ΕΘΝΙΚΟ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ΤΜΗΜΑ ΜΑΘΗΜΑΤΙΚΩΝ
ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
« ΣΤΑΤΙΣΤΙΚΗ ΚΑΙ ΕΠΙΧΕΙΡΗΣΙΑΚΗ ΕΡΕΥΝΑ »

Εκτιμήσεις της Απόστασης Ολικής Κύμανσης και Εφαρμογές

ΒΑΣΙΛΙΚΗ ΘΕΟΔΩΡΟΠΟΥΛΟΥ

ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ: Ν. ΠΑΠΑΔΑΤΟΣ

ΑΘΗΝΑ, 2018

ΕΥΧΑΡΙΣΤΙΕΣ

Πρώτα απ' όλα θα ήθελα να ευχαριστήσω θερμά τον καθηγητή μου κ. Ν. Παπαδάτο για την ανάθεση αυτής της μεταπτυχιακής εργασίας. Τον ευχαριστώ επίσης για την πολύτιμη επιστημονική καθοδήγησή του, τις συμβουλές του και τη στήριξή του καθ' όλη τη διάρκεια της συνεργασίας μας. Οι γνώσεις του, η συνεχής παρακολούθηση της πορείας μου, αλλά και η διάθεσή του να βοηθήσει ουσιαστικά, έπαιξαν καθοριστικό ρόλο στην εκπόνηση της παρούσας μεταπτυχιακής εργασίας.

Τέλος, θα ήθελα να ευχαριστήσω την οικογένειά μου, για την κατανόηση, την υπομονή και τη συμπαράσταση που μου προσέφερε καθ' όλη τη διάρκεια των μεταπτυχιακών μου σπουδών.

Π Ε Ρ Ι Ε Χ Ο Μ Ε Ν Α

ΕΘΝΙΚΟ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ΤΜΗΜΑ ΜΑΘΗΜΑΤΙΚΩΝ

ΜΕΤΑΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

Βασιλική Θεοδωροπούλου

Επιβλέπων: Ν.Παπαδάτος, Καθηγητής Ε.Κ.Π.Α.

Μάιος 2018

Περίληψη

Σε αυτή την εργασία θα ασχοληθούμε με την απόσταση ολικής κύμανσης μεταξύ δύο κατανομών πιθανοτήτων.

Στο πρώτο κεφάλαιο αναφέρουμε κάποιες γενικές ιδιότητες και βασικά θεωρήματα.

Στο δεύτερο κεφάλαιο ασχολούμαστε με μια εφαρμογή της απόστασης ολικής κύμανσης στο πρόβλημα των συμπτώσεων. Επίσης μελετάμε μια νέα απόσταση που καλείται ¨Απόσταση Παραγοντικών Ροπών¨.

Στο τρίτο κεφάλαιο αναφερόμαστε στην προσέγγιση *Poisson*, που αποτελεί έναν από τους πιο σημαντικούς τομείς των Εφαρμοσμένων Πιθανοτήτων και πλήθος ερευνητών την έχει μελετήσει συστηματικά. Κάτι τέτοιο είναι εφικτό με μια εναλλακτική μεθοδο που βασίζεται στην έννοια των *w*-συναρτήσεων. Η εύρεση κατάλληλων άνω φραγμάτων (¨κοντά¨ στο 0) της απόστασης ολικής κύμανσης μεταξύ δύο κατανομών είναι ένας τρόπος μέτρησης του σφάλματος που προκύπτει από την προσέγγιση.

Στο τέταρτο κεφάλαιο αποδεικνύεται μια συνελιξιακή ανισότητα για το τυποποιημένο άθροισμα ανεξάρτητων απόλυτα συνεχών τυχαίων μεταβλητών. Με βάση την ανισότητα αυτή, που ισχύει με την προυπόθεση ότι το στήριγμα –support των τυχαίων μεταβλητών είναι διάστημα, αποδεικνύεται ότι η τάξη σύγκλισης κατά ολική κύμανση των τυποποιημένων μερικών αθροισμάτων $\frac{S_n - n\mu}{\sigma\sqrt{n}}$ των ανεξάρτητων και ισόνομων τυχαίων μεταβλητών $\mathrm{X}, \mathrm{X}_1, \mathrm{X}_2, \ldots$ προς την τυποποιημένη κανονική Z είναι τουλάχιστον $O\left(n^{-\frac{1}{2}}\right)$.

Επίσης δίνεται τιμή της σταθεράς c_X για την οποία $d_{TV}\left(\frac{S_n - n\mu}{\sigma\sqrt{n}}, Z\right) \le \frac{c_X}{\sqrt{n}}$ και εφαρμογές σε τυχαία αθροίσματα.

Στο πέμπτο κεφάλαιο αναφέρουμε ότι υπάρχει μια νέα τυχαία μεταβλητή X^*, η οποία θεωρείται ως μετασχηματισμός της X με μονοκόρυφη πυκνότητα, που ικανοποιεί την ταυτότητα συνδιακύμανσης

$$Cov[\mathrm{X}, g(\mathrm{X})] = \sigma^2 E[g'(\mathrm{X}^*)]$$

για κάθε απόλυτα συνεχής συνάρτηση g με παράγωγο g', δοθέντος ότι $E[g'(\mathrm{X}^*)] < \infty$.

Αναφέρουμε τις ιδιότητες της τυχαίας μεταβλητής X^* και εφαρμογή σε φράγματα διασποράς.

Στο έκτο κεφάλαιο παρουσιάζουμε άνω φράγματα για την απόσταση ολικής κύμανσης μεταξύ δύο απόλυτα συνεχών τυχαίων μεταβλητών X και Y.

Επίσης, δίνουμε μια απόδειξη του τοπικού οριακού θεωρήματος.

ΕΘΝΙΚΟ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ΤΜΗΜΑ ΜΑΘΗΜΑΤΙΚΩΝ

ΜΕΤΑΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

Βασιλική Θεοδωροπούλου

Επιβλέπων: Ν.Παπαδάτος, Καθηγητής Ε.Κ.Π.Α.

Μάιος 2018

Abstract

In this postgraduate thesis we study the total variation distance between two probability distributions.

In the first chapter we mention some general properties and basic theorems.

In the second chapter we give an application of the total variation distance to the matching problem. Also, we study a new distance which is called ¨Factorial Moment Distance¨.

In the third chapter we refer to the Poisson approach, which is one of the most important areas of Applies Probabilities and many researchers have studied it systematically. This is possible with an alternative method based on the meaning of *w*-functions. Finding appropriate upper bounds (near to 0) of the total variation distance between two distributions is a way of calculating the error which is derived from the approach.

In the fourth chapter is proved a convolution inequality for the standardized sum of independent absolutely continuous random variables.

Based on the inequality, provided that the support of the random variables is an interval,

is proved that the rate of convergence in total variation of the standardized sums $\frac{S_n - n\mu}{\sigma\sqrt{n}}$ of independent and identically distributed random variables $X, X_1, X_2, \ldots$ to the standard normal Z is of order, at least $O\left(n^{-\frac{1}{2}}\right)$. Also, it is given the value of the constant c_X such that $d_{TV}\left(\frac{S_n - n\mu}{\sigma\sqrt{n}}, Z\right) \leq \frac{c_X}{\sqrt{n}}$ and applications in random sums.

In the fifth chapter we mention that there exists a new random variable X^* which can be viewed as a transformation of X with a unimodal density, satisfying the covariance identity

$$Cov[X, g(X)] = \sigma^2 E[g'(X^*)]$$

for any absolutely continuous function g with derivative g', provided that $E[g'(X^*)] < \infty$. We mention the properties of the random variable X^* and an application to variance bounds.

In the sixth chapter upper bounds for the total variation distance between two absolutely continuous random variables X and Y are obtained. Also, we give a proof of the local limit theorem.

Κ Ε Φ Α Λ Α Ι Ο 1

Απόσταση Ολικής Κύμανσης

Υπάρχουν πολλοί τρόποι να ορίσουμε απόσταση μεταξύ δύο κατανομών πιθανοτήτων. Σε αυτή την εργασία θα ασχοληθούμε κυρίως με την απόσταση ολικής κύμανσης.

