Variáveis aleatórias

Variáveis aleatórias discretas e contínuas; esperança e variância

Prof. Me. Lineu Alberto Cavazani de Freitas

Departamento de Estatística
Laboratório de Estatística e Geoinformação

UFPR
UNIVERSIDADE FEDERAL DO PARANÁ

Introdução

Variáveis aleatórias

- Variáveis aleatórias são uma continuação do assunto sobre experimentos aleatórios.
- Uma **variável aleatória** (v.a) associa um **valor numérico** a cada resultado de um fenômeno aleatório.
- **Letras maiúsculas** denotam a variável aleatória (X, Y, etc).
- **Letras minúsculas** denotam os **valores observados** da variável aleatória (x, y, etc).

Variáveis aleatórias

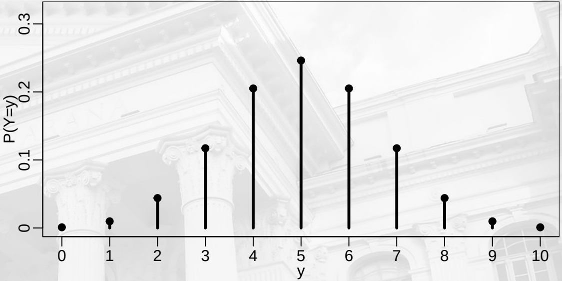
- Os tipos de variáveis aleatórias estão associados ao tipo do espaço amostral.
- Uma variável aleatória é **discreta** se seu **espaço amostral** é discreto.
- Uma variável aleatória é **contínua** se seu **espaço amostral** é contínuo.
- Alguns possíveis espaços amostrais são:
  - Toda a reta real.
  - Valores reais estritamente positivos.
  - Valores positivos mas com zeros sendo possíveis resultados.
  - Intervalos restritos (proporções, índices).
  - Contagens (limitadas ou ilimitadas).
  - Dentre outras.

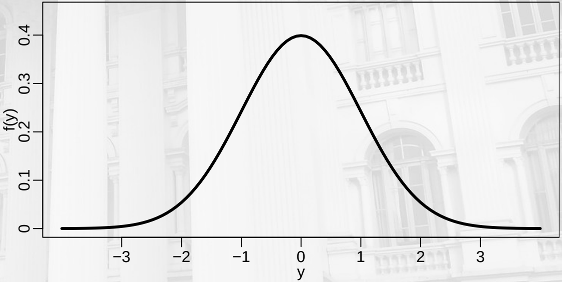
Exemplos

- Y_1: Qualidade de uma peça produzida.
  - $Y_1 = \begin{cases} 1, \text{se defeituosa.} \\ 0, \text{se nao defeituosa.} \end{cases}$
- Y_2: Número de caras observadas no lançamento de 3 moedas.
  - $Y_2 = \{0{,}1{,}2{,}3\}$.
- Y_3: Km em que ocorre um defeito em uma pista de 20 km.
  - $Y_3 = [0{,}20]$.
- Y_4: Proporção de indivíduos que apresenta determinada característica.
  - $Y_4 = [0{,}1]$.
- Y_5: Peso de um indivíduo.
  - $Y_5 = (0,\infty)$.

Variáveis aleatórias

- Existem **funções** que atribuem **probabilidades** aos possíveis valores de uma variável aleatória.
- No caso discreto esta função é chamada de **função de probabilidade** (fp).
- No caso contínuo esta função é chamada de **função densidade de probabilidade** (fdp).
- Podemos descrever o comportamento destas funções por meio de medidas como **média**, **variância** e **covariância**.





Variáveis aleatórias discretas

Variáveis aleatórias discretas

- Considere uma **variável aleatória discreta** Y, que assume os valores $y_1, y_2, \ldots, y_n$.
- A **função de probabilidade** (fp) atribui probabilidades a cada possível valor de Y. É dada por:

$$P(Y = y_i) = p(y_i) = p_i, \ i = 1{,}2, \ldots$$

- As propriedades da função seguem as regras de probabilidades, ou seja:
  - A probabilidade de cada valor deve estar entre 0 e 1.

$$0 \leq p(y_i) \leq 1, \ \forall \ i = 1{,}2, \ldots$$

  - A soma de todas as probabilidades é igual a 1.

$$\sum_i p(y_i) = 1.$$

Exemplo

Considere o experimento aleatório que consiste no lançamento de 3 moedas honestas. Seja Y a variável aleatória que representa o número de caras nos 3 lançamentos:

a) Defina os eventos.
b) Qual é o espaço amostral?
c) Encontre a função probabilidade de Y.
d) Verifique se as propriedades de uma função de probabilidades são atendidas.

