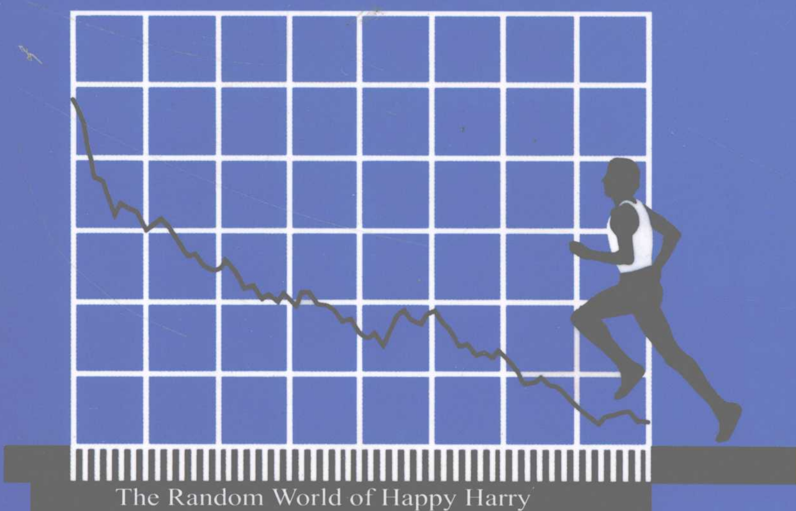
Sidney I. Resnick

Adventures in Stochastic Processes

随机过程探究



Birkhäuser

世界图书出版公司
www.wpcbj.com.cn

Sidney Resnick

Adventures in Stochastic Processes

with Illustrations

Birkhäuser
Boston · Basel · Berlin

图书在版编目（CIP）数据

随机过程探究 = Adventures in Stochastic Processes：英文/（美）雷斯尼克（Resnick，S. I.）著．—影印本．—北京：世界图书出版公司北京公司，2011. 1
ISBN 978-7-5100-2972-1

Ⅰ. ①随… Ⅱ. ①雷… Ⅲ. ①随机过程—英文 Ⅳ. ①O211. 6

中国版本图书馆 CIP 数据核字（2010）第 212139 号

书　　名：Adventures in Stochastic Processes
作　　者：Sidney I. Resnick

中 译 名：随机过程探究
责任编辑：高蓉　刘慧

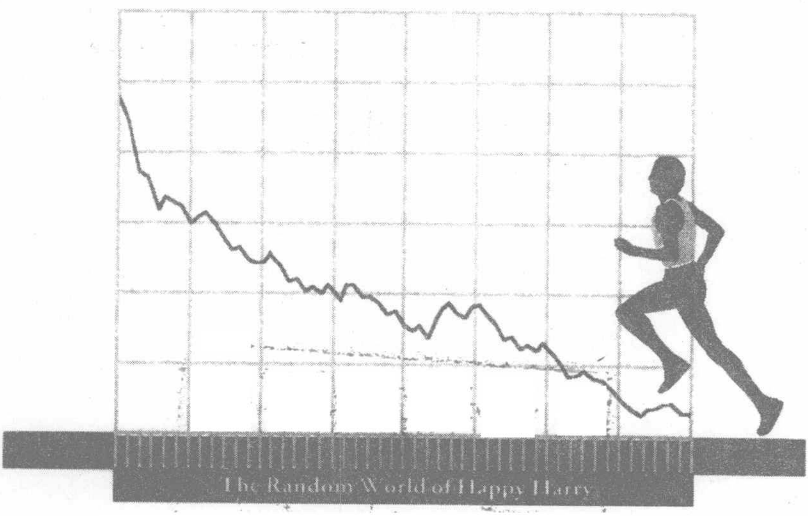
出 版 者：世界图书出版公司北京公司
印 刷 者：三河国英印务有限公司
发　　行：世界图书出版公司北京公司（北京朝内大街 137 号　100010）
联系电话：010-64021602，010-64015659
电子信箱：kjb@ wpcbj. com. cn

开　　本：24 开
印　　张：27
版　　次：2011 年 01 月
版权登记：图字：01-2010-1414

书　　号：978-7-5100-2972-1/O・846　　定　　价：69. 00 元

Adventures in Stochastic Processes

WELCOME TO THE RANDOM WORLD OF HAPPY HARRY—famed restaurateur, happy hour host, community figure, former semi-pro basketball player, occasional software engineer, talent agent, budding television star, world traveller, nemesis of the street gang called the Mutant Creepazoids, theatre patron, supporter of precise and elegant use of the English language, supporter of the war on drugs, unsung hero of the fairy tale *Sleeping Beauty*, and the target of a vendetta by the local chapter of the Young Republicans. Harry and his restaurant are well known around his Optima Street neighborhood both to the lovers of fine food and the public health service. Obviously this is a man of many talents and experiences who deserves to have a book written about his life.



