Markus Laubenthal
Lennard Alms
Bilal Kizilkaya

Foundations of Graphics

Universität Bonn
Institut für Informatik II
November 12, 2019

Winter term 2019
Prof. Dr. Matthias Hullin

Sheet 4 - Monte Carlo Sampling

Please upload your solutions to
`https://uni-bonn.sciebo.de/s/XfRkWhpWDEeR8Y7`
by **Sa, 16.11.2019, 23:59 CET**.
Make sure to **list all group members** on all pages / source files.
Theoretical solutions must be texed or scanned, **photos will not be accepted**.

Practical Part

Assignment 1) Monte-Carlo Integration *(2+1=3Pts)*

This exercise will focus on the approximation of integrals using Monte Carlo integration.

a) Calculate the area of a circle, which is centered at the origin and has the radius $r = 3$, using Monte Carlo integration. Additionally the error between the real area and the approximated area should be considered, i.e. plot this error with respect to the number of points that are used for the Monte Carlo integration.

b) Extend the framework by additionally calculating the volume of a sphere, which is centered at the origin and has the radius $r = 3$, using Monte Carlo integration. Again, plot the error as a function of the number of points that is used for the approximation.

Assignment 2) Importance Sampling *(2+2+2+1=7Pts)*

In this exercise, we will use importance sampling to integrate the function

$$f(x) = x^{0.9} \cdot \exp\left(-\frac{x^2}{2}\right)$$

on the interval $[0 \ldots 4]$. For this we consider three different distributions (a uniform distribution p_u, a normal distribution p_n and a polynomial p_p that was fit to f) and will compare how their choice affects the resulting integral value.

a) Implement the functions f, p_n and p_p and plot them. The normal distribution p_n is given by the formula

$$p_n(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{(\mu - x)^2}{2 \cdot \sigma^2}\right)$$

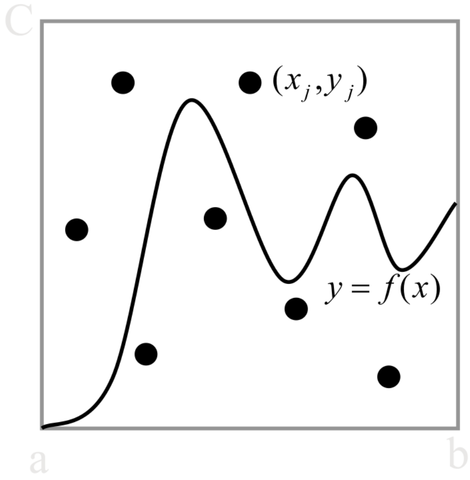
with $\sigma^2 = 0.3$ and $\mu = 1.2$.
The polynomial p_p is defined by being a 4th degree polynomial, with $p_p(x) = f(x)\, \forall x \in \{0, 1, 2, 3, 4\}$.
Hint: Familiarize yourself with `polyfit` and `polyval`.
Add a short comment on how you would expect the three different distributions to perform and why.

Markus Laubenthal
Lennard Alms
Bilal Kizilkaya

b) Rejection sampling is used when you cannot calculate the CDF in a simple way. To calculate the integral

$$I = \int_a^b f(x)dx = \iint\limits_{y<f(x)} dxdy$$

you construct a rectangle with known area $A_{rect} = C \cdot (b-a)$ and count the number of samples



fulfilling the criterion $y < f(x)$. The ratio of accepted samples to the total number of drawn samples gives the fraction of the rectangle area which approximates the area under the curve:

$$I \approx A_{rect} \cdot \frac{|\{samples_{accept}\}|}{|\{samples_{total}\}|}$$

Implement rejection sampling. Write a function that can generate an arbitrary amount of samples for a given distribution. You can assume that $x \in [0 \ldots 4]$ throughout this exercise.
To verify that your function works, create two scatter plots (showing p_n and p_p respectively), which show the graph of the distribution and all samples that were accepted. Use $n = 200$ samples (that is accepted samples!).

c) Integrate f using the three different distributions. Perform an analysis of how each one converges for an increasing amount of samples (up to 500) and plot the results. Find a good way to put all three graphs into one plot while still remaining distinguishable. Add a short comment on how the distributions perform and whether this fits your expectations.

d) You should have realized that the value of the integral slightly differs for the different kernels. Which one was correct? Explain what was done wrong in the calculation and how this could be fixed.

Theoretical Part

Assignment 3) Cumulative Distribution Function *(1+2=3Pts)*

a) Explain the relation between the cumulative distribution function and the probability density function.

b) Calculate the probability $P(0 < X \leq 1.5)$ that a real-valued random variable X will be found in the interval $(0, 1.5]$ for the probability density function

$$f(x) = \begin{cases} x & \text{if } x \in [0,1] \\ -x+2 & \text{if } x \in [1,2] \\ 0 & \text{else} \end{cases} \tag{1}$$

Markus Laubenthal
Lennard Alms
Bilal Kizilkaya

using the definition of the cumulative distribution function.

Good luck!

Assignment 3) Cumulative Distribution Function *(1+2=3Pts)*

a) Explain the relation between the cumulative distribution function and the probability density function.

b) Calculate the probability $P(0 < X \leq 1.5)$ that a real-valued random variable X will be found in the interval $(0, 1.5]$ for the probability density function

$$f(x) = \begin{cases} x & \text{if } x \in [0,1] \\ -x+2 & \text{if } x \in [1,2] \\ 0 & \text{else} \end{cases} \tag{1}$$

using the definition of the cumulative distribution function.

a) cumulative distribution is the integral of the PDF

At a point x cdf(x) gives the probability such that $P(\text{pdf}(x') < \text{pdf}(x)) = \text{cdf}(x) \; \forall x'$

b)

$$f(x) = \begin{cases} x & \text{if } x \in [0,1] & c_1 \\ -x+2 & \text{if } x \in (1,2] & c_2 \\ 0 & \text{else} \end{cases}$$

$F_1(x) = \frac{1}{2}x^2 \qquad (x=1) - (x=0)$

$F_2(x) = -\frac{1}{2}x^2 + 2x \qquad 1 \cdot 1{,}5$

$$\int_0^1 f(x) + \int_1^{1{,}5} f(x) = \frac{1}{2} - \frac{1}{2}1.5^2 + 3 + \frac{1}{2} - 2 = \underline{0.875}$$

Markus Laubenthal
Lennard Alms
Bilal Kizilkaya

2a) We expect p_p to perform second best because it has the lowest variance w.r.t. $f(x)$

$f(x)$ as pdf should perform best since the variance is 0

p_u should deliver the worst result.

2c)

Error of different distributions behaves as expected.

d)

d) You should have realized that the value of the integral slightly differs for the different kernels. Which one was correct? Explain what was done wrong in the calculation and how this could be fixed.

Markus Laubenthal
Lennard Alms
Bilal Kizilkaya

The ψ_p function approximates f best. Therefor, the weights used for the integral calculation are more accurate and the result will be better.