Probability : Theory and Examples

-- by Richard Durrett (Cornell University)

We (the members of Shigekawa's laboratory who enter in 2007) use this textbook in our seminar and I read through from Chapter 4 to Chatpter 7 in this book. I'll write down what I notice , consider and want to correct etc.

Chapter 4. Martingale

Section 4.1 Conditional Expectation

[Comments]

• We feel clear if we show this lemma.

Lemma 1. *(Lemma of A_ε)*
Let $\mathcal{F}$ is σ-field. Then,

$$\int_A X\,dP = \int_A Y\,dP \ (^\forall A \in \mathcal{F}) \Longrightarrow X = Y \text{ a.s.}$$

Proof.
It is enough to show in case $Y=0$ and $X \geqslant 0$ since it can be raplaced X as $X-Y$ and X is decomposed as $X = X^+ - X^-$. Let ε is arbitrary, >0 and $\in \mathbb{Q}$. Let $A_\varepsilon := \{\omega | X(\omega) \geqslant \varepsilon\}$. Then $0 = \int_A X\,dP \geqslant \varepsilon P(A_\varepsilon)$ so that $P(A_\varepsilon) = 0$. So $P(X>0) = P(\bigcup_{\varepsilon>0} A_\varepsilon) \leqslant \sum_{\varepsilon>0} P(A_\varepsilon) = 0$. It means $X=0$ a.s.

□

[Outlines]

• Theorem 1.6

We should make 3 steps.

1. $F(x,\omega) := \inf\{G(q,\omega) \,|\, q>x\} \Longrightarrow F(x,\omega) = P(\varphi(X) \leqslant x \,|\, \mathcal{G})$
2. $^{\exists 1}\nu(\omega,\cdot)$ on $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ for every $\omega \in \Omega_0$
3. $A \in \mathcal{S}$ is defined uniquely for every $B \in \mathfrak{B}(\mathbb{R})$ and r.c.d. $\mu(\omega, A)$ is defined as : $\mu(\omega,A) = \nu(\omega,B)$

Section 4.3 Examples

[Corrections]

• p.239 l.3

$$P(G_n = m+2) = \frac{n!}{m!(n-m)!} \frac{2(m+1)!(n-m)!}{(n+2)!} \sim \frac{2x}{n}$$

• p.241 l.26 Let μ and ν be probability measures on sequence space

[Comments]

• Does $X_n = \prod_{1 \leqslant m \leqslant n} q_m(\omega_m)$ satisfy the assumption of theorem 3.3 ?

Section 4.4 Doob's Inequality, Convergence in L^p

[Corrections]
• p.253 l.11 Kronecker's lemma, (8.5) in Chapter 1.

[Comments]
• Theorem 4.1
We'll often use the result as follows :
$$E(X_0) \leqslant E(X_{N\wedge k}) \leqslant E(X_k)$$
We find that if $\{X_n\}_n$ is martingale, then $E(X_0)=E(X_{N\wedge k})=E(X_k)$

Section 4.5 Uniform Integrability, Convergence in L^1

[Comments]
• Theorem 5.7
Why " $\int_A X\,dP = \int_A Y_\infty\,dP$ (forall $A \in \mathcal{F}_n$) $\Longrightarrow Y_\infty = E(X|\mathcal{F}_\infty)$ " ?

Section 4.6 Backwards Martingales

[Corrections]
• p.268 l.7 (6.6) occurs

[Outlines]
• Steps of the proof of theorem 6.6

[Details]
• Theorem 6.8.
Noting that $E(1_{(X_1=1)}|\mathcal{E}),\cdots,E(1_{(X_n=0)}|\mathcal{E})$ are i.i.d.,
$$\begin{aligned}
& P(X_1=\cdots=X_k=1, X_{k+1}=\cdots=X_n=0) \\
= \; & E(1_{(X_1=\cdots=X_k=1, X_{k+1}=\cdots=X_n=0)}) \\
= \; & E(E(1_{(X_1=\cdots=X_k=1, X_{k+1}=\cdots=X_n=0)}|\mathcal{E})) \\
= \; & E[E(1_{(X_1=1)}|\mathcal{E})\cdots E(1_{(X_k=1)}|\mathcal{E})E(1_{(X_{k+1}=0)}|\mathcal{E})\cdots E(1_{(X_n=0)}|\mathcal{E})] \\
= \; & E[P(X_1=1|\mathcal{E})\cdots P(X_k=1|\mathcal{E})P(X_{k+1}=0|\mathcal{E})\cdots P(X_n=0|\mathcal{E})] \\
= \; & \int_0^1 \theta^k(1-\theta)^{n-k}\,dF(\theta)
\end{aligned}$$

Section 4.7 Optional Stopping Theorems

[Corrections]
• p.272 l.9 $P(\inf_n S_n \leqslant a)$
• p.272 l.30 a is an integer
• p.273 l.10 Use (7.5)

Chapter 5. Markov Chains

Section 5.1 Definitions and Examples

Section 5.2 Extensions of the Markov Property

• A comment for the Markov property and the strong Markov property
These two properties are applied to Brownian Motion $(B_t)_{t\geqslant 0}$.

Let $(X_t)_{t\geqslant 0}$ be a (continuous time) Markov process (i.e. $P(X_s\in B|X_s)=P(X_t\in B|\mathcal{F}_s)$ for $t>s$ and for every $B\in\mathfrak{B}(\mathbb{R})$, where $\mathcal{F}_t:=\sigma(X_u;u<t)$). Then, the two properties hold just as discrete time Markov chains.

[the Markov property]

$$E(Y\circ\theta_t|\mathcal{F}_t)=E_{B_t}(Y)$$

[the strong Markov property]
We define $\mathcal{C}:=\{\rightarrow;\mathbb{R}_{\geqslant 0}\ni t\rightarrow\omega(t)\in C[0,\infty)\}$. Let Y_S be a bounded $\mathcal{C}$-measureable function and let S be a stopping time. Then

$$E(Y_S\circ\theta_S|\mathcal{F}_S)=E_{X_S}(Y_S)\text{ on }\{S<\infty\}$$

Section 5.3 Recurrence and Transience

• Comments
[Decomposition theorem (Thm 3.6)]
Note that C_x is a equivarent class for $x\in R$. Since if we define $x\sim y\Leftrightarrow\rho_{xy}>0$, $\sim$ is a equivarence relation.

(i) $\rho_{xx}>0$ is trivial by the definition of R, (ii) if $\rho_{xy}>0$, then thm 3.4 implies $\rho_{yx}=1>0$, and (iii) if $\rho_{xy},\rho_{yz}>0$, then exercise 3.4 implies $\rho_{xz}\geqslant\rho_{xy}\rho_{yz}>0$.

