Math 206A Probability: Chapter 3 Central Limit Theorems written by Rick Durrett

This article is transcribed by Charlie Seager

January 29, 2024

Chapter 3.1 The De Moivre-Laplace Theorem

Lemma 3.1.1 If $c_j \to 0, a_j \to \infty$ and $a_j c_j \to \lambda$ then $(1+c_j)^{aj} \to e^{\lambda}$.

Theorem 3.1.2 If 2k/ $\sqrt{2n} \to x$ then $P(S_{2n} = 2k)$ $(\pi n)^{-1/2} e^{-x^2/2}$.

Theorem 3.1.3 The De Moivre-Laplace Theorem If $a < b$ then as $m \to \infty$

$$P(a \leq S_m/\sqrt{m} \leq b) \to \int_a^b (2\pi)^{-1/2} e^{-x^2/2} dx$$

Chapter 3.2 Weak Convergence

Lemma 3.2.7 V_{n+1} has density function

$$fv_{n+1}(x) = (2n+1)\binom{2n}{n} x^n (1-x)^n$$

Chapter 3.2.2 Theory

Theorem 3.2.8 If $F_n \Rightarrow F_\infty$ then there are random variables $Y_n, 1 \leq n \leq \infty$ with distribution F_n so that $Y_n \to Y_\infty$ a.s.

Theorem 3.2.9 $X_n \Rightarrow X_\infty$ if and only if for every bounded continous function g we have $Eg(X_n) \to Eg(X_\infty)$

Theorem 3.2.10 Continous mapping theorem Let g be a measurable function and $D_g = \{x : g$ is discontinous at x$\}$. If $X_n \Rightarrow X_\infty$ and $P(X_\infty \in D_g) = 0$ then $g(X_n) \Rightarrow g(X)$. If in addition g is bounded then $Eg(X_n) \to Eg(X_\infty)$

Theorem 3.2.11 The following statements are equivalent: (i) $X_n \Rightarrow X_\infty$
(ii) For all open sets G, lim $inf_{n\to\infty} P(X_n \in G) \geq P(X_\infty \in G)$
(iii) For all closed sets K, lim $sup_{n\to\infty} P(X_n \in K) \leq P(X_\infty \in K)$
(iv) For all Borel sets A with $P(X_\infty \in \partial A) = 0, lim_{n\to\infty} P(X_n \in A) = P(X_\infty \in A)$.

Theorem 3.2.12 Helly's selection theorem For every sequence F_n of distribution functions, there is a subsequence $F_{n(k)}$ and a right continuous non-decreasing function F so that $lim_{k\to\infty} F_{n(k)}(y) = F(y)$ at all continuity points y of F.

Theorem 3.2.13 Every subsequential limit is the distribution function of a probability measure if and only if the sequence F_n is tight, i.e. for all $\epsilon > 0$ there is an M, so that

$$limsup_{n\to\infty} 1 - F_n(M_\epsilon) + F_n(-M_\epsilon) \le \epsilon$$

Theorem 3.2.14 If there is a $\varphi \ge 0$ so that $\varphi(x) \to \infty$ as $|x| \to \infty$ and

$$C = sup_n \int \varphi(x) dF_n(x) < \infty$$

Theorem 3.2.15 If each subsequence of X_n has a further subsequence that converges to X then $X_n \Rightarrow X$.

Chapter 3.3 Characteristic Functions

Chapter 3.3.1 Definition, Inversion Formula

Theorem 3.3.1 All characteristic functions have the following properties:
(a) $\varphi(0) = 1$
(b) $\varphi(-t) = \bar{\varphi(t)}$
(c) $|\varphi(t)| = |Ee^{itX}| \le E|e^{itX}| = 1$
(d) $|\varphi(t+h) - \varphi(t)| \le E|e^{ihX} - 1|$, so $\varphi(t)$ is uniformly continous on $(-\infty, \infty)$
(e) $Ee^{it(aX+b)} = e^{itb}\varphi(at)$

