Hypothesis Testing and Interval Estimation of a Single Lognormal Mean

Let $y_1, ..., y_n$ be a sample of observations from a lognormal distribution with parameters $\mu$ and $\sigma^2$. Let $x_i = \ln(y_i)$, $i = 1, 2, ..., n$. The sample mean and the variance of the $x_i$'s are respectively given by

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \quad \text{and} \quad s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2.$$  

Note that $\exp(\mu + \sigma^2/2)$ is the mean of the lognormal distribution, and testing the mean is equivalent to testing $\eta = \mu + \sigma^2/2$. Consider

$$H_0 : \eta \geq \eta_0 \quad \text{vs.} \quad H_a : \eta < \eta_0. \quad (1)$$

Algorithm 1

For a given logged data set, compute the observed sample mean and variance, namely, $\bar{x}$ and $s^2$, respectively.

For $i = 1$ to $m$

- Generate a standard normal variate $Z$
- Generate a chi-square random variate $V^2$ with degrees of freedom $n - 1$
- Set $T_{2i} = \bar{x} - \frac{Z}{\sqrt{n-1}} \frac{s}{\sqrt{n}} + \frac{1}{2} \frac{V^2}{(n-1)}$
- Set $K_i = 1$ if $T_{2i} > \eta_0$, else $K_i = 0$
  (end i loop)

$\frac{1}{m} \sum_{i=1}^{m} K_i$ is the generalized p-value for testing the hypotheses in (1).

If the computed generalized p-value is less than the nominal level $\alpha$, then the null hypothesis in (1) will be rejected.

The $100(1-\alpha)$th percentile of $T_{21}, ..., T_{2m}$, denoted by $T_{2,1-\alpha}$, is the $100(1-\alpha)\%$ generalized upper confidence limit for $\eta = \mu + \sigma^2/2$. Furthermore, $\exp(T_{2,1-\alpha})$ is the $100(1-\alpha)\%$ generalized upper limit for the lognormal mean. Appropriate quantiles of $T_{2i}$'s can be used to construct two-sided limits for $\eta$ or $\exp(\eta)$.

In order to get consistent results regardless of the initial seed used for random number generation, the number of iteration in the above algorithm (i.e., the value of $m$) should be at least 100,000.