# Practical: Metropolis-Hastings Algorithm
Instructors: Kari Auranen, Elizabeth Halloran and Vladimir Minin
July 17 – July 19, 2017

## Sampling from the standard normal distribution

Suppose our target is a univariate standard normal distribution with density $f(x) = 1/(\sqrt{2\pi})e^{-x^2/2}$. Given current state $x^{(t)}$, we generate two uniform r.v.s $U_1 \sim U[-\delta, \delta]$ and $U_2 \sim U[0, 1]$. Then set

$$
x^{(t+1)} = \begin{cases} x^{(t)} + U_1 & \text{if } U_2 \leq \min\left\{ e^{\left[ \left( x^{(t)} \right)^2 - \left( x^{(t)} + U_1 \right)^2 \right]/2}, 1 \right\} \\ x^{(t)} & \text{otherwise.} \end{cases}
$$

$\delta$ is a tuning parameter. Large $\delta$ leads to small acceptance rate, small $\delta$ leads to slow exploration of the state space. The rule of thumb for random walk proposals is to keep acceptance probabilities around 30-40%. If your proposal is close to the target, then higher acceptance rates are favorable.

### Your task

Implement the above algorithm. Experiment with the tuning parameter $\delta$ and report empirically estimate acceptance probabilities for different values of this parameter.

## Distribution of the time of infection

Consider a two state continuous-time Markov SIS model, where the disease status $X_t$ cycles between the two states: 1=susceptible, 2=infected. Let the infection rate be $\lambda_1$ and clearance rate be $\lambda_2$. Suppose that an individual is susceptible at time 0 ($X_0 = 1$) and infected at time $T$ ($X_T = 2$). We don't know anything else about the disease status of this individual during the interval $[0, T]$. If $T$ is small enough, it is reasonable to assume that the individual was infected only once during this time interval. We would like to obtain the distribution of the time of infection $I$, conditional on the information we have.

### Your task

Implement a Metropolis-Hastings sampler to draw realizations from the distribution

$$
\Pr(I \mid X_0 = 1, X_t = 2, N_t = 1) \propto \Pr(0 < t < I : X_t = 1, I < t < T : X_t = 2),
$$

where $N_t$ is the number of infections. Since $X_t$ is a continuous-time Markov chain, the last probability (it is actually a density) can be written as

$$\Pr(0 < t < I : X_t = 1, I < t < T : X_t = 2) = \underbrace{\lambda_1 e^{-\lambda_1 I}}_{\text{density of waiting time until infection}} \times \overbrace{e^{-\lambda_2(T-I)}}^{\text{prob of staying infected}} .$$

To make things concrete, set $\lambda_1 = 0.1$, $\lambda_2 = 0.2$ and $T = 1.0$. For your proposal distribution, use a uniform random walk with reflective boundaries $0$ and $T$. In other words, given a current value of the infection time $t_c$, generate $u = \text{Unif}_{[t_c-\delta,t_c+\delta]}$ $(2\delta < T)$ and then make a proposal value

$$t_p = \begin{cases} u & \text{if } 0 < u < T, \\ 2T - u & \text{if } u > T, \\ -u & \text{if } u < 0. \end{cases}$$

This is a symmetric proposal, so your M-H ratio will contain only the ratio of target densities:

$$\frac{\lambda_1 e^{-\lambda_1 t_p} e^{-\lambda_2(T-t_p)}}{\lambda_1 e^{-\lambda_1 t_c} e^{-\lambda_2(T-t_c)}} = e^{-\lambda_1(t_p-t_c)-\lambda_2(t_c-t_p)} = e^{(t_p-t_c)(\lambda_2-\lambda_1)}.$$

Plot the histogram of the posterior distribution of the infection time. Try a couple of sets of values for $\lambda_1$ and $\lambda_2$ and examine the effect of these changes on the posterior distribution of the infection time.