Analiza greški

- Kako je procijeniti kad ne znamo egzaktnu vrijednost integrala?
- Varijanca uzorka?

$$\tilde{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}[f(x_i) - < f >]^2$$

$$< f >= \frac{1}{n}\sum_{i=1}^{n} f(x_i)$$

- *n* tipično veliki pa se često računa varijanca

$$\sigma^2 =< f^2 > - < f >^2,$$

$$< f^2 >= \frac{1}{n}\sum_{i=1}^{n}[f(x_i)]^2.$$

- Na ovaj način dobivamo grešku koja je veća od stvarne

- Bitna je varijanca srednjih vrijednosti

$$\sigma_m^2 = \frac{\sigma^2}{n}$$

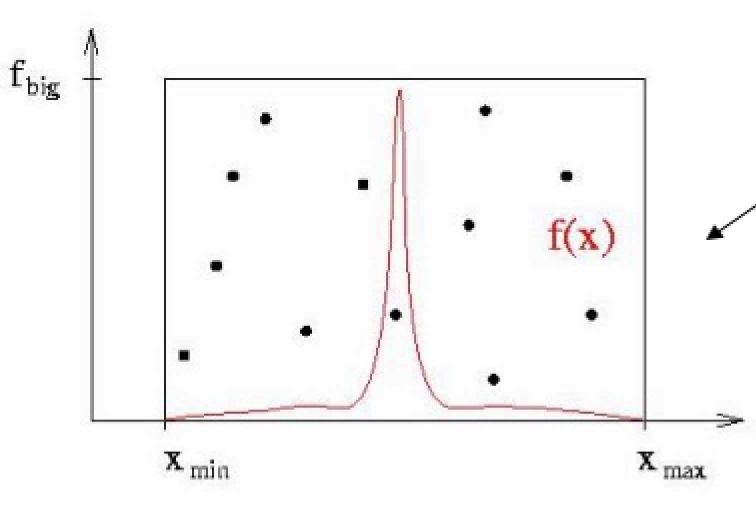
- Ona procjenjuje koliko prosjek od *n* mjerenja odstupa od egzaktnog prosjeka
- Sa 68% vjerojatnosti smo unutar jedne σ_m od prave vrijednosti, 97% unutar dvije σ_m

- Ponekad su susjedni pomaci korelirani, npr.

$$x_0, \ldots, x_i = x_{i-1} + (2r-1)\delta \qquad r \in [0,1],\ \delta = 0.01$$

- $\sigma/\sqrt{n}$ nije prava mjera greške

Značajnost odabira



Neefikasno

- Potrebno birati točke češće u područjima gdje je iznos podintegralne funkcije veći ili se brzo mijenja
- Potrebne metode za generiranje pseudoslučajnih brojeva koji nisu jednoliko raspoređeni

Nejednolike razdiobe
– inverzna transformacija

Vjerojatnost odabira i-tog događaja kada znamo vjerojatnosti pojedinih događaja u nizu od n diskretnih događaja

$$\sum_{j=0}^{i-1} p_j \leq r \leq \sum_{j=0}^{i} p_j.$$

Ako imamo kontinuiranu funkciju gustoće vjerojatnosti $p(x)$, definiramo kumulativnu funkciju razdiobe vjerojatnosti

$$P(x) \equiv \int_{-\infty}^{x} p(x')dx' = r. \qquad\qquad \int_{-\infty}^{\infty} p(x)dx = 1$$

Generiranje slučajnih brojeva iz $[a, b]$ jednolike razdiobe

$$p(x) = \begin{cases} 1/(b-a) & a \leq x \leq b \\ 0 & \text{inače} . \end{cases} \quad \Rightarrow \quad P(x) = \frac{x-a}{b-a} \quad \xRightarrow{P(x) = r} \quad x = a + (b-a)r$$

