Resumo – Elementos de Probabilidades e Estatística

Capítulo 1 – Dados Estatísticos

1.2 Amostra, População

- **Colecção de Dados** : conjunto de observações de certo(s) atributo(s), qualquer que seja a forma como foram recolhidas
- **Categorias Básicas de Dados** :
  a) **Dados temporais** : dados cujas observações se referem a uma mesma entidade, para vários momentos ou periodos de tempo ; caracterizados essencialmente pela sua ordem cronológica ; a frequência temporal das observações é outro aspecto importante ; descrição da evolução no tempo → série temporal / cronograma ;
  b) **Dados seccionais** : dados cujas observações se referem a determinadas entidades (unidades estatísticas) em certo momento ou periodo de tempo ; a característica fundamental é que a ordem das observações é irrelevante ;
  c) **Dados seccionais combinados** : dados com aspectos seccionais e temporais ; conjunto de dados seccionais, cada um referente a certa data ; comparação de comportamentos entre as datas ;
  d) **Dados de painel** : conjunto fixo de entidades observadas em várias datas ; característica essencial → o conjunto de entidades a observar é sempre o mesmo para todas as observações temporais (dificulta a obtenção) / várias observações temporais para a mesma entidade
- **População** : conjunto dos elementos cujos atributos são objecto de um determinado estudo → conjuntos fundamentais para efectuar análises estatísticas
- **Unidade Estatística** : facto ou entidade elementar que é objecto de observação, qualquer que seja a sua natureza
- **Censo / Indagação Completa** : análise de todos os elementos de uma população
- **Amostra** : subconjunto finito de uma população sobre o qual incide o estudo dos atributos da população (conjunto de unidades estatísticas)
- **Processo de Amostragem** : forma de selecção de uma amostra a partir da população ; determinante para a qualidade das inferências ; geralmente aleatório
- **Tipos de Atributos** :
  1) **Atributos Quantitativos** : os seus valores determinam as suas intensidades
  2) **Atributos Qualitativos** :
     a) **Representação em Escalas Nominais** : atribuição dos valores não tem qualquer significado (ex: “mulher” = 0 ; “homem” = 1)
     b) **Representação em Escalas Ordinais** : numeração deve respeitar a ordem das várias modalidades (ex: graus de especialização de 1 a 5)
- **Variável e Dimensão** : O valor númerico de um atributo é representado por uma variável x. Se a amostra observada tem dimensão n (ou seja, n elementos), tem-se $x_1, x_2, \dots, x_n$ onde x_i $(i = 1,2, \dots, n)$ é o valor do atributo da i-ésima observação.
- **Tipos de Variáveis** :
  a) **Variáveis Discretas** : variáveis que podem tomar somente um número finito ou uma infinidade numerável de valores (ex: contagens, escalas nominais/ordinais, etc)
  b) **Variáveis Contínuas** : variáveis que podem tomar qualquer valor dentro de um intervalo de números reais (ex: altura, peso, taxas de juro, etc)

1.3 Gráfico caule-e-folhas

- **Gráfico caule-e-folhas** : método semigráfico usado quando a amostra é pequena ; permite que o observador se torne mais sensível ao aspecto global dos dados e comece a explorar a existência de padrões
- **Construção/Elementos dos gráficos caule-e-folhas** :
  - **Número de folhas** : número de dígitos dominados de cada dígito dominante
  - **Caule** : dígitos dominantes
  - **Folhas** : dígitos dominados
  - n : número total de folhas (dimensão da amostra)
  - **unidade** : primeiro dígito dominado

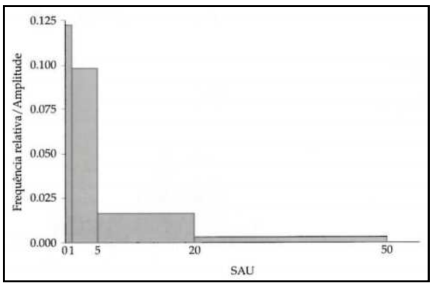
N.º de folhas	Caule	Folhas
2	5	47
10	6	2356668889
14	7	00123334457899
8	8	11346789
7	9	0134479
2	10	59

$n = 43$; unidade = 0.1.

- **Excepções** :
  - **Caules com demasiadas folhas** : $12+$ e $12-$ / <u>dígito dominado</u> : $0^*, t, f, s, 0\cdot, 1^*, \ldots$
  - **Valores anormalmente pequenos/grandes** : novas linhas rotuladas com “Pe” e “Gr”
- **Número de linhas de um gráfico caule-e-folhas** : parte inteira de $2\sqrt{n}$
- **Propriedades que se podem identificar graças à construção de gráficos caule-e-folhas** :
  a) Existência de valores em torno dos quais se concentram os restantes
  b) Maior ou menor dispersão dos valores
  c) Simetrias ou assimetrias
  d) Presença de valores extremos (“pequenos” e “grandes”)

1.4 Distribuições de Frequências, Histogramas

- **Frequência Absoluta** : número de vezes que um valor de x é observado ; notação: F_j ; propriedade : $\sum_j F_j = n$
- **Frequência Relativa** : $f_1 = \frac{F_j}{n}$; propriedade : $\sum_j f_j = 1$
- **Variável x contínua → classes de valores** :
  - → os intervalos de classe não têm pontos em comum : $I_j \cap I_k = \emptyset \, , j \neq k$
  - → o domínio A da variável deve ser igual à união de todos os intervalos : $A = \bigcup_{i=1}^{m} I_i$
  - → os intervalos devem ser abertos à esquerda e fechados à direita : $I_j = \left]l_{j-1}, l_j\right]$
  - → a amplitude deve ser constante : $h_j = l_j - l_{j-1}$ $(j = 1,2,\ldots,m)$
- **Representação gráfica de distribuições de frequências de variáveis discretas** : diagramas de barras
- **Representação gráfica de distribuições de frequências de variáveis contínuas** :
  - diagrama de áreas / histograma, formado por uma sucessão de rectângulos adjacentes
  - áreas dos rectângulos = resp. frequências relativas
  - **Intervalos de classe com amplitude constante** :
    - base do rectângulo : $h = 1$
    - altura do rectângulo : f_1 ou F_j
  - **Intervalos de classe com amplitudes diferentes** :
    - base do rectângulo : $h_j = l_j - l_{j-1}$
    - altura do rectângulo : $\frac{f_1}{h_j}$ ou $\frac{F_j}{h_j}$



1.5 Características Numéricas : Média e Desvio-Padrão

- **Média** : $\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i$ (medida de localização)
- **Ponto médio de uma classe** : $x_j' = l_{j-1} + \frac{1}{2}h_j = \frac{l_{j-1} + l_j}{2}$
- **Média (de dados agrupados em classes)** : $\bar{x} = \frac{1}{n}\sum_{j=1}^{m} F_j x_j' = \sum_{j=1}^{m} f_j x_j'$
- **Propriedade da Média** : A soma dos desvios em relação à média é nula. $\sum_{i=1}^{n}(x_i - \bar{x}) = 0$

- **Moda (M_*)** : valor com maior frequência
- **Classe modal** : classe com maior frequência
- **Variância** : $s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$ (medida de dispersão)
- **Variância corrigida** (quando a amostra tem dimensão pequena) : $s'^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$
- **Desvio-Padrão** : $s = \sqrt{s^2} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$ ou $s' = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$
- **Variância (para dados classificados)** : $s^2 = \frac{1}{n}\sum_{j=1}^{m} F_j\left(x_j' - \bar{x}\right)^2 = \sum_{j=1}^{m} f_j\left(x_j' - \bar{x}\right)^2$
- **Variância corrigida (para dad. class.)** : $s'^2 = \frac{1}{n-1}\sum_{j=1}^{m} F_j\left(x_j' - \bar{x}\right)^2 = \sum_{j=1}^{m} \frac{n}{n-1} f_j\left(x_j' - \bar{x}\right)^2$
- **Variância (forma simplicifada)** : $s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i{}^2) - \bar{x}^2$
- **Desvio-absoluto médio** : $d = \frac{1}{n}\sum_{i=1}^{n}|x_i - \bar{x}|$ ou $d = \frac{1}{n}\sum_{j=1}^{m} F_j\left|x_j' - \bar{x}\right|$
- **Coeficiente de Variação** : $CV = \frac{s}{\bar{x}}$; observações com sinal positivo / independente da unidade em que se exprime a variável / permite comparação entre distribuições

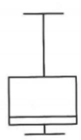
1.6 Características Numéricas : Estatísticas da Ordem

- **Extremos** : $x_{(1)} = min(x_i)$; $x_{(n)} = max(x_i)$
- **Mediana** : valor da colecção que tem 50% de observações inferiores e 50% de observações superiores ; $M = \begin{cases} x_{(k+1)} & se\ n = 2k+1 \\ \frac{x_{(k)} + x_{(k+1)}}{2} & se\ n = 2k \end{cases}$
- **Quantil de ordem α** : valor da colecção que tem αn observações inferiores e $(1-\alpha)n$ observações superiores ; calcula-se a ordem: $r = 1 + (n-1)\alpha$;
  $\rightarrow$ se $r \in \mathbb{Z}$, $q_\alpha = x_{(r)}$
  $\rightarrow$ se $r \notin \mathbb{Z}$, $q_\alpha = x_{(r)} + \varepsilon\left(x_{(r+1)} - x_{(r)}\right) = (1-\varepsilon)x_{(r)} + \varepsilon x_{(r+1)}$, onde $0 \le \varepsilon < 1$
- **Quartis** :

Quartis	Ordem aproximada
1º quartil : Q_1	$\frac{n+3}{4}$
Mediana : M	$\frac{n+1}{2}$
3º quartil : Q_3	$\frac{3n+1}{4}$

r é a parte inteira
ε é a parte fraccionária

- **Amplitude do Intervalo de Variação** : $AIV = x_{(n)} - x_{(1)}$ (medida de dispersão absoluta)
- **Amplitude do Intervalo de Interquartis** : $AIQ = Q_3 - Q_1$ (medida de dispersão absoluta)
- Uma medida de dispersão relativa (qdo todas as obs. têm o mesmo sinal) : $\frac{AIQ}{M}$
- **Resumo dos 5 números** : mínimo, 1º quartil, mediana, 3º quartil, máximo
  - Gráfico : caixa-de-bigodes
    - caixa formada por 1º e 3º quartis
    - mediana atravessa a caixa
    - "bigodes" → mínimo e máximo
    - caixa contém 50% das observações



- **Estimativa da estatística de ordem $x_{(i)} \in I_j$** : $\hat{x}_{(i)} = l_{j-1} + \left[\left(i - F_{j-1}^*\right) - \frac{1}{2}\right]\frac{h_j}{F_j}$, onde:

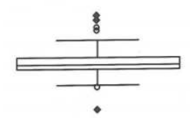
- l_{j-1} : limite inferior da classe I_j ;
- F_{j-1}^* : frequência absoluta acumulada até à classe I_{j-1}: $F_1 + F_2 + \cdots + F_{j-1}$;
- h_j : amplitude da classe I_j ;
- F_j : frequência absoluta da classe I_j ;

- **Estimativa do quantil $q_\alpha \in I_j$** : $q_\alpha = l_{j-1} + \frac{\alpha - f_{j-1}^*}{f_j} h_j$
- **Amplitude dos intervalos de classe** : $h = \frac{2 \times AIQ}{\sqrt[3]{n}}$
- **Principais medidas de localização e de dispersão** :

Medidas	$\bar{x}$	s	CV	$m.$	M	AIQ	AIQ/M
Localização	×			×	×		
Dispersão absoluta		×				×	
Dispersão relativa			×				×
Atributo quantitativo	×	×	×	×	×	×	×
Atributo qualitativo nominal				×			
Atributo qualitativo ordinal				×	×	×	×

1.7 Outliers

- **Outlier Severo** : x_i é um outlier severo quando $x_i < Q_1 - 3(Q_3 - Q_1)$
  ou $x_i > Q_3 + 3(Q_3 - Q_1)$
- **Outlier Moderado** : x_i é um outlier moderado quando
  $Q_1 - 3(Q_3 - Q_1) < x_i < Q_1 - 1{,}5(Q_3 - Q_1)$
  ou $Q_3 + 1{,}5(Q_3 - Q_1) < x_i < Q_3 + 3(Q_3 - Q_1)$
- **Barreira externa inferior** : $Q_1 - 3(Q_3 - Q_1)$
- **Barreira externa superior** : $Q_3 + 3(Q_3 - Q_1)$
- **Barreira interna inferior** : $Q_1 - 1{,}5(Q_3 - Q_1)$
- **Barreira interna exterior** : $Q_3 + 1{,}5(Q_3 - Q_1)$
- **Caixa-de-bigodes (modificada)** :



  → bigodes : barreiras internas superior e inferior
  → outliers moderados/severos → representados por símbolos gráficos diferentes (∘, ■)

1.8 Correlação

- **Covariância** : $s_{xy} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$ ou $s_{xy} = \frac{1}{n}\sum_{i=1}^{n}(x_i y_i) - \bar{x}\bar{y}$
- **Coeficiente de correlação de Pearson** : $r = \frac{s_{xy}}{s_x s_y} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}}$

  → $-1 \le r \le -1$
  → CC mede o grau de associação linear
  → sinal → correlação positiva ou negativa
  → valor absoluto → intensidade da associação linear
  → CC próximo de 1 → dados dispostos essencialmente nos 1º e 3º quadrantes (linear)
  → CC próximo de -1 → dados dispostos essencialmente nos 2º e 4º quadrantes (linear)
- **Coeficiente de correlação de Spearman** : Substitui-se cada par (x_i, y_i) é substituído pelo par $[ord(x_i), ord(y_i)]$, onde $ord(x_i)$ representa a ordem de x_i na colecção. Calcula-se o coeficiente de correlação com a fórmula habitual.
  – mesmos valores → média das ordens
  – outliers podem influenciar CC → retirar outliers

Capítulo 2 – Probabilidades

2.1 Introdução

- **Experiência aleatória** : Diz-se que uma experiência é aleatória se apresentar as seguintes características:
  a) O conjunto dos resultado possíveis é fixado antecipadamente;
  b) O resultado da experiência nunca pode ser previsto de forma exacta, mesmo que se desenvolvam todos os esforços para manter sob controlo as circunstâncias relevantes para o resultado.