1.1. Βασικές Έννοιες-Ιδιότητες

Δίνονται οι ακόλουθοι ορισμοί:

Ορισμός 1.1 (Απόσταση Kolmogorov (Kolmogorov Distance))

Η απόσταση Kolmogorov μεταξύ των τυχαίων μεταβλητών Χ και Υ ορίζεται ως

$$d_K(\mathrm{X},\mathrm{Y}) = \sup_{x\in R}\left|P(\mathrm{X}\le x) - P(\mathrm{Y}\le x)\right| = \sup_{x}\left|F_1(x) - F_2(x)\right|,$$

όπου F_1 η συνάρτηση κατανομής της τυχαίας μεταβλητής Χ και F_2 η συνάρτηση κατανομής της τυχαίας μεταβλητής Υ. Η απόσταση Kolmogorov εστιάζεται στις ημιευθείες $\{(-\infty, x], x\in R\}$. Αν επιτρέψουμε το suprenum να λαμβάνεται ως προς όλα τα Borel υποσύνολα του R παίρνουμε την απόσταση ολικής κύμανσης:

Ορισμός 1.2 (Απόσταση Ολικής Κύμανσης (Total Variation Distance))

Η απόσταση ολικής κύμανσης μεταξύ των τυχαίων μεταβλητών Χ και Υ ορίζεται ως

$$d_{TV}(\mathrm{X},\mathrm{Y}) = \sup_{A\in B(R)}\left|P(\mathrm{X}\in \mathrm{A}) - P(\mathrm{Y}\in \mathrm{A})\right|,$$

όπου $P(\mathrm{X}\in \mathrm{A})$ το μέτρο πιθανότητας που επάγει η τυχαία μεταβλητή Χ στα Borel-σύνολα Α και $P(\mathrm{Y}\in \mathrm{A})$ το μέτρο πιθανότητας που επάγει η τυχαία μεταβλητή Υ.

Παρατήρηση 1.1

Οι αποστάσεις $d = d_K(\mathrm{X},\mathrm{Y})$ και $d = d_{TV}(\mathrm{X},\mathrm{Y})$ αποτελούν μετρικές πιθανοτήτων διότι πληρούν τις παρακάτω ιδιότητες:

i) $d(\mathrm{X},\mathrm{Y}) \ge 0$ για κάθε Χ, Υ και $d(\mathrm{X},\mathrm{Y}) = 0$ αν και μόνο αν $\mathrm{X} \overset{d}{=} \mathrm{Y}$.

ii) $d(\mathrm{X},\mathrm{Y}) = d(\mathrm{Y},\mathrm{X})$ (αντιμεταθετικότητα).

iii) $d(\mathrm{X},\mathrm{Y}) \le d(\mathrm{X},\mathrm{Z}) + d(\mathrm{Z},\mathrm{Y})$ (τριγωνική ανισότητα).

Παρατήρηση 1.2

Για τις αποστάσεις $d_K(\mathrm{X},\mathrm{Y})$ και $d_{TV}(\mathrm{X},\mathrm{Y}),$ που είναι suprenum των απόλυτων τιμών της διαφοράς δύο μέτρων πιθανοτήτων, ισχύει ότι:

$$0 \le d_K(\mathrm{X},\mathrm{Y}) \le d_{TV}(\mathrm{X},\mathrm{Y}) \le 1.$$

Παρατήρηση 1.3

Έστω X, Y τυχαίες μεταβλητές. Χρησιμοποώντας την οικογένεια των δεξιά κλειστών ημιευθειών (του R):

$D_1 = \{(-\infty, x],\ x \in R\} \subseteq B(R),$ προκύπτει ότι:

$$\begin{aligned} d_K(\mathrm{X},\mathrm{Y}) &= \sup_{x\in R}\left|P(\mathrm{X}\le x) - P(\mathrm{Y}\le x)\right| \\ &= \sup_{\mathrm{A}\in D_1}\left|P(\mathrm{X}\in \mathrm{A}) - P(\mathrm{Y}\in \mathrm{A})\right| \\ &\le \sup_{\mathrm{A}\in B(R)}\left|P(\mathrm{X}\in \mathrm{A}) - P(\mathrm{Y}\in \mathrm{A})\right| = d_{TV}(\mathrm{X},\mathrm{Y}). \end{aligned}$$

Επομένως, αν $\{\mathrm{X}_n,\ n\ge 1\}$ ακολουθία τυχαίων μεταβλητών και Y τυχαία μεταβλητή, από τα παραπάνω προκύπτει ότι: $d_{TV}(\mathrm{X}_n,\mathrm{Y}) \to 0 \Rightarrow d_K(\mathrm{X}_n,\mathrm{Y}) \to 0,$ καθώς $n\to\infty.$

Το αντίστροφο δεν ισχύει, όπως φαίνεται παρακάτω.

Παράδειγμα 1.1

Στο παράδειγμα αυτό πετυχαίνεται η χειρότερη περίπτωση απόστασης ολικής κύμανσης, δηλαδή ίση με την μονάδα, ενώ η απόσταση Kolmogorov τείνει στο 0, καθώς $n\to\infty.$

Εστω $\mathrm{X}_1, \mathrm{X}_2, \ldots, \mathrm{X}_n$~Bernoulli(p), $n = 1, 2, \ldots$ $(0 < p < 1).$

Χρησιμοποιώντας το Κεντρικό Οριακό Θεώρημα έχουμε ότι:

$$\mathrm{X}_n = \frac{\sqrt{n}(\overline{\mathrm{X}}_n - p)}{\sqrt{p(1-p)}} \xrightarrow{d} \Phi.$$

Εφαρμόζοντας το θεώρημα Pólya που λέει ότι αν $F_n \to F$ και F είναι συνεχής συνάρτηση κατανομής, τότε $\sup_x\left|F_n(x) - F(x)\right| \to 0,$ έχουμε ότι $d_K(\mathrm{X}_n, Z) = \sup_x\left|F_n(x) - \Phi(x)\right| \to 0,$ καθώς $n\to\infty,$ όπου Z η τυποποιημένη κανονική τυχαία μεταβλητή. Όμως

$$\mathrm{X}_n = \frac{\sqrt{n}(\overline{\mathrm{X}}_n - p)}{\sqrt{p(1-p)}} \in \left\{ \frac{\sqrt{n}(0-p)}{\sqrt{p(1-p)}}, \frac{\sqrt{n}(1/n-p)}{\sqrt{p(1-p)}}, \dots, \frac{\sqrt{n}(n/n-p)}{\sqrt{p(1-p)}} \right\} \overset{o\rho}{=} \mathrm{A}_n, \text{ αφού}$$

$$\overline{\mathrm{X}}_n \in \left\{0, \frac{1}{n}, \dots, \frac{n}{n}\right\}.$$

Επομένως, μπορώ να ορίσω το αριθμήσιμο σύνολο:

$\mathrm{A} = \bigcup_{n=1}^{\infty} \mathrm{A}_n = \bigcup_{n=1}^{\infty} \{x_0{}^n, x_1{}^n, \dots, x_n{}^n\}$, όπου $P(\mathrm{X}_n \in \mathrm{A}) = 1$, ενώ $P(Z \in \mathrm{A}) = 0$. Άρα:

$$d_{TV}(\mathrm{X}_n, Z) = \sup_{\mathrm{A} \in B(R)} \left| P(\mathrm{X}_n \in \mathrm{A}) - P(Z \in \mathrm{A}) \right| = 1 \quad \text{για κάθε} \quad n \geq 1.$$

Παρατήρηση 1.4

Η απόσταση ολικής κύμανσης $d_{TV}(\mathrm{X}, \mathrm{Y})$ έχει την ακόλουθη στατιστική ερμηνεία. Ας υποθέσουμε ότι θέλουμε να ελέγξουμε αν μια διακριτή τυχαία μεταβλητή Ζ ακολουθεί την κατανομή της Χ ή της Υ. Δηλαδή, έχουμε τον έλεγχο

$$H_0 : P(\mathrm{Z} \in \mathrm{A}) = P(\mathrm{X} \in \mathrm{A}) \text{ έναντι } H_1 : P(\mathrm{Z} \in \mathrm{A}) = P(\mathrm{Y} \in \mathrm{A}).$$

Έστω Κ η κρίσιμη περιοχή του ελέγχου (περιοχή απόρριψης της υπόθεσης H_0).