Metoda inverzne transformacije

- Eksponencijalna razdioba

$$p(x) = \begin{cases} (1/\lambda)e^{-x/\lambda} & 0 \le x \le \infty \\ 0 & x < 0\,. \end{cases} \Rightarrow r = P(x) = 1 - e^{-x/\lambda}. \Rightarrow \begin{array}{l} x = -\lambda \ln(1-r) \\ x = -\lambda \ln r \end{array}$$

- Gaussova gustoća vjerojatnosti (nije praktično izvući r pa idemo u 2D)

$$p(x) = \frac{1}{(2\pi\sigma^2)^{1/2}} e^{-x^2/2\sigma^2} \Rightarrow p(x,y)dxdy = \frac{1}{(2\pi\sigma^2)} e^{-(x^2+y^2)/2\sigma^2} dxdy.$$

$$R = (x^2+y^2)^{1/2}\,, \qquad \theta = arctg(\tfrac{y}{x}) \qquad \begin{array}{l} \rho = R^2/2 \\ \sigma = 1 \end{array} \Rightarrow p(\rho,\theta)d\rho d\theta = \frac{1}{2\pi} e^{-\rho} d\rho d\theta$$

ρ odaberemo prema eksponencijalnoj razdiobi $p(\rho)$, a θ prema jednolikoj

$$p(\rho, \lambda = 1) = \begin{cases} e^{-\rho}, & 0 \le \rho < \infty \\ 0, & \rho < 0 \end{cases} \Rightarrow \begin{array}{l} x = \sqrt{2\rho}\cos\theta \\ y = \sqrt{2\rho}\sin\theta \end{array}$$

$$0 \le \theta < 2\pi$$

x i y generirani prema Gaussovoj (normalnoj) razdiobi

Z5 Značajnost odabira

$$I = \int_a^b f(x)dx = \int_a^b \frac{f(x)}{g(x)} g(x)dx = \int_a^b \frac{f(x)}{g(x)} dG(x)$$

$$G(x) = \int_a^x g(x)dx$$

$$r = G(x)$$

$$I = \int_{G(a)}^{G(b)} \frac{f(G^{-1}(r))}{g(G^{-1}(r))} dr$$

$$I = \frac{1}{N} \sum_{i=1}^{N} \frac{f(G^{-1}(r_i))}{g(G^{-1}(r_i))}$$

$$I = \frac{1}{N} \sum_{i=1}^{N} \frac{f(x_i^{(g)})}{g(x_i^{(g)})}$$

$$I = \int_0^1 e^{-x^2} dx$$

$$\text{erf}(1) * \sqrt{(}\pi)/2 = 0.746824$$

$$\int_0^1 e^{-x} dx = -\frac{1}{e} + 1 = 1 - \frac{1}{e} = \frac{e-1}{e}$$

$$g(x) = \frac{e^{-x}e}{(e-1)}$$

$$G(x) = \int_0^x \frac{e^{-x}e}{e-1} dx = \frac{(1-e^{-x})e}{e-1}$$

$$G^{-1}(u) = -\ln\left(1 - u\frac{e-1}{e}\right)$$

$$I = \frac{1}{N} \sum_{i=1}^{N} \frac{f(G^{-1}(r_i))}{g(G^{-1}(r_i))}$$

Z5 Značajnost odabira

Monte Carlo simulacijama usporedite vrijednosti integrala I za različite razdiobe $p(x)$ argumenta x, odnosno producirajte rezultate iz donje tablice (ne treba CPU vrijeme). Priložite kod (rezultate neka sprema u datoteku pa je priložite). [Skripta, pgl. 2.7.3]

$$I = \int_0^1 e^{-x^2} dx \cdot \frac{p(x)}{p(x)} \qquad F_n = \frac{1}{n} \sum_{i=1}^{n} \frac{f(x_i)}{p(x_i)}$$

	$p(x) = 1$	$p(x) = Ae^{-x}$
n (uzoraka)	5×10^6	4×10^5
F_n	0.74684	0.74689
σ	0.2010	0.0550
$\sigma/\sqrt{n}$	0.00009	0.00009
ukupno CPU vrijeme	20	2.5
CPU vrijeme po uzorku	4×10^{-6}	6×10^{-6}