Κεφάλαιο 6

Σύγκλιση ακολουθιών τυχαίων μεταβλητών

6.1 (1 Ιουνίου 2009)

6.1α΄ Στοχαστική σύγκλιση και σύγκλιση κατά κατανομή

Ορισμός 6.1.1. Έστω $(X_n)_{n\in\mathbb{N}}$ ακολουθία τυχαίων μεταβλητών στον χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$ και έστω X τυχαία μεταβλητή στον ίδιο χώρο πιθανότητας. Γράφουμε F_{X_n} και F_X για τις συναρτήσεις κατανομής των X_n και X αντίστοιχα.

(α) Λέμε ότι $X_n \to X$ **στοχαστικά** (ή **κατά πιθανότητα**) αν για κάθε $\varepsilon > 0$ ισχύει

$$\lim_{n\to\infty} \mathbb{P}(\,|X_n - X| \geq \varepsilon\,) = 0.$$

(α) Λέμε ότι $X_n \to X$ **κατά κατανομή** (ή **κατά νόμο**) αν, για κάθε $x \in \mathbb{R}$ στο οποίο η F_X είναι συνεχής, ισχύει

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x).$$

Με άλλα λόγια,

$$\lim_{n\to\infty} F_{X_n}(x) = F(x).$$

Θεώρημα 6.1.2. (α) *Αν $X_n \to X$ στοχαστικά, τότε $X_n \to X$ κατά κατανομή.*

(β) *Αν $X_n \to c$ κατά κατανομή, τότε $X_n \to c$ στοχαστικά (γράφοντας c εννοούμε μια τυχαία μεταβλητή X με την ιδιότητα $\mathbb{P}(X = x) = 1$).*

Θεώρημα 6.1.3. *Υποθέτουμε ότι $X_n \to X$ κατά κατανομή και $Y_n \to c$ στοχαστικά. Τότε:*

(i) $X_n \pm Y_n \to X \pm c$ *κατά κατανομή.*

(ii) $X_nY_n \to cX$ *κατά κατανομή.*

(iii) *Αν $c \neq 0$ και $\mathbb{P}(Y_n \neq 0) = 1$ για κάθε $n \in \mathbb{N}$, τότε $\frac{X_n}{Y_n} \to \frac{X}{c}$ κατά κατανομή.*

Θεώρημα 6.1.4 (Slutsky–Fréchet). *Αν $X_n \to X$ στοχαστικά και η $g : \mathbb{R} \to \mathbb{R}$ είναι συνεχής συνάρτηση, τότε $g(X_n) \to g(X)$ στοχαστικά.*

Θεώρημα 6.1.5 (Lévy–Cramér). *Έστω $(X_n)_{n\in\mathbb{N}}$ ακολουθία τυχαίων μεταβλητών στον χώρο πιθανότητας $(\Omega, \mathcal{A}, \mathbb{P})$, (F_n) η ακολουθία των συναρτήσεων κατανομής τους και (φ_n) η ακολουθία των χαρακτηριστικών τους συναρτήσεων.*

(α) *Αν $F = F_X$ είναι μια συνάρτηση κατανομής και $F_n(x) \to F(x)$ για κάθε $x \in \mathbb{R}$ στο οποίο η F είναι συνεχής, τότε $\varphi_n(t) \to \varphi_X(t)$ για κάθε $t \in \mathbb{R}$.*

(β) *Αν $\varphi : \mathbb{R} \to \mathbb{C}$ είναι συνάρτηση συνεχής στο $t = 0$ και $\varphi_n(t) \to \varphi(t)$ για κάθε $t \in \mathbb{R}$, τότε η φ είναι χαρακτηριστική συνάρτηση κάποιας τυχαίας μεταβλητής και $F_n(x) \to F(x)$ για κάθε $x \in \mathbb{R}$ στο οποίο η F_X είναι συνεχής.*

6.1β′ Νόμος των μεγάλων αριθμών του Bernoulli

Θεώρημα 6.1.6. *Έστω X_k το πλήθος των επιτυχιών στην k-οστή δοκιμή μιας ακολουθίας ανεξάρτητων δοκιμών Bernoulli με πιθανότητα επιτυχίας p. Τότε, η ακολουθία των δειγματικών μέσων*

$$\overline{X}_n = \frac{X_1 + \cdots + X_n}{n}$$

συγκλίνει στοχαστικά στη σταθερά p.

