

Universidade Federal do Rio Grande do Norte
Instituto Metr pole Digital
IMD0601 - Bioestat stica

Probabilidade

Prof. Dr. Tetsu Sakamoto
Instituto Metr pole Digital - UFRN
Sala A224, ramal 182
Email: tetsu@imd.ufrn.br



Baixe a aula (e os arquivos)

- Para aqueles que não clonaram o repositório:

```
> git clone https://github.com/tetsufmbio/IMD0601.git
```

- Para aqueles que já tem o repositório local:

```
> cd /path/to/IMD0601
```

```
> git pull
```

O que é probabilidade?

Probabilidade é um campo da matemática e da estatística que foca em tentar descrever e quantificar a chance de um evento acontecer.

Entender probabilidade nos permite analisar dados e extrair conclusões válidas cientificamente.

Dois pensamentos

- Frequentista
 - Experimentos aleatórios onde é possível repetição;
 - Probabilidade de dar cara em uma jogada de moeda;
- Bayesiana
 - Permite a incorporação da intuição do investigador ou de um conhecimento subjetivo (**probabilidade a priori**) como:
 - Informações históricas de experimentos semelhantes;
 - Informações de experimentos relacionados a este estudo;

Teoria da probabilidade

Baseado nos estudos da aleatoriedade dos resultados de um experimento.

Aleatório → O resultado do evento não é fixo, portanto, se o experimento é repetido, o resultado pode ser diferente, resultando em uma variabilidade.

Para analisar um experimento em termos de probabilidade:

- **Espaço amostral** (Ω) → Conjunto de todos os resultados possíveis de um experimento;
- **Eventos** → subconjunto do espaço amostral na qual atribuímos a probabilidade.

Operações de conjunto

Probabilidade lida com operações básicas da teoria de conjunto:

- União;
- Interseção;
- Complemento.

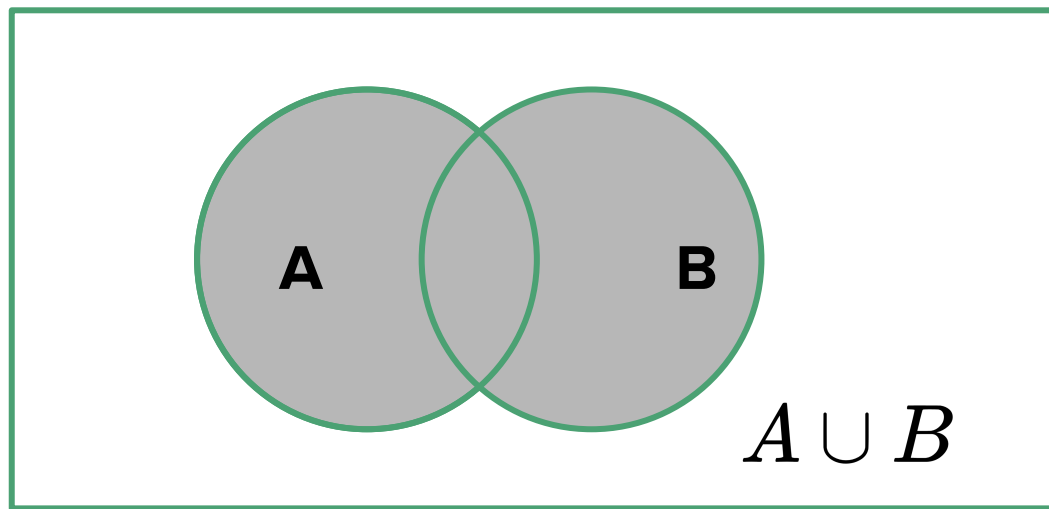
Diagrama de Venn



Operações de conjunto (União)

União de dois eventos (A e B) → Evento que contém todos os resultados de A e B.