Preface

While this is a book about Harry and his adventurous life, it is primarily a serious text about stochastic processes. It features the basic stochastic processes that are necessary ingredients for building models of a wide variety of phenomena exhibiting time varying randomness.

The book is intended as a first year graduate text for courses usually called *Stochastic Processes* (perhaps amended by the words "Applied" or "Introduction to ... ") or *Applied Probability*, or sometimes *Stochastic Modelling*. It is meant to be very accessible to beginners, and at the same time, to serve those who come to the course with strong backgrounds. This flexiblity also permits the instructor to push the sophistication level up or down. For the novice, discussions and motivation are given carefully and in great detail. In some sections beginners are advised to skip certain developments, while in others, they can read the words and skip the symbols in order to get the content without more technical detail than they are ready to assimilate. In fact, with the numerous readings and variety of problems, it is easy to carve a path so that the book challenges more advanced students, but remains instructive and manageable for beginners. Some sections are starred and come with a warning that they contain material which is more mathematically demanding. Several discussions have been modularized to facilitate flexible adaptation to the needs of students with differing backgrounds. The text makes crystal clear distinctions between the following: proofs, partial proofs, motivations, plausibility arguments and good old fashioned hand-waving.

Where did Harry, Zeke and the rest of the gang come from? Courses in Stochastic Processes tend to contain overstuffed curricula. It is, therefore, useful to have quick illustrations of how the theory leads to techniques for calculating numbers. With the Harry vignettes, the student can get in and out of numerical illustrations quickly. Of course, the vignettes are not meant to replace often stimulating but time consuming real applications. A variety of examples with applied appeal are sprinkled throughout the exposition and exercises. Our students are quite fond of Harry and enjoy psychoanalyzing him, debating whether he is "a polyester sort of guy" or the "jeans and running shoes type." They seem to have no trouble discerning the didactic intent of the Harry stories and accept the need for some easy numerical problems before graduating to more serious ones. Student culture has become so ubiquitous that foreign students who are not native English speakers can quickly get into the swing. I think Harry is a useful and entertaining guy but if you find that you loathe him, he is easy to avoid in the text.

Where did they come?
I can't say.
But I bet they have come a long long way.[1]

To the instructor: The discipline imposed during the writing was that the first six chapters should not use advanced notions of conditioning which involve relatively sophisticated ideas of integration. Only the elementary definition is used: $P(A|B) = P(A \cap B)/P(B)$. Instead of conditioning arguments we find independence where we need it and apply some form of the product rule: $P(A \cap B) = P(A)P(B)$ if A and B are independent. This maintains rigor and keeps the sophistication level down.

No knowledge of measure theory is assumed but it is assumed that the student has already digested a good graduate level pre–measure theoretic probability course. A bit of measure theory is discussed here and there in starred portions of the text. In most cases it is simple and intuitive but if it scares you, skip it and you will not be disadvantaged as you journey through the book. If, however, you know some measure theory, you will understand things in more depth. There is a sprinkling of references throughout the book to *Fubini's theorem, the monotone convergence theorem* and *the dominated convergence theorem*. These are used to justify the interchange of operations such as summation and integration. A relatively unsophisticated student would not and should not worry about justifications for these interchanges of operations; these three theorems should merely remind such students that somebody knows how to check the correctness of these interchanges.

Analysts who build models are supposed to know how to build models. So for each class of process studied, a construction of that process is included. Independent, identically distributed sequences are usually assumed as primitives in the constructions. Once a concrete version of the process is at hand, many properties are fairly transparent. Another benefit is that if you know how to construct a stochastic process, you know how to *simulate* the process. While no specific discussion of simulation is included here, I have tried to avoid pretending the computer does not exist. For instance, in the Markov chain chapters, formulas are frequently put in matrix form to make them suitable for solution by machine rather than by hand. Packages such as Minitab, Mathematica, Gauss, Matlab, etc., have been used successfully as valuable aids in the solution of problems but local availability of computing resources and the rapidly changing world of hardware and software make specific suggestions unwise. Ask your local guru

[1] Dr. Seuss, *One Fish, Two Fish, Red Fish, Blue Fish*

for suggestions. You need to manipulate some matrices, and find roots of polynomials; but nothing too fancy. If you have access to a package that does symbolic calculations, so much the better. A companion disk to this book is being prepared by Douglas McBeth which will allow easy solutions to many numerical problems.

