

Chapter 1

Recurrence and Regeneration

In its most basic form the concept of recurrence was first formulated by H. Poincaré in his theory of dynamical systems, specifically in the following statement quoted by Chandrasekhar [24]:

In a system of material particles under the influence of forces which depend only on the spatial coordinates, a given initial state must, in general, recur, not exactly, but to any desired degree of accuracy, infinitely often, provided the system always remains in the finite part of the phase space.

This statement caused controversies between classical and statistical physicists, but was clarified by M.V. Smoluchowski and P. and T. Ehrenfest in probabilistic terms; for historical remarks on this subject see Kac [47, 48].

An equivalent notion is that of regeneration, the first clear formulation of which was given by Palm [83]. Roughly speaking, the idea is that a random process sometimes possesses a set of points (called regeneration points) such that at each such point the process starts from scratch, that is, it evolves anew independently of its past. For a Markov process every point is a regeneration point.

A closely related concept is that of recurrent events, for which W. Feller developed a systematic theory in 1949-1950. A recurrent event is an attribute of a sequence of repeated trials, indicating a pattern such that after each occurrence of this pattern the trials start from scratch. An elementary example is equalization of the cumulated number of heads and tails in coin tossing and success runs in Bernoulli trials, but perhaps the most important application of the theory is to time-homogeneous Markov chains with a countable state space (Example 1.1 below), where the successive returns

¹This chapter is an updated version of material included in N.U. Prabhu's Technical Report: 'Random Walks, Renewal Processes and Regenerative Phenomena' (June 1994), Uppsala University, Department of Mathematics.

of the chain to a fixed state constitute a recurrent event (this idea had in fact been used earlier by W. Doeblin in 1938). The preceding definition of recurrent events is unsatisfactory because it does not distinguish the pattern (or rather, the phenomenon) from the events that constitute it, and moreover, it does not indicate a possible extension of the concept to continuous time. K.L. Chung's 1967 formulation of repetitive patterns suffers from the same drawbacks. As a consequence, in applied literature one very often finds loose descriptions of recurrent phenomena in discrete as well as continuous time. A continuous time theory was proposed by M.S. Bartlett in 1953, but it deals with a particular case which could be treated by other techniques.

Replacing Feller's framework of an infinite sequence of trials with a countable set of outcomes by a sequence $\{X_n, n \geq 1\}$ of random variables taking values in $S = \{x_1, x_2, \dots\}$ we may describe a recurrent phenomenon as follows. Let $\sigma\{X_1, X_2, \dots, X_n\}$ be the σ -field induced by X_1, X_2, \dots, X_n ($n \geq 1$). Denote by Z_n a random variable taking values in $\{0, 1\}$ and such $Z_0 = 1$ a.s. The sequence $\{Z_n, n \geq 0\}$ is a recurrent phenomenon if it satisfies the following conditions:

(a) For each $n > 1$,

$$\{Z_n = 1\} \in \sigma\{X_1, X_2, \dots, X_n\}. \quad (1.1)$$

(b) For $\alpha \in \{0, 1\}$ and $m \geq 1, n \geq 1$,

$$\begin{aligned} P\{X_1 = x_1, X_2 = x_2, \dots, X_{m+n} = x_{m+n}, Z_{m+n} = \alpha | Z_m = 1\} \\ = P\{X_1 = x_1, X_2 = x_2, \dots, X_m = x_m | Z_m = 1\} \cdot \\ P\{X_1 = x_{m+1}, X_2 = x_{m+2}, \dots, X_n = x_{m+n}, Z_n = \alpha\}. \end{aligned} \quad (1.2)$$

However, this definition is really that of a recurrent process $\{X_n\}$, namely, a process which has imbedded in it a recurrent phenomenon.

While it is true that recurrent phenomena rarely occur by themselves, it is possible to define a recurrent phenomenon $\{Z_n, n \geq 0\}$ rather simply as follows. Let $\{Z_n, n \geq 0\}$ be a sequence of random variables on a probability space (Ω, \mathcal{F}, P) , taking values in $\{0, 1\}$ and such that $Z_0 = 1$ a.s. and

$$\begin{aligned} P\{Z_{n_1} = Z_{n_2} = \dots = Z_{n_r} = 1\} \\ = P\{Z_{n_1} = 1\}P\{Z_{n_2-n_1} = 1\} \cdots P\{Z_{n_r-n_{r-1}} = 1\} \end{aligned} \quad (1.3)$$

for every set of integers n_i such that

$$0 < n_1 < n_2 < \dots < n_r \quad (r \geq 1).$$

Example 1.1. Suppose $\{X_n, n \geq 0\}$ is a time-homogeneous Markov chain on the state space $S = \{0, 1, 2, \dots\}$. For $a \in S$, assume $X_0 = a$ a.s., and define the random variables Z_n as follows:

$$Z_n = \begin{cases} 1 & \text{if } X_n = a \\ 0 & \text{otherwise.} \end{cases}$$

It is seen that the property (1.3) is satisfied and hence $\{Z_n, n \geq 0\}$ is a recurrent phenomenon, with

$$u_n = P\{Z_n = 1\} = P_{aa}^{(n)},$$

where $P_{ij}^{(n)}$ ($i, j \in S, n \geq 0$) are the n -step transition probabilities of the chain. Conversely, the sequence $\{u_n\}$ associated with a recurrent phenomenon can arise only in this way (K.L. Chung). Actually we can do better, as indicated in the following.