Section 5.4 Stationary Measure

• Correction
p.306 l.19 Exercise 3.10

• Comments
[The check of that X_{n+1} is Poisson with mean $\mu p+\lambda$ in example 4.7]

Section 5.5 Asymptotic Behavior

• Correction
p.306 l.19 Exercise 3.10
p.318 l.25

$$Y_n=\begin{cases}Y_{n-1}+Z_n & \text{if }|Z_n|>M\\ Y_{n-1}+W_n & \text{if }|Z_n|\leqslant M\end{cases}$$

Section 5.6 General State Space

• Correction
p.330 l.15 Exercise 6.13
p.330 l.23 $z=0$

• The check for $\bar{p}$ is a trans. proba. on $\bar{S}$

• Detail of lemma 6.3

• Detail of example 6.6 Poisson arrivals

Chapter 6. Ergodic Theorems

Section 6.1 Definitions and Examples

Section 6.2 Birkhoff's Ergodic Theorem

• Correction
p.342 l.12

$$\frac{1}{n}\sum_{m=0}^{n-1} 1_{A_k}(\varphi^m(0)) \to \log_{10}\left(\frac{k+1}{k}\right)$$

• Outlines of theorem 6.1

Step 1. $\frac{1}{n}\sum_{m=0}^{n-1} X(\varphi^m(\omega)) \to E(X|\mathcal{I})$ a.s.

Step 1.1. Preparation
Step 1.2. $F=D$
Step 1.3. $P(D)=0$
Step 1.4. Conclusion

Step 2. $\frac{1}{n}\sum_{m=0}^{n-1} X(\varphi^m(\omega)) \to E(X|\mathcal{I})$ in L^1

Step 2.1. Preparation
Step 2.2. Obsevation about X_M' as $M \to \infty$
Step 2.3. Estimation of X_M''
Step 2.4. Conclusion

• Outlines of Lemma 6.2
Step 1. $X(\omega) = S_j(\omega) - M_k(\omega)$ for $1 \leqslant j \leqslant k+1$
Step 2. $E(X(\omega); M_k > 0) \geqslant 0$ with a little calculation.

Section 6.3 Recurrence

• Outlines of Thm 3.1
Step 1. $R_n \geqslant \sum_{1 \leqslant m \leqslant n} 1_A(\varphi^m(\omega))$, hence $\liminf_{n\to\infty} R_n/n \geqslant E(1_A|\mathcal{I})$ a.s.
Step 2. $R_n \geqslant k + \sum_{1 \leqslant m \leqslant n-k} 1_{A_k}(\varphi^m(\omega))$, hence $\limsup_{n\to\infty} R_n/n \leqslant E(1_A|\mathcal{I})$ a.s.
, where $A_k := \{S_1 \neq 0, \cdots, S_k \neq 0\}$
Step 3. Combining the results of Step 1 and Step 2.

• Outlines of Thm 3.3
Step 1. $t_1, t_2, \cdots$ is stationary under $P(\cdot | X_0 \in A)$.
Step 1.1. $P(\bigcup_{k \geqslant 1} C_k) = 1$
Step 1.2 Show that

$$P(t_1=m, t_2=n | X_0 \in A) = P(t_2=m, t_3=n | X_0 \in A) \text{ for all } m, n \in \mathbb{N}_0$$

Step 2. Show that

$$E(T_1 | X_0 \in A) = \frac{1}{P(X_0 \in A)}$$

Section 6.4 Mixing

• Correction
p.347 l.18 It should be $C \equiv \{\omega\colon \varphi^n(\omega) \in A\ i.o.\}$ so that all of ''$P(B)$'' in the same paragraph should be ''$P(C)$''.
p.347 l.18 Not (b) but (ii).

Section 6.5 Entropy

• Correction
p.355 l.8 $=$ should be $\equiv$ or $\stackrel{\Delta}{=}$.
p.356 l.1 $\sum \exp(-\varepsilon n) < \infty$ or $\sum e^{-\varepsilon n} < \infty$.
p.357 l.17 Not ''log a'' but ''$\log a$''.
p.357 l.24 ''$a e^{-\lambda}$'' should be $(\sharp A_j) e^{-\lambda}$
p.357 l.28 (the last line)

$$-\frac{1}{n} \log p(\omega_0, \cdots, \omega_{n-1})$$

• Error of estimation in the proof of thm 5.2
$P(A_j) \leqslant (\sharp A_j) e^{-\lambda}$ is right, since we have to estimate as $P(\sup_n g_n > \lambda) < (\text{const.}) e^{-\lambda}$.
[Detail of estimation of $E(\sup_n g_n)$]
As we know, since $\{\sup_j g_j > \lambda\} \subset \cup_j A_j$ and A_j are disjoint by their def. ,

$$P(\sup_j g_j > \lambda) \leqslant P(\coprod_j A_j) = \sum_j P(A_j) \leqslant \sum_j (\sharp A_j) e^{-\lambda} \leqslant e^{-\lambda} (\sharp S) = a e^{-\lambda}$$

Lem. 5.7 in Chap. 1 implies,

$$E(\sup_n g_n) = \int_0^\infty P(\sup_n g_n > \lambda) d\lambda = a \int_0^\infty e^{-\lambda} d\lambda = a < \infty$$

Section 6.6 A subadditive Ergodic Theorem

• Correction
p.358 l.21 ergodic theorem (2.1)
p.359 l.10 ''$k < m$'' should be ''$0 < k < m$''
p.359 l.29 $L_{0,m} + L_{m,n} \leqslant L_{0,n}$
p.362 l.20

$$(6.6) \quad \underline{Y} \equiv \liminf_{n \to \infty} Y_{0,n}/n \leqslant -\varepsilon$$

p.364 l.8

$$E\left|\frac{X_{0,m} + \cdots + X_{(k-1)m,km}}{km} - \Gamma_m\right| \to 0$$

• Comments
1. The assumption (i) in the subadditive ergodic theorem is called ''subadditiveness''.

• Outline of the proof of the subadditive ergodic theorem.
I believe that I shouldn't forget to make strategy with many small steps when I follow a very long proof.

Step 1. Show the property (a).
♠$_1$ (i) implies $X_{0,n}^{+} \leqslant X_{0,m}^{+} + X_{m,n}^{+}$.
♠$_2$ $|x| = x^{+} + x^{-} = 2x^{+} - x$
♠$_3$ $X_{0,n}^{+} \leqslant X_{0,1}^{+} + \cdots + X_{n-1,n}^{+}$
♠$_4$ $E|X_{0,n}| \leqslant (2E(X_{0,1}^{+}) - \gamma_0) n < \infty$
♠$_5$ $\liminf_{n\to\infty} a_n/n \geqslant \gamma (\equiv \inf_{m\geqslant 1} a_m/m)$
♠$_6$ $\limsup_{n\to\infty} a_n/n \leqslant a_m/m$ for all $m \geqslant 1$
♠$_7$ Show (6.3) :

$$\frac{a_n}{n} \to \inf_{m\geqslant 1} \frac{a_m}{m} \equiv \gamma$$

Step 2. Show $E(\overline{X}) \leqslant \gamma$ where $\overline{X} := \limsup_{n\to\infty} X_{0,n}/n$.
♠$_1$ Prove that

$$\frac{X_{0,m} + \cdots + X_{(k-1)m,km}}{k} \to A_m \text{ a.s. and in } L^1$$

by the ergodic theorem.
♠$_2$ $E(A_m) = E(E(X_{0,m}|\mathcal{I}_m)) = E(X_{0,m})$
♠$_3$ If we fix $l \geqslant 0$ and we let $\varepsilon > 0$, then we can claim that

$$\sum_{k\geqslant 1} P(X_{km,km+l} > (km+l)\varepsilon) \leqslant \sum_{k\geqslant 1} P(X_{0,l} > k\varepsilon) \leqslant \frac{l}{\varepsilon} E(X_{0,1}^{+}) < \infty$$

♠$_4$ Prove that

$$\frac{X_{km,km+l}}{n} \to 0 \text{ a.s.}$$

by reduction to the absurd.
♠$_5$ $\overline{X} \leqslant A_m/m$
♠$_6$ $E(\overline{X}) \leqslant \gamma$
♠$_7$ In particular, if all seq.'s in (ii) are ergodic, then $\overline{X} \leqslant \gamma$ since $E(X_{0,m}|\mathcal{I}_m) = E(X_{0,m}) = A_m$.

