ACH2053 – Introdução à Estatística

Revisão Teórica:
01 – Introdução à Probabilidade

Marcelo de Souza Lauretto
EACH
www.each.usp.br/lauretto

Referência:
Morris DeGroot, Mark Schervish.
Probability and Statistics. 4th Ed. – 1º capítulo

1.2 Interpretações de probabilidade

- Não existe uma interpretação única para o termo *probabilidade* por todos os estatísticos e filósofos

- Ao longo dos anos, cada interpretação de probabilidade proposta por alguns autores tem sido criticada por outros

- O “verdadeiro” significado de probabilidade é ainda um assunto controverso

- Apresentamos a seguir três interpretações de probabilidade

1.2 Interpretações de probabilidade

- **Interpretação clássica**
  - Baseada no conceito de *resultados equiprováveis*
    - Exemplo: Quando uma moeda é lançada, há dois resultados possíveis: uma cara ou uma coroa. Se for possível assumir que esses dois resultados são igualmente possíveis de ocorrer, então devem ter a mesma probabilidade. Uma vez que a soma das probabilidades precisa ser 1, então a probabilidade de cara e a probabilidade de coroa devem ser ½.
  - Em termos mais gerais, se o resultado de algum processo precisa ser um dentre k possíveis resultados, e se esses k resultados são equiprováveis, então a probabilidade de cada resultado é $1/k$

1.2 Interpretações de probabilidade

- **Críticas à interpretação clássica:**
  - O conceito de *resultados equiprováveis* (igualmente possíveis) é essencialmente baseado no conceito de probabilidade que estamos tentando definir
    - Resultados igualmente possíveis = resultados com mesma *probabilidade*
  - Não é fornecido um método sistemático para atribuir probabilidades para eventos que não sejam assumidos como equiprováveis (e que ocorrem na maioria dos problemas reais)
    - Quando uma moeda ou um dado honesto são lançados, os diferentes resultados possíveis podem ser considerados como equiprováveis por causa da natureza do processo
    - Contudo, poucos problemas são dessa natureza:
      - Probabilidade de um indivíduo se casar dentro de 2 anos
      - Probabilidade de sucesso de uma *startup*

1.2 Interpretações de probabilidade

- **Interpretação frequentista**
  - Em diversos problemas, a probabilidade de se obter um resultado específico de um processo pode ser interpretado por meio da *frequência relativa* com que aquele resultado seria obtido se o processo fosse repetido um número grande de vezes, sob condições similares
    - Exemplo: Considera-se que a probabilidade de se obter cara quando uma moeda honesta é lançada seja ½ porque, se a moeda for lançada um grande número de vezes sob condições similares, a frequência relativa de caras deve ser de aproximadamente 50% (ou ½)

1.2 Interpretações de probabilidade

- **Críticas à interpretação frequentista:**
  - Condições vagas: O que significam as expressões:
    - “um número grande de vezes”?
    - “sob condições similares”?
  - No exemplo dado, dizemos que a frequência relativa de caras deveria ser “aproximadamente ½”, mas não há um limite especificado para a variação dessa frequência
    - Ex: em 1.000.000 de lançamentos de uma moeda, não é razoável esperar que a moeda dê *exatamente* 500.000 caras, mas também não é razoável esperar um grande desvio em relação a esse valor
    - Definição do limite de variação demanda o estabelecimento preciso das verossimilhanças das possíveis quantidades de caras, o que dependeria justamente do conceito de probabilidade que se está tentando definir → *argumento circular*

1.2 Interpretações de probabilidade

- **Interpretação subjetiva (ou pessoal)**
  - A probabilidade que uma pessoa atribui para um possível resultado representa seu próprio julgamento da possibilidade de obtenção do resultado. Outra pessoa, com outras crenças ou informações, pode atribuir uma probabilidade diferente para o mesmo resultado
  - Sob essa visão, seria apropriado mais apropriado falar na *probabilidade subjetiva* de um resultado do que na *probabilidade verdadeira* daquele resultado
  - Se os julgamentos pessoais das verossimilhanças relativas de várias combinações de resultados satisfazem certas condições de consistência, então pode-se provar que suas probabilidades subjetivas podem ser determinadas unicamente
  - Em contraste com a estatística frequentista, permite combinar novas evidências (baseados em novos experimentos) com crenças anteriores (a priori), que podem ser baseadas em resultados de experimentos passados ou em conhecimentos subjetivos
    - probabilidade *a priori* → dados → probabilidade *a posteriori*
  - Fortemente baseada na Estatística Bayesiana (embora parte da comunidade Bayesiana se oponha à interpretação subjetiva)

1.2 Interpretações de probabilidade

- **Críticas à interpretação subjetiva**
  - Não é trivial satisfazer as condições de consistência para um número infinito de exemplos
  - A interpretação subjetiva não provê uma base "objetiva" para que dois ou mais cientistas possam chegar a uma avaliação comum do estado atual do conhecimento em uma área de mesmo interesse
- **Algumas respostas às críticas**
  - A avaliação particular de um cientista a respeito da probabilidade de um resultado deve ser, em última instância, sua própria avaliação baseada em todas as evidências disponíveis.
  - A natureza subjetiva da ciência se revela em vários aspectos:
    - Na escolha do problema pelo cientista (entre a classe de problemas que poderiam ter sido escolhidos)
    - Na definição do experimento
    - Na escolha dos modelos sobre os dados
    - Nas conclusões obtidas a partir dos dados experimentais

1.3 Experimentos e eventos

- **Definição 1.3.1: Experimento e evento**
  Um **experimento** é qualquer processo, real ou hipotético, no qual os possíveis resultados podem ser identificados antecipadamente.
  Um **evento** é um conjunto bem definido de possíveis resultados do experimento.
- A definição acima é bastante ampla e comporta quase qualquer processo imaginável, sejam seus resultados previamente conhecidos ou não.
- Um tipo comum de experimento hipotético é repetir uma tarefa bem definida infinitamente sob condições similares
  - Visão frequentista
  - Simulações

1.3 Experimentos e eventos

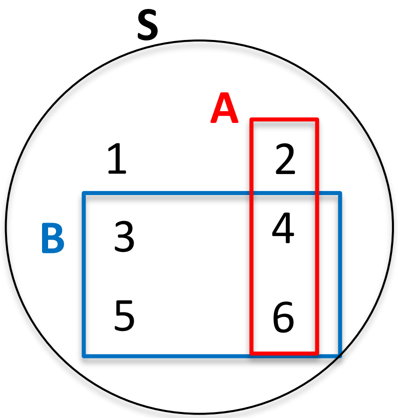
- Exemplos:
  - Em um experimento no qual uma moeda é lançada 10 vezes, queremos determinar a probabilidade de *se obter pelo menos 4 caras*
  - Em um experimento no qual uma amostra de 1000 transístores deve ser selecionada de um grande carregamento de itens similares e cada item deve ser inspecionado, queremos determinar a probabilidade de *não encontrar mais do que um transistor defeituoso*
  - Em um ensaio clínico, queremos estimar a probabilidade de um novo medicamento para uma certa doença ser mais efetivo como tratamento do que um medicamento já existente
  - Avaliando um projeto de P&D industrial em um certo momento, queremos calcular a probabilidade de que *o projeto resultará no desenvolvimento de um novo produto dentro de um determinado número de meses*

1.3 Experimentos e eventos

- Robustez da teoria da probabilidade:
  - Como mencionado anteriormente, existem controvérsias sobre o significado e interpretação do termo *probabilidade*
  - Contudo, uma vez que as probabilidades tenham sido atribuídas para os resultados de um experimento, há um consenso de que a moderna teoria da probabilidade fornece a metodologia apropriada para o estudo posterior dessas probabilidades
  - Quase todo o trabalho na teoria matemática de probabilidade tem sido relacionado a dois problemas:
    1. Métodos para determinar as probabilidades de certos eventos a partir das probabilidades especificadas para cada possível resultado de um experimento
    2. Métodos para revisar as probabilidades dos eventos quando uma nova informação relevante adicional é obtida

