Princípios de Bioestatística
Variável Aleatória/Modelos Binomial e Poisson

Enrico A. Colosimo/UFMG
http://www.est.ufmg.br/~enricoc/

Depto. Estatística - ICEx - UFMG

Variável Aleatória

- Conceito importante em Estatística.
- Uma função que associa cada possível resultado do espaço amostral a um valor numérico.
- Pode ser discreta ou contínua.

Variável Aleatória (discreta)

Definição

Uma variável X (função), associada a cada possível resultado do espaço amostral, é denominada **variável aleatória discreta**, se assume valores num conjunto enumerável ou contável.

- Um dado é lançado duas vezes. O espaço amostral deste experimento é:

$$\Omega = \left\{ \begin{array}{cccccc} (1,1), & (1,2), & (1,3), & (1,4), & (1,5), & (1,6), \\ (2,1), & (2,2), & (2,3), & (2,4), & (2,5), & (2,6), \\ (3,1), & (3,2), & (3,3), & (3,4), & (3,5), & (3,6), \\ (4,1), & (4,2), & (4,3), & (4,4), & (4,5), & (4,6), \\ (5,1), & (5,2), & (5,3), & (5,4), & (5,5), & (5,6), \\ (6,1), & (6,2), & (6,3), & (6,4), & (6,5), & (6,6) \end{array} \right\}$$

- Seja X a variável *soma dos dois lançamentos*, ou seja, $X = \text{“face do primeiro lançamento"} + \text{“face do segundo lançamento"}$.

- Quando o evento $(1,1)$ ocorre, X associa a este resultado o valor 2. Da mesma forma temos:

$$\Omega = \left\{ \begin{array}{cccccc} (1,1), & (1,2), & (1,3), & (1,4), & (1,5), & (1,6), \\ (2,1), & (2,2), & (2,3), & (2,4), & (2,5), & (2,6), \\ (3,1), & (3,2), & (3,3), & (3,4), & (3,5), & (3,6), \\ (4,1), & (4,2), & (4,3), & (4,4), & (4,5), & (4,6), \\ (5,1), & (5,2), & (5,3), & (5,4), & (5,5), & (5,6), \\ (6,1), & (6,2), & (6,3), & (6,4), & (6,5), & (6,6) \end{array} \right\} \overset{X}{\Rightarrow} \left\{ \begin{array}{cccccc} 2, & 3, & 4, & 5, & 6, & 7, \\ 3, & 4, & 5, & 6, & 7, & 8, \\ 4, & 5, & 6, & 7, & 8, & 9, \\ 5, & 6, & 7, & 8, & 9, & 10, \\ 6, & 7, & 8, & 9, & 10, & 11, \\ 7, & 8, & 9, & 10, & 11, & 12 \end{array} \right\}$$

- X assume valores no conjunto $\{2,3,4,5,6,7,8,9,10,11,12\}$.

Definição

Seja X uma variável aleatória discreta e $x_1, x_2, x_3, \ldots$, seus diferentes valores. A função que atribui a cada valor da variável aleatória sua probabilidade é denominada **função discreta de probabilidade** ou, simplesmente **função de probabilidade**.

Retomando o exemplo do lançamento dos dois dados.

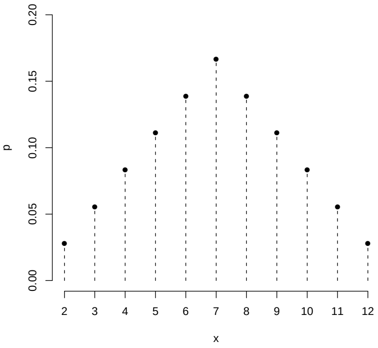
- X assume valores no conjunto $\{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$.
- A probabilidade dos possíveis resultados em Ω é $P[(1, 1)] = P[(1, 2)] = \ldots = P[(2, 1)] = P[(2, 2)] = \ldots = P[(6, 6)] = 1/36$.
- $P[X = 2] = P[(1, 1)] = 1/36$,
  $P[X = 3] = P[(1, 2) \cup (2, 1)] = 1/36 + 1/36 = 2/36$.

