PROBABILITY THEORY AND EXAMPLES. 2nd edition

Updated typo list, December, 2002; * indicates typo not on May 5, 2000 list. Thanks to spell checking I have found many more misspelled words. These and other small points of grammar have not been added to the list.

Contributions from Manel Baucells, Eric Blair, Zhen-Qing Chen, Ted Cox, Bradford R. Crain, Winston K. Crandall, Amir Dembo, Neil Falkner, Changyong Feng, Boris Granovsky, Jan Hannig, Andrew Hayen, Martin Hildebrand, Kyoungmun Jang, Anatole Joffe, Daniel Kifer, Steve Krone, T.Y. Lee, Shlomo Levental, Torgny Lindvall, Arif Mardin, Carl Mueller, Robin Pemantle, Yuval Peres, Mark Pinsky, Ross Pinsky, Boris Pittel, David Pokorny, Vinayak Prabhu, Brett Presnell, Jim Propp, Yossi Schwarzfuchs, Rami Shakarchi, Lian Shen, Marc Shivers, Bob Strain, Rich Sowers, Tsachy Weissman, Hao Zhang

Preface page iv, line -7: hardly fail to adopt it

page v, line -6: before the final assault

Table of Contents

page viii, line 8: Erdős (also in other places in the book)

page viii, Stopping Times 173

page x, line 3: Carathéodory (also in the headings of A.2)

Chapter 1

p.5, Remark at bottom: this is wrong. $(\inf, \geq)$ also works exactly. The remark is sort of silly since all of the four pairs miss only on a countable set so any of them does the job

p.6, Figure 1.1.2: $F(y_1)$ and $F(y_2)$ instead of $F^{-1}(y_1)$ and $F^{-1}(y_2)$.

p.7, line -6: no closed form expression for $F(x)$

p.11, line 3: $f : (\mathbf{R}^n, \mathcal{R}^n) \to (\mathbf{R}, \mathcal{R})$ is meant, as the proof makes it clear.

* p.12, line 10: can take the value $+\infty$ or $-\infty$.

p.14, line 8: (5.1) and (5.2)

p.15, line 1: the proof of (5.1)

p.16, line -4

p.16, Exercise 3.6: Suppose $a > -b$.

p.16, Exercise 3.8: $0 \leq a < EY$

p.16, line -11: (5.4)–(5.6)

p.16, line -4: and (3.1c) imply

p.17, (3.8). there is no reason to suppose $h \geq 0$

p.17, (b): $E(|h(\bar{Y})|; Y > M) \leq \epsilon_M Eg(Y) + |h(0)|P(Y > M)$ This requires a number of subsequent changes.

p.22, Exercise 3.16: $\lim_{y\to 0+} yE(1/X; X > y) = 0$

p.23, line -9: The first definition above is, in turn, a special case of the second p.24, line 12: $i \in I$ and $B_i = \mathbf{R}$

p.25, line 11: so are $\overline{\mathcal{A}_i} = \mathcal{A}_i \cup \{\Omega\}$, so

p.25, line -8: $\Omega \in \mathcal{A}_i$

* p.25, line -1: $B_k \in \mathcal{L}$ not $A, B \in \mathcal{L}$

p.28, line -3: If $X_1, \dots X_n$ are independent, and $X_i \geq 0$ for $1 \leq i \leq n$ or $E|X_i| < \infty$ for $1 \leq i \leq n$

p.31, lines -6 to -2: the calculation is unnecessary, noting that $f_{X,Y}(x)$ is a density function, so the constants must work out right. (Neil Falkner)

p.33, line 9: so that $\varphi(S)$ is a Borel subset of $\mathbf{R}$ and both φ and φ^{-1} are measurable

p.33, line -7: see Exercise 4.10

p.34, line 2: the number (i) is not necessary

p.36, line -11: Z_n

p.39, last three lines. $X_{n,k} = 0$ otherwise

* p.40, line 6 of proof: let $i_k = \inf(\{1, 2, \dots n\} - \{i_1, \dots, i_{k-1}\})$

p.40, (*): add: in probability

p.40, line -6. We will see in Example 4.6

p.43, line -11. $E(\bar{X}_{n,1}^2)/n \leq$, i.e., bar is missing and we have an inequality

p.43, line -6: $[K, n]$

p.44, line 10: it follows that $P(|S_n/n - \mu| > \varepsilon) \to 0$

p.45, header: Weaks Laws

* p.46, Problem 5.7: Generalize (5.7)

p.47, Exercise 5.9: Assume $X_1, X_2, \dots$ are i.i.d., $P(0 \leq X_i < \infty) = 1$, and $P(X_i > 0) > 0$.