1.4 Teoria dos conjuntos – Espaço amostral

- **Definição 1.4.1: Espaço amostral**
  A coleção de todos os possíveis resultados de um experimento é denominada **espaço amostral** do experimento.
  - O espaço amostral de um experimento pode ser interpretado como um **conjunto**, ou coleção, de diferentes resultados possíveis
  - Cada resultado pode ser interpretado como um **ponto**, ou **elemento**, do espaço amostral
  - Eventos podem ser interpretados como subconjuntos do espaço amostral
- Exemplo: lançamento de um dado de 6 faces
  - Espaço amostral: **S = {1,2,3,4,5,6}**
  - Evento A: obtenção de um número par
    A = {2,4,6}
  - Evento B: Obtenção de um número estritamente maior do que 2 **B = {3,4,5,6}**



1.4 Teoria dos conjuntos – Espaço amostral

- Outros exemplos de espaços amostrais:
  - Se o resultado de um experimento consiste na determinação do sexo de um bebê recém-nascido, então

    S = {masculino, feminino}
  - Se o resultado de um experimento é a ordem de chegada de uma corrida entre 7 cavalos numerados de 1 a 7, então

    $$S = \{\text{todas as } 7! \text{ permutações de } (1,2,3,4,5,6,7)\}$$
  - Se o experimento consiste em jogar duas moedas independentes, então o espaço amostral é formato pelos quatro pontos a seguir:

    $$S = \{(H,H),(H,T),(T,H),(T,T)\}$$

    onde H representa cara e T representa coroa
  - Se o experimento consiste em jogar dois dados independentes, então

    $$S = \{(i,j) \mid i,j = 1,2,3,4,5,6\}$$
  - Se o experimento consiste em medir (em horas) o tempo de vida de um transistor, então o espaço amostral é formado por todos os números reais não negativos:

    $$S = \{x \ \in \mathbb{R} \mid x > 0\}$$

1.4 Teoria dos conjuntos – Espaço amostral

- Seja S o espaço amostral de um experimento. Então cada possível resultado s do experimento é dito **ser um membro** do conjunto S, ou **pertencer** a S. Notação: $s \in S$

- Dizer que um evento E ocorreu na realização de um experimento tem dois significados:
  - O resultado do experimento satisfaz às condições que especificam aquele evento E;
  - O resultado, considerado um ponto do espaço amostral, é um elemento de E.

1.4 Teoria dos conjuntos – Espaço amostral

- Para que se possa realizar todos os cálculos de probabilidade razoáveis, há três condições simples que devem ser satisfeitas pela coleção de conjuntos que denominamos eventos:
  - **Condição I:**
    O espaço amostral S deve ser um evento.
  - **Condição II:**
    Se A é um evento, então A^c também é um evento.
  - **Condição III:**
    Se $A_1, A_2, \ldots$ é uma coleção contável de eventos, então $\bigcup_{i=1}^{\infty} A_i$ também é um evento.

- Tarefa: Revisar Seção 1.4 do DeGroot – Teoria dos conjuntos

1.5 Definição de probabilidade

- Apresentaremos a seguir a definição matemática, ou axiomática, de probabilidade

- Em um dado experimento, é necessário atribuir, para cada evento *A* no espaço amostral *S*, um número Pr(*A*) indicando a probabilidade de que *A* ocorrerá

- Para se enquadrar na definição matemática de probabilidade, o número Pr(*A*) precisa satisfazer três axiomas específicos, que asseguram que Pr(*A*) terá certas propriedades intuitivas esperadas que uma probabilidade tenha, sob cada uma das várias interpretações descritas na Seção 1.2

1.5 Definição de probabilidade

- **Axioma I:** Para todo evento A, $\Pr(A) \geq 0$.
A probabilidade do resultado do experimento ser um elemento de *A* deve ser um número não negativo
- **Axioma II:** $\Pr(S) = 1$.
Com probabilidade 1, o resultado do experimento será um ponto do espaço amostral S.
- **Axioma III:** Para toda sequência infinita de eventos mutuamente exclusivos (disjuntos) $A_1, A_2, \ldots$,

$$\Pr\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \Pr(A_i)\,.$$

Se dois eventos são disjuntos, é natural assumir que a probabilidade de um ou outro ocorrer seja a soma das probabilidades de cada um. O mesmo deve valer para uma união finita ou mesmo infinita de eventos disjuntos.

1.5 Definição de probabilidade

- Exemplos:
  - No experimento de lançar um dado não viciado, para cada subconjunto A de S = {1,2,3,4,5,6}, seja Pr(A) o número de elementos de A dividido por 6.
    - É trivial notar que essa atribuição satisfaz os dois primeiros axiomas.
    - Há um número finito de coleções de conjuntos não vazios disjuntos. É trivial ver que a probabilidade da sequência finita de eventos disjuntos não vazios é igual à soma das probabilidades dos eventos. Para estender para sequencias infinitas (axioma 3) basta usar o resultado (demonstrado adiante) $\Pr(\emptyset) = 0$.
  - O mesmo valeria para um dado viciado: Suponha que o lado 6 seja duas vezes mais provável do que cada um dos demais. Podemos estabelecer $p_i = 1/7$ para i=1,2,3,4,5 e $p_6 = 2/7$. Então, para todo evento A, definimos $\Pr(A) = \sum_{i \in A} p_i$.
    - Ex: Se A = {1,3,5}, então Pr(A) = $p_1 + p_3 + p_5$ = 3/7.
    - Essa definição também satisfaz os três axiomas

1.5 Definição de probabilidade

- A definição matemática de probabilidade é baseada nos três axiomas apresentados:

- **Definição 1.5.1:**
Uma **medida de probabilidade**, ou simplesmente **probabilidade**, em um espaço amostral S é uma especificação de números $\Pr(A)$ para todos os eventos A que satisfazem os Axiomas 1, 2 e 3.

- Apresentamos a seguir alguns resultados importantes derivados dos axiomas.

1.5 Definição de probabilidade - Resultados

- **Teorema 1.5.1:** $\Pr(\emptyset) = 0.$
  - Dem: Considere a sequência infinita $A_1, A_2, \ldots$ tal que $A_i = \emptyset$ para i=1,2,...
    Note que os eventos são todos disjuntos e que $\bigcup_{i=1}^{\infty} A_i = \emptyset$. Portanto, segue do Axioma 3 que:

$$\Pr(\emptyset) = \Pr\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \Pr(A_i) = \sum_{i=1}^{\infty} \Pr(\emptyset).$$

O único número real que satisfaz $\Pr(\emptyset) = \sum_{i=1}^{\infty} \Pr(\emptyset)$ é $\Pr(\emptyset) = 0.$

1.5 Definição de probabilidade - Resultados

- **Teorema 1.5.2:** Para qualquer sequência finita de *n* eventos disjuntos $A_1, A_2, \dots, A_n$,

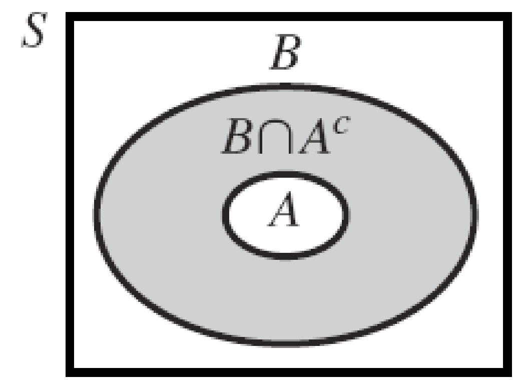
$$\Pr\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \Pr(A_i).$$

  - Dem: Considere a sequência infinita de eventos $A_1, A_2, \dots$, na qual $A_1, A_2, \dots, A_n$ são os *n* eventos disjuntos dados, e $A_i = \emptyset$ para *i* > *n*. Os eventos nessa sequência são claramente disjuntos; além disso, $\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{n} A_i$. Portanto, pelo Axioma 3, temos: →

$$\begin{aligned}
\Pr\left(\bigcup_{i=1}^{n} A_i\right) &= \Pr\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \Pr(A_i) \\
&= \sum_{i=1}^{n} \Pr(A_i) + \sum_{i=n+1}^{\infty} \Pr(A_i) \\
&= \sum_{i=1}^{n} \Pr(A_i) + 0 \\
&= \sum_{i=1}^{n} \Pr(A_i).
\end{aligned}$$

