

SimulationVarCheck

Naveed Merchant

October 23, 2018

We're curious to see if our proof has the wrong idea. We examine two functions that are completely different but have the same mean and variance.

We draw from a normal and a normal mixture which have the same variance and same mean, but clearly are not the same function.

```
set.seed(10000)
normdraw <- rnorm(1000,sd = 1)
unifdraw <- runif(1000)
mixturenormdraw <- rnorm(1000,mean = 3, sd = 1)*(unifdraw > 0.5) + rnorm(1000,mean = 3, sd = 1)*(unifdraw < 0.5)

dlength <- 1000

dataset1 <- normdraw
dataset2 <- mixturenormdraw

XT1 <- dataset1[1:(dlength*.3)]
XV1 <- dataset1[-(1:dlength*.3)]
XT2 <- dataset2[1:(dlength*.3)]
XV2 <- dataset2[-(1:dlength*.3)]
source("MarginalLikIntfunctions.R")
ExpectedKernML <- logmarg.kern(XT1,XV1)[[2]] + logmarg.kern(XT2,XV2)[[2]]
ExpectedKernML2 <- logmarg.kern(c(XT1,XT2),c(XV1,XV2))[[2]]
ExpectedKernML - ExpectedKernML2

## [1] 752.3072
```

The log BF overwhelmingly favors the alternative hypothesis.

We do this again for a t-distribution.

```
set.seed(10000)
normdraw <- rnorm(1000,sd = sqrt(3))
tdraw <- rt(1000,df = 3)
dlength <- 1000

dataset1 <- normdraw
dataset2 <- tdraw

XT1 <- dataset1[1:(dlength*.3)]
XV1 <- dataset1[-(1:dlength*.3)]
XT2 <- dataset2[1:(dlength*.3)]
XV2 <- dataset2[-(1:dlength*.3)]
source("MarginalLikIntfunctions.R")
ExpectedKernML <- logmarg.kern(XT1,XV1)[[2]] + logmarg.kern(XT2,XV2)[[2]]
ExpectedKernML2 <- logmarg.kern(c(XT1,XT2),c(XV1,XV2))[[2]]
ExpectedKernML - ExpectedKernML2

## [1] 23.6025
```

It still favors the Alternate hypothesis, not quite as harshly as the mixture case, but it still does.

Things don't run as expected with the cauchy distribution (this gives me an idea how to proceed actually!).

If we check to see if two distributions are common with the cauchy,

We notice that the LBF varies harshly and the integral function suggests the integral is divergent.

We require a modest sample size to run into this error!

I think I will examine higher order moments. I am under the impression that the variance idea is correct, but it is an approximation that is raised to a high power.

My current impression is that we are examining more than just the variance, but rather are examining how far apart higher order moments are. I will verify my intuition. If this is correct, I think this test is powerful but not infallible. I am under the impression that if I match enough moments I may be able to falsify the distribution that it is approximating. This is a result of the Hamburger Moment problem (on an unbounded interval, the moments do not uniquely define the pdf and in addition, we are not looking at all the moments! Only a fair number of them (I think)).

““