Spurious correlations

Bob Week

```r
require(matlab) # ones and zeros
```

```
## Loading required package: matlab

##
## Attaching package: 'matlab'

## The following object is masked from 'package:stats':
##
##     reshape

## The following objects are masked from 'package:utils':
##
##     find, fix

## The following object is masked from 'package:base':
##
##     sum
```

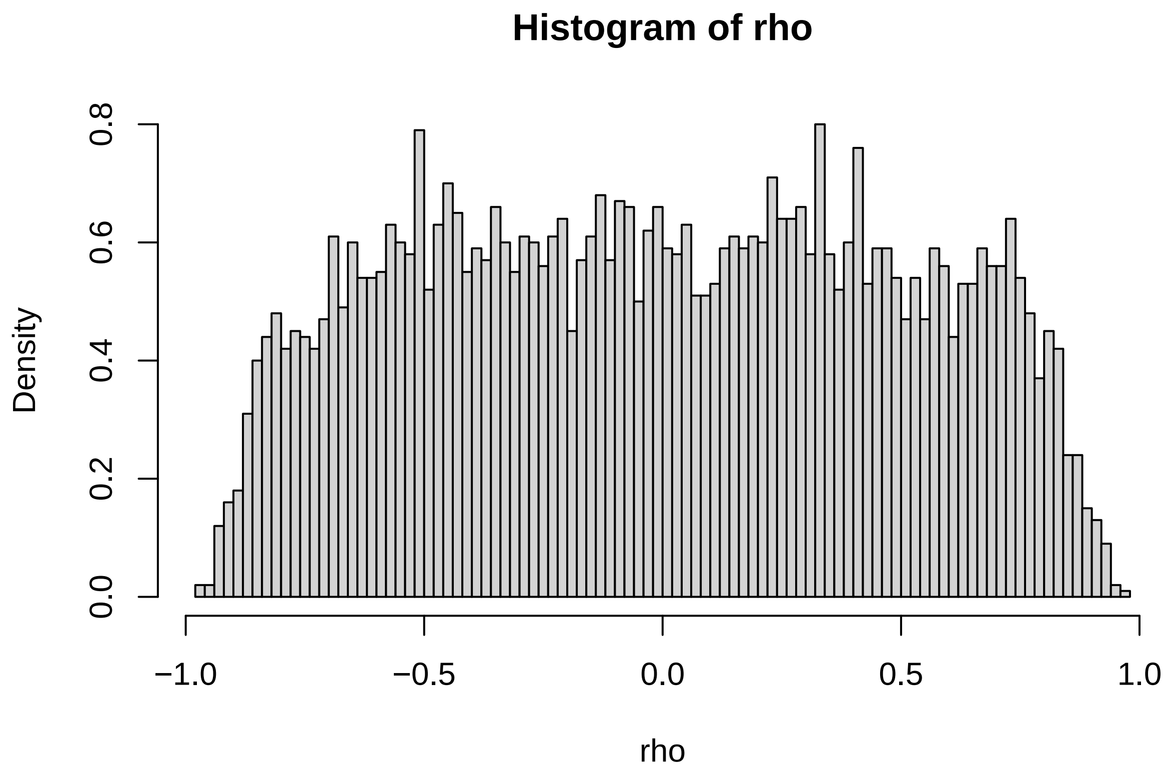
```r
require(Sim.DiffProc)
```

```
## Loading required package: Sim.DiffProc

## Package 'Sim.DiffProc', version 4.8
## browseVignettes('Sim.DiffProc') for more informations.
```

```r
k = 5000
rho = zeros(k,1)
dt = 0.001
for(i in 1:k){
  bms = BM(N=(1/dt),M=2)
  bm1bar = sum(bms[,1])*dt
  bm2bar = sum(bms[,2])*dt
  num = sum((bms[,1]-bm1bar)*(bms[,2]-bm2bar)*dt)
  den = sqrt(sum((bms[,1]-bm1bar)^2*dt)*sum((bms[,2]-bm2bar)^2*dt))
  rho[i] = num/den
}

hist(rho,freq=F,breaks = 100)
```



Set X_1 and X_2 two iid Brownian motions on $[0,1]$. We consider the spurious correlation of sample paths, measured by

$$\rho = \frac{\int_0^1 (X_1(t) - \bar{X}_1)(X_2(t) - \bar{X}_2)dt}{\sqrt{\int_0^1 (X_1(t) - \bar{X}_1)^2 dt \int_0^1 (X_2(t) - \bar{X}_2)^2 dt}}$$

where $\bar{X}_i = \int_0^1 X_i(t)dt$. Let's start with the covariance: $C = \int_0^1 (X_1(t) - \bar{X}_1)(X_2(t) - \bar{X}_2)dt$. Using independence and properties of Brownian motion, we have $\mathbb{E}C = 0$. To get the variance, we can use the fact that Brownian motion is continuous an hence C can be written as a Reimann integral. That is,

$$C = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} (X_1(\tfrac{i}{n}) - \bar{X}_1)(X_2(\tfrac{i}{n}) - \bar{X}_2).$$

Then

$$\mathbb{V}C = \lim_{n\to\infty} \frac{1}{n^2} \left(\sum_{i=1}^{n} \mathbb{V}(X_1(\tfrac{i}{n}) - \bar{X}_1)\mathbb{V}(X_2(\tfrac{i}{n}) - \bar{X}_2) - 2\sum_{k,l} \mathbb{C} \right).$$

Using properties of Brownian motion, $\mathbb{V}(X_j(\tfrac{i}{n}) - \bar{X}_j) = \tfrac{i}{n} + \mathbb{V}\bar{X}_j - 2\mathbb{C}(X_j(\tfrac{i}{n}), \bar{X}_j)$. Capitalizing again on the fact that integrals of Brownian motion are Riemann,

$$\mathbb{V}\bar{X}_j = \mathbb{V}\int_0^1 X_j(\tau)d\tau = \mathbb{V}\lim_{k\to\infty} \frac{1}{k} \sum_{l=1}^{k} X_j(\tfrac{l}{k})$$