Theorem 3.3.2 If X_1 and X_2 are independent and have ch.f.'s φ_1 and φ_2 then $X_1 + X_2$ has ch.f. $\varphi_1(t)\varphi_2(t)$

Lemma 3.3.9 If $F_1, ..., F_n$ have ch.f. $\varphi_1, ..., \varphi_n$ and $\lambda_i \ge 0$ have $\lambda_1 + \cdots + \lambda_n = 1$ then $\sum_{i=1}^n \lambda_i F_i$ has ch.f. $\sum_{i=1}^n \lambda_i \varphi_i$

Theorem 3.3.11 The inversion formula Let $\varphi(t) = \int e^{itx}\mu(dx)$ where μ is a probability measure. If $a < b$ then

$$lim_{T\to\infty} (2\pi)^{-1} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \varphi(t) dt = \mu(a,b) + \frac{1}{2}\mu(\{a,b\})$$

Corollary 3.3.12 If φ is real then X and -X have the same distribution.

Corollary 3.3.13 If X_i i =1,2 are independent and have normal distributions with mean 0 and variance σ_i^2 then $X_1 + X_2$ has a normal distribution with mean 0 and variance $\sigma_1^2 + \sigma_2^2$

Theorem 3.3.14 If $\int |\varphi(t)| dt < \infty$ then μ has bounded continous density

$$f(y) = \frac{1}{2\pi} \int e^{-ity} \varphi(t) dt$$

Chapter 3.3.2 Weak Convergence

Theorem 3.3.17 Continuity theorem Let $\mu_n, 1 \le n \le \infty$ be probability measures with ch.f. φ_n.
(i) If $\mu_n \Rightarrow \mu_\infty$ then $\varphi_n(t) \to \varphi_\infty(t)$ for all t.
(ii) If $\varphi_n(t)$ converges pointwise to a limit $\varphi(t)$ that is continuous at 0, then the associated sequence of distributions μ_n is tight and converges to the measure μ with characteristic function φ

Chapter 3.3.3 Moments and Derivatives

Theorem 3.3.18 If $\int |x|^n \mu(dx) < \infty$ then its characteristic function φ has a continuous derivative of order n given by $\varphi^{(n)}(t) = \int (ix)^n e^{itx} \mu(dx)$

Lemma 3.3.19

$$|e^{ix} - \sum_{m=0}^{n} \frac{(ix)^m}{m!}| \le min(\frac{|x|^{n+1}}{(n+1)!'} \frac{2|x|^n}{n!})$$

The first term on the right is the usual order of magnitude we expect in the correction term. The second is better for large $|x|$ and will help us prove the central limit theorem without assuming finite third moments.

Theorem 3.3.20 If $E|X|^2 < \infty$ then

$$\varphi(t) = 1 + itEX - t^2E(X^2)/2 + o(t^2)$$

Theorem 3.3.21 If $\lim sup_{h|0}\{\varphi(h) - 2\varphi(0) + \varphi(-h)\}/h^2 > -\infty$ then $E|X|^2 < \infty$

Chapter 3.3.4 Polya's Criterion*

Theorem 3.3.22 Polya's criterion Let $\varphi(t)$ be real nonnegative and have $\varphi(0) = 1, \varphi(t) = \varphi(-t)$ and φ is decreasing and convex on $(0, \infty)$ with

$$lim_{t\downarrow 0}\varphi(t) = 1, \qquad lim_{t\uparrow\infty}\varphi(t) = 0$$

Then there is a probability measure v on $(0, \infty)$, so that (*)

$$\varphi(t) = \int_0^\infty (1 - |\tfrac{t}{s}|)^+ v(ds)$$

and hence φ is a characteristic function.

Chapter 3.3.5 The Moment Problem*

Theorem 3.3.25 If $\lim sup_{k\to\infty}\mu_{2k}^{1/2k}/2k = r < \infty$ then there is at most one d.f. F with $\mu_k = \int x^k dF(x)$ for all positive integers k.