Απόδειξη. Έχουμε $\mathbb{P}(X_k = 1) = p$ και $\mathbb{P}(X_k = 0) = q = 1 - p$. Άρα, $\mathbb{E}(X_k) = p$ και $V(X_k) = pq$. Συνεπώς,

$$\mathbb{E}(\overline{X}_n) = \frac{1}{n}\sum_{k=1}^{n}\mathbb{E}(X_k) = \frac{np}{n} = p$$

και, λόγω της ανεξαρτησίας των X_k,

$$V(\overline{X}_n) = \frac{1}{n^2}\sum_{k=1}^{n}V(X_k) = \frac{npq}{n^2} = \frac{pq}{n}.$$

Έστω $\varepsilon > 0$. Από την ανισότητα του Chebyshev,

$$\mathbb{P}(|\overline{X}_n - p| \geq \varepsilon) = \mathbb{P}(|\overline{X}_n - \mathbb{E}(\overline{X}_n)| \geq \varepsilon) \leq \frac{V(\overline{X}_n)}{\varepsilon^2} = \frac{pq}{n\varepsilon^2} \to 0$$

όταν $n \to \infty$. Άρα, $\overline{X}_n \to p$ στοχαστικά.

6.1γ′ Νόμος των μεγάλων αριθμών του Chebyshev

Θεώρημα 6.1.7. *Έστω (X_k) ακολουθία ανεξάρτητων τυχαίων μεταβλητών με $\mathbb{E}(X_k) = \mu_k$ και $V(X_k) = \sigma_k^2 < \infty$. Αν*

$$\lim_{n\to\infty}\frac{1}{n^2}\sum_{k=1}^{n}\sigma_k^2 = 0,$$

τότε η ακολουθία $\overline{X}_n - \overline{\mu}_n$, όπου

$$\overline{X}_n = \frac{X_1 + \cdots + X_n}{n}$$

και

$$\overline{\mu}_n = \frac{\mu_1 + \cdots + \mu_n}{n},$$

συγκλίνει στοχαστικά στο 0.

Απόδειξη. Έχουμε

$$\mathbb{E}(\overline{X}_n) = \frac{1}{n}\sum_{k=1}^{n}\mathbb{E}(X_k) = \frac{1}{n}\sum_{k=1}^{n}\mu_k = \overline{\mu}_n$$

και, λόγω της ανεξαρτησίας των X_k,

$$V(\overline{X}_n) = \frac{1}{n^2}\sum_{k=1}^{n}V(X_k) = \frac{1}{n^2}\sum_{k=1}^{n}\sigma_k^2.$$

Έστω $\varepsilon > 0$. Από την ανισότητα του Chebyshev,

$$\mathbb{P}(|\overline{X}_n - \overline{\mu}_n| \geq \varepsilon) = \mathbb{P}(|\overline{X}_n - \mathbb{E}(\overline{X}_n)| \geq \varepsilon) \leq \frac{V(\overline{X}_n)}{\varepsilon^2} = \frac{1}{n\varepsilon^2}\sum_{k=1}^{n}\sigma_k^2 \to 0$$

όταν $n \to \infty$. Άρα, $\overline{X}_n - \overline{\mu}_n \to 0$ στοχαστικά.

Σημείωση. Όπως φαίνεται από την απόδειξη, η υπόθεση ότι οι X_k είναι ανεξάρτητες μπορεί να αντικατασταθεί από την υπόθεση ότι οι X_k είναι ασυσχέτιστες.

Πόρισμα 6.1.8. *Έστω (X_k) ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών με $\mathbb{E}(X_k) = \mu$ και $V(X_k) = \sigma^2 < \infty$. Τότε, $\overline{X}_n \to \mu$ στοχαστικά.*

Απόδειξη. Προκύπτει άμεσα από το προηγούμενο θεώρημα αν παρατηρήσουμε ότι $\overline{\mu}_n = \mu$ για κάθε $n \in \mathbb{N}$ και

$$\lim_{n\to\infty} \frac{1}{n^2} \sum_{k=1}^{n} V(X_k) = \lim_{n\to\infty} \frac{n\sigma^2}{n^2} = \lim_{n\to\infty} \frac{\sigma^2}{n} = 0.$$

Παράδειγμα 6.1.9. *Νόμος των μεγάλων αριθμών του Poisson*: Έστω X_k το πλήθος των επιτυχιών στην k-οστή δοκιμή μιας ακολουθίας ανεξάρτητων δοκιμών Bernoulli με πιθανότητα επιτυχίας p_k. Τότε, αν θέσουμε

$$\overline{X}_n = \frac{X_1 + \cdots + X_n}{n}$$

και

$$\overline{p}_n = \frac{p_1 + \cdots + p_n}{n},$$

έχουμε ότι $\overline{X}_n - \overline{p}_n \to 0$ στοχαστικά.

