várias variáveis

- probabilidade condicional (Fórmula de Bayes) $p_{X,Y}(x,y)$
  probabilidade conjunta
  $p_{X|Y}(x|y) = \frac{p_{X,Y}(x,y)}{p_Y(y)}$.
  marginal $p_Y(y) = \sum_{x \in \mathscr{X}} p_{X,Y}(x,y)$ e portanto $\sum_{x \in \mathscr{X}} p_{X|Y}(x|y) = 1$ $\forall y$
- variáveis independentes
  $p_{X,Y}(x,y) = p_X(x)p_Y(y)$ ou seja
  $p_{X|Y}(X = x|Y = y) = p_X(x)$;
  $Var(X + Y) = Var(X) + Var(Y)$
- independentes e idênticamente distribuídas -
  $p_N(x_1, x_2, \cdots x_N) = \prod_{i=1,N} p(x_i)$; $x \in A_1$ $x \in A_2$
- Acontecimentos mutuamente exclusivos, $\mathscr{A}_1 \cap \mathscr{A}_2 \equiv \oslash$ temos
  $prob(A_1 \cup \mathscr{A}_2) = prob(\mathscr{A}_1) + prob(\mathscr{A}_2)$
  Ex: $X \in \mathscr{X} \equiv \mathscr{R}$ e queremos calcular $prob(|X| \geq \varepsilon)$ com $\varepsilon > 0$.
  Podemos escrever $prob(|X| \geq \varepsilon) = \int_{\varepsilon}^{\infty} p_X(x)dx + \int_{-\infty}^{-\varepsilon} p_X(x)dx$

teorema de Chebyshev e lei dos grandes números

- Para um conjunto de N, variáveis aleatórias, $X_1, X_2, \cdots, X_N$, independentes e idênticamente distribuídas, com variância finita, $Var\, X \leq C$, verifica-se para $Y = \frac{1}{N}\sum_{k=1}^{N} X_k$ e para qualquer $\varepsilon > 0$:

  $\overline{Y} = 1/N \, \overline{\sum_k X_k} = \frac{1}{N} \sum_k \overline{X_k} = \frac{1}{N} N \overline{X} = \overline{X}$

  - $\lim_{N\to\infty} prob\left(\left|Y - \overline{X}\right| < \varepsilon\right) = 1$. Isto significa que a variável Y se aproxima do valor médio exato, $\overline{X}$ quando N cresce.
  - demonstração

    $Var\, Y = 1/N^2 \, Var\left(\sum_k X_k\right) = \frac{1}{N^2} \sum_k Var\, X_k = \frac{1}{N^2} N \, Var\, X = \frac{Var\, X}{N}$

    - como $\overline{Y} = \overline{X}$, $\boxed{Var\, Y = \frac{Var\, X}{N}}$ e $Var\, Y \leq \frac{C}{N}$, temos pela desigualdade de Chebyshev $prob\left(\left|Y - \overline{X}\right| \geq \varepsilon\right) \leq \frac{Var\, Y}{\varepsilon^2} \leq \frac{C}{N\varepsilon^2}$
    - Como $prob\left(\left|Y - \overline{X}\right| < \varepsilon\right) = 1 - prob\left(\left|Y - \overline{X}\right| \geq \varepsilon\right) \geq 1 - \frac{C}{N\varepsilon^2}$ depois de tomarmos o limite $N \to \infty$ obtemos, $\lim_{N\to\infty} prob\left(\left|Y - \overline{X}\right| < \varepsilon\right) \geq 1$. Dado que a probabilidade não pode ser superior a 1 demonstra-se o pretendido.

