ABC simple example

Grace Yan

Coin toss example

Scenario: toss the coin 10 times. Coin lands heads with probability θ.

Setup:

- *Prior*: $\theta \sim U[0, 1]$.
- *True likelihood* (pretend this is unknown): $X_i \sim bin(10, \theta)$. The simulator simulates from this likelihood, which we don't know in closed form.
- *True posterior*: We know (by conjugacy) that the posterior is $beta(7, 5)$ in this example, but we pretend we don't know this and so we're trying to find the posterior through ABC.
- *Observed data*: one experiment resulting in 6 heads out of 10 tosses (1 experiment = 10 tosses). This is `x_obs` .

Algorithm:

We simulate a θ from the prior, then plug it into the simulator to get a simulated datapoint x_i. If $x_i = 6$ we accept the θ, otherwise reject θ. Do this to n θ's to get n datapoints which we call `x_sim` . The accepted θ's are the approximate posterior samples we want.

```
set.seed(147)
n <- 1e6
thetas = runif(n,0,1) # draw n thetas from prior
x_obs = 6 # observed data is 6 heads - just 1 experiment

simulator <- function(theta){
  x_sim <- rbinom(1,10,theta)
}

# Use 6 cores
library(parallel)
all_x_sim_parallel <- mclapply(thetas, simulator, mc.cores=6)

# Only accept thetas that give x_sim = x_obs (i.e. = 6), otherwise reject
ind = all_x_sim_parallel==x_obs
post = thetas[ind]
length(post) # Looks like the right proportion (should be around 1/10, since there
are 10 possible # heads)
```

```
## [1] 90253
```

Plot:

```
pts <- 0:1000 / 1000
hist(post, prob=TRUE, main="Approximated vs. Exact posterior distributions",xlab
='theta',ylab='p(theta|x=6)',xlim=c(0,1),col='#CEA9FB')
lines(pts, dbeta(pts,7,5), lwd=2, col='#B10000')
legend(0.03, 2.5, legend=c("ABC posterior", "Exact posterior"),
       col=c("#CEA9FB", "#B10000"), lwd=2, cex=0.8)
```