

# Integração Numérica: Funções univariadas

Prof. Caio Azevedo

- Objetivo: calcular aproximações analíticas numérica para integrais uni e multivariadas.

- Relembrando:  $\int_a^b f(x)dx = F(b) - F(a), F(x) = \frac{d}{dx}f(x).$

- Relembrando:  $\int_a^b \left( \int_c^d f(x, y)dx \right) dy = F(b, d) - F(a, d) - F(b, c) + F(a, c), F(x, y) = \frac{\partial^2}{\partial x \partial y} f(x, y) = \frac{\partial^2}{\partial y \partial x} f(x, y).$

- Soluções:

- Aproximações analíticas (Laplace).
- Aproximações numéricas não estocásticas (Quadratura, Quadratura adaptativa).
- Aproximações numéricas estocásticas (não-iterativas) (Monte Carlo).
- Aproximações numéricas estocásticas iterativas (Monte Carlo via Cadeias de Markov).

# Aproximações analíticas

- Sob certas condições, para um dado  $x_0$ ,

$$f(x) = \sum_{i=0}^{\infty} \frac{f^{(i)}(x_0)}{i!} (x - x_0)^i, \quad f^{(i)} = \frac{d^i}{dx^i} f(x).$$

- Sob certas condições,

$$\begin{aligned} \int_a^b f(x) dx &= \sum_{i=0}^{\infty} \frac{f^{(i)}(x_0)}{i!} \int_a^b (x - x_0)^i dx \\ &= \sum_{i=0}^{\infty} \frac{f^{(i)}(x_0)}{i!} \left[ \frac{(b - x_0)^{i+1}}{i+1} - \frac{(a - x_0)^{i+1}}{i+1} \right] \end{aligned}$$

# Aproximação de Laplace (AL)

- Suponha que desejamos calcular  $\int_a^b e^{Mf(x)} dx$ . Suponha que existe um único máximo global para  $f(\cdot)$
- Seja  $x_0$  o máximo global de  $f(\cdot)$ . Então :

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2}f''(x_0)(x - x_0)^2 + R$$

$$R = O((x - x_0)^3)$$

- Então

$$f(x) \approx f(x_0) - \frac{1}{2}|f''(x_0)|(x - x_0)^2$$

# Cont.

## ■ Aproximação de Laplace:

$$\begin{aligned}\int_a^b e^{Mf(x)} dx &\approx e^{Mf(x_0)} \int_a^b e^{-M|f''(x_0)|(x-x_0)^2/2} dx \\ &= e^{Mf(x_0)} \sigma \left\{ \Phi\left(\frac{b-x_0}{\sigma}\right) - \Phi\left(\frac{a-x_0}{\sigma}\right) \right\}\end{aligned}$$

$\sigma = \sqrt{\sigma^2}$ ,  $\sigma^2 = (M|f''(x_0)|)^{-1}$ ,  $\Phi(\cdot)$  é a fda da normal padrão.

# Característica da AL

- Relativamente rápida.
- Adequada quando o integrando é unimodal e o máximo pode ser obtido facilmente (ainda que seja necessário utilizar métodos numéricos).
- Inadequada: em integrais múltiplas à medida que a dimensão aumenta e/ou o máximo é complicado de ser obtido.
- Mesmo no caso univariado para funções complicadas: vários máximos locais, assimetria, etc.

## Exemplo: função beta incompleta

- Função:  $f(x) = \int_a^b \frac{x^{\alpha-1}(1-x)^{\beta-1}}{\beta(\alpha,\beta)} dx$
- Pode se provar que  $x_0 = \frac{\alpha-1}{\alpha+\beta+2}$  e  $\sigma^2 = \left( \frac{\alpha-1}{x_0^2} + \frac{\beta-1}{(1-x_0)^2} \right)^{-1}$ .
- Para  $\alpha = \beta = 3$ , temos que  $x_0 = 0,5$  e  $\sigma^2 = 0,0625$ . Assim:

Intervalo [a,b]	R	AL
[0, 50; 0, 55]	0,09313	0,09313
[0, 45; 0, 65]	0,35796	0,35837
[0, 30; 0, 70]	0,67384	0,67712
[0, 20; 0, 80]	0,88416	0,90457
[0, 10; 0, 90]	0,98288	1,04621

## Exemplo: posteriori da distribuição $\text{gama}(r, \lambda = 1)$

- $f(x) = \frac{1}{\Gamma(r)} e^{-x} x^{r-1} \mathbb{1}_{(0, \infty)}(x)$
- Amostra aleatória de tamanho  $n$  e  $p(r) \propto \mathbb{1}_{(0, \infty)}(r)$ .
- Posteriori :

$$p(r|\mathbf{x}) = \frac{\prod_{i=1}^n x_i^{r-1} \Gamma(r)^{-n}}{\int_0^\infty \prod_{i=1}^n x_i^{r-1} \Gamma(r)^{-n} dr}$$

- Seja  $g(r) = \prod_{i=1}^n x_i^{r-1} \Gamma(r)^{-n}$ .
- O máximo de  $g(\cdot)$  tem de ser obtido numericamente (veremos mais a frente).