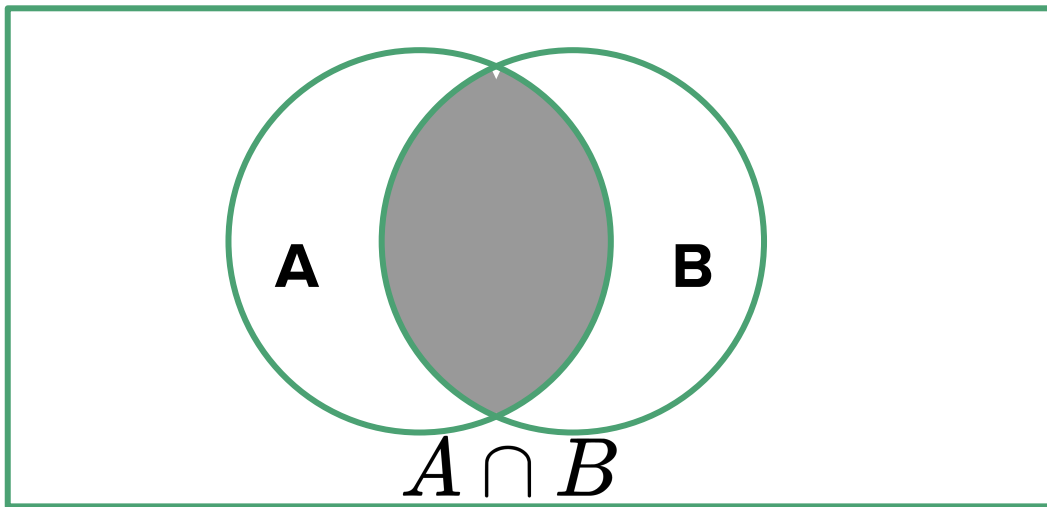
“A **ou** B ocorre”



Operações de conjunto (Interseção)

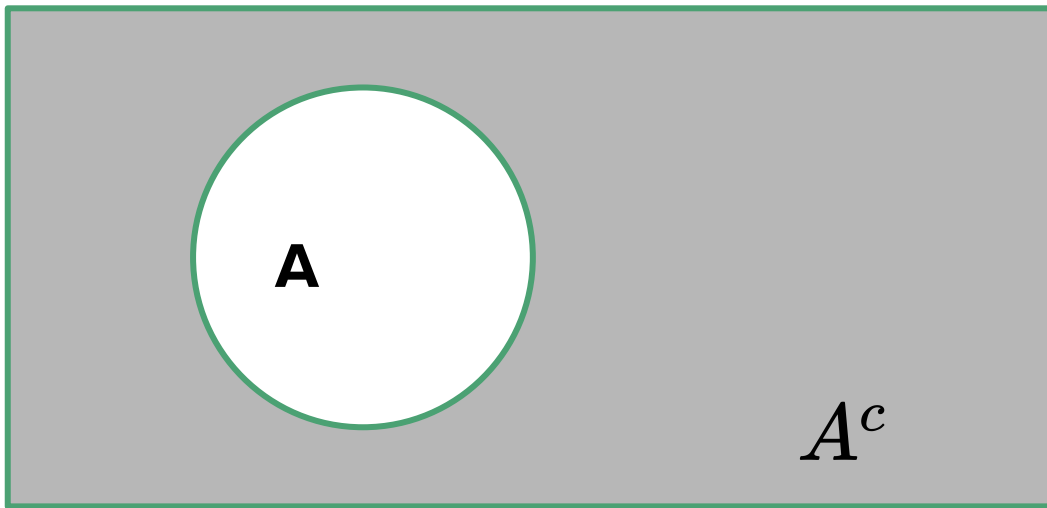
Interseção de dois eventos (A e B) → Evento que contém todos os resultados em comum de A e B.

“A e B ocorre”



Operações de conjunto (Complemento)