Exemplo

a) Defina os eventos.
  - C: cara, K: cora.

b) Qual é o espaço amostral?

$$\Omega = \{(C,C,C),(C,C,K),(C,K,C),(K,C,C),\\ (C,K,K),(K,C,K),(K,K,C),(K,K,K)\}$$

c) Encontre a função probabilidade de Y.
  - Y: número de caras.
  - $y = \{0,1,2,3\}$.
  - Função de probabilidade:

y	0	1	2	3
Prob.	1/8	3/8	3/8	1/8

d) Verifique se as propriedades de uma função de probabilidades são atendidas.
  - As probabilidades estão entre 0 e 1?
    - Sim.
  - A soma das probabilidades é igual a 1?
    - $1/8 + 3/8 + 3/8 + 1/8 = 1$.

Função de distribuição de probabilidade

- Em algumas situações pode ser interessante calcular a **probabilidade acumulada** até um certo valor da variável aleatória.
- Para isso, podemos usar a **função de distribuição**, também chamada de **função acumulada de probabiliade**.
- Trata-se de um análogo à distribuição de frequência acumulada.
- A função de distribuição ou função acumulada de probabilidade de uma variável aleatória Y é dada pela expressão:

$$F(y) = P(Y \leq y)$$

Exemplo

Retomando o exemplo do lançamento de 3 moedas honestas:

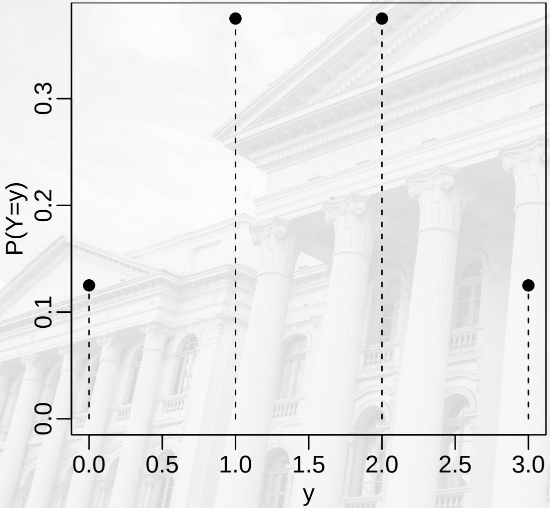
- Y: número de caras.
- $y = \{0,1,2,3\}$.
- Função de probabilidade:

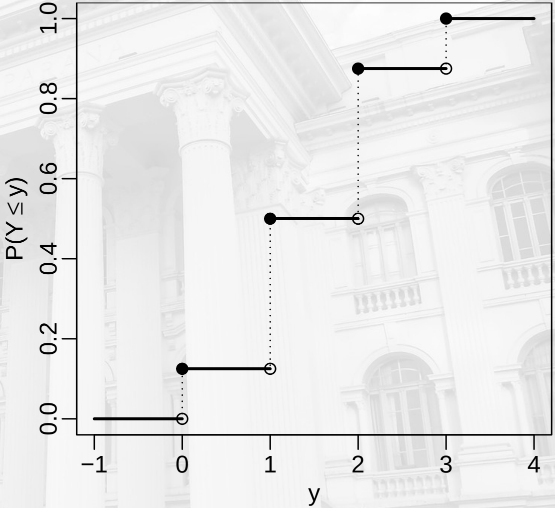
y	0	1	2	3
Prob.	1/8	3/8	3/8	1/8
Prob.Ac.	1/8	4/8	7/8	8/8

$$F(y) = \begin{cases} 0, & se\ y < 0 \\ 1/8, & se\ 0 \leq y < 1 \\ 4/8, & se\ 1 \leq y < 2 \\ 7/8, & se\ 2 \leq y < 3 \\ 1, & se\ y \geq 3 \end{cases}$$

Exemplo

De forma gráfica:





Esperança e variância para variáveis aleatórias discretas

- Sabemos que a descrição completa do comportamento de uma variável aleatória discreta é feita por meio da sua função de probabilidade.
- Desta forma presume-se que qualquer quantidade destinada a **resumir** os valores observados da variável devem envolver esta função.
- Podemos obter resumos numéricos da função de probabilidade como média, mediana, moda, variância, desvio padrão.

Esperança

- A **média**, também chamada de **valor esperado** ou **esperança**, representa o ponto de equilíbrio da distribuição dos valores da variável aleatória.

- É dada pelo somatório do produto entre cada valor e sua respectiva probabilidade (tal como um média ponderada).

$$\mu = E(Y) = \sum_{y} y \times p(y).$$

Variância

- Para variáveis aleatórias, a medida de dispersão mais utilizada é a **variância**.
- No caso discreto é dada pelo somatório dos desvios em relação a média ao quadrado, vezes a respectiva probabilidade.

$$\sigma^2 = E[(Y-\mu)^2] = \sum_y (y-\mu)^2 \times p(y).$$

- De forma alternativa:

$$\sigma^2 = V(Y) = E(Y^2) - \mu^2.$$

- A raiz quadrada positiva da variância é chamada de desvio padrão de Y, denotado por σ.