1.5 Definição de probabilidade - Resultados

- **Teorema 1.5.3:** Para todo evento A, $\Pr(A^c) = 1 - \Pr(A)$.
  - Dem: Como A e A^c são disjuntos e $A \cup A^c = S$, segue do Teorema 1.5.2 que $Pr(S) = \Pr(A) + \Pr(A^c)$.
    Como $Pr(S) = 1$ pelo Axioma 2, segue que $\Pr(A^c) = 1 - \Pr(A)$.
- **Teorema 1.5.4:** Se $A \subseteq B$, então $\Pr(A) \leq \Pr(B)$.
  - Dem: Como $A \subseteq B$, B pode ser tratado como a união de dois eventos disjuntos A e $B \cap A^c$ (ver diagrama ao lado).
    Logo, $\Pr(B) = \Pr(A) + \Pr(B \cap A^c)$.
    Como $\Pr(B \cap A^c) \geq 0$ (Axioma 1),
    então $\Pr(B) \geq \Pr(A)$.



- **Teorema 1.5.5:** Para todo evento A, $0 \leq \Pr(A) \leq 1$.
  - Dem: Pelo Axioma 1, $\Pr(A) \geq 0$; como todo evento $A \subseteq S$, o Teorema 1.5.4 e o Axioma 2 implicam que $\Pr(A) \leq \Pr(S) = 1$.

1.5 Definição de probabilidade - Resultados

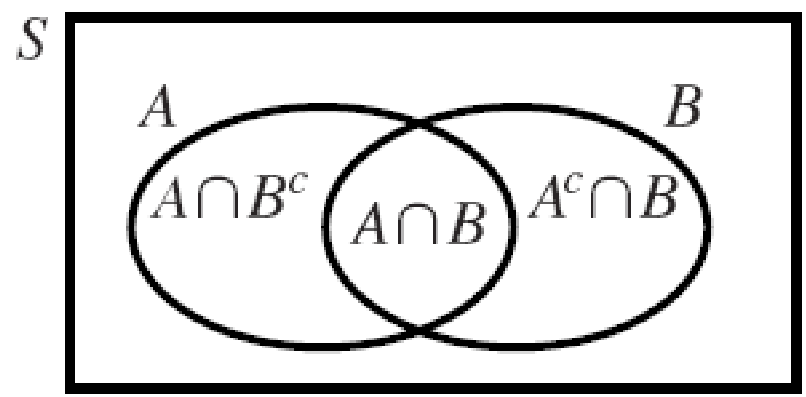
- Para os próximos teoremas, usamos o resultado abaixo:
- Teorema 1.4.11:

Partitioning a Set. For every two sets A and B, $A \cap B$ and $A \cap B^c$ are disjoint and

$$A = (A \cap B) \cup (A \cap B^c).$$

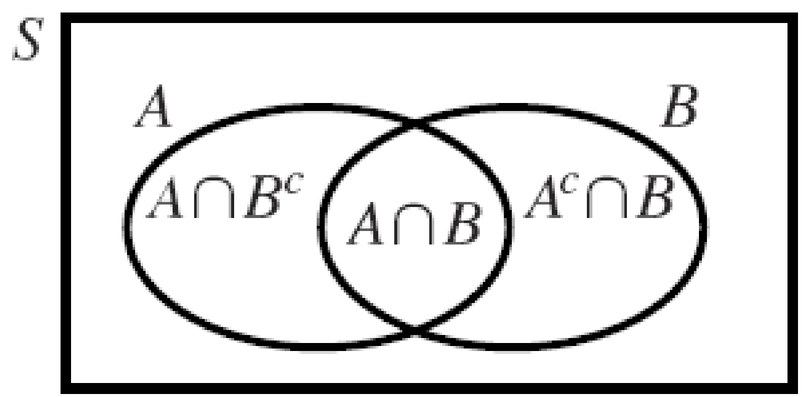
In addition, B and $A \cap B^c$ are disjoint, and

$$A \cup B = B \cup (A \cap B^c). \qquad \blacksquare$$



1.5 Definição de probabilidade - Resultados

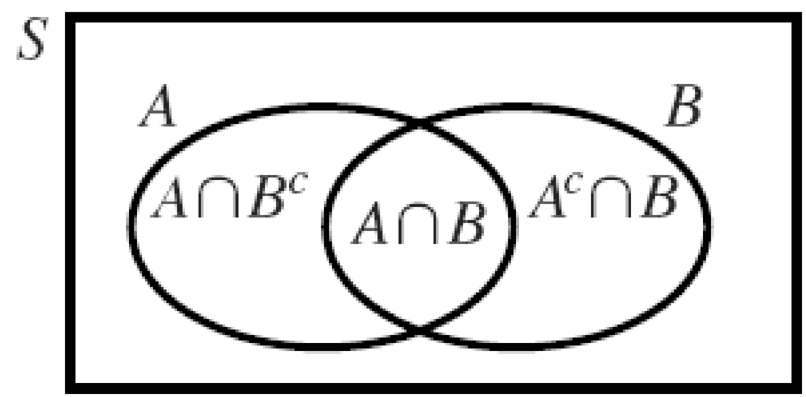
- **Teorema 1.5.6:** Para quaisquer dois eventos A e B, $\Pr(A \cap B^c) = \Pr(A) - \Pr(A \cap B)$.
  Dem: Pelo Teorema 1.4.11, os eventos $A \cap B^c$ e $A \cap B$ são disjuntos, e $A = (A \cap B) \cup (A \cap B^c)$ (ver diagrama abaixo).
  Segue do Teorema 1.5.2 que:
  $$\Pr(A) = \Pr(A \cap B) + \Pr(A \cap B^c).$$
  Subtraindo-se $\Pr(A \cap B)$ dos dois lados da última equação, obtemos
  $$\Pr(A \cap B^c) = \Pr(A) - \Pr(A \cap B).$$



1.5 Definição de probabilidade - Resultados

- **Teorema 1.5.7:** Para quaisquer dois eventos A e B,
  $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$.
  Dem: Pelo Teorema 1.4.11, temos $A \cup B = B \cup (A \cap B^c)$,
  e os dois eventos do lado direito dessa equação são disjuntos.
  Logo, temos

$$\begin{aligned} \Pr(A \cup B) &= \Pr(B) + \boxed{\Pr(A \cap B^c)} && \text{(Teor. 1.5.2)} \\ &= \Pr(B) + \boxed{\Pr(A) - \Pr(A \cap B)} && \text{(Teor. 1.5.6)} \end{aligned}$$



1.5 Definição de probabilidade - Resultados

- Exemplo: Considere a seguinte situação hipotética*:
  - Imagine uma mulher chamada Linda, de 31 anos de idade, solteira, sincera e muito inteligente. Cursou filosofia na universidade. Quando estudante, preocupava-se profundamente com discriminação e justiça social e participou de protestos contra as armas nucleares.
  - Questão: Classifique cada uma das afirmações abaixo, em uma escala de 1 (altamente improvável) a 5 (altamente provável):
    - Linda participa de um coletivo feminista.
    - Linda é bancária e participa de um coletivo feminista.
    - Linda é bancária.

*Adaptado de:
Leonard Mlodinow. O Andar do Bêbado: Como o Acaso Determina Nossas Vidas. Ed. Jorge Zahar, 2009.

1.5 Definição de probabilidade - Resultados

- Exemplo: Considere a seguinte situação hipotética*:
  - Imagine uma mulher chamada Linda, de 31 anos de idade, solteira, sincera e muito inteligente. Cursou filosofia na universidade. Quando estudante, preocupava-se profundamente com discriminação e justiça social e participou de protestos contra as armas nucleares.
  - Questão: Classifique cada uma das afirmações abaixo, em uma escala de 1 (altamente improvável) a 5 (altamente provável):
    - Linda participa de um coletivo feminista.
    - Linda é bancária e participa de um coletivo feminista.
    - Linda é bancária.
  - Agora verifique se sua classificação respeita o Teorema 1.5.4 (Se $A \subseteq B$, então $\Pr(A) \leq \Pr(B)$.)