Example 1.2. Let $\{X_k, k \geq 1\}$ be a sequence of independent and identically distributed random variables on the state space $\{1, 2, \dots\}$. Denote $S_0 \equiv 0, S_k = X_1 + X_2 + \dots + X_k$ ($k \geq 1$) and

$$R = \{n \geq 0 : S_k = n \text{ for some } k \geq 0\}.$$

That is, R is the range of the renewal process $\{S_k, k \geq 0\}$. For $n \geq 0$ let

$$Z_n = \begin{cases} 1 & \text{if } n \in R \\ 0 & \text{otherwise.} \end{cases}$$

Then it can be shown that $\{Z_n, n \geq 0\}$ is a recurrent phenomenon. Conversely, it turns out that every recurrent phenomenon is the indicator function of the range of a renewal process induced by a distribution on the positive integers. Then X_k are identified as the recurrence times (Poincaré cycles).

The definition (1.3) suggests the appropriate extension of the concept of recurrent phenomena to continuous time, but we might now prefer the term regenerative to recurrent as a matter of fashion. Thus the family of random variables $\{Z(t), t \geq 0\}$ is a regenerative phenomenon if for each t , $Z(t)$ takes values in $\{0, 1\}$ and

$$\begin{aligned} P\{Z(t_1) = Z(t_2) = \dots = Z(t_r) = 1\} \\ = P\{Z(t_1) = 1\}P\{Z(t_2 - t_1) = 1\} \dots P\{Z(t_r - t_{r-1}) = 1\} \end{aligned} \quad (1.4)$$

whenever $0 < t_1 < t_2 < \dots < t_r$ ($r \geq 1$). The definitions (1.3) and (1.4) were proposed by J.F.C. Kingman in 1963 and the theory of continuous-time regenerative phenomena was developed by him in a series of papers

published during 1964-1970. He also gave a streamlined account of the theory [55].

Example 1.3. Let $\{B(t), t \geq 0\}$ be a Brownian motion on $(-\infty, \infty)$ and define

$$Z(t) = \begin{cases} 1 & \text{if } B(t) = 0 \\ 0 & \text{if } B(t) \neq 0 \end{cases}$$

Thus $Z(t)$ is the indicator function of the state 0. Clearly, the property (1.4) is satisfied and $\{Z(t), t \geq 0\}$ is a regenerative phenomenon. However, we have

$$p(t) = P\{Z(t) = 1\} = 0.$$

Kingman considers regenerative phenomena satisfying the condition

$$p(t) \rightarrow 1 \text{ as } t \rightarrow 0^+. \quad (1.5)$$

If this condition is satisfied, then $p(t) > 0$ for all $t \geq 0$, and we may take $p(0) = 1$. Such p -functions and the corresponding phenomena are called standard.

Example 1.4. Let $\{X(t), t \geq 0\}$ be a time-homogeneous Markov process on the state space $S = \{0, 1, 2, \dots\}$. Let $a \in S$, $X_0 = a$ a.s. and for $t \geq 0$ define

$$Z(t) = \begin{cases} 1 & \text{if } X(t) = a \\ 0 & \text{otherwise.} \end{cases}$$

Then, as in Example 1.1, it is seen that $\{Z(t), t \geq 0\}$ is a regenerative phenomenon, with

$$p(t) = P\{Z(t) = 1\} = P_{aa}(t),$$

where $P_{ij}(t)$ ($i, j \in S, t \geq 0$) is the transition probability function of the process. The usual assumptions on these functions lead to the fact that the phenomenon is standard. Unlike the situation in discrete time, the function p of a standard regenerative phenomenon need not arise only in this manner. Kingman [55] investigates the so-called Markov characterization problem, namely, the problems of establishing a criterion for a standard p -function to be a diagonal element of some Markov transition matrix.

Example 1.2 shows the connection between recurrent phenomena and renewal processes. In continuous time a similar connection exists between regenerative phenomena and Lévy processes with non-decreasing sample functions (subordinators); the survey paper by Fristedt [39] sheds more light on this connection. Maisonneuve [67] has studied regenerative phenomena from a slightly different point of view, and a theory of stochastic regenerative processes developed earlier by Smith [107, 108] has considerably influenced research workers in this area.

For the theory of recurrent phenomena, despite the drawbacks in its formulation indicated above, the best sources are still Feller's (1949) paper [34] and his book ([35], Chapter XIII).

We shall use the term regenerative phenomena to cover the discrete time as well as the continuous-time cases. For a more recent treatment see Prabhu ([92], Chapter 9).

Semiregenerative phenomena are regenerative phenomena that have, in addition, a Markov component. Such phenomena are important in the study of Markov-additive processes, and arise in several important applications. In the following chapters we study semiregenerative phenomena, including, in particular, their connections with Markov-additive processes.