Τότε ισχύει

$$\begin{aligned} a &= P(\text{σφάλμα τύπου I}) \\ &= P(\text{ απόρριψη της } H_0 \mid \text{η } H_0 \text{ ισχύει}) \\ &= P_{H_0}(\mathrm{Z} \in \mathrm{K}) \end{aligned}$$

και

$$\begin{aligned} \beta &= P(\text{σφάλμα τύπου II}) \\ &= P(\text{ αποδοχή της } H_0 \mid \text{η } H_1 \text{ ισχύει}) \\ &= P_{H_1}(\mathrm{Z} \notin \mathrm{K}). \end{aligned}$$

Επομένως,

$$\begin{aligned} \alpha + \beta &= P_{H_0}(\mathrm{Z} \in \mathrm{K}) + P_{H_1}(\mathrm{Z} \notin \mathrm{K}) \\ &= P_{H_0}(\mathrm{Z} \in \mathrm{K}) + 1 - P_{H_1}(\mathrm{Z} \in \mathrm{K}) \\ &= 1 - \left| P_{H_1}(\mathrm{Z} \in \mathrm{K}) - P_{H_0}(\mathrm{Z} \in \mathrm{K}) \right|. \end{aligned}$$

(Η τελευταία ισότητα προκύπτει από το γεγονός ότι για την κρίσιμη περιοχή Κ, έχουμε $P_{H_1}(\mathrm{Z} \in \mathrm{K}) \geq P_{H_0}(\mathrm{Z} \in \mathrm{K})$). Παίρνοντας το inf του αθροίσματος $\alpha+\beta$ για όλες τις δυνατές επιλογές της κρίσιμης περιοχής Κ, έχουμε

$$\inf_{\mathrm{K}}(a+\beta) = 1 - \sup_{\mathrm{K}}\left|P_{H_1}(\mathrm{Z} \in \mathrm{K}) - P_{H_0}(\mathrm{Z} \in \mathrm{K})\right|$$

$$= 1 - d_{TV}(\mathrm{X}, \mathrm{Y}).$$

Τέλος, αφού $0 \leq d_{TV}(\mathrm{X}, \mathrm{Y}) \leq 1$, είναι προφανές ότι $0 \leq \inf_{\mathrm{K}}(\alpha+\beta) \leq 1$.

1.2. Βασικά Θεωρήματα

Θεώρημα 1.1

Αν οι τυχαίες μεταβλητές Χ και Υ έχουν πυκνότητες f_1, f_2 (ως προς το μέτρο Lebesgue) τότε:

$$d_{TV}(\mathrm{X}, \mathrm{Y}) = \frac{1}{2}\int_{-\infty}^{+\infty}\left|f_1(x) - f_2(x)\right|dx = \int_{\{x: f_1(x) > f_2(x)\}}(f_1(x) - f_2(x))dx = \int_{\{x: f_2(x) > f_1(x)\}}(f_2(x) - f_1(x))dx$$

και $d_{TV}(\mathrm{X}, \mathrm{Y}) = 1 - \int_{-\infty}^{+\infty}\min\{f_1(x), f_2(x)\}dx = -1 + \int_{-\infty}^{+\infty}\max\{f_1(x), f_2(x)\}dx.$

Απόδειξη:

Θέτουμε $\mathrm{A}_+ = \{x \in R : f_1(x) > f_2(x)\}$ και $\mathrm{A}_- = \{x \in R : f_2(x) > f_1(x)\}$. Τότε για τυχόν $\mathrm{A} \in B(R)$:

$$P(\mathrm{X} \in \mathrm{A}) - P(\mathrm{Y} \in \mathrm{A}) = \int_{\mathrm{A}}(f_1(x) - f_2(x))dx$$

$$= \int_{(\mathrm{A}\cap\mathrm{A}_+)\cup(\mathrm{A}\cap\mathrm{A}_-)}(f_1(x) - f_2(x))dx$$

$$= \int_{\{x\in\mathrm{A}: f_1(x) > f_2(x)\}\bigcup\{x\in\mathrm{A}: f_2(x) \geq f_1(x)\}}(f_1(x) - f_2(x))dx$$

$$= \int_{\{x\in\mathrm{A}: f_1(x) > f_2(x)\}}(f_1(x) - f_2(x))dx + \int_{\{x\in\mathrm{A}: f_2(x) \geq f_1(x)\}}(f_1(x) - f_2(x))dx$$

$$\leq \int_{\{x\in\mathrm{A}: f_1(x) > f_2(x)\}}(f_1(x) - f_2(x))dx \leq \int_{\{x\in R: f_1(x) > f_2(x)\}}(f_1(x) - f_2(x))dx. \qquad (1.1)$$

Ομοίως

$$P(\mathrm{Y}\in \mathrm{A})-P(\mathrm{X}\in \mathrm{A})=\int_{\mathrm{A}}(f_2(x)-f_1(x))dx$$

$$=\int_{\{x\in \mathrm{A}: f_1(x)\geq f_2(x)\}\bigcup\{x\in \mathrm{A}: f_2(x)>f_1(x)\}}(f_2(x)-f_1(x))dx$$

$$=\int_{\{x\in \mathrm{A}: f_1(x)\geq f_2(x)\}}(f_2(x)-f_1(x))dx+\int_{\{x\in \mathrm{A}: f_2(x)>f_1(x)\}}(f_2(x)-f_1(x))dx$$

$$\leq\int_{\{x\in \mathrm{A}: f_2(x)>f_1(x)\}}(f_2(x)-f_1(x))dx\leq\int_{\{x\in R: f_2(x)>f_1(x)\}}(f_2(x)-f_1(x))dx. \qquad (1.2)$$

Ισχύει ότι, για οποιοδήποτε $\mathrm{A}\in B(R)$,

$$\int_R(f_1(x)-f_2(x))dx=0$$

$$\Rightarrow\int_{\mathrm{A}}(f_1(x)-f_2(x))dx+\int_{\mathrm{A}^c}(f_1(x)-f_2(x))dx=0$$

$$\Rightarrow\int_{\mathrm{A}}(f_1(x)-f_2(x)dx=-\int_{\mathrm{A}^c}(f_1(x)-f_2(x))dx.$$

Από τις σχέσεις (1.1) και (1.2) προκύπτει ότι:

$$\left|P(\mathrm{X}\in \mathrm{A})-P(\mathrm{Y}\in \mathrm{A})\right|\leq\int_{\{x\in \mathrm{A}: f_1(x)>f_2(x)\}}(f_1(x)-f_2(x))dx.$$

Επομένως, $d_{TV}(\mathrm{X},\mathrm{Y})=\sup\limits_{\mathrm{A}\in B(R)}\left|P(\mathrm{X}\in \mathrm{A})-P(\mathrm{Y}\in \mathrm{A})\right|$

$$=\int_{\{x\in R: f_1(x)>f_2(x)\}}(f_1(x)-f_2(x))dx$$

$$=-\int_{\{x\in R: f_1(x)\leq f_2(x)\}}(f_1(x)-f_2(x))dx$$

$$=\int_{\{x\in R: f_1(x)\leq f_2(x)\}}(f_2(x)-f_1(x))dx.$$

Το πρώτο μέλος της ισότητας αποδεικνύεται ως εξής:

$$\int_{-\infty}^{+\infty}\left|f_1(x)-f_2(x)\right|dx=\int_{\{x: f_1(x)>f_2(x)\}}(f_1(x)-f_2(x))dx+\int_{\{x: f_2(x)\geq f_1(x)\}}(f_2(x)-f_1(x))dx$$

$$=\int_{\{x: f_1(x)>f_2(x)\}}(f_1(x)-f_2(x))dx-\int_{\{x: f_2(x)\geq f_1(x)\}^c}(f_2(x)-f_1(x))dx$$

$$=\int_{\{x: f_1(x)>f_2(x)\}}(f_1(x)-f_2(x))dx-\int_{\{x: f_1(x)>f_2(x)\}}(f_2(x)-f_1(x))dx$$

$$=\int_{\{x: f_1(x)>f_2(x)\}}(f_1(x)-f_2(x))dx+\int_{\{x: f_1(x)>f_2(x)\}}(f_1(x)-f_2(x))dx$$

$$=2\int_{\{x: f_1(x)>f_2(x)\}}(f_1(x)-f_2(x))dx=2d_{TV}(\mathrm{X},\mathrm{Y}).$$

