# Rejected

Our first look at Monte-Carlo integration showed that when we could sample a uniform random variable from our domain $\Omega$, then

$$\text{volume}(\Omega)\ f_{ave} = \int_\Omega f(X)\ d\Omega$$

could be used to approximate $\int_\Omega f(X)\ d\Omega$ as

$$\int_\Omega f(X) d\Omega = \text{volume}(\Omega)\hat{f} + \text{O}(1/\sqrt{N}) = S + \text{O}(1/\sqrt{N})$$

where $\hat{f} = (\sum_{i=1}^N f(x_i))/N$ is our computed sample mean.

Typically $\Omega \subset \mathbb{R}^d$ and $d$ is large. If $\Omega$ is a box, say $\Omega = [a_1, b_1] \times [a_2, b_2] \times \ldots \times [a_d, b_d]$, then $\text{volume}(\Omega) = \prod_{j=1}^d (b_j - a_j)$, and sampling from $\Omega$ is a matter of creating a vector $x = (x_1, x_2, \ldots, x_n)$, where $x_i$ is a uniform (pseudo-)random variable on $[a_i, b_i]$. But usually $\Omega$ isn't so simple, and it is a challenge to compute volume($\Omega$) and/or to sample from a uniform distribution on $\Omega$. One technique that addresses both challenges is *rejection sampling.*

Consider a domain $D \subset \mathbb{R}^d$ for which (i) $\Omega \subset D$, (ii) volume($D$) is known (or at least, easy to compute), and (iii) we can sample from a uniform distribution on $D$. If we sample $x \in D$, then either $x \in \Omega$ or not. The ratio of the number of points in $\Omega$ to those in $D$ approaches the ratio volume($\Omega$)/volume($D$). Furthermore (and this is vital): the sampling of those $x \in \Omega$ is from a uniform distribution *on* $\Omega$. Therefore, if we sample $M$ times from $D$ and $N$ of those are in $\Omega$, then $\lim_{N\to\infty} N/M = \text{volume}(\Omega)/\text{volume}(D)$ and in fact

$$\int_\Omega f(X) dV = \left[\frac{N}{M}\text{volume}(D)\right] \left[\frac{\sum_{i=1}^N f(x)}{N}\right] \ +\ \text{O}(1/\sqrt{N}).$$

Here is the basic rejection sampling, giving $\text{I} \approx \int_\Omega f(X) d\Omega$ :

```
s = 0;  j = 0;  M = 0;
while j < N,
    sample x from uniform distribution on D
    if x is in Omega then
        j = j + 1
        s = s + f(x)
    end_if
    M = M + 1
end_while
I = s*volume(D)/M
```

It is important that $\Omega$ fit "snugly" in $D$: The smaller $D - \Omega$, the fewer samples are rejected. Imagine sampling air in a classroom and rejecting points not in you, versus sampling from a box around your desk and rejecting points not in you.

Often the biggest computation in the loop is to decide if $x \in \Omega$ or not, and this question must be answered for each $x \in D$.