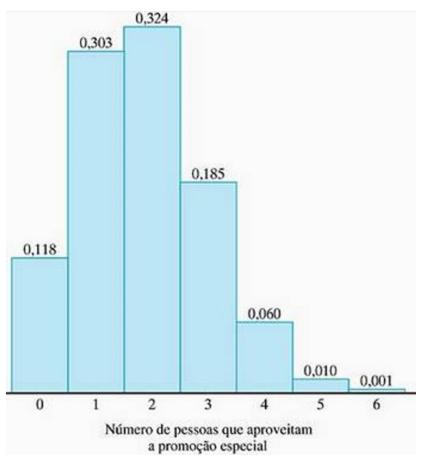
$$f(3) = {6 \choose 3} (0,30)^{3} (0,70)^{3} = 0,185$$

$$f(4) = {6 \choose 4} (0,30)^4 (0,70)^2 = 0,060$$

$$f(5) = {6 \choose 5} (0,30)^5 (0,70)^1 = 0,010$$

$$f(6) = {6 \choose 6} (0,30)^6 (0,70)^0 = 0,001$$







★ Vale para qualquer cenário que teríamos p = 0,3 e escolhas de 0 a 6:

Relógio tem p = 0,3 de durar dois anos, então qual a probabilidade dentre 6 (0, 1, 2, 3, 4, 5, 6) durem 2 anos.

★ Pessoa tem p = 0,3 de sobreviver com uma doença por mais de 10 anos, então qual a probabilidade dentre 6 (0, 1, 2, 3, 4, 5, 6) vivam mais que 10 anos.



**★ Exemplo**: A probabilidade de um eclipse lunar ser ocultado por nuvens num observatório é de 0,6. Encontre as probabilidades de serem ocultados: **a)** no máximo 3 dentre 10 eclipses; **b)** no mínimo 7 dentre 10.

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

a) n = 10 e p = 0,6 com x = 0, depois 1, 2 e 3:  

$$0.000 + 0.002 + 0.011 + 0.042 = 0.055$$

**b)** n = 10 e p = 0,6 com x = 7, depois 8, 9 e 10: 
$$0.215 + 0.121 + 0.040 + 0.006 = 0.382$$



**★ Exemplo**: A **probabilidade** de um eclipse lunar ser ocultado por nuvens num observatório **é de 0,63**. Encontre as probabilidades de serem ocultados: **a)** no máximo 3 dentre 10 eclipses; **b)** no mínimo 7 dentre 10.

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

a) n = 10 e p = 0,63 com x = 0, depois 1, 2 e 3:  

$$0.0000 + 0.0008 + 0.0063 + 0.0285 = 0.0356$$

**b)** n = 10 e p = 0,63 com x = 7, depois 8, 9 e 10: 
$$0.2394 + 0.1529 + 0.0578 + 0.0098 = 0.4599$$



★ Ao observar um valor da variável que possui distribuição binomial, como em quantas caras em 25 lançamentos de uma moeda ou número de acidentes (dentre 20 investigados) causados por embriaguez, dizemos que estamos extraindo uma **amostra de uma população binomial**.



- ★ A distribuição binomial vale quando temos amostras com reposição e as provas são todas independentes, como ao jogar uma moeda.
- ★ Exemplo (sem reposição): uma floricultura envia limoeiros de 3 anos em lotes de 24 e, ao chegar, seleciona ao acaso 3 de cada lote. Se tudo ok, é aceito, se não, as 21 restantes são inspecionadas. Podemos escolher 3 ok e o resto ruim, temos um risco considerável. Veja que ao escolher 1, ela é retirada do todo!



- ★ Escolher n objetos de um conjunto em que a objetos são do tipo sucesso e b objetos são do tipo fracasso, a amostragem é sem reposição, queremos a probabilidade de obter x sucessos e n x fracassos.
- ★ Temos combinação de **a** + **b** com **n** de escolher.

$$\binom{a+b}{n}$$

Temos **x** com **a** combinações de sucessos e **b** com **n** - **x** combinações de fracassos.  $\binom{a}{x} \cdot \binom{b}{n-x}$ 



★ A probabilidade de x sucessos em n provas é:

$$f(x) = \frac{\binom{a}{x} \cdot \binom{b}{n-x}}{\binom{a+b}{n}} \quad para \ x = 0, 1, 2, \dots, ou \ n$$

x não pode ser maior que a e n - x não pode ser maior que b.