Step 3. Show $E(\underline{X}) \geqslant \gamma$ where $\underline{X} := \liminf_{n\to\infty} X_{0,n}/n$, (b) except L^1 convergence and (c).
♠$_1$ Prove $\underline{X} = \underline{X}_m$ where $\underline{X}_m := \liminf_{n\to\infty} X_{m,m+n}/n$ by some tricks.
♠$_2$ Prove (6.6) :

$$\underline{Y} \leqslant -\varepsilon \text{ where } \underline{Y} := \liminf_{n\to\infty} Y_{0,n}/n$$

♠$_3$ Show that for all $\varepsilon > 0$, there are some $N(\varepsilon)$ s.t.

$$E(Y_{m,m+1}\,; T_m > N) = E(Y_{0,1}\,; T_0 > N) \leqslant \varepsilon$$

, where $T_m := \min\{n \geqslant 1; Y_{m,m+n} \leqslant 0\}$.
♠$_4$ Show that

$$Y_{0,n} \leqslant \sum_{m=0}^{n-1} \xi_m + \sum_{j=1}^{N} |Y_{n-j,n-j+1}|$$

, where

$$S_m := \begin{cases} T_m & \text{on } \{T_m \leqslant N\} \\ 1 & \text{on } \{T_m > N\} \end{cases} , \quad \xi_m := \begin{cases} 0 & \text{on } \{T_m \leqslant N\} \\ Y_{m,m+1} & \text{on } \{T_m > N\} \end{cases} \quad \text{and}$$

$$R_n := \begin{cases} 0 & \text{if } n=0 \\ R_{n-1}+S_{R_{n-1}} & \text{if } n>0 \end{cases}$$

♠$_5$ $\gamma \leqslant E(\underline{X})$

♠$_6$ In particular, if all seq.'s in (ii) are ergodic, then $X=\gamma$.

Step 4. Show $\lim_{n\to\infty} X_{0,n}/n =^{\exists} X$ in L^1 .

Let Γ_m be A_m/m and let Γ be $\inf_{m\geqslant 1}\Gamma_m$.

♠$_1$ Show that

$$E\left|\frac{X_{0,n}}{n}-\Gamma\right| = 2E\left(\frac{X_{0,n}}{n}-\Gamma\right)^+ - E\left(\frac{X_{0,n}}{n}-\Gamma\right) \leqslant 2E\left(\frac{X_{0,n}}{n}-\Gamma\right)^+$$

♠$_2$ Show that

$$E\left(\frac{X_{0,n}}{n}-\Gamma\right)^+ \leqslant E\left(\frac{X_{0,n}}{n}-\Gamma_m\right)^+ + E(\Gamma_m-\Gamma)^+$$

♠$_3$ $E(\Gamma_m-\Gamma)^+<\varepsilon$ for sufficiently large m .

♠$_4$ $E(X_{0,n}/n-\Gamma_m)^+\to 0$ as $m\to\infty$ after $n\to\infty$ (or $k\to\infty$).

♠$_5$ Conclude by combining ♠$_1$, ♠$_2$, ♠$_3$ and ♠$_4$

Section 6.7 Applications

• Correction

p.365 l.1 $E(\log A_m(1,1))^- < \infty$

p.365 l.3 $E(X_{0,n}) \geqslant -\gamma_0 n$ $(\exists\, \gamma_0<\infty)$

p.366 l.2 the eigenvalues of A.

p.368 l.12

$$\gamma = \inf\{a : \log\mu - c(a) < 0\}$$

p.369 l.7 $E(t(0,u))<\infty$

p.369 l.10 Not ν but nu.

p.369 l.27 to apply (6.1)

p.370 l.24 $p_2=1$

• Comments

[The check of that $X_{m,n}$ satisfies the assumptions (i) -- (iv) in example 7.1]

That $\{A_i\}_{i\geqslant 1}$ is stationary is defined as ;

We regard each A_i as a $\mathbb{R}^{k^2}$ vector and we define that ''$\{A_i\}_{i\geqslant 1}$ is stationary'' is

$$(A_1, A_2, \cdots) =_d (A_m, A_{m+1}, \cdots) \text{ for all } m\geqslant 1$$

(i) : Since $\alpha_{0,m}(1,1)\alpha_{m,n}(1,1) \leqslant \alpha_{0,n}(1,1)$,

$$X_{0,m}+X_{m,n} = -\log\alpha_{0,m}(1,1) - \log\alpha_{m,n}(1,1) \geqslant -\log\alpha_{0,n}(1,1) = X_{0,n}$$

(ii)(iii) : Note that

$$\begin{aligned} \alpha_{m,m+k}(1,1) &= \textstyle\sum_{i_1,\cdots,i_{k-1}} A_{m+1}(1,i_1)\cdots A_{m+k}(i_{k-1},1) \\ &=_d \textstyle\sum_{i_1,\cdots,i_{k-1}} A_1(1,i_1)\cdots A_k(i_{k-1},1) \\ &= \alpha_{0,k}(1,1) \end{aligned}$$

So,

$$X_{m,m+k} = -\log\alpha_{m,m+k}(1,1) =_d -\log\alpha_{0,k}(1,1) = X_{0,k}$$

(iv) : If $E(\log A_m(1,1))^-<\infty$, then $E(X_{0,1}^+) = E(\log A_1(1,1))^- =_d E(\log A_m(1,1))^- < \infty$

And if $E[\log(\sup_{i,j} A_m(i,j))] (=: \gamma_0) < \infty$ (recall that $A_m =_d A_1$ for all m), then

$$
\begin{aligned}
E(X_{0,n}) &= -E(\log \alpha_{0,n}(1,1)) \\
&= -E[\log \textstyle\sum_{i_1,\cdots,i_{n-1}} A_1(1,i_1)\cdots A_n(i_{n-1},1)] \\
&\geqslant -E[\log\{\sup_{i,j} A_1(i,j) \textstyle\sum_{i_1,\cdots,i_{n-1}} A_2(i_1,i_2)\cdots A_n(i_{n-1},1)\}] \\
&\geqslant \cdots \\
&\geqslant -E[\log\{\textstyle\prod_{1\leqslant l\leqslant n} \sup_{i,j} A_l(i,j)\}] \\
&= -\textstyle\sum_{1\leqslant l\leqslant n} E[\log\{\sup_{i,j} A_l(i,j)\}] \\
&= -\gamma_0 n
\end{aligned}
$$