p.48, line 10: subsequence

p.48, Remark: Exercises 6.14 and 6.15

p.48, line -6: convergence

p.49, Exercise 6.3: $|x|/g(x) \to 0$

p.52, line -2: see Exercise 4.7

p.53, line -7: delete $P(A_{m_2} \cap \dots \cap A_{m_k})$

* p.54, line 9: middle = should be $\geq$

p.55, Exercise 6.13: $\sup_n X_n < \infty$ a.s.

p.55, Exercise 6.16: the σ-field $\mathcal{F}$ consists of all subsets

p.55, Exercise 6.18: $\limsup_{n\to\infty} X_n / \log n = 1$ a.s.

p.55, line -1: end with a comma

p.56, line 7: proof of the weak law

p.58, line 1: $\to 0$ a.s.

p.59, line 1: By (7.1) and (7.2), ...

p.61, line 3: measure

p.61, line 6: missing). For clarity it should be written as $P\{\ldots\} \to 1$

p.61, Exercise 7.3: $n^{-1} \log |X_n|$

p.64, lines 2 and 4: $\sup_{m \geq M}$

p.64, line -9: Example 4.7

p.67, lines -7 to -5: the term $m^{1/p} P(|X_i| > m^{1/p})$ is missing from the formulas. The following alternative argument was suggested by Neil Falkner.

$$\begin{aligned} |\mu_m| &\leq E(|X_i|; |X_i| > m^{1/p}) = m^{1/p} E(|X_i|/m^{1/p}; |X_i| > m^{1/p}) \\ &\leq m^{1/p} E\left((|X_i|/m^{1/p})^p; |X_i| > m^{1/p} \right) = m^{-1+1/p} E(|X_i|^p; |X_i| > m^{1/p}). \end{aligned}$$

p.67, line -2: of the last result

* p.69, Exercise 8.4: $\sum_n X_n/n$ converges a.s., and hence $n^{-1} \sum_{m=1}^n X_m \to 0$ a.s.

p.69, Exercise 8.5: (missing parentheses)

$$\sum_{n=1}^{\infty} \left(P(X_n > 1) + E(X_n 1_{(X_n \leq 1)}) \right)$$

p.70, Exercise 8.12: item number (i) is missing: (i) Use (*) to show that...

in (ii) $\varepsilon > 0$

p.71, second line of Exercise 9.2. delete)/2 from end of formula

p.71, last line: and $\theta > 0$ is small, then

p.72, line 3: > 0

p.72, line -3: whenever

p.75, line 5: dF_λ^n

p.75, line -10: and $\nu > a$ are

p.76, line -7: $\varphi'(\theta_a)/\varphi(\theta_a)$

p.77, line 14: $\varphi(\lambda)$ should be $\log \varphi(\lambda)$

Chapter 2

p.80, Exercise 1.1: change the condition to: $\max_{1 \leq i \leq n} |c_{j,n}| \to 0$ and $\sup_n \sum_{j=1}^n |c_{j,n}| < \infty$. (The second one follows automatically if all $c_{j,n}$ have the same sign.)

p.84, line -2: Scheffé's theorem

p.85, Exercise 2.3: $P(X_i > x) \sim \frac{1}{\sqrt{2\pi}} \frac{1}{x} e^{-x^2/2}$

p.87, line -9: f is bounded and continuous

p.89, line 6: that occurs in (2.5)

* p.89, final display in proof of (2.5): n_k not n

p.91, Exercise 2.10: c is a constant. Then p.91, line -6: the metric of Exercise 6.4

p.93, line 1: $X + Y$ should be $X_1 + X_2$

p.96, line 15: see Exercise 6.6

p.98, line -8: Exercise 4.4 in the Appendix

p.100, line -17: sequence

p.101, Exercise 3.13: $t = 3^k\pi$, $k = 0, 1, 2, \ldots$

p.103, line -9: don't need to go to Arzela-Ascoli

p.103, line -2: in combination with Exercise 5.4

p.104, Exercise 3.21: by Exercise 8.10 in Chapter 1

p.105, line 6: $\lim_{t\to 0} \varphi(t) = 1$

p.107, Exercise 3.26: $Y = 0$ a.s.

p.107, line -8: tight by (2.7)