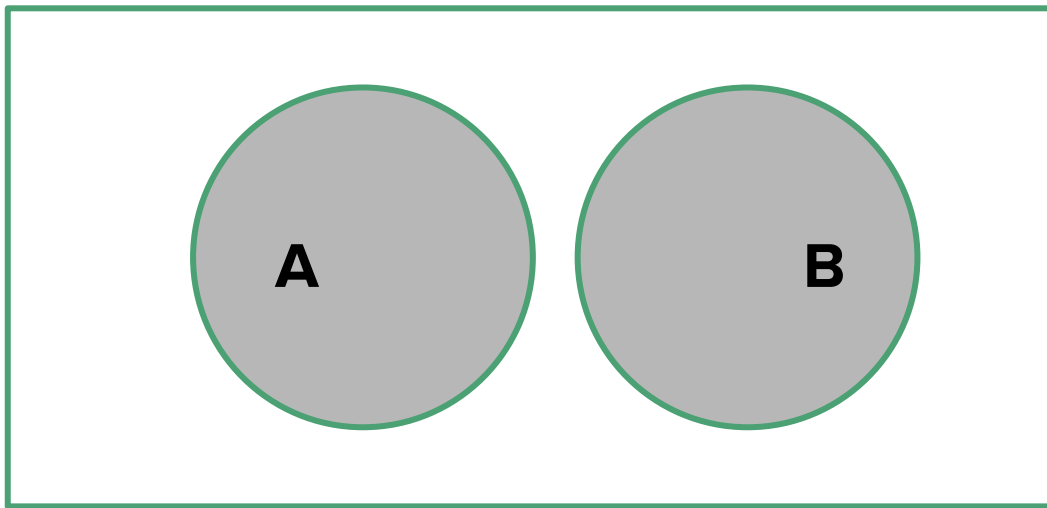
Complemento de um evento (A) → Evento em que o evento A não ocorre. Conjunto de todos os resultados no espaço amostral em que não está contido em A



Eventos disjuntos

Dois eventos (A e B) são considerados disjuntos se a sua interseção é vazia. Eventos que não possuem elementos em comum.

Eventos mutuamente exclusivos (não confundir com eventos independentes).



Axiomas da teoria de probabilidade

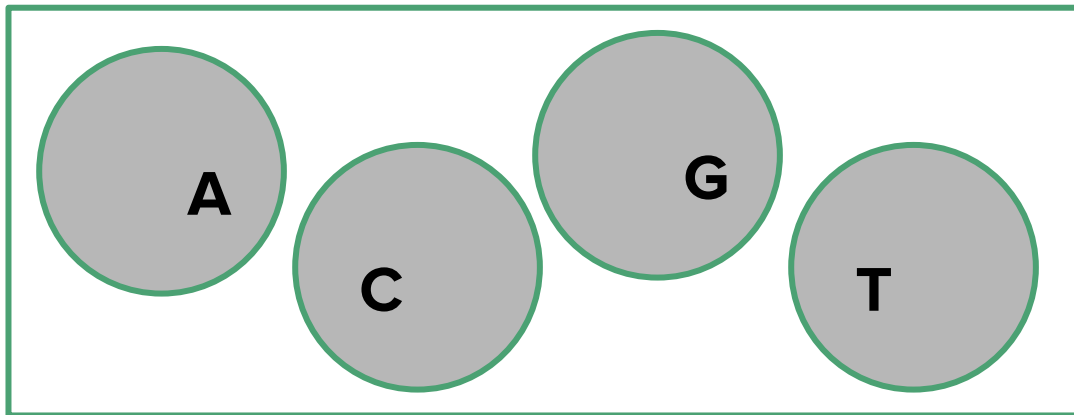
1. Probabilidade é sempre positiva

- Conceito de probabilidade negativa não existe;
- Probabilidade de $A \rightarrow P(A)$
- $P(A)$ deve ser um valor entre 0 e 1 \rightarrow medida de proporção;
- Se $P(A) = 0$, A é um evento impossível;
- Se $P(A) = 1$, A é um evento que sempre ocorre.