[The check of that $Y_{m,n}$ satisfies the assumptions (i) -- (iv) in example 7.1]
(i) :

$$
\begin{aligned}
Y_{0,n} &= \log \beta_{0,n} \\
&= \log \|A_1 \cdots A_n\| \\
&\leqslant \log \|A_1 \cdots A_m\| + \log \|A_{m+1} \cdots A_n\| \\
&= \log \beta_{0,m} + \log \beta_{m,n} \\
&= Y_{0,m} + Y_{m,n}
\end{aligned}
$$

(ii)(iii) : Since $\beta_{0,k} = \|A_1 \cdots A_k\| =_d \|A_{m+1} \cdots A_{m+k}\| = \beta_{m,m+k}$, $Y_{0,k} = \log \beta_{0,k} =_d \log \beta_{m,m+k} = Y_{m,m+k}$.
(iv) : Since the definition of operator norms implies that $\|A_1\| < \infty$, $E(Y_{0,1}^+) = E(\log \beta_{0,1})^+ = E(\log \|A_1\|)^+ < \infty$.
While, if we let $\gamma_0 := \min_{1\leqslant l\leqslant n} E(\log \inf_{i,j} |A_l(i,j)|)$,

$$
\begin{aligned}
E(Y_{0,n}) &= E(\log \|A_1 \cdots A_n\|) \\
&= E(\log \max_i \textstyle\sum_j |(A_1 \cdots A_n)(i,j)|) \\
&\geqslant E(\log |(A_1 \cdots A_n)(i,j)|) \\
&\geqslant E(\log \left| \textstyle\sum_{i_1,\cdots,i_{n-1}} (A_1(i,i_1)\cdots A_n(i_{n-1},j) \right|) \\
&\geqslant E(\log \inf_{i,j} |A_1(i,j)| \left| \textstyle\sum_{i_1,\cdots,i_{n-1}} (A_2(i_1,i_2)\cdots A_n(i_{n-1},j) \right|) \\
&\geqslant \cdots \\
&\geqslant E(\log \textstyle\prod_{1\leqslant l\leqslant n} \inf_{i,j} |A_l(i,j)|) \\
&\geqslant \textstyle\sum_{1\leqslant l\leqslant n} E(\log \inf_{i,j} |A_l(i,j)|) \\
&\geqslant \gamma_0 n
\end{aligned}
$$

• Details
[The check of that $\|A\|$ is a operator norm]
At first, we should show that :

$$
\max_i \sum_j |A(i,j)| = \max_{\|x\|_1=1} \|xA_1\|_1
$$

$$
\begin{aligned}
\max_{\|x\|_1=1} \|xA\|_1 &= \max_{\|x\|_1=1} \sum_j |(xA)_j| \\
&= \max_{\|x\|_1=1} \sum_j |\sum_i x_i A(i,j)| \quad \text{for } \|x\|_1 = 1 \\
&\geqslant \max_{1\leqslant i\leqslant k} \sum_j |A(i,j)|
\end{aligned}
$$

Note that we take x such as $x_i = 1$ in the last inequality.

While, we can claim that , for $x; \|x\|_1 = 1$

$$
\begin{aligned}
\|xA\|_1 &= \sum_j \left| \sum_i x_i A(i,j) \right| \\
&\leqslant \sum_j \sum_i |x_i||A(i,j)| \\
&= \sum_i \sum_j |x_i||A(i,j)| \\
&= \sum_i |x_i| \sum_j |A(i,j)| \\
&\leqslant \max_i \sum_j |A(i,j)|
\end{aligned}
$$

If we take x such as $\|x\|_1 = 1$ and $\|xA\|_1$ is maximal ,

$$
\max_{\|x\|_1=1} \|xA\|_1 \leqslant \max_i \sum_j |A(i,j)|
$$

So, we can claim that $\|\cdot\|$ is a operator norm by the definition of operator norms in functional analysis. Then, we can claim that $\|AB\| \leqslant \|A\|\|B\|$ etc. by basic properties of operator norms.

[$\lim_{n\to\infty} (1/n) \log \|A_1 \cdots A_n\| = -X$]
It suffices to show that $\lim_{n\to\infty} (1/n) \max_i \log \sum_j \alpha_{0,n}(i,j) = -X$. Since $(1/n) \log \alpha_{0,n}(i,j) \to -X$ a.s. ,

$$
\begin{aligned}
\max_i \frac{\log \sum_j \alpha_{0,n}(i,j)}{n} &\leqslant \max_i \max_j \frac{\log k\alpha_{0,n}(i,j)}{n} \\
&= \max_i \max_j \frac{\log k + \log \alpha_{0,n}(i,j)}{n} \\
&= \frac{\log k}{n} + \max_i \max_j \frac{\log \alpha_{0,n}(i,j)}{n} \\
&\to -X
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
\max_i \frac{\log \sum_j \alpha_{0,n}(i,j)}{n} &\geqslant \max_i \min_j \frac{\log k\alpha_{0,n}(i,j)}{n} \\
&= \max_i \min_j \frac{\log k + \log \alpha_{0,n}(i,j)}{n} \\
&= \frac{\log k}{n} + \max_i \min_j \frac{\log \alpha_{0,n}(i,j)}{n} \\
&\to -X
\end{aligned}
$$

[The first equality at p.368]
Note that if we let Z_n be the number of individuals in the generation n , then

$Z_n\ (an) = Z_n\ 1_{(T_n \leqslant an)}$. Since $1_{(T_n \leqslant an)}$ is independent of Z_n (so that $E(1_{(T_n \leqslant an)}\ |Z_n) = E(1_{(T_n \leqslant an)})$) ,

$$\begin{aligned} E(Z_n(an)) &= E[E(Z_n(an)|Z_n)] \\ &= E[E(Z_n 1_{(T_n \leqslant an)}|Z_n)] \\ &= E[Z_n E(1_{(T_n \leqslant an)}|Z_n)] \\ &= E[Z_n E(1_{(T_n \leqslant an)})] \\ &= E[Z_n P(T_n \leqslant an)] \\ &= E(Z_n)P(T_n \leqslant an) \\ &= \mu^n P(T_n \leqslant an) \end{aligned}$$

Recall that $\{Z_n/\mu^n\}$ is a martingale so that $E(Z_n)/\mu^n = E(Z_0) = 1$.

