

# Probability: Theory and Examples

## Solutions Manual

The creation of this solution manual was one of the most important improvements in the second edition of *Probability: Theory and Examples*. The solutions are not intended to be as polished as the proofs in the book, but are supposed to give enough of the details so that little is left to the reader's imagination. It is inevitable that some of the many solutions will contain errors. If you find mistakes or better solutions send them via e-mail to [rtd1@cornell.edu](mailto:rtd1@cornell.edu) or via post to Rick Durrett, Dept. of Math., 523 Malott Hall, Cornell U., Ithaca NY 14853.

*Rick Durrett*

# Contents

## 1 Laws of Large Numbers 1

1. Basic Definitions 1
2. Random Variables 3
3. Expected Value 4
4. Independence 7
5. Weak Laws of Large Numbers 12
6. Borel-Cantelli Lemmas 15
7. Strong Law of Large Numbers 19
8. Convergence of Random Series 20
9. Large Deviations 24

## 2 Central Limit Theorems 26

1. The De Moivre-Laplace Theorem 26
2. Weak Convergence 27
3. Characteristic Functions 31
4. Central Limit Theorems 35
6. Poisson Convergence 39
7. Stable Laws 43
8. Infinitely Divisible Distributions 45
9. Limit theorems in  $\mathbf{R}^d$  46

## 3 Random walks 48

1. Stopping Times 48
4. Renewal theory 51

## 4 Martingales 54

1. Conditional Expectation 54
2. Martingales, Almost Sure Convergence 57
3. Examples 43
4. Doob's Inequality,  $L^p$  Convergence 64
5. Uniform Integrability, Convergence in  $L^1$  66
6. Backwards Martingales 68
7. Optional Stopping Theorems 69

## 5 Markov Chains 74

1. Definitions and Examples 74
2. Extensions of the Markov Property 75
3. Recurrence and Transience 79
4. Stationary Measures 82
5. Asymptotic Behavior 84
6. General State Space 88

## 6 Ergodic Theorems 91

1. Definitions and Examples 91
2. Birkhoff's Ergodic Theorem 93
3. Recurrence 95
6. A Subadditive Ergodic Theorem 96
7. Applications 97

## 7 Brownian Motion 98

1. Definition and Construction 98
2. Markov Property, Blumenthal's 0-1 Law 99
3. Stopping Times, Strong Markov Property 100
4. Maxima and Zeros 101
5. Martingales 102
6. Donsker's Theorem 105
7. CLT's for Dependent Variables 106
8. Empirical Distributions, Brownian Bridge 107
9. Laws of the Iterated Logarithm 107

## **Appendix: Measure Theory 108**

1. Lebesgue-Stieltjes Measures 108
2. Carathéodary's Extension Theorem 109
3. Completion, etc. 109
4. Integration 109
5. Properties of the Integral 112
6. Product Measures, Fubini's Theorem 114
8. Radon-Nikodym Theorem 116

# 1 Laws of Large Numbers

## 1.1. Basic Definitions

1.1. (i)  $A$  and  $B-A$  are disjoint with  $B = A \cup (B-A)$  so  $P(A) + P(B-A) = P(B)$  and rearranging gives the desired result.

(ii) Let  $A'_n = A_n \cap A$ ,  $B_1 = A'_1$  and for  $n > 1$ ,  $B_n = A'_n - \cup_{m=1}^{n-1} A'_m$ . Since the  $B_n$  are disjoint and have union  $A$  we have using (i) and  $B_m \subset A_m$

$$P(A) = \sum_{m=1}^{\infty} P(B_m) \leq \sum_{m=1}^{\infty} P(A_m)$$

(iii) Let  $B_n = A_n - A_{n-1}$ . Then the  $B_n$  are disjoint and have  $\cup_{m=1}^{\infty} B_m = A$ ,  $\cup_{m=1}^n B_m = A_n$  so

$$P(A) = \sum_{m=1}^{\infty} P(B_m) = \lim_{n \rightarrow \infty} \sum_{m=1}^n P(B_m) = \lim_{n \rightarrow \infty} P(A_n)$$

(iv)  $A_n^c \uparrow A^c$  so (iii) implies  $P(A_n^c) \uparrow P(A^c)$ . Since  $P(B^c) = 1 - P(B)$  it follows that  $P(A_n) \downarrow P(A)$ .

1.2. (i) Suppose  $A \in \mathcal{F}_i$  for all  $i$ . Then since each  $\mathcal{F}_i$  is a  $\sigma$ -field,  $A^c \in \mathcal{F}_i$  for each  $i$ . Suppose  $A_1, A_2, \dots$  is a countable sequence of disjoint sets that are in  $\mathcal{F}_i$  for all  $i$ . Then since each  $\mathcal{F}_i$  is a  $\sigma$ -field,  $A = \cup_m A_m \in \mathcal{F}_i$  for each  $i$ .