Άρα $d_{TV}(\mathrm{X},\mathrm{Y}) = \frac{1}{2}\int_{-\infty}^{+\infty} |f_1(x) - f_2(x)| dx.$

Για την δεύτερη σχέση του θεωρήματος, παρατηρούμε ότι για κάθε $x \in R$,

$$\max\{f_1(x), f_2(x)\} + \min\{f_1(x), f_2(x)\} = f_1(x) + f_2(x)$$

και

$$\max\{f_1(x), f_2(x)\} - \min\{f_1(x), f_2(x)\} = |f_1(x) - f_2(x)|.$$

Αφαιρώντας και προσθέτοντας κατά μέλη αυτές τις δύο ισότητες, προκύπτει ότι

$$2\min\{f_1(x), f_2(x)\} = f_1(x) + f_2(x) - |f_1(x) - f_2(x)|$$

και $$2\max\{f_1(x), f_2(x)\} = f_1(x) + f_2(x) + |f_1(x) - f_2(x)|,$$

αντίστοιχα. Ολοκληρώνοντας τώρα καθεμία από τις παραπάνω σχέσεις ως προς x και λαμβάνοντας υπόψη ότι

$$\int_{-\infty}^{+\infty} f_1(x)dx = 1, \quad \int_{-\infty}^{+\infty} f_2(x)dx = 1 \text{ και } d_{TV}(\mathrm{X},\mathrm{Y}) = \frac{1}{2}\int_{-\infty}^{+\infty} |f_1 x) - f_2(x)| dx,$$

ολοκληρώνουμε την απόδειξη.

Θεώρημα 1.2

Για τις διακριτές τυχαίες μεταβλητές X και Y με τιμές στο $N = \{0,1,\ldots\}$ και τις αντίστοιχες συναρτήσεις πιθανότητας p και q ισχύει:

$$d_{TV}(\mathrm{X},\mathrm{Y}) = \frac{1}{2}\sum_{x \in N} |p(x) - q(x)| = \sum\nolimits_{\{x:p(x)>q(x)\}} (p(x) - q(x)) = \sum\nolimits_{\{x:q(x)>p(x)\}} (p(x) - q(x)) \text{ και}$$

$$d_{TV}(\mathrm{X},\mathrm{Y}) = 1 - \sum_{x \in N} \min\{p(x), q(x)\} = -1 + \sum_{x \in N} \max\{p(x), q(x)\}.$$

Απόδειξη

Όμοια με την περίπτωση των συνεχών τυχαίων μεταβλητών.

Παράδειγμα 1.2

Ας θεωρήσουμε τις τυχαίες μεταβλητές X, Y με τιμές στο $\{0,1,2\}$ και συναρτήσεις πιθανότητας p και q αντίστοιχα, τέτοιες ώστε

$$p(0) = P(\mathrm{X}=0) = \frac{1}{3}, \quad p(1) = P(\mathrm{X}=1) = \frac{1}{3}, \quad p(2) = P(\mathrm{X}=2) = \frac{1}{3}$$

και

$$q(0) = P(\mathrm{Y}=0) = \frac{1}{4}, \; q(1) = P(\mathrm{Y}=1) = \frac{1}{2}, \; q(2) = P(\mathrm{Y}=2) = \frac{1}{4}.$$

Από τα δεδομένα είναι φανερό ότι $\mathrm{A}_+ = \{x \in N : p(x) > q(x)\} = \{0,2\}$ και $\mathrm{A}_- = \{x \in N : p(x) < q(x)\} = \{1\}$. Έτσι λοιπόν, η απόσταση ολικής κύμανσης είναι

$$d_{TV}(\mathrm{X},\mathrm{Y}) = \sum_{x \in \mathrm{A}_+} (p(x) - q(x)) = (p(0) - q(0)) + (p(2) - q(2)) = \frac{1}{12} + \frac{1}{12} = \frac{1}{6}.$$

$$\left(d_{TV}(\mathrm{X},\mathrm{Y}) = \sum_{x \in \mathrm{A}_-} (q(x) - p(x)) = q(1) - p(1) = \frac{1}{6} \right).$$

Κ Ε Φ Α Λ Α Ι Ο 2

Το Πρόβλημα των Συμπτώσεων

Σε αυτό το κεφάλαιο θα ασχοληθούμε με μια εφαρμογή της απόστασης ολικής κύμανσης στο πρόβλημα των συμπτώσεων. Επίσης θα μελετήσουμε μια νέα απόσταση που καλείται ¨Απόσταση Παραγοντικών Ροπών¨.

2.1. Κλασσικό Πρόβλημα Συμπτώσεων

Από μια κάλπη που περιέχει n σφαιρίδια αριθμημένα από το 1 μέχρι το n εξάγουμε το ένα μετά το άλλο χωρίς επανάθεση όλα τα σφαιρίδια. Η εξαγωγή του j σφαιριδίου κατά την j δοκιμή, j=1,2,…,n καλείται συνάντηση (σύμπτωση). Να υπολογιστεί η πιθανότητα να πραγματοποιηθούν ακριβώς k συναντήσεις στις n δοκιμές.

Παρατηρούμε ότι σε κάθε εξαγωγή και των n σφαιριδίων αντιστοιχεί μια μετάθεση των n θετικών ακεραίων $\{1,2,\ldots,n\}$. Επομένως, ο αριθμός των δειγματικών σημείων του δειγματικού χώρου Ω του τυχαίου πειράματος της εξαγωγής και των n σφαιριδίων είναι ίσος με $n!$, τον αριθμό των μεταθέσεων των n αριθμών $\{1,2,\ldots,n\}$. Έστω A_i το ενδεχόμενο να πραγματοποιηθεί συνάντηση (σύμπτωση) στην i δοκιμή, i=1,2,…,n. Ας θεωρήσουμε μια οποιαδήποτε συλλογή r δεικτών (σφαιριδίων) $\{i_1, i_2, \ldots, i_r\}$ από το σύνολο n δεικτών (σφαιριδίων) $\{1,2,\ldots,n\}$. Τότε

$$p_r = P(\mathrm{A}_{i_1}\mathrm{A}_{i_2}\cdots\mathrm{A}_{i_r}) = \frac{(n-r)!}{n!}, \quad r=1,2,\ldots,n$$

επειδή οι $n-r$ υπόλοιποι αριθμοί (σφαιρίδια), $\{1,2,\ldots,n\}-\{i_1, i_2, \ldots, i_r\}$, μπορούν να εξαχθούν σε οποιαδήποτε από τις υπόλοιπες $n-r$ δοκιμές κατά $(n-r)!$ τρόπους. Επομένως τα ενδεχόμενα A_i, i=1,2,…,n είναι ανταλλάξιμα. Εισάγοντας τις πιθανότητες p_r, r=1,2,…,n στη σχέση $P_{n,k} = \binom{n}{k}\sum_{r=0}^{n-k}(-1)^r\binom{n-k}{r}p_{k+r}$, συνάγουμε ότι η πιθανότητα $P_{n,k}$ να πραγματοποιηθούν ακριβώς k συναντήσεις στις n δοκιμές είναι ίση με

$$P_{n,k} = \frac{1}{k!}\sum_{r=0}^{n-k}\frac{(-1)^r}{r!}, \qquad k=0,1,2,\ldots,n.$$

Το κλασικό αυτό πρόβλημα των συναντήσεων (συμπτώσεων) αποδίδεται στον Montmort (1708) και μπορεί να περιγραφεί κατά διάφορους τρόπους.