Axiomas da teoria de probabilidade

2. Para um dado espaço amostral, a soma das probabilidades é 1

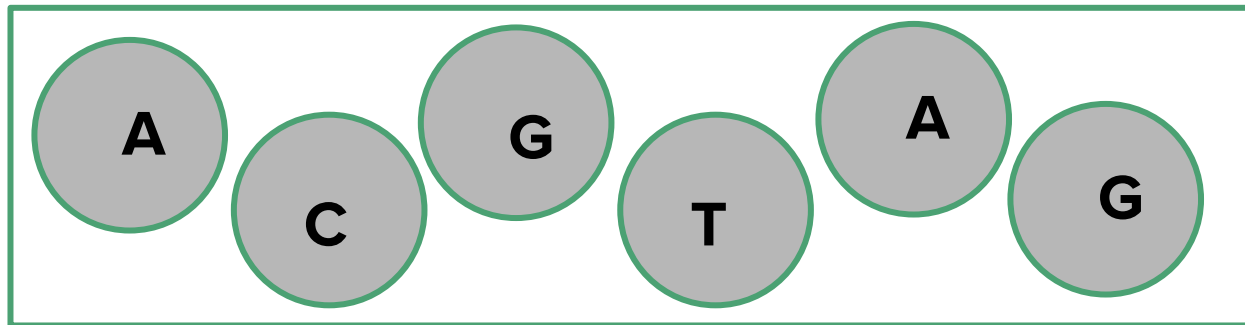
- Considere um espaço amostral constituído de 4 nucleotídeos {A, C, G, T}
- A probabilidade de ocorrer cada nucleotídeo é $\frac{1}{4}$ ($P(A)=P(C)=P(G)=P(T)$);



Axiomas da teoria de probabilidade

2. Para um dado espaço amostral, a soma das probabilidades é 1

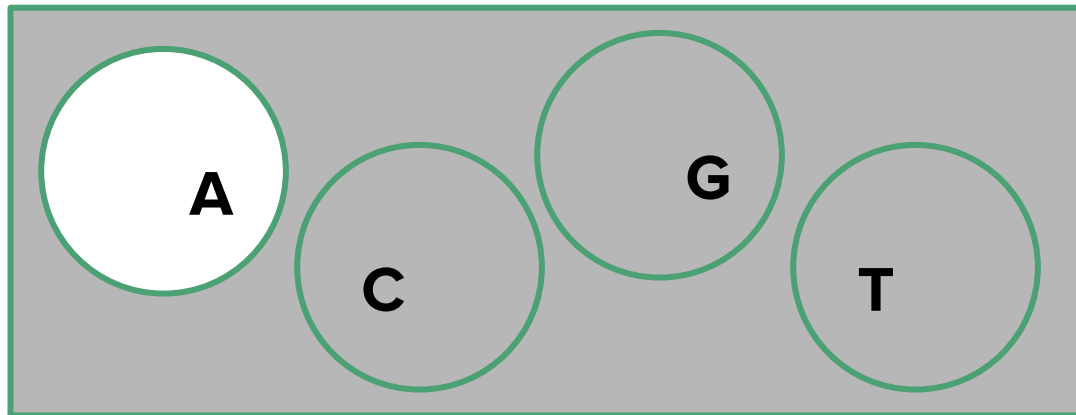
- Para cada evento em um espaço amostral que possuam a mesma probabilidade, a probabilidade dos eventos corresponde a frequência do evento no espaço amostral.



Axiomas da teoria de probabilidade

2. Para um dado espaço amostral, a soma das probabilidades é 1

- Este axioma é o que permite o cálculo da probabilidade do complemento.
- Se $P(A)$ é conhecido, $P(A^c) = 1 - P(A)$

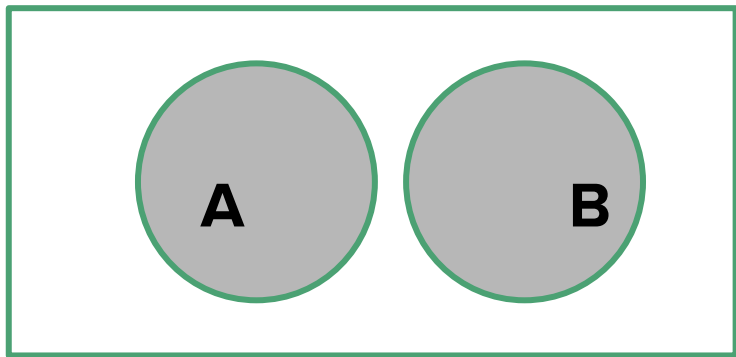


Axiomas da teoria de probabilidade

3. Para eventos mutuamente exclusivos, a $P(A \cup B) = P(A) + P(B)$

$$P(E1) \cup P(E2) \cup \dots P(E_n) = P(E1) + P(E2) + \dots + P(E_n)$$

$$= \sum_{i=1}^n P(E_i)$$



Métodos de Contagem

Em cálculos de probabilidade, você pode precisar recorrer a um desses métodos.

