Hypothesis Testing and Interval Estimation for the Difference Between Two Lognormal Means

Let x_{1i}, $i = 1, 2, ..., n_1$, and x_{2i}, $i = 1, 2, ..., n_2$, denote random samples from the lognormal(μ_1, σ_1^2) and lognormal(μ_2, σ_2^2) respectively. Let $y_{1i} = \ln(x_{1i})$, $i = 1, 2, ..., n_1$, and $y_{2i} = \ln(x_{2i})$, $i = 1, 2, ..., n_2$. Define

$$\bar{y}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} y_{ij} \text{ and } s_i^2 = \frac{1}{n_i - 1}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_i)^2,\ i = 1, 2. \tag{1}$$

Let $\eta_i = \mu_i + \sigma_i^2/2$, $i = 1, 2$. The mean of the lognormal(μ_i, σ_i^2) is given by $\exp(\eta_i)$, $i = 1, 2$.

Hypothesis Testing about $\exp(\eta_1) - \exp(\eta_2)$

The problem of testing

$$H_0 : \exp(\eta_1) \le \exp(\eta_2) \text{ vs } H_1 : \exp(\eta_1) > \exp(\eta_2) \tag{2}$$

is equivalent to

$$H_0 : \eta_1 \le \eta_2 \text{ vs } H_1 : \eta_1 > \eta_2. \tag{3}$$

Let

$$T_i = \bar{y}_i - \frac{Z_i}{U_i/\sqrt{n_i - 1}}\frac{s_i}{\sqrt{n_i}} + \frac{1}{2}\frac{s_i^2}{U_i^2/(n_i - 1)},\ i = 1, 2, \tag{4}$$

where $Z_i = \sqrt{n_i}(\bar{Y}_i - \mu_i)/\sigma_i \sim N(0, 1)$ and $U_i^2 = (n_i - 1)S_i^2/\sigma_i^2 \sim \chi^2_{n_i - 1}$, for $i = 1, 2$, and these random variables are also independent. The generalized test variable for testing (3) is given by

$$T = T_1 - T_2 - (\eta_1 - \eta_2). \tag{5}$$

Let

$$T_* = T_1 - T_2, \tag{6}$$

so that $T = T_* - (\eta_1 - \eta_2)$. The generalized p-value for testing the hypotheses in (3) is given by

$$P(T \le 0 | \eta_1 - \eta_2 = 0) = P(T_* \le 0). \tag{7}$$

Confidence Interval for $\exp(\eta_1) - \exp(\eta_2)$

Let $Q = \exp(T_1) - \exp(T_2)$, where T_i is given in (4). Appropriate quantiles of Q can be used to construct confidence limits for $D = \exp(\eta_1) - \exp(\eta_2)$. Let Q_p denote the pth quantile of Q. Then, $(Q_{\alpha/2}, Q_{1-\alpha/2})$ is a $1 - \alpha$ confidence limit for D. One-sided limits for D can be similarly obtained.

The following algorithm can be used to compute the p-value for testing (2) or (3), and to construct a $1 - \alpha$ confidence limit for $D = \exp(\eta_1) - \exp(\eta_2)$.

Algorithm

For given sample sizes n_1 and n_2, and the summary statistics $(\bar{y}_1, s_1^2)$ and $(\bar{y}_2, s_2^2)$, the following Monte Carlo method can be used to compute the generalized p-value and the generalized confidence limit.

For j = 1, m_2
  Generate Z_1 and Z_2 from $N(0,1)$
  Generate $U_1 \sim \chi^2_{n_1-1}$ and $U_2 \sim \chi^2_{n_2-1}$
  Compute $T_j = \exp(T_1) - \exp(T_2)$; T_i given in (4)
(end j loop)
Find the proportion of T_j's which are less than 0. This proportion is the generalized p-value for testing hypotheses in (2). Let T_p denote the pth quantile of $T_1, ..., T_{m_2}$. Then $(T_{\alpha/2}, T_{1-\alpha/2})$ is a $1-\alpha$ confidence limit for $D = \exp(\eta_1) - \exp(\eta_2)$.