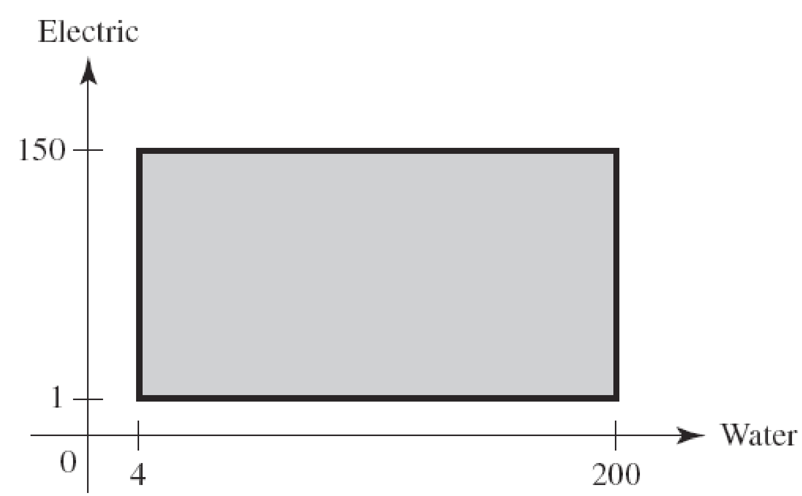
*Adaptado de:
Leonard Mlodinow. O Andar do Bêbado: Como o Acaso Determina Nossas Vidas. Ed. Jorge Zahar, 2009.

1.5 Definição de probabilidade - Resultados

- Exemplo: Diagnóstico médico
  - Um paciente chega ao consultório médico com uma inflamação na garganta e febre baixa. Depois de um exame, o médico decide que o paciente tem uma <u>infecção bacteriana</u> ou uma <u>infecção viral</u> ou <u>ambas</u>. O médico atribui probabilidade 0,7 para a infecção bacteriana e 0,4 para a infecção viral.
    Qual a probabilidade do paciente ter as duas formas de infecção?
  - Resp: Seja B o evento de que o paciente tem uma infecção bacteriana, e seja V o evento de que o paciente tem uma infecção viral. Sabemos que $\Pr(B) = 0{,}7$ e $\Pr(V) = 0{,}4$. Devemos calcular $\Pr(B \cap V)$.
    Usamos o Teorema 1.5.7, que afirma que
    $$\Pr(B \cup V) = \Pr(B) + \Pr(V) - \Pr(B \cap V)\,.$$
    Como $S = B \cup V$, o lado esquerdo dessa equação é 1, e portanto
    $$1 = 0{,}7 + 0{,}4 - \Pr(B \cap V)$$
    o que leva a $\Pr(B \cap V) = 0{,}1$.
    A probabilidade do paciente ter **somente a infecção bacteriana** é
    $$\Pr(B \cap V^c) = \Pr(B) - \Pr(B \cap V) = 0{,}6.$$

1.5 Definição de probabilidade - Resultados

- Exemplo: Dimensionamento elétrico e hidráulico
  - Um empreiteiro está construindo um complexo de escritórios e precisa planejar a demanda de água e de eletricidade, a fim de especificar as dimensões dos cabos, conduítes e cabos elétricos. Após consultar potenciais inquilinos e examinar dados históricos, o empreiteiro considera que a demanda por eletricidade variará entre 1 milhão e 150 milhões de KWH por dia, e a demanda de água será entre 4 e 200 mil litros/dia. Todas as combinações de demanda elétrica e de água são consideradas possíveis.
  - O espaço amostral será o conjunto de pares ordenados $\{(x, y) : 4 \leq x \leq 200 \text{ e } 1 \leq y \leq 150\}$.
  - Há muitas maneiras possíveis de distribuir a probabilidade no espaço amostral



1.5 Definição de probabilidade - Resultados

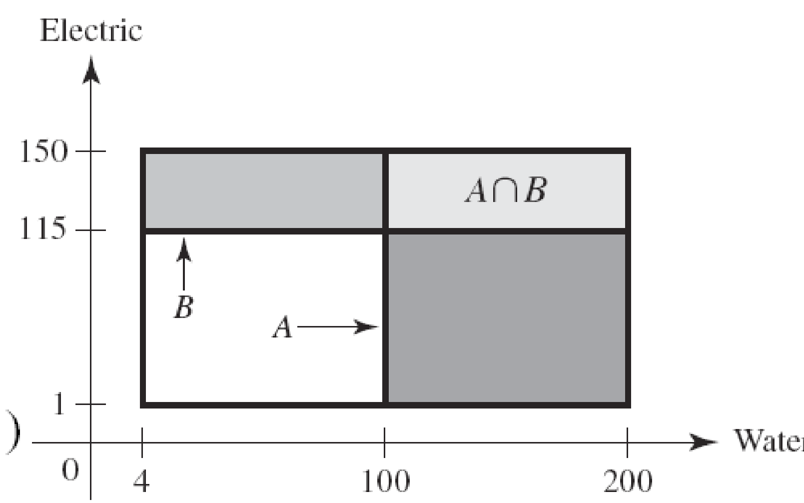
- Exemplo: Dimensionamento elétrico e hidráulico (cont.)
  - Uma escolha simples é fixar a probabilidade de um evento E proporcional à área de E.
    A área de S é $(150 - 1) \times (200 - 4) = 29.204$,
    e portanto $\Pr(E)$ é igual à área de E dividida por 29.204.
  - Seja A o conjunto onde a demanda de água seja de pelo menos 100, e seja B o evento de que a demanda elétrica seja de pelo menos 115, e suponha que esses valores sejam considerados uma alta demanda.

$$\Pr(A) = \frac{14,900}{29,204} = 0.5102$$

$$\Pr(B) = \frac{6,860}{29,204} = 0.2349$$

$$\Pr(A \cap B) = 3,500/29,204 = 0.1198$$

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B) = 0.6253$$



1.5 Definição de probabilidade - Resultados

- **Observação importante:**
  probabilidade zero ≠ impossível
  - Quando um evento possui probabilidade 0, não significa que o evento seja impossível. No exemplo anterior, há incontáveis eventos com probabilidade 0, mas que não são impossíveis.
  - Por exemplo, para todo *x*, o evento de que a demanda de água seja *x* corresponde a um segmento de linha vertical no espaço amostral *S*. Uma vez que segmentos de linha têm área 0, sua probabilidade é 0, embora o evento não seja impossível de ocorrer.
    (De fato, se todos os eventos na forma {demanda de água = *x*} fossem impossíveis, então a demanda de água não poderia assumir nenhum valor)
  - Outra forma de entender: Dada uma constante $\varepsilon > 0$, o evento
    $$\{\text{demanda de água está entre } x - \varepsilon \text{ } e \text{ } x + \varepsilon\}$$
    terá probabilidade positiva, que convergirá para 0 à medida em que ε vai para 0.