★ **Exemplo**: Um funcionário deveria remeter 6 de 15 pacotes, mas mistura aleatoriamente e manda 6. Qual é a probabilidade de que apenas três dos pacotes que deveriam ir, realmente foram?

a = 6 (tipo sucesso), b = 9 (tipo fracasso), n = 6 (tentativas) e = 3 (sucessos obtidos)

$$f(3) = \frac{\binom{6}{3} \cdot \binom{9}{6-3}}{\binom{15}{6}} = \frac{20 \cdot 84}{5.005} \approx 0,336$$



★ **Exemplo**: De 16 ambulâncias, 5 emitem muitos poluentes. Se 8 são escolhidas ao acaso, qual é a probabilidade de que na amostra inclua

pelo menos 3 que emitem muitos poluentes?

pelo menos 3, pode ser 3, 4 ou 5! f(3) + f(4) + f(5), com a = 5, b = 11 e n = 8:

$$f(3) = \frac{\binom{5}{3} \cdot \binom{11}{5}}{\binom{16}{8}} = \frac{10 \cdot 462}{12.870} = 0,359$$

$$f(4) = \frac{\binom{5}{4} \cdot \binom{11}{4}}{\binom{16}{8}} = \frac{5 \cdot 330}{12.870} = 0,128$$

$$f(5) = \frac{\binom{5}{5} \cdot \binom{11}{3}}{\binom{16}{8}} = \frac{1 \cdot 165}{12.870} \approx 0,013$$



★ Muitas vezes podemos utilizar a binomial no lugar da hipergeométrica, por ser mais simples, binomial 2 parâmetros (n e p), já a hipergeométrica 3 parâmetros (a, b e n).

★ Se n não exceder 5% de a + b, podemos utilizar a binomial sem problemas:

$$n \le (0,05)(a+b)$$



★ **Exemplo**: 120 dos 300 internos, têm pena por crimes com drogas. Se 8 forem escolhidos ao acaso, qual é a probabilidade de que 3 dentre 8 estejam cumprindo pena por crimes com drogas?

n = 8, a + b = 300, (0,05)\*(300) = 15, logo podemos utilizar a binomial:

n = 8 e p = 120/300 = 0,40, com x = 3, utilizando a expressão da binomial, encontramos probabilidade de 0,279

se utilizar a hipergeométrica encontramos que o erro é de apenas 0,003



★ Quando n é grande e p é pequena, as distribuições binomiais são aproximadas para:

$$f(x) = \frac{(np)^x \cdot e^{-np}}{x!}$$
 para  $x = 0, 1, 2, 3, ...$ 

- ★ Número de Euler (irracional) e = 2,71828..., base dos logaritmos naturais (ln).
- **★** Quão grande: **n** ≥ **100** e **n\*p** < **10**
- ★ Alguns casos, a hipergeométrica pode ser aproximada para binomial que pode ser aproximada para a Poisson



★ **Exemplo**: 2% dos livros de uma gráfica apresentam defeitos. Com a distribuição de Poisson, encontre a probabilidade de que 5 apresentem defeito em um lote de 400.

$$n = 400 \ge 100 \text{ e } n*p = 400*(0,02) = 8 < 10, \text{ e } x = 5$$
:

$$f(5) = \frac{8^5 \cdot e^{-8}}{5!} = \frac{(32.768)(0,00033546)}{120} \approx 0,0916$$



**★ Exemplo**: Probabilidade de 0,00006 de um pneu furar em um túnel. Com a distribuição de Poisson, encontre a probabilidade de que ao menos 2 dentre 10.000 carros tenham o pneu furado nesse túnel.

n =  $10.000 \ge 100$  e n\*p = 10.000\*(0,00006) = 0,6 < 10, aqui teríamos x = 2, 3, 4,..., 10.000, melhor subtrair 1 do x = 0 e x = 1:

$$f(0) = \frac{(0,6)^0 \cdot e^{-0,6}}{0!} = \frac{1(0,5488)}{1} = 0,5488$$
  
$$f(1) = \frac{(0,6)^1 \cdot e^{-0,6}}{1!} \approx \frac{(0,6)(0,5488)}{1} = 0,3293$$

Então: 1 - (0,5488 + 0,3293) = 0,1219



★ **Exemplo**: De 4.000 faturas, 28 possuem erros. Escolhendo 150 aleatoriamente das 4.000, qual a probabilidade de encontrar 2 com erros nessas 150.

