TN2011-2: R-G Error Propagation

George Rebane
27 November 2020

References:
1. Rebane, G.J., *TN0708-1: Predicting the Lifetime of Minimally Known Processes – Gott Extended*
2. Rebane, G.J., *TN1902-1: Predicting Termination of Minimally Known Ongoing Processes*
3. Rebane, G.J., *TN2011-1: R-G Probabilities for Discrete Events*

The references document the development of R-G theory for minimally known processes (MKPs) of age T that may terminate at any (continuous) time in the future. Termination probabilities are calculated for the next time interval ΔT from now, and also for some arbitrary ΔT in the future. The theory has been extended in reference [3] to processes consisting of a sequence of discrete events, of which N have already been observed. For discrete event processes the probability can be calculated for termination within the sequence of the next n events. Reference [2] extends the theory from knowing only the age T of the MKP to knowing only its lifetime T_L, and reference [3] treats the case when we know only that the process contains only N_L events.

In this technical note we consider how imperfect knowledge of T and T_L will impact the P_E, the termination probability in the next ΔT. Suppose we only know T to within a probability distribution from which we compute its expected value μ_T and variance $\sigma_T{}^2$. Then starting with the deterministic P_E from [1], we apply the standard error propagation protocols of a scalar random variable within a defined function.

$$P_E\left(T,\Delta T\right)=\frac{\Delta T}{T+\Delta T}\xrightarrow{\text{for random } T}E\left(P_E\right)\simeq P_E\left(\mu_T,\Delta T\right)+\frac{1}{2}\frac{\partial^2 P_E}{\partial T^2}\sigma_T^2 \tag{1}$$

The approximate expected value is obtained by a series expansion of P_E around μ_T, and truncating after the third term [Papoulis (2002), p150]. Taking the derivatives

$$\frac{\partial P_E}{\partial T}=\frac{-\Delta T}{\left(T+\Delta T\right)^2},\qquad \frac{\partial^2 P_E}{\partial T^2}=\frac{2\Delta T}{\left(T+\Delta T\right)^3}\ , \tag{2}$$

lets us write

$$\begin{aligned} E\left(P_E\right) &\simeq P_E\left(\mu_T,\Delta T\right)+\frac{1}{2}\frac{\partial^2 P_E}{\partial T^2}\bigg|_{\mu_T}\sigma_T^2 \\ &=\frac{\Delta T}{\mu_T+\Delta T}+\frac{\Delta T\sigma_T^2}{\left(\mu_T+\Delta T\right)^3}=\Delta T\left[\frac{\left(\mu_T+\Delta T\right)^2+\sigma_T^2}{\left(\mu_T+\Delta T\right)^3}\right]\end{aligned} \tag{3}$$

The variance of P_E is computed as

$$\sigma_{P_E}^2 = \left(\left. \frac{dP_E}{dT} \right|_{\mu_T} \right)^2 \sigma_T^2 = \left[\frac{-\Delta T}{\left(\mu_h + \Delta T\right)^2} \right]^2 \sigma_T^2 \tag{4}$$

The scaled version of (1) is written with the scaling ratio $h = \Delta T / T$ as

$$P_E(h) = \frac{h}{1+h} \tag{5}$$

Once more we take the derivatives

$$\frac{\partial P_E}{\partial h} = \frac{1}{(1+h)^2}, \quad \frac{\partial^2 P_E}{\partial h^2} = \frac{-2}{(1+h)^3}\ , \tag{6}$$

and assume we know h to within a probability distribution with mean $\mu_h = \Delta T / \mu_T$ and standard deviation σ_h^2. Here the analogue of (1) gives

$$\begin{aligned} &P_E(h) = \frac{h}{1+h} \xrightarrow{\text{for random } h} E(P_E) \simeq \frac{\mu_h}{1+\mu_h} + \frac{1}{2} \left. \frac{\partial^2 P_E}{\partial h^2} \right|_{\mu_h} \sigma_h^2 = \frac{\mu_h}{1+\mu_h} - \frac{1}{(1+\mu_h)^3} \sigma_h^2, \\ &\text{or } E(P_E) \simeq \frac{\mu_h (1+\mu_h)^2 - (\mu_h h_{\sigma_T})^2}{(1+\mu_h)^3}, \text{ when } \sigma_h^2 \text{ is related to } \sigma_T^2 \text{ through } h_{\sigma_T} \text{ (see below)}. \end{aligned} \tag{7}$$

Here we have to be careful how we derive σ_h^2. Using the standard propagation of error approach for variance, we proceed as follows and note that in the derivation we have identified a new variance related scaling factor that supports a compact formulation of the scaled variance.