Variável aleatória discreta - Função de Probabilidade

Exemplo: lançamento de dois dados

- A função de probabilidade de X é dada por:

X	2	3	4	5	6	7	8	9	10	11	12
$p_X = P(X = x)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36



- **Observações:**
  - $P(X = x)$ pertence ao intervalo $(0, 1)$, para $X = 2, \ldots, 12$,
  - $P(X = 2) + P(X = 3) + \ldots P(X = 12) = 1/36 + 2/36 + \ldots + 1/36 = 1$
- **Pergunta:** Qual a probabilidade da soma dos resultados dos dois lançamentos ser menor do que 6?

$$\begin{aligned} P(X < 6) &= P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5) \\ &= 1/36 + 2/36 + 3/36 + 4/36 = 10/36. \end{aligned}$$

Importante:

- Probabilidade é sempre maior ou igual a zero.
- A soma de todas as probabilidades tem que totalizar um.

Exemplo:

- Observa-se o sexo das crianças em famílias com três filhos.
- Denotamos *m* para o sexo masculino e *f* para o sexo feminino.
- Existem oito possibilidades para uma família de três filhos. Estas possibilidades são listadas no espaço amostral:

$$\Omega = \{(mmm), (mmf), (mfm), (fmm), (mff), (fmf), (ffm), (fff)\}.$$

Variável aleatória discreta

Definimos:

- X : número de crianças do sexo masculino (m).
- A cada possível resultado do espaço amostral, X associa um valor numérico

Ω	*mmm*	*mmf*	*mfm*	*fmm*	*mff*	*fmf*	*ffm*	*fff*
X	3	2	2	2	1	1	1	0

- Note que X assume valores no conjunto $\{0, 1, 2, 3\}$ que é enumerável, portanto X é variável aleatória discreta.

Variável aleatória discreta

- **Pergunta:** qual é a função de probabilidade de X?

- Cada resultado possível do espaço amostral tem probabilidade $\frac{1}{8}$ de acontecer, então:

$$P(X=0) = P(fff) = \frac{1}{8}$$

- A probabilidade da variável aleatória X assumir o valor zero é a mesma probabilidade do evento (fff) ocorrer. Da mesma forma:

$$\begin{aligned}
P(X=1) &= P(\{mff\} \cup \{fmf\} \cup \{ffm\}) \\
&= P(mff) + P(fmf) + P(ffm) = \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{3}{8} \\
P(X=2) &= P(\{mmf\} \cup \{mfm\} \cup \{fmm\}) = \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{3}{8} \\
P(X=3) &= P(mmm) = \frac{1}{8}
\end{aligned}$$

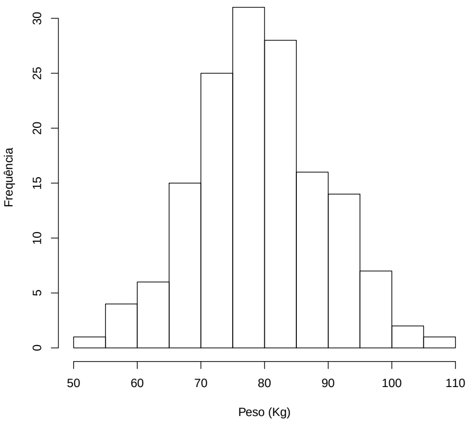
- **Resumindo:**

X	0	1	2	3
$P(X = x)$	1/8	3/8	3/8	1/8

Nomenclatura: Esta Variável Aleatória é chamada de Binomial

Variável Aleatória Contínua

Exemplo: Histograma de uma amostra de pesos de 150 homens de uma certa população.



Os valores de peso não estão em um conjunto enumerável.

Variável Aleatória Contínua

- A probabilidade, no caso contínuo, é caracterizada a partir de uma função positiva denominada densidade de probabilidade.
- Definição: f(x)é uma função de densidade de probabilidade de uma variável aleatória contínua X, se:
  - $f(x) \geq 0,$ para todo x,
  - A área definida por $f(x)$ é igual a 1: $\int_x f(x) = 1$
  - O cálculo de probabilidade é definido como: $P(a \leq X \leq b) = \int_a^b f(x)dx$ onde $a < b$

Principais modelos probabilísticos

1. Discreto
    - Uniforme;
    - Binomial;
    - Poisson.
2. Contínuo
    - Normal
    - Outros: exponencial, gama, etc

Principais modelos discretos

- Algumas variáveis aleatórias aparecem com bastante frequência em situações práticas e justificam um estudo mais cuidadoso.
- Em geral nesses casos, a função de probabilidade pode ser escrita de uma maneira mais compacta, isto é, existe uma **lei** para atribuir as probabilidades.
- Por exemplo, se uma variável aleatória W tem função de probabilidade dada por

X	1	2	3	4	5	6
$P(X = x)$	1/21	2/21	3/21	4/21	5/21	6/21

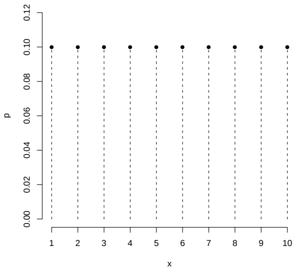
- Uma forma abreviada de apresentar a variável e sua função de probabilidade é $P(X = x) = x/21, x = 1, 2, \ldots, 6$.