* p.110, line 9: Taking $x = tX$ in (3.6) and multiplying by $e^{i\theta X}$

p.113, line 7: $(1 + \frac{\gamma}{n})^{n-1}$ has to be changed to $(e^{\frac{\gamma}{n}})^{n-1}$

p.115, line -10: Assume also that $\text{Var}(X_i) \in (0, \infty)$

p.116, Exercise 4.6: (three errors). Suppose $a_n \to \infty$; Kolmogorov's inequality is (8.2); in the hint use the notation Z_n instead of X_n.

p.116, Exercise 4.7: $\text{Var}(Y_i) = \sigma^2 \in (0, \infty)$

p.118, line 3: apply Exercise 1.1 with $c_{m,n} = -t^2\sigma^2_{n,m}$

p.118, line 5: $\sum_{m=1}^n c_{m,n} \to -\sigma^2 t^2/2$

p.119, line -6: (8.3) in Chapter 1

p.121, line 9: (ii) of Theorem (4.5)

p.121, Exercise 4.13: $\beta > 0$

p.122, line 2: deriving

* p.143, line 5: fraction of empty boxes

* p.144, (e) $ne^{-r/n} \to \lambda$

p.147, line 4: T'_2 should be T'_k

p.148, Exercise 6.7: assume $U_1, \ldots, U_n$ are independent

p.149, line 14: $m(A) = \{j \le N : X_j \in A\}$. The problems with the other formula are: 1) if ν has atoms the X_i may coincide; 2) even worse, if $\mathcal{S}$ is ugly, $|A \cap \{X_1, \ldots, X_n\}|$ may be non-measurable. The latter cannot happen, if there is a countable set $B_1, B_2, \ldots \in \mathcal{S}$ that separates the points of S.

p.154, line 1: $n^{1/\alpha}$ should be a_n

p.155, line 7: $n^{1/\alpha}$ should be a_n

p.157, lines 6: when $\alpha = 1/2$, $\kappa = 1$ and $c = 0$, $b = 1$, the density is

Chapter 3

p.174, discussion preceding (1.1): interchange the notations A and B (the meaning of A on the preceding page is quite different)

p.176, line 3: in Exercise 1.11

p.176, Example 1.1: $N = \inf\{k : |S_k| \geq x\}$

p.177, line 7: $S = \mathbf{R}$

p.178, line 7: cemetery

some inconsistency in writing or not writing out the argument ω:

p.178, line -12: $\tau_n(\omega) = \tau_{n-1}(\omega) + \dots$

p.178, line -9: $T_n(\omega) = T_{n-1}(\omega) + \dots$

p.178, line -4: $T(\theta^{T_{n-1}}) < \infty$ is independent of $T_{n-1} < \infty$

p.179, line 9: $\alpha_k(\omega) = \alpha_{k-1}(\omega) + \dots$

p.179, Exercise 1.9: β instead of $\beta(\omega)$

p.179, Exercise 1.10: $\overline{\beta}$ instead of $\overline{\beta}(\omega)$

p.184, (2.1): closed subgroup of $\mathbf{R}^d$

p.184, line -4: $p_{\delta,m}(z)$

* p.189, line 4: conflicting use of m. $T_k = \inf\{\ell \geq 0 : S_\ell \in k\varepsilon + [0,\varepsilon)^d\}$. This changes propagates to the end of the proof and one has to change the ℓ's in the last display to j's

* p.189, line 7: $\sum_{n=0}^{\infty} \sum_{\ell=0}^{n}$

p.191, line 5: $\varphi(t) = E\exp(it \cdot X_j)$, $t \in \mathbf{R}^d$, is

p.191, line -2: $e^{it \cdot x}$

p.196, line 3: Exercise (misspelled)

p.197, lines -11 and -5: conflicting uses of x; $(0, x)$ vs. x-axis. Change x-axis to horizontal axis

p.201, line -4: $u_{2k} = \sum_{m=1}^{k} f_{2m} u_{2k-2m}$

p.204, line -10: see (7.3)

p.204, Theorem (4.2): in the proof S_n and S_{N_t} should be T_n and T_{N_t}.

* p.205, Exercise 4.1: replace X_i by ξ_i

Chapter 4

* p.224, line 9: $\int_A g(X)dP$ (previously listed as p. 124)

p.229, Exercise 1.10: positive integer valued r.v. with $EN^2 < \infty$

p.229, Exercise 1.11: $EX^2 = EY^2 < \infty$

p.234, line -3: defines a

p.235, proof of (2.9): 1) in the first line: Let $Y_m = a + (X_m - a)^+$. By (2.9) Y_m is a...