1.6 Espaços amostrais finitos

- Os experimentos mais simples para determinação e derivação de probabilidades são aqueles que envolvem apenas uma quantidade finita de resultados.
- Nesta seção, consideraremos experimentos para os quais existe apenas um número finito de possíveis resultados.
  - Espaço amostral *S* possui apenas um número finito de pontos $s_1, s_2, \dots, s_n$.
- Em um experimento desse tipo, uma medida de probabilidade em *S* é especificada atribuindo-se uma probabilidade p_i para cada ponto $s_i \in S$.
  - $p_i =$ probabilidade do resultado do experimento ser s_i $(i = 1, \dots, n)$
- A fim de satisfazer os axiomas de probabilidade, os números $p_1, p_2, \dots, p_n$ precisam atender às duas condições abaixo:

$$p_i \geq 0 \quad \text{for } i = 1, \dots, n \qquad \text{e} \qquad \sum_{i=1}^{n} p_i = 1.$$

1.6 Espaços amostrais finitos

- Exemplo: Rupturas de fibras
  - Considere um experimento no qual 5 fibras com diferentes comprimentos são submetidas a um teste para aprender qual fibra se romperá primeiro. Suponha também que:
    - O comprimentos das fibras são 1,2,3,4 e 5 polegadas;
    - A probabilidade de uma fibra ser a primeira a se romper é proporcional ao comprimento daquela fibra.
  - Devemos determinar a probabilidade de que o comprimento da fibra que se romperá primeiro não seja superior a 3".
  - Resp: Denotaremos por s_i o resultado no qual a fibra de comprimento i é a primeira a se romper ($i = 1, \dots, 5$). Então $\mathrm{S} = \{s_1, s_2, s_3, s_4, s_5\}$ e $p_i = \alpha i$ para $i = 1, \dots, 5$, onde α é um fator de proporcionalidade (usualmente chamado também *constante de normalização*). Como é necessário que $p_1 + \cdots + p_5 = 1$, e sabemos que $p_1 + \cdots + p_5 = 15\alpha$, então $\alpha = 1/15$. Se A é o evento de que o comprimento da fibra a quebrar primeiro não seja maior do que 2 polegadas, então $\mathrm{A} = \{s_1, s_2, s_3\}$. Portanto,

$$\Pr(A) = p_1 + p_2 + p_3 = \frac{1}{15} + \frac{2}{15} + \frac{3}{15} = \frac{2}{5}.$$

1.6 Espaços amostrais finitos

- **Definição: Espaços amostrais simples**
  Um espaço amostral finito *S* contendo *n* resultados $s_1, s_2, \ldots, s_n$ é denominado **espaço amostral simples** se a probabilidade atribuída para cada um dos resultados é 1/n. Se um evento *A* nesse espaço amostral possui exatamente *m* resultados, então

$$\Pr(A) = \frac{m}{n}.$$

1.6 Espaços amostrais finitos

- Exemplo: Lançamento de dois dados
  - Consideraremos o experimento de lançar dois dados não viciados, e devemos calcular a probabilidade de cada um dos possíveis valores das somas dos dois números que podem aparecer.
  - Para simplificar, podemos assumir que os dois dados são distinguíveis, e portanto cada resultado pode ser representado por um par (*x*,*y*), onde *x* é o número que aparece no 1º dado e y é o número que aparece no 2º dado. O espaço amostral S e as probabilidades das somas são, respectivamente:

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

$$P_2 = P_{12} = \frac{1}{36}, \qquad P_5 = P_9 = \frac{4}{36},$$

$$P_3 = P_{11} = \frac{2}{36}, \qquad P_6 = P_8 = \frac{5}{36},$$

$$P_4 = P_{10} = \frac{3}{36}, \qquad P_7 = \frac{6}{36}.$$

1.7 Métodos de contagem

- Vimos anteriormente que, para um **espaço amostral simples** *S* (que tem número finito de resultados equiprováveis), a probabilidade de um evento *A* é a razão entre o número de resultados em *A* e o número de resultados em *S*.

- Em muitos experimentos, a quantidades de resultados em S é tão grande que sua enumeração exaustiva pode ser muito cara, lenta ou terá alta probabilidade de estar incorreta.

- Existem métodos comuns para contagem do número de elementos em conjuntos, que exploram a estrutura existente em muitos experimentos: cada resultado consiste de diversas partes e é relativamente fácil calcular quantas possibilidades existem para cada uma das partes.

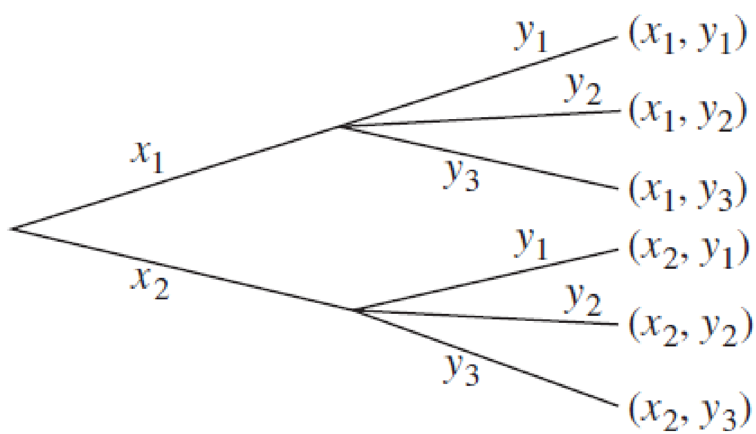
1.7 Métodos de contagem – Regra da multiplicação

- **Teorema 1.7.1 – Regra da multiplicação para experimentos de duas partes:**
  Suponha que um experimento deve ser realizado em duas partes, de tal forma que a 1ª parte do experimento pode resultar em qualquer um dos m resultados possíveis $x_1, x_2, \dots, x_m$ e, para cada resultado da 1ª parte, há n resultados possíveis $y_1, y_2, \dots, y_n$ para a 2ª parte. Então, o espaço amostral S possui exatamente mn resultados.
  - Dem: Essa regra pode ser provada enumerando-se todos os possíveis resultados das duas partes do experimentos como na tabela abaixo. Como a tabela possui m linhas e n colunas, o resultado é imediato.

$$
\begin{array}{cccc}
(x_1, y_1) & (x_1, y_2) & \cdots & (x_1, y_n) \\
(x_2, y_1) & (x_2, y_2) & \cdots & (x_2, y_n) \\
\vdots & \vdots & & \vdots \\
(x_m, y_1) & (x_m, y_2) & \cdots & (x_m, y_n)
\end{array}
$$

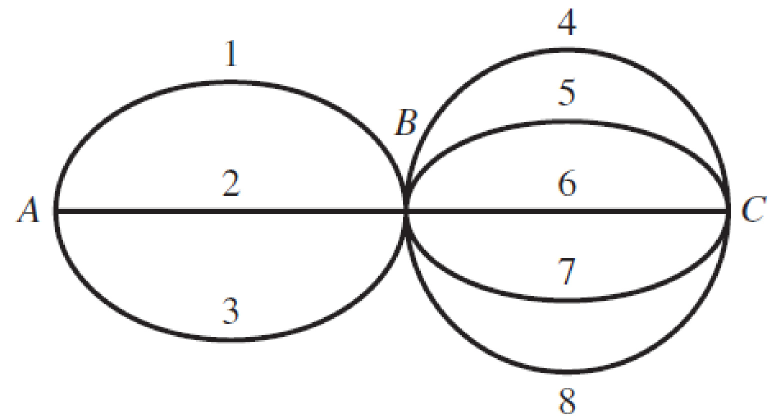
1.7 Métodos de contagem – Regra da multiplicação

- Outra forma conveniente de visualizar o resultado do Teorema 1.7.1 é através de um diagrama de árvore, conforme ilustrado na figura abaixo (m=2, n=3). Cada nó terminal (folha) representa um resultado do experimento, e as arestas são rotuladas pelos resultados obtidos em cada parte do experimento.



1.7 Métodos de contagem – Regra da multiplicação

- Exemplo 1: Rotas entre cidades
  - Suponha que haja 3 rotas diferentes da cidade A para a cidade B e que haja 5 rotas diferentes da cidade B para a cidade C, e queremos calcular a quantidade de possíveis rotas de A para C.
  - Resp: Representando as cidades e rotas de acordo com a figura ao lado, e numerando-se as rotas de 1 a 8, podemos representar cada rota de A para C pelo par (x,y), onde x denota o número da rota de A para B e y denota o número da rota de B para C. É imediato perceber que o número total de rotas de A para C é $3 \times 5 = 15$.



- Exemplo 2: Lançamento de 2 dados:
  - No lançamento independente de 2 dados, é imediato que o número de possíveis resultados é $6 \times 6 = 36$.