There is much more material here than can be covered in one semester. Some selection according to the needs of the students is required. Here is the core of the material: Chapter 1: 1.1–1.6. Skip the proof of the continuity theorem in 1.5 if necessary but mention Wald's identity. Some instructors may prefer to skip Chapter 1 and return later to these topics, as needed. If you are tempted by this strategy, keep in mind that Chapter 1 discusses the interesting and basic random walk and branching processes and that facility with transforms is worthwhile. Chapter 2: 2.1–2.12,2.12.1. In Section 2.13, a skilled lecturer is advised to skip most of the proof of Theorem 2.13.2, explain coupling in 15 minutes, and let it go at that. This is one place where hand-waving really conveys something. The material from Section 2.13.1 should be left to the curious. If time permits, try to cover Sections 2.14 and 2.15 but you will have to move at a brisk pace. Chapter 3: In renewal theory stick to basics. After all the discrete state space theory in Chapters 1 and 2, the switch to the continuous state space world leaves many students uneasy. The core is Sections 3.1–3.5, 3.6, 3.7, and 3.7.1. Sections 3.8 and 3.12.3 are accessible if there is time but 3.9–3.12.2 are only for supplemental reading by advanced students. Chapter 4: The jewels are in Sections 4.1 to 4.7. You can skip 4.3.1. If you have a group that can cope with a bit more sophistication, try 4.7.1, 4.8 and 4.9. Once you come to know and love the Laplace functional, the rest is incredibly easy and short. Chapter 5: The basics are 5.1–5.8. If you are pressed for time, skip possibly 5.6 and 5.8; beginners may avoid 5.2.1, 5.3.1 and 5.5.1. Section 5.7.1 is on queueing networks and is a significant application of standard techniques, so try to reserve some time for it. Section 5.9 is nice if there is time. Despite its beauty, leave 5.11 for supplemental reading by advanced students. Chapter 6: Stick to some easy path properties, strong independent increments, reflection, and some explicit calculations. I recommend 6.1, 6.2, 6.4, 6.5, 6.6, 6.7, and 6.8. For beginners, a quick survey of 6.11–6.13 may be adequate. If there is time and strong interest in queueing, try 6.9. If there is strong interest in statistics, try 6.10. I like Chapter 7, but it is unlikely it can be covered in a first course. Parts of it require advanced material.

In the course of teaching, I have collected problems which have been inserted into the examples and problem sections; there should be a good supply. These illustrate a variety of applied contexts where the skills mastered in the chapter can be used. Queueing theory is a frequent context for many exercises. Many problems emphasize calculating numbers which seems to be a skill most students need these days, especially considering the wide clientele who enroll for courses in stochastic processes. There is a big payoff for the student who will spend serious time working out the problems. Failure to do do will relegate the novice reader to the status of voyeur.

Some acknowledgements and thank you's: The staff at Birkhäuser has been very supportive, efficient and colleagal, and the working relationship could not have been better. Minna Resnick designed a stunning cover and logo. Cornell's Kathy King probably does not realize how much cumulative help she intermittently provided in turning scribbled lecture notes into something I could feed the TEX machine. Richard Davis (Colorado State University), Gennady Sammorodnitsky (Cornell) and Richard Serfozo (Georgia Institute of Technology) used the manuscript in classroom settings and provided extensive lists of corrections and perceptive suggestions. A mature student perspective was provided by David Lando (Cornell) who read almost the whole manuscript and made an uncountable number of amazingly wise suggestions about organization and presentation, as well as finding his quota of mistakes. Douglas McBeth made useful comments about appropriate levels of presentation and numerical issues. David Lando and Eleftherios Iakavou helped convince me that Harry could become friends with students whose mother tongue was different from English. Joan Lieberman convinced me even a lawyer could appreciate Harry. Minna, Rachel and Nathan Resnick provided a warm, loving family life and generously shared the home computer with me. They were also very consoling as I coped with two hard disk crashes and a monitor melt-down.

