Math 206A Probability: Chapter 2 Laws of Large Numbers written by Rick Durrett

This article is transcribed by Charlie Seager

January 29, 2024

Chapter 2.1 Independence

Theorem 2.1.1
(i) If X and Y are independent then $\sigma(X)$ and $\sigma(Y)$ are
(ii) Conversely, if $\mathcal{F}$ and $\mathcal{G}$ are independent, $X \in \mathcal{F}$ and $Y \in \mathcal{G}$ then X and Y are independent.

Theorem 2.1.2 (i) If A and B are independent then so are A^c and B, A and B^c, and A^c and B^c.
(ii) Conversely events A and B are independent if and only if their indicator random variables 1_A and 1_B are independent.

Theorem 2.1.3 Let $A_1, A_2, ..., A_n$ be independent.
(i) $A_1^c, A_2, ..., A_n$ are independent;
(ii) $1_{A_1}, ..., 1_{A_n}$ are independent.

Chapter 2.1.1 Sufficient Conditions for Independence

Lemma 2.1.5 Without loss of generality we can suppose A_i contains Ω. In this case the condition is equivalent to

$$P(\cap_{i=1}^n A_i) = \prod_{i=1}^n P(A_i) \qquad \text{whenever } A_i \in \mathcal{A}_i$$

since we can set $A_i = \Omega$ for $i \notin I$.

Theorem 2.1.6 $\pi - \lambda$ Theorem. If $\mathcal{P}$ is a $\pi - system$ and $\mathcal{L}$ is a $\lambda-$system that contains $\mathcal{P}$ then $\sigma(\mathcal{P}) \subset \mathcal{L}$.

Theorem 2.1.7 Suppose $\mathcal{A}_1, \mathcal{A}_2, ..., \mathcal{A}_n$ are independent and each $\mathcal{A}_i$ is a π-system. Then $\sigma(\mathcal{A}_1), \sigma(\mathcal{A}_2), ..., \sigma(\mathcal{A}_n)$ are independent.

Theorem 2.1.8 In order for $X_1, ..., X_n$ to be independent, it is sufficient that for all $x_1, ..., x_n \in (-\infty, \infty]$

$$P(X_1 \leq x_1, ..., X_n \leq x_n) = \prod_{i=1}^n P(X_i \leq x_i)$$

Theorem 2.1.9 Suppose $\mathcal{F}_{i,j}, 1 \leq i \leq n, 1 \leq j \leq m(i)$ are independent and let $\mathcal{G}_i = \sigma(\cup_j \mathcal{F}_{i,j})$. Then $\mathcal{G}_1, ..., \mathcal{G}_n$ are independent.

Theorem 2.1.10 If for $1 \leq i \leq n, 1 \leq j \leq m(i)$, $X_{i,j}$ are independent and $\{_i : R^{m(i)} \to R$ are measurable then $f_i(X_{i,1}, ..., X_{i,m(i)})$ are independent.

Chapter 2.1.2 Independence, Distribution and Expectation

Theorem 2.1.11 Suppose $X_1, ..., X_n$ are independent random variable and X_i has distribution μ_i, then $(X_1, ..., X_n)$ has distribution $\mu_1 x \dots x \mu_n$.

Theorem 2.1.12 Suppose X and Y are independent and have distributions μ and v. If $h : R^2 \to \mathcal{R}$ is a measurbale function with $h \geq 0$ or $E|h(X,Y)| < \infty$ then

$$Eh(X,Y) = \int\int h(x,y)\mu(dx)v(dy)$$

In particular, if h(x,y) = f(x)g(y) where $f, g : R \to \mathcal{R}$ are measurable functions with f, $g \geq 0$ or $E|f(X)|$ and $E|g(Y)| < \infty$ then

$$Ef(X)g(Y) = Ef(X) \cdot Eg(Y)$$

Theorem 2.1.13 If $X_1 ... X_n$ are independent and have (a) $X_i \geq 0$ for all i, or (b) $E|X_i| < \infty$ for all i then

$$E(\prod_{i=1}^n X_i) = \prod_{i=1}^n EX_i$$

i.e. the expectation on the left exists and has the value given on the right.

Chapter 2.1.3 Sums of Independent Random Variables Theorem 2.1.15 If X and Y are independent, $F(x) = P(X \leq x)$, and $G(y) = P(Y \leq y)$, then

$$P(X + Y \leq z) = \int F(z - y)dG(y)$$

The integral on the right hand side is called the convolution of F and G and is denoted F * G(z). The meaning of dG(y) will be explained in the proof.