(ii) We take the intersection of all the  $\sigma$ -fields containing  $\mathcal{A}$ . The collection of all subsets of  $\Omega$  is a  $\sigma$ -field so the collection is not empty.

1.3. It suffices to show that if  $\mathcal{F}$  is the  $\sigma$ -field generated by  $(a_1, b_1) \times \dots \times (a_n, b_n)$ , then  $\mathcal{F}$  contains (i) the open sets and (ii) all sets of the form  $A_1 \times \dots \times A_n$  where  $A_i \in \mathcal{R}$ . For (i) note that if  $G$  is open and  $x \in G$  then there is a set of the form  $(a_1, b_1) \times \dots \times (a_n, b_n)$  with  $a_i, b_i \in \mathbf{Q}$  that contains  $x$  and lies in  $G$ , so any open set is a countable union of our basic sets. For (ii) fix  $A_2, \dots, A_n$  and

6.5. Pick  $N_k$  so that if  $m, n \geq N_k$  then  $d(X_m, X_n) \leq 2^{-k}$ . Given a subsequence  $X_{n(m)}$  pick  $m_k$  increasing so that  $n(m_k) \geq N_k$ . Using Chebyshev's inequality with  $\varphi(z) = z/(1+z)$  we have

$$P(|X_{n(m_k)} - X_{n(m_{k+1})}| > k^{-2}) \leq (k^2 + 1)2^{-k}$$

The right hand side is summable so the Borel-Cantelli lemma implies that for large  $k$ , we have  $|X_{n(m_k)} - X_{n(m_{k+1})}| \leq k^{-2}$ . Since  $\sum_k k^{-2} < \infty$  this and the triangle inequality imply that  $X_{n(m_k)}$  converges a.s. to a limit  $X$ . To see that the limit does not depend on the subsequence note that if  $X_{n'(m'_k)} \rightarrow X'$  then our original assumption implies  $d(X_{n(m_k)}, X_{n'(m'_k)}) \rightarrow 0$ , and the bounded convergence theorem implies  $d(X, X') = 0$ . The desired result now follows from (6.2).

6.6. Clearly,  $P(\cup_{m \geq n} A_m) \geq \max_{m \geq n} P(A_m)$ . Letting  $n \rightarrow \infty$  and using (iv) of (1.1), it follows that  $P(\limsup A_m) \geq \limsup P(A_m)$ . The result for  $\liminf$  can be proved by imitating the proof of the first result or applying it to  $A_m^c$ .

6.7. Using Chebyshev's inequality we have for large  $n$

$$P(|X_n - EX_n| > \delta EX_n) \leq \frac{\text{var}(X_n)}{\delta^2 (EX_n)^2} \leq \frac{Bn^\beta}{\delta^2 (a^2/2)n^{2\alpha}} = Cn^{\beta-2\alpha}$$

If we let  $n_k = [k^{2/(2\alpha-\beta)}] + 1$  and  $T_k = X_{n_k}$  then the last result says

$$P(|T_k - ET_k| > \delta ET_k) \leq Ck^{-2}$$

so the Borel Cantelli lemma implies  $T_k/ET_k \rightarrow 1$  almost surely. Since we have  $ET_{k+1}/ET_k \rightarrow 1$  the rest of the proof is the same as in the proof of (6.8).

6.8. Exercise 4.16 implies that we can subdivide  $X_n$  with large  $\lambda_n$  into several independent Poissons with mean  $\leq 1$  so we can suppose without loss of generality that  $\lambda_n \leq 1$ . Once we do this and notice that for a Poisson  $\text{var}(X_m) = EX_m$  the proof is almost the same as that of (6.8).

6.9. The events  $\{\ell_n = 0\} = \{X_n = 0\}$  are independent and have probability  $1/2$ , so the second Borel Cantelli lemma implies that  $P(\ell_n = 0 \text{ i.o.}) = 1$ . To prove the other result let  $r_1 = 1$ ,  $r_2 = 2$  and  $r_n = r_{n-1} + [\log_2 n]$ . Let  $A_n = \{X_m = 1 \text{ for } r_{n-1} < m \leq r_n\}$ .  $P(A_n) \geq 1/n$ , so it follows from the second Borel Cantelli lemma that  $P(A_n \text{ i.o.}) = 1$ , and hence  $\ell_{r_n} \geq [\log_2 n]$  i.o. Since  $r_n \leq n \log_2 n$  we have

$$\frac{\ell_{r_n}}{\log_2(r_n)} \geq \frac{[\log_2 n]}{\log_2 n + \log_2 \log_2 n}$$

## 2.9. Limit Theorems in $\mathbf{R}^d$

9.1.

$$F_i(x) = P(X_i \leq x)$$

$$\lim_{n \rightarrow \infty} P(X_1 \leq n, \dots, X_{i-1} \leq n, X_i \leq x, X_{i+1} \leq n, \dots, X_d \leq n)$$

$$\lim_{n \rightarrow \infty} F(n, \dots, n, x, n, \dots, n)$$

where the  $x$  is in the  $i$ th place and  $n$ 's in the others.