Για την απόδειξη, εφαρμόζουμε το νόμο των μεγάλων αριθμών του Chebyshev: χρησιμοποιώντας την ανισότητα $4x(1-x) \leq 1$ για κάθε $x \in [0,1]$, βλέπουμε ότι $4p_k(1-p_k) \leq 1$, άρα

$$\frac{1}{n^2} \sum_{k=1}^{n} V(X_k) = \frac{1}{4n^2} \sum_{k=1}^{n} 4p_k(1-p_k) \leq \frac{n}{4n^2} = \frac{1}{4n} \to 0$$

όταν το $n \to \infty$. Άρα, μπορούμε να εφαρμόσουμε το θεώρημα 6.1.7.

6.2 (3 Ιουνίου 2009)

6.2α΄ Νόμος των μεγάλων αριθμών του Khintchine

Θεώρημα 6.2.1. *Έστω (X_k) ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών με $\mathbb{E}(X_k) = \mu$. Τότε, $\overline{X}_n \to \mu$ στοχαστικά.*

Απόδειξη. Γράφουμε $\varphi_n = \varphi_{\overline{X}_n}$ και $\varphi = \varphi_{X_i}$. Γνωρίζουμε ότι

$$\varphi_n(t) = \prod_{i=1}^{n} \varphi_{X_i/n}(t) = \left[\varphi_{X_1/n}(t)\right]^n = [\varphi(t/n)]^n.$$

Αφού οι X_i έχουν μέση τιμή, το θεώρημα 5.4.2 μας εξασφαλίζει ότι

$$\varphi(s) = 1 + i\mu s + o(s).$$

Παίρνοντας λογαρίθμους, έχουμε

$$\log \varphi_n(t) = n \log \varphi(t/n) = n\left(\varphi(t/n) - 1\right) \frac{\log \varphi(t/n)}{\varphi(t/n) - 1}.$$

Όμως,

$$\frac{\log \varphi(t/n)}{\varphi(t/n) - 1} \to 1$$

όταν $n \to \infty$, και

$$n\left(\varphi(t/n) - 1\right) = n\left(i\mu\frac{t}{n} + o(t/n)\right) = i\mu t + t\frac{o(t/n)}{t/n} \to i\mu t$$

όταν $n \to \infty$. Συνεπώς,

$$\log \varphi_n(t) \to i\mu t,$$

δηλαδή

$$\varphi_n(t) \to e^{i\mu t}.$$

Η συνάρτηση $\varphi(t) = e^{i\mu t}$ είναι η χαρακτηριστική συνάρτηση της $X = \mu$. Από το θεώρημα 6.1.5 βλέπουμε ότι $\overline{X}_n \to \mu$ κατά κατανομή, και από το θεώρημα 6.1.2 έπεται ότι $\overline{X}_n \to \mu$ στοχαστικά.

6.2β′ Κεντρικό οριακό θεώρημα Lindeberg–Lévy

Θεώρημα 6.2.2. *Έστω (X_k) ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών με $\mathbb{E}(X_k) = \mu$ και $V(X_k) = \sigma^2 < \infty$. Θεωρούμε τους τυποποιημένους μέσους*

$$Z_n = \frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma}.$$

Τότε, η ακολουθία των συναρτήσεων κατανομής $F_n = F_{Z_n}$ συγκλίνει στη συνάρτηση κατανομής

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2} dt$$

της τυπικής κανονικής. Δηλαδή, $Z_n \to Z$ κατά κατανομή, όπου Z τυπική κανονική τυχαία μεταβλητή.