2.2 Espaço de resultados / Acontecimentos

- **Espaço de resultados ou Espaço-amostra (Ω)** : conjunto fundamental (não vazio) formado por todos os resultados que hipoteticamente é possível obter quando se efectua uma determinada experiência aleatória
- **Acontecimento** : subconjunto do espaço Ω
  → Acontecimentos elementares : subconjuntos $\{\omega\} \subset \Omega$ formados por um só elemento
- **Realização de um acontecimento** : Ao efectuar a experiência aleatória associada com Ω, diz-se que o acontecimento $A \subset \Omega$ se realiza, se o resultado da experiência é um ponto que pertence a A: $\omega \in A$
- **Álgebra dos acontecimentos** :
  - Implicação de acontecimentos : $(A \subset B) <=> (x \in A => x \in B)$
  - Identidade de acontecimentos : $(A = B) <=> (x \in A <=> x \in B)$
  - União de acontecimentos : $A \cup B = \{x : x \in A \vee x \in B\}$ $(A \cup \emptyset = A\,;\, A \cup \Omega = \Omega)$
  - Intersecção de acontecimentos : $A \cap B = \{x : x \in A \wedge x \in B\}$ $(A \cap \emptyset = \emptyset\,;\, A \cap \Omega = A)$
  - Acontecimentos incompatíveis / mutuamente exclusivos : A e B são incompatíveis sse a realização de um implica a não realização do outro (acontecimentos disjuntos)
  - Acontecimento impossível : $\emptyset$ (exemplo : A e B são incompatíveis sse $A \cap B = \emptyset$)
  - Acontecimento certo : Ω (para qualquer ω da experiência aleatória tém-se $\omega \in \Omega$)
  - Diferença de acontecimentos : $A - B = A \cap \bar{B} = \{x : x \in A \wedge x \notin B\}$
  - Acontecimento contrário/complementar : $\bar{A} = \{x : x \notin A\}$ $(A \cup \bar{A} = \Omega\,;\, A \cap \bar{A} = \emptyset\,;\, \bar{\bar{A}} = A)$
- **Propriedades das operações definidas sobre acontecimentos** :
  1) Associatividade : $A \cup (B \cup C) = (A \cup B) \cup C$; $A \cap (B \cap C) = (A \cap B) \cap C$
  2) Comutatividade : $A \cup B = B \cup A$; $A \cap B = B \cap A$
  3) Distributividade : $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$; $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
  4) Leis de Morgan : $\overline{A \cup B} = \bar{A} \cap \bar{B}$; $\overline{A \cap B} = \bar{A} \cup \bar{B}$
- **Operações sobre infinidades numeráveis de acontecimentos** :
  1) União numerável : $\bigcup_{j=1}^{+\infty} A_j$
  2) Intersecção numerável : $\bigcap_{j=1}^{+\infty} A_j$

2.3 Medida de Probabilidade. Axiomática de Kolmogorov

- **Medida de probabilidade** : função P que a cada acontecimento $A, A \subset \Omega$, faz corresponder um número real, $P(A)$, probabilidade do acontecimento A, que verifica os 3 axiomas seguintes :
  P1) $P(A) \geq 0$
  P2) $P(\Omega) = 1$
  P3) Se A e B forem acont. incompatíveis, $A \cap B = \emptyset$, então $P(A \cup B) = P(A) + P(B)$
  → P3*) Se $A_1, A_2, A_3, \dots, A_n, \dots$ forem acontecimentos em número finito numerável, dois a dois incompatíveis, $A_i \cap A_j = \emptyset\ (i \neq j)$, então $P\left(\bigcup_{j=1}^{+\infty} A_j\right) = \sum_{j=1}^{+\infty} P(A_j)$

- **Propriedades da medida de probabilidade** :
  1) $P(A) = 1 - P(\bar{A})$
  2) $P(\emptyset) = 0$
  3) $P(B - A) = P(B) - P(A \cap B)$
  4) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
  5) $A \subset B \Rightarrow P(A) \leq P(B)$
  6) $P(A) \leq 1$

2.4 Interpretações do conceito de probabilidade

- **Regra de Laplace** : Sejam os acontecimentos elementares equiprováveis, então temos

$$P(A) = \frac{m}{n} = \frac{\#(A)}{\#(\Omega)} = \frac{n^{\underline{o}}\ de\ casos\ favoráveis\ a\ A}{n^{\underline{o}}\ de\ casos\ possíveis}$$

- **Como reconhecer que os casos são igualmente possíveis?**
  1) O princípio da indiferença faz apelo às propriedades de simetria ou de homogeneidade da situação experimental (se o dado é perfeito, porque seriam umas faces preferidas em detrimento de outras?);
  2) O princípio da razão insuficiente diz que se não há razão para crer em que qualquer dos casos é mais provável do que os outros, pode actuar-se como se todos os casos fossem igualmente prováveis.
- **Lei dos grandes números** : $\lim_{n \to +\infty} P(|f_n(A) - P(A)| < \varepsilon) = 1$
- As propriedades das frequências relativas verificam a axiomática de Kolmogorov :

  E1) $f_n(A) \geq 0$

  E2) $f_n(\Omega) = 1$

  E3) Se A e B forem acont. incompatíveis, então $f_n(A \cup B) = f_n(A) + f_n(B)$
- **Probabilidade subjectiva** (grau de credibilidade) : aproximação quantitativa à credibilidade que se atribui a um determinado acontecimento. (deve respeitar o princípio da coerência)

2.5 Métodos de contagem

- **Regra fundamental da contagem** : Suponha-se que uma experiência aleatória é composta por k etapas. Se na primeira etapa há m_1 casos possíveis, na segunda etapa há m_2 casos possíveis, …, na k-ésima etapa há m_k casos possíveis, então o número total de resultados da experiência aleatória, combinando as k etapas é $m_1 \times m_2 \times \ldots \times m_k$.
- **Amostra ordenada** : Se k elementos são seleccionados de um conjunto de n elementos e se a ordem da selecção é registada, o conjunto de k elementos forma uma amostra ordenada de dimensão k.
- **Amostragem sem reposição** : Quando se selecciona um elemento e o mesmo não é reposto no conjunto, diz-se que se seguiu um processo de amostragem sem reposição.
- **Amostragem com reposição** : Um processo de amostragem com reposição ocorre quando se selecciona um elemento e o mesmo é reposto no conjunto antes de se seleccionar o elemento seguinte.
- **Conceitos fundamentais de Análise Combinatória** :
  - **Arranjos (sem repetição)** : $A_k^n = \frac{n!}{(n-k)!}$
  - **Arranjos completos (com repetição)** : $\alpha_k^n = n^k$
  - **Permutações** : $P_n = A_n^n = n!$
  - **Combinações** : $C_k^n = \binom{n}{k} = \frac{A_k^n}{P_k} = \frac{n!}{k!(n-k)!}$ (coeficiente binomial)
  - **Permutações de elementos não todos distintos** : $\frac{n!}{k_1!k_2!\ldots k_r!}$ (coeficiente multinomial)

2.6 Probabilidade Condicionada. Teorema de Bayes

- **Probabilidade Condicionada** : $P(A|B) = \frac{P(A\cap B)}{P(B)}$ se $P(B) > 0$
- **Regra da multiplicação de probabilidades** : $P(A \cap B) = P(B) \times P(A|B) = P(A) \times P(B|A)$
  $\rightarrow$ Generalizável para 3 ou + acont. (Ex: $P(A \cap B \cap C) = P(A) \times P(B|A) \times P(C|A \cap B)$)
- **Partição do espaço de resultados** : Diz-se que a classe de acontecimentos $\{A_1, A_2, A_3, \dots, A_m, \dots\}$ é uma partição de Ω quando $A_i \cap A_j = \emptyset\ (i \neq j)$ e $\bigcup_j A_j = \Omega$
  $\rightarrow$ Consequência : $P(\bigcup_j A_j) = \sum_j P(A_j) = 1$
- **Regra da Probabilidade Total** : Se $\{A_1, A_2, A_3, \dots, A_m, \dots\}$ é uma partição de Ω e se $P(A_j) > 0\ (j = 1,2, \dots, m, \dots)$, vem, para qualquer acontecimento B,
  $$P(B) = \sum_j P(A_j) \times P(B|A_j)$$
- **Teorema de Bayes** : Se $\{A_1, A_2, A_3, \dots, A_m, \dots\}$ é uma partição de Ω e se $P(A_j) > 0$ $(j = 1,2, \dots, m, \dots)$, vem, para qualquer acontecimento B, a verificar $P(B) > 0$,
  $$P(A_j|B) = \frac{P(A_j) \times P(B|A_j)}{\sum_{i=1} P(A_j) \times P(B|A_j)} \quad (j = 1,2, \dots, m, \dots)$$
  Nota 1 : $\sum_j P(A_j|B) = 1$
  Nota 2 : $\sum_{i=1} P(A_j) \times P(B|A_j) = P(B)$
- **Interpretação do Teorema de Bayes** :
  $\rightarrow$ Probabilidade *a priori* : $P(A_j)$
  $\rightarrow$ Probabilidade *a posteriori* : $P(A_j|B)$
  $\rightarrow$ Verosimilhanças de B para cada A_j : $P(B|A_j)$
  - $[\text{probabilidade } a\ priori] \propto [\text{probabilidade } a\ posteriori] \times [\text{verosimilhança}]$
    $\propto$ → "proporcional a"

2.7 Acontecimentos Independentes

- **Acontecimentos independentes** : Dois acontecimentos, A e B, do mesmo espaço de resultados, dizem-se independentes, se e só se $P(A \cap B) = P(A) \times P(B)$
- **Teorema** : Se A e B forem acontecimentos independentes, então
  $$P(A|B) = P(A) \text{ se } P(B) > 0 \quad ; \quad P(B|A) = P(B) \text{ se } P(A) > 0$$
- **Nota sobre o Exemplo 2.38 (Lançamento de uma moeda regular/viciada)** : A igualdade $P(A \cap B) = P(A) \times P(B)$ verifica-se apenas nos casos triviais, $p = 0$ e $p = 1$, e no caso simétrico, $p = \frac{1}{2}$ (moeda "regular"). Assim, A e B podem ser ou não independentes, consoante a natureza da moeda.
- **Casos com 3 (ou mais) acontecimentos** : Os acontecimentos A, B e C podem ser independentes dois a dois e não se verificar $P(A \cap B \cap C) = P(A) \times P(B) \times P(C)$ ou vice-versa.
- **Independência completa ou mútua** : Os acontecimentos A, B e C do mesmo espaço de resultados dizem-se completamente independentes ou mutuamente independentes se e só se: $P(A \cap B) = P(A) \times P(B)$, $P(A \cap C) = P(A) \times P(C)$, $P(B \cap C) = P(B) \times P(C)$, $P(A \cap B \cap C) = P(A) \times P(B) \times P(C)$
  Nota : Definição extensível a quatro ou mais acontecimentos.
- **Independência condicional** : Dois acontecimentos, A e B, são independentes condicionalmente em relação a um acontecimento C se e só se
  $$P[(A \cap B)|C] = P(A|C) \times P(B|C)$$
  Nota : A independência condicional não implica independência no sentido corrente a não ser, obviamente, quando $C = \Omega$.

Capítulo 3 – Variável Aleatória / Função de Distribuição

3.1 Variável Aleatória

- **Variável Aleatória** : Uma variável aleatória X, é uma função com domínio Ω e com contradomínio $\mathbb{R}$. Assim: $X: \omega \in \Omega \rightarrow X(\omega) \in \mathbb{R}$

3.2 Função de Distribuição

- **Função de Distribuição** : A função real de variável real, F_X, com domínio $\mathbb{R}$, definida por
$$F_X(x) = P(X \leq x)$$
designa-se por função de distribuição da variável aleatória X.
- **Propriedades da Função de Distribuição** :
  1) $0 \leq F(x) \leq 1$
  2) F é não decrescente: $\Delta x > 0 \Rightarrow F(x) \leq F(x + \Delta x)$
  3) $F(-\infty) = \lim_{x \to -\infty} F(x) = 0$ e $F(+\infty) = \lim_{x \to +\infty} F(x) = 1$
  4) $P(a < X \leq b) = F(b) - F(a), \quad \forall a, b: b > a$
  5) F é contínua à direita: $F(a + 0) = \lim_{x \to a+0} F(x) = F(a)$
  6) $P(X = a) = F(a) - F(a - 0), \quad \forall a \in \mathbb{R}$
- **Acontecimentos com probabilidade zero** : Estes acontecimentos têm probabilidade zero, mas não são impossíveis. Exemplo: "Obter π na escolha de um número real qualquer no intervalo $[3,4]$". Voltar a calhar π é praticamente impossível, daí ser um acontecimento com probabilidade zero.
- **Mais Propriedades da Função de Distribuição** :
  a) $P(X < b) = F(b - 0)$
  b) $P(X > a) = 1 - F(a)$
  c) $P(X \geq a) = 1 - F(a - 0)$
  d) $P(a < X < b) = F(b - 0) - F(a)$
  e) $P(a \leq X < b) = F(b - 0) - F(a - 0)$
  f) $P(a \leq X \leq b) = F(b) - F(a - 0)$
- **Teorema** : A função real de variável real, F, é uma função de distribuição se e só se verifica as propriedades 2, 3 e 5.

3.3 Classificação das Variáveis Aleatórias

- **Conjunto dos Pontos de Descontinuidade** : $D = \{x : P(X = x) = F(X) - F(X - 0) > 0\}$
- **Classificação das Variáveis Aleatórias** :
  - X é uma **variável aleatória discreta** se e só se $P(X \in D) = 1$, ou seja, se e só se a soma dos saltos da função de distribuição for igual a 1. Assim:
  $$P(X \in D) = \sum_{x \in D} P(X = x) = 1$$
  - X é uma **variável aleatória contínua** se e só se $D = \emptyset$ e se existe uma função real de variável real f_X, não negativa, tal que
  $$F_X(x) = \int_{-\infty}^{x} f_X(u) du, \quad \forall x \in \mathbb{R}$$
  - Seja F_{X_1} uma função de distribuição associada a uma variável aleatória discreta, X_1, e F_{X_2} outra função de distribuição carterizadora de uma variável aleatória contínua, X_2. Diz-se que X é uma **variável aleatória mista** se e só se a respectiva função de distribuição puder ser escrita do seguinte modo:
  $$F_X(x) = \lambda F_{X_1}(x) + (1 - \lambda) F_{X_2}(x), \quad \text{onde } 0 < \lambda < 1$$
  - Existem variáveis aleatórias que <u>não são nem discretas, nem contínuas nem mistas</u>.

A) Variáveis Aleatórias Discretas

- <u>Nota</u> : Uma variável aleatória é discreta quando toda a massa de probabilidade está concentrada nos pontos de descontinuidade da função de distribuição.