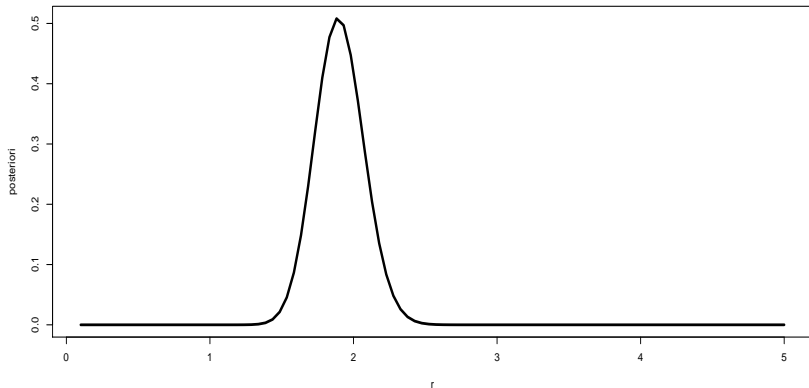
# Ingredientes

- $h(r) = e^{(\ln(g(r)))}$ ,  $M = 1$ .
- $[\ln g(r)]' = -n\Psi(r) + \sum_{i=1}^n \ln x_i$ ,  $\Psi(r) = \frac{\Gamma'(r)}{\Gamma(r)}$
- $[\ln g(r)]'' = -n[\Psi'(r)]$
- $\int_0^\infty h(r) dr \approx e^{g(r_0)} \sqrt{2\pi|\Psi'(r_0)|} \{1 - \Phi(\frac{-r_0}{\sigma})\}$ ,  $\sigma^2 = (|\Psi'(r_0)|)^{-1}$ .

# Exemplo numérico

- Dados simulados,  $n = 50$  e  $r = 8$ .
- O ponto de máximo foi obtido através do algoritmo Escore de Fisher.
- $r_0 = 1,893$ ,  $\int_0^\infty h(r)dr \approx 102743122$ ,  $\sigma = 1,203$

# Aproximação da posteriori



# Integração por Quadratura

- Substituir (no caso univariado) o cálculo da área sob a curva pela soma das áreas de um número finito de retângulos.
- Substituir (no caso multivariado) o cálculo da área sob a curva pela soma das áreas de um número finito de paralelepípedos.
- Ponto-chave: Definir os pontos e pesos de quadratura.
- Formas de cálculo dos pontos e pesos:
  - Não-adaptativa: mantem-se fixo os pontos e pesos.
  - Adaptativa: muda-se os pontos e pesos.

# Características

- Em geral, na Estatística, deseja-se calcular  $\mathcal{E}(w(X)) = \int_a^b w(x)f(x)dx$ , onde  $f(\cdot)$  é alguma fdp.
- Sendo assim,  $f(\cdot)$  terá massa relevante em apenas um subconjunto de  $\mathcal{R}$ .
- Define-se um conjunto de pontos,  $x_1, x_2, \dots, x_m$  e os respectivos pesos associados  $A_1, A_2, \dots, A_m, A_i = A(x_i)$ . Em geral, os pesos correspondem aos comprimentos dos intervalos determinados pelos pontos de quadraturas.
- Pode-se imaginar que cada ponto corresponde ao valor médio de intervalos de mesmo comprimento.

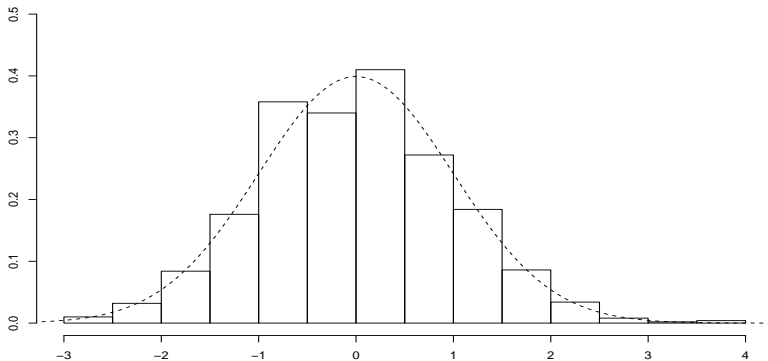
# Cont.