Theorem 3.3.26 Suppose $\int x^k dF_n(x)$ has a limit μ_k for each k and

$$\lim sup_{k\to\infty}\mu_2 k^{1/2k}/2k < \infty$$

then F_n converges weakly to the unique distribution with these moments.

Chapter 3.4 Central Limit Theorems

Chapter 3.4.1 i.i.d. Sequences

Theorem 3.4.1 Let $X_1, X_2, ...$ be i.id. with $EX_i = \mu$, $var(X_i) = \sigma^2 \in (0, \infty)$. If $S_n = X_1 + \cdots + X_n$ then

$$(S_n - n\mu)/\sigma n^{1/2} \Rightarrow \mathcal{X}$$

where $\mathcal{X}$ has the standard normal distribution.

Theorem 3.4.2 If $c_n \to c \in C$ then $(1 + c_n/n)^n \to e^c$

Lemma 3.4.3 Let $z_1, ..., z_n$ and $w_1, .., w_n$ be complex numbers of modulus $\leq \theta$. Then

$$|\prod_{m=1}^n z_m - \prod_{m=1}^n w_m| \leq \theta^{n-1}\sum_{m=1}^n |z_m - w_m|$$

Lemma 3.4.4 If b is a complex number with $|b| \leq 1$ then $|e^b - (1+b)| \leq |b|^2$.

Chapter 3.4.2 Triangular Arrays

Theorem 3.4.10 The Lindeberg-Feller Theorem For each n, let $X_{n,m}, 1 \leq m \leq n$, be independent random variables with $EX_{n,m} = 0$. Suppose
(i) $\sum_{m=1}^n EX_{n,m}^2 \to \sigma^2 > 0$
(ii) For all $\epsilon > 0$, $lim_{n\to\infty}\sum_{m=1}^n E(|X_{n,m}|^2 : |X_{n,m}| > \epsilon) = 0$
Then $S_n = X_{n,1} + \cdots + X_{n,n} \Rightarrow \sigma\mathcal{X}$ as $n \to \infty$

THeorem 3.4.14 Let $X_1, X_2, ...$ be i.i.d. and $S_n = X_1 + \cdots + X_n$. In order that there exist constants a_n and $b_n > 0$ so that $(S_n - a_n)/b_n \Rightarrow \mathcal{X}$ it is necessary and sufficient that

$$y^2P(|X_1| > y)/E(|X_1|^2; |X_1| \leq y) \to 0$$

Chapter 3.4.3 Prime Divisors (Erdos-Kac)*

Lemma 3.4.15 $h_n(\epsilon) \to 0$ for each fixed $\epsilon > 0$ so we can pick $\epsilon_n \to 0$ so that $h_n(\epsilon_n) \to 0$

Theorem 3.4.16 Erdos-Kac central limit theorem As $n \to \infty$

$$P_n(m \leq n : g(m) - loglogn \leq x(loglogn)^{1/2}) \to P(\mathcal{X} \leq x)$$

Chapter 3.4.4 Rates of Convergence (Berry-Esseen)*

Theorem 3.4.17 Let $X_1, X_2, ...$ be i.i.d. with $EX_i = 0, EX_i^2 = \sigma^2$, and $E|X_i|^3 = p < \infty$. If $F_n(x)$ is the distribution of $(X_1 + \cdots + X_n)/\sigma\sqrt{n}$ and $\mathcal{N}(x)$ is the standard normal distribution, then

$$|F_n(x) - \mathcal{N}(x)| \leq 3p/\sigma^3\sqrt{n}$$

Lemma 3.4.18 Let F and G be distribution functions with $G^{'}(x) \leq \lambda < \infty$. Let $\Delta(x) = F(x) - G(x)$, $\mathfrak{N} = sup|\Delta(x)|, \Delta_L = \Delta * H_L$, and $\mathfrak{N}_L = sup|\Delta_L(x)|$. Then

$$\mathfrak{N}_L \geq \tfrac{\mathcal{N}}{2} - \tfrac{12\lambda}{\pi L} \qquad \text{or} \qquad \mathcal{N} \leq 2\mathcal{N}_L + \tfrac{24\lambda}{\pi L}$$