Aqui x = 2, a = 28, b = 4000 - 28 = 3972 e n = 150, mas  $150 \le 0.05*4000 = 200$ , podemos utilizar a binomial com n = 150 e p = 28/4000 = 0.007, porém  $n \ge 100$  e n\*p = 150\*(0.007) = 1.05 < 10 podemos utilizar a Poisson com x = 2:

$$f(2) = \frac{1,05^2 \cdot e^{-1.05}}{2!} = \frac{1,1025(0,349938)}{2} = 0,1929$$



★ Muitos casos não tem nenhuma ligação com a binomial, assim **np** passa a ser o parâmetro λ (letra grega minúscula *lambda*), então a probabilidade de se obter x sucessos é:

$$f(x) = \frac{\lambda^x \cdot e^{-\lambda}}{x!}$$
 para  $x = 0, 1, 2, 3, ...$ 

- $\star$  O  $\lambda$  é interpretado como o número esperado, ou médio, de sucessos.
- ★ Situações onde se espera número fixo de "sucessos" por alguma unidade, como quanto são esperados 1,6 acidentes por dia em um cruzamento ou são esperados 8 azeitonas em uma pizza média portuguesa, etc...



**Exemplo**: Dado que um banco recebe em média  $\lambda = 6$  cheques sem cobertura por dia, qual é a probabilidade de receber 4 em um dia?

Aqui x = 4, e  $\lambda = 6$ :

$$f(4) = \frac{6^4 \cdot e^{-6}}{4!} = \frac{1.296(0,002479)}{24} = 0,1339$$



**Exemplo**: Se podemos antecipar  $\lambda = 5.6$  imperfeições por peça de um material, qual é a probabilidade de uma peça conter x = 3 imperfeições?

Aqui x = 3, e  $\lambda$  = 5,6:

$$f(3) = \frac{5.6^3 \cdot e^{-5.6}}{3!} = \frac{175,616(0,003698)}{6} \approx 0,1082$$



- ★ A esperança matemática deve ser interpretada como uma média.
- ★ Valor esperado de uma variável aleatória interpretada como sua média.

$$\mu = \sum x \cdot f(x)$$

- ★ µ letra grega minúscula (chamada mu)
- $\star$  Média de uma distribuição binomial:  $\mu = n \cdot p$
- $\star$  Para **distribuição de Poisson**, temos o parâmetro  $\lambda = \mu$  que deve ser interpretado como uma média.



★ **Exemplo**: Encontre a média da distribuição que se refere ao número de pontos obtidos na jogada de um dado equilibrado.

$$\mu = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = 3\frac{1}{2}$$

Número de pontos na jogada de um dado	Probabilidade
1	16
2	16
3	16
4	16
5	16
6	16

notação da fração indica 3 + 1/2 = 3,5



**Exemplo**: Encontre o número médio de pessoas, dentre 6 que estejam fazendo compras, aproveitem a promoção.

$$\mu = 0(0,118) + 1(0,303) + 2(0,324) + 3(0,185) + 4(0,060) + 5(0,010) + 6(0,001)$$
$$= 1.802$$

$$f(0) = {6 \choose 0} (0,30)^0 (0,70)^6 = 0,118$$

$$f(1) = {6 \choose 1} (0.30)^1 (0.70)^5 = 0.303$$

$$f(2) = {6 \choose 2} (0,30)^2 (0,70)^4 = 0,324$$

$$f(3) = {6 \choose 3} (0,30)^3 (0,70)^3 = 0,185$$

$$f(4) = {6 \choose 4} (0.30)^4 (0.70)^2 = 0.060$$

$$f(5) = {6 \choose 5} (0,30)^5 (0,70)^1 = 0,010$$

$$f(6) = {6 \choose 6} (0,30)^6 (0,70)^0 = 0,001$$



★ **Exemplo**: Encontre o número médio de pessoas, dentre 6 que estão fazendo compras no mercado e aproveitam a promoção, utilizando a expressão da média para distribuição binomial.

nesse problema tínhamos n = 6 e p = 0,3, então:

$$\mu = 6*0,3 = 1,8$$



- $\star$  Para a <u>distribuição hipergeométrica</u>:  $\mu = \frac{\pi}{a+b}$
- ★ **Exemplo**: Dentre 12 ônibus, 5 estão com os freios gastos. Se 6 desses 12 são escolhidos ao acaso, quantos podemos esperar que tenham freios gastos?

amostragem sem reposição, a = 5, b = 7 e n = 6: 
$$\mu = \frac{6 \cdot 5}{5 + 7} = 2.5$$

★ Como selecionamos metade do total, é de se esperar que metade dos que estão com os freios ruins estejam incluídos.