Χαρακτηριστικά είναι τα επόμενα παραδείγματα.

α) Ας θεωρήσουμε μια δεσμίδα n παιγνιοχάρτων τα οποία είναι τοποθετημένα σε μια οποιοδήποτε διάταξη το ένα δίπλα στο άλλο, και ας τοποθετήσουμε τα n παιγνιόχαρτα μιας άλλης όμοιας δεσμίδας σε μια τυχαία διάταξη κάτω από τα παιγνιόχαρτα της πρώτης δεσμίδας. Το ενδεχόμενο ένα παιγνιόχαρτο να καταλάβει την ίδια θέση και στις δύο διατάξεις των δεσμίδων είναι μια συνάντηση.

β) Ας θεωρήσουμε την τυχαία τοποθέτηση n διακεκριμένων επιστολών μέσα σε n διακεκριμένους φακέλους κατά την οποία σε κάθε φάκελο τοποθετείται μόνο μία αποστολή. Το ενδεχόμενο η j επιστολή να τοποθετηθεί στον j φάκελο είναι μια συνάντηση.

Η μαθηματική διατύπωση του προβλήματος:

Έστω $\pi_n = (\pi_n(1), \ldots, \pi_n(n))$ μια τυχαία μετάθεση του $T_n=\{1,\ldots,n\}$ και π_n είναι ομοιόμορφα κατανεμημένη πάνω στις $n!$ μεταθέσεις του T_n. Ο αριθμός j καλείται σταθερό σημείο της π_n αν $\pi_n(j) = j$. Ορίζουμε ως Z_n τον συνολικό αριθμό των σταθερών σημείων της π_n, $Z_n = \sum_{j=1}^{n} \mathrm{I}\{\pi_n(j) = j\}$, όπου Ι δείκτρια συνάρτηση. Η μελέτη του Z_n αντιστοιχεί στο γνωστό πρόβλημα συμπτώσεων.

Προφανώς, η Z_n μπορεί να πάρει τις τιμές $0,1,\ldots,n-2,n,$ και η κατανομή της είναι

$$P(Z_n = j) = \frac{1}{j!}\sum_{k=0}^{n-j}\frac{(-1)^k}{k!}, \quad j=0,1,\ldots,n-2,n.$$

Είναι φανερό ότι η Z_n συγκλίνει στη Z που ακολουθεί την Poisson(1) κατανομή.

Υπενθυμίζουμε ότι η απόσταση ολικής κύμανσης μεταξύ δύο τυχαίων μεταβλητών X_1 και X_2 ορίζεται ως

$$d_{TV}(\mathrm{X}_1, \mathrm{X}_2) = \sup_{\mathrm{A}\in B(R)} \left| P(\mathrm{X}_1 \in \mathrm{A}) - P(\mathrm{X}_2 \in \mathrm{A}) \right|,$$

όπου $B(R)$ η Borel σ-άλγεβρα του R.

Ο Diaconis απέδειξε ότι $d_{TV}(Z_n, Z) \le \frac{2^n}{n!}$. Το φράγμα αυτό βελτιώθηκε από τον DasGupta:

$$d_{TV}(Z_n, Z) \le \frac{2^n}{(n+1)!}.$$

Όπως θα δειχθεί παρακάτω (Θεώρημα 2.3), ισχύει η σχέση $d_{TV}(Z_n, Z) \sim \frac{2^n}{(n+1)!}$, όπου

$\alpha_n \sim b_n$ σημαίνει ότι $\lim_n \frac{\alpha_n}{b_n} = 1.$

Θεωρούμε τώρα τα σύνολα των διακριτών τυχαίων μεταβλητών

$$D_n := \{X : P(X \in \{0,1,\dots,n\}) = 1\},$$

$$D_\infty := \{X : P(X \in \{0,1,\dots\}) = 1\}.$$

Ένα εύλογο ερώτημα είναι το εξής:

¨ Είναι αλήθεια ότι $\min_{X \in D_n} d_{TV}(X, Z) \sim d_{TV}(Z_n, Z);$ ¨

Πρόταση 2.1

Έστω $Z \sim Poisson(1)$. Τότε

$$\min_{X \in D_n} \{d_{TV}(X, Z)\} = 1 - e^{-1} \sum_{j=0}^{n} \frac{1}{j!} \sim \frac{e^{-1}}{(n+1)!}. \tag{2.1}$$

Απόδειξη

Για κάθε $X_1, X_2 \in D_\infty$ με συναρτήσεις μάζας πιθανότητας p_1 και p_2 αντίστοιχα, η απόσταση ολικής κύμανσης γράφεται ως

$$d_{TV}(X_1, X_2) = \frac{1}{2} \sum_{j=0}^{\infty} |p_1(j) - p_2(j)| = \sum_{j=0}^{\infty} (p_1(j) - p_2(j))^+,$$

όπου $x^+ = \max\{x, 0\}$. Επομένως, για κάθε $X_1 \in D_n$ θα είναι $p_1(j) = 0$ για όλα τα $j > n$, και συνεπώς, αφού $(p_1(j) - p_2(j))^+ = 0$ για $j > n$,

$$d_{TV}(X_1, X_2) = \sum_{j=0}^{n} (p_1(j) - p_2(j))^+ \ge \sum_{j=0}^{n} (p_1(j) - p_2(j)) = 1 - \sum_{j=0}^{n} p_2(j) = P(X_2 > n).$$

Έχουμε ισότητα αν και μόνο αν $p_1(j) \ge p_2(j)$, $j = 0,1,\dots,n$.

Εφαρμόζοντας την προηγούμενη ανισότητα για $p_2(j)=P(Z=j)=\dfrac{e^{-1}}{j!}$, παίρνουμε την ισότητα στην (2.1) και το ελάχιστο επιτυγχάνεται για κάθε τυχαία μεταβλητή $\mathrm{X}\in D_n$ με

$$P(\mathrm{X}=j)\geq \frac{e^{-1}}{j!},\ \ j=0,1,\ldots,n.$$

Τώρα θα δείξουμε ότι η ποσότητα $1-e^{-1}\sum_{j=0}^{n}\dfrac{1}{j!}$ είναι τάξεως $\dfrac{e^{-1}}{(n+1)!}$ για $n\to\infty$.

Το υπόλοιπο κατά Cauchy στο ανάπτυγμα Taylor είναι

$$f(x)-\sum_{j=0}^{n}\frac{f^{(j)}(0)}{j!}x^j=\frac{1}{n!}\int_0^x (x-y)^n f^{(n+1)}(y)dy. \qquad (2.2)$$

Αν θέσουμε $f(x)=e^x$ παίρνουμε

$$e^x-\sum_{j=0}^{n}\frac{x^j}{j!}=\frac{1}{n!}\int_0^x (x-y)^n e^y dy$$

και για $x=1$,

$$e-\sum_{j=0}^{n}\frac{1}{j!}=\frac{1}{n!}\int_0^1 (1-y)^n e^y dy$$

$$\Rightarrow 1-e^{-1}\sum_{j=0}^{n}\frac{1}{j!}=e^{-1}\left(e-\sum_{j=0}^{n}\frac{1}{j!}\right)=\frac{e^{-1}}{n!}\int_0^1 (1-y)^n e^y dy$$

Από τις ανισότητες $1<e^y<1+(e-1)y,\ \ 0<y<1$ αν τις πολλαπλασιάσουμε με $(1-y)^n$ και τις ολοκληρώσουμε ως προς y έχουμε

$$\int_0^1 (1-y)^n dy<\int_0^1 (1-y)^n e^y dy<\int_0^1 (1-y)^n dy+(e-1)\int_0^1 y(1-y)^n dy.$$