2) in the second line: the same number of times that X_m does

p.236, line -14: submartingale

p.236, line -7: $P(\xi_i = 1) = P(\xi_i = -1) = 1/2$

p.237, Example 2.3: it is not necessary to write k instead of n after the first line

p.237, (2.12): A_n is a predictable increasing sequence

p.238, Exercise 2.6: $\text{var}(\xi_m) = \sigma_m^2 < \infty$

p.241, line -7: In Example 4.5

p.244, line -4: $X_n \to X$ ν-a.s.

* p.245, line 11: To complete the proof without using (3.3): X_n converges to X in $L^1(\nu)$, therefore for any $A \in \mathcal{F}_m$ and $n \geq m$ we have

$$\mu_n(A) = \int_A X_n d\nu \overset{n\to\infty}{\longrightarrow} \int_A X d\nu,$$

therefore $\mu(A) = \int_A X d\nu$. Now apply the π-λ theorem.

p.246, line 3: If $0 < \mu < \infty$ then Z_n/μ^n is a martingale

p.247, line 7: martingale

p.250, line -11: $\max_{1\leq m\leq n} |S_m|$

p.250, line -8: Example 4.7

p.251, Theorem (4.3): the way the theorem is stated, it requires $p < \infty$

p.251, lines 8,10 and -3: $(p/(p-1))^p$

p.253, line -2: generalize (8.3) and (8.7) from Chapter 1

p.264, line -7: of (5.7) is

p.268, proof of the ballot theorem: actually if one lets $H = \{S_n < n\}$ then one has

$$P(G|S_n) = 1_H(n - S_n)/n$$

always, i.e., one does not need to assume $P(\xi_j \leq 2)$. The reasoning here is that nonnegativity of the ξ_j implies $S_{j-1} \leq S_j$ so if $S_j < j$, the worst that can happen is $S_{j-1} = j - 1$. This observation is due to Ted Cox.

p.273, line 7: Exercise 1.7

p.273, line preceding Theorem (7.5): Wald's equation is (1.6) in Chapter 3

p.274, Example 7.1: The assumption $p > 1/2$ is not needed in part (a), but it is needed in parts (c) and (d)

p.275, (c): $\inf_n S_n \leq a$

* p.275, line 3 of proof of (d): add parentheses after E and at end of formula for clarity

p.275, line 4 of the proof of (d): use (c) not Exercise 7.3

p.276, Exercise 7.8: $P(Z_n = 0 \text{ for some } n \geq 1 | Z_0 = x) = \rho^x$. ($P_x$ has not been introduced yet.)

Chapter 5

p.286, line 5: $\mathcal{A}$ should be A

* p.286, The three references to (1.4) in the last paragraph should all be (1.5)

* p.287, line -7: (5.9) in Chapter 4

p.288, Reflection principle: for the proof to work assume that $\xi_1, \xi_2, \dots$ are i.i.d. Otherwise "Formal proof" on p.289 has to be modified to deal with a nonhomogeneous Markov chain.

* p.288, proof of (2.4): using (2.1) now

p.289, line 6: distribution

p.293, line 5: $y_i \neq x$

p.299, line -13: exercise should be theorem

* p.300, line 7: use (3.9)

p.304, line -5: By (3.4)

* p.304, line -1: Example 1.3

* p.305, line -7, -6: μ_a instead of μ twice

* p.308, Exercise 4.7: Example 1.3

* p.312, first and third display of proof: P_y-a.s.

* p.312, next to last display: P_x-a.s.

p.317, line 10: Example 1.3

p.318, line -17: Example 5.3

p.319, Convergence theorem: S_d should be S_{d-1}

* p.321, last line: $i = 0, 1$

* p.323, Exercise 5.4: Assume the a's are defined by the formula in Example 1.4. In particular $a_j > 0$ for all $j \geq 0$.

p.323, line -12: Example 1.4

* p.323, line -9: ξ_m not ξ_n

p.326, line 5: $\mu = \delta_b$ should be $\rho = \delta_b$

Chapter 6

p.343, line -4: $X' = X - \alpha$

p.345, line -11: $\log_{10}(k+1)$

p.345, line -9: 1_A should be 1_{A_k}

p.348, Exercise 3.2: Assume also that $X_1, X_2, \dots$ is ergodic.

