

Distribuições Discretas: Bernoulli e Binomial

Prof. Alejandra, Cristian, Paula e Rafael

Introdução à Probabilidade e Estatística



Estes slides tem por objetivo apenas complementar e esclarecer os textos principais indicados para a disciplina, e não devem ser usados como único texto de referência.

Eles tratam apenas dos principais aspectos da teoria abordada, deixando para os textos e vídeos os detalhes teóricos e exemplos mais elaborados.

Junto à essas notas o aluno deve ler os textos e ver os vídeos indicados no site do curso no Moodle, além da referência principal do curso.

- Dizemos que as variáveis aleatórias discretas são aquelas que assumem valores em um conjunto finito ou *enumerável*. Por exemplo $\{x_1, x_2, x_3, \dots\}$.
- Que se X é uma variável aleatória discreta podemos definir a função de probabilidade $p(a)$ de X como $p(a) = P(X = a)$.
- Duas v.a.'s com a mesma função distribuição são chamadas de *igualmente distribuídas*.
- Que a função de distribuição F para uma variável aleatória discreta X pode ser escrita em termos de $p(a)$:

$$F(a) = P(X \leq a) = \sum_{\text{todo } x \leq a} p(x).$$

- E que se X é uma variável aleatória com função de probabilidade $p(x)$ então a esperança de X , denotada por $E[X]$, é definida por

$$E[X] = \sum_{x:p(x)>0} xp(x).$$

Considere as seguintes situações:

- uma peça é classificada como boa ou defeituosa;
- o resultado de um exame médico para detecção de uma doença é positivo ou negativo;
- um paciente submetido a um tratamento, durante um período de tempo fixo, cura-se ou não da doença;
- um entrevistado concorda ou não com a afirmação feita;
- no lançamento de um dado, ocorre ou não a face “5”.

As situações anteriores apresentam só duas alternativas e cada uma delas podem ser representada, genericamente, como *sucesso* ou *fracasso*.

Esses experimentos recebem o nome de *Ensaio de Bernoulli* e dão origem às chamadas v.a.'s com distribuição de Bernoulli.

Suponha que realizamos um experimento e estamos interessados em observar algum evento pré-fixado. Se o evento em questão for observado, declaramos que aconteceu um *sucesso*, e caso contrário, um *fracasso*.

Se X é a variável que assume 1 quando obtemos um sucesso, e 0 no caso de um fracasso, então a função de probabilidade de X é dada por

$$\begin{aligned}p(1) &= P(X = 1) = p \\p(0) &= P(X = 0) = 1 - p,\end{aligned}$$

onde $0 \leq p \leq 1$, é a chamada probabilidade de sucesso.

Definição 1

Uma variável aleatória X é chamada de variável aleatória de *Bernoulli* se sua função de probabilidade for dada por

$$\begin{aligned}p(1) &= P(X = 1) = p \\p(0) &= P(X = 0) = 1 - p,\end{aligned}$$

para algum $p \in (0,1)$.

Alternativamente, podemos dizer que X tem distribuição *Bernoulli* com parâmetro p .

Denotaremos isso como $X \sim \text{Ber}(p)$ ou $X \sim \text{Bernoulli}(p)$.

Considere o lançamento de uma moeda e uma variável aleatória X associada a esse experimento dada por

$$X = \begin{cases} 1, & \text{se ocorre cara} \\ 0, & \text{se ocorre coroa} \end{cases}$$

Seja p a probabilidade de cara, $0 < p < 1$. Neste caso, “sucesso” é definido como o lançamento dar cara, que ocorre com probabilidade p , e “fracasso” é observarmos coroa. Assim

$$P(X = 1) = p(1) = p, \text{ e } P(X = 0) = p(0) = 1 - p.$$

Logo, $X \sim \text{Ber}(p)$ para qualquer que seja p .

O que podemos falar da moeda se $p = 1/2$?