Chapter 7. Brownian Motion

Section 7.1 Definitions and Construction

• Correction

• Comments

[Def. of Gaussian process]

''$(X_t)_{t\geqslant 0}$ is Gaussian prosess'' if and only if for every finite set of indices $0 \leqslant t_1 < \cdots < t_k$, there are $\sigma_{lj} > 0$ and $\mu_j \in \mathbb{R}$ such that

$$E\left(\exp\left(i\sum_{1\leqslant l\leqslant k} t_l X_{t_l}\right)\right) = \exp\left(-\frac{1}{2}\sum_{l,j}\sigma_{lj}t_l t_j + i\sum_{l}\mu_l t_l\right)$$

And in the case of def. (a$'$) -- (c$'$) , $\mu_l = E(B_{t_l}) = 0$ and $\sigma_{lj} = E(B_{t_l}B_{t_j}) = t_l \wedge t_j$ so that

$$E\left(\exp\left(i\sum_{1\leqslant l\leqslant k} t_l B_{t_l}\right)\right) = \exp\left(-\frac{1}{2}\sum_{l,j} t_l t_j (t_l \wedge t_j)\right)$$

[The last equation in p.373]

$\int p_{t_j - t_{j-1}}(x, dy) p_{t_{j+1}-t_j}(y, z) = p_{t_{j+1}-t_{j-1}}(x, z)$ is trivial by Chapman-Kolmogorov's equation in continuous time. (Then we have to show that the Brownian path is Markov chain for any subseqence $\{t_j\}_{j\geqslant 0} : t_0 < t_1 < \cdots < t_n < \cdots$. But it is also trivial by the extension from $\mu_{x,s_1,\cdots,s_{n-1}}$ to $\mu_{x,t_1,\cdots,t_{n-1},t_n}$)

[Steps of construction of Brownian path]

Step 1. Enumerate elements in $\mathbb{Q}_2$ as $q_1, q_2, \cdots$

Step 2. Construct a probability measure ν_x on $(\Omega_q, \mathcal{F}_q)$ s.t. $\nu_x(\omega; \omega_0 = x) = 1$

Step 3. $(*)$ in Thm 1.3 holds for $t_i \in \mathbb{Q}_2$

[Outlines of Thm 1.4]

Step 1. We consider a prob. meas. ν_x on a m'ble space $(C, \mathcal{C})$.

Step 2. If we let $P_x = \nu_x \circ \psi^{-1}$, then $B_t(\omega) = \omega_t$ has the finite dim. dist. for $t \in \mathbb{Q}_2$.

Step 3. Construction of $(B_t)_{t\geqslant 0}$.

Step 4. We can suppose $B_0 = 0$ by thm 1.1 and thm 1.2 and prove the result for $T = 1$ without loss of generality.

[We can compute $E|B_t|^m$]

$$\begin{aligned} E|B_t|^m &= \int |B_t|^m dP \\ &= \int t^{\frac{m}{2}} |B_1|^m dP \\ &= t^{\frac{m}{2}} \int_{-\infty}^{\infty} |x|^m P(B_1 \in dx) \\ &= t^{\frac{m}{2}} (2\pi)^{-\frac{1}{2}} 2 \int_0^{\infty} x^m e^{-\frac{x^2}{2}} dx \\ &= t^{\frac{m}{2}} \pi^{-\frac{1}{2}} 2^{\frac{1}{2}} \int_0^{\infty} 2^{\frac{m-1}{2}} y^{\frac{m-1}{2}} e^{-y} dy \\ &= t^{\frac{m}{2}} \pi^{-\frac{1}{2}} 2^{\frac{m}{2}} \Gamma\left(\frac{m+1}{2}\right) \\ &= \begin{cases} t^{\frac{m}{2}} \pi^{-\frac{1}{2}} 2^{\frac{m}{2}} \left(\frac{m-1}{2}\right)! & \text{if } m \text{ is odd} \\ t^{\frac{m}{2}} 2^{\frac{m-2}{2}} (m-1)!! & \text{if } m \text{ is even} \end{cases} \end{aligned}$$

[Outlines of Thm 1.5]
Step 1. $P(G_n^c) = K \cdot 2^{-n\{(1-\eta)(1+\alpha-\beta\gamma)-(1+\eta)\}}$
, where $G_n := \{|X_{j\cdot 2^{-n}} - X_{i\cdot 2^{-n}}| \leqslant \{(j-i)2^{-n}\}^{\gamma}$ for all $(i,j) \in I_n\}$ and
$I_n := \{(i,j); 0 \leqslant i \leqslant j \leqslant 2^n, 0 < j-i \leqslant 2^{n\eta}\}$
Step 2. = Lemma 1.6
Step 3. Step 2. completes the proof of thm 1.5.
Step 4. Thm 1.5 implies thm 1.4.

[Definition (Holder continuous)]
''$(B_t)_{t\geqslant 0}$ is Holder continuous with exponent $\beta \in (0,1]$'' if and only if

$$|B_t - B_s| \leqslant {}^{\exists}C|t-s|^{\beta} \text{ for } t,s \geqslant 0$$

[Definition (Lipschitz continuous)]
''$(B_t)_{t\geqslant 0}$ is Lipschitz continuous'' if and only if

$$|B_t - B_s| \leqslant {}^{\exists}C|t-s| \text{ for } t,s \geqslant 0$$

[Outlines of Thm 1.8]
Step 1. $A_n \subset B_n$
Step 2. Evaluating $P(A_n)$

• Details
[The proof of Thm 1.2 in the case of general n]
It suffices to show the result in the case of $n = 2$. We can show the result in general n as follows. First, we define $D \subset \mathbb{R}^2$ for all $\alpha, \beta \in \mathbb{R}$:

$$D:=\{(x,y)\in\mathbb{R}^2; x\leqslant\alpha, y-x\leqslant\beta\}$$

Then, it suffices to show that $P((B_{st_1}, B_{st_2})\in D)=P((\sqrt{s}B_{t_1}, \sqrt{s}\,B_{t_2})\in D)$ since $\mathfrak{B}(\mathbb{R}^2)=\mathfrak{B}(D)$. Using the result in $n=1$ and that B_{st_1} and $B_{st_2}-B_{st_1}$ are independent,

$$\begin{aligned}
P((B_{st_1},B_{st_2})\in D) &= P(B_{st_1}\leqslant\alpha, B_{st_2}-B_{st_1}\leqslant\beta)\\
&= P(B_{st_1}\leqslant\alpha)P(B_{st_2}-B_{st_1}\leqslant\beta)\\
&= P(\sqrt{s}B_{t_1}\leqslant\alpha)P(\sqrt{s}\,(B_{t_2}-B_{t_1})\leqslant\beta)\\
&= P(\sqrt{s}B_{t_1}\leqslant\alpha, \sqrt{s}\,B_{t_2}-\sqrt{s}\,B_{t_1})\leqslant\beta)\\
&= P((\sqrt{s}B_{t_1}, \sqrt{s}B_{t_2})\in D)
\end{aligned}$$

[The reason why ‘‘(a$'$) , (b$'$) and (c$'$) $\Longrightarrow$(a)’’]
It suffices to show that the covariance of $(B_{t_1}, B_{t_2}-B_{t_1}, \cdots, B_{t_n}-B_{t_{n-1}})=0$. For example,

$$E(\{B_{t_1}-E(B_{t_1})\}\{B_{t_2}-B_{t_1}-E(B_{t_2}-B_{t_1})\})=E(B_{t_1}B_{t_2}-B_{t_1}^2)=t_1-t_1=0$$

We get the same result : $E[(B_{t_i}-B_{t_{i-1}})(B_{t_j}-B_{t_{j-1}})]=0$ for all $1\leqslant i,j\leqslant n$ (, where $B_{t_0}=B_0=0$).

[Thm 1.4]
It will be better to rewrite with some notations. Of course, we use notations in textbook as they are.
First, we observe Ω_q more strictly ;

$$\Omega_q^0:=\{\omega\in\Omega_q;\ \text{uniformly continuous on } [0,T]\cap\mathbb{Q}_2\}$$

So, we can rewrite statements of Thm 1.4 :

Thm 1.4
Let $T<\infty$ and $x\in\mathbb{R}$. Then, $\nu_x(\Omega_q^0)=1$ (i.e. $\nu_x($ ω is uniformly continuous on $[0,T]\cap\mathbb{Q}_2$) = 1).