$$\begin{aligned} h = \frac{\Delta T}{T} \rightarrow \sigma_h^2 &= \left(\left. \frac{dh}{dT} \right|_{\mu_T} \right)^2 \sigma_T^2 = \left(\left. -\frac{\Delta T}{T^2} \right|_{\mu_T} \right)^2 \sigma_T^2 \cdots \\ &= \left(-\frac{\Delta T}{\mu_T^2} \right)^2 \sigma_T^2 = \left(-h \frac{\sigma_T}{\mu_T} \right)^2 = h^2 \left(-\frac{\sigma_T}{\mu_T} \right)^2, \\ \sigma_h^2 = h^2 h_{\sigma_T}^2 &\xrightarrow{T-\text{dispersion scaling constant}} h_{\sigma_T} = -\frac{\sigma_T}{\mu_T} \end{aligned} \tag{8}$$

The variance in P_E is then expressed as

$$\sigma^2_{P_E} = \left(\frac{dP_E}{dh}\bigg|_{\mu_h} \right)^2 \sigma^2_h = \left[\frac{1}{\left(1+\mu_h\right)^2} \right]^2 \sigma^2_h \tag{9}$$

$P_E(T, \Delta T_1, \Delta T_2) = P_E(h_1, h_2)$

Next, we examine $E(P_E)$ for the more general case when the interval of interest begins at ΔT_1 from now and ends at $\Delta T_2 > \Delta T_1$. From [2] equation (2) we have the deterministic version with scaling ratios $h_i = \Delta T_i / T$ for i = 1,2.

$$P_E\left(T, \Delta T_1, \Delta T_2\right) = \frac{T\left(\Delta T_2 - \Delta T_1\right)}{\left(T+\Delta T_1\right)\left(T+\Delta T_2\right)} = \frac{h_2 - h_1}{\left(1+h_1\right)\left(1+h_2\right)} \ , \tag{10}$$

and the relevant derivatives wrt T are

$$\frac{\partial P_E}{\partial T} = \frac{\Delta T_1}{\left(T+\Delta T_1\right)^2} - \frac{\Delta T_2}{\left(T+\Delta T_2\right)^2}, \quad \frac{\partial^2 P_E}{\partial T^2} = 2\left[\frac{\Delta T_2}{\left(T+\Delta T_2\right)^3} - \frac{\Delta T_1}{\left(T+\Delta T_1\right)^3} \right] \tag{11}$$

From (1) we can again write

$$\begin{aligned} P_E\left(T, \Delta T_1, \Delta T_2\right) &\xrightarrow{\text{for random } T} E\left(P_E\right) \simeq P_E\left(T, \Delta T_1, \Delta T_2\right)\Big|_{\mu_T} + \frac{1}{2}\frac{\partial^2 P_E}{\partial T^2}\bigg|_{\mu_T} \sigma^2_T \ \ldots \\ &= \frac{\mu_T\left(\Delta T_2 - \Delta T_1\right)}{\left(\mu_T+\Delta T_1\right)\left(\mu_T+\Delta T_2\right)} + \left[\frac{\Delta T_2}{\left(\mu_T+\Delta T_2\right)^3} - \frac{\Delta T_1}{\left(\mu_T+\Delta T_1\right)^3} \right] \sigma^2_T \end{aligned} \tag{12}$$

The calculation of the variance is straightforward for the unscaled version of P_E. As in (4), we express the variance in the usual way for a function of a single random variable.

$$\boxed{\sigma^2_{P_E} = \left(\frac{dP_E}{dT}\bigg|_{\mu_T} \right)^2 \sigma^2_T = \left[\frac{\Delta T_1}{\left(\mu_T+\Delta T_1\right)^2} - \frac{\Delta T_2}{\left(\mu_T+\Delta T_2\right)^2} \right]^2 \sigma^2_T} \tag{13}$$

For the scaled version of the general case, we need to first develop $\sigma^2_{P_E(\underline{h})}$. This requires some extra consideration because we are now dealing with a function of two random variables. Moreover, since the random scaling ratios $h_i = \Delta T_i / T$ for i = 1,2 are not independent, but instead are perfectly correlated since they both derive from the same T distribution, we need appeal to $Cov\left(\underline{h}\right)$, the covariance matrix of the random column vector $\underline{h} = \left[h_1, h_2\right]^T$. The needed partial derivatives are

$$\frac{\partial P_E(\underline{h})}{\partial h_1} = -\frac{1}{(1+h_1)^2}, \quad \frac{\partial P_E(\underline{h})}{\partial h_2} = \frac{1}{(1+h_2)^2}$$
$$\frac{\partial^2 P_E(\underline{h})}{\partial h_1^2} = \frac{2}{(1+h_1)^3}, \quad \frac{\partial^2 P_E(\underline{h})}{\partial h_2^2} = -\frac{2}{(1+h_2)^3} \tag{14}$$

