# Lab 5

*Wiktor Soral*

*October 31 2017*

## Monte Carlo simulations

### Calculating evidence

$P(\theta|\mathcal{D}) = \frac{P(\mathcal{D}|\theta) \times P(\theta)}{P(\mathcal{D})}$

posterior = likelihood $\times$ prior / evidence

How to calculate evidence?

### Calculating evidence

With only 2 or 3 values of $\theta$ it is fairly simple (recall HW 1):

$P(\mathcal{D}) = [P(\mathcal{D}|\theta = 1) \times P(\theta = 1)] + [P(\mathcal{D}|\theta = 2) \times P(\theta = 2)]$

In general we could write:

$P(\mathcal{D}) = \sum_\theta P(\mathcal{D}|\theta) \times P(\theta)$

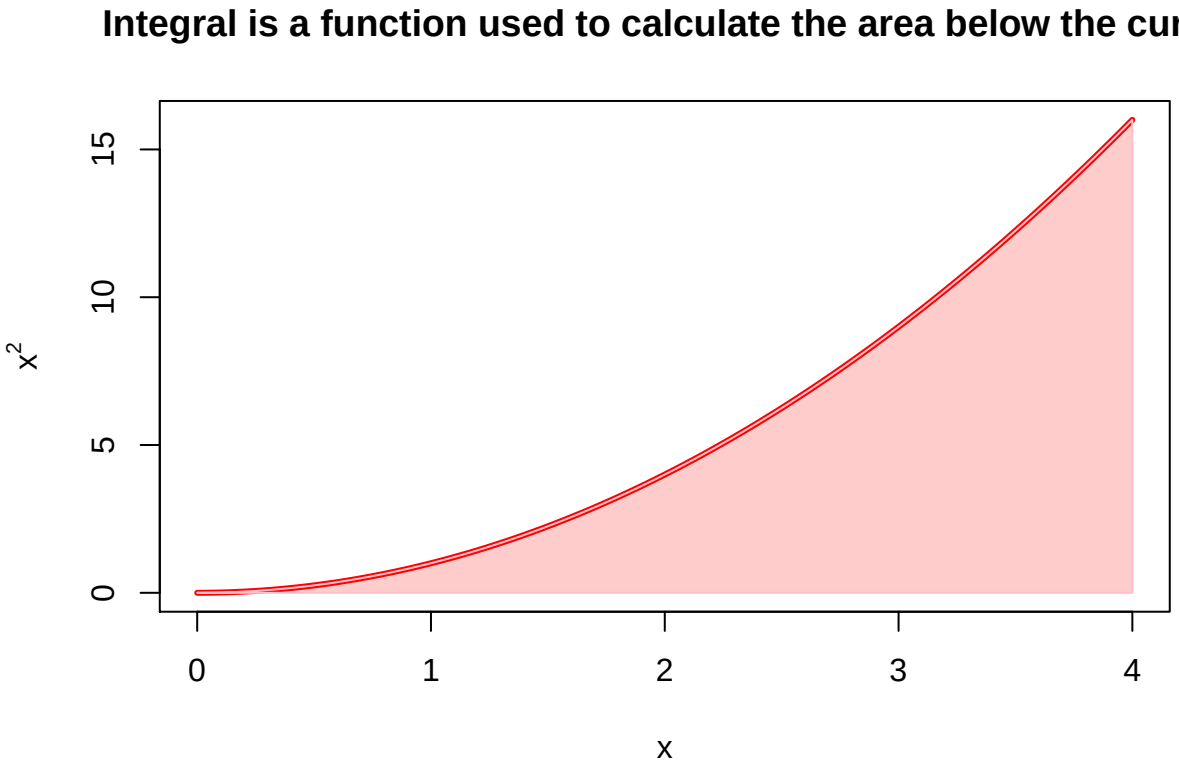
### Calculating evidence

Though, usually we will use $\theta$ with continuous scale (i.e. infinite number of values).

In math we denote summing across values of continuous scale as **integrating**:

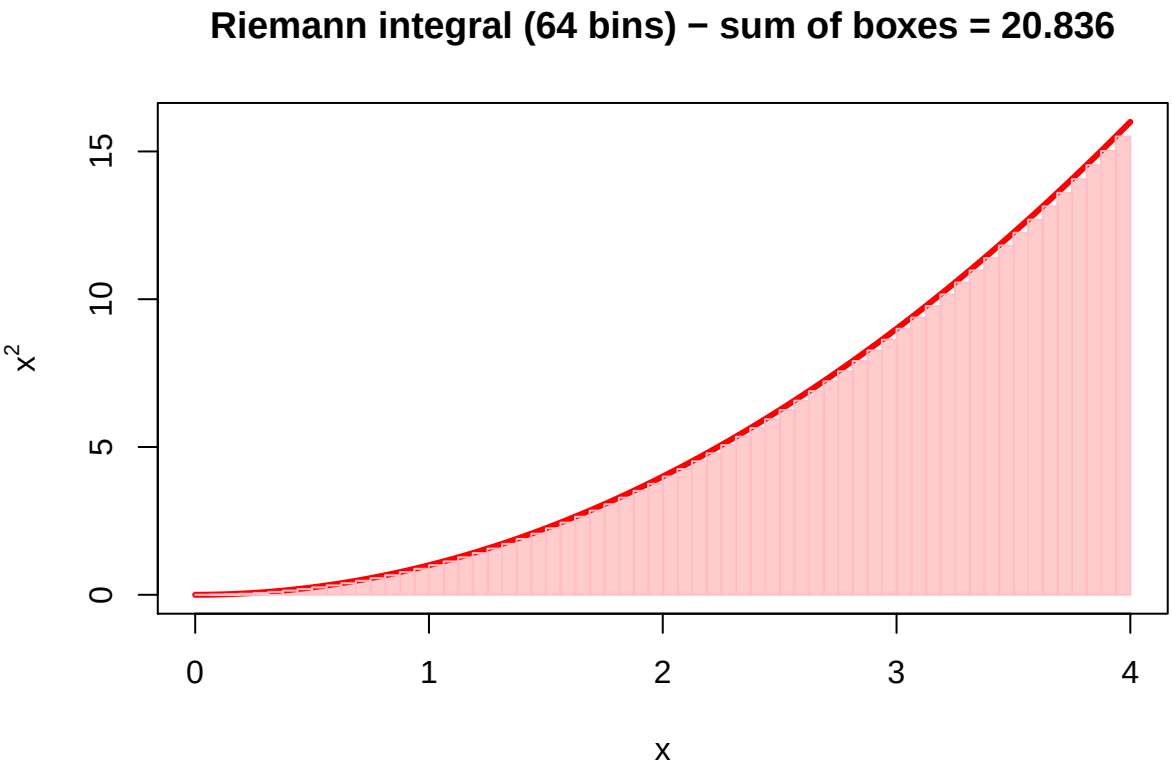
$P(\mathcal{D}) = \int_\theta P(\mathcal{D}|\theta) \times P(\theta) \mathrm{d}\theta$

**Calculating integrals - analytical app.**




We know from analytical solution: $\int x^2 \mathrm{dx} = \frac{x^3}{3} + const$

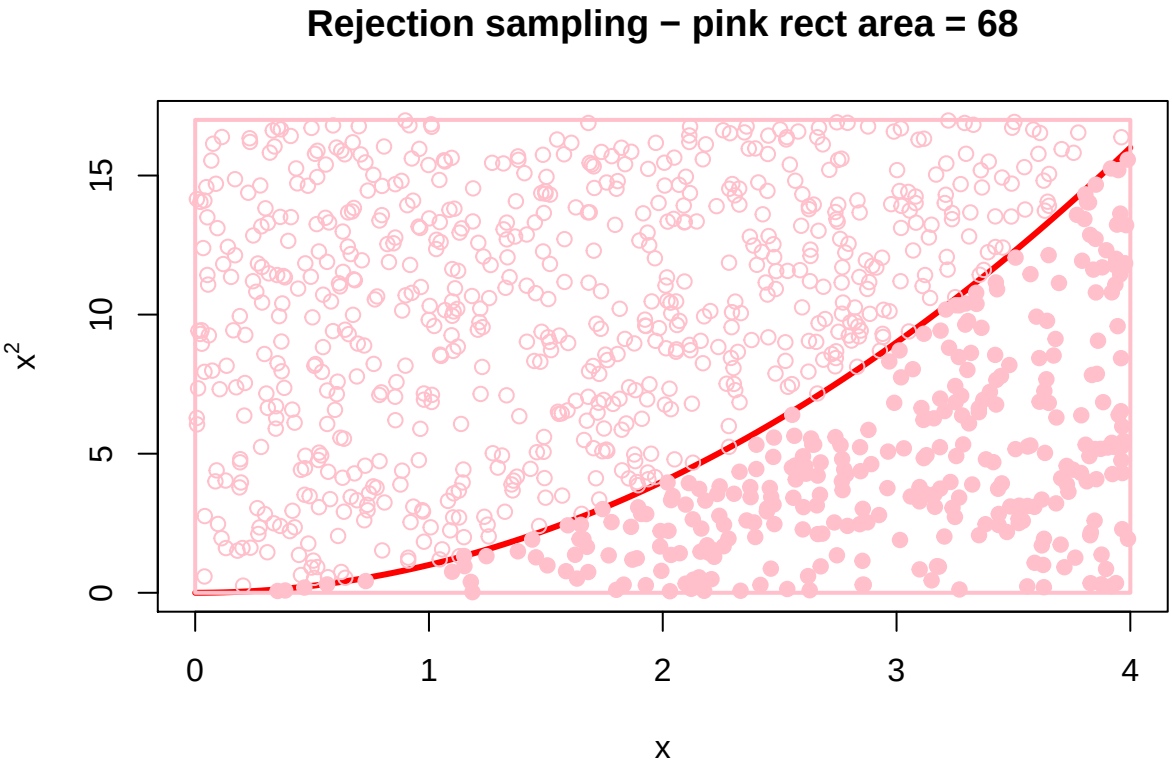
## Calculating integrals - numerical app.




## Computing integrals - numerical app.

- With 1 parameter/dimension CPU has to perform 100 operations
- With 2 parameters/dimensions CPU has to perform $100^2 = 10000$ operations
- With 3 parameters/dimensions CPU has to perform $100^3 = 1000000$ operations
- With 10 parameters/dimensions CPU has to perform $100^{10} = 10000000000000000000$ operations
- Numerical approach quickly becomes too hard even for fast computers

**Sampling**




E.g. 318 of 1000 sampled points are below the curve - 0.318. $68 \times 0.318 = 21.624$