Modelo Uniforme Discreto

- Seja X uma variável aleatória cujos possíveis valores são representados por $x_1, x_2, \ldots, x_k$.
- X segue o modelo *Uniforme Discreto* se atribui a **mesma probabilidade** $1/k$ a cada um desses k valores.
- A função de probabilidade é dada por

$$P(X = x) = 1/k, \quad \text{para} \quad x = 1, 2, \ldots, k.$$

- Para $k = 10$ o gráfico da função de probabilidade é



- **Observação:**
  1. Todos os valores têm a mesma probabilidade.
  2. Modelo do sorteio aleatório.

- Em muitas situações práticas a variável de interesse assume somente dois valores.
  - Em um levantamento do sistema de saúde, as pessoas são perguntadas se tem ou não um seguro de hospitalização.
  - Amostras de sangue são testadas para a presença ou ausência de um antígeno.
  - Ratos alimentados com uma substância cancerígena em potencial são examinados para tumores.
  - Crianças recém-nascidas são classificadas como tendo ou não a síndrome de Down.
  - Etc, etc...

- Estas situações têm alternativas dicotômicas, que genericamente podem ser representadas por respostas do tipo *sucesso-fracasso*.
- Esses experimentos recebem o nome de *Ensaios de Bernoulli* e dão origem a uma variável aleatória com o mesmo nome.

Modelo Bernoulli

X	0	1
$P(X = x)$	$1 - p$	p

Modelo Binomial

- Considere a repetição de n ensaios de Bernoulli independentes.
- Todos com a mesma probabilidade de sucesso p.
- X: número total de sucessos (Binomial com parâmetros n e p).

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, k = 0, 1, \ldots, n,$$

com $\binom{n}{k}$ representando o coeficiente binomial calculado por

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

- O coeficiente binomial $\binom{n}{k}$ é o número de diferentes maneiras que k objetos podem ser selecionados a partir de n objetos.

- **Exemplo:** De 10 residentes, três são escolhidos para cobrir um serviço hospitalar em um feriado. De quantas maneiras podem ser escolhidos os residentes?

$$\binom{10}{3} = \frac{10!}{3!7!} = \frac{1 \times 2 \times 3 \times 4 \times 5 \times 6 \times 7 \times 8 \times 9 \times 10}{(1 \times 2 \times 3)(1 \times 2 \times 3 \times 4 \times 5 \times 6 \times 7)} = 120$$

Retomando o Exemplo

- Observa-se o sexo das crianças em famílias com três filhos.
- X : número de crianças do sexo masculino (m).

X	0	1	2	3
$P(X = x)$	1/8	3/8	3/8	1/8

$$P(X = k) = \binom{3}{k} 0,5^k (1 - 0,5)^{3-k}, k = 0, 1, \ldots, 3,$$

com $\binom{3}{k}$ representando o coeficiente binomial calculado por

$$\binom{3}{k} = \frac{3!}{k!(3-k)!}.$$

Notação: $X \sim b(3, 0, 5)$ indica que a variável aleatória X segue o modelo Binomial com parâmetros $n = 3$ e $p = 0, 5$.

As seguintes condições caracterizam um ensaio binomial:

1. Cada ensaio assume uma e apenas uma de duas possibilidades (desfecho dicotômico).
2. O número de ensaios é fixo. Cada observação é um *ensaio de Bernoulli*.
3. A probabilidade de *sucesso* é a mesma (constante) em cada ensaio. Ou dito de forma equivalente, os ensaios são independentes.

Exemplo: Modelo Binomial

- **Exemplo:** A prevalência de um certa doença em uma população é 10%. Uma amostra de 100 pessoas foi retirada desta população. Defina X como sendo o número de doentes na amostra.

  1. X tem distribuição binomial?
  2. Qual é a probabilidade de 5 doentes na amostra?

Exemplo: Modelo Binomial

- **Exemplo:** A prevalência de um certa doença em uma população é 10%. Uma amostra de 100 pessoas foi retirada desta população. Defina X como sendo o número de doentes na amostra.