Recalling from (8) for the scalar case, we can write the variances for the h_i as

$$\sigma_{h_i}^2 = h_i^2 h_{\sigma_T}^2, \quad i = 1, 2 \ , \tag{15}$$

which gives the covariance

$$Cov(\underline{h}) = \begin{bmatrix} \sigma_{h_1}^2 & \rho_{1,2}\sigma_{h_1}\sigma_{h_2} \\ \rho_{1,2}\sigma_{h_1}\sigma_{h_2} & \sigma_{h_2}^2 \end{bmatrix} = \begin{bmatrix} h_1^2 h_{\sigma_T}^2 & 1\bullet h_1 h_2 h_{\sigma_T}^2 \\ 1\bullet h_1 h_2 h_{\sigma_T}^2 & h_2^2 h_{\sigma_T}^2 \end{bmatrix}$$
$$= \begin{bmatrix} h_1^2 & h_1 h_2 \\ h_1 h_2 & h_2^2 \end{bmatrix} h_{\sigma_T}^2 = \begin{bmatrix} \left(\frac{\Delta T_1}{\mu_T}\right)^2 & \frac{\Delta T_1 \Delta T_2}{\mu_T^2} \\ \frac{\Delta T_1 \Delta T_2}{\mu_T^2} & \left(\frac{\Delta T_2}{\mu_T}\right)^2 \end{bmatrix} h_{\sigma_T}^2 = \begin{bmatrix} \Delta T_1^2 & \Delta T_1 \Delta T_2 \\ \Delta T_1 \Delta T_2 & \Delta T_2^2 \end{bmatrix} \frac{\sigma_T^2}{\mu_T^4} \tag{16}$$

Using (14), we may now express $\sigma_{P_E(\underline{h})}^2$ in its vector-matrix form as

$$\sigma_{P_E(\underline{h})}^2 = \left(\frac{\partial P_E(\underline{h})}{\partial \underline{h}}\right)^T Cov(\underline{h}) \left(\frac{\partial P_E(\underline{h})}{\partial \underline{h}}\right) = \begin{pmatrix} -(1+h_1)^{-2} \\ (1+h_2)^{-2} \end{pmatrix}^T \begin{bmatrix} h_1^2 & h_1 h_2 \\ h_1 h_2 & h_2^2 \end{bmatrix} \begin{pmatrix} -(1+h_1)^{-2} \\ (1+h_2)^{-2} \end{pmatrix} h_{\sigma_T}^2 \tag{17}$$

Note that the quadratic form in the h_i evaluates to a scalar scaling constant that multiplies $h_{\sigma_T}^2$, the normalized error in T as shown in (8). Were we to desire the scaled version of $\sigma_{P_E(\underline{h})}^2$ independently of the 'T version' in (17), then from (16) we would compute the version that accepts the scaling vector $\underline{h}$ from its own binomial distribution as

$$\sigma^2_{P_E(\underline{h})} = \left(\frac{\partial P_E(\underline{h})}{\partial \underline{h}} \bigg|_{\underline{\mu}_h} \right)^T Cov(\underline{h}) \left(\frac{\partial P_E(\underline{h})}{\partial \underline{h}} \bigg|_{\underline{\mu}_h} \right)$$
$$= \begin{pmatrix} -(1+\mu_{h_1})^{-2} \\ (1+\mu_{h_2})^{-2} \end{pmatrix}^T \begin{bmatrix} \sigma^2_{h_1} & \rho_{1,2}\sigma_{h_1}\sigma_{h_2} \\ \rho_{1,2}\sigma_{h_1}\sigma_{h_2} & \sigma^2_{h_2} \end{bmatrix} \begin{pmatrix} -(1+\mu_{h_1})^{-2} \\ (1+\mu_{h_2})^{-2} \end{pmatrix} \tag{18}$$

This error propagation derivation concludes with the graphic below that illustrates the correctness of $p(h)$, the theoretical pdf of h shown in cyan, calculated from the $P(h)$ in (5), the related cdf of h shown in red. The normalized histogram is a sample run of the Monte Carlo simulation of the pdf - RGcdfsimulation03.m - based on the Copernican arguments of its derivation in [2].

$$p(h) = \frac{dP(h)}{dh} = \frac{d}{dh}\left[\frac{h}{1+h}\right] = \frac{1}{(1+h)^2} \tag{19}$$