

# Using acor to compute error bars and autocorrelation time

Jonathan Goodman  
Courant Institute of Mathematical Sciences, NYU  
goodman@cims.nyu.edu  
May 7, 2009  
Funded by DOE

The program `acor` estimates the errors and autocorrelation time of a time series. It could be applied to real data or to data from Markov chain Monte Carlo. It is intended to be used on fairly long time series. See the README file for more details of the use.

The algorithm combines a direct calculation of the autocovariance function  $C(t)$  with an averaging procedure. If the time series is  $X(t)$ , the code first computes and subtracts out the sample mean:

$$\overline{X} = \frac{1}{n} \sum_{t=1}^n X(t) ,$$

and

$$X(t) \leftarrow X(t) - \overline{X} .$$

Next it computes the first few values of the empirical autocovariance function

$$\hat{C}(t) = \frac{1}{n-t} \sum_{s=1}^n X(t+s)X(s) .$$

(Warning, the actual code differs from this in a way that will be unimportant for large  $n$ . I did this to make it run faster by having better cache performance.) It estimates the diffusion coefficient

$$D = \sum_{-\infty}^{\infty} C(t) ,$$

and the autocorrelation time

$$\tau = \frac{D}{C(0)} ,$$

using those values. See the Monte Carlo lecture notes by Alan Sokal (Google it) for background.

If the estimated  $\tau$  at this point is too large, it computes a reduced time series

$$Y(t) = X(2t) + X(2t+1) .$$

An estimate of  $D$  from the  $Y$  series can be used to estimate  $\tau$  and  $D$  for the  $X$  series. The difference is  $\tau$  for the  $Y$  series should be half as long.

This process is applied recursively until the autocorrelation time is short enough that the direct calculation works. This is  $\tau \leq 2$  in the parameters that the code has “out of the box”.

The code comes with a tester `acorTest` that generates a time series with a simple geometric autocovariance function. It creates a file of the format that `acor` can read. The included *makefile* can create both: type `make acorTest` or `make acor`. When I ran it on my system, I typed `acorTest 4000000 ao` to make a file named `ao` with a time series of length four million. At this time, the parameter `a` in `acorTest` was set to `.9`. Then I ran the program and got output (edited for clarity, the output is all on one line):

```

jg> ./acor ao
sample mean = 5.0011,  standard deviation = 0.00144837,
autocorrelation time = 19.1301, series length = 4000000
```

Warning: the program `acorTest` sets the seed for the random number generator to a fixed value. This makes it so that if you generate two time series with the same `a` you get the same time series. The exact answer is  $\tau = 19$  for `a = .9`.

There also is a C++ procedure called `acor` that receives the time series as an argument.

Please let me know if this routine is helpful, and especially if it is not helpful.