Reappraisal on the estimation of parameters of the hybrid lognormal distribution

Shigeru KUMAZAWA, former JAERI (Predecessor of JAEA)

The hybrid log-normal (HLN) distribution is the continuous probability distribution of the variate X defined in the positive region, which bridges the log-normal distribution and the normal distribution. Giving the hybrid function hyb $(x) = x + \ln(x)$, it is denoted as hyb $(\rho X) \sim \mathrm{N}(\mu, \sigma^2)$, where μ is E [hyb (ρX)] and σ^2 is V [hyb (ρX)]. The HLN distribution model has been applied to various data of biology and social statistics as well as radiological statistics that has made the model developed, with the linear regression of $z_i = \alpha + \beta\, x_i + \gamma \ln x_i + \varepsilon_i$ for the ascending ordered data $\{x_i \mid i = 1, n\}$, where the normal cumulative distribution function $\Phi(z_i) \approx$ prob $\{X \le x_i\}$, which is equal to (i – 0.375) / (n + 0.25) by Blom (1958).

The paper presents a simple regression method of the HLN model $h_i = \mu + \sigma z_i + \varepsilon_{hi}$, where $h_i = \mathrm{hyb}(\rho x_i)$, basically using the EXCEL functions LINEST and SOLVER to maximize the R-squared with respect to the parameter $\rho > 0$. With the estimate $\hat{\rho}$, $R^2 = {S_{zh}}^2/S_{zh}S_{zz}$, the SSE $S_{he} = S_{hh}(1 - R^2)$ and the standard error $\mathrm{se}(\hat{\rho}) = (\rho_2 - \rho_1)/2$ of two solutions $\rho_1 < \rho_2$ in the equation of $R^2(\rho) = R^2 - (1 - R^2)/f_e$ are obtained, where $S_{hh}, S_{zz}\ and\ S_{zh}$ are sums of squared deviations and deviation products of h_i and z_i, respectively, and f_e is $n - 3$ degrees of freedom.

The derived results are $\hat{\rho} = A_{zx}/A_{zy}$ and $se(\hat{\rho}) = \sqrt{c_z A_{zz}(S_{zz}R^2(\hat{\rho}) - c_z)}/\left(c_z S_{xx} - {S_{xz}}^2\right)$: S_{xx}, S_{yy}, S_{xy}, S_{xz}, S_{yz} are the sum of squared deviations and deviation products for x_i, $y_i = \ln x_i$ *and* z_i ; $A_{zx} = S_{xy}S_{yz} - S_{xz}S_{yy}$, $A_{zy} = S_{xy}S_{xz} - S_{xx}S_{yz}$, $A_{zz} = S_{xx}S_{yy} - {S_{xy}}^2$ and $c_z = S_{zz}\left(R^2(\hat{\rho})(n - 2) - 1\right)/f_e$.

Analyzing the data on the number of words per sentence in 60 sentences taken from Toynbee's "A Study of History" (Wilks, 1948), the graph of $R^2(\rho)$ is presented to show so that it should change from the log-normal model ($\rho \to 0$) to the normal model ($\rho \to +\infty$) via the HLN model by increasing the value of ρ. The value of $R^2(\rho \to 0)$ is larger than that of $R^2(\rho \to +\infty)$ but the largest is the value of $R^2(\hat{\rho})$. To estimate $\mathrm{se}(\hat{\rho})$ is graphically shown in terms of $R^2(\rho)$, compared to $S_{he}(\rho)$ and $S_{he}(\rho)/S_{hh}(\rho) = 1 - R^2(\rho)$. The results of data analysis are $\hat{\mu} = 0.4084 \pm 0.0096$, $\hat{\sigma} = 0.7972 \pm 0.00985$, $\hat{\rho} = 0.01294 \pm 0.00236$, and $R^2 = 0.9912$. The simple regression of the HLN distribution works well.

A similar approach is applicable to data $\{x_i, y_i | i = 1, n\}$ with the transformation of $u_i = \mathrm{hyb}(\tau x_i)$ and $v_i = \mathrm{hyb}(\nu y_i)$ as the simple regression $v_i = \alpha + \beta u_i + \varepsilon_{2Di}$ to attain the global maximum of $R^2(\tau, \nu)$ using LINEST and SOLVER, where $\tau,\ \nu$ are positive scaling parameters with the inverse unit of x and y, respectively, α, β are regression coefficients and ε_{2Di} is the residual. This approach is to find the best linearity of plotting data on a comprehensive section paper, called "the hybrid-hybrid section paper," that contains not only four conventional section papers (linear-linear, linear-log, log-linear and log-log) but also five new section papers (linear-hybrid, hybrid-log, log-hybrid, hybrid-linear and hybrid-hybrid) that all serve to connect four conventional ones smoothly as their interfaces. Several data of dose-response on the chromosome aberrations per cell irradiated to ionizing radiation do not always show a linearity on any conventional section paper but these show a good linearity on the hybrid-hybrid section paper.