- Assim

$$\int_a^b w(x)f(x)dx \approx \sum_{i=1}^m A_i w(x_i)f(x_i)$$

- Existem várias formas de se determinar os pontos e pesos. Em geral, faz-se, em torno do máximo (moda) de  $f(\cdot)$ . Pontos ótimos, em geral, são obtidos através de certos polinômios.
- Polinômios de Gauss-Hermite, Gauss-Legendre, Jacobi, Legendre, Gauss-Laguerre etc. Veja Abramowitz & Stegun (1972).

# Graficamente



# Exemplo: família de localização escala

- Seja  $Y = \mu + \sigma Z$ ,  $Z \sim D(\theta)$ ,  $\mu \in \mathcal{R}$ ,  $\sigma \in \mathcal{R}^2$ ,  $\theta \in \mathcal{R}^p$ .
- Exemplos:
  - 1  $Z_1 \sim N(0, 1)$ .
  - 2  $Z_2 \sim t_\nu$ .
  - 3  $Z_3 \sim SN(0, 1, \lambda)$ .
- Particularmente, o máximo de  $Z_1$  e  $Z_2$  são iguais à 0.
- Qualquer integral envolvendo estas variáveis podem ser facilmente resolvidas.



# Distribuição Normal

- Seja  $f(x) \sim N(\mu, \sigma^2)$  e defina  $h = \int_a^b w(x)f(x)dx$ .
- Defina  $m$  pontos igualmente espaçados,  $x_1, x_2, \dots, x_m$ ,  $x_1 = \mu - 4\sigma$  e  $x_m = \mu + 4\sigma$ .
- Se  $x_1 < a$ , faça  $x_1 = a$ , se  $x_m > b$ , faça  $x_m = b$ .
- Calcule  $h \approx \sum_{i=1}^m A_i w(x_i) f(x_i)$ .

# Distribuição t de Student

- Seja  $f(x) \sim t_\nu(\mu, \sigma^2)$  e defina  $h = \int_a^b w(x)f(x)dx$ .
- Defina  $p_1$  e  $p_2$  ( $p_1 < p_2$ ) probabilidades de interesse.
- Dado  $\nu$ , calcule  $y_1, y_2$  tais que  $p_1 = F(y_1)$ ,  
 $p_2 = F(y_m)$ ,  $Y \sim t_\nu(0, 1)$ .
- Defina  $m$  pontos igualmente espaçados,  $y_1, y_2, \dots, y_m$ . Depois,  
calcule  $x_i = \sigma y_i + \mu$ .
- Se  $x_1 < a$ , faça  $x_1 = a$ , se  $x_m > b$ , faça  $x_m = b$ .
- Calcule  $h \approx \sum_{i=1}^m A_i w(x_i) f(x_i)$ .

# Distribuição $\text{gama}(r, \lambda)$

- Seja  $f(x) \sim \text{gama}(r, \lambda)$  e defina  $h = \int_a^b w(x)f(x)dx$ .
- Defina  $p_1$  e  $p_2$  probabilidades de interesse.
- Dado  $\text{gama}(r, \lambda)$ , calcule  $x_1, x_2$  tais que  $p_1 = F(x_1)$ ,  $p_2 = F(x_m)$ .
- Defina  $m$  pontos igualmente espaçados,  $x_1, x_2, \dots, x_m$ .
- Se  $x_1 < a$ , faça  $x_1 = a$ , se  $x_m > b$ , faça  $x_m = b$ .
- Calcule  $h \approx \sum_{i=1}^m A_i w(x_i) f(x_i)$ .

# Quadratura Adaptativa

- Atualização dos pontos e pesos de quadratura.
- Nosso interesse continua sendo  $h = \int_a^b w(x)f(x)dx$ .
- Algoritmo
  - Calcule duas aproximações para  $h$ ,  
$$S_1 = S(f(x); a, b) = \sum_{i=1}^m A_i w(x_i) f(x_i)$$
 e  
$$S_2 = S(f(x); a, (a+b)/2) + S(f(x); (a+b)/2, b)$$
  - Calcule  $\epsilon = |S_1 - S_2|$ .
  - Se  $\epsilon < \tau$ , pare, caso contrário, aplique o passo 1, recursivamente, até que a precisão requerida seja atingida.