Theorem 2.1.16 Suppose that X and density f and Y with distribution function G are independent. Then X + Y has density

$$h(x) = \int f(x - y)dG(y)$$

When Y has density g, the last formula can be written as

$$h(x) = \int f(x - y)g(y)dy$$

Theorem 2.1.18 If X = gamma (α, λ) and Y = gamma(β, λ) are independent then X + Y is gamma$(\alpha+\beta, \lambda)$. Consequently if $X_1, ..., X_n$ are independent exponential(λ) r.v.'s then $X_1 + \cdots + X_n$ has a gamma(n, λ) distribution.

Theorem 2.1.20 If X = normal(μ, a) and Y = normal(v, b) are independent then X + Y = normal$(\mu + v, a + b)$.

Chapter 2.1.4 Constructing Independent Random Variables

Theorem 2.1.21 Kolmogorov's extension theorem Suppose we are given probability measures μ_n on $(R^n, \mathcal{R}^n)$ that are consistent that is,

$$\mu_{n+1}((a_1, b_1]x \ldots x(a_n, b_n]xR) = \mu_n((a_1, b_1]x \ldots x(a_n, b_n])$$

Then there is a unique probability measure P on $(R^N, \mathcal{R}^N)$ with

$$P(\omega : \omega_i \in (a_i, b_i], 1 \leq i \leq n) = \mu_n((a_1, b_1]x \ldots x(a_n, b_n])$$

Theorem 2.1.22 If S is a Borel subset of a complete seperable metric space M, and S is the collection of Borel subsets of S, then (S, $\mathcal{S}$) is nice.

Chapter 2.2 Weak Laws of Large Numbers

Chapter 2.2.1 L^2 Weak Laws

Theorem 2.2.1 Let $X_1, ..., X_n$ have $E(X_i^2) < \infty$ and be uncorrelated. Tehn

$$var(X_1 + \cdots + X_n) = var(X_1) + \cdots + var(X_n)$$

where var(Y) = the variance of Y.

Lemma 2.2.2 If $p > 0$ and $E|Z_n|^p \to 0$ then $Z_n \to 0$ in probability.

Theorem 2.2.3 L^2 **weak law** Let $X_1, X_2, ...$ be uncorrelated random variables with $EX_i = \mu$ and $var(X_i) \le C < \infty$. If $S_n = X_1 + \cdots + X_n$ then as $n \to \infty$, $S_n/n \to \mu$ in L^2 and in probability.

Chapter 2.2.2 Triangular Arrays

Theorem 2.2.6 Let $\mu_n = ES_n, \sigma_n^2 = var(S_n)$. If $\sigma_n^2/b_n^2 \to 0$ then

$$\frac{S_n - \mu_n}{b_n} \to 0 \text{ in probability}$$

Lemma 2.2.9 $X_{n,1}, ..., X_{n,n}$ are independent and $P(X_{n,j} = 1) = \frac{1}{n-j+1}$

Chapter 2.2.3 Truncation

Theorem 2.2.11 Weak law for triangular arrays For each n let $X_{n,k}, 1 \le k \le n$ be independent. Let $b_n > 0$ with $b_n \to \infty$ and let $X_{n,k} = X_{n,k}1(|X_{n,k}| \le b_n)$. Suppose that as $n \to \infty$

$$\text{(i) } \sum_{k=1}^n P(|X_{n,k}| > b_n) \to 0 \text{ and}$$
$$\text{(ii) } b_n^{-2} \sum_{k=1}^n EX_{n,k}^2 \to 0$$

If we let $S_n = X_{n,1} + \cdots + X_{n,n}$ and put $a_n = \sum_{k=1}^n E\bar{X}_{n,k}$ then

$$(S_n - a_n)/b_n \to 0 \text{ in probability}$$

Theorem 2.2.12 Weak Law of Large Numbers Let $X_1, X_2 \ldots$ be i.i.d. with

$$xP(|X_i| > x) \to 0 \text{ as } x \to \infty$$

Lemma 2.2.13 If $Y \ge 0$ and $p > 0$ then $E(Y^p) = \int_0^\infty py^{p-1}P(Y > y)dy$

Theorem 2.2.14 Let $X_1, X_2, ...,$ be i.i.d. with $E|X_i| < \infty$. Let $S_n = X_1 + \cdots + X_n$ and let $\mu = EX_1$. Then $S_n/n \to \mu$ in probability.

Chapter 2.3 Borel-Cantelli Lemmas

Theorem 2.3.1 Borel-Cantelli Lemma If $\sum_{n=1}^\infty P(A_n) < \infty$ then

$$P(A_n i.o.) = 0$$

Theorem 2.3.2 $X_n \to X$ in probability if and only if for every subsequence $X_{n(m)}$ there is a further subsequence $X_{n(m_k)}$ that converges almost surely to X.