- **Função Probabilidade** : A função f_X chama-se função probabilidade da variável aleatória discreta X se e só se:
$$f_X(x) = \begin{cases} P(X = x) > 0 & se \quad x \in D \\ P(X = x) = 0 & se \quad x \notin D \end{cases}$$
  → Suporte da Distribuição : Subdomínio de f_X ; valores de x que pertencem a D ; habitualmente, refere-se a função probabilidade apenas para os valores do suporte
  → $P(X \in B) = P(B) = \sum_{x \in B \cap D} f(x)$
  → $F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i)$, $\forall x \in \mathbb{R}$ com $D = \{x_1, x_2, \dots\}$
  → É indiferente representar-se uma v.a. discreta por $F(x)$ ou $f(x)$.

B) Variáveis Aleatórias Contínuas

- **Função Densidade de Probabilidade** : A função f_X, tal que $F_X(x) = \int_{-\infty}^{x} f_X(u)du$, chama-se função densidade de probabilidade ou simplesmente função densidade.
  → $f(x) = \begin{cases} F'(x) & \text{(nos pontos onde existe derivada)} \\ 0 & \text{(nos outros pontos)} \end{cases}$
  → $\int_{-\infty}^{+\infty} f(x)dx = F(+\infty) - F(-\infty) = 1$
  → $P(B) = P(X \in B) = \int_B f(x)dx$
  → Se $B =]a, b]$, com $a < b$, tem-se $P(B) = P(a < X \leq b) = \int_a^b f(x)dx = F(b) - F(a)$
  → Devido à continuidade da função de distribuição, $P(X = a) = P(X = b) = 0$, e pode escrever-se:
  $P(a < X \leq b) = P(a \leq X < b) = P(a < X < b) = P(a \leq X \leq b)$, com $a < b$
  → É indiferente representar-se uma v.a. discreta por $F(x)$ ou $f(x)$.

C) Variáveis Aleatórias Mistas

- **Variável Aleatória Mista** : Seja F_{X_1} uma função de distribuição associada a uma variável aleatória discreta, X_1, e F_{X_2} outra função de distribuição carterizadora de uma variável aleatória contínua, X_2. Diz-se que X é uma variável aleatória mista se e só se a respectiva função de distribuição puder ser escrita do seguinte modo:
$$F_X(x) = \lambda F_{X_1}(x) + (1 - \lambda)F_{X_2}(x), \quad \text{onde } 0 < \lambda < 1$$

3.4 Funções de uma variável aleatória

- **Mudança de Variável** : Seja X uma variável aleatória, e considere-se uma função $Y = \psi(X)$ é também uma variável aleatória que assume o valor $y = \psi(x)$ quando $X = x$. Conhecida a função de distribuição de X, F_X, pretende-se determinar a função de distribuição de Y, F_Y. Ao substituir a variável aleatória X pela variável aleatória Y, está a efectuar-se uma mudança de variável.
  → Utilidade : facilita a interpretação e o tratamento quando da variável aleatória original
- **Caso 1 : Variável Aleatória Discreta** :
  – **Conjunto dos pontos que Y assume com probabilidade positiva** : $D_Y = \psi(D_X)$;
  $$A_y = \{x : \ \psi(x) = y, x \in D_X\}$$
  – **Função Probabilidade de Y** : $f_Y(y) = P(Y = y) = P(X \in A_Y) = \sum_{x \in A_y} f_X(x)$, para $y \in D_Y$
- **Caso 2 : Variável Aleatória Qualquer** :
  – **Função de Distribuição de Y** : $F_Y(y) = P(Y \leq y) = P(A_y^*)$ onde $A_y^* = \{x : \ \psi(x) \leq y\}$
- **Condição suficiente para que uma variável aleatória $Y = \psi(X)$ seja também contínua** : Seja $Y = \psi(X)$, onde X é uma variável aleatória contínua e ψ é uma função real de variável real com domínio em A; supõe-se que este domínio contém o conjunto dos valores assumidos por X com densidade positiva. Se ψ tem derivada não nula em todos os pontos do seu domínio e é estritamente monótona em A, então Y é uma variável aleatória contínua.

3.5 Variáveis Aleatórias Bidimensionais

- **Variável Aleatória Bidimensional** : Uma variável aleatória bidimensional, (X,Y), é uma função com domínio Ω e com contradomínio $\mathbb{R}^2$. Assim,
$$(X,Y):\ \omega \in \Omega \to \big(X(\omega),Y(\omega)\big) \in \mathbb{R}^2$$
- **Função de Distribuição Conjunta** : Seja (X,Y) uma variável aleatória bidimensional. A função real de duas variáveis reais, $F_{X,Y}$, com domínio $\mathbb{R}^2$, definida por
$$F_{X,Y}(x,y) = P(X \leq x, Y \leq y),$$
é a função de distribuição de (X,Y), ou função de distribuição conjunta das variáveis X e Y.
- **Propriedades da Função de Distribuição Conjunta** :
  1) $0 \leq F(x,y) \leq 1$
  2) F é não decrescente, separadamente, em relação a x e em relação a y :
  $$\Delta x > 0 \Rightarrow F(x,y) \leq F(x+\Delta x, y) \quad \text{e} \quad \Delta y > 0 \Rightarrow F(x,y) \leq F(x, y+\Delta y)$$
  3) $F(-\infty,y) = F(x,-\infty) = 0$ e $F(+\infty,+\infty) = 1$
  4) $P(I) = P(x_1 < x \leq x_2, y_1 < y \leq y_2) = F(x_2,y_2) - F(x_1,y_2) - F(x_2,y_1) + F(x_1,y_1)$
  5) F é contínua à direita, separadamente, em relação a x e em relação a y :
  $$F(x+0,y) = F(x,y), \quad F(x,y+0) = F(x,y)$$
- **Função de Distribuição Marginal** : A função F_X definida por $F_X(x) = F_{X,Y}(x,+\infty)$ designa-se por função de distribuição marginal da variável aleatória X. A função F_Y dada por $F_Y(y) = F_{X,Y}(+\infty,y)$ é a função de distribuição marginal da variável aleatória Y.
- **Variáveis Aleatórias Independentes** : Considere-se uma variável aleatória bidimensional (X,Y). Sejam B_1 e B_2 dois acontecimentos quaisquer tais que B_1 envove apenas X, e B_2 refere-se apenas a Y. As variáveis aleatórias X e Y são independentes se e só se
$$P(X \in B_1, Y \in B_2) = P(X \in B_1)P(Y \in B_2)$$
$$\text{ou então} \quad F_{X,Y}(x,y) = F_X(x)F_Y(y)$$
- **Teorema** : Se X e Y são variáveis aleatórias independentes, e se ψ e φ são duas funções, então as variáveis aleatórias $U = \psi(X)$ e $V = \varphi(Y)$ são também independentes.

3.6 Variáveis Bidimensionais Discretas

- **Variável Aleatória Bidimensional Discreta** : (X,Y) é variável aleatória bidimensional discreta se e só se X e Y são variáveis aleatórias discretas.
- **Conjunto dos Pontos de Descontinuidade** : $D = \{(x,y) : P(X=x, Y=y) > 0\}$
- **Função Probabilidade Conjunta** : Seja (X,Y) é variável aleatória bidimensional discreta. A função real de duas variáveis reais, $f_{X,Y}$, com domínio $\mathbb{R}^2$, definida por
$$f_{X,Y}(x,y) = P(X=x, Y=y)$$
é a função probabilidade de (X,Y), ou função probabilidade conjunta das variáveis X e Y.
  → $f(x,y) = \begin{cases} P(X=x,Y=y) > 0 & se \quad (x,y) \in D \\ P(X=x,Y=y) = 0 & se \quad (x,y) \notin D \end{cases}$
  → $\sum_{(x,y)\in D} f(x,y) = 1$; $P[(X,Y) \in B] = \sum_{(x,y)\in B\cap D} f(x,y)$
  → $F(x,y) = P(X \leq x, Y \leq y) = \sum_{x_i \leq x} \sum_{y_i \leq y} f(x_i,y_i)$
- **Função Probabilidade Marginal** : Considere-se uma variável bidimensional discreta, (X,Y). Seja F_X a função de distribuição marginal de X, e
$$D_X = \{x : P(X=x) = F_X(x) - F_X(x-0) > 0\}$$
o conjunto dos pontos de descontinuidade de F_X.
A função probabilidade marginal de X, f_X, é definida da seguinte maneira:
$$f_X(x) = \begin{cases} P(X=x) > 0 & se \quad x \in D_X \\ P(X=x) = 0 & se \quad x \notin D_X \end{cases}$$
Do mesmo modo se define a função probabilidade marginal de Y, f_Y,
$$f_Y(y) = \begin{cases} P(Y=y) > 0 & se \quad y \in D_Y \\ P(Y=y) = 0 & se \quad y \notin D_Y \end{cases}$$
onde $D_Y = \{y : P(Y=y) = F_Y(y) - F_Y(y-0) > 0\}$

→ A função probabilidade marginal de X obtém-se somando, para cada x, a função probabilidade conjunta para todos os y :

$$f_X(x) = P(X = x) = \sum_{y \in D_Y} P(X = x, Y = y) = \sum_{y \in D_Y} f_{X,Y}(x, y)$$

→ idem para Y

- **Teorema** : As variáveis aleatórias discretas, X e Y, são independentes se e só se
$$f_{X,Y}(x, y) = f_X(x) f_Y(y)$$
isto é, se e só se a função probabilidade conjunta for igual ao produto das funções probabilidade marginais.
- **Função Probabilidade Condicionada** : Seja $f_{X,Y}$ a função probabilidade conjunta de X e Y. A função probabilidade de X condicionada pela realização do acontecimento $\{Y = y\}$, com $P(Y = y) > 0$, é dada por:
$$f_{X|Y=y}(x) = P(X = x|Y = y) = \frac{P(X=x,Y=y)}{P(Y=y)} = \frac{f_{X,Y}(x,y)}{f_Y(y)} \quad (y \text{ fixo})$$
De modo análogo, a função probabilidade de Y condicionada pela realização do acontecimento $\{X = x\}$, com $P(X = x) > 0$, é dada por:
$$f_{Y|X=x}(y) = P(Y = y|X = x) = \frac{P(X=x,Y=y)}{P(X=x)} = \frac{f_{X,Y}(x,y)}{f_X(x)} \quad (x \text{ fixo})$$

→ $\sum_{x \in D_X} f_{X|Y=y}(x) = 1$ e $\sum_{y \in D_Y} f_{X|Y=y}(y) = 1$

→ Se X e Y são independentes, então

$$f_{X|Y=y}(x) = f_X(x), \text{ se } f_Y(y) > 0 \quad ; \quad f_{Y|X=x}(y) = f_Y(y), \text{ se } f_X(x) > 0$$

3.7 Variáveis Bidimensionais Contínuas

- **Variável Aleatória Bidimensional Contínua** : A variável aleatória (X, Y), com função de distribuição $F_{X,Y}$, é uma variável aleatória bidimensional contínua se e só se existe uma função real de duas variáveis reais, não negativa, $f_{X,Y}$, tal que
$$F_{X,Y}(x, y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f_{X,Y}(u, v) du dv, \quad \forall (x, y) \in \mathbb{R}^2$$
- **Função Densidade Conjunta** : A função $f_{X,Y}$ (introduzida na definição anterior) chama-se função densidade de (X, Y) ou função densidade conjunta de X e Y

→ $F_{X,Y}(+\infty, +\infty) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = 1$

→ Se a função densidade $f(x, y)$ for contínua no ponto (x, y), então $f(x, y) = \frac{\partial^2 F(x,y)}{\partial x \partial y}$

- **Funções Densidade Marginais** : A função densidade marginal de X, f_X, é dada por
$$f_X(x) = \frac{dF_X(x)}{dx} = \int_{-\infty}^{+\infty} f_{X,Y}(x, y) dy$$
De forma análoga,
$$f_Y(y) = \frac{dF_Y(y)}{dy} = \int_{-\infty}^{+\infty} f_{X,Y}(x, y) dx$$
é a função densidade marginal de Y, f_Y.
- **Teorema** : As variáveis aleatórias X e Y dizem-se independentes se e só se
$$f_{X,Y}(x, y) = f_X(x) f_Y(y)$$
isto é, se e só se a função densidade conjunta for igual ao produto das funções densidade marginais.
- **Função Densidade Condicionada** : Seja $f_{X,Y}(x, y)$ a função densidade conjunta de X e Y. A função densidade de X condicionada por $Y = y$, com $f_Y(y) > 0$, é definida da seguinte maneira:
$$f_{X|Y=y}(x) = \frac{f_{X,Y}(x,y)}{f_Y(y)} \quad (y \text{ fixo})$$
A função densidade de Y condicionada por $X = x$, com $f_X(x) > 0$, é dada por:
$$f_{Y|X=x}(y) = \frac{f_{X,Y}(x,y)}{f_X(x)} \quad (x \text{ fixo})$$

→ $\int_{-\infty}^{+\infty} f_{X|Y=y}(x) dx = 1$

→ $P(a < X < b|Y = c) = \int_a^b f_{X|Y=c}(x) dx = \frac{\int_a^b f_{X,Y}(x,c) dx}{f_Y(c)}$

- → Factorização da Funç. de Dens. Conj. : $f_{X,Y}(x,y) = f_X(x) f_{Y|X=x}(y) = f_Y(y) f_{X|Y=y}(x)$
- → Fórmula de Bayes para Densidades :

$$f_{X|Y=y}(x) = \frac{f_X(x) f_{Y|X=x}(y)}{f_Y(y)} = \frac{f_X(x) f_{Y|X=x}(y)}{\int_{-\infty}^{+\infty} f_X(x) f_{Y|X=x}(y) dx}, \text{ com } f_Y(y) > 0$$

3.8 Distribuições Multidimensionais

- **Vectores Aleatórios** : Variáveis Aleatórias k-dimensionais (com $k \geq 2$)
- **Função de Distribuição** : A função de distribuição de $X = (X_1, \dots, X_k)$, F_X, é a função real de k variáveis reais definida por

$$F_X(x_1, \dots, x_k) = P(X_1 \leq x_1, \dots, X_k \leq x_k)$$

  - → F_X é não decrescente e contínua à direita em relação a cada variável
  - → $F_X(-\infty, x_2, \dots, x_k) = \cdots = F_X(x_1, x_2, \dots, -\infty) = 0 \quad ; \quad F_X(+\infty, +\infty, \dots, +\infty) = 1$
  - → A noção de <u>função de distribuição marginal</u> abrange agora não só as situações unidimensionais mas também todos os possíveis subconjuntos das k variáveis.
    - – Exemplo 1 : Seja $U = (X_1, \dots, X_k)$, com $m < k$, então

$$F_U(u) = F_X(x_1, \dots, x_m, +\infty, \dots, +\infty)$$

    - – Exemplo 2 : $F_X(x_1, +\infty, \dots, +\infty) = P(X_1 \leq x_1) = F_{X_1}(x_1)$
- **Vectores Aleatórios Independentes** : Os vectors aleatórios U e V dizem-se independentes se e só se $F_X(x_1, \dots, x_k) = F_{U,V}(u, v) = F_U(u) F_V(v)$
ou seja, se e só se a função de distribuição conjunta é igual ao produto das funções de distribuição marginais.
- **Variáveis Aleatórias Independentes** : As variáveis aleatórias $X_1, \dots, X_k$ dizem-se independentes se e só se $F_X(x_1, \dots, x_k) = F_{X_1}(x_1) \dots F_{X_k}(x_k)$
- **Variável Aleatória k-dimensional Discreta** : $X = (X_1, \dots, X_k)$ é variável aleatória k-dimensional discreta se e só se as variáveis aleatórias X_j, com $j = 1, \dots, k$, são discretas.
  - → As restantes propriedades das variáveis aleatórias discretas são generalizáveis.
- **Variável Aleatória k-dimensional Contínua** : A variável aleatória k-dimensional $X = (X_1, \dots, X_k)$, onde F_X é a respectiva função de distribuição, é contínua se e só se existe uma função real de k variáveis reais, não negativa, f_X, a verificar

$$F_X(X_1, \dots, X_k) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_k} f_X(u_1, \dots, u_k) du_1 \dots du_k$$

  - → As restantes propriedades das variáveis aleatórias discretas são generalizáveis.