# Quadratura Adptativa no R

- Função **integrate** (já vem instalada nas versões mais recentes).
- Permite calcular integrais, usando a quadratura adaptativa, para funções pré-definidas pelo usuário, em um dado intervalo.

# Integração por Monte Carlo

- Lei Forte dos Grandes Números: Seja  $X_1, X_2, \dots$  uma sequência iid, tal que  $\mathcal{E}(X_i) = \mu, \forall i$ . Então

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{q.c.} \mu$$

- Integração por Monte Carlo.

- 1 Gere um conjunto de  $m$  variáveis aleatórias i.i.d, digamos

$x_1, x_2, \dots, x_m$ .

- 2 Calcule  $h \approx \frac{1}{n} \sum_{i=1}^m w(x_i)$ .

## Exemplo: Integral trigonométrica

- Seja  $h = \int_0^{100} \cos(x)e^{-\text{seno}(x)} dx$ .
- Pode-se provar que  $h = e^{-0} - e^{0,5063656} = -0,6592498$
- Faça  $Y = X/100$ , assim  $h = 100 \int_0^1 \cos(100y)e^{-\text{seno}(100y)} dy = 100 * E(\cos(100y)e^{-\text{seno}(100y)})$ ,  $Y \sim U(0, 1)$ .
- Simular  $m$  valores iid de  $Y \sim U(0, 1)$  e calcule  $100 * \frac{1}{m} \sum_{i=1}^m \cos(100y_i)e^{-\sin(100*y_i)}$ .

## Exemplo: Integral trigonométrica

- Seja  $w = -\text{sen}(x)$ . Assim  $h = -\int_0^{0,5063656} e^u du$ . Assim, por exemplo  $-0,5063656\mathcal{E}(e^X), \sim U(0; 0,5063656)$  ou  $-e^{0,5063656}\mathcal{E}(e^{2X}), X \sim \text{exp}_{(0,5063656)}(1)$ .



# Integração por Amostragem por Importância

- Muito esforço para simular pontos com massa desprezível
- Pode ser complicado definir o valor esperado e/ou a distribuição apropriadas.
- Refinar a geração dos números aleatórios: método da amostragem por importância.

# Algoritmo

- Gere uma amostra aleatória,  $X_1, \dots, X_m$  de uma densidade candidata  $g(\cdot)$  com o mesmo suporte  $f(\cdot)$ .
- Assim  $h = \int_a^b w(x)f(x)dx = \int_a^b w(x)\frac{f(x)}{g(x)}g(x)dx$ .
- Construa pesos de importância

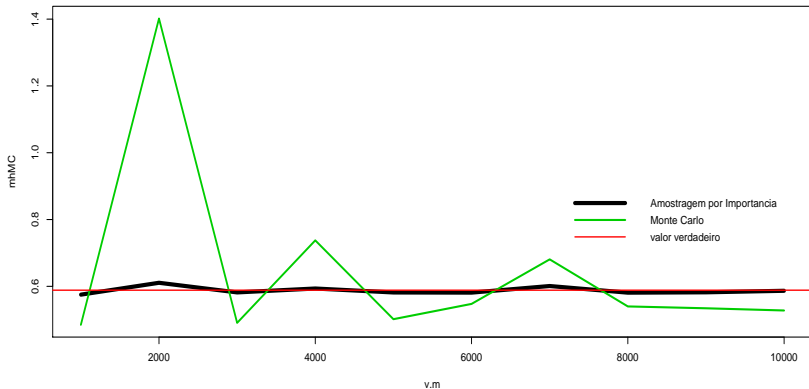
$$W(X_i) = \frac{f(X_i)}{g(X_i)}$$

- Calcule  $h \approx \frac{1}{m} \sum_{i=1}^m W(X_i)w(X_i)$

# Exemplo: Integral uniforme

- $h = \int_{-1}^1 e^{-90(x_1-0,5)^4} dx.$
- Densidade proposta:  $N_{[-1,1]}(0, 5; 180^{-1}).$
- Comparar com IMC, ou seja, estimar  $\mathcal{E}(e^{-90(X_1-0,5)^4}), X_1 \sim U[-1, 1].$

# Comparação: API x MC



# Alternativa API: estimador ponderado

- Estimador alternativo

$$\hat{\mu} = \sum_{i=1}^m \omega_i w(X_i)$$

$$\omega_i = \frac{W(X_i)}{\sum_{i=1}^m W(X_i)} = \frac{f(X_i)/g(X_i)}{\sum_{i=1}^m f(X_i)/g(X_i)}$$

- Assim, basta conhecer  $f(\cdot)$  e  $g(\cdot)$  a menos de constantes.