Princípio fundamental da contagem

Permutação e Arranjos simples

Combinação

Princípio Fundamental da Contagem

Também conhecido como Princípio da multiplicação;

Aplicado a dois ou mais experimentos independentes;

Se em dois experimentos, o número de resultados do primeiro for “m” e do segundo for “n”, o número total de possíveis resultados é $m \cdot n$.



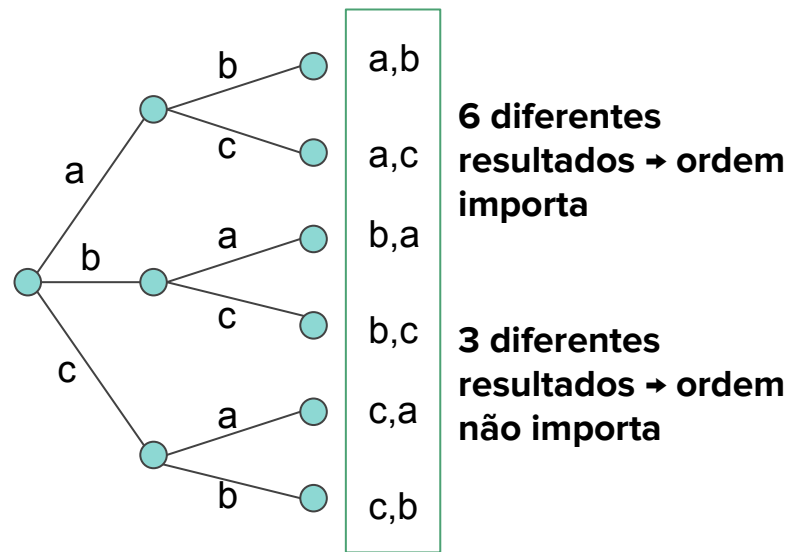
$$\text{Número de possíveis resultados} = 3 \cdot 5 = 15$$

Permutação, Arranjos e Combinação

Em ocasiões onde você está amostrando objetos distintos de um mesmo conjunto, então existem dois cenários a considerar:

- A ordem importa;
- A ordem não importa.

$L = \{a, b, c\}$, quero saber de quantas formas possíveis posso pegar duas letras:



Permutação e Arranjos

A ordem do rearranjo importa.

Permutação → Rearranjamento envolvendo todos os elementos de um conjunto;

- n = número de elementos em um conjunto.

$$P(n) = n!$$

Arranjos → Rearranjamento envolvendo parte dos elementos de um conjunto.

- n = número de elementos em um conjunto;
- r = número de elementos no arranjo.

$$A(n, r) = \frac{n!}{(n-r)!}$$

Combinação

A ordem do rearranjo não importa.

- n = número de elementos em um conjunto;
- r = número de elementos no arranjo.

$$C(n, r) = \frac{n!}{r!(n-r)!}$$

Métodos de Contagem e Probabilidade

Em um experimento onde os possíveis resultados possuem as mesmas probabilidades, a probabilidade de um evento pode ser calculado pelo número de resultados que caracteriza um evento dividido pelo número total de possíveis resultados em um espaço amostral.

Modelos probabilísticos

Modelos matemáticos que representam a relação entre as variáveis;

Representados na forma de equações matemáticas;

Um modelo simples seria a relação linear entre a variável independente x e a variável dependente y .

$$y = mx + b \qquad f(x) = mx + b$$

Onde “ m ” é o slope da linha e “ b ” o local onde a linha intercepta do eixo y .

Resultados de um experimento podem ser matematicamente modelados, fornecendo uma forma de organizar e estruturar a distribuição dos dados.