Um auditor da Receita Federal examina declarações de Imposto de Renda de pessoas físicas, cuja variação patrimonial ficou acima do limite considerado aceitável. De dados históricos, sabe-se que 10% dessas declarações são fraudulentas. Considere o experimento correspondente ao sorteio aleatório de uma dessas declarações. Seja X a v.a. que indica que a declaração é fraudulenta. X tem distribuição Bernoulli?

Solução: Consideremos como sucesso o evento de que a declaração sorteada seja fraudulenta, escrevendo que $X = 1$ quando o sucesso aconteça. Logo $p(1) = 0,1$.

O complementar deste evento é o fracasso, que denotaremos por $X = 0$. Assim, $p(0) = 0,9$.

Desta maneira, $X \sim Ber(0,1)$

Se $X \sim \text{Ber}(p)$ então X só pode assumir dois valores, 0 ou 1. A Esperança será então:

$$\begin{aligned} E(X) &= 0 \times p(0) + 1 \times p(1) \\ &= p(1) \\ &= p. \end{aligned}$$

Calcularemos a variância valendo-nos da fórmula

$$Va(X) = E(X^2) - (E(X))^2.$$

Observe que a quantidade em vermelho não é conhecida e devemos calculá-la.

$$\begin{aligned} E(X^2) &= 0^2 \times p(0) + 1^2 \times p(1) \\ &= p(1) \\ &= p. \end{aligned}$$

Logo,

$$Va(X) = p - p^2 = p(1 - p)$$

Continuemos com o exemplo do auditor da receita. Lá definimos a variável X que indicava a ocorrência ou não de uma declaração fraudulenta em um sorteio aleatório, e vimos que $X \sim \text{Ber}(0,1)$.

Calculando a esperança e a variância para esta v.a., como $p = 0,1$, temos que

$$E(X) = 0,1$$

$$\text{Var}(X) = 0,1 \times 0,9 = 0,09.$$

E se ao invés de jogar uma moeda, jogarmos 5 moedas? Qual seria a probabilidade de obter 3 caras?

Se selecionarmos 10 declarações, com reposição, de imposto de renda, qual é a probabilidade de ter 8 fraudulentas?

Observe que ambos exemplos acima consistem no mesmo “tipo” de problema: repetimos um ensaio de Bernoulli n vezes, de maneira independente, e estamos interessados em contar o total de sucessos observados.

Exemplo

Considere o lançamento de uma moeda, não honesta, 3 vezes de maneira independente. Seja $p = 1/3$ a probabilidade que a moeda tem de sair cara. Considere X a v.a que conta o número de caras nesses 3 lançamentos. Logo X assume valores em $\{0,1,2,3\}$. Vejamos a função de probabilidade $p(k)$ de X .

$\{X = 0\} = \{(\text{coroa}, \text{coroa}, \text{coroa})\}$.  Quantos resultados favoráveis temos? Só 1 = $\binom{3}{0}$ Logo,

$$P(X = 0) = \left(\frac{1}{3}\right)^0 \left(\frac{2}{3}\right)^3 = \frac{8}{27}.$$

$\{X = 1\} = \{(\text{cara}, \text{coroa}, \text{coroa}) ; (\text{coroa}, \text{cara}, \text{coroa}) ; (\text{coroa}, \text{coroa}, \text{cara})\}$

Quantos resultados favoráveis temos? 3 = $\binom{3}{1}$  Logo,

$$P(X = 1) = 3 \times \left(\frac{1}{3}\right)^1 \left(\frac{2}{3}\right)^2 = \frac{12}{27}.$$

$\{X = 2\} = \{(\text{cara}, \text{cara}, \text{coroa}) ; (\text{coroa}, \text{cara}, \text{cara}) ; (\text{cara}, \text{coroa}, \text{cara})\}$

Quantos resultados favoráveis temos? 3 = $\binom{3}{2}$  Logo,

$$P(X = 2) = 3 \times \left(\frac{1}{3}\right)^2 \left(\frac{2}{3}\right)^1 = \frac{6}{27}.$$

