

Seminar Talk, March 31, 2000
Department of Statistics
Case Western Reserve University

**Estimation with Recurrent Data Accruing from a Sum-Quota
Sampling Scheme¹**

Edsel A. Peña

Bowling Green State University
(currently Visiting the University of Michigan)

Consider a study where successive occurrences of a recurrent event (e.g., failures of a mechanical/electronic system or admission to a hospital of a patient with a chronic disease) is to be monitored for an experimental unit in a reliability or engineering study, a clinical trial, or in other settings. Denote by T_1, T_2, T_3, \dots the successive inter-occurrence times, so T_1 is the time to occurrence of the first event, T_2 is the time to occurrence of the second event measured from T_1 , and so on. Assume that the T_i 's are independent and identically distributed nonnegative variables with common distribution function F , which may be parametrically or nonparametrically specified. In such studies there is usually a random termination time, τ , with distribution function G . Thus, monitoring of the unit only occurs in the interval $[0, \tau]$. Let $S_0 = 0$ and $S_k = T_1 + T_2 + \dots + T_k$ for $k = 1, 2, \dots$. Define the variable K to be the largest index such that $S_K \leq \tau$, so K is the number of events observed over $[0, \tau]$. Then, for this unit, the observable random vector is $D = (K, T_1, T_2, \dots, T_K, \tau - S_K)$. Note that K has distribution which depends on F (and G), so it is informative with respect to making inference about F . Furthermore, the elements of D are dependent, even when viewed conditionally on $K = k$, and $\tau - S_K$ is a right-censored observation. In this talk I will consider the problem of estimating parameters of F , or estimating F itself, based on n independent replicates D_1, D_2, \dots, D_n of D associated with n experimental units. Finite and asymptotic properties of the estimators will be discussed. Dangers and pitfalls of ignoring the informativeness of K , the information contained in the right-censored observation, or when portions of the vectors D_j are not included in the statistical analysis, will be indicated. Existing open problems and model generalizations will be described.

¹This work is joint with R. Strawderman and M. Hollander.