Theorem 2.3.3 Let y_n be a sequence of elements of a topological space. If every subsequence $y_{n(m)}$ has a further subsequence $y_{n(m_k)}$ that converges to y then $y_n \to y$

Theorem 2.3.4 If f is continous and $X_n \to X$ in probability then $f(X_n) \to f(X)$ in probability. If, in addition, f is bounded then $Ef(X_n) \to Ef(X)$.

Theorem 2.3.5 Let $X_1, X_2...$ be i.i.d. with $EX_i = \mu$ and $EX_i^4 < \infty$. If $S_n = X_1 + \cdots + X_n$ then $S_n/n \to \mu$ a.s.

Theorem 2.3.7 The second Borel-Cantelli lemma If the events A_n are independent then $\sum P(A_n) = \infty$ implies $P(A_n \text{ i.o. }) = 1$

Theorem 2.3.8 If $X_1, X_2, ...,$ are i.i.d. with $E|X_i| = \infty$ then $P(|X_n| \geq n$ i.o.$) = 1$. So if $S_n = X_1 + \cdots + X_n$ then P(lim S_n/n exists $\in (-\infty, \infty)) = 0$

Theorem 2.3.9 If $A_1, A_2....$ are pairwise independent and $\sum_{n=1}^{\infty} P(A_n) = \infty$ then as $n \to \infty$

$$\sum_{m=1}^{n} 1A_m / \sum_{m=1}^{n} P(A_m) \to 1 \text{ a.s.}$$

Theorem 2.3.11 If $R_n = \sum_{m=1}^{n} 1A_m$ is the number of records at time n then as $n \to \infty$

$$R_n/logn \to 1 \text{ a.s.}$$

Chapter 2.4 Strong Law of Large Numbers

Theorem 2.4.1 Strong law of large numbers. Let $X_1, X_2, ...,$ be pairwise independent identically distributed random variables with $E|X_i| < \infty$. Let $EX_i = \mu$ and $S_n = X_1 + \cdots + X_n$. Then $S_n/n \to \mu$ a.s. as $n \to \infty$

Lemma 2.4.2 Let $Y_k = X_k 1_{(|X_k| \leq k)}$ and $T_n = Y_1 + \ldots Y + Y_n$. It is sufficient to prove that $T_n/n \to \mu$ a.s.

Lemma 2.4.3 $\sum_{k=1}^{\infty} var(Y_k)/k^2 \leq 4E|X_1| < \infty$.

Lemma 2.4.4 If $y \geq 0$ then $2y \sum_{k>y} k^{-2} \leq 4$

Theorem 2.4.5 Let $X_1, X_2, ...,$ be i.i.d. with $EX_i^+ = \infty$ and $EX_i^- < \infty$. If $S_n = X_1 + \cdots + X_n$ then $S_n/n \to \infty$ a.s.

Theorem 2.4.7 If $EX_1 = \mu \leq \infty$ then as $t \to \infty$,

$$N_t/t \to 1/\mu \qquad \text{a.s. } (1/\infty = 0)$$

Theorem 2.4.9 The Glivenko-Cantelli theorem As $n \to \infty$,

$$sup_x |F_n(x) - F(x)| \to 0 \qquad \text{a.s.}$$

Chapter 2.5 Convergence of Random Series*

Theorem 2.5.3 Komogorov's 0-1 law If $X_1, X_2, ...$ are independent and $A \in \mathcal{T}$ then P(A) = 0 or 1.

Theorem 2.5.4 Hewitt-Savage 0-1 law. If $X_1, X_2, ...$ are i.i.d. and $A \in \epsilon$ then $P(A) \in \{0, 1\}$

Theorem 2.5.5 Komogorov's maximal inequality. Suppose $X_1, ..., X_n$ are independent with $EX_i = 0$ and var$(X_i) < \infty$. If $S_n = X_1 + \cdots + X_n$ then

$$P(max 1 \leq k \leq n |S_k| \geq x) \leq x^{-2} var(S_n)$$

Theorem 2.5.6 Suppose $X_1, X_2, ...$ are independent and have $EX_n = 0$. If

$$\sum_{n=1}^{\infty} var(X_n) < \infty$$

then with probability one $\sum_{n=1}^{\infty} X_n(\omega)$ converges.