* p.348, Exercise 3.3: assume the X_i are integer valued.

p.363, line 11: $\sup_{m \geq 1} E(L_{0,m}/m)$

p.365, line 2: $=$ should be $\leq$

p.369, line -10: Exercise 7.2

p.371, line 12: $P(Z_n(an) \geq 1, \text{ i.o.})$

Chapter 7

p.378, line -3: $|X(q) - X(r)| \le A|q - r|^\gamma$

p.378, line -2: (1.6) implies (1.5)

p.380, line -2: change 2^{n-1} to 2^{-n} in both places

p.382, line -7: (4.1) in Chapter 1

p.383, line 5: (1.5) in Chapter 5

* p.384. In the proof of (2.5) the reference to (2.5) should be (2.4)

p.385, line 5: Use (2.5)

* p.386, line 3 of proof of (2.9): 1_A not A

p.387, Last rites: add the comment that $\mathcal{F}_s$ is also right continuous

* p.390, (3.7): for consistency we should write $Y_s(\omega)$

p.391, line -11: the indices are missing from f

$$f_0(s) \int dy_1 p_{t_1}(x, y_1) f_1(y_1) \cdots \int dy_n p_{t_n - t_{n-1}}(y_{n-1}, y_n) f_n(y_n)$$

p.392, line 5: $A = G_0 \times \{\omega : \omega(s_j) \in G_j, 1 \le j \le k\}$

p.392, line 9: $Y_s^n(\omega) = f_0^n(s) \prod_{j=1}^k f_j^n(\omega(s_j))$

p.392, line -9: (2.7) instead of (2.9)

* p.395, Exercise 4.1: $M_t = \max_{0 \le s \le t} B_s$

p.400, proof of Theorem (5.7): $E_0 \exp(\theta B(T \wedge t) - \theta^2 (T \wedge t)/2)$

* p.400, bottom: Let $f(x, t, \theta) =$, $f_k(x, t, \theta)$ be the kth derivative, and $h_k(x, t) = f_k(x, t, 0)$

* p.401, top: $h_k(B_t, t)$ is a martingale. We have seen $h_1(B_t, t)$ and $h_2(B_t)$. In table $h_k(\theta, t)$ and $f(x, t, \theta)$.

p.402, Exercise 5.6: Two conflicting uses of a. Let $T = \inf\{t : B_t \notin (-\ell, \ell)\}$

p.402, Exercise 5.7: $\exp(B_t^2/(2(1+t)))$

p.404, line -7: Root (1969)

p.407, line 6: Example 4.1

p.407, line 10: given in (4.9)

p.407, line -1 (two errors): in (3.8) of Chapter 3 that if ... and $P(S_m = 0) = 0$ for all $m \ge 1$

p.410, line -13: $E\varphi(B(\cdot))$ (E is missing)

p.411, line -3: $S_j = s_j$, $0 \le j \le k - 1$

p.414, (ii): for all $\epsilon > 0$

p.417, line -9: on the right hand side $|X_m| > \varepsilon\sqrt{n}$ should be $|X_m| > \varepsilon\sigma\sqrt{n}$

p.419, line 5: $\theta^{-1}\mathcal{F}_m = \mathcal{F}_{m-1}$ (?)

p.419, line -9: X_{n+K} should be X_{n+k}

p.422, line 4: the right hand side should be

$$\|f\|_\infty^2 \sum_x \pi(x) \|p^n(x,\cdot) - \pi(\cdot)\|^2$$

p.425, line 8: last two Y_2's should be Y

p.425, line 10: $\|Y\|_q^{1/\theta+1/q}$ should be $\|Y\|_q$

p.426, line 10: ')' missing from $E(YE(X|\mathcal{F}))$

p.429, In (8.2): $\inf_{0<y<1} \widehat{G}_n(y) - y$

p.434, line 10: (1.4) from Chapter 1

p.437, line 3: By (6.3)

p.437, line -13: $\max_{t_{k-1} \le s \le t_{k+2}} |B(s) - B(t_{k-1})|$

Appendix

p.447, line 7: tackling

p.459, line -13: member of the class

p.460, Exercise 4.2: Let $0 \le f < \infty$ a.e.

p.462, line 3: fairly intuitive applications of these ideas

p.480, (8.6): measures

p.483, Example 9.2: the right hand side should be $-\sum_{n=1}^{\infty} n(1-p)^{n-1}$

References

p.485, line 7: Convergence of probability measures

p.486, Derriennic (1983) Un not Une