1.7 Métodos de contagem – Regra da multiplicação

- **Teorema 1.7.2 – Regra da multiplicação:**
  Suponha que um experimento tem k partes $(k \geq 2)$, de forma que a i-ésima parte possui n_i possíveis resultados $(i = 1,2,\ldots,k)$, e que todos os resultados de cada parte podem ocorrer independentemente dos resultados que tenham ocorrido nas demais partes.
  Então o espaço amostral S do experimento conterá todos os vetores da forma $(u_1, u_2, \ldots, u_k)$, onde u_i é um dos possíveis resultados da parte i, $(i = 1,2,\ldots,k)$.
  O número total de resultados desses possíveis vetores será igual a $n_1 n_2 \ldots n_k$.

1.7 Métodos de contagem – Regra da multiplicação

- Exemplo: Lançamento de várias moedas
  Suponha o lançamento de 6 moedas. Cada resultado em S consiste na sequência de 6 caras e coroas, p.ex. HTTHHH. Como há dois resultados possíveis para cada moeda, o número total de resultados em S é $2^6 = 64$.
  Se cara e coroa são consideradas equiprováveis, então S é um espaço amostral simples.
  - A probabilidade de obtenção de caras nos 6 lançamentos é 1/64, já que há apenas 1 resultado com essa característica.
  - A probabilidade de obtenção de uma cara (e cinco coroas) é 6/64 = 3/32, já que existem 6 resultados possíveis.

- Exemplo: Trava de cofre: cofres tradicionais usualmente possuem um disco com 40 números de 0 a 39. A combinação consiste em uma sequência de três números (que precisam ser escolhidos na ordem correta). Cada número pode aparecer em cada uma das três posições da combinação, e portanto o número total de combinações é $40^3 = 64.000$.

1.7 Métodos de contagem – Regra da multiplicação

- ***Amostragem com reposição:***
  Considere uma urna com *n* bolas numeradas $1, \dots, \mathrm{n}$.
  - Primeiro, uma bola é sorteada na urna, e seu número anotado. Essa bola é então colocada de volta na urna;
    Esse processo se repete, com um total de *k* bolas sorteadas da mesma maneira; cada resultado consiste na sequência de *k* bolas, na ordem em que foram selecionadas.
  - Um processo desse tipo é denominado ***amostragem com reposição***.
  - O espaço amostral S desse experimento é o conjunto de todos os vetores $(x_1, \dots, x_k)$, onde x_i é o número da i-ésima bola sorteada $(i = 1, \dots, k)$.
  - Como há *n* possibilidades de sorteio para cada uma das *k* bolas, o número total de vetores em *S* é n^k.
  - Além disso, assumindo-se que cada uma das bolas é igualmente provável de ser sorteada em cada passo, e que todos os sorteios são independentes dos demais, a probabilidade de cada vetor em S é $1/n^k$.

- Obs: Lançamentos sucessivos de moedas e de dados honestos podem ser vistos como problemas de amostragem com reposição

1.7 Métodos de contagem – Regra da multiplicação

- Exemplo: Amostragem com reposição:
  Considere um experimento no qual:
  - uma carta é sorteada de um baralho de *n* cartas diferentes, e devolvida ao baralho;
  - uma segunda carta é então sorteada do baralho e também devolvida;
  - uma terceira carta é sorteada.
  - Cada resultado consiste nas 3 cartas na ordem selecionada.
  - O número de possíveis sequências de cartas selecionadas é n^3, e a probabilidade de cada sequência de cartas é $1/n^3$.

1.7 Métodos de contagem – Permutações e arranjos

- **Definição 1.7.1: Arranjos**
  Suponha que um conjunto tenha *n* elementos. Suponha que um experimento consista em selecionar *k* dos elementos, um de cada vez sem reposição.
  Seja cada resultado a sequência dos *k* elementos na ordem em que foram selecionados.
  Cada resultado desse experimento é denominado um **arranjo de *n* elementos tomados *k* a *k*.** Denotamos o número de tais permutações pelo símbolo $P_{n,k}$.

- **Obs:** No livro do DeGroot, esse termo é definido como **permutação de *n* elementos tomados *k* a *k***.
  Nos slides que seguem, usaremos a terminologia usual em português: **permutação** denota cada uma das sequências possíveis dos *n* elementos originais. Quando se escolhem *k* elementos ($k \leq n$) dos n originais, tal sequência formada é denominada **arranjo de *n* elementos tomados *k* a *k***.

1.7 Métodos de contagem – Permutações e arranjos

- **Teorema 1.7.3: Número de arranjos.**
  O número de arranjos de n elementos tomados *k* a *k*
  (0 ≤ *k* ≤ *n*) é:

$$P_{n,k} = n(n-1)\ldots(n-k+1) = \frac{n!}{(n-k)!}$$

- Obs: $0! \equiv 1$. Por quê?
  - Várias razões (ver livro E. Scheinerman, Matemática Discreta: Uma Introdução, Seção: "*Muita confusão em torno de 0!*")
  - Duas razões intuitivas:
    a) Se n! é o número de listas ordenadas que podem ser formadas com *n* números, o número de listas ordenadas possíveis com 0 números é $0! = 1$ (a lista vazia).
    b) $0! = 1$ permite casos particulares da equação acima: $P_{n,n} = \frac{n!}{0!} = n!$

1.7 Métodos de contagem – Permutações e arranjos

- Exemplo: ***Amostragem sem reposição:***
  Considere um experimento no qual:
  - uma carta é sorteada e removida de um baralho de *n* cartas diferentes;
  - uma segunda carta é então sorteada e removida das $n - 1$ cartas restantes;
  - uma terceira carta é sorteada das $n - 2$ cartas restantes.
  - Cada resultado consiste nas 3 cartas na ordem selecionada.
  - O número de sequências de cartas possíveis é $P_{n,3} = n(n-1)(n-2)$
  - A probabilidade de cada sequência é $1/P_{n,3}$

1.7 Métodos de contagem – Permutações e arranjos

- Exemplo: Escolha de diretores:
  Suponha que um clube consiste de 25 membros, e que um presidente e um secretário devem ser escolhidos para a diretoria do clube. Devemos determinar o número total de maneiras em que essas duas posições podem ser escolhidas.
  - Resp: Como as posições devem ser preenchidas escolhendo-se primeiramente um dentre os 25 membros para ser o presidente, e então escolhendo-se um entre os 24 membros restantes para ser o secretário, o número de possíveis escolhas é $P_{25,2} = 25 \times 24 = 600$.

1.7 Métodos de contagem – Permutações e arranjos

- Voltemos ao problema da amostragem com reposição:
- Considere uma urna com *n* bolas numeradas $1, \dots, \mathrm{n}$.
  - Primeiro, uma bola é sorteada na urna, e seu número anotado. Essa bola é então colocada de volta na urna; Esse processo se repete, com um total de *k* bolas sorteadas da mesma maneira, e cada resultado consiste na sequência de *k* bolas, na ordem em que foram selecionadas.
  - O espaço amostral S desse experimento é o conjunto de todos os vetores $(x_1, \dots, x_k)$, onde x_i é o número da i-ésima bola sorteada $(i = 1, \dots, k)$. Como há *n* possibilidades de sorteio para cada uma das *k* bolas, o número total de vetores em *S* é n^k.
  - Além disso, assumindo-se que cada uma das bolas é igualmente provável de ser sorteada em cada passo, e que todos os sorteios são independentes dos demais, a probabilidade de cada vetor em S é $1/n^k$.
- Problema: Qual a probabilidade do evento *E* em que cada uma das *k* bolas sorteadas terão um número diferente?