X_k, $k=1,\dots,N$ são v.a. independentes e identicamente distribuidas e $Y = \frac{1}{N}\sum_{k=1}^{N} X_k$

$$\operatorname{var} Y = \overline{Y^2} - \overline{Y}^2 \qquad \overline{Y} = \frac{1}{N}\sum_{k=1}^{N}\overline{X_k} = \overline{X}$$

$$\overline{X_k X_\ell} = \overline{X_k}\,\overline{X_\ell} = \overline{X}^2$$

$$\overline{Y^2} = \frac{1}{N^2}\overline{\left(\sum_k X_k\right)^2} = \frac{1}{N^2}\sum_k\sum_\ell \overline{X_k X_\ell}$$

$$= \frac{1}{N^2}\left[\sum_{k\neq\ell}\overline{X_k X_\ell} + \sum_k \overline{X_k^2}\right] = \frac{1}{N^2}\left[N(N-1)\overline{X}^2 + N\overline{X^2}\right]$$

Como $k \neq \ell$ são independentes

$$\operatorname{var} Y = \frac{1}{N^2}\left[N(N-1)\overline{X}^2 + N\overline{X^2}\right] - \overline{X}^2 = \frac{1}{N}\left(\overline{X^2} - \overline{X}^2\right)$$

função característica

Definição

$e^{ikx} = \cos kx + i \sin kx$ $\quad \frac{d}{dk} e^{ikx} = ix\, e^{ikx}$

- A função característica, $Q(k)$, de uma variável aleatória X define-se como: $Q(k) = \overline{\exp(i\,k\,X)}$ ou seja $Q_X(k) = \int_{x\in\mathscr{X}} \exp(i\,k\,x)\, p_X(x)\, dx$

$Q_X(0) = 1$ $\quad \frac{d^n}{dk^n} e^{ikx} = i^n x^n e^{ikx}$

- derivando $Q_X(k)$ em ordem a k, n vezes: $\left|\frac{d^n}{dk^n} Q_X(k)\right|_{k=0} = i^n \overline{X^n}$

ver exercício 6

- para uma variável Gaussiana, $Q_X(k) = \exp\left(i\mu k - \frac{\sigma^2 k^2}{2}\right)$

Uma distribuição de probabilidade é determinada pela sua função caraterística. Em particular para $x \in \mathscr{X} = [-\infty, \infty]$ e $p_X(x)$ contínua no intervalo, temos: $p_X(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp(-i\,k\,x)\, Q_X(k) dk$.

transformada de Fourier

função característica

$$\overline{e^{ikX_1} e^{ikX_2}} = \overline{e^{ikX_1}}\ \overline{e^{ikX_2}}$$ Se X_1 e X_2 forem independentes

- Para um conjunto de N, variáveis aleatórias, $X_1, X_2, \cdots, X_N$, independentes, a variável aleatória $Y = \sum^N{}_{k=1} X_k$ tem uma função característica, $Q_Y(k) = \prod_{k=1}^N Q_{X_k}(k)$.
  - demonstração:
    $Q_Y(k) = \overline{\exp(i\,k\,Y)} = \overline{\exp\left(i\,k\sum^N{}_{k=1} X_k\right)} = \overline{\prod_{k=1}^N \exp(i\,k\,X_k)}$.
    Como as variáveis aleatórias são independentes a média de um produto é igual ao produto das médias o que demonstra o resultado.
- Se forem identicamente distribuidas: $Q_Y(k) = [Q_X(k)]^N$

função característica

$Re\left[Q_X(k)\right] = \int_0^{+\infty} f(k,x)\,dx$

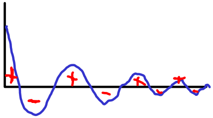
$f(k,x) = \cos(kx)\, p_X(x)$

exemplo:

$p_X(x) = \frac{1}{1+x^3}$

$x \geq 0$

$Re\left[Q_X(k)\right] \xrightarrow[k \text{ cresce}]{} 0$



quanto maior k
mais oscila

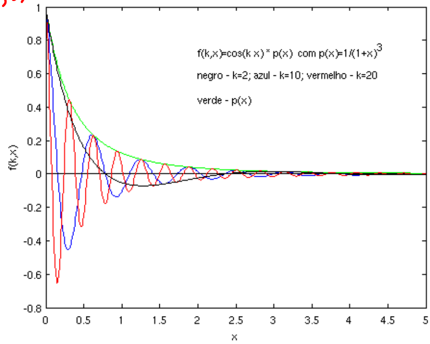


Figure: A figura mostra a função $f(k,x)$a integrar em x para calcular $Q(k)$.Quando k aumenta $Q_X(k)$ tende para zero. A função a integrar oscila em função de x cada vez mais quando k aumenta

Teorema do Limite Central

- No exemplo da figura $Q_X(k) = 0.2152, 0.031$ e 0.0121 para $k = 2, 10$ e 20. Se $N = 10$ temos $Q_Y(k) = 2.1 \times 10^{-7}, 7 \times 10^{-16}$ e 6.6×10^{-20} para os mesmos k. Conclui-se que $Q_Y(k)$ diminui rapidamente quando k aumenta!
- Para um conjunto de N, variáveis aleatórias, $X_1, X_2, \cdots, X_N$, independentes e idênticamente distribuídas (caso especial), com variância finita, a variável aleatória $Y = \sum^N{}_{k=1} X_k$ tem uma distribuição Gaussiana de média $\overline{Y} = N\overline{X}$ e variância, $Var\,Y = N\,Var\,X$.

Teorema do limite central

Teorema do Limite Central

$Q_Y(k) = [Q_X(k)]^N \qquad \ln Q_Y(k) = N \ln Q_X(k)$

ver exercício 7

- demonstração não rigorosa:
  - temos $\ln Q_Y(k) = N \ln Q_X(k)$ e expandimos até segunda ordem em k, ou seja $\ln Q_Y(k) =$
    $N \ln\left(1 + \left|\frac{d}{dk} Q_X(k)\right|_{k=0} k + \frac{1}{2}\left|\frac{d^2}{dk^2} Q_X(k)\right|_{k=0} k^2 + \cdots\right) =$
    $N \ln\left(1 + i\overline{X}k - \frac{1}{2}\overline{X^2}k^2 + \cdots\right)$.
  - usamos, $\ln(1+z) \simeq z - \frac{1}{2}z^2$ com $\boxed{z = i\overline{X}k - \frac{1}{2}\overline{X^2}k^2}$ e $z^2/2 = -\frac{1}{2}\overline{X}^2 k^2 + \cdots$
  - Obtemos $\ln Q_Y(k) = N\left[\underbrace{i\overline{X}k - \frac{1}{2}\left(\overline{X^2} - \overline{X}^2\right)k^2}_{z - \frac{z^2}{2}} + \cdots\right] \not=$

$Q_Y(k) =$ $\exp\left(i\overline{Y}k - Var\, Y \frac{k^2}{2}\right)$

$Y = \sum_k X_k \Rightarrow \overline{Y} = N\overline{X}\ ;\ Var\, Y = N\, Var\, X$

Expansão de Q(k) em série de Taylor até 2a ordem à volta de k=0

$$Q_x(k) = Q_x(0) + \left(\frac{dQ_x(k)}{dk}\right)_{k=0} k + \frac{1}{2!}\left(\frac{d^2Q_x(k)}{dk^2}\right)_{k=0} k^2 + \ldots$$

$$= 1 + i\overline{x}\,k + \frac{1}{2} i^2 \overline{x^2}\,k^2$$

$$= 1 + ik\overline{x} - \frac{1}{2}\overline{x^2}\,k^2$$

sabemos que a função se anula para k grande pelo que queremos apenas aproximar a função próximo de k=0