Απόδειξη. Γράφουμε $\varphi_n = \varphi_{Z_n}$ και $\varphi = \varphi_{Y_1}$, όπου $Y_k = \frac{X_k - \mu}{\sigma}$. Έχουμε

$$Z_n = \frac{Y_1 + \cdots + Y_n}{\sqrt{n}}.$$

Γνωρίζουμε ότι

$$\varphi_n(t) = \prod_{i=1}^{n} \varphi_{Y_i/\sqrt{n}}(t) = \left[\varphi_{Y_1/\sqrt{n}}(t)\right]^n = [\varphi(t/\sqrt{n})]^n.$$

Αφού οι Y_i έχουν μέση τιμή 0 και διασπορά 1, το θεώρημα 5.4.2 μας εξασφαλίζει ότι

$$\varphi(s) = 1 - \frac{s^2}{2} + o(s^2).$$

Παίρνοντας λογαρίθμους, έχουμε

$$\log \varphi_n(t) = n \log \varphi(t/\sqrt{n}) = n\left(\varphi(t/\sqrt{n}) - 1\right) \frac{\log \varphi(t/\sqrt{n})}{\varphi(t/\sqrt{n}) - 1}.$$

Όμως,

$$\frac{\log \varphi(t/\sqrt{n})}{\varphi(t/\sqrt{n}) - 1} \to 1$$

όταν $n \to \infty$, και

$$n\left(\varphi(t/\sqrt{n}) - 1\right) = n\left(-\frac{t^2}{2n} + o(t^2/n)\right) = -\frac{t^2}{2} + t^2 \frac{o(t^2/n)}{t^2/n} \to -\frac{t^2}{2}$$

όταν $n \to \infty$. Συνεπώς,

$$\log \varphi_n(t) \to -\frac{t^2}{2},$$

δηλαδή

$$\varphi_n(t) \to e^{-t^2/2}.$$

Η συνάρτηση $\varphi(t) = e^{-t^2/2}$ είναι η χαρακτηριστική συνάρτηση της Z. Από το θεώρημα 6.1.5 συμπεραίνουμε ότι $Z_n \to Z$ κατά κατανομή.

Σημείωση. Οι τυχαίες μεταβλητές Z_n γράφονται ισοδύναμα στη μορφή

$$Z_n = \frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma} = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

όπου

$$S_n = X_1 + \cdots + X_n.$$

Έτσι, το κεντρικό οριακό θεώρημα διατυπώνεται ισοδύναμα ως εξής: για κάθε $z \in \mathbb{R}$,

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq z\right) = \Phi(z).$$

Μπορούμε λοιπόν, προσεγγιστικά (και για μεγάλα n) να χρησιμοποιούμε την

$$\mathbb{P}(a < S_n \leq b) = \mathbb{P}\left(\frac{a - n\mu}{\sigma\sqrt{n}} < \frac{S_n - n\mu}{\sigma\sqrt{n}} \leq \frac{b - n\mu}{\sigma\sqrt{n}}\right) \approx \Phi\left(\frac{b - n\mu}{\sigma\sqrt{n}}\right) - \Phi\left(\frac{a - n\mu}{\sigma\sqrt{n}}\right).$$

Πόρισμα 6.2.3 (κεντρικό οριακό θεώρημα de Moivre–Laplace). *Αν S_n είναι το πλήθος των επιτυχιών σε μια ακολουθία n ανεξάρτητων δοκιμών Bernoulli με πιθανότητα επιτυχίας p, τότε, για κάθε $z \in \mathbb{R}$,*

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{S_n - np}{\sqrt{npq}} \leq z\right) = \Phi(z).$$

Παραδείγματα 6.2.4. (α) Η διάρκεια ζωής των ηλεκτρικών λαμπτήρων που παράγει ένα εργοστάσιο ακολουθεί την εκθετική κατανομή με μέση τιμή 200 ώρες. Κατά τον έλεγχο ενός δείγματος 49 λαμπτήρων, ποιά είναι η πιθανότητα το πολύ 14 λαμπτήρες να έχουν ζωή μικρότερη από 140 ώρες;

(β) Έστω (X_k) ακολουθία ανεξάρτητων και ισόνομων τυχαίων μεταβλητών που ακολουθούν την ομοιόμορφη κατανομή στο διάστημα $[-0.05, 0.05]$. Να υπολογιστεί κατά προσέγγιση η πιθανότητα $\mathbb{P}(|S_{100}| \leq 1)$.

6.3 (10 Ιουνίου 2009)

Συζητήθηκαν τα θέματα της εξέτασης της περιόδου Ιουνίου 2005.

6.4 (15 Ιουνίου 2009)

Συζητήθηκαν θέματα εξετάσεων των περιόδων Ιουνίου 2008 και Σεπτεμβρίου 2008.

6.5 (17 Ιουνίου 2009)

Συζητήθηκαν θέματα εξετάσεων από τις περιόδους Σεπτεμβρίου 2003, Σεπτεμβρίου 2005 και Ιουλίου 2007.