Lemma 3.4.19 Let K_1 and K_2 be d.f. with mean 0 whose ch.f. $\|_i$ are integrable

$$K_1(x) - K_2(x) = (2\pi)^{-1}\int -e^{itx}\tfrac{\|_1(t) - \|_2(t)}{it}dt$$

Chapter 3.5 Local Limit Theorems*

Theorem 3.5.2 Let $\varphi(t) = Ee^{itX}$. There are only three possibilities
(i) $|\varphi(t)| < 1$ for all $t \neq 0$
(ii) There is a $\lambda > 0$ so that $|\varphi(\lambda)| = 1$ and $|\varphi(t)| < 1$ for $0 < t < \lambda$. In this case, X has a lattice distribution with span $2\pi/\lambda$
(iii) $|\varphi(t)| = 1$ for all t. In this case, X = b a.s. for some b.

Theorem 3.5.3 Under the hypotheses above, as $n \to \infty$

$$sup_{x\in\mathcal{L}_n}|\tfrac{n^{1/2}}{h}p_n(x) - n(x)| \to 0$$

Theorem 3.5.4 Under the hypotheses above, if $x_n/\sqrt{n} \to x$ and $a < b$

$$\sqrt{n}P(S_n \in (x_n + a,\, x_n + b)) \to (b - a)n(x)$$

Poisson Convergence

Chapter 3.6.1 The Basic Limit Theorem

Theorem 3.6.1 For each n let $X_{n,m}, 1 \leq m \leq n$ be independent random variables with $P(X_{n,m} = 1) = p_{n,m}, P(X_{n,m} = 0) = 1 - p_{n,m}$. Suppose
(i) $\sum_{m=1}^{n} p_{n,m} \to \lambda \in (0, \infty)$,
and (ii) $max_{1\leq m\leq n}p_{n,m} \to 0$ If $S_n = X_{n,1} + \ldots X_{n,n}$ then $S_n \Rightarrow Z$ where Z is Poisson(λ).

Lemma 3.6.4 (i) $d(\mu, v) = ||\mu - v||$ defines a metric on probability measures on Z and

(ii) $||\mu_n - \mu|| \to 0$ if and only if $\mu_n(x) \to \mu(x)$ for each $x \in Z$, which by Excercise 3.2.11 is equivalent to $\mu_n \Rightarrow \mu$

Lemma 3.6.5 If $\mu_1 x \mu_2$ denotes the product measure on Z x Z that has $(\mu_1 x \mu_2)(x, y) = \mu_1(x)\mu_2(y)$ then

$$||\mu_1 x \mu_2 - v_1 x v_2|| \leq ||\mu_1 - v_1|| + ||\mu_2 - v_2||$$

Lemma 3.6.6 If $\mu_1 * \mu_2$ denotes the convolution of μ_1 and μ_2, that is

$$\mu_1 * \mu_2(x) = \sum_y \mu_1(x-y)\mu_2(y)$$

Lemma 3.6.7 Let μ be the measure with $\mu(1) = p$ and $\mu(0) = 1-p$. Let v be a Poisson distribution with mean p. Then $||\mu - v|| \leq p^2$.

Chapter 3.6.2 Two Examples with Dependence

Theorem 3.6.10 If $ne^{-r/n} \to \lambda \in [0, \infty)$ the number of empty boxes approaches a Poisson distribution with mean λ.

Chapter 3.7 Poisson Processes

Theorem 3.7.1 Let $X_{n,m}, 1 \leq m \leq n$ be independent nonnegative integer valued random variables with $P(X_{n,m} = 1) = p_{n,m}, P(X_{n,m} \geq 2) = \epsilon_{n,m}$.
(i) $\sum_{m=1}^{n} p_{n,m} \to \lambda \in (0, \infty)$,
(ii) $max_{1 \leq m \leq n} p_{n,m} \to 0$,
and (iii) $\sum_{m=1}^{n} \epsilon_{n,m} \to 0$
If $S_n = X_{n,1} + \cdots + X_{n,n}$ then $S_n \Rightarrow Z$ where Z is Poisson(λ).