Αφού $\displaystyle\int_0^1 (1-y)^n dy=\left[-\frac{(1-y)^{n+1}}{n+1}\right]_0^1=\frac{1}{n+1}$ και

$$\int_0^1 y(1-y)^n dy=\int_0^1 y\left(-\frac{(1-y)^{n+1}}{n+1}\right)' dy$$

$$=\left[-y\frac{(1-y)^{n+1}}{n+1}\right]_0^1-\int_0^1\left(-\frac{(1-y)^{n+1}}{n+1}\right)dy$$

$$= \frac{1}{n+1}\int_0^1 (1-y)^{n+1}dy =$$

$$= \frac{1}{n+1}\left[-\frac{(1-y)^{n+2}}{n+2}\right]_0^1 = \frac{1}{(n+1)(n+2)}$$

προκύπτει ότι

$$\frac{1}{n+1} < \int_0^1 (1-y)^n e^y dy < \frac{1}{n+1} + (e-1)\frac{1}{(n+1)(n+2)}$$

ή $$\frac{1}{n+1} < \int_0^1 (1-y)^n e^y dy < \frac{1}{n+1}\left(1+\frac{e-1}{n+2}\right).$$

Πολλαπλασιάζοντας με το e^{-1} και διαιρώντας με το $n!$ προκύπτει ότι

$$\frac{e^{-1}}{(n+1)!} < \min_{\mathrm{X}\in D_n}\{d_{TV}(\mathrm{X},Z)\} = 1 - e^{-1}\sum_{j=0}^{n}\frac{1}{j!} < \frac{e^{-1}}{(n+1)!}\left(1+\frac{e-1}{n+2}\right),$$

και επομένως $\min_{\mathrm{X}\in D_n}\{d_{TV}(\mathrm{X},Z)\} \sim \frac{e^{-1}}{(n+1)!}$.

Το προηγούμενο αποτέλεσμα δείχνει ότι η κατανομή της Z_n, παρότι είναι εξαιρετικά κοντά σε αυτήν της $\mathrm{Z} \sim Poisson(1)$, δεν πετυχαίνει την βέλτιστη τάξη σύγκλισης, αφού προφανώς

$$\lim_n \frac{\frac{e^{-1}}{(n+1)!}}{\frac{2^n}{(n+1)!}} = \lim_n \frac{e^{-1}}{2^n} = 0.$$

2.2. Απόσταση Παραγοντικών Ροπών

Ξεκινάμε με την ακόλουθη παρατήρηση:

Για τις τυχαίες μεταβλητές Z και Z_n, $E(Z)_k = 1$ και $E(Z_n)_k = \mathrm{I}_{\{k \le n\}}$, k=0,1,…,

όπου $E(\mathrm{X})_k$ η k-τάξης καθοδική παραγοντική ροπή της τυχαίας μεταβλητής X

(για κάθε $x \in R$, $(x)_0 = 1$ και $(x)_k = x(x-1)\cdots(x-k+1)$, k=1,2,…).

Η απόσταση παραγοντικών ροπών ορίζεται σε μια κατάλληλη υποκλάση των διακριτών τυχαίων μεταβλητών. Συγκεκριμένα, για κάθε $t \ge 0$, ορίζουμε

$$\mathcal{X}(t) := \{\mathrm{X} \in D_\infty : \text{υπάρχει } t' > t \text{ τέτοιο ώστε } P_{\mathrm{X}}(1+t') < \infty\},$$

όπου $P_{\mathrm{X}}(u) = E(u^{\mathrm{X}})$ η πιθανογεννήτρια συνάρτηση της X.

Επίσης, ορίζουμε

$$\mathcal{X}(\infty) := \bigcap_{t \in [0,\infty)} \mathcal{X}(t) = \{\mathrm{X} \in D_\infty : P_{\mathrm{X}}(1+t') < \infty \text{ για κάθε } t' > 0\}.$$

Ορισμός 2.1

A) Έστω $\mathrm{X}_1, \mathrm{X}_2 \in \mathcal{X}(0)$. Για $\alpha > 0$, ορίζουμε την απόσταση παραγοντικών ροπών τάξης α των $\mathrm{X}_1, \mathrm{X}_2$ ως

$$d_\alpha(\mathrm{X}_1, \mathrm{X}_2) := \sum_{k=1}^{\infty} \frac{a^{k-1}}{k!} \left| E(\mathrm{X}_1)_k - E(\mathrm{X}_2)_k \right|. \tag{2.3}$$

B) Έστω $\mathrm{X} \in \mathcal{X}(0)$ και $\{\mathrm{X}_n\}_{n=1}^{\infty} \subset \mathcal{X}(0)$. Λέμε ότι η X_n συγκλίνει κατά απόσταση παραγοντικών ροπών τάξης α στο X, συμβολικά $\mathrm{X}_n \xrightarrow{\alpha} \mathrm{X}$ αν $d_\alpha(\mathrm{X}_n, \mathrm{X}) \to 0$ καθώς $n \to \infty$.

Λήμμα 2.1

Αν $\mathrm{X} \in \mathcal{X}(1)$, τότε η συνάρτηση πιθανότητας p μπορεί να γραφτεί ως

$$p(j) = \sum_{k=j}^{\infty} \frac{(-1)^{k-j}}{k!} \binom{k}{j} E(\mathrm{X})_k, \quad j=0,1,\ldots. \tag{2.4}$$

Απόδειξη

Από την υπόθεση $\mathrm{X} \in \mathcal{X}(1)$, μπορούμε να βρούμε έναν αριθμό $t' > 1$ τέτοιο ώστε $E\left[(1+t')^{\mathrm{X}}\right] = \sum_{j=0}^{\infty} (1+t')^j p(j) < \infty$. Επειδή Χ είναι μη-αρνητική, η πιθανογεννήτρια συνάρτηση έχει ανάπτυγμα Taylor γύρω από το 0 με ακτίνα σύγκλισης $R' \geq 1 + t' > 2$,

$$P(u) = \sum_{j=0}^{\infty} u^j p(j) \in R, \quad |u| < R.$$

Είναι γνωστό ότι $\left.\frac{d^k}{du^k} P(u)\right|_{u=1} = E(\mathrm{X})_k$ και επειδή P έχει ανάπτυγμα Taylor γύρω από το 1 με ακτίνα σύγκλισης $R' \geq t' > 1$, έχουμε

$$P(u) = \sum_{k=0}^{\infty} \frac{E(\mathrm{X})_k}{k!} (u-1)^k, \quad |u-1| < R'.$$

Επομένως, από την προηγούμενη σχέση και επειδή $0 \in (1-R', 1+R')$ έχουμε

$$P(j) = \frac{1}{j!} \frac{d^j}{du^j} P(u)\bigg|_{u=0} = \frac{1}{j!} \sum_{k=j}^{\infty} \frac{(u-1)^{k-j}}{(k-j)!} E(\mathrm{X})_k \bigg|_{u=0} = \sum_{k=j}^{\infty} \frac{(-1)^{k-j}}{j!(k-j)!} E(\mathrm{X})_k.$$

Θεώρημα 2.1

Αν $\mathrm{X}_1, \mathrm{X}_2 \in \mathcal{X}(1)$, τότε $d_{TV}(\mathrm{X}_1, \mathrm{X}_2) \leq d_\alpha(\mathrm{X}_1, \mathrm{X}_2)$ για κάθε $\alpha \geq 2$. (2.5)

Απόδειξη

$$d_{TV}(\mathrm{X}_1, \mathrm{X}_2) = \frac{1}{2} \sum_{j=0}^{\infty} |p_1(j) - p_2(j)| = \frac{1}{2} \sum_{j=0}^{\infty} \left| \sum_{k=j}^{\infty} \frac{(-1)^{k-j}}{k!} \binom{k}{j} \left(E(\mathrm{X}_1)_k - E(\mathrm{X}_2)_k\right) \right|$$

$$\leq \frac{1}{2} \sum_{j=0}^{\infty} \sum_{k=j}^{\infty} \frac{1}{k!} \binom{k}{j} |E(\mathrm{X}_1)_k - E(\mathrm{X}_2)_k|.$$

Εναλλάσσοντας τη σειρά των αθροίσεων σύμφωνα με το Θεώρημα του Tonelli, έχουμε

$$d_{TV}(\mathrm{X}_1, \mathrm{X}_2) \leq \sum_{k=0}^{\infty} \sum_{j=0}^{k} \frac{1}{k!} \binom{k}{j} |E(\mathrm{X}_1)_k - E(\mathrm{X}_2)_k|$$

$$= \sum_{k=0}^{\infty} \frac{2^{k-1}}{k!} |E(\mathrm{X}_1)_k - E(\mathrm{X}_2)_k|.$$

Η απόδειξη ολοκληρώνεται από το γεγονός ότι $E(\mathrm{X}_1)_0 = E(\mathrm{X}_2)_0 = 1$.