<u>Capítulo 4 – Distribuições Discretas</u>

4.2 Valor Esperado

- **Valor Esperado de uma Variável Aleatória Discreta** : Se X é uma variável aleatória discreta com função probabilidade f_X, a expressão

$$\mu = E(X) = \sum_{x \in D} x f_X(x)$$

quando

$$\sum_{x \in D} |x| f_X(x) < +\infty \quad \textbf{(4.2)}$$

define o valor esperado, média ou esperança matemática de X.
  - → centro de gravidade da distribuição
  - → parâmetro de localização
- **Valor Esperado de uma Função de uma Variável Aleatória Discreta** : Se X é uma variável aleatória discreta com função probabilidade f_X, a expressão

$$E\big(\psi(X)\big) = \sum_{x \in D} \psi(x) f_X(x)$$

é o valor esperado de $\psi(x)$, impondo-lhe uma condição semelhante a **(4.2)**
  - → $\sum_{y \in D_Y} y f_Y(y) = \sum_{x \in D_X} \psi(x) f_X(x) <=> E(Y) = E\big(\psi(X)\big)$

- **Propriedades do Valor Esperado** :
  1) Se c é uma constante, então $E(c) = c$
  2) Se c é uma constante e se existe $E[\psi(X)]$, então $E[c\psi(X)] = cE[\psi(X)]$
  3) Se existirem $E[\psi_1(X)]$ e $E[\psi_2(X)]$, então $E[\psi_1(X) + \psi_2(X)] = E[\psi_1(X)] + E[\psi_2(X)]$
- **Outra Propriedade** : O valor esperado de uma combinação linear de funções de variável aleatória X é a respectiva combinação linear dos valores esperados:

$$E\left[\sum_{j=1}^{n} c_j \psi_j(X)\right] = \sum_{j=1}^{n} c_j E\left[\psi_j(X)\right]$$

- **Resultados Importantes** :
  - Se X e Y são duas variáveis aleatórias discretas, então $E(X + Y) = E(X) + E(Y)$
  - Se X e Y são duas variáveis aleatórias discretas independentes, então $E(XY) = E(X)E(Y)$

4.3 Momentos

- **Momento de Ordem k em Relação à Origem** : O valor esperado,

$$\mu'_k = E\left(X^k\right) = \sum_{x \in D} x^k f_X(x)$$

se existir, é o momento de ordem k em relação à origem, ou momento ordinário de ordem k da variável aleatória X.

→ Nota : $\mu'_1 = E(X) = \sum_{x \in D} x f_X(x)$ → valor esperado

- **Teorema** : Se existe o momento ordinário de ordem k de uma variável aleatória, então também existe o momento ordinário de ordem s, com $s < k$, da mesma variável aleatória.
- **Momento de Ordem k em Relação à Média** : O valor esperado,

$$\mu_k = E[(X - \mu)^k] = \sum_{x \in D} (X - \mu)^k f_X(x)$$

se existir, é o momento de ordem k em relação à média, ou momento central de ordem k da variável aleatória X.

- **Variância** : Designa-se por variância de X o valor esperado, se existir, de $(X - \mu)^k$,

$$Var(X) = \sigma_X^2 = E[(X - \mu)^2] = \sum_{x \in D} (X - \mu)^2 f_X(x)$$

→ $Var(X) \geq 0$; $Var(X) = 0 <=> P(X = c) = 1$

→ parâmetro de dispersão da distribuição de X em torno da média

- **Propriedades da Variância** :
  1) Se c é uma constante, $Var(c) = 0$
  2) Se c é uma constante, $Var(cX) = c^2 Var(X)$
  3) Se c e d são constantes, $Var(cX + d) = c^2 Var(X)$
- **Fórmula mais operacional para o cálculo da variância** : $Var(X) = E(X^2) - \mu^2$
- **Desvio Padrão** : Ao valor positivo da raíz quadrada da variância, $\sigma_X = +\sqrt{Var(X)}$, chama-se desvio padrão da variável aleatória X.

  → medida de dispersão absoluta
- **Coeficiente de Variação** : Considere-se uma variável aleatória X com suporte contido no cpnjunto dos números reais positivos. O coeficiente de variação de X é igual, se existir, ao quociente entre o desvio padrão e a média, $CV = \frac{\sigma}{\mu}$

  → medida de dispersão relativa
- **Coeficiente de Assimetria** : O quociente $\gamma_1 = \frac{E(X-\mu)^3}{\sigma^3} = \frac{\mu_3}{\sigma^3}$, se exisir, é designado por coeficiente de assimetria da variável aleatória X.
  - $\gamma_1 = 0$ sse a distribuição é simetrica
  - $\gamma_1 > 0$ sse os desvios positivos dominam
  - $\gamma_1 < 0$ sse os desvios negativos dominam
- **Variância de Variáveis Aleatórias Independentes** : Se X e Y são duas variáveis aleatórias discretas independentes, então $Var(X + Y) + Var(X) + Var(Y)$

4.4 Funções Geradoras

- **Sucessões Importantes** :
  - Sucessão das probabilidades : $p_0, p_1, p_2, \ldots$
  - Sucessão dos momentos em relação à origem : $\mu'_1 = \mu, \mu'_2, \ldots, \mu'_k, \ldots$
  - Sucessão dos momentos em relação à média : $\mu_1 = 0, \mu_2, \ldots, \mu_k, \ldots$
- **Função Geradora das Probabilidades** : Seja X uma variável aleatória discreta assumindo, com probabilidade positiva, os valores do conjunto $D = \{0,1,2,\ldots\}$. Seja
$$p_x = P(X = x) \text{ com } x = 0,1,2,\ldots \text{ a verificar } \sum_{x=0}^{+\infty} p_x = 1$$
A função geradora das probabilidades de X é a função real de variável real, Π_X, definida por $$\Pi_X(z) = E(z^X) = \sum_{x=0}^{+\infty} p_x z^x$$
$\rightarrow$ $p_x = P(X = x) = \frac{\Pi^{(x)}(0)}{x!}$
$\rightarrow$ $E(X) = \Pi'(1) = \sum_{x=1}^{+\infty} x p_x$
$\rightarrow$ $Var(X) = \Pi''(1) + \Pi'(1) - [\Pi'(1)]^2$
- **Função Geradora dos Momentos** : Seja X uma variável aleatória discreta com função probabilidade f_X. Se existir um número real $s_0 > 0$, tal que
$$E(e^{sX}) = \sum_{x \in D} e^{sx} f_X(x)$$
existe para $-s_0 < s < s_0$, então
$$M_X(s) = E(e^{sX})$$
define a função geradora dos momentos de X, M_X.
$\rightarrow$ $M'(0) = \sum_{x \in D} x f_X(x) = E(X)$
- **Teorema** : A função geradora dos momentos, se existir numa vizinhança de 0, determina univocamente a função de distribuição ; inversamente, se existir, a função geradora dos momentos é única.
Formalizando: Se para cada uma das variáveis aleatórias X e Y, existe função geradora dos momentos, e se $M_X(s) = M_Y(s)$ para algum intervalo $-s_0 < s < s_0$, então X e Y possuem a mesma função de distribuição (e a mesma probabilidade, no caso de as variáveis aleatórias serem discretas).
- **Relação entre a Função Geradora dos Momentos com a Função Geradora das Probabilidades** : $M(s) = E(e^{sX}) = \Pi(e^s) <=> \Pi(z) = M[\ln(z)]$
- **Teorema** : Se X e Y são variáveis aleatórias discretas independentes, com funções geradoras dos momentos $M_X(s)$ e $M_Y(s)$, então a função geradora dos momentos $X + Y$ existe, e é dada pela expressão:
$$M_{X+Y}(s) = M_X(s) M_Y(s)$$

4.5 Distribuição Uniforme

- **Distribuição Uniforme** : Seja X uma variável aleatória discreta com $D = \{x_1, x_2, \ldots, x_n\}$. Diz-se que X segue uma distribuição uniforme nos n pontos x_j, com $j = 1,2,\ldots,n$, se e só se a função probabilidade é dada por
$$f_X(x) = \begin{cases} \frac{1}{n} & se\ x \in D \\ 0 & se\ x \notin D \end{cases}$$
$\rightarrow$ a função probabilidade é constante para um número finito de pontos
$\rightarrow$ a distribuição é simetrica
- **Principais características da Distribuição Uniforme** :
  - Valor Esperado : $E(X) = \mu = \frac{1}{n}\sum_{j=1}^{n} x_j$
  - Momento de Ordem k em relação à Origem : $E\left(X^k\right) = \frac{1}{n}\sum_{j=1}^{n} x_j^k$
  - Variância : $Var(X) = \frac{1}{n}\sum_{j=1}^{n}\left(x_j - \mu\right)^2$
  - Função Geradora dos Momentos : $M(s) = \frac{1}{n}\sum_{j=1}^{n} e^{sx}$

- Casos Particulares Importantes :
  1) Ex: Lançamento de um dado perfeito : $D = \{1,2,\dots,n\}$
  2) Ex: Obtenção de um dos 10 dígitos : $D = \{0,1,2,\dots,m\}$

4.6 Distribuição de Bernoulli / Distribuição Binomial

- Prova de Bernoulli : Experiência aleatória em que se observa a realização (sucesso; $x = 1$) ou a não realização (insucesso; $x = 0$) de determinado acontecimento A com probabilidade $P(A) = \theta$.

$$f_X(x|\theta) = \begin{cases} 1-\theta & se\ x = 1 \\ \theta & se\ x = 0 \\ 0 & para\ outros\ x \end{cases}$$

- Distribuição de Bernoulli : Se a variável aleatória X tem função probabilidade dada por

$$f_X(x|\theta) = \theta^x(1-\theta)^{1-x} \text{ com } x = \{0,1\} \text{ para } 0 < \theta < 1,$$

  diz-se que X tem distribuição de Bernoulli. Simbolicamente, $X \sim B(1;\theta)$
- Principais características da Distribuição de Bernoulli :
  - Valor Esperado : $E(X) = \mu = \theta$
  - Momento de Ordem 2 em relação à Origem : $E(X^2) = \theta$
  - Variância : $Var(X) = \theta(1-\theta)$
  - Desvio Padrão : $\sigma = \sqrt{\theta(1-\theta)}$
  - Coeficiente de Assimetria : $\gamma_1 = \frac{1-2\theta}{\sigma}$
  - Função Geradora dos Momentos : $M(s) = E(e^{sX}) = (1-\theta) + \theta e^s$
- Sucessão de Provas de Bernoulli Independentes : Sucessão de experiências aleatórias independentes em cada uma das quais se observa a realização (sucesso; $x = 1$) ou a não realização (insucesso; $x = 0$) de determinado acontecimento A com probabilidade $P(A) = \theta$, constante de prova para prova.
- Distribuição Binomial : Se a variável aleatória X tem função probabilidade dada por

$$f_X(x|\theta) = \begin{cases} \binom{n}{x}\theta^x(1-\theta)^{1-x} & se\ x = 0,1,2,\dots,n \\ 0 & para\ outros\ x \end{cases}, \text{ para } 0 < \theta < 1,$$

  diz-se que X tem distribuição binomial. Simbolicamente, $X \sim B(n;\theta)$
- Principais características da Distribuição Binomial :
  - Valor Esperado : $E(X) = \mu = n\theta$
  - Momento de Ordem 2 em relação à Origem : $E(X^2) = n(n-1)\theta^2 + n\theta$
  - Variância : $Var(X) = n\theta(1-\theta)$
  - Desvio Padrão : $\sigma = \sqrt{n\theta(1-\theta)}$
  - Coeficiente de Assimetria : $\gamma_1 = \frac{1-2\theta}{\sigma}$
  - Função Geradora dos Momentos : $M(s) = E(e^{sX}) = [(1-\theta) + \theta e^s]^n$
- Teorema : Sejam Y_1 e Y_2 duas variáveis aleatórias independentes. Então

$$Y_1 \sim B(n_1;\theta),\ \ Y_2 \sim B(n_2;\theta) \ \Rightarrow \ Y = Y_1 + Y_2 \sim B(n;\theta)$$

  onde $n = n_1 + n_2$.

4.7 Distribuição Geométrica / Distribuição Binomial Negativa

- Nova perspectiva : Realizar provas de Bernoulli até que se dê um sucesso. O número de provas é que passa a ser aleatório.
- Distribuição Geométrica : Quando a variável aleatória X tem função probabilidade dada por

$$f_X(x) = \begin{cases} (1-\theta)^{x-1}\theta & se\ x = 1,2,\dots \\ 0 & para\ outros\ x \end{cases}$$

  diz-se que tem distribuição geométrica ou distribuição de Pascal.

  $\rightarrow$ A sucessão das probabilidades forma uma progressão geométrica de razão $1-\theta$.