$\{X = 3\} = \{(cara, cara, cara)\}$. $\leftarrow$ Quantos resultados favoráveis temos? Só 1 = $\binom{3}{3}$ Logo,

$$P(X = 3) = 1 \times \left(\frac{1}{3}\right)^3 \left(\frac{2}{3}\right)^0 = \frac{1}{27}.$$

Como fica a probabilidade? A função de probabilidade $p(k)$ para qualquer $k \in \{0, 1, 2, 3\}$

$$p(k) = \binom{3}{k} \left(\frac{1}{3}\right)^k \left(\frac{2}{3}\right)^{3-k}$$

Veja que se tivéssemos lançado a moeda uma vez só, teríamos um ensaio Bernoulli com parâmetros $1/3$, mas lançamos a moeda 3 vezes. Ou seja, tivemos a repetição de 3 ensaios Bernoulli independentes. Este tipo de experimento é chamado *experimento Binomial*.

Definição 2

A v.a. X que conta o número de sucessos em n ensaios de Bernoulli independentes, cada um com probabilidade de sucesso p , é chamada v.a. *Binomial* de parâmetros n e p .

A função de probabilidade de X é dada por

$$p(k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, 2, \dots, n.$$

Também dizemos que X tem *distribuição Binomial* de parâmetros n e p

Notação: $X \sim \text{Bin}(n, p)$ ou $X \sim \text{Binomial}(n, p)$.

Observe que uma variável aleatória de Bernoulli é tão somente uma variável aleatória binomial com parâmetros 1 e p .

Exemplo

Quatro moedas honestas são jogadas. Se os resultados são por hipótese independentes, determine a função de probabilidade do número de caras obtido.

Se X é igual ao número de caras (sucessos) que aparecem, então $X \sim \text{Bin}(4, 1/2)$. Portanto,

$$p(0) = P(X=0) = \binom{4}{0} \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^4 = \frac{1}{16}$$

$$p(1) = P(X=1) = \binom{4}{1} \left(\frac{1}{2}\right)^1 \left(\frac{1}{2}\right)^3 = \frac{4}{16}$$

$$p(2) = P(X=2) = \binom{4}{2} \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^2 = \frac{6}{16}$$

$$p(3) = P(X=3) = \binom{4}{3} \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^1 = \frac{4}{16}$$

$$p(4) = P(X=4) = \binom{4}{4} \left(\frac{1}{2}\right)^4 \left(\frac{1}{2}\right)^0 = \frac{1}{16}$$

Note que

$$\sum_{k=0}^4 p(k) = \frac{1}{16} + \frac{4}{16} + \frac{6}{16} + \frac{4}{16} + \frac{1}{16}.$$

e que a função distribuição de X é dada por

$$F(x) = \begin{cases} 0 & , x < 0 \\ \sum_{k=0}^m \binom{n}{k} p^k (1-p)^{n-k} & , m \leq x < m+1, m = 0, \dots, n-1. \\ 1 & , x \geq n \end{cases}$$

Em um teste de múltipla escolha com 3 respostas possíveis para cada uma das 5 questões, qual é a probabilidade de que um estudante acerte 4 questões ou mais apenas chutando?

Solução: Se X é o número acertos do estudante, então X é uma variável aleatória binomial com parâmetros $(5, 1/3)$. Portanto, a probabilidade de 4 ou mais questões é

$$\begin{aligned}P(X \geq 4) &= P(X = 4) + P(X = 5) \\&= \binom{5}{4} \left(\frac{1}{3}\right)^4 \left(\frac{2}{3}\right)^1 + \binom{5}{5} \left(\frac{1}{3}\right)^5 \left(\frac{2}{3}\right)^0 \\&= \frac{11}{243}.\end{aligned}$$

Vamos agora examinar as propriedades de uma variável aleatória binomial com parâmetros n e p .