    1. X tem distribuição binomial?
    2. Qual é a probabilidade de 5 doentes na amostra?

Respostas:

1. Binomial para populações grandes;
2. $P(X = 5) = \binom{100}{5} 0,1^5 (1 - 0,1)^{100-5} = 0,034$

- **Exemplo:** Dez pacientes são tratados cirurgicamente, com uma probabilidade de 70% da cirurgia ser bem sucedida ($X \sim b(10; 0,7)$).
  - Qual é a probabilidade de apenas cinco cirurgias serem bem sucedidas?

$$\begin{aligned} P(X=5) &= \binom{10}{5} 0,7^5 (1-0,7)^{10-5} \\ &= \frac{10!}{5!5!} \times 0,7^5 \times 0,3^5 \\ &= 252 \times 0,16807 \times 0,00243 = 0,1029 \end{aligned}$$

  - Qual é a probabilidade de apenas cinco ou menos cirurgias serem bem sucedidas?

$$\begin{aligned} P(X \leq 5) &= P(X=0) + P(X=1) + P(X=2) \\ &+ P(X=3) + P(X=4) + P(X=5) \\ &= 0,0000 + 0,0001 + 0,0014 + 0,0090 + 0,0368 + 0,1029 \\ &= 0,1503 \end{aligned}$$

Média

Valor Esperado (Média)

Dada a variável aleatória X, número de sucessos em n ensaios de Bernoulli com probabilidade p de sucesso. A média de X é

$$E(X) = np.$$

- No exemplo da cirurgia, a média de X é

  $E(X) = 10 \times 0,7 = 7$ cirurgias bem sucedidas em 10 realizadas.

- X : número de ocorrências de um evento em um período de tempo ou unidade de área ou volume.
- Exemplos:
  - número de óbitos mensais em um hospital,
  - número de ovos por unidade de volume de fezes,
  - número de células infectadas por mm^2,
  - número de pacientes que chegam diariamente em uma unidade de pronto atendimento.
- A distribuição de X é caracterizada pelo número médio de ocorrências que será denominada por λ.

Exemplo: X: número de chegadas de pacientes em um pronto socorro no período de 0 as 8 hs.

X :0,1,2,3,... e $\lambda = 3$ (pacientes/perídodo).

X	0	1	2	3	4	5	6	7	8	9	10
$P(X = x)$	0,05	0,15	0,22	0,22	0,17	0,10	0,05	0,02	0,01	0,003	0,001

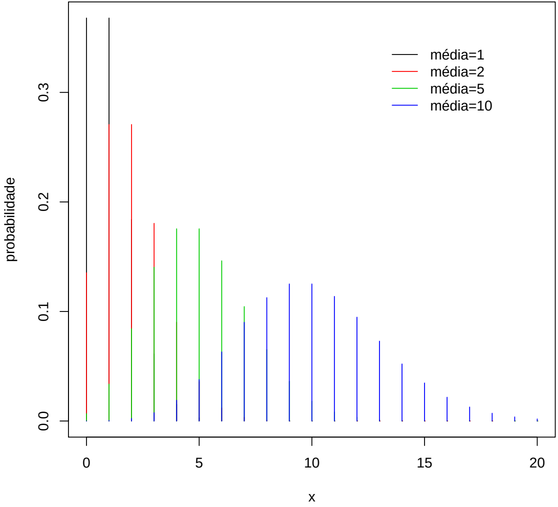
- É muito provável que chegue pelo menos um paciente $P(X > 0) = 0,95$.
- Em particular, é bastante provável que chegue entre 1 e 4 pacientes: $P(0 < X < 5) = 0,77$.
- É muito pouco provável que chegue mais de 8 pacientes.

Modelo de Poisson

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

em que $x = 0, 1, 2, 3, 4, 5, ...$

Modelo de Poisson



Exemplo: Modelo de Poisson

Em um certo plano de saúde, o número médio de consultas por associado é 2,8 por ano. A administração do plano gostaria de saber qual é a probabilidade de um determinado associado ao longo de um ano:

- não fazer nenhuma consulta ao longo de um ano?
- fazer uma única consulta.
- fazer pelo menos duas consultas?

Respostas:

$$P(X = 0) = \frac{e^{-2,8} 2,8^0}{0!} = e^{-2,8} = 0,06$$

Exemplo: Modelo de Poisson

$$P(X = 1) = \frac{e^{-2,8} 2,8^1}{1!} = 0,17$$

$$P(X \geq 2) = 1 - (P(X = 0) + P(X = 1)) = 1 - (0,06 + \frac{e^{-2,8} 2,8^1}{1!}) = 0,77.$$