- **Propriedades da Série Geométrica de razão z** :
  a) $S(z) = \sum_{k=0}^{\infty} z^k = \frac{1}{1-z}$ (série absolutamente convergente para $|z| < 1$)
  b) $S'(z) = \sum_{k=1}^{\infty} kz^{k-1} = \frac{1}{(1-z)^2}$
  c) $S''(z) = \sum_{k=2}^{\infty} k(k-1)z^{k-2} = \frac{1}{(1-z)^3}$
- **Principais características da Distribuição Geométrica** :
  - $\sum_{k=1}^{\infty} f_X(x|\theta) = 1$
  - Valor Esperado : $E(X) = \mu = \frac{1}{\theta}$
  - Variância : $Var(X) = \sigma^2 = \frac{1-\theta}{\sigma^2}$
  - Desvio Padrão : $\sigma = \frac{\sqrt{1-\theta}}{\sigma}$
  - Coeficiente de Assimetria : $\gamma_1 = \frac{2-\theta}{\sqrt{1-\theta}} > 0$ (enviesamento positivo)
  - Função Geradora dos Momentos : $M(s) = E(e^{sX}) = \frac{\theta e^s}{1-(1-\theta)e^s}$, com $(1-\theta)e^s < 1$
- **Algumas Propriedades** :
  → $P(a \leq X \leq b) = (1-\theta)^{a-1} - (1-\theta)^b$
  → Se $b \to +\infty$, $P(X \geq a) = (1-\theta)^{a-1}$
  → Se $a = 1$, $P(X \leq b) = 1 - (1-\theta)^b$
- **Função de Distribuição de uma Variável Aleatória com Distribuição Geométrica** :

$$F_X(x) = \begin{cases} 0 & se\ x < 1 \\ 1-(1-\theta)^k & se\ k \leq x \leq k+1 \end{cases}$$

  para $k = 1,2, \ldots$
- **Propriedade da "Falta de Memória"** : $P(X > s | X > t) = P(X > s-t)$
- **Nova perspectiva** : Realizar provas de Bernoulli independentes <u>até</u> obter r sucessos.
- Distribuição Binomial Negativa : Uma variável aleatória X com função probabilidade dada por

$$f_X(x|\theta) = \begin{cases} \binom{x-1}{r-1} \theta^r (1-\theta)^{x-r} & se\ x = r, r+1, r+2, \ldots \\ 0 & para\ outros\ x \end{cases}, \text{ para } 0 < \theta < 1$$

  diz-se que tem distribuição binomial negativa. Simbolicamente, $X \sim BN(r;\theta)$.
- **Principais características da Distribuição Binomial Negativa** :
  - Valor Esperado : $E(X) = \mu = \frac{r}{\theta}$
  - Variância : $Var(X) = \sigma^2 = \frac{r(1-\theta)}{\sigma^2}$
  - Desvio Padrão : $\sigma = \frac{\sqrt{r(1-\theta)}}{\sigma}$
  - Coeficiente de Assimetria : $\gamma_1 = \frac{r\theta}{\sqrt{1-\theta}} > 0$
  - Função Geradora dos Momentos : $M(s) = E(e^{sX}) = \left(\frac{\theta e^s}{1-(1-\theta)e^s}\right)^r$, com $(1-\theta)e^s < 1$
- **Teorema** : Sejam Y_1 e Y_2 duas variáveis aleatórias independentes. Então

$$Y_1 \sim BN(n_1;\theta),\ \ Y_2 \sim BN(n_2;\theta) \ \Rightarrow \ Y = Y_1 + Y_2 \sim BN(n;\theta)$$

  onde $n = n_1 + n_2$.
- **Corolário** : $X_i \sim BN(1;\theta)$ se $i = 1,2,\ldots,r$, independentes $\Rightarrow \sum_{i=1}^{r} X_i \sim BN(r;\theta)$
  → <u>Interpretação do Corolário</u> : Realizar provas até aparecerem r sucessos é o mesmo que realizar provas até aparecer o primeiro sucesso, continuar até aparecer o segundo, e assim sucessivamente, até aparecer o r-ésimo sucesso.

4.8 Distribuição Hipergeométrica

- **Distribuição Hipergeométrica** : Se a variável aleatória X tem função probabilidade dada por

$$f_X(x|N,M,n) = \begin{cases} \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} & se\ x \in D \\ 0 & se\ x \notin D \end{cases}$$

com $D = \{x : max\{0, n-(N-M) \leq x \leq min\{n,M\}, x\ inteiro\}\}$

diz-se que tem distribuição hipergeométrica. Simbolicamente $X \sim H(N,M,n)$

- **Principais características da Distribuição Hipergeométrica** : $\left(\theta = \frac{M}{N}\right)$
  - $\sum_{x=0}^{n} f_X(x|N,M,n) = 1$
  - Valor Esperado : $E(X) = n\frac{M}{N} = n\theta$
  - Variância : $Var(X) = n \cdot \frac{M}{N} \cdot \frac{N-M}{N} \cdot \frac{N-n}{N-1} = n\theta(1-\theta)\frac{N-n}{N-1}$

4.9 Distribuição de Poisson

- **Processo de Poisson** : Suponha-se que se procede à contagem do número de acontecimentos (chegada de doentes, "chegada" de avarias, chegada de navios, etc) ocorridos ao longo do tempo. Tem-se um processo de Poisson com parâmetro $\lambda > 0$ quando se verificam as seguintes condições:
  - → O número de acontecimentos que ocorrem em dois intervalos disjuntos são independentes;
  - → A probabilidade de ocorrer exactamente um acontecimento em qualquer intervalo de amplitude Δt arbirtariamente pequena é aproximadamente $\lambda\Delta t$;
  - → A probabilidade de ocorrerem dois ou mais acontecimentos em qualquer intervalo de amplitude Δt arbirtariamente pequena é aproximadamente igual a zero.
- **Distribuição de Poisson** : Uma variável aleatória X com função probabilidade dada por

$$f_X(x|\lambda) = \begin{cases} \frac{e^{-\lambda}\lambda^x}{x!} & se\ x = 0,1,2,\ldots \\ 0 & para\ outros\ x \end{cases}, \text{ com } \lambda > 0,$$

diz-se que tem distribuição de Poisson. Simbolicamente, $X \sim Po(\lambda)$

- **Principais características da Distribuição de Poisson** :
  - $\sum_{x=0}^{+\infty} \frac{e^{-\lambda}\lambda^x}{x!} = 1$
  - Função Geradora dos Momentos : $M(s) = E(e^{sX}) = exp[\lambda(e^s - 1)]$
  - Valor Esperado : $E(X) = \lambda$
  - Variância : $Var(X) = \lambda$
  - Desvio Padrão : $\sigma = \sqrt{\lambda}$
  - Coeficiente de Assimetria : $\gamma_1 = \lambda^{-1/2}$
- **Distribuição de Poisson e Intervalos de Tempo** :

$$f_X(x|\lambda t) = P[X(t) = x|\lambda t] = \frac{e^{-\lambda t}(\lambda t)^x}{x!}, se\ x = 0,1,2,\ldots, \text{ para } \lambda > 0$$

- **Teorema** : Sejam X_1 e X_2 duas variáveis aleatórias independentes. Então

$$X_1 \sim Po(\lambda_1),\ \ X_2 \sim Po(\lambda_2) \ \Rightarrow\ X = X_1 + X_2 \sim Po(\lambda)$$

onde $\lambda = \lambda_1 + \lambda_2$.

- **Lei dos Acontecimentos Raros** : A génese da distribuição de Poisson, a partir do processo de Poisson, mostra que, quando $\theta = \frac{\lambda}{n} \to 0$ mantendo-se fixo $n\theta = \lambda$, a binomial tende para Poisson, $\lim_{n \to +\infty} \binom{n}{x}\left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} = \frac{e^{-n\theta}(n\theta)^x}{x!}$

4.10 Valor Esperado e Momentos de Variáveis Aleatórias Bidimensionais Discretas

- **Valor Esperado de Função de Variável Aleatória Bidimensional** : Se (X, Y) é uma variável aleatória bidimensional discreta com função probabilidade $f_{X,Y}$, e se ψ é uma função de (X, Y), a expressão
$$E[\psi(X,Y)] = \sum_{(x,y)\in D} \psi(x,y) f_{X,Y}(x,y)$$
é o valor esperado de $\psi(x, y)$. Tal como no caso unidimensional, tem de se verificar
$$\sum_{(x,y)\in D} |\psi(x,y)| f_{X,Y}(x,y) < +\infty$$
- **Teorema** : Se (X, Y) for uma variável aleatória bidimensional discreta com função probabilidade $f_{X,Y}$, e se existirem $E(X)$ e $E(Y)$, então
$$E(X+Y) = E(X) + E(Y)$$
- **Teorema** : Se X e Y forem variáveis aleatórias discretas e independentes, com função probabilidade conjunta $f_{X,Y}$, e se existirem $E(X)$ e $E(Y)$, então
$$E(XY) = E(X)E(Y)$$
- **Momentos de Ordem $r + s$ em Relação à Origem** : Seja (X, Y) uma variável aleatória bidimensional discreta. O valor esperado,
$$\mu'_{rs} = E(X^r Y^s) = \sum_{(x,y)\in D} x^r y^s f_{X,Y}(x,y)$$
define, se existir, um momento de ordem $r + s$ em relação à origem (ordinário) de variável aleatória (X, Y).
  - $\rightarrow$ Momentos de primeira ordem :
    - $r = 1, s = 0 \Longrightarrow \mu'_{10} = E(X) = \mu_X$
    - $r = 0, s = 1 \Longrightarrow \mu'_{01} = E(Y) = \mu_Y$
  - $\rightarrow$ Momentos de segunda ordem :
    - $r = 2, s = 0 \Longrightarrow \mu'_{20} = E(X^2)$
    - $r = 1, s = 1 \Longrightarrow \mu'_{11} = E(XY)$
    - $r = 0, s = 2 \Longrightarrow \mu'_{02} = E(Y^2)$
- **Momentos de Ordem $r + s$ em Relação à Média** : Seja (X, Y) uma variável aleatória bidimensional discreta. O valor esperado,
$$\mu_{rs} = E[(X-\mu)^r (Y-\mu)^s] = \sum_{(x,y)\in D} (x-\mu_X)^r (y-\mu_Y)^s f_{X,Y}(x,y)$$
define, se existir, um momento de ordem $r + s$ em relação à média (central) de variável aleatória (X, Y).
  - $\rightarrow$ Momentos de primeira ordem (são nulos) :
    - $r = 1, s = 0 \Longrightarrow \mu_{10} = 0$
    - $r = 0, s = 1 \Longrightarrow \mu_{01} = 0$
  - $\rightarrow$ Momentos de segunda ordem :
    - $r = 2, s = 0 \Longrightarrow \mu_{20} = E[(X-\mu)^2] = Var(X) = \sigma_X^2$
    - $r = 1, s = 1 \Longrightarrow \mu_{11} = E[(X-\mu)(Y-\mu)]$
    - $r = 0, s = 2 \Longrightarrow \mu_{02} = E[(Y-\mu)^2] = Var(Y) = \sigma_Y^2$
- **Covariância** : A covariância das variáveis aleatórias X e Y é
$$Cov(X,Y) = \sigma_{X,Y} = E[(X-\mu_X)(Y-\mu_Y)]\,,$$
se este valor esperado existir.
  - $\rightarrow$ $Cov(X,Y) = E(XY) - E(X)E(Y)$
  - $\rightarrow$ $Cov(c,X) = E(cX) - cE(X) = 0$
- **Interpretação do Sinal e da Grandeza da Covariância** :
  - $\rightarrow$ Centro de gravidade da distribuição conjunta : (μ_X, μ_Y)
  - $\rightarrow$ $(X-\mu_X)(Y-\mu_Y) > 0$ → 1º e 3º quadrantes
  - $\rightarrow$ $(X-\mu_X)(Y-\mu_Y) < 0$ → 2º e 4º quadrantes
- **Coeficiente de Correlação** : O coeficiente de correlação entre as variáveis aleatórias X e Y é dado por
$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}} = \frac{\sigma_{X,Y}}{\sigma_X \sigma_Y}$$

- **Teorema** : Se duas variáveis aleatórias, X e Y, são independentes, então a respectiva covariância é igual a zero.
  $\rightarrow$ A independência implica não correlação, mas a recíproca não é verdadeira.
- **Teorema** : Se existem segundos momentos para as variáveis aleatórias X e Y, então
  $$Var(X \pm Y) = Var(X) \pm Cov(X,Y) + Var(Y)$$
  Em particular, se as variáveis são independentes,
  $$Var(X \pm Y) = Var(X) + Var(Y)$$
- **Valor Esperado Condicionado** : Seja a variável aleatória $Z = \psi(X,Y)$, função das variáveis aleatórias discretas X e Y. O valor esperado de $Z = \psi(X,Y)$ condicionado por $X = x$, é definido por
  $$E(Z|X=x) = E[\psi(X,Y)|X=x] = \sum_{y \in D_Y} \psi(x,y) f_{Y|X=x}(y)$$
  se se verificar a condição habitual de existência do valor esperado.
  De modo análogo se define
  $$E(Z|Y=y) = E[\psi(X,Y)|Y=y] = \sum_{x \in D_X} \psi(x,y) f_{X|Y=y}(x)$$
  $\rightarrow$ Se $\psi(x,y) = Y$, $E(Y|X=x) = \sum_{y \in D_Y} y f_{Y|X=x}(y)$
  $\rightarrow$ Se $\psi(x,y) = X$, $E(X|Y=y) = \sum_{x \in D_X} x f_{X|Y=y}(x)$
- **Teorema** : O valor esperado de $Z = \psi(X,Y)$, se existir, é igual ao valor esperado do seu valor esperado condicionado por X,
  $$E(Z) = E[E(Z|X)]$$
  $\rightarrow$ Se $Z = Y$, $E(Y) = E[E(Y|X)]$ → regra do valor esperado iterado (→ v.e. marginal)
- **Variância Condicionada** : A variância de Y condicionada por $X = x$ é definida por
  $$Var(Y|X=x) = E[\{Y - E(Y|X=x)\}^2|X=x] = \sum_{y \in D_Y} \{y - E(Y|X=x)\}^2 f_{Y|X=x}(y)$$
  Do mesmo modo se define a variância de X condicionada por $Y = y$,
  $$Var(X|Y=y) = E[\{X - E(X|Y=y)\}^2|Y=y] = \sum_{x \in D_X} \{x - E(X|Y=y)\}^2 f_{X|Y=y}(x)$$
  $\rightarrow$ $Var(Y|X=x) = E(Y^2|X=x) - E(Y|X=x)^2$
  $\rightarrow$ $Var(X|Y=y) = E(X^2|Y=y) - E(X|Y=y)^2$
- **Teorema** : Se existe $E(Y^2)$, então
  $$Var(Y) = Var[E(Y|X)] + E[Var(Y|X)] \quad \text{(Variância Marginal)}$$
- **Teorema** : A covariância entre X e Y é igual à covariância entre X e o valor esperado de Y condicionado por X,
  $$Cov(X,Y) = Cov[X, E(Y|X)]$$
- **Independência em Média** : A variável aleatória Y é independente em média da variável aleatória X se e só se
  $$E(Y|X=x) = E(Y), \text{ qualquer que seja } x$$
  A variável aleatória X é independente em média da variável aleatória Y se e só se
  $$E(X|Y=y) = E(X), \text{ qualquer que seja } y$$