Secondly, we can define $\psi:\Omega_q^0\to C$ as follows ;

$$\psi(\omega)(t)=\begin{cases}\omega(t) & \text{if } t\in[0,T]\cap\mathbb{Q}_2\\ \lim\limits_{t_j\to t}\omega(t_j) & \text{if } t\in[0,T]\cap\mathbb{Q}_2^c\end{cases}$$

, where $\{t_j\}_{j\geqslant1}$ is a sequence (in $[0,T]\cap\mathbb{Q}_2$) which converges to t.

Then, it will be more clear that ψ is measurable ;

If $t\in[0,T]\cap\mathbb{Q}_2$, then $\{\omega;\psi(\omega)(t)\leqslant a\}=\{\omega;\omega(t)\leqslant a\}\in\mathcal{F}_q$. If $t\in[0,T]\cap\mathbb{Q}_2^c$, then $\{\omega;\psi(\omega)(t)\leqslant a\}=\lim_{j\to\infty}\{\omega;\omega(t_j)\leqslant a\}=\{\omega;\omega(t)\leqslant a\}\in\mathcal{F}_q$. Then, ψ is $\mathcal{F}_q$-measurable.

Section 7.2 Markov Property, Blumenthal's 0-1 law

• Correction

• Outlines

[Theorem 2.1 (= the Markov property of BM)]

The proof is very similar to one of the Markov property of discrete-time Markov chains.

Step 1. Prove that

$$E_x(Y\circ\theta_s;A)=E_x(E_{X_s}(Y);A) \text{ for all } A\in\mathcal{F}_s^+ \qquad (*)$$

Step 1.1. Show $(**)$

Step 1.2. $\varphi(x_h,h)\to\varphi(x,0)$ as $h\to 0$

Step 1.3. Show $(*)$

Step 2. General cases

• Comments

• Details

[The proof of theorem 2.7]

Suppose $P(T_0>0)>0$. Then there must be $n\geqslant 1$ such that $P(T_0\geqslant n^{-1})>0$. Since

$$P(T_0>0)=P\left(\bigcup_{n\geqslant 1} T_0\geqslant n^{-1}\right)\leqslant\sum_{n\geqslant 1} P(T_0\geqslant n^{-1})$$

and if $P(T_0\geqslant n^{-1})=0$ for all $n\geqslant 1$ then $P(T_0>0)=0$ i.e. $P(T_0=0)=1$. We take such a n $(\geqslant 1)$.

Then it must be either $B_t>0$ for all $0<t<n^{-1}$ or $B_t<0$ for all $0<t<n^{-1}$ on $\{T_0\geqslant n^{-1}\}$ $(*)$ since any Brownian path is continuous a.s.

Theorem 2.6 implies that $B_t>0$ for all $0<t<n^{-1}$ on $\{T_0\geqslant n^{-1}\}$. But the symmetricity of BM also implies that $B_t<0$ for all $0<t<n^{-1}$ on $\{T_0\geqslant n^{-1}\}$, since if we define $\tau^{'}:=\inf\{t\geqslant 0;B_t>0\}$ then we obtain $P_0(\tau^{'}=0)=1$ just as theorem 2.6.

Thus, the last two results contradict to $(*)$. That means $P(T_0=0)=1$

□

Section 7.3 Stopping Times, Strong Markov Property

• Correction

p.388 l.27 $\rho(x,K):=\inf\{|x-y|;y\in K\}$

• Outlines of Thm 3.7 (= the strong Markov property of BM)

The proof is very similar to one of the strong Markov property of discrete-time MC.

Step 1. Prove the strong Markov property under these assumption :

$$\{t_n\}:t_n\uparrow\infty \text{ and } P_x(S<\infty)=\sum_{n\geqslant 1}P_x(S=t_n)$$

Step 2. Extend the result of Step 1 to the general cases.

Step 2.1. The cases of $Y_s(\omega)=f_0(s)\prod_{1\leqslant i\leqslant n}f_i(\omega_{t_i})$.

Step 2.2. The cases of bounded $Y_S(\omega)$ on $\{S<\infty\}$.
We use Monotone class theorem (Thm 1.5 in Chap 5) here.

• Details
[Theorem 3.7]
The reason why $x \mapsto \int dy p_t(x,y) f(y)$ is continuous when f is bounded and continuous.
Note that

$$p_t(x,dy)=\frac{1}{\sqrt{2\pi t}}e^{-\frac{(y-x)^2}{2t}}dy$$

and we have only to show $x \mapsto \int p_t(x,dy)$ is continuous. The mean value theorem implies that there is a $c \in (0,h)$ and

$$\begin{aligned}\left|\int p_t(x+h,dy)-\int p_t(x,dy)\right| &\leqslant \int |p_t(x+h,dy)-p_t(x,dy)| \\ &\leqslant \frac{1}{\sqrt{2\pi t}}\int |e^{-\frac{(y-x-h)^2}{2t}}-e^{-\frac{(y-x)^2}{2t}}|dy \\ &\leqslant \frac{h}{\sqrt{2\pi t}}\int \left|\frac{y-x-c}{t}\right|e^{-\frac{(y-x-c)^2}{2t}}dy \\ &\leqslant \frac{2h}{\sqrt{2\pi t}}\int_{y\geqslant x+c}\frac{y-x-c}{t}e^{-\frac{(y-x-c)^2}{2t}}dy \\ &\leqslant \sqrt{2}h\end{aligned}$$

Section 7.4 Maxima and Zeros

• What does the author mean in the paragraph beginning ''The scaling relation ...'' ?
• Outlines of Thm 4.10 in Example 4.5 (Lévy's modulus of continuity)
Let $I_{m,n}:=[m/2^n,(m+1)/2^n]$ and $\Delta_{m,n}:=\sup\{|B_t-B_{m/2^n}|: t\in I_{m,n}\}$ for all $m,n\in\mathbb{N}_0$.
Step 1. The reflection principle implies that

$$P_\mu\left(\Delta_{m,n}\geqslant\frac{a}{2^{n/2}}\right)<4\exp\left(-\frac{a^2}{2}\right)$$

Step 2. We estimate $\Delta_{m,n}$ by the Borel Canteli lemma ;

$$\Delta_{m,n}<(bn)^{1/2}\cdot 2^{-n/2} \text{ for large } n$$

, where $\varepsilon>0$ is arbitorary and $b:=2(1+\varepsilon)\log 2$.
Step 3. We estimate $|B_s-B_t|$ by the triangle inequality and complete the proof.