1.7 Métodos de contagem – Permutações e arranjos

- Problema: Qual a probabilidade do evento *E* em que cada uma das *k* bolas sorteadas terão um número diferente?
  - Resp: Se $k > n$, isso é impossível.
  - Logo, assumiremos $k \leq n$.
  - O número de resultados em *E* é o número de vetores para os quais todos as k componentes são diferentes.
    Esse número é $P_{n,k}$, pois a componente x_1 de cada vetor tem n valores possíveis, a componente x_2 tem n-1 valores possíveis, e assim por diante.
  - Como *S* é um espaço amostral simples contendo n^k vetores, a probabilidade p de que k números diferentes serão sorteados é:

$$p = \frac{P_{n,k}}{n^k} = \frac{n!}{(n-k)!\, n^k}.$$

1.7 Métodos de contagem – Permutações e arranjos

- Problema do aniversário: enunciado
  - Determinar a probabilidade p de que **ao menos duas pessoas** em um grupo de k fazem aniversário no mesmo dia.
  - Algumas simplificações:
    - As datas de aniversário são independentes (não há gêmeos, por exemplo);
    - Todos os 365 dias do ano típico são equiprováveis
    - Aniversários em 29/02 seriam considerados no dia anterior ou no dia seguinte
      - Ou então poderíamos assumir que nenhuma das pessoas do grupo faz aniversário no dia 29/02

1.7 Métodos de contagem – Permutações e arranjos

- Problema do aniversário: solução
  - S = conjuntos de vetores $(x_1, \ldots, x_k)$, onde x_i é a data de nascimento do i-ésimo indivíduo $(i = 1, \ldots, k)$.
    $$|\mathrm{S}| = 365^k$$
  - Queremos determinar quantos vetores $(x_1, \ldots, x_k)$ podem ter **pelo menos** duas datas de aniversário iguais → trabalhoso porque:
    $\Pr(\geq 2 \text{ datas iguais}) = \Pr(2 \text{ datas iguais}) + \Pr(3 \text{ datas iguais}) + \cdots$
  - Solução mais simples:
    $\Pr(\geq 2 \text{ datas iguais}) = 1 - \Pr(\text{todas as datas diferentes})$
  - a quantidade de resultados $(x_1, \ldots, x_k)$ em que todas as datas de aniversário **são diferentes** é $P_{365,k}$.
    Assim, a probabilidade de todas as datas de aniversário em grupo de *k* pessoas serem diferentes é $P_{n,k}/365^k$.
  - Portanto, a probabilidade de ao menos duas pessoas em um grupo de *k* fazerem aniversário no mesmo dia será
    $$p = 1 - \frac{P_{365,k}}{365^k} = 1 - \frac{365!}{(365-k)!\,365^k}.$$

1.7 Métodos de contagem – Permutações e arranjos

- Problema do aniversário: exemplos numéricos

Table 1.1 The probability p that at least two people in a group of k people will have the same birthday

k	p	k	p
5	0.027	25	0.569
10	0.117	30	0.706
15	0.253	40	0.891
20	0.411	50	0.970
22	0.476	60	0.994
23	0.507		

1.7 Métodos de contagem – Cômputo de fatoriais

- Obs: Cômputo de razões entre fatoriais ou potências de números grandes:
  - O cálculo direto dos fatoriais e potências na expressão abaixo é computacionalmente inviável para valores elevados de n e k:

$$p = 1 - \frac{P_{n,k}}{n^k} = 1 - \frac{n!}{(n-k)!\,n^k}$$

Na linguagem R: 365! não é computável

  - Soluções:
    - Aproximação por fórmula de Stirling (ver livro do DeGroot)
    - Uso de logaritmos
  - Usando logaritmos: ex:
    - $lpd \equiv \log\left(\frac{n!}{(n-k)!n^k}\right) = \log(n!) - \log[(n-k)!\,n^k] =$

      $= \sum_{i=1}^{n} \log(i) - \sum_{i=1}^{n-k} \log(i) - k\log(n)$
    - $p = 1 - \exp(lpd)$
    - Sintaxe em R para o exemplo acima:

```
> n = 365; k=20
> lpd = sum(log(1:n)) - sum(log(1:(n-k))) - k*log(n)
> p = 1 - exp(lpd)
```

1.8 Métodos combinatórios - Combinações

- Exemplo: Escolha de subconjuntos de tamanho 2
  Considere o conjunto {a, b, c, d} contendo as quatro letras diferentes. Desejamos contar o número de subconjuntos distintos de tamanho 2.
  - Neste caso, podemos listar todos os subconjuntos de tamanho 2:
    {a, b}, {a, c}, {a, d}, {b, c}, {b, d}, {c, d}
    o que resulta em 6 subconjuntos distintos.
  - Esse problema é diferente de contar os arranjos, porque {a, b} e {b, a} são o mesmo subconjunto. O mesmo vale para todos os demais subconjuntos.
  - Se fôssemos contar o número de arranjos de 4 elementos, tomados 2 a 2, teríamos $P_{n,k} = n!/(n-k)!$ = 4!/2! = 12
    (a,b), (b,a), (a,c), (c,a), (a,d), (d,a), (b, c), (c,b), (b, d), (b,d), (c, d), (d,c)
  - Para encontrar o número de subconjuntos distintos de tamanho 2 em *n* elementos, podemos dividir $P_{n,k}$ pelo tamanho de cada classe de equivalência (número de pares equivalentes em cada subconjunto).

1.8 Métodos combinatórios - Combinações

- Exemplo: Escolha de subconjuntos de tamanho 3
  Considere o conjunto {a, b, c, d} contendo as quatro letras diferentes. Desejamos contar o número de subconjuntos distintos de tamanho 3.
  - Neste caso, podemos listar todos os subconjuntos de tamanho 3:

    {a, b, c}, {a, b, d}, {a,c,d}, {b, c, d}

    o que resulta em 3 subconjuntos distintos.
  - Outra forma de responder: Contando o número de arranjos de 4 elementos, tomados 3 a 3: $P_{4,3} = \frac{4!}{(4-3)!} = 24$

    (a,b,c), (a,c,b), (b,a,c), (b,c,a), (c,a,b), (c,b,a),
    (a,b,d), (a,d,b), (b,a,d), (b,d,a), (d,a,b), (d,b,a),
    (a,c,d), (a,d,c), (c,a,d), (c,d,a), (d,a,c), (d,c,a),
    (b,c,d), (b,d,c), (c,b,d), (c,d,b), (d,b,c), (d,c,b)

    Para encontrar o número de subconjuntos distintos de tamanho 3 em *n* elementos, podemos dividir $P_{n,3}$ pelo tamanho de cada classe de equivalência (número de tuplas equivalentes em cada subconjunto).
    - Número de tuplas equivalentes para cada subconjunto = ?

1.8 Métodos combinatórios - Combinações

- **Definição 1.8.1: Combinações**
  Considere um conjunto com *n* elementos. Cada subconjunto de tamanho *k* selecionado desse conjunto é denominado uma **combinação de *n* elementos tomados *k* a *k***.

- **Teorema 1.8.1: Combinações**
  O número de subconjuntos de tamanho k que podem ser formados de um conjunto de tamanho n é
  $$C_{n,k} = \frac{P_{n,k}}{k!} = \frac{n!}{k!\,(n-k)!}$$
  – Dem: $C_{n,k}$ deve ser igual ao número de arranjos de n tomados *k* a *k* ($P_{n,k}$) dividido pelo número de sequências possíveis formadas com k elementos, isto é, $k!$.

1.8 Métodos combinatórios - Combinações

- Exemplo (cont): Escolha de subconjuntos de tamanho 3
  - Considere o conjunto {a, b, c, d} contendo as quatro letras diferentes. Desejamos contar o número de subconjuntos distintos de tamanho 3.
  - Arranjos de 4 elementos tomados 3 a 3: 24

    $\boxed{(a,b,c),\ (a,c,b),\ (b,a,c),\ (b,c,a),\ (c,a,b),\ (c,b,a)}$,
    $\boxed{(a,b,d),\ (a,d,b),\ (b,a,d),\ (b,d,a),\ (d,a,b),\ (d,b,a)}$,
    $\boxed{(a,c,d),\ (a,d,c),\ (c,a,d),\ (c,d,a),\ (d,a,c),\ (d,c,a)}$,
    $\boxed{(b,c,d),\ (b,d,c),\ (c,b,d),\ (c,d,b),\ (d,b,c),\ (d,c,b)}$

    Para encontrar o número de subconjuntos distintos de tamanho 3 em *n* elementos, podemos dividir $P_{n,3}$ pelo tamanho de cada classe de equivalência (número de tuplas equivalentes em cada subconjunto).