Exemplo

Retomando o exemplo do lançamento de 3 moedas honestas:

- Y: número de caras.
- $y = \{0,1,2,3\}$.
- Função de probabilidade:

y	0	1	2	3
Prob.	1/8	3/8	3/8	1/8
Prob.Ac.	1/8	4/8	7/8	8/8

$E(Y) = (0 \times 1/8) + (1 \times 3/8) + (2 \times 3/8) + (3 \times 1/8) = 1{,}5.$

$V(Y) = [(0 - 1{,}5)^2 \times 1/8] + [(1 - 1{,}5)^2 \times 3/8] + [(2 - 1{,}5)^2 \times 3/8] + [(3 - 1{,}5)^2 \times 1/8] = 0{,}75$

Variáveis aleatórias contínuas

Variáveis aleatórias contínuas

- Considere uma **variável aleatória contínua** Y.

- A **função densidade de probabilidade** (fdp) atribui probabilidades a intervalos de valores (a,b) da variável Y:

$$P(a < Y < b) = \int_a^b f(y)dy$$

- As propriedades da função seguem as regras de probabilidades, ou seja:
  - A função é não negativa.

$$f(y) \geq 0$$

  - A área total sob a curva é igual a 1.

$$\int_{-\infty}^{+\infty} f(y)dy = 1.$$

Variáveis aleatórias contínuas

- Note que para variáveis aleatórias contínuas **não é possível atribuir probabilidades a valores pontuais** do domínio.
- A função que descreve a probabilidade dos possíveis valores da variável é dada por uma **curva**.
- Para obter probabilidades calculamos **áreas** abaixo desta curva, e **a área abaixo de um ponto é zero**.
- Por esta razão atribuímos probabilidades a intervalos de valores.
- Portanto:
  - $P(Y = y) = 0$.
  - $P(a \leq Y \leq b) = P(a < Y \leq b) = P(a \leq Y < b) = P(a < Y < b)$.

Exemplo

Considere o experimento aleatório que consiste em verificar o surgimento de um defeito em uma pista de um trecho de rodovia com extensão de 20 km. Considere que existem razões para crer que a probabilidade da ocorrência de um defeito é constante para todo o trecho. Seja Y a variável aleatória que representa o km em que ocorreu o defeito:

a) Encontre a função densidade de probabilidade de Y.

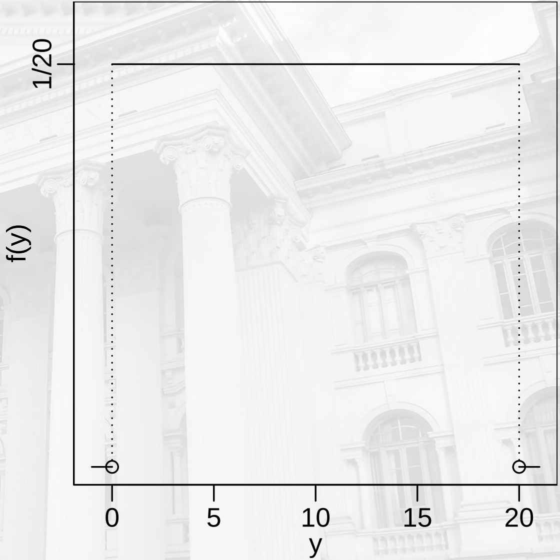
b) Verifique se as propriedades de uma função densidade de probabilidade são atendidas.

Exemplo

a) Encontre a função densidade de probabilidade de Y.

- Y: km em que ocorre o defeito.
- $y = [0{,}20]$.
- Função densidade de probabilidade:

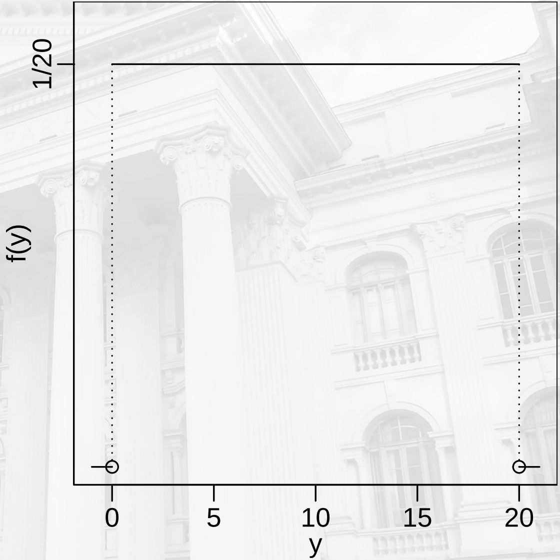
$$f(y) = \begin{cases} 1/20, & 0 < y < 20 \\ 0, & \text{caso contrário} \end{cases}$$



Exemplo

b) Verifique se as propriedades de uma função de probabilidades são atendidas.