Section 7.5 Martingales

Section 7.6 Donsker's Theorem

• Correction
p.403 l.23 Not ''*a* limit'' but ''a limit''.

p.403 l.26 (the last line) $[j/3,(j+1)/3]$.
p.405 l.11 Though this isn't a fatal error (as example 6.5 isn't example),

$$|x^k-y^k|\leqslant\int_x^y\frac{|z|^{k-1}}{k-1}dz\leqslant\frac{\varepsilon M^{k-1}}{k-1}$$

• Outlines of Skorokhod's representation theorem (Thm 6.1)
As we have proved the strong Markov property, first we prove this result in the simple cases, and then we prove in the general cases.
Step 1. Suppose X concentrates on $\{a,b\}$ $(a<0<b)$.
Step 2. Extend the result of Step 1 to general X -- construct a distribution function F.
Step 2.1. Construct the analogue of Step 1.
Step 2.2. Invent a r.v. (U,V) on $\mathbb{R}^2$ which satisfies formula (6.2) .
Step 2.3. Check $E(T_{U,V})=E(X^2)$ and $B_{T(U,V)}=_d X$.

• Outlines of Donsker's representation theorem (Thm 6.7)
Step 0. Preparation of some notations.
Step 1. = Lemma 6.8
Step 1.1. Preparation for the proof of Lem. 6.8.
Step 1.2. Show that forall $\varepsilon>0$, there are some l s.t.

$$P(|B_t-B_s|<\varepsilon \text{ for all } 0\leqslant s\leqslant 1 \text{ and } |t-s|\leqslant 2/l)>1-\varepsilon$$

Step 1.3. Show that there is an integer N_k s.t. if $n\geqslant N_k$, then

$$P\left(\sup_{0\leqslant s\leqslant 1}|T^n_{[ns]}-s|<\frac{2}{l}\right)\geqslant 1-\varepsilon$$

Step 1.4. Show that if $2/l<\varepsilon$ and $n\geqslant N_k$, then

$$P(\|S_{n,(nt)}-B_t\|>2\varepsilon)<2\varepsilon$$

, where $\|S_{n,(n\cdot)}-B(\cdot)\|:=\sup_{0\leqslant v\leqslant 1}|S_{n,(nv)}-B_v|$
Step 2. = Lemma 6.9
Step 3. Conclusion

Section 7.7 CLT's for Dependent Variables

• Correction
p.409 l.6 $\mu_k(S_1,\cdots,S_{k-1};\cdot)$ should be S_k-S_{k-1} .
p.418 l.5 $(a-b)^2\leqslant 2a^2+2b^2$.
p.418 l.10 (7.5) should be (7.6)
p.419 l.4

$$\|E(X_0|\mathcal{F}_{-n})\|_2^2=\sum_x\pi(x)\left(\sum_y p^n(x,y)f(y)\right)^2$$

p.421 l.19 C should be C_1 .
p.422 l.2 C should be C_2 .
p.424 l.25 $\beta(n)=\beta(\mathcal{F}_{-n},\sigma(X_0))$

• Outlines of Thm 7.6
Step 0. Preparation

Step 1. $S_n = T_n + \theta Z_0 - \theta^{n+1} Z_0$ and $T_{(n\cdot)}/\sqrt{n} \Rightarrow \sigma B(\cdot)$
Step 2. Conclusion
Step 2.1. Check $\sigma^2 = E(X_0^2) + 2\sum_{n\geqslant 1} E(X_0 X_n)$
Step 2.2. $\| S_n - T_n \|_2 \precsim n^{-1/2}$

• Three Cs in the proof of lemma 7.7
As you know, C in p.421 l.13 which means a set, C in p.421 l.19 which means a constant and C in p.422 l.2 which means a constant are different from each other.

Section 7.8 Empirical Distributions, Brownian Bridge

• Correction
p.426 l.24 (the last line) It should be $\lambda = 1$ or λ should be omitted.
p.429 l.2 Not ''the proof of (4.6)'' but ''the formula (4.7)''
p.430 (8.11) should be

$$P_0\left(\max_{0\leqslant t\leqslant 1} |B_t^0| < b\right) = 1 + 2\sum_{k\geqslant 1} (-1)^k e^{-2k^2b^2} \text{ for } b > 0$$

Section 7.9 Laws of the Iterated Logarithm

• Correction
p.433 l.9 Not (2.6) but (2.5)
p.433 l.14 $h(t)^2$
p.433 l.15 Not (4.1) but (4.8)
p.434 l.10 Not (9.6) but (9.7)
p.435 l.11 $f(t) = t, -t$
p.435 l.24 ''We have $2n$ instead of $n/2$'' is right.

• Check (9.4)
We use L'Hôpital's rule ;

$$\begin{aligned}\lim_{x\to\infty} \frac{\int_x^\infty \exp(-y^2/2)\,dy}{x^{-1}\exp(-x^2/2)} &= \lim_{x\to\infty} \frac{-\exp(-x^2/2)}{-x^{-2}\exp(-x^2/2) + x^{-1}\cdot(-x)\exp(-x^2/2)} \\ &= \lim_{x\to\infty} \frac{1}{x^{-2}+1} \\ &= 1\end{aligned}$$

• Details of the proof of thm 9.1
[Outline of the prrof]
Step 1. Preparation = Check of (9.2), (9.3) and (9.4).
Step 2. Prove (a) :

$$\limsup_{t\to\infty} \frac{B_t}{\sqrt{2t\log\log t}} \leqslant 1$$

Step 3. Prove the following inequality :

$$\limsup_{t_n\to\infty} \frac{B_{t_{n+1}}}{\sqrt{2t_{n+1}\log\log t_{n+1}}} \geqslant \frac{\alpha-1}{\alpha} - \frac{1}{\alpha^{1/2}}$$

so that

$$\limsup_{t\to\infty} \frac{B_t}{\sqrt{2t\log\log t}} \geqslant 1$$

where $t_n = \alpha^n$ for $\alpha > 1$.

• Comments of Kolmogorov's test (9.6)

Comment 1.

I guess that h is a continuous function.

Comment 2.

I invented the proof of theorem 9.6. But it may contain something wrong.

(i) If $P_0(B_t < h(t)$ for $0 \leqslant t \leqslant \delta) = 1$ for sufficiently small $\delta > 0$ i.e. h is upper class, then $P_0(T_{h(t)} > \delta) = 1 - P_0(T_{h(t)} \leqslant \delta) = 1$ i.e. $P_0(T_{h(t)} \leqslant \delta) = 0$. This implies,

$$\int_0^\delta t^{-3/2} h(t) \exp\left(-\frac{h(t)^2}{2t}\right) dt = 0$$

So,

$$\int_0^1 t^{-3/2} h(t) \exp\left(-\frac{h(t)^2}{2t}\right) dt = \int_\delta^1 t^{-3/2} h(t) \exp\left(-\frac{h(t)^2}{2t}\right) dt \leqslant (1-\delta) K < \infty$$

where

$$K := \max_{\delta \leqslant t \leqslant 1} \left| t^{-3/2} h(t) \exp\left(-\frac{h(t)^2}{2t}\right) \right|$$

(ii) If $P_0(B_t \geqslant h(t)$ for some $(0 \leqslant) t \leqslant \delta) = 0$ for all $\delta > 0$ i.e. h is lower class,

By the reflection principle,

$$2P_0(B_t \geqslant h(t)) = P_0(T_{h(t)} < t) = (2\pi)^{-1/2} \int_0^t t^{-3/2} h(t) \exp\left(-\frac{h(t)^2}{2t}\right) dt$$