While writing a previous book in 1985, I wore out two mechanical pencils. The writing of this book took place on four different computers. Financial support for modernizing the computer equipment came from the National Science Foundation, Cornell's Mathematical Sciences Institute and Cornell's School of Operations Research and Industrial Engineering. Having new equipment postponed the arrival of bifocals and made that marvellous tool called TEX almost fun to use.

Table of Contents

*This section contains advanced material which may be skipped on first reading by beginning readers.

*This section contains advanced material which may be skipped on first reading by beginning readers.

CHAPTER 5. CONTINUOUS TIME MARKOV CHAINS

CHAPTER 6. BROWNIAN MOTION

*This section contains advanced material which may be skipped on first reading by beginning readers.

CHAPTER 7. THE GENERAL RANDOM WALK*

*This section contains advanced material which may be skipped on first reading by beginning readers.

CHAPTER 1

Preliminaries

Discrete Index Sets and/or Discrete State Spaces

THIS CHAPTER eases us into the subject with a review of some useful techniques for handling non-negative integer valued random variables and their distributions. These techniques are applied to some significant examples, namely, the simple random walk and the simple branching process. Towards the end of the chapter stopping times are introduced and applied to obtain Wald's identity and some facts about the random walk. The beginning student can skip the advanced discussion on sigma-fields and needs only a primitive understanding that sigma fields organize information within probability spaces.

Section 1.7, intended for somewhat advanced students, discusses the *distribution of a process* and leads to a more mature and mathematically useful understanding of what a stochastic process is rather than what is provided by the elementary definition: *A stochastic process is a collection of random variables $\{X(t), t \in T\}$ defined on a common probability space indexed by the index set T which describes the evolution of some system.* Often $T = [0, \infty)$ if the system evolves in continuous time. For example, $X(t)$ might be the number of people in a queue at time t, or the accumulated claims paid by an insurance company in $[0, t]$. Alternatively, we could have $T = \{0, 1, \dots\}$ if the system evolves in discrete time. Then $X(n)$ might represent the number of arrivals to a queue during the service interval of the nth customer, or the socio-economic status of a family after n generations. When considering stationary processes, $T = \{\dots, -1, 0, 1, \dots\}$ is a common index set. In more exotic processes, T might be a collection of regions, and $X(A)$, the number of points in region A.

1.1. Non-negative Integer Valued Random Variables.

Suppose X is a random variable whose range is $\{0, 1, \dots, \infty\}$. (Allowing a possible value of ∞ is a convenience. For instance, if X is the waiting time for a random event to occur and if this event never occurs, it is natural to think of the value of X as ∞.) Set

$$P[X = k] = p_k, \quad k = 0, 1, \dots,$$

so that

$$P[X < \infty] = \sum_{k=0}^{\infty} p_k, \quad P[X = \infty] = 1 - \sum_{k=0}^{\infty} p_k =: p_\infty.$$

(Note that the notation "=:" means that a definition is being made. Thus $1 - \sum_{k=0}^{\infty} p_k =: p_\infty$ means that p_∞ is defined as $1 - \sum_{k=0}^{\infty} p_k$. In general $A =: B$ or equivalently $B := A$ means B is defined as A.) If $P[X = \infty] > 0$, define $E(X) = \infty$; otherwise

$$E(X) = \sum_{k=0}^{\infty} k p_k.$$

If $f : \{0, 1, \dots, \infty\} \mapsto [0, \infty]$ then in an elementary course you probably saw the derivation of the fact that

$$Ef(X) = \sum_{0 \le k \le \infty} f(k) p_k.$$

If $f : \{0, 1, \dots, \infty\} \mapsto [-\infty, \infty]$ then define two positive functions f^+ and f^- by

$$f^+ = \max\{f, 0\}, \quad f^- = -\min\{f, 0\}$$

so that $Ef^+(X)$ and $Ef^-(X)$ are both well defined and

$$Ef^\pm(X) = \sum_{0 \le k \le \infty} f^\pm(k) p_k.$$

Now define

$$Ef(X) = Ef^+(X) - Ef^-(X)$$

provided at least one of $Ef^+(X)$ and $Ef^-(X)$ is finite. In the contrary case, where both are infinite, the expectation does not exist. The expectation is finite if $\sum_{0 \le k \le \infty} |f(k)| p_k < \infty$.