4.11 Valor Esperado de Momentos de Variáveis Aleatórias k-Dimensionais Discretas

- **Momento de Ordem $r_1 + r_2 + \cdots + r_k$ em Relação à Média** :
  $$E[(X_1 - \mu_1)^{r_1}(X_2 - \mu_2)^{r_2} \ldots (X_k - \mu_k)^{r_k}]$$
- **Momentos Centrais de 2ª Ordem** :
  - Variância : $Var(X_i) = \sigma_{ii} = \sigma_i^2$
  - Covariância : $Cov(X_i, X_j) = \sigma_{ij} \quad (i \neq j)$
- **Matriz das Covariâncias** : $Cov(\boldsymbol{X}) = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1k} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2k} \\ \vdots & \vdots & & \vdots \\ \sigma_{k1} & \sigma_{k2} & \cdots & \sigma_{kk} \end{bmatrix}$
- **Coeficiente de Correlação entre X_i e X_j** : $\rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}\sigma_{jj}}} = \frac{\sigma_{ij}}{\sigma_i \sigma_j}$

- **Matriz das Correlações do Vector Aleatório $\boldsymbol{X}$** : $Corr(\boldsymbol{X}) = \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1k} \\ \rho_{21} & 1 & \cdots & \rho_{2k} \\ \vdots & \vdots & & \vdots \\ \rho_{k1} & \rho_{k2} & \cdots & 1 \end{bmatrix}$
- **Relação entre a Variância e a Matriz das Covariâncias** :
$$Var(Y) = Var(\boldsymbol{a}^T\boldsymbol{X}) = \boldsymbol{a}^T Cov(\boldsymbol{X})\boldsymbol{a}$$
- **Relação entre a Matriz das Covariâncias e a Matriz das Correlações** :
$$Cov(\boldsymbol{Y}) = Cov(\boldsymbol{A}X) = \boldsymbol{A}Cov(\boldsymbol{X})\boldsymbol{A}^T$$

➢ Nota: $\boldsymbol{a}$ e $\boldsymbol{A}$ são matrizes de coeficientes

4.12 Distribuição Multinomial

- **Distribuição Multinomial** : Se uma variável aleatória k-dimensional tem função probabilidade dada por
$$f_X(x_1, x_2, \dots, x_k) = \frac{n!}{x_1!, x_2!, \dots, x_k!(n - x_1 - x_2 - \dots - x_k)!}$$
$$\times\, \theta_1^{x_1}\theta_2^{x_2} \dots \theta_k^{x_k}(1 - \theta_1 - \theta_2 - \dots - \theta_k)^{n - x_1 - x_2 - \dots - x_k},$$
para as probabilidades positivas, diz-se que tem distribuição multinomial.
- **Principais características da Distribuição Multinomial** :
  - ➢ Valor Esperado : $E(X_i) = n\theta_i$
  - ➢ Variância : $Var(X_i) = n\theta_i(1 - \theta_i)$
  - ➢ Desvio Padrão : $\sigma = \sqrt{n\theta_i(1 - \theta_i)}$
  - ➢ Covariância : $Cov\left(X_i, X_j\right) = \sigma_{ij} = -n\theta_i\theta_j$
  - ➢ Coeficiente Correlação : $\rho_{ij} = -\sqrt{\frac{\theta_i\theta_j}{(1-\theta_i)(1-\theta_j)}}\quad (i \neq j)$

Capítulo 5 – Distribuições Contínuas

5.2 Valor Esperado

- **Valor Esperado de uma Variável Aleatória Discreta** : Se X é variável aleatória contínua com função densidade f_X, a expressão
$$E(X) = \int_{-\infty}^{+\infty} x f_X(x)dx$$
quando
$$\int_{-\infty}^{+\infty} |x| f_X(x)dx < +\infty \quad \textbf{(5.2)}$$
define o valor esperado, média ou esperança matemática de X.
  - → centro de gravidade da distribuição
  - → parâmetro de localização

- **Valor Esperado de uma Função de uma Variável Aleatória Discreta** : Se X é uma variável aleatória contínua com função densidade f_X, e ψ é uma função real de variável real. A expressão
$$E\big(\psi(X)\big) = \int_{-\infty}^{+\infty} \psi(x) f_X(x)dx$$
é o valor esperado de $\psi(x)$, impondo-lhe uma condição semelhante a **(5.2)**
- **Propriedades do Valor Esperado** : As propriedades do valor esperado para variáveis aleatórias contínuas são as mesmas do que as propriedades para variáveis aleatórias discretas. Apenas é necessário substituir as somas por integrais.

5.3 Momentos

- **Momento de Ordem k em Relação à Origem** : O valor esperado,

$$\mu_k' = E\left(X^k\right) = \int_{-\infty}^{+\infty} x^k f_X(x)dx$$

se existir, define o momento de ordem k em relação à origem, ou momento ordinário de ordem k da variável aleatória X.

Nota : $\mu_1' = E(X) = \int_{-\infty}^{+\infty} x f_X(x)dx$ → valor esperado

- **Teorema** : Se existe o momento ordinário de ordem k de uma variável aleatória, então também existe o momento ordinário de ordem s, com $s < k$, da mesma variável aleatória.
- **Momento de Ordem k em Relação à Média** : O valor esperado,

$$\mu_k = E[(X-\mu)^k] = \int_{-\infty}^{+\infty} (X-\mu)^k f_X(x)dx$$

se existir, é o momento de ordem k em relação à média, ou momento central de ordem k da variável aleatória X.

- **Variância** : A variância de uma variável aleatória contínua X é dada por

$$Var(X) = \sigma_X^2 = E[(X-\mu)^2] = \int_{-\infty}^{+\infty} (X-\mu)^2 f_X(x)dx$$

$\rightarrow$ parâmetro de dispersão da distribuição de X em torno da média

- **Propriedades da Variância** :
  - $\rightarrow$ $Var(cX) = c^2 Var(X)$
  - $\rightarrow$ $Var(X) = E(X^2) - \mu^2$ → fórmula mais operacional para o cálculo da variância
  - $\rightarrow$ $Var(X+Y) = Var(X) + Var(Y)$, se X e Y são independentes
- **Desvio Padrão** : $\sigma_X = +\sqrt{Var(X)}$ → medida de dispersão absoluta
- **Coeficiente de Variação** : $CV = \frac{\sigma}{\mu}$ → medida de dispersão relativa
- **Coeficiente de Assimetria** : $\gamma_1 = \frac{E(X-\mu)^3}{\sigma^3} = \frac{\mu_3}{\sigma^3}$
- **Coeficiente de Kurtosis** : O coeficiente de *kurtosis* de uma variável aleatória X, se existir, é definido por $\gamma_1 = \frac{\mu_4}{\sigma^4}$ → excesso de *kurtosis* : $\gamma_1 - 3$

5.4 Parâmetros da Ordem

- **Quantil** : Seja X uma variável aleatória contínua, e α, um número real a verificar $0 < \alpha < 1$. Quantil de ordem α, $\mathcal{E}_\alpha$, é um valor de X que satisfaz a condição

$$\int_{-\infty}^{\mathcal{E}_\alpha} f_X(x)dx = \alpha <=> F_X(\mathcal{E}_\alpha) = \alpha$$

  - ➢ medida de localização
  - $\rightarrow$ Mediana : $\alpha = 0{,}5$ → $\mathcal{E}_{0,5} = \mu_c = med(X)$ → $\int_{-\infty}^{\mathcal{E}_{0,5}} f_X(x)dx = 0{,}5$
  - $\rightarrow$ 1º Quartil : $\alpha = 0{,}25$ → $\mathcal{E}_{0,25}$ → $\int_{-\infty}^{\mathcal{E}_{0,25}} f_X(x)dx = 0{,}25$
  - $\rightarrow$ 3º Quartil : $\alpha = 0{,}75$ → $\mathcal{E}_{0,75}$ → $\int_{-\infty}^{\mathcal{E}_{0,75}} f_X(x)dx = 0{,}75$
- **Moda** : Seja X uma variável aleatória contínua. Moda, μ_*, é um valor de X que verifica a condição

$$f_X(x) \leq f_X(\mu_*), \forall x \in \mathbb{R}$$

Nota: A moda não é um parâmetro da ordem, contudo trata-se de uma importante medida de localização.

- **Amplitude Interquartis** : $AIQ = \mathcal{E}_{0,75} - \mathcal{E}_{0,25}$ (medida de dispersão absoluta)
- Uma medida de dispersão relativa : $\frac{AIQ}{\mu_c}$
- **Principais medidas de localização e de dispersão** :

Medidas	μ	σ	CV	μ_*	μ_c	AIQ	AIQ/μ_c
Localização	×			×	×		
Dispersão absoluta		×				×	
Dispersão relativa			×				×

5.5. Função Geradora dos Momentos

- **Função Geradora dos Momentos** : Seja X uma variável aleatória contínua com função densidade f_X. Se existir um número real $s_0 > 0$, tal que

$$E(e^{sX}) = \int_{-\infty}^{+\infty} e^{sx} f_X(x) dx$$

existe e é finito para $-s_0 < s < s_0$, então

$$M(s) = E(e^{sX})$$

define a função geradora dos momentos de X.

a) Existe uma correspondência biunívoca entre função geradora dos momentos (quando existe) e função densidade

b) $M_{X+Y}(s) = M_X(s) M_Y(s)$

5.6 Distribuição Uniforme

- **Distribuição Uniforme** : Diz-se que uma variável aleatória contínua, X, tem distribuição uniforme no intervaalo (α, β), com $\alpha < \beta$, quando a função densidade é dada por

$$f_X(x|\alpha,\beta) = \begin{cases} \frac{1}{\beta-\alpha} & se\ \alpha < x < \beta \\ 0 & para\ outros\ x \end{cases}$$

Simbolicamente, $X \sim U(\alpha, \beta)$.

- **Principais características da Distribuição Uniforme** :
  - Função de Distribuição : $F_X(x|\alpha,\beta) = \begin{cases} 0 & se\ x < \alpha \\ \frac{x-\alpha}{\beta-\alpha} & se\ \alpha < x < \beta \\ 1 & se\ x \geq \beta \end{cases}$
  - Função Geradora dos Momentos : $M(s) = E(e^{sX}) = \int_\alpha^\beta \frac{e^{sx}}{\beta-\alpha} dx = \begin{cases} \frac{e^{s\beta}-e^{s\alpha}}{s(\beta-\alpha)} & se\ s \neq 0 \\ 1 & se\ s = 0 \end{cases}$
  - Expressão Geral dos Momentos em Relação à Origem :

$$E(X^k) = \int_\alpha^\beta \frac{x^k}{\beta-\alpha} dx = \frac{1}{\beta-\alpha}\left[\frac{x^{k+1}}{k+1}\right]_\alpha^\beta = \frac{\beta^{k+1}-\alpha^{k+1}}{(k+1)(\beta-\alpha)}$$

  - Média : $E(X) = \frac{\alpha+\beta}{2}$
  - Variância : $Var(X) = \frac{(\beta-\alpha)^2}{12}$
  - Coeficiente de Assímetria é nulo : $\gamma_1 = 0$

- **Transformação Uniformizante** :

a) Seja X uma variável aleatória contínua com função de distribuição F_X estritamente crescente no intervalo aberto $]a,b[$ onde X tem densidade positiva (este intervalo pode ser ilimitado, isto é, pode acontecer $a = -\infty$ ou $b = +\infty$; tem-se $F_X(a) = 0$ e $F_X(b) = 1$). Então a variável aleatória $Y = F_X(X)$ tem distribuição $U(0,1)$.

b) Seja Y uma variável aleatória contínua com distribuição $U(0,1)$, e F_X a função de distribuição de uma variável aleatória contínua. Supõe-se que F_X é estritamente crescente no intervalo aberto $]a,b[$ onde esta variável aleatória tem densidade positiva (o intervalo pode ser ilimitado, isto é, pode acontecer $a = -\infty$ ou $b = +\infty$; tem-se $F_X(a) = 0$ e $F_X(b) = 1$). Então a variável aleatória $X = F_X^{-1}(Y)$ é contínua com função de distribuição F_X.