9.2. It is clear that  $F$  has properties (ii) and (iii). To check (iv) let  $G(x) = \prod_{i=1}^d F_i(x_i)$  and  $H(x) = \prod_{i=1}^d F_i(x_i)(1 - F_i(x_i))$ . Using the notation introduced just before (iv)

$$\sum_v \operatorname{sgn}(v)G(v) = \prod_{i=1}^d F_i(b_i) - F_i(a_i)$$

$$\sum_v \operatorname{sgn}(v)H(v) = \prod_{i=1}^d \{F_i(b_i)(1 - F_i(b_i)) - F_i(a_i)(1 - F_i(a_i))\}$$

To show  $\sum_v \operatorname{sgn}(v)(G(v) + \alpha H(v)) \geq 0$  we note

$$F_i(b_i)(1 - F_i(b_i)) - F_i(a_i)(1 - F_i(a_i))$$

$$= \{F_i(b_i) - F_i(a_i)\}(1 - F_i(a_i))$$

$$+ F_i(a_i)\{(1 - F_i(b_i)) - (1 - F_i(a_i))\}$$

$$= \{1 - F_i(b_i) - F_i(a_i)\}(F_i(b_i) - F_i(a_i))$$

and  $|1 - F_i(b_i) - F_i(a_i)| \leq 1$ .

9.3. Each partial derivative kills one intergal.

9.4. If  $K$  is closed,  $H = \{x : x_i \in K\}$  is closed. So

$$\limsup_{n \rightarrow \infty} P(X_{n,i} \in K) = \limsup_{n \rightarrow \infty} P(X_n \in H) \leq P(X \in H) = P(X_i \in K)$$

9.5. If  $X$  has ch.f.  $\varphi$  then the vector  $Y = (X, \dots, X)$  has ch.f.

$$\psi(t) = E \exp \left( i \sum_j t_j X \right) = \varphi \left( \sum_j t_j \right)$$

3.5. First note that (3.3) implies  $\bar{E}T_1 = 1/P(X_0 = 1)$ , so the right hand side is  $P(X_0 = 1, T_1 \geq n)$ . To compute the left now we break things down according to the position of the first 1 to the left of 0 and use translation invariance to conclude  $P(T_1 = n)$  is

$$\begin{aligned} &= \sum_{m=0}^{\infty} P(X_{-m} = 1, X_j = 0 \text{ for } j \in (-m, n), X_n = 1) \\ &= \sum_{m=0}^{\infty} P(X_0 = 1, X_j = 0 \text{ for } j \in (0, m+n), X_{m+n} = 1) \\ &= P(X_0 = 1, T_1 \geq n) \end{aligned}$$

## 6.6. A Subadditive Ergodic Theorem

6.1. (1.3) implies that the stationary sequences in (ii) are ergodic. Exercise 3.1 implies  $EX_{0,n} = \sum_{m=1}^n P(S_1 \neq 0, \dots, S_n \neq 0)$ . Since  $P(S_1 \neq 0, \dots, S_n \neq 0)$  is decreasing it follows easily that  $EX_{0,n}/n \rightarrow P(\text{no return to } 0)$ .

6.2. (a)  $EL_1 = P(X_1 = Y_1) = 1/2$ . To compute  $EL_2$  let  $N_2 = |\{i \leq 2 : X_i = Y_i\}|$  and note that  $L_2 - N_2 = 0$  unless  $(X_1, X_2, Y_1, Y_2)$  is  $(1, 0, 0, 1)$  or  $(0, 1, 1, 0)$ . In these two cases which have probability  $1/16$  each  $L_2 - N_2 = 1$  so  $EL_2 = EN_2 + 1/8 = 9/8$  so  $EL_2/2 = 9/16$

(b) The expected number of sequences of length  $K$  is  $\binom{n}{K} 2^{-K}$ . Taking  $K = an$  using Stirling's formula  $m! \sim m^m e^{-m} \sqrt{2\pi m}$  without the term under the square root we have that the above

$$\approx \frac{n^{2n} 2^{-an}}{(an)^{2an} ((1-a)n)^{2(1-a)n}} = (a^{2a} (1-a)^{2(1-a)} 2^a)^{-n}$$

From the last computation it follows that

$$\frac{1}{n} \log \left( \binom{n}{na}^2 2^{-na} \right) \rightarrow -2a \log a - 2(1-a) \log(1-a) - a \log 2$$

When  $a = 1$  the right hand side is  $-\log 2 < 0$ . By continuity it is also negative for  $a$  close to 1.