Modelos probabilísticos

Conceitos

Variáveis aleatórias

- Discretos
- Contínuos

Distribuição de probabilidade

Densidade de probabilidade

Estatística univariada X multivariada

Variáveis aleatórias

Função que mapeia o espaço amostral em números reais

- $f: \Omega \rightarrow \mathbb{R}$

Representação dos resultados na forma numérica;

Variáveis aleatórias

Considere a sequência de RNA abaixo:

AUGCUUCGAAUGCUGUAUGAUGUC

A: 5
U: 9
G: 6
C: 4
Total: 24

$X \rightarrow$ variável aleatório representando o nucleotídeo;

$x \rightarrow$ representa o resultado de uma variável aleatório;

Residue	Value of $X (=x)$	$P (X=x)$
A	0	$5/24=0.208$
C	1	$4/24=0.167$
G	2	$6/24=0.25$
U	3	$9/24=0.375$

Modelos probabilísticos simplesmente utiliza o conceito de variáveis aleatórios para representar matematicamente os resultados de um experimento.

Variáveis aleatórias discretas X contínuas

Discreta

- Conjunto dos possíveis resultados de um experimento é contável;
- Não quer dizer que o espaço amostral seja finito;
- Exemplo: número de pessoas, número de A

Contínua

- Conjunto dos possíveis resultados de um experimento não é contável;
- Uma variável aleatória que pode assumir qualquer valor, normalmente dentro de um intervalo.
- Exemplo: Características mensuráveis como altura, peso, intensidade do sinal...

Funções em modelos probabilísticos

Função de Probabilidade

- Função massa de probabilidade;
- Função densidade de probabilidade;

Função de Distribuição Acumulada (fda)

Função de Probabilidade

São funções que tentam seguir a distribuição dos dados;

Frequentemente, os dados seguem uma forma conhecida de distribuição, como a distribuição normal, gama ou beta;

Toda variável aleatória possui uma função de probabilidade associada.

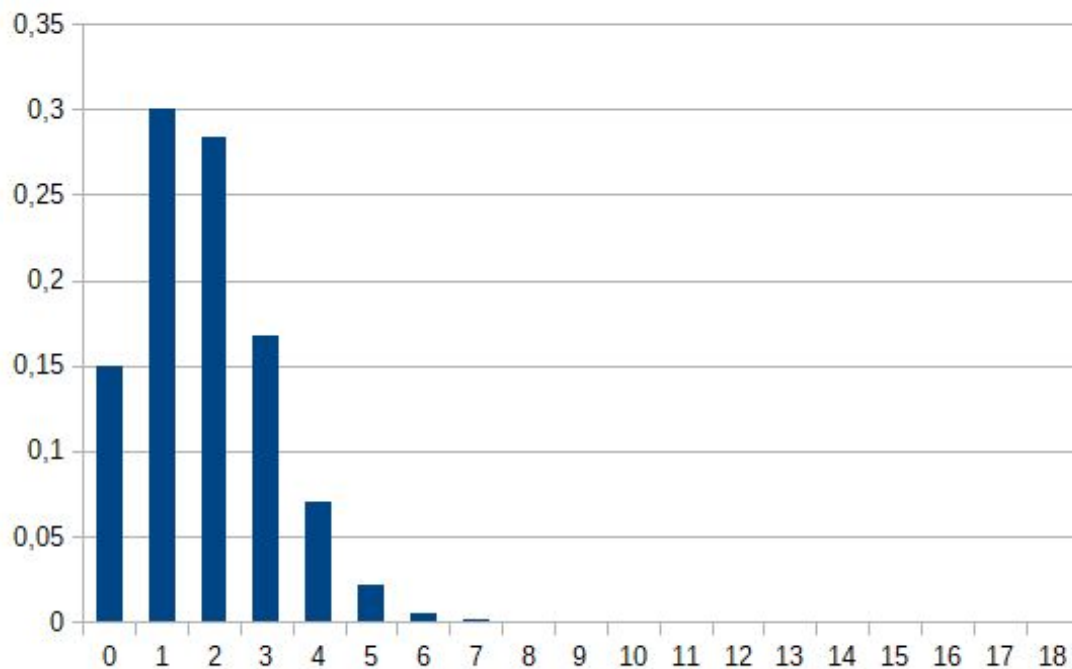
- Variáveis discretas → **função massa de probabilidade;**
- Variáveis contínuas → **função densidade de probabilidade;**