5.7 Distribuição Normal

- **Distribuição Normal** : Diz-se que a variável aleatória X tem distribuição normal com parâmetros μ e σ^2 quando a função densidade é da forma
$$f(x|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right) \text{ com } -\infty < x < +\infty$$
onde $\mu \in \mathbb{R}$ e $\sigma^2 > 0$. Simbolicamente $X \sim N(\mu,\sigma^2)$.
- **Função de Distribuição de $X \sim N(\mu,\sigma^2)$** : $F(x|\mu,\sigma^2) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{1}{2\sigma^2}(t-\mu)^2\right) dt$
- **Caso Particular Importante** : $\mu = 0$ e $\sigma^2 = 1$
  - Função Densidade : $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$
  - Função de Distribuição : $\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2} dt$
- **Distribuição Normal Estardadizada** : A distribuição $N(0,1)$ é uma distribuição normal estandardizada se $X \sim N(\mu,\sigma^2)$ e $Z = \frac{X-\mu}{\sigma}$.
- **Simetria** :
  → $N(\mu,\sigma^2)$ é simetrica em relação à recta $x = \mu$
  → $N(0,1) \rightarrow \phi(x) = \phi(-x)$; $\Phi(x) = 1 - \Phi(-x)$
- **Função Geradora dos Momentos de $Z \sim N(0,1)$** : $M_Z(s) = e^{-s^2/2}$
- **Função Geradora dos Momentos de $X \sim N(\mu,\sigma^2)$** : $M_Z(s) = exp\left(\mu s + \frac{\sigma^2 s^2}{2}\right)$
- **Valor Esperado e Variância** : $E(X) = \mu$; $Var(X) = \sigma^2$
- **Expressão Geral dos Momentos Centrais de Ordem Par** : $\mu_{2r} = \frac{(2r)!\sigma^{2r}}{2^r r!}$
- **Coeficiente de Kurtosis** : $\gamma_2 = \frac{\mu_4}{\sigma^4} = 3$
- **Interpretação do excesso de kurtosis** :
  a) Distribuição leptokurtica : $\gamma_2 > 3 \rightarrow$ caudas mais "espessas"
  b) Distribuição platilurtica : $\gamma_2 < 3 \rightarrow$ caudas mais "finas"
  c) Distribuição mesokurtica : $\gamma_2 = 3 \rightarrow$ distribuição normal
- **Intervalos e Probabilidades da Distribuição Normal** :

Intervalos	Probabilidades
$\mu \pm \sigma$	0,6826
$\mu \pm 2\sigma$	0,9544
$\mu \pm 3\sigma$	0,9973
$\mu \pm 0{,}6745\sigma$	0,5000
$\mu \pm 1{,}6450\sigma$	0,9000
$\mu \pm 1{,}9600\sigma$	0,9500
$\mu \pm 2{,}5758\sigma$	0,9900

- **Teorema** : Se as variáveis aleatórias X_i $(i = 1,2,\dots,k)$ são independentes, então
$$X_i \sim N(\mu_i,\sigma_i^2) \quad \Longrightarrow \quad Y = \sum_{i=1}^{k} \alpha_i X_i \sim N(\mu_Y,\sigma_Y^2)$$
onde
$$\mu_Y = \sum_{i=1}^{k} \alpha_i \mu_i \quad \text{e} \quad \sigma_Y^2 = \sum_{i=1}^{k} \alpha_i^2 \sigma_i^2$$
→ Qualquer combinação linear de variáveis aleatórias independentes com distribuição normal tem ainda distribuição normal.
- **Corolário** : Se as variáveis aleatórias X_i $(i = 1,2,\dots,k)$ são independentes e identicamente distribuídas, então
$$X_i \sim N(\mu,\sigma^2) \quad \Longrightarrow \quad Y = \sum_{i=1}^{k} X_i \sim N(k\mu,k\sigma^2)$$
- **Corolário** : Se as variáveis aleatórias X_i $(i = 1,2,\dots,k)$ são independentes e identicamente distribuídas, então
$$X_i \sim N(\mu,\sigma^2) \quad \Longrightarrow \quad \bar{X} = \frac{1}{k}\sum_{i=1}^{k} X_i \sim N\left(\mu,\frac{\sigma^2}{k}\right)$$

- **Teorema** : Se as variáveis aleatórias X_i $(i = 1,2,\dots,k)$ são normais tais que
$$E(X_i) = \mu_i\ ,\ \ Var(X_i) = \sigma_i^2 = \sigma_{ii}\ ,\ \ Cov(X_i, X_j) = \sigma_{ij}\ (i \neq j)$$
então
$$Y = \sum_{i=1}^{k} \alpha_i X_i \sim N(\mu_Y, \sigma_Y^2)$$
onde
$$\mu_Y = \sum_{i=1}^{k} \alpha_i \mu_i$$
e
$$\begin{aligned}\sigma_Y^2 = \alpha_1^2\sigma_{11} + 2\alpha_1\alpha_2\sigma_{12} + 2\alpha_1\alpha_3\sigma_{13} + \cdots + 2\alpha_1\alpha_k\sigma_{1k} \\ +\alpha_2^2\sigma_{22} + 2\alpha_2\alpha_3\sigma_{23} + \cdots + 2\alpha_2\alpha_k\sigma_{2k} \\ \cdots \\ +\alpha_k^2\sigma_{kk}\end{aligned}$$

5.8 Distribuição Exponencial

- **Distribuição Exponencial** : A variável aleatória X com função densidade definida por
$$f(x|\lambda) = \begin{cases} 0 & se\ x \leq 0 \\ \lambda e^{-\lambda x} & se\ x > 0 \end{cases}$$
diz-se que tem distribuição exponencial (ou exp. negativa). Simbolicamente, $X \sim Ex(\lambda)$.
- **Principais Características da Distribuição Exponencial** :
  - Função de Distribuição : $F(x|\lambda) = \begin{cases} 0 & se\ x \leq 0 \\ 1 - e^{-\lambda x} & se\ x > 0 \end{cases}$
  - Função Geradora dos Momentos : $M(s) = \left(1 - \frac{s}{\lambda}\right)^{-1} = \frac{\lambda}{\lambda - s},\ \ (s < \lambda)$
  - Valor Esperado : $E(X) = \frac{1}{\lambda}$
  - Variância : $Var(X) = \frac{1}{\lambda^2}$
  - Coeficiente de Variação : $CV = 1$
  - Coeficiente de Assimetria : $\gamma_1 = 2$
  - Coeficiente de Kurtosis : $\gamma_2 = 9$
  - Não tem Moda porque $f(x|\lambda)$ é decrescente com x e o domínio é aberto.
  - Mediana : $\mu_c = \frac{\ln 2}{\lambda}$ → enviesamento positivo → $\mu_c < \mu$
- **Propriedade da "Falta de Memória"** :
$$X \sim Ex(\lambda) \Longrightarrow P(X > x + h | X > x) = P(X > x)$$
- **Teorema** : Se $X_1, X_2, \dots, X_k$ são variáveis aleatórias independentes, então
$$X_i \sim Ex(\lambda) \Longrightarrow \min_i X_i \sim Ex(k\lambda)$$

5.9 Distribuição Gama ; Distribuição do Qui-Quadrado

- **Função Gama (factorial generalizado)** : Função real de variável real que faz corresponder a cada número real positivo, $\alpha > 0$, o número real definido por
$$\Gamma(\alpha) = \int_0^{+\infty} e^{-x} x^{\alpha-1} dx\ \ (\alpha > 0)$$
- **Propriedades da Função Gama** :
  a) Substituindo x por λx, obtém-se $\Gamma(\alpha) = \lambda^\alpha \int_0^{+\infty} e^{-\lambda x} x^{\alpha-1} dx$
  b) Com $\alpha > 1$, integrando por partes, obtém-se $\Gamma(\alpha) = (\alpha - 1) \cdot \Gamma(\alpha - 1)$
  c) Se $\alpha = n$ (inteiro positivo), tem-se $\Gamma(n) = (n-1)!$
  mas $\Gamma(1) = \int_0^{+\infty} e^{-x} dx = 1$
  d) $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$
- **Distribuição Gama** : Uma variável aleatória X com função densidade dada por
$$f(x|\alpha,\lambda) = \begin{cases} 0 & se\ x \leq 0 \\ \frac{\lambda^\alpha e^{-\lambda x} x^{\alpha-1}}{\Gamma(\alpha)} & se\ x > 0 \end{cases},\ \ \text{com } \alpha > 0 \text{ e } \lambda > 0$$
diz-se ter distribuição gama de parâmetros α e λ. Simbolicamente, $X \sim G(\alpha, \lambda)$.
→ Nota : Quando $\alpha = 1$ cai-se no caso particular da distribuição exponencial.

- Principais Características da Distribuição Exponencial :
  - Função Geradora dos Momentos : $M(s) = \left(1-\frac{s}{\lambda}\right)^{-n} = \left(\frac{\lambda}{\lambda-s}\right)^{n}$, $(s < \lambda)$
  - Momentos em Relação à Origem : $E\left(X^k\right) = \frac{\alpha(\alpha+1)\dots(\alpha+k-1)}{\lambda^k}$
  - Valor Esperado : $E(X) = \mu = \frac{\alpha}{\lambda}$
  - Variância : $Var(X) = \sigma^2 = \frac{\alpha}{\lambda^2}$
  - Coeficiente de Variação : $CV = \frac{1}{\sqrt{\alpha}}$
  - Coeficiente de Assimetria : $\gamma_1 = \frac{2}{\sqrt{\alpha}}$
  - Coeficiente de Kurtosis : $\gamma_2 = 3 + \frac{6}{\alpha}$
  - Se $\alpha \leq 1$, não tem Moda ; Se $\alpha > 1$, Moda : $moda(X) = \mu_* = \frac{\alpha-1}{\lambda}$
  - Se $\alpha > 1$, verifica-se sempre $\mu_* < \mu_c < \mu$
- Teorema : Sejam X_1 e X_2 duas variáveis aleatórias independentes. Então

  $$X_1 \sim G(\alpha_1;\lambda)\ ,\ \ X_2 \sim G(\alpha_2;\lambda) \ \implies\ X = X_1 + X_2 \sim G(\alpha;\lambda)\ ,\ \ \alpha = \alpha_1 + \alpha_2$$

  Generalizando para k variáveis aleatórias independentes, tem-se

  $$X_i \sim G(\alpha_i;\lambda), i = 1,2,\dots,k \ \implies\ X = \sum_{i=1}^{k} X_i \sim G(\alpha;\lambda),\ \ \alpha = \sum_{i=1}^{k} \alpha_i$$

  → Quando, em particular, $\alpha_i = 1$ $(i = 1,2,\dots,k)$, tem-se

  $$X_i \sim Ex(\lambda), i = 1,2,\dots,k \ \implies\ X = \sum_{i=1}^{k} X_i \sim G(k;\lambda)$$

  ou seja, a soma de k variáveis aleatórias indepenentes com distribuição exponencial, com o mesmo λ, tem distribuição gama com parâmetros k e λ.
- Tempo de Espera pelo n-ésimo acontecimento (Poisson) : $f(x|n,\lambda) = \frac{\lambda^n e^{-\lambda x} x^{n-1}}{(n-1)!}$
- Distribuição do Qui-Quadrado : Diz-se que a variável aleatória X tem distribuição do qui-quadrado com n graus de liberdade (n inteiro positivo), simbolicamente $X \sim \chi^2(n)$, quando a respectiva função densidade é dada por

  $$f(x|n) = \begin{cases} 0 & se\ x \leq 0 \\ \frac{e^{-\frac{x}{2}} x^{\frac{n}{2}-1}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} & se\ x > 0 \end{cases} ,\ \ \text{para } n = 1,2,\dots$$

  Nota : A distribuição do qui-quadrado é uma distribuição gama com $\alpha = \frac{n}{2}$ e $\lambda = \frac{1}{2}$,

  $$X \sim \chi^2(n) <=> X \sim G\left(\frac{n}{2};\frac{1}{2}\right)$$

- Métodos de Cálculo :

  → $X \sim \chi^2(n) \implies \sqrt{2X} - \sqrt{2n-1} \overset{a}{\sim} N(0,1)$

  → $X \sim G(n;\lambda) \Leftrightarrow 2\lambda X - \chi^2(2n)$
- Principais Características da Distribuição do Qui-Quadrado :
  - Valor Esperado : $E(X) = n$
  - Variância : $Var(X) = 2n$
  - Coeficiente de Variação : $CV = \sqrt{\frac{2}{n}}$
  - Coeficiente de Assimetria : $\gamma_1 = \sqrt{\frac{8}{n}}$
  - Coeficiente de Kurtosis : $\gamma_2 = 3 + \frac{12}{n}$
  - Função Geradora dos Momentos : $M(s) = (1-2s)^{-\frac{n}{2}}$ com $s < \frac{1}{2}$
- Teorema : Sejam X_1 e X_2 duas variáveis aleatórias independentes. Então

  $$X_1 \sim \chi^2(n_1)\ ,\ \ X_2 \sim \chi^2(n_2) \ \implies\ X = X_1 + X_2 \sim \chi^2(n)\ ,\ \ n = n_1 + n_2$$

  Generalizando para k variáveis aleatórias independentes, tem-se

  $$X_i \sim \chi^2(n_i), i = 1,2,\dots,k \ \implies\ X = \sum_{i=1}^{k} X_i \sim \chi^2(n),\ \ \alpha = \sum_{i=1}^{k} n_i$$

- **Teorema** : Seja X uma variável aleatória. Então
$$X \sim N(0,1) \Longrightarrow Y = X^2 \sim \chi^2(1)$$
ou
$$X_i \sim N(\mu_i, \sigma_i^2) \Longrightarrow \sum_{i=1}^{n} \left(\frac{X_i - \mu_i}{\sigma_i}\right)^2 \sim \chi^2(n)$$

5.10 Distribuição Beta

- Função Beta : Função real de duas variáveis reais, fazendo corresponder a cada par de números reais positivos (α, β), um número real definido por
$$B(\alpha, \beta) = \int_0^1 u^{\alpha-1}(1-u)^{\beta-1} du \quad \text{com } \alpha > 0 \text{ e } \beta > 0$$
- Propriedades da Função Beta
  a) $B(\alpha, \beta) = B(\beta, \alpha)$
  b) $B(1,1) = 1$
  c) As funções gama e beta estão relacionadas do seguinte modo : $B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$
  d) $B\left(\frac{1}{2}, \frac{1}{2}\right) = \pi$
- Distribuição Beta : A variável aleatória X com função densidade dada por
$$f_X(x|\alpha, \beta) = \begin{cases} \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1} & se\ 0 < x < 1 \\ 0 & para\ outros\ x \end{cases}, \quad \text{para } \alpha > 0 \text{ e } \beta > 0,$$
diz-se que tem distribuição beta. Simbolicamente $X \sim Be(\alpha, \beta)$.
- **Principais Características da Distribuição Beta** :
  - Momentos em Relação à Origem : $E(X^k) = \frac{\alpha(\alpha+1)\ldots(\alpha+k-1)}{(\alpha+\beta)(\alpha+\beta+1)\ldots(\alpha+\beta+k-1)}$
  - Valor Esperado : $E(X) = \frac{\alpha}{\alpha+\beta}$
  - Variância : $Var(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$
  - Coeficiente de Variação : $CV = \sqrt{\frac{\beta}{\alpha(\alpha+\beta+1)}}$
  - Coeficiente de Assimetria : $\gamma_1 = \frac{2(\beta-\alpha)\sqrt{\alpha+\beta+1}}{(\alpha+\beta+2)\sqrt{\alpha\beta}}$
  - Coeficiente de Kurtosis : $\gamma_2 = \frac{3(\alpha+\beta+1)[2(\alpha+\beta)^2+\alpha\beta(\alpha+\beta-6)]}{\alpha\beta(\alpha+\beta+2)(\alpha+\beta+3)}$
- **Forma/Simetria** :
  → $\alpha = \beta$ → distribuição simétrica
  → $\alpha < \beta$ → distribuição assimétrica positiva → cauda longa à direita
  → $\alpha > \beta$ → distribuição assimétrica negativa → cauda longa à esquerda
  → $\alpha > 0$ e $\beta > 0$ → distribuição unimodal → $moda(X) = \mu_c = \frac{\alpha-1}{\alpha+\beta-2}$
  → $\alpha = 1$ e $\beta = 1$ → distribuição uniforme
  → outros valores de α e β → não tem moda

5.11 Teorema do Limite Central

- **Interpretação do Teorema do Limite Central** : O Teorema do Limite Central garante que a soma de n variáveis aleatórias independentes, todas com a mesma média e a mesma variância, tem, depois de estandardizada e para valores de n suficientemente grandes, distribuição aproximada, $N(0,1)$.
- **Teorema da Continuidade** : Considere-se a sucessão de variáveis aleatórias,
$$X_1, X_1 + X_2, \ldots, X_1 + X_2 + \cdots + X_n, \ldots, \text{ ou } \sum_{i=1}^{n} X_i \quad \text{com } n = 1,2, \ldots$$
Sejam $\{F_n\}$ e $\{M_n\}$, com $n = 1,2, \ldots$, as respectivas sucessões de funções de distribuição e de funções geradoras dos momentos [supõe-se que existem para $|s| < s_0$ $(s_0 > 0)$]. Então, se $M_n(s) \to M(s)$ $(s > 0)$ e $M_n(0) \to 1$, a sucessão $\{F_n\}$ tende para a função de distribuição F que tem M como função geradora dos momentos.

