

Modification of the log-normal distribution model based on the small sample theory

*Masami Okada¹

1.Meteorological Research Institute

Introduction

Log-normal distribution model based on the small sample theory is a statistical sophisticated model to calculate the probability of forthcoming repeating events on the renewal process. We prospectively forecasted probabilities for small interplate repeating earthquakes along the Japan Trench (Okada et al., 2012). The number of forecasts in four experiments from 2006 through 2010 was 528 of which 249 cases were filled with qualifying event. Total of probabilities of forecast was 212.9 which was surely less than 249 of observation and the probabilities was rejected by the N-test. The bias of lower probability is confirmed by numerical simulation with random numbers, too. Hence I tried to modify the LN-SST for better forecasting.

Method

Suppose $n+1$ random variables $X_i = \log(T_i)$ and $X_f = \log(T_f)$ obey a normal distribution $N(\mu, \sigma^2)$.

$X_f = \log(T_f)$ represents the interval from the last event to the forthcoming one. Take the variable, as follows;

$$Z = \sqrt{(n-1)/(n+1)} * (X_f - X_{\text{mean}}) / S.$$

It is well known that the Z-variable follows a t-distribution with the $n-1$ degree of freedom. Here X_{mean} and S are mean and standard deviation of n variable of X_i . At forecasting time we can calculate the values of X_{mean} and S , then the expected distribution of X_f is calculated, too. The probability of events in the forecast period is given with the conditional probability from the distribution of X_f .

Possible reason of lower probability are as follows;

- (1) The t-distribution spreads wider than standard normal distribution and has a lower peak of the distribution.
- (2) The expression of the conditional probability is not linear, then the forecast probability may tend to lower.

Modification

It is possible improvements for LN-SST to correct the bias of lower probability and to improve forecast score, as follows;

- (1) Keep a definition of Z mentioned above intact and increase the degree of freedom.

If we raise one degree of freedom for the data of the experiment, the total of forecast probabilities increases to 217.0 from 212.9, and the results improve somewhat, too.

- (2) Increase the probabilities with some quantity depending on calculated probability.

The original probability from 0 through 1 is converted with a formulae, $y = \log(p/(1-p))$ into infinite interval, then suitable value (e.g., $c=0.3$) is added on y . Revised value is given by inverse conversion from y to probability. The total of forecast probabilities becomes 241.0, and the results is considerably improved (figure 1).

Keywords: repeating earthquake, forecast of earthquake, log-normal distribution, small sample theory