Função massa de probabilidade

- $P(x) \geq 0$;
- $\sum P(x_i) = 1$;

$$B_{n,p}(k) = \binom{n}{k} p^k q^{n-k}$$

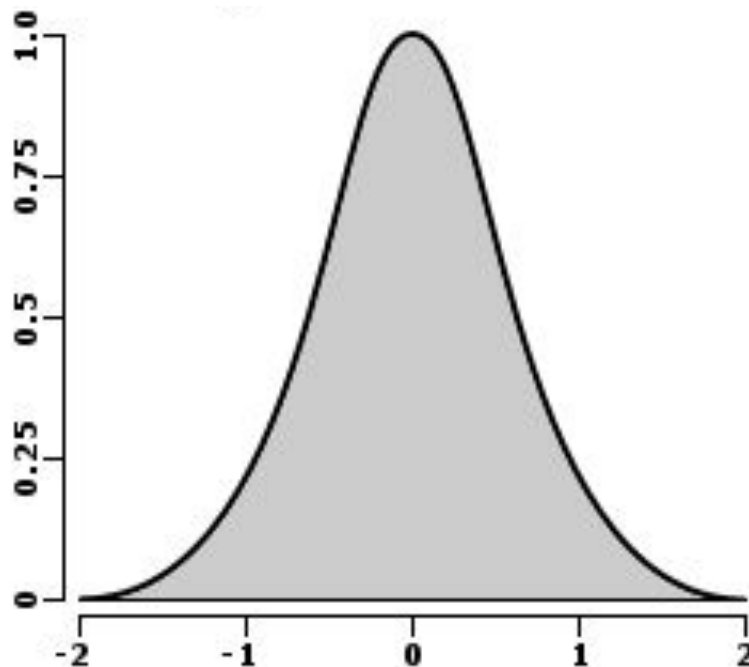
$n = 18$ e $p = 0.1$



Função densidade de probabilidade

- $f(x) \geq 0$;
- Área sob a curva = 1;

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



Função de probabilidade acumulada

Uma função que descreve a probabilidade de uma variável aleatório X se menor ou igual a x para todos os valores x .

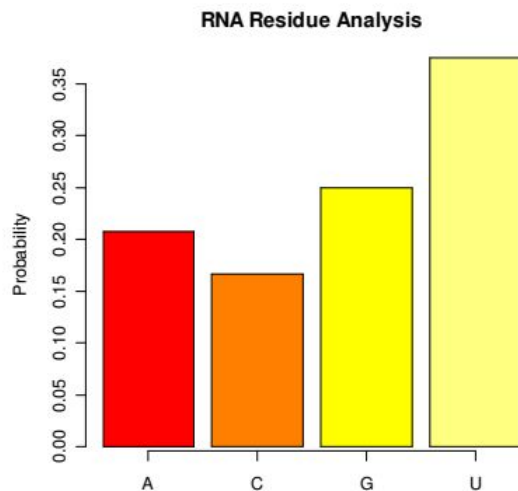
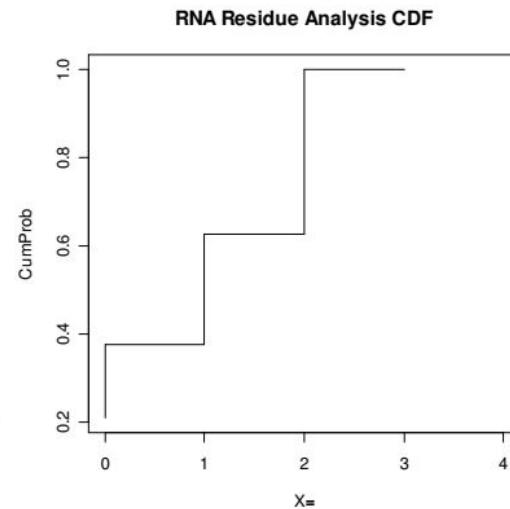


Table 6-2: Probability Distribution and Cumulative Probability Distribution for RNA Residue Analysis Example

Residue	Value of X (=x)	P (X=x)	F(x)= P(X ≤ x)
A	0	5/24=0.208	0.208
C	1	4/24=0.167	0.375
G	2	9/24=0.375	0.75
U	3	6/24=0.25	1



Função de probabilidade acumulada

Uma função que descreve a probabilidade de uma variável aleatório X se menor ou igual a x para todos os valores x .