- **Teorema do Limite Central** : Dada a sucessão de variáveis aleatórias independentes e identicamente distribuídas (*iid*), $X_1, X_2, \dots, X_n, \dots$, com média μ e variância σ^2, então, quando $n \to +\infty$, a função de distribuição da variável aleatória,
$$Z_n = \frac{\sum_{i=1}^{n}(X_i) - n\mu}{\sqrt{n}\sigma}$$
tende para uma função de distribuição $N(0,1)$, ou seja, a distribuição assintótica ou aproximada de Z_n é $N(0,1)$. Simbolicamente, $Z_n \overset{a}{\sim} N(0,1)$.

  - <u>Outras formulações</u> :

    $\to$ $\lim_{n\to\infty} P(Z_n \leq x) = \lim_{n\to\infty} P\left(\frac{\sum_{i=1}^{n}(X_i) - n\mu}{\sqrt{n}\sigma} \leq x\right) = \Phi(x)$

    $\to$ $P(Z_n \leq x) = P\left(\frac{\sum_{i=1}^{n}(X_i) - n\mu}{\sqrt{n}\sigma} \leq x\right) \approx \Phi(x)$ (n grande)

- **Corolário** : Dada uma sucessão de variáveis aleatórias *iid*, $X_1, X_2, \dots, X_n, \dots$, com distribuição de Bernoulli de média $E(X_i) = \theta$ e $Var(X_i) = \theta(1-\theta)$, vem
$$\frac{\sum_{i=1}^{n}(X_i) - n\theta}{\sqrt{n\theta(1-\theta)}} \overset{a}{\sim} N(0,1)$$
$\to$ $P(a \leq X \leq b) = \Phi\left(\frac{b - n\theta}{\sqrt{n\theta(1-\theta)}}\right) - \Phi\left(\frac{a - n\theta}{\sqrt{n\theta(1-\theta)}}\right)$

- **Correcção de Continuidade** : para $n > 20$,
$$P(a \leq X \leq b) \approx \Phi\left(\frac{b + \frac{1}{2} - n\theta}{\sqrt{n\theta(1-\theta)}}\right) - \Phi\left(\frac{a - \frac{1}{2} - n\theta}{\sqrt{n\theta(1-\theta)}}\right)$$
<u>Caso particular ($c = a = b$)</u> :
$$P(X = c) = P\left(c - \frac{1}{2} \leq X \leq c + \frac{1}{2}\right) = \Phi\left(\frac{c + \frac{1}{2} - n\theta}{\sqrt{n\theta(1-\theta)}}\right) - \Phi\left(\frac{c - \frac{1}{2} - n\theta}{\sqrt{n\theta(1-\theta)}}\right)$$

- **Regras práticas orientadoras do cálculo de probabilidades que envolvem a distribuição binomial** :

  $\to$ Sempre que possível, deve utilizar-se directamente a distribuição binomial, o que permite o cálculo exacto das probabilidades recorrendo a meios computacionais.

  $\to$ Quando necessário (sempre para $n > 20$), o cálculo aproximado das probabilidades deve atender aos seguintes casos:
    - Se $\theta \leq 0{,}1$, deve utilizar-se a aproximação da binomial pela Poisson;
    - Se $\theta \geq 0{,}9$, deve usar-se a Poisson considerando o acontecimento complementar
    - Se $0{,}1 < \theta < 0{,}9$, deve recorrer-se à fórmula da correcção da continuidade

  - <u>Nota importante</u> : À medida que θ se afasta de 0,5, são necessárias amostras com dimensão cada vez maior, uma vez que a distribuição binomial se torna mais assimétrica.

- **Corolário** : Se X é uma variável com distribuição de Poisson, $X \sim Po(\lambda)$, então
$$\lim_{n\to\infty} P\left(\frac{X - \lambda}{\sqrt{\lambda}} \leq x\right) = \Phi(x)$$
  - <u>Outra formulação</u> : … $\frac{X-\lambda}{\sqrt{\lambda}} \overset{a}{\sim} N(0,1)$ quando $\lambda \to +\infty$

- **Correcção de Continuidade** : para $\lambda > 20$,
$$P(a \leq X \leq b) \approx \Phi\left(\frac{b + \frac{1}{2} - \lambda}{\sqrt{\lambda}}\right) - \Phi\left(\frac{a - \frac{1}{2} - \lambda}{\sqrt{\lambda}}\right)$$

5.12 Valor Esperado e Momentos de Variáveis Aleatórias Contínuas

- **Valor Esperado de Função de Variável Aleatória Bidimensional** : Se (X, Y) for uma variável aleatória bidimensional contínua, com função densidade $f_{X,Y}$, e se ψ é uma função de (X, Y), a expressão
$$E[\psi(X,Y)] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \psi(x,y) f_{X,Y}(x,y)dxdy$$
define o valor esperado de $\psi(x, y)$, desde que se verifique a já referida condição de existência do valor esperado.
- **Teorema** : … $E(X + Y) = E(X) + E(Y)$
- **Teorema** : … $E(XY) = E(X)E(Y)$
- **Momentos de Ordem $r + s$ em Relação à Origem** : Seja (X, Y) uma variável aleatória bidimensional contínua. O valor esperado,
$$\mu'_{rs} = E(X^r Y^s) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x^r y^s f_{X,Y}(x,y)dxdy$$
define um momento de ordem $r + s$ em relação à origem de (X, Y).
  - → Momentos de primeira ordem :
    - – $r = 1, s = 0 \Longrightarrow \mu'_{10} = E(X) = \mu_X$
    - – $r = 0, s = 1 \Longrightarrow \mu'_{01} = E(Y) = \mu_Y$
  - → Momentos de segunda ordem :
    - – $r = 2, s = 0 \Longrightarrow \mu'_{20} = E(X^2)$
    - – $r = 1, s = 1 \Longrightarrow \mu'_{11} = E(XY)$
    - – $r = 0, s = 2 \Longrightarrow \mu'_{02} = E(Y^2)$
- **Momentos de Ordem $r + s$ em Relação à Média** : Seja (X, Y) uma variável aleatória bidimensional contínua. O valor esperado,
$$\mu_{rs} = E[(X-\mu)^r(Y-\mu)^s] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x-\mu_X)^r(y-\mu_Y)^s f_{X,Y}(x,y)dxdy$$
define um momento de ordem $r + s$ em relação à média de (X, Y).
  - → Momentos de primeira ordem (são nulos) :
    - – $r = 1, s = 0 \Longrightarrow \mu_{10} = 0$
    - – $r = 0, s = 1 \Longrightarrow \mu_{01} = 0$
  - → Momentos de segunda ordem :
    - – $r = 2, s = 0 \Longrightarrow \mu_{20} = E[(X-\mu)^2] = Var(X) = \sigma_X^2$
    - – $r = 1, s = 1 \Longrightarrow \mu_{11} = E[(X-\mu)(Y-\mu)]$
    - – $r = 0, s = 2 \Longrightarrow \mu_{02} = E[(Y-\mu)^2] = Var(Y) = \sigma_Y^2$
- **Covariância** : A covariância das variáveis aleatórias X e Y é
$$Cov(X,Y) = \sigma_{X,Y} = E[(X-\mu_X)(Y-\mu_Y)]\,,$$
se este valor esperado existir.
  - → $Cov(X,Y) = E(XY) - E(X)E(Y)$
  - → $Cov(c,X) = E(cX) - cE(X) = 0$
- **Coeficiente de Correlação** : O coeficiente de correlação entre as variáveis aleatórias X e Y é dado por
$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}} = \frac{\sigma_{X,Y}}{\sigma_X\sigma_Y}$$
- **Teorema** : Se duas variáveis aleatórias, X e Y, são independentes, então a respectiva covariância é igual a zero.
  - → A independência implica não correlação, mas a recíproca não é verdadeira.
- **Teorema** : Se existem segundos momentos para as variáveis aleatórias X e Y, então
$$Var(X \pm Y) = Var(X) \pm Cov(X,Y) + Var(Y)$$
Em particular, se as variáveis são independentes,
$$Var(X \pm Y) = Var(X) + Var(Y)$$

- **Valor Esperado Condicionado** : Seja a variável aleatória $Z = \psi(X, Y)$, função das variáveis aleatórias contínuas X e Y. O valor esperado de $Z = \psi(X, Y)$ condicionado por $X = x$, é definido por
$$E(Z|X = x) = E[\psi(X, Y)|X = x] = \int_{-\infty}^{+\infty} \psi(x, y) f_{Y|X=x}(y) dy$$
se se verificar a condição habitual de existência do valor esperado.
De modo análogo se define
$$E(Z|Y = y) = E[\psi(X, Y)|Y = y] = \int_{-\infty}^{+\infty} \psi(x, y) f_{X|Y=y}(x) dx$$
$\rightarrow$ Se $\psi(x, y) = Y$, $E(Y|X = x) = \int_{-\infty}^{+\infty} y f_{Y|X=x}(y) dy$
$\rightarrow$ Se $\psi(x, y) = X$, $E(X|Y = y) = \int_{-\infty}^{+\infty} x f_{X|Y=y}(x) dx$
- **Teorema** : O valor esperado de $Z = \psi(X, Y)$, se existir, é igual ao valor esperado do seu valor esperado condicionado por X,
$$E(Z) = E[E(Z|X)]$$
$\rightarrow$ Se $Z = Y$, $E(Y) = E[E(Y|X)]$ → regra do valor esperado iterado (→ v.e. marginal)
- **Variância Condicionada** : A variância de Y condicionada por $X = x$ é definida por
$$Var(Y|X = x) = E[\{Y - E(Y|X = x)\}^2 | X = x]$$
$$= \int_{-\infty}^{+\infty} \{y - E(Y|X = x)\}^2 f_{Y|X=x}(y) dy$$
Do mesmo modo se define a variância de X condicionada por $Y = y$,
$$Var(X|Y = y) = E[\{X - E(X|Y = y)\}^2 | Y = y]$$
$$= \int_{-\infty}^{+\infty} \{x - E(X|Y = y)\}^2 f_{X|Y=y}(x) dx$$
$\rightarrow$ $Var(Y|X = x) = E(Y^2|X = x) - E(Y|X = x)^2$
$\rightarrow$ $Var(X|Y = y) = E(X^2|Y = y) - E(X|Y = y)^2$

➢ **Os resultados da secção 4.11 ainda são válidos para variáveis aleatórias k-dimensionais contínuas (matrizes das covariâncias e das correlações, etc).**

5.13 Distribuição Normal Bidimensional

- **Distribuição Normal Bidimensional** : Se a variável aleatória bidimensional (X, Y) tem função densidade é da forma
$$f_{X,Y}(x, y|\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2, \rho_{X,Y}) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{X,Y}^2}}$$
$$\times\, exp\left\{-\frac{1}{2(1-\rho_{X,Y}^2)}\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho_{X,Y}\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right\}$$
com $(x, y) \in \mathbb{R}^2$, $\mu_X \in \mathbb{R}$, $\mu_Y \in \mathbb{R}$, $\sigma_X^2 > 0$, $\sigma_Y^2 > 0$ e $-1 \leq \rho_{X,Y} \leq 1$, diz-se que tem distribuição normal bidimensional.
- **Principais Características da Distribuição Normal Bidimensional** :
  - ➢ $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_{X,Y}(x, y) dx\, dy = 1$
  - ➢ Valor Esperado (marginal) : $E(X) = \mu_X$; $E(Y) = \mu_Y$
  - ➢ Variância (marginal) : $Var(X) = \sigma_X^2$; $Var(Y) = \sigma_Y^2$
  - ➢ Coeficiente de Correlação : $\rho_{X,Y}$
  - ➢ Covariância : $Cov(X, Y) = \rho_{X,Y}\sigma_X\sigma_Y$
- **Distribuição Normal Bidimensional Estandardizada** : Se $\mu_X = \mu_Y = 0$ e $\sigma_X^2 = \sigma_Y^2 = 1$,
$$f_{X,Y}(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}} exp\left\{-\frac{1}{2(1-\rho_{X,Y}^2)}(x^2 - 2\rho xy + y^2)\right\}$$
- **Função Densidade Marginal** :
$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma_X^2}} exp\left\{-\frac{1}{2\sigma_X^2}(x - \mu_X)^2\right\} \;\; ; \;\; f_Y(y) = \frac{1}{\sqrt{2\pi\sigma_Y^2}} exp\left\{-\frac{1}{2\sigma_Y^2}(y - \mu_Y)^2\right\}$$
- **Teorema** : Se (X, Y) é uma variável aleatória normal bidimensional, então X e Y são independentes se e so se X e Y forem não correlacionados ($\rho = 0$).

- **Função Densidade de X condicionada por $Y = y$** :

$$f_{X|Y=y}(x,y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{1}{\sqrt{2\pi\sigma_X^2(1-\rho^2)}} \times exp\left\{-\frac{1}{2\sigma_X^2(1-\rho^2)}\left[x - \left(\mu_X + \rho\frac{\sigma_X}{\sigma_Y}(y-\mu_Y)\right)\right]^2\right\}$$

- **Valor Esperado de X condicionado por $Y = y$** : $E(X|Y=y) = \mu_X + \rho\frac{\sigma_X}{\sigma_Y}(y-\mu_Y)$
- **Variância de X condicionada por $Y = y$** : $Var(X|Y=y) = \sigma_X^2(1-\rho^2)$

5.14 Distribuição Normal k-Dimensional

- **Os resultados obtidos na secção 5.13 são generalizáveis.**

…