If we take two seq.'s $\{m_k\}_{k\geqslant 1}$ and $\{m_k'\}_{k\geqslant 1}$ s.t. $m_1 < m_1' < m_2 < m_2' < \cdots$ and

$$P_0(B_t \geqslant h(t) \text{ for some } t \in (2^{-m_k}, 2^{-m_k'}]) \geqslant \frac{1}{2}$$

This implies

$$\int_{2^{-m_k'}}^{2^{-m_k}} t^{-3/2} h(t) \exp\left(-\frac{h(t)^2}{2t}\right) dt \geqslant \sqrt{2\pi}$$

So,

$$\int_0^1 t^{-3/2} h(t) \exp\left(-\frac{h(t)^2}{2t}\right) dt \geqslant \sum_{k\geqslant 1} \int_{2^{-m_k'}}^{2^{-m_k}} t^{-3/2} h(t) \exp\left(-\frac{h(t)^2}{2t}\right) dt = \sum_{k\geqslant 1} \sqrt{2\pi} = \infty$$

Comment 3. The meaning of ''a little calculus'' in l.18, p.433

Don't forget that what we want is that

$$\int_0^\delta t^{-3/2} h(t) \exp\left(-\frac{h(t)^2}{2t}\right) dt$$

converges or diverges for sufficiently small $\delta > 0$.

Note that

$$\lg_2(t^{-1}) + \frac{3}{2}\lg_3(t^{-1}) + \sum_{m=4}^{n-1} \lg_m(t^{-1}) + (1+\varepsilon)\lg_n(t^{-1}) \sim \lg_2(t^{-1}) \quad (t \searrow 0)$$

We use this fact :

$$\int \frac{1}{s\,\mathrm{lg}_1 s \cdots \mathrm{lg}_{n-1} s}\, ds = \mathrm{lg}_n s + C$$

where C is an arbitrary constant. You'll get if you differentiate w.r.t. s in both of the sides.

Comment 4.
But h which the author suggests is NOT suitable as an example of h in the result of Kolmogorov's test, since $h\uparrow$ and $t^{-1/2}h\uparrow$ as $t\downarrow 0$.

• Details of the proof of theorem 9.7
[The reason why it suffices to show (a)]
Note that

$$\begin{aligned}
\limsup_{n\to\infty} \frac{S_n}{\sqrt{2n\log\log n}} &= \limsup_{t\to\infty} \frac{S_{[t]}}{\sqrt{2[t]\log\log[t]}} \\
&= \limsup_{t\to\infty} \frac{S_{[t]}}{\sqrt{2t\log\log t}} \left(\frac{2t\log\log t}{2[t]\log\log[t]}\right) \\
&= \limsup_{t\to\infty} \frac{S_{[t]}}{\sqrt{2t\log\log t}}
\end{aligned}$$

So, (a) implies that

$$\limsup_{t\to\infty} \frac{S_{[t]}-B_t}{\sqrt{2t\log\log t}} = \limsup_{t\to\infty} \frac{S_{[t]}}{\sqrt{2t\log\log t}} - \limsup_{t\to\infty} \frac{B_t}{\sqrt{2t\log\log t}} = 0$$

and recall that $\limsup_{t\to\infty} B_t/\sqrt{2t\log\log t} = 1$ by thm 9.1.
The last two observation imply that

$$\limsup_{n\to\infty} \frac{S_n}{\sqrt{2n\log\log n}} = 1$$

[Detail of $P(\limsup_{t\to\infty} (S_{[t]}-B_t)/\sqrt{t\log\log t} \geqslant \sqrt{3\delta}) < (4/\sqrt{6\pi})\varepsilon$]
Note that

$$\begin{aligned}
&P\left(\limsup_{t\to\infty} \frac{S_{[t]}-B_t}{\sqrt{t\log\log t}} \geqslant \sqrt{3\delta}\right) \\
&\leqslant P\left(\bigcup_{k>l} \max_{t_k \leqslant t \leqslant t_{k+1}} \frac{M_t}{\sqrt{t\log\log t}} \geqslant \sqrt{3\delta}\right) \\
&\leqslant \sum_{k>l} P\left(\max_{t_k \leqslant t \leqslant t_{k+1}} \frac{M_t}{\sqrt{t\log\log t}} \geqslant \sqrt{3\delta}\right) \\
&\leqslant \sum_{k>l} P\left(\max_{t_k \leqslant t \leqslant t_{k+1}} \frac{M_t}{\sqrt{t_{k-1}\log\log t_{k-1}}} \geqslant \sqrt{3\delta}\right) \\
&\leqslant \sum_{k>l} P\left(\max_{t_k \leqslant t \leqslant t_{k+1},\, t_{k-1} \leqslant s \leqslant t_{k+2}} \frac{|B_s-B_t|}{\sqrt{t_{k-1}\log\log t_{k-1}}} \geqslant \sqrt{3\delta}\right) \\
&\leqslant \frac{4}{\sqrt{6\pi}} \sum_{k>l} (\log\log t_{k-1})^{-1/2} (\log t_{k-1})^{-3/2}
\end{aligned}$$

for large l .
Now we have to decide an integer $L(\varepsilon)$ for all $\varepsilon>0$ such that if $l \geqslant L$, then $(\log\log t_l)^{-1/2} < \sqrt{\varepsilon}$ and $\sum_{k>l} (\log t_{k-1})^{-3/2} < \sqrt{\varepsilon}$.
First, note that

$$
\begin{aligned}
(\log\log t_l)^{-1/2} < \sqrt{\varepsilon} &\iff \log\log t_l > \varepsilon^{-1} \\
&\iff t_l > \exp\exp\varepsilon^{-1} \\
&\iff (1+\varepsilon)^l > \exp\exp\varepsilon^{-1} \\
&\iff l\log(1+\varepsilon) > \exp\varepsilon^{-1} \\
&\iff l > \exp\varepsilon^{-1}/\log(1+\varepsilon)
\end{aligned}
$$

Let $L_1(\varepsilon) := \min\{l; l > \exp\varepsilon^{-1}/\log(1+\varepsilon)\}$.
Secondly, note that

$$
\begin{aligned}
\sum_{k>l} (\log t_{k-1})^{-3/2} < \sqrt{\varepsilon} &\iff \sum_{k\geqslant l} (\log t_k)^{-3/2} < \sqrt{\varepsilon} \\
&\iff \sum_{k\geqslant l} \{k\log(1+\varepsilon)\}^{-3/2} < \sqrt{\varepsilon} \\
&\iff \{\log(1+\varepsilon)\}^{-3/2} \sum_{k\geqslant l} k^{-3/2} < \sqrt{\varepsilon} \\
&\Longleftarrow (l-1)^{-1/2} < \sqrt{\varepsilon}\{\log(1+\varepsilon)\}^{3/2}/2 \\
&\iff l > 1 + 4/[\varepsilon\{\log(1+\varepsilon)\}^3]
\end{aligned}
$$

Let $L_2(\varepsilon) := \min\{l; l > 1 + 4/[\varepsilon\{\log(1+\varepsilon)\}^3]\}$.
Then, it suffices to let $L(\varepsilon) := L_1(\varepsilon) \vee L_2(\varepsilon)$.