Theorem 2.5.8 Kolmogorov's three-series theorem Let $X_1, X_2, ...$ be independent. Let $A > 0$ and let $Y_i = X_i 1_{(|X_i| \leq A)}$. In order that $\sum_{n=1}^{\infty} X_n$ converges a.s. it is necessary and sufficient that

$$\text{(i) } \sum_{n=1}^{\infty} P(|X_n| > A) < \infty, \text{ (ii) } \sum_{n=1}^{\infty} EY_n \text{ converges, and (iii) } \sum_{n=1}^{\infty} var(Y_n) < \infty$$

Theorem 2.5.9 Kronecker's lemma If $a_n \uparrow \infty$ and $\sum_{n=1}^{\infty} x_n/a_n$ converges then

$$a_n^{-1} \sum_{m=1}^{n} x_m \to 0$$

Theorem 2.5.10 The strong law of large numbers Let $X_1, X_2, ...$ be i.i.d. random variables with $E|X_i| < \infty$. Let $EX_i = \mu$ and $S_n = X_1 + ... + X_n$. Then $S_n/n \to \mu$ a.s. as $n \to \infty$

Chapter 2.5.1 Rates of Convergence

Theorem 2.5.11 Let $X_1, X_2, ...$ be i.i.d. random variables with $EX_i = 0$ and $EX_i^2 = \sigma^2 < \infty$. Let $S_n = X_1 + \cdots + X_n$. If $\epsilon > 0$ then

$$S_n/n^{1/2}(logn)^{1/2+\epsilon} \to 0 \text{ a.s.}$$

Theorem 2.5.12 Let $X_1, X_2, ...$ be i.i.d. with $EX_1 = 0$ and $E|X_1|^p < \infty$ where $1 < p < 2$. If $S_n = X_1 + \cdots + X_n$ then $S_n/n^{1/p} \to 0$ a.s.

Chapter 2.5.2 Infinite Mean

Theorem 2.5.13 Let $X_1, X_2, ...$ be i.i.d. with $E|X_1| = \infty$ and let $S_n = X_1 + \cdots + X_n$. Let a_n be a sequence of positive numbers with a_n/n increasing. Then $\lim sup_{n\to\infty} |S_n|/a_n = 0$ or ∞ according as $\sum_n P(|X_1| \geq a_n) < \infty$ or $= \infty$

Chapter 2.6 Renewal Theorey*

Theorem 2.6.1 As $t \to \infty$, $N_t/t \to 1/\mu$ a.s. where $\mu = E\xi_i \in (0, \infty]$ and $1/\infty = 0$

Theorem 2.6.2 Wald's equation. Let $X_1, X_2, ...$ be i.i.d. with $E|X_i| < \infty$. If N is a stopping time with $EN < \infty$ then $ES_N = EX_1 EN$.

Theorem 2.6.3 As $t \to \infty, U(t)/t \to 1/\mu$

Theorem 2.6.4 Blackwell's renewal theorem If F is nonarithmetic then

$$U([t, t+h]) \to h/\mu \qquad \text{as } t \to \infty$$

Theorem 2.6.9 If h is bounded then the function

$$H(t) = \int_0^t h(t-s) dU(s)$$

is the unique solution of the renewal equation that is bounded on bounded intervals.

Theorem 2.6.12 The renewal theorem If F is nonarithmetic and h is directly Riemann integrable then as $t \to \infty$

$$H(t) \to \frac{1}{\mu} \int_0^{\infty} h(s) ds$$

Lemma 2.6.13 If $h(x) \geq 0$ is decreasing with $h(0) < \infty$ and $\int_0^{\infty} h(x)dx < \infty$, then h is directly Riemann integrable.

Chapter 2.7 Large Deviations*

Lemma 2.7.1 If $\gamma_{m+n} \geq \gamma_m + \gamma_n$ then as $n \to \infty, \gamma_n/n \to sup_m \gamma_m/m$.

Lemma 2.7.2 If $a > \mu$ and $\theta > 0$ is small then $a\theta - k(\theta) > 0$

Theorem 2.7.7 Suppose in addition to (H1) and (H2) that there is $\theta_a \in (0, \theta_+)$ so that $a = \varphi'(\theta_a)/\varphi(\theta_a)$. Then as $n \to \infty$

$$n^{-1} logP(S_n \geq na) \to -a\theta_a + log\varphi(\theta_a)$$

Lemma 2.7.8 $\frac{dF^n}{dF^n_\lambda} = e^{-\lambda x}\varphi(\lambda)^n$

Theorem 2.7.9 Suppose $x_o = sup\{x : F(x) < 1\} < \infty$ and F is not a point mass at $x_0.\phi(\theta) < \infty$ for all $\theta > 0$ and $\phi^{'}(\theta)/\phi(\theta) \to x_o$ as $\theta \uparrow \infty$

Theorem 2.7.10 Suppose $x_o = \infty, \theta_+ < \infty$, and $\varphi^{'}(\theta)/\varphi(\theta)$ increases to a finite limit a_0 as $\theta \uparrow \theta_+$. If $a_0 \leq a < \infty$

$$n^{-1} log P(S_n \geq na) \to -a\theta_+ + log\varphi(\theta_+)$$

i.e. $\gamma(a)$ is linear for $a \geq a_0$.