- A função é não negativa?
  - Sim.
- A área total sob a curva é igual a 1?
  - Neste caso não se faz necessário o uso de integrais, basta calcular a área do retângulo.
  - $A = (b - a) \times f(y) = 20/(1/20) = 1.$



Função de distribuição acumulada

- Da mesma forma que no caso discreto existe a chamada função de distribuição acumulada.

- A função de distribuição acumulada $F(y)$ de uma v.a. contínua Y com densidade $f(y)$ é:

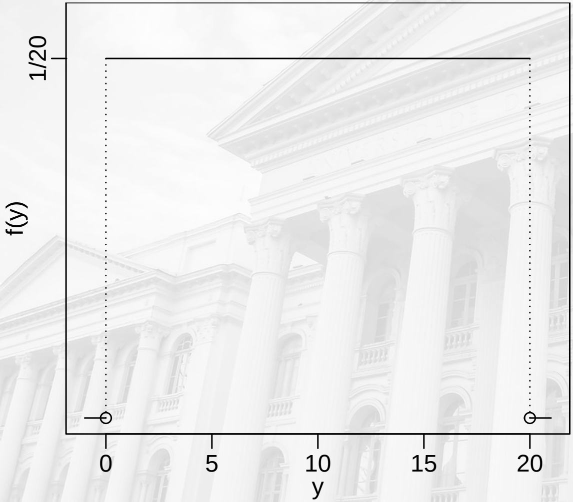
$$F(y) = P(Y \leq y) = \int_{-\infty}^{y} f(t)dt, \quad -\infty < y < +\infty.$$

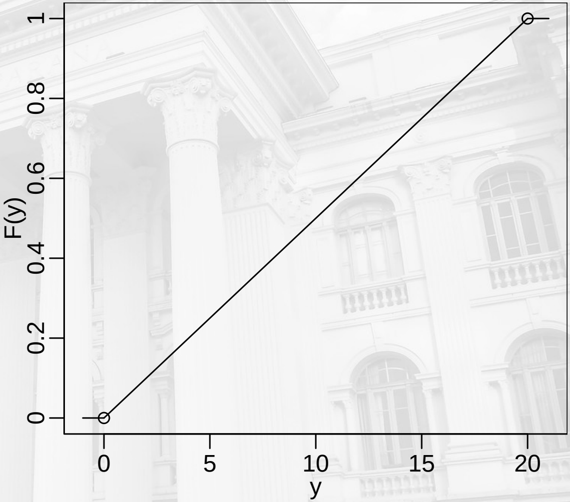
- Como consequência, temos que:

$$P(a < Y < b) = F(b) - F(a).$$

Exemplo

De forma gráfica:





Esperança de uma v.a contínua

- A média, também chamada de valor esperado ou esperança, representa o ponto de equilíbrio da distribuição dos valores da variável aleatória.

- No caso contínuo, adapta-se a expressão discreta pois passamos a usar a integral e não mais o somatório.

$$\mu = E(Y) = \int_{-\infty}^{+\infty} y f(y) dy.$$

Variância de uma v.a contínua

- A variância, no caso contínuo, é dada pela integral dos desvios em relação a média ao quadrado, vezes a função de probabilidade.

$$\sigma^2 = E[(Y-\mu)^2] = \int_{-\infty}^{+\infty} (y-\mu)^2 f(y)dy.$$

- De forma alternativa:

$$\sigma^2 = V(Y) = E(Y^2) - \mu^2.$$

- A raiz quadrada positiva da variância é chamada de desvio padrão de Y, denotado por σ.

Exemplo

Retomando o exemplo do defeito na pista:

- Y: km em que ocorre o defeito.
- $y = [0{,}20]$.
- Função de probabilidade: $f(y) = \frac{1}{20}$.

$$E(Y) = \int_{-\infty}^{+\infty} y f(y) dy = \int_a^b \frac{y}{20} dy = \left.\frac{y^2}{2 \times 20}\right|_0^{20} = 400/40 = 10$$

$$\sigma^2 = \left[\int_{-\infty}^{+\infty} y^2 f(y) dy\right] - \mu^2 = \left[\int_0^{20} \frac{y^2}{20} f(y) dy\right] - 10^2 = \left.\frac{y^3}{60}\right|_0^{20} - 100 = 133{,}33 - 100 = 33{,}33$$

O que foi visto:

- Variáveis aleatórias discretas.
- Variáveis aleatórias contínuas.
- Esperança de variáveis aleatórias.
- Variância de variáveis aleatórias.

Próximos assuntos:

- Distribuições de probabilidades.