Esperança e Variância

Pode-se provar que

$$E(X) = np, \quad E(X^2) = np[(n-1)p + 1] \quad \text{e} \quad \text{Var}(X) = np(1-p).$$

(*Desafio: Tentar calcular a esperança!*)

Exemplo

Continuando com o exemplo do teste de múltipla escolha

O número médio de acertos do estudante, em 5 questões cada uma delas com 3 opções, seria,

$$E(X) = 5 \times \frac{1}{3} = \frac{5}{3} \approx 1,66.$$

E a variância seria,

$$Var(X) = 5 \times \frac{1}{3} \times \frac{2}{3} = \frac{10}{9} \approx 1,11,$$

com desvio padrão de $\sqrt{1,11} \approx 1,05$.

Para verificar se os lotes de peças recebidos em uma determinada fábrica são aceitos, é realizado um plano de controle que consiste em selecionar 10 itens aleatoriamente, com reposição, e determinar o número de peças defeituosas. Um lote é rejeitado se duas ou mais peças defeituosas forem encontradas. Qual é a probabilidade de aceitar lotes com 5% de peças defeituosas?

Solução: Seja X a v.a que conta o número de peças defeituosas em um lote. Como a probabilidade de uma peça ser defeituosa do lote será 0,05, então $X \sim \text{Bin}(10; 0,05)$. Para aceitar o lote, devemos obter, entre as 10 peças selecionadas, menos de duas peças defeituosas. Pelo tanto, a probabilidade de aceitar o lote é $P(X < 2)$.

$$\begin{aligned}P(X < 2) &= p(0) + p(1) \\&= \binom{10}{0} (0,05)^0 (0,95)^{10} + \binom{10}{1} (0,05)^1 (0,95)^9 \\&= 0,599 + 0,315 \\&= 0,914.\end{aligned}$$

Qual é q quantidade esperada de peças defeituosas?

$$E(X) = 10 \times 0,05 = 0,5.$$

E o desvio padrão?

$$Var(X) = 10 \times 0,05 \times 0,95 = 0,475,$$

logo o desvio padrão é $\sigma = \sqrt{Var(X)} = 0,689$.

Exemplo

Em uma distribuição binomial, sabe-se que a média é 4,5 e a variância é 3,15. Encontre os valores dos parâmetros da distribuição.

Solução: O exercício pede para encontrar n e p . Para isso, temos que

$$E(X) = np = 4,5$$

$$Var(X) = np(1 - p) = 3,15.$$

Substituindo a primeira equação na segunda, resulta

$$4,5(1 - p) = 3,15 \Rightarrow p = 0,3$$

Substituindo na primeira equação, obtemos que

$$n \times 0,3 = 4,5 \Rightarrow n = 15.$$

Uma urna contém 3 bolas numeradas (1, 2 e 3). Primeiro uma bola é retirada da urna. Se sair a bola i , seleciona-se i peças com reposição de um lote contendo 60% de peças defeituosas. Seja X o número de peças defeituosas na amostra.

- 1 Ache o conjunto de valores possíveis de X .

Se $i = 1$, então $X \in \{0, 1\}$. Se $i = 2$, então $X \in \{0, 1, 2\}$ e finalmente se $i = 3$, então $X \in \{0, 1, 2, 3\}$. Logo, X pode assumir valores no conjunto $\{0, 1, 2, 3\}$.

- 2 Porque não temos distribuição binomial?

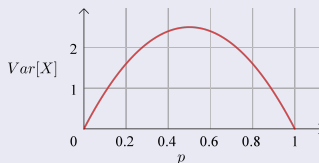
Porque o número de peças defeituosas em cada retirada do lote **não é** constante, e depende do número sorteado na urna! Logo, X não pode ser uma binomial. Mas, dado o número da bola sorteada na urna temos sim uma distribuição binomial. Ou seja, conhecido o número sorteado i , $X \sim \text{Bin}(i; 0,6)$.

Desafio: Você consegue calcular $E(X)$?

