X é uma V.A. definida no conjunto $[1, \dots, N]$

$$\mathbb{E}[g(X)] = \underbrace{\sum_{i=1}^{N} g(i)}_{G_N} \overbrace{f_X(i)}^{P[X=i]}$$

distribuição uniforme

Para que $f_X(i)$ independa de i, fazemos $P[X=i] = \frac{1}{N}$.

Assim, o somatório passa a ser

$$\mathbb{E}[g(x)] = \sum_{i=1}^{N} g(i) \cdot \frac{1}{N} = \frac{1}{N} \cdot \sum_{i=1}^{N} g(i) = \frac{1}{N} \cdot G_N$$

$$G_N = N \cdot \mathbb{E}[g(X)]$$ → Não há nada de aleatório aqui, pois N é um parâmetro.

Nosso problema agora é estimar $\mathbb{E}[g(X)]$ que pode ser feito através da média amostral M_N

→ Realização da variável aleatória

$$M_N = \frac{1}{N} \cdot \sum_{j=1}^{N} g(X_j)$$

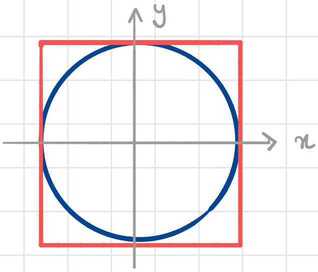
Sabemos que $M_n \xrightarrow[n \to \infty]{} \mathbb{E}\, g(X)$. Perceba que M_N é aleatório, pois M_N é calculado a partir de n amostras aleatórias de uma V.A.. Se M_n é aleatório, então ele possui um valor esperado:

$$\mathbb{E}[M_n] = \mathbb{E}\left[\frac{1}{n}\sum_{j=1}^{n} g(x)\right] = \mathbb{E}[g(x)]$$

Assim, o valor esperado do meu estimador é o valor esperado do que se quer estimar - dizemos então que o estimador é não-enviesado.

No entanto, temos que $\mathrm{Var}[M_n] = \dfrac{\mathrm{Var}[g(X)]}{n}$. Portanto, o número de amostras n impacta diretamente na variância de M_n.

Calculando π



$$\frac{A_c}{A_q} = \frac{\pi}{4}$$

$$g(x,y) = \begin{cases} 1, & \text{se } x^2+y^2 \leq 1 \\ 0, & \text{caso contrário} \end{cases}$$

Definimos 2 V.A.: $X \sim \mathcal{U}[-1,1]$ e $Y \sim \mathcal{U}[-1,1]$
$f_X(x) = 1/2$ $\quad$ $f_Y(y) = 1/2$

$$\mathbb{E}[g(X,Y)] = \int_{-1}^{1}\int_{-1}^{1} g(x,y)\cdot f_{XY}(x,y)\,dx\,dy$$

Assumindo X e Y independentes: $f_{XY}(x,y) = f_X(x)\,f_Y(y)$

Portanto: $\mathbb{E}[g(x,y)] = \dfrac{1}{4}\cdot\displaystyle\int_{-1}^{1}\int_{-1}^{1} g(x,y)\,dx\,dy$

Resolvendo a integral dupla: $\mathbb{E}[g(x,y)] = \dfrac{\pi}{4}$

Assim, $\pi = 4 \cdot \mathbb{E}[g(x,y)]$ e chegamos em um momento em que podemos encontrar o que queremos estimar a partir de um valor esperado, resolvível por amostragem.

$$M_n = \frac{1}{n} \sum_{i=1}^{n} g(x,y)$$

■ O Erro de π

Pela desigualdade de Chebyschev:

$$P\left[\pi_M \in [\mu - \varepsilon, \mu + \varepsilon]\right] \geq 1 - \frac{\sigma^2}{\varepsilon^2 n} = \beta$$

Para calcularmos a confiança β, precisamos apenas calcular σ^2 (ε e n são parâmetros).

Considerando $I = \begin{cases} 1 & \text{com probabilidade } p \\ 0 & \text{" \quad " } 1-p \end{cases}$, temos

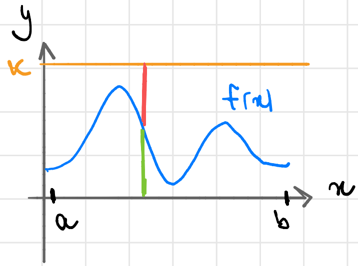
$\sigma^2 = \mathrm{VAR}(I) = p(1-p) = p - p^2 \rightarrow$ máximo em $p = 1/2$

Em seu valor máximo, $\sigma^2(p = 1/2) = 1/4$. Assim, para $p = 1/2$ garantimos um número mínimo de n que nos garanta uma confiança β.

$$\mathbb{E}[(M_n-\mu)^2] = \mathbb{E}[(M_n - \mathbb{E}[M_n])^2] \doteq \text{VAR}(M_n) = \frac{\sigma^2}{n}$$

Assim, o MSE de uma média amostral é igual à sua variância, que, como vimos, é $\frac{\sigma^2}{n}$

Integração numérica



Novamente, temos 2 V.A.: $X \sim \mathcal{U}[a,b]$ e $Y \sim [0,K]$.

Definimos uma indicadora

$$g(x,y) = \begin{cases} 1, & \text{se } y \leq f(x) \\ 0, & \text{caso contrário} \end{cases}$$

$$\mathbb{E}[g(x,y)] = \int_y \int_x g(x,y) \cdot f_{XY}(x,y)\, dx\, dy$$

$$= \int_y \int_x g(x,y) \cdot f_{\underline{XY}}(x,y)\, dx\, dy$$

$$= K \cdot \int_x f(x)\, dx$$

↳ sem ser conjunta

TRABALHAR A INTEGRAL DUPLA DA INDICADORA PARA CHEGAR NA INTEGRAL DE $F(x)$.

Dica: TRABALHAR NOS LIMITES DE INTEGRAÇÃO

Método Alternativo

$$I = \int f(x)\,dx \qquad\qquad X \sim \mathcal{U}[0,1]$$

$$\mathbb{E}_{f_X(x)} f(x) = \int f(x) \cdot \overbrace{f_X(x)}^{\text{densidade da V.A. } x}_{=1} \, dx = \int f(x)\,dx = I$$

Para aplicar MC, pegamos $X_i \sim [0, \dots, 1]$, sorteamos n valores e aplicamos a média amostral de $f(x)$. Assim, encontramos uma estimativa em cima de uma V.A. uniforme.

Genericamente, se pensarmos $X \sim \mathcal{U}[a,b]$, então

$$f_X(x) = \frac{1}{b-a} \longrightarrow \mathbb{E}_{f_X(x)} f(x) = \frac{1}{(b-a)} I.$$