Teorema do Limite Central

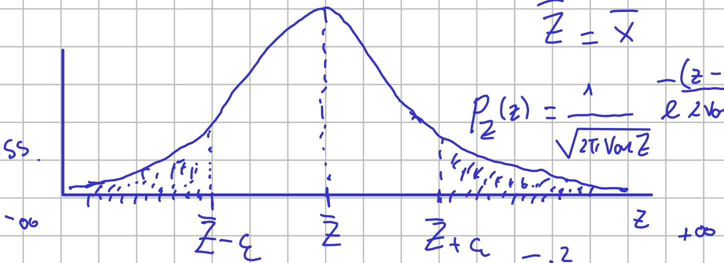
- Portanto, para N grande, Y é uma variável aleatória Gaussiana com média $\overline{Y} = N\overline{X}$ e variância, $Var\, Y = N\, Var\, X$.
- Se considerarmos a variável $Z = \frac{Y}{N} = \frac{1}{N}\sum^{N}{}_{k=1} X_k$ vemos que Z é uma variável Gaussiana de média $\overline{Z} = \overline{X}$ e variância, $Var\, Z = \frac{Var\, X}{N}$. $Var\, Z = Var\left[\frac{1}{N} Y\right] = \frac{1}{N^2} Var\, Y = \frac{Var\, X}{N} \xrightarrow[N\to\infty]{} 0$
- O teorema do limite central permite estimar a probabilidade de grandes desvios quando N é grande: $prob\left(\left|Z - \overline{X}\right| \geq \varepsilon\right) = \sqrt{\frac{2N}{\pi\, Var\, X}} \int_{\varepsilon}^{\infty} \exp\left(-\frac{N}{2\, Var\, X} x^2\right) dx$. A probabilidade de se observar um desvio, $\varepsilon > 0$, relativamente à média, diminui rápidamente com o aumento de N. (devido ao fator N na exponencial)

$$\int_{-\infty}^{\bar{x}-\varepsilon} e^{-\frac{(z-\bar{z})^2}{2\,\mathrm{Var}\,Z}} \frac{dz}{\sqrt{2\pi\,\mathrm{Var}\,Z}} \qquad \bar{x}-\varepsilon \qquad \bar{x} \qquad \bar{x}+\varepsilon \qquad \int_{\bar{x}+\varepsilon}^{\infty} e^{-\frac{(z-\bar{z})^2}{2\,\mathrm{Var}\,Z}} \frac{dz}{\sqrt{2\pi\,\mathrm{Var}\,Z}} \qquad z$$

distribuição de Gauss é simétrica relativamente à média

$p_Z(z)$

v.a. Gauss.



$$\bar{Z} = \bar{X}$$

$$p_Z(z) = \frac{1}{\sqrt{2\pi\,\mathrm{Var}\,Z}}\, e^{-\frac{(z-\bar{z})^2}{2\,\mathrm{Var}\,Z}}$$

$$\mathrm{prob}\left(|Z-\bar{Z}| \geq \varepsilon\right) = 2\int_{\bar{Z}+\varepsilon}^{\infty} \frac{e^{-\frac{(z-\bar{Z})^2}{2\,\mathrm{Var}\,Z}}}{\sqrt{2\pi\,\mathrm{Var}\,Z}}\, dz$$

fazemos a substituição $x = Z - \bar{Z}$

$$\text{prob}\left(|Z - \bar{Z}| \geq \varepsilon\right) = 2\int_{\varepsilon}^{\infty} \frac{e^{-\frac{x^2}{2\,\text{var}\,Z}}}{\sqrt{2\pi\,\text{var}\,Z}}\,dx$$

Como $\text{Var}\,Z = \dfrac{\text{Var}\,X}{N}$

$$= \sqrt{\frac{2N}{\pi\,\text{var}\,X}}\int_{\varepsilon}^{\infty} e^{-\frac{N x^2}{2\,\text{var}\,X}}\,dx$$