Observações

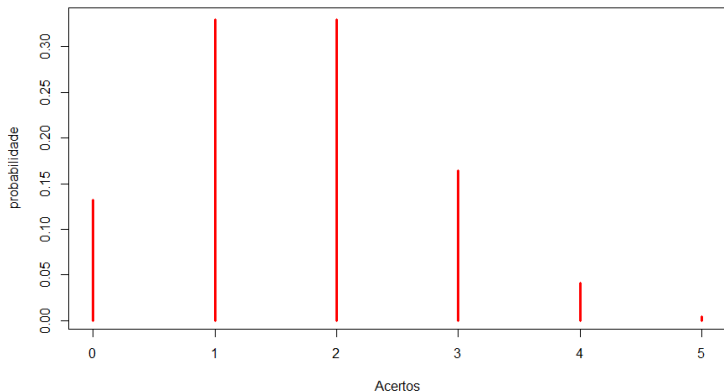
Se $X \sim \text{Bin}(n, p)$, então

- À medida que k varia de 0 a n , $P(X = k) = p(k)$ primeiro cresce monotonicamente e depois decresce monotonicamente, atingindo seu maior valor quando k é o maior inteiro menor ou igual a $(n+1)p$;
- A esperança $E[X] = np$ cresce com n e com p ;
- A variância $\text{Var}[X] = np(1-p)$ também cresce com n , mas a medida que p cresce de 0 a 1, ela se comporta como uma parábola, atingindo o seu máximo quando $p = 0,5$.



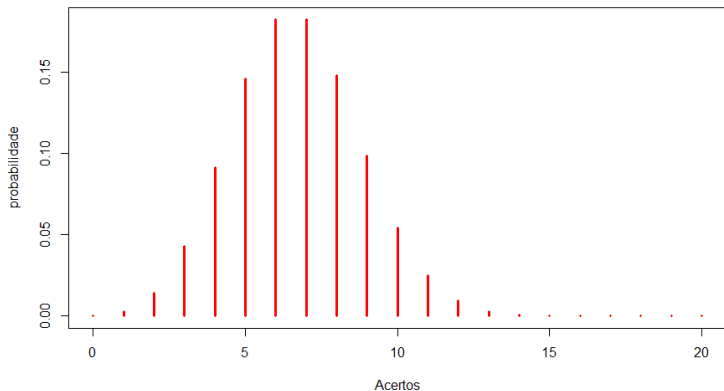
Exemplo: Acertos parte 1

A seguinte figura mostra a distribuição da probabilidade da v.a número de acertos em 5 questões, dado que cada questão tem 3 opções, $X \sim (5, 1/3)$.



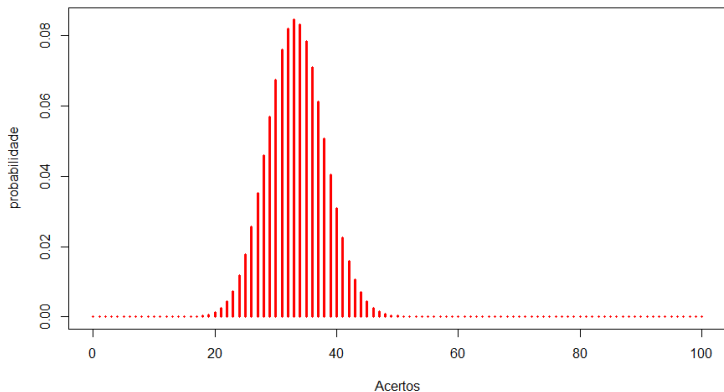
Exemplo: Acertos parte 2

Se se agora tivermos 20 questões ao invés de 5?, Ou seja, $X \sim (20, 1/3)$.



Exemplo: Acertos parte 3

Se se agora tivermos 100 questões ao invés de 20?, Ou seja, $X \sim (100, 1/3)$.



O que você pode concluir?

Suponha que uma moeda viciada que dê cara com probabilidade p seja jogada 10 vezes. Dado que um total de 6 caras foi observado, determine a probabilidade condicional de que

- ❶ o primeiro resultado tenha sido cara, o segundo coroa, e o terceiro coroa
- ❷ o primeiro resultado tenha sido coroa, o segundo cara, e o terceiro coroa