Επομένως η d_α-σύγκλιση (στο $\mathcal{X}(1)$) εξασφαλίζει και τη σύγκλιση κατά ολική κύμανση και δίνει πολύ καλά φράγματα για τον ρυθμό σύγκλισης κατά ολική κύμανση.

2.3. Εφαρμογή στο Γενικευμένο Πρόβλημα Συμπτώσεων

Θεωρούμε το κλασσικό πρόβλημα συμπτώσεων, όπου, τώρα, καταγράφουμε μόνο το ποσοστό των συμπτώσεων, εξαιτίας ενός τυχαίου μηχανισμού. Ο μηχανισμός αποφασίζει ανεξάρτητα σε κάθε σύμπτωση. Ειδικότερα, όταν μια σύμπτωση συμβαίνει, ο μηχανισμός την μετράει με πιθανότητα λ, ανεξάρτητα από τις άλλες συμπτώσεις, και την αγνοεί με πιθανότητα $1-\lambda$, όπου $0<\lambda\le 1$. Εμείς ενδιαφερόμαστε για τον αριθμό $Z_n(\lambda)$ των καταμετρημένων συμπτώσεων.

Η περίπτωση $\lambda=1$ αντιστοιχεί στο κλασσικό πρόβλημα συμπτώσεων, όπου καταγράφονται όλες οι συμπτώσεις, έτσι ώστε $Z_n = Z_n(1)$.

Η μαθηματική διατύπωση του προβλήματος:

Έστω $\pi_n = (\pi_n(1),\ldots,\pi_n(n))$ μια τυχαία μετάθεση του $\{1,\ldots,n\}$, όπως και στην αρχή.

Έστω, επίσης, $J_1(\lambda),\ldots,J_n(\lambda)$ ανεξάρτητες, ισόνομες τυχαίες μεταβλητές που ακολουθούν την $Bernoulli(\lambda)$ και είναι ανεξάρτητες από την π_n.

Ο αριθμός $Z_n(\lambda)$ των καταγεγραμμένων συμπτώσεων μπορεί να γραφτεί ως

$$Z_n(\lambda) = \sum_{i=1}^{n} J_i(\lambda) I\{\pi_n(i)=i\}.$$

Έστω $A_i = \{J_i(\lambda)=1\}$, $B_i = \{\pi_n(i)=i\}$, $E_i = A_i \bigcap B_i$, $i=1,\ldots,n$.

Τότε $Z_n(\lambda)$ είναι ο αριθμός των γεγονότων E που θα συμβούν.

$$P(Z_n(\lambda)=j) = P(\text{ακριβώς } j \text{ μεταξύ } E_1,\ldots,E_n \text{ συμβαίνουν})$$

$$= \sum_{i=j}^{n} (-1)^{i-j} \binom{i}{j} S_{i,n},$$

όπου $S_{0,n}=1$, $S_{i,n} = \sum_{1\le k_1<\ldots<k_i<n} P(E_{k_1}\bigcap\ldots\bigcap E_{k_i})$, $i=1,\ldots,n$.

Επειδή τα A είναι ανεξάρτητα από τα B έχουμε

$$P(E_{k_1}\bigcap\ldots\bigcap E_{k_i}) = P(A_{k_1}\bigcap\ldots\bigcap A_{k_i})\, P(B_{k_1}\bigcap\ldots\bigcap B_{k_i}) = \lambda^i \frac{(n-i)!}{n!},$$

έτσι ώστε

$$S_{i,n}=\binom{n}{i}\lambda^i\frac{(n-i)!}{n!}=\frac{n!}{i!(n-i)!}\lambda^i\frac{(n-i)!}{n!}=\frac{\lambda^i}{i!},\quad i=0,1,\dots,n.$$

Τότε η συνάρτηση πιθανότητας της $Z_n(\lambda)$ δίνεται από τον τύπο

$$\begin{aligned}
p_{n;\lambda}(j):=P(Z_n(\lambda)=j)&=\sum_{i=j}^{n}(-1)^{i-j}\binom{i}{j}\frac{\lambda^i}{i!}\\
&=\sum_{i=j}^{n}(-1)^{i-j}\frac{i!}{j!(i-j)!}\frac{\lambda^i}{i!}\\
&=\sum_{i=j}^{n}(-1)^{i-j}\frac{\lambda^i}{j!(i-j)!}\\
&=\frac{1}{j!}\sum_{i=j}^{n}(-1)^{i-j}\frac{\lambda^i}{(i-j)!}\\
&=\frac{\lambda^j}{j!}\sum_{i=j}^{n}(-1)^{i-j}\frac{\lambda^{i-j}}{(i-j)!}\\
&=\frac{\lambda^j}{j!}\sum_{i=0}^{n-j}(-1)^{i}\frac{\lambda^{i}}{i!}\\
&=\frac{\lambda^j}{j!}\sum_{i=0}^{n-j}\frac{(-\lambda)^{i}}{i!},\quad j=0,1,\dots,n.
\end{aligned}\tag{2.6}$$

Η κατανομή του αριθμού συμπτώσεων στο γενικευμένο πρόβλημα (2.6) έχει εισαχθεί από τον Niermann, ο οποίος έδειξε ότι $p_{n,\lambda}$ είναι μια συνάρτηση μάζας πιθανότητας για όλα τα $\lambda\in(0,1]$. Παρόλα αυτά δεν έδωσε μια πιθανοθεωρητική ερμηνεία στη συνάρτηση $p_{n,\lambda}$ και βρήκε τις ιδιότητες της αναλυτικά.

Επειδή

$$\lim_{n\to\infty}\sum_{i=0}^{n-j}\frac{(-\lambda)^i}{i!}=e^{-\lambda}$$

για κάθε σταθερό *j*, βλέπουμε ότι $p_{n,\lambda}$ συγκλίνει κατά σημείο στη συνάρτηση μάζας πιθανότητας του $Z(\lambda)$, όπου $Z(\lambda)$ μια *Poisson* τυχαία μεταβλητή με μέση τιμή λ, *Poisson*(λ). Αρκετά ενδιαφέρον είναι ότι η προσέγγιση *Poisson* είναι εξαιρετικά ακριβής. Αριθμητικά αποτελέσματα παρουσιάζονται στην εργασία του Niermann. Επίσης, ο Niermann απέδειξε ότι $E[Z_n(\lambda)]=Var[Z_n(\lambda)]=\lambda$ για κάθε $n\geq 2$ και $\lambda\in(0,1]$. Στην πραγματικότητα, το ακόλουθο γενικό αποτέλεσμα δείχνει ότι οι πρώτες n ροπές των $Z_n(\lambda)$ και $Z(\lambda)$ είναι ίδιες, δίνοντας κάποιο φως στην καταπληκτική ταχύτητα της προσέγγισης *Poisson*.

Λήμμα 2.2

Για κάθε $\lambda \in (0,1]$, $E(Z_n(\lambda))_k = \lambda^k I\{k \le n\}$, $k=1,2,\ldots$.