Theorem 3.7.2 If (i)-(iv) hold then N(0,t) has a Poisson distribution with mean λt.

Chapter 3.7.1 Compound Poisson Procces

Theorem 3.7.3 Let $Y_1, Y_2, ...$ be independent and identically distributed let N be an independent nonnegative integer valued random variable and let $S = Y_1 + \cdots + Y_n$ with S=0 when N=0
(i) If $E|Y_i|, EN < \infty$, then $ES = EN \cdot EY_i$
(ii) If $EY_i^2, EN^2 < \infty$, then var(S) = EN$var(Y_i) + var(N)(EY_i)^2$
(iii) If N is Poisson (λ), then var(S) = λEY_i^2.

Chapter 3.7.2 Thinning

Theorem 3.7.4 $N_j(t)$ are independent rate $\lambda P(Y_i = j)$ Poisson processes.

Theorem 3.7.5 Suppose that a Poisson process with rate λ we keep a point that lands at s with probability p(s). Then the result is a nonhomogeneous Poisson process with rate λp(s).

Chapter 3.7.3 Conditioning

Theorem 3.7.9 Let T_n be the time of the nth arrival in a rate λ Poisson process. Let $U_1, U_2, ..., U_n$ be independent uniform on (0,t) and let V_k^n be the kth smallest number in $\{U_1, ..., U_n\}$. If we condition on N(t) = n. Thevectors $V = (V_1^n, \ldots, V_n^n)$ and $T = (T_1, \ldots T_n)$ have the same distribution.

Corollary 3.7.10 If $0 < s < t$ then

$$P(N(s) = m|N(t) = n) = \binom{n}{m}(\tfrac{s}{t})^m(1 - \tfrac{s}{t})^{n-m}$$

Theorem 3.7.11 Let T_n be the time of the nth arrival in a rate λ Poisson process. Let $U_1, U_2, ..., U_n$ be independent uniform on (0,1) and let V_k^n

be the kth smallest number in $\{U_1, ..., U_n\}$. The vectors $(V_1^n, \ldots, V_n^n)$ and $(T_1/T_{n+1}, \ldots, T_n/T_{n+1})$ have the same distribution.

Chapter 3.8 Stable Laws*

Lemma 3.8.1 If $h_n(\epsilon) \to g(\epsilon)$ for each $\epsilon > 0$ and $g(\epsilon) \to g(0)$ as $\epsilon \to 0$ then we can pick $\epsilon_n \to 0$ so that $h_n(\epsilon_n) \to g(0)$

Theorem 3.8.2 Suppose $X_1, X_2, ...$ are i.i.d. with a distribution that satisfies
(i) $lim_{x\to\infty} P(X_1 > x)/P(|X_1| > x) = \theta \in [0, 1]$
(ii) $P(|X_1| > x) = x^{-\alpha} L(x)$
where $\alpha < 2$ and L is slowly varying. Let $S_n = X_1 + \cdots + X_n$
$a_n = inf\{x : P(|X_1| > x) \leq n^{-1}\}$ and $b_n = nE(X_1 1_{(|X_1| \leq a_n)})$
As $n \to \infty, (S_n - b_n)/a_n \Rightarrow Y$ where Y has a nondegenerate distribution.

Lemma 3.8.3 For any $\delta > 0$ there is C so that for all $t \geq t_0$ and $y \leq 1$

$$P(|X_1| > yt)/P(|X_1| > t) \leq Cy^{-\alpha-\delta}$$

Theorem 3.8.8 Y is a limit of $(X_1 + \cdots + X_k - b_k)/a_k$ for some i.i.d. sequence X_i if and only if Y has a stable law.