    - Número de tuplas equivalentes para cada subconjunto = 3!

$$C_{4,3} = \frac{P_{4,3}}{3!} = \frac{4!}{3!\ (4-3)!} = 4$$

1.8 Métodos combinatórios – Coeficientes binomiais

- **Definição 1.8.2: Coeficientes binomiais**

  O número $C_{n,k}$ é também denotado pelo símbolo $\binom{n}{k}$.

  Para $k = 0,1,\dots,n$,

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

  Quando essa notação é usada, esse número é denominado **coeficiente binomial**.

1.8 Métodos combinatórios – Coeficientes binomiais

- O nome *coeficiente binomial* é derivado do *teorema binomial*:
- **Teorema 1.8.2: Teorema Binomial**

  Para quaisquer números *x* e *y* e para todo inteiro positivo *n*,

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}.$$

  – Ideia: a potência $(x+y)^n$ é igual a uma soma com 2^n termos, cada qual correspondendo a um produto de x's e y's. P.Ex:

  $$\begin{aligned}(x+y)^3 &= xxx + xxy + xyx + xyy + yxx + yxy + yyx + yyy \\ &= x^3y^0 + x^2y + x^2y + xy^2 + x^2y + xy^2 + xy^2 + x^0y^3\end{aligned}$$

  Para cada valor de $k = 0,1,\ldots,n$, $\binom{n}{k}$ denota o número de termos contendo o produto de *k x*'s por (*n*-*k*) *y*'s, independentemente da ordem em que os *x*'s e *y*'s aparecem.

  No exemplo acima, o termo x^2y aparece $\binom{3}{2} = 3$ vezes.

1.8 Métodos combinatórios – Coeficientes binomiais

- **Teorema 1.8.3:**
  Para todo *n*,
  $$\binom{n}{0} = \binom{n}{n} = 1.$$
  Para todo *n* e todo $k = 0,1, \dots, n$,
  $$\binom{n}{k} = \binom{n}{n-k}.$$
  – Dem: Equação 1 resulta de 0! = 1. Equação 2 decorre da equação descrita na definição 1.8.2.

1.8 Métodos combinatórios – Coeficientes binomiais

- Outra identidade útil (livro Ross):

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$$

  - Dem 1: analiticamente: basta mostrar que $\binom{n-1}{k-1} + \binom{n-1}{k} = \binom{n}{k}$
  - Dem 2: Considere um grupo de n objetos, e fixe atenção em um desses objetos – chamemos de objeto 1.
    Há $\binom{n-1}{k-1}$ grupos de tamanho k que contêm o objeto 1 (pois cada grupo é formado selecionando k-1 dentre os n-1 objetos restantes);
    Também há $\binom{n-1}{k}$ grupos de tamanho k que não contêm o objeto 1;
    Como existe um total de $\binom{n}{k}$ grupos de tamanho k, a equação vale.
  - Resultado base para o triângulo de Pascal (https://www.mathsisfun.com/pascals-triangle.html)

1.8 Métodos combinatórios – Coeficientes multinomiais

- Exemplo: Escolha de comitês
  Suponha que 20 membros de uma organização devem ser divididos em 3 comitês A, B e C, de tal maneira que o comitê A deve ter 8 membros, o comitê B deve ter 7 membros e o comitê C deve ter 5 membros. Determinar o número de possíveis divisões.
  - Para formar o comitê A, precisamos escolher 8 dos 20 membros, o que pode ser feito de $\binom{20}{8}$ maneiras. Em seguida, para dividir os 12 membros restantes nos comitês B e C existem $\binom{12}{7}$ maneiras.
  - Logo, a resposta é: $\binom{20}{8}\binom{12}{7} = \frac{20!}{8!12!}\,\frac{12!}{7!5!} = \frac{20!}{8!7!5!}$.
  - Note que o 12! que aparece no numerador de $\binom{20}{8}$ se cancela com o 12! que aparece no denominador de $\binom{12}{7}$ ➔ chave para o resultado geral a seguir.

1.8 Métodos combinatórios – Coeficientes multinomiais

- No caso geral, suponha que n elementos distintos devem ser divididos em k grupos diferentes ($k \geq 2$) de maneira que, para $k = 1, \dots, k$, o j-ésimo grupo contém exatamente n_j elementos, onde $n_1 + n_2 + \cdots + n_k = n$.
  Deseja-se determinar o número de maneiras diferentes nas quais os *n* elementos podem ser divididos nos *k* grupos.
  - Os n_1 elementos no 1º grupo podem ser escolhidos dentre os n elementos disponíveis de $\binom{n}{n_1}$ maneiras diferentes.
  - Depois da formação do 1º grupo, os n_2 elementos do 2º grupo podem ser selecionados dos $n - n_1$ elementos restantes de $\binom{n - n_1}{n_2}$ maneiras distintas.
  - Para a formação do 3º grupo, há $\binom{n - n_1 - n_2}{n_3}$ maneiras distintas.
  - O número total de maneiras para formar os 3 primeiros grupos será então:

$$\binom{n}{n_1}\binom{n - n_1}{n_2}\binom{n - n_1 - n_2}{n_3}$$

1.8 Métodos combinatórios – Coeficientes multinomiais

- No caso geral (cont):
  - O processo continua até a escolha do grupo (k-1), que terá $\binom{n - n_1 - \ldots - n_{k-2}}{n_{k-1}}$.
    Quando esse grupo tiver sido formado, os n_k elementos restantes irão obrigatoriamente para o k-ésimo grupo.
  - Logo, o número total de maneiras diferentes de dividir os n elementos nos k grupos será

$$\binom{n}{n_1}\binom{n-n_1}{n_2}\binom{n-n_1-n_2}{n_3}\cdots\binom{n-n_1-\ldots-n_{k-2}}{n_{k-1}} =$$

$$\frac{n!}{n_1!\ (n-n_1)!}\ \frac{(n-n_1)!}{n_2!\ (n-n_1-n_2)!} \cdots \frac{(n-n_1-\ldots-n_{k-2}\ )!}{n_{k-1}!\ n_k!} =$$

$$\frac{n!}{n_1!\, n_2! \ldots n_k!}\ .$$

1.8 Métodos combinatórios – Coeficientes multinomiais

- **Definição 1.9.1: Coeficientes multinomiais**
  O número $\frac{n!}{n_1!n_2!\ldots n_k!}$, denotado por $\binom{n}{n_1, n_2, \ldots, n_k}$, é denominado **coeficiente multinomial**.

- O nome *coeficiente multinomial* deriva do *teorema multinomial* (próximo slide).

1.8 Métodos combinatórios – Coeficientes multinomiais

- **Teorema 1.9.1: Teorema multinomial**
  Para quaisquer números $x_1, x_2, \dots, x_k$ e qualquer inteiro positivo n,

$$(x_1 + x_2 + \cdots + x_k)^n = \sum_{\substack{n_1,\dots,n_k \mid \\ n_1+\cdots+n_k=n}} \binom{n}{n_1, n_2, \dots, n_k} (x_1)^{n_1} (x_2)^{n_2} \dots (x_k)^{n_k}$$

1.8 Métodos combinatórios – Coeficientes multinomiais

- O coeficiente multinomial é uma generalização do coeficiente binomial.

$$\binom{n}{k, n-k} \equiv \binom{n}{k} .$$

- Exemplo: Escolha de comitês
  Suponha que 20 membros de uma organização devem ser divididos em 3 comitês A, B e C, de tal maneira que o comitê A deve ter 8 membros, o comitê B deve ter 7 membros e o comitê C deve ter 5 membros. Determinar o número de possíveis divisões.
  - Resp: $n = 20,\ k = 3,\ n_1 = 8,\ n_2 = 7,\ n_3 = 5$

$$\binom{n}{n_1, n_2, n_3} = \binom{20}{8, 7, 5} = \frac{20!}{8!\,7!\,5!} .$$