Desigualdade de Markov

- Seja Y uma v. a. qualquer que toma valores não negativos. Demonstra-se que $prob(Y \geq \varepsilon) \leq \frac{\overline{Y}}{\varepsilon}$.
  - demonstração:
    $prob(Y \geq \varepsilon) = \int_\varepsilon^\infty p_Y(x)dx \leq \int_\varepsilon^\infty \frac{x}{\varepsilon} p_Y(x)dx$ dado $\frac{x}{\varepsilon} > 1$, ou seja, $prob(Y \geq \varepsilon) \leq \frac{\int_0^\infty x p_Y(x)dx}{\varepsilon} = \frac{\overline{Y}}{\varepsilon}$ dado que $\frac{x}{\varepsilon} > 0$
- Caso particular $Y = \left(X - \overline{X}\right)^2 \geq 0$. Temos $prob(\left(X - \overline{X}\right)^2 \geq \varepsilon^2) \leq \frac{\overline{\left(X - \overline{X}\right)^2}}{\varepsilon^2}$(desigualdade de Markov). Como $prob(\left(X - \overline{X}\right)^2 \geq \varepsilon^2) = prob(\left|X - \overline{X}\right| \geq \varepsilon)$ obtemos a desigualdade de Chebyshev:$prob(\left|X - \overline{X}\right| \geq \varepsilon) \leq \frac{var\, X}{\varepsilon^2}$

processo estocástico

ocorre no tempo aleatório $X_1, X_2, X_3, \cdots$ $X_t \in \mathcal{X}$

- $\{X_t\}$ $t \in \mathbb{N}$ em que cada $X_t$toma valores em $\mathscr{X}$
- $p_N(x_1, x_2, \cdots x_N)$ é a probabilidade conjunta das variáveis, $X_1, X_2, \cdots, X_N$

  $N = 5$ $\mathcal{A} = \{1,2\}$ $\bar{\mathcal{A}} = \{3,4,5\}$
- definindo um sub-conjunto $\mathscr{A} \subset \{1,2,\cdots,\mathrm{N}\}$e o complementar $\overline{\mathscr{A}} \subset \{1,2,\cdots,\mathrm{N}\}\backslash A$ $\mathcal{A}$
  - a probabilidade marginal é $p_{\mathscr{A}}(x_{\mathscr{A}}) = \sum_{x_{\overline{\mathscr{A}}}} p_N(x_1, x_2, \cdots, x_N)$

    $x_{\mathcal{A}} = (x_1, x_2)$ $x_{\bar{\mathcal{A}}} = (x_3, x_4, x_5)$
- cadeia de Markov
  - $p_N(x_1, x_2, \cdots x_N) = p_1(x_1) \prod_{t=1}^{N-1} w(x_t \to x_{t+1})$ onde $w(x \to y)$ é a probabilidade de transição entre $x \in \mathscr{X}$ e $y \in \mathscr{X}$ com $\sum_{y \in \mathscr{X}} w(x \to y) = 1$ e $p_1(x_1)$ é a distribuição de probabilidade inicial.

$p_N(x_1, x_2, \dots, x_N) \equiv$ prob. conjunta

$$p_{X_1}(x_1) = \sum_{x_2, x_3, \dots, x_N} p_N(x_1, x_2, \dots, x_N)$$

$\downarrow$ marginal de X_1

$X_t \equiv$ posição de uma partícula no instante t

$X_t \equiv$ índice de uma bola.

$$p_{X_2}(x_2) = \sum_{x_1, x_3, \dots, x_N} p_N(x_1, x_2, x_3, \dots, x_N) = \text{marginal de } X_2$$

$$p_{X_1, X_2}(x_1, x_2) = \sum_{x_3, x_4, \dots, x_N} p_N(x_1, x_2, \dots, x_N) = \text{marginal de } X_1, X_2$$

$$p_{X_N}(x_N) = \sum_{x_1, x_2, \dots, x_{N-1}} p_N(x_1, x_2, \dots, x_N) = \text{marginal de } X_N$$