University of Sydney
School of Information Technologies

Simulation of Networks

Basics of Confidence Intervals – a very pragmatic approach

Simulation results are meaningless if we do not have some idea of their accuracy, and furthermore, each simulation run will give a different estimate. How do we know what the correct value is? Generally speaking, we can't know – but we can use the simulation data to make an estimate.

Independent Samples

Let's first assume we have a sequence of *n* data points $\{z_1, z_{2,} \cdots, z_n\}$ which are ***known*** to be independent. We estimate the mean as:

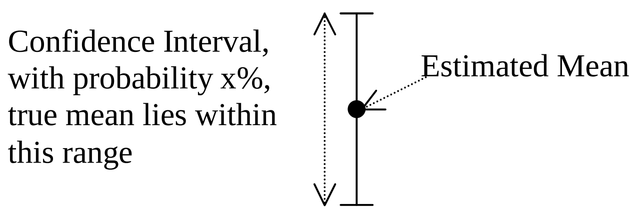
$$\hat{z} = \frac{\sum_{i=1}^{n} z_i}{n}$$

and the variance as:

$$\hat{V} = \frac{\sum_{i=1}^{n} (z_i)^2 - n * (\hat{z})^2}{n(n-1)}$$

We can now construct a confidence interval for the mean. An approximate way to do this is to:

- Assume that there are enough samples so that we can use the Central Limit Theorem (i.e. approx Gaussian distributed)
- Make a hypothesis test – with (say) 95% probability, the true mean lies within some interval
- From a Gaussian distribution, the confidence interval is the estimated mean, ± 1.96 estimated standard deviations (standard deviation is the square root of the variance)
- Then we say that, with probability 95%, the true mean lies within the range $[\hat{z} - 1.96\sqrt{\hat{V}}, \hat{z} + 1.96\sqrt{\hat{V}}$



Effect of Correlation

Of course, simulation data is never independent – there is always correlation between successive samples. It is always a good idea to estimate the correlation – the autocovariance of lag *k* is defined as:

$$R(k) = E\{(z_i - \bar{z})(z_{i-k} - \bar{z})\} \text{ where } \bar{z} \text{ is the true mean}$$

If the autocovariances are low, then an approximation of independence is reasonable, otherwise we need to do better. A couple of simple methods are listed below – there are many more in the literature.

1 Estimates based on estimates of autocovariance (see Ripley)

Use $$\hat{V} = \frac{n}{(n-L)(n-L-1)} \sum_{|s|<L} \left(1 - \frac{|s|}{n}\right)\left(c_s - \hat{z}^2\right)$$

with $$c_s = \sum_{i=1}^{n-s} \frac{z_i z_{i+s}}{n-s}$$

and with L chosen so that the autocorrelation is close to zero for lags greater than L.

2 Method of Batch Means (see Ripley, Pawlikowski)

The idea here is to group the data into batches, to that successive batches are approximately independent. For example, if we have n data points $\{z_1, z_{2,} \cdots, z_n\}$, group them into m batches of size p (assume $n = m * p$):

$$\{z_1, z_{2,} \cdots, z_p\}\{z_{p+1}, z_{2,} \cdots, z_{2p}\}\cdots\{z_{n-p+1}, z_{2,} \cdots, z_n\}$$

Now take the batch means: $x_1 = \frac{\sum_{i=1}^{p} z_i}{p}, x_2 = \frac{\sum_{i=p+1}^{2p} z_i}{p}, \cdots, x_m = \frac{\sum_{i=n-p+1}^{n} z_i}{p}$

Now our data sequence in the set of m observations $\{x_1, x_{2,} \cdots, x_m\}$, assumed to be uncorrelated. This then gives less bias in the estimator, at the expense of more variability in the estimator of precision. The difficult part is choosing the batch size – too large a batch produces confidence intervals which are much larger than is justified by the actual data – too small a batch produces excessive correlation, which produces misleadingly small confidence intervals. However, the method is easy to implement, and is probably the most widely-used method of estimating confidence intervals.

A very similar idea is the method of independent replications – run the simulation m times, with p observations in each run, then treat the outputs of each of the runs in the same way that the batch estimates were treated above.

Transients in the simulation

Everything above assumes that the simulation data is collected once the simulation has achieved steady state, otherwise the tests are meaningless. There are many ways to determine this (see e.g. Pawlikowski) – but a pragmatic approach is to ensure that the simulation runs for sufficiently long so that the transients have negligible effect on the estimates of confidence interval.

However, if data is available from only short runs (i.e. few data points, which is not necessarily the same as a short run times), then you will need to ensure that transient data is discarded.

References

B.D. Ripley, "Uses and abuses of statistical simulation", *Mathematical Programming,* Vol 42, pp 53-68, 1988

K. Pawlikowski, "Steady-state simulation of queueing processes: A survey of problems and solutions", *ACM Computing Surveys,* Vol 22, No 2, pp123-170, June 1990

A.M. Law and W.D. Kelton, *Simulation Modelling and Analysis,* McGraw-Hill, 1991