If $p_\infty = 0$ and

$$f(k) = k^n, \text{then} Ef(X) = EX^n = n\text{th moment};$$

$$f(k) = (k - E(X))^n, \text{ then } Ef(X) = E(X - E(X))^n$$
$$= n\text{th central moment}.$$

In particular, when $n = 2$ in the second case we get

$$\mathrm{Var}(X) = E(X - E(X))^2 = EX^2 - (E(X))^2.$$

Some examples of distributions $\{p_k\}$ that you should review and that will be particularly relevant are

1. Binomial, denoted $b(k; n, p)$, which is the distribution of the number of successes in n Bernoulli trials when the success probability is p. Then

$$P[X = k] = b(k; n, p) := \binom{n}{k} p^k (1-p)^{n-k}, \quad 0 \le k \le n, 0 \le p \le 1,$$

and $E(X) = np, \mathrm{Var}(X) = np(1-p)$.

2. Poisson, denoted $p(k; \lambda)$. Then for $k = 0, 1, \dots, \lambda > 0$

$$P[X = k] = p(k, \lambda) := e^{-\lambda}\lambda^k / k!$$

and $E(X) = \lambda, \mathrm{Var}(X) = \lambda$.

3. Geometric, denoted $g(k; p)$, so that for $k = 0, 1, \dots$

$$P[X = k] = g(k, p) := (1-p)^k p, \quad 0 \le p \le 1,$$

which is the distribution of the number of failures before the first success in repeated Bernoulli trials. The usual notation is to set $q = 1 - p$. Then

$$\begin{aligned} E(X) &= \sum_{k=0}^{\infty} k q^k p = p \sum_{k=1}^{\infty} k q^k \\ &= p \sum_{k=1}^{\infty} \left(\sum_{j=1}^{k} 1 \right) q^k, \end{aligned}$$

and reversing the order of summation yields

$$\begin{aligned} &= p \sum_{j=1}^{\infty} \sum_{k=j}^{\infty} q^k = p \sum_{j=1}^{\infty} q^j / (1-q) \\ &= \sum_{j=1}^{\infty} q^j = q/(1-q) = q/p. \end{aligned} \tag{1.1.1}$$

Alternatively, we could have computed this by summing tail probabilities:

Lemma 1.1.1. *If X is non-negative integer valued then*

$$E(X) = \sum_{k=0}^{\infty} P[X > k]. \tag{1.1.2}$$

Proof. To verify this formula involves reversing the steps of the previous computation:

$$\sum_{k=0}^{\infty} P[X>k] = \sum_{k=0}^{\infty}\sum_{j=k+1}^{\infty} p_j = \sum_{j=1}^{\infty}\left(\sum_{k=0}^{j-1} 1\right) p_j$$
$$= \sum_{j=1}^{\infty} j p_j = E(X). \blacksquare$$

In the multivariate case we have a random vector with non-negative integer valued components $\mathbf{X}' = (X_1, \dots, X_k)$ with a mass function

$$P[X_1 = j_1, \dots, X_k = j_k] = p_{j_1,\dots,j_k}$$

for non-negative integers $j_1, \dots, j_k$. If

$$f : \{0, 1, \dots, \infty\}^k \mapsto [0, \infty]$$

then

$$Ef(X_1, \dots, X_k) = \sum_{(j_1,\dots,j_k)} f(j_1, \dots, j_k) p_{j_1,\dots,j_k}.$$

If $f : \{0, 1, \dots, \infty\}^k \mapsto R$ then

$$Ef(X_1, \dots, X_k) = Ef^+(X_1, \dots, X_k) - Ef^-(X_1, \dots, X_k)$$

as in the case $k = 1$ if at least one of the expectations on the right is finite. Now recall the following *properties:*

1. For $a_1, \dots, a_k \in R$

$$E\left(\sum_{i=1}^{k} a_i X_i\right) = \sum_{i=1}^{k} a_i E(X_i)$$

(provided the right side makes sense; no $\infty - \infty$ please!).

2. If $X_1, \dots, X_k$ are independent so that the joint mass function of $X_1, \dots, X_k$ factors into a product of marginal mass functions, then for any bounded functions $f_1, \dots, f_k$ with domain $\{0, 1, \dots, \infty\}$ we have

$$E\prod_{i=1}^{k} f_i(X_i) = \prod_{i=1}^{k} Ef_i(X_i). \tag{1.1.3}$$