- $\bigstar$  Para uma distribuição qualquer a variância é:  $\sigma^2 = \sum (x \mu)^2 \cdot f(x)$
- ★ Também chamada de variância da variável aleatória ou a variância de sua distribuição de probabilidade, a sua raiz quadrada nos fornece o desvio-padrão populacional.

**\star** Variância da distribuição binomial:  $\sigma^2 = np(1-p)$ 



**Exemplo**: Encontre o desvio-padrão para o exemplo do mercado e da promoção. (Calculamos  $\mu = 6*0.3 = 1.8$ ; aqui o n = 6 e p = 0.3)

Número de pessoas	Probabilidade	Desvio da média	Quadrado do desvio da média	$(x-\mu)^2 f(x)$	
0	0,118	-1,8	3,24	0,38232	
1	1 0,303	-0.8	0,64	0,19392	
2	0.324	0.2	0.04	0.01296	1.26604
3	0.185	1,2	1,44	0,26640	$\sigma = \sqrt{1,26604} =$
4	0,060	2,2	4,84	0,29040	
5	0,010	3,2	10,24	0,10240	
6	0,001	4,2	17.64	0,01764	
				$\sigma^2 = 1,26604$	

OU: 
$$\sigma^2 = np(1-p)$$
  $\sigma^2 = 6(0,30)(0,70) = 1,26$   $\sigma = 1,12$ 



- ★ Vale lembrar: O desvio-padrão mede o tamanho esperado das flutuações ao acaso de uma variável aleatória, quando pequeno, existe uma alta probabilidade de que resulte em um valor próximo da média, quando grande é mais provável que resulte bem afastado da média.
- ★ Isso foi visto quando falamos do Teorema de Tchebichev, que continua válido:

A probabilidade de uma variável aleatória tomar um valor a menos de k desvios-padrão de qualquer um dos dois lados da média é pelo menos  $1 - \frac{1}{k^2}$ .

 $\star$  A menos de k desvios-padrão da média:  $\mu - k\sigma e \mu + k\sigma$ 



**Exemplo**: A quantidade de chamadas telefônicas recebidas por um serviço de atendimento entre as 9 e as 10 horas da manhã é uma variável aleatória cuja distribuição tem a média  $\mu$  = 27,5 e desvio-padrão  $\sigma$  = 3,2. O que o teorema de Tchebichev, com k = 3, nos diz sobre o número de chamadas telefônicas que o serviço pode esperar receber nesse horário?

$$\mu - 3\sigma = 27.5 - 3(3.2) = 17.9 \text{ e } \mu + 3\sigma = 27.5 + 3(3.2) = 37.1$$

podemos garantir com probabilidade de pelo menos 1 -  $1/3^2 = 8/9 \approx 0.89$  que receberá entre 17,9 (18) e 37,1 (37) chamadas.



**★ Exemplo**: O que o teorema de Tchebichev, com k = 5, pode nos dizer sobre o número de caras e, portanto, sobre a proporção de caras que podemos obter em 400 jogadas de uma moeda equilibrada?

Distr. Binomial: n = 400 e p = 0,50, então  $\mu$  = 400\*0,50 = 200 e  $\sigma$  =  $\sqrt{(400*0,50*0,50)}$  = 10.

$$\mu$$
 - 5 $\sigma$  = 200 - 5\*10 = 150  $\mu$  + 5 $\sigma$  = 200 + 5\*10 = 250

podemos garantir com probabilidade de pelo menos  $1 - 1/5^2 = 24/25 \approx 0,96$  que vamos obter entre 150 e 250 caras ou que a proporção de caras vai ficar entre 150/400 = 0,375 e 250/400 = 0,625.



#### **EXERCÍCIOS**

**★** Lista 3 de Exercícios → Parte 1

https://docs.google.com/document/d/1hebfg13HiDFVPoanhcPTj5a4571tU6vUA1wA1QQtxCl/edit2usp=share\_link



★ Muito obrigado pela atenção!