Theorem 3.8.9 Convergence of types theorem If $W_n \Rightarrow W$ and there are constants $\alpha_n > 0, \beta_n$ so that $W_n' = \alpha_n W_n + \beta_n \Rightarrow W'$ where W and W' are nondegenerate, then there are constants α and β so that $\alpha_n \to \alpha$ and $\beta_n \to \beta$

Chapter 3.9 Infinitely Divisible Distributions*

Theorem 3.9.1 Z is a limit of sums of type (*) if and only if Z has an infinitely divisible distribution.

Theorem 3.9.6 Levy-Khinchin Theorem Z has an infinitely divisible distribution if and only if its characteristic function has

$$log\varphi(t) = ict - \frac{\sigma^2 t^2}{2} + \int (e^{itx} - 1 - \frac{itx}{1+x^2})\mu(dx)$$

where μ is a measure with $\mu(\{0\}) = 0$ and $\int \frac{x^2}{1+x^2}\mu(dx) < \infty$

Theorem 3.9.7 Kolmogorov's Theorem Z has an infinitely divisible distribution with mean 0 and finite variance if and only if its ch.f. has

$$log\varphi(t) = \int (e^{itx} - 1 - itx)x^{-2} v(dx)$$

Here the integrand is $-t^2/2$ at 0, v is called the canonical measure and var(Z) = v(R).

Chapter 3.10 Limit Theorems in R^d

Theorem 3.10.1 The following statements are equivalent:
(i) $Ef(X_n) \to Ef(X_\infty)$ for all bounded continuous f.
(ii) $Ef(X_n) \to Ef(X_\infty)$ for all bounded Lipshitz continuous f.
(iii) For all closed sets K, $limsup_{n\to\infty} P(X_n \in K) \leq P(X_\infty \in K)$.
(iv) For all open sets G, $liminf_{n\to\infty} P(X_n \in G) \geq P(X_\infty \in G)$
(v) For all sets A with $P(X_\infty \in \partial A) = 0$, $lim_{n\to\infty} P(X_n \in A) = P(X_\infty \in A)$
(vi) Let D_f = the set of discontinuities of f. For all bounded functions f with $P(X_\infty \in D_f) = 0$ we have $Ef(X_n) \to Ef(X_\infty)$.

Theorem 3.10.2 On R^d weak convergence defined in terms of convergence of distribution $F_n \Rightarrow F$ is equivalent to notion of weak convergence defined for a general metric space.

Theorem 3.10.3 If μ_n is tight, then there is a weakly convergent subsequence.

Theorem 3.10.4 Inversion formula if $A = [a_1, b_1]x...x[a_d, b_d]$ with $\mu(\partial A) = 0$ then

$$\mu(A) = lim_{T\to\infty}(2\pi)^{-d}\int_{[-T,T]^d}\prod_{j=1}^{d}\psi_j(t_j)\varphi(t)dt$$

where $\psi_j(s) = (exp(-isa_j) - exp(-isb_j))/is.$

Theorem 3.10.5 Convergence theorem Let X_n, $1 \le n \le \infty$ be random vectors with ch.f. φ_n. A necessary and sufficient condition for $X_n \Rightarrow X_\infty$ is that $\varphi_n(t) \to \varphi_\infty(t)$

Theorem 3.10.6 Cramer Wold Device A sufficient condition for $X_n \Rightarrow X_\infty$ is that $\theta \cdot X_n \Rightarrow \theta \cdot X_\infty$ for all $\theta \in R^d$

Theorem 3.10.7 The central limit theorem in R^d let $X_1, X_2, ...$ be i.i.d. random vectors with $EX_n = \mu$ and finite covariances

$$\Gamma_{i,j} = E((X_{n,j} - \mu_i)(X_{n,j} - \mu_j))$$

If $S_n = X_1 + \cdots + X_n$ then $(S_n - n\mu)/n^{1/2} \Rightarrow \mathcal{X}$ where $\mathcal{X}$ has a multivariate normal distribution with mean 0 and covariance Γ, i.e.

$$Eexp(i\theta\cdot\mathcal{X}) = exp(-\sum_i\sum_j\theta_i\theta_j\Gamma_{i,j}/2)$$