Chapter 1 - Random Number Generation

Generation of Random Vectors, A geometric interpretation

Prof. Alex Alvarez, Ali Raisolsadat

School of Mathematical and Computational Sciences
University of Prince Edward Island

The next result (Lemma 1.33 from the textbook) establishes a connection between distributions with general densities on $\mathbb{R}^d$ and uniform distributions on $\mathbb{R}^{d+1}$.

Lemma: Let $f : \mathbb{R}^d \to [0, \infty)$ be a probability density and let

$$A = \left\{(x, y) \in \mathbb{R}^d \times [0, \infty) : 0 \leq y < f(x)\right\} \subseteq \mathbb{R}^{d+1}$$

Then $|A| = 1$ and the following two statements are equivalent:

a) (X, Y) is uniformly distributed on A

b) X is distributed with density f on $\mathbb{R}^d$ and $Y = f(X)U$ where $U \sim U[0, 1]$ independently of X.

Interpretation for One-Dimensional Densities

In the case $d = 1$ we can easily interpret the previous result.

This Lemma could be used in both directions as needed:

- **First implication**: Starting with a sample uniformly distributed over A, we can look at the first component and get a sample that follows density f(this is very related to the usual sampling rejection algorithm). Problem 2 from the homework uses this result.

Interpretation for One-Dimensional Densities

In the case $d = 1$ we can easily interpret the previous result.

This Lemma could be used in both directions as needed:

- **First implication**: Starting with a sample uniformly distributed over A, we can look at the first component and get a sample that follows density f(this is very related to the usual sampling rejection algorithm). Problem 2 from the homework uses this result.
- **Second implication**: We can sample from density f and the uniform distribution to get a uniformly distributed sample over set A.

We could use the second implication in order to get uniformly distributed samples over an unbounded set A.

Example: Generate a sample of 100 random points uniformly distributed under the curve of the probability density function of a standard normal random variable.

Algorithm:

1. Generate $X \sim N(0, 1)$
2. Compute $Z = pdf(X)$
3. Generate $Y \sim U[0, Z]$
4. Return (X, Y)

Code

Code in R

```r
n <- 1000;
X <- rnorm(n);
Z <- dnorm(X)
Y <- runif(n,min=0, max=Z)
plot(X,Y)
```

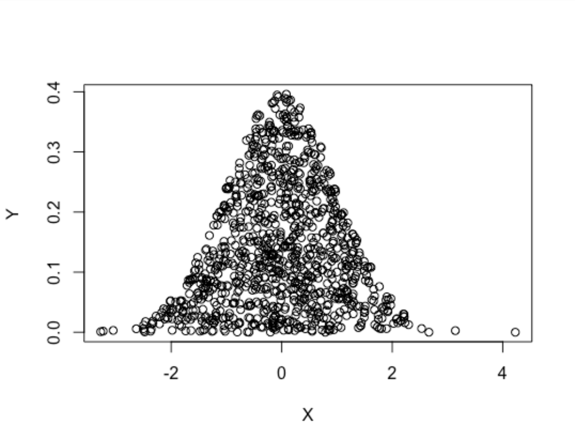
Python Code

```python
import numpy as np
import matplotlib.pyplot as plt
from scipy.stats import norm

n = 1000
X = np.random.normal(size=n)
Z = norm.pdf(X)
Y = np.random.uniform(low=0, high=Z)

plt.scatter(X, Y)
plt.show()
```

You can program this with a ”for” loop instead, but it would be less efficient.

Visualizing the Uniform Sample Under a PDF

Generated Sample uniformly distributed over the region under the density function of a standard normal random variable.



Remarks:

- This is a very straightforward method as long as we know how to sample from the density f.
- It can be applied also when $d \geq 1$, to both bounded and unbounded sets.
- By design the method only applies to sets A with $|A| = 1$ (meaning that area equals 1 for $d = 1$, volume equals 1 for $d = 2$, etc.) but we can modify it to also take into account other sets where $|A| \neq 1$.

Homework

1. Generate a sample of random points uniformly distributed under the graph of the density of the chi-square distribution with 4 degrees of freedom.
2. Consider the probability density function given by $f(x) = \frac{1}{(x+1)\ln 2}$ on $[0,1]$ and 0 otherwise. Define B as follows

$$B = \left\{(x,y) \in \mathbb{R}^2 : 0 \leq x \leq 1, 0 \leq y < f(x)\right\}$$

and let A be the rectangle $A = \left\{(x,y) \in \mathbb{R}^2 : 0 \leq x \leq 1, 0 \leq y < 1/\ln 2\right\}$
We can see that $B \subset A$.
a) Write a computer program that generates a random sample of 1000 uniformly distributed random points in B, by applying rejection sampling to uniformly distributed random points in A.
b) Use the outcome from part a) to get a random sample of 1000 random numbers with probability density f.