Απόδειξη

Για $k > n$ η σχέση είναι προφανής, αφου $Z_n(\lambda) \in D_{n.}$. Για $k=1,\ldots,n-1$,

$$E(Z_n(\lambda))_k = \sum_{j=k}^{n} \frac{\lambda^j}{(j-k)!} \sum_{i=0}^{n-j} \frac{(-\lambda)^i}{i!}$$

$$= \lambda^k \sum_{j=k}^{n} \frac{\lambda^{j-k}}{(j-k)!} \sum_{i=0}^{n-j} \frac{(-\lambda)^i}{i!}$$

$$\overset{r=j-k}{=} \lambda^k \sum_{r=0}^{n-k} \frac{\lambda^r}{r!} \sum_{i=0}^{(n-k)-r} \frac{(-\lambda)^i}{i!}$$

$$= \lambda^k \sum_{r=0}^{n-k} p_{n-k;\lambda}(r),$$

και αφού $p_{n-k;\lambda}$ είναι μια συνάρτηση πιθανότητας με στήριγμα στο $\{0,1,\ldots,n-k\}$, παίρνουμε το επιθυμητό αποτέλεσμα. Για $k=n$, $E(Z_n(\lambda)_n) = n!\, p_{n;\lambda}(n) = \lambda^n$ και αυτό ολοκληρώνει την απόδειξη.

Πόρισμα 2.1

Για κάθε $\lambda \in (0,1]$ και $\alpha > 0$,

$$\inf_{\mathrm{X} \in D_n} \{d_\alpha(\mathrm{X}, Z(\lambda))\} = d_\alpha(Z_n(\lambda), Z(\lambda)).$$

Επομένως, για $\lambda \in (0,1]$, $Z_n(\lambda)$ ελαχιστοποιεί την απόσταση παραγοντικών ροπών από τη $Z(\lambda)$ ως προς όλες τις τυχαίες μεταβλητές που έχουν στήριγμα ένα υποσύνολο του $\{0,1,\ldots,n\}$. Χρησιμοποιώντας την (2.5), προκύπτει εύκολα ότι η $Z_n(\lambda)$ είναι μοναδική με αυτήν την ιδιότητα. Αξίζει να σημειωθεί ότι για $\lambda \ge 1$, δεν υπάρχει τυχαία μεταβλητή $\mathrm{X} \in D_n$ τέτοια ώστε $E(\mathrm{X})_k = \lambda^k I\{k \le n\}$ για όλα τα k.

Πράγματι, αφού $D_n \subset \mathcal{X}(\infty) \subset \mathcal{X}(1)$, υποθέτοντας $\mathrm{X} \in D_n$ και $E(\mathrm{X})_k = \lambda^k I\{k \le n\}$, προκύπτει ότι

$$0 \le P(\mathrm{X} = n-1) = \lambda^{n-1} \frac{(1-\lambda)}{(n-1)!}$$

που υποδηλώνει ότι $\lambda \le 1$.

Τώρα υπολογίζουμε κάποια ακριβή και ασυμπτωτικά αποτελέσματα για την απόσταση παραγοντικών ροπών και την απόσταση ολικής κύμανσης μεταξύ $\mathrm{Z}_n(\lambda)$ και $Z(\lambda)$ όταν $\lambda \in (0,1]$.

Θεώρημα 2.2

Έστω σταθερό $\alpha > 0$, $\lambda \in (0,1]$ και $d_\alpha(n) := d_\alpha(\mathrm{Z}_n(\lambda), Z(\lambda))$.

Τότε
$$d_\alpha(n) = \frac{\alpha^n \lambda^{n+1}}{n!} \int_0^1 (1-y)^n e^{\alpha\lambda y} dy. \tag{2.7}$$

Επίσης, έχουμε την ακόλουθη διπλή ανισότητα

$$1 + \frac{\alpha\lambda}{n+2} + \frac{\alpha^2\lambda^2}{(n+2)(n+3)} < \frac{(n+1)!}{\alpha^n \lambda^{n+1}} d_\alpha(n) < 1 + \frac{\alpha\lambda}{n+2} + \frac{\alpha^2\lambda^2 e^{\alpha\lambda}}{(n+2)(n+3)}. \tag{2.8}$$

Επομένως, καθώς $n \to \infty$,

$$d_\alpha(n) \sim \frac{\alpha^n \lambda^{n+1}}{(n+1)!} \quad \text{και, ακριβέστερα,}$$

$$d_\alpha(n) = \frac{\alpha^n \lambda^{n+1}}{(n+1)!}\left(1 + \frac{\alpha\lambda}{n+2} + o\left(\frac{1}{n}\right)\right). \tag{2.9}$$

Απόδειξη

Από τον ορισμό της d_α, την (2.2) και το Λήμμα 2.2,

$$d_\alpha(n) = \frac{1}{\alpha} \sum_{k=n+1}^{\infty} \frac{(\alpha\lambda)^k}{k!} = \frac{1}{\alpha}\left(\sum_{k=0}^{\infty} \frac{(\alpha\lambda)^k}{k!} - \sum_{k=0}^{n} \frac{(\alpha\lambda)^k}{k!}\right)$$

$$= \frac{1}{\alpha}\left(e^{\alpha\lambda} - \sum_{k=0}^{n} \frac{(\alpha\lambda)^k}{k!}\right)$$

$$= \frac{1}{\alpha n!} \int_0^{\alpha\lambda} (\alpha\lambda - x)^n e^x dx,$$

και αν θέσουμε $x = \alpha\lambda y$ οδηγούμαστε στην (2.7).

Πράγματι $d_\alpha(n) = \dfrac{1}{\alpha n!}\int_0^{\alpha\lambda}(\alpha\lambda - \alpha\lambda y)^n e^{\alpha\lambda y}\alpha\lambda dy$

$$= \frac{\alpha\lambda}{\alpha n!}\int_0^{\alpha\lambda}(\alpha\lambda)^n(1-y)^n e^{\alpha\lambda y}dy$$

$$= \frac{\lambda}{n!}\int_0^{\alpha\lambda}\alpha^n\lambda^n(1-y)^n e^{\alpha\lambda y}dy$$

$$= \frac{\alpha^n\lambda^{n+1}}{n!}\int_0^{\alpha\lambda}(1-y)^n e^{\alpha\lambda y}dy.$$

Τώρα η (2.8) προκύπτει από τις ανισότητες

$$1+\alpha\lambda y+\frac{1}{2}\alpha^2\lambda^2 y^2 < e^{\alpha\lambda y} < 1+\alpha\lambda y+\frac{1}{2}e^{\alpha\lambda}\alpha^2\lambda^2 y^2, \quad 0<y<1,$$

ενώ η (2.9) είναι προφανής από (2.8).

Το Θεώρημα 2.2 αποδείχτηκε.

Το Θεώρημα 2.2 και το Θεώρημα 2.1 δίνουν την ακόλουθη πρόταση.

Πόρισμα 2.2

Ένα άνω φράγμα για την $d_{TV}(Z_n, Z)$ δίνεται από

$$d_{TV}(Z_n, Z) < \frac{2^n}{(n+1)!}\left(1+\frac{2}{n+2}+\frac{4e^2}{(n+2)(n+3)}\right) \sim \frac{2^n}{(n+1)!}. \qquad (2.10)$$

Το φράγμα $d_{TV}(Z_n, Z) \le \dfrac{2^n}{n!}$, που δόθηκε από τον Diaconis δεν είναι ασυμπτωτικά βέλτιστο, επειδή $\dfrac{2^n}{(n+1)!} = o\left(\dfrac{2^n}{n!}\right)$. Επομένως, η απόσταση παραγοντικών ροπών δίνει ένα βέλτιστης τάξης άνω φράγμα για την απόσταση ολικής κύμανσης στο πρόβλημα σταθερών σημείων.

Θεώρημα 2.3

Για κάθε $\lambda \in (0,1]$, έστω $d_{TV}(Z_n(\lambda), Z(\lambda))$ η απόσταση ολικής κύμανσης μεταξύ $Z_n(\lambda)$ και $Z(\lambda)$. Τότε

$$d_{TV}(n) = \frac{\lambda^{n+1}}{2n!}\int_0^1 [y^n + (2-y)^n] e^{-\lambda y} dy. \tag{2.11}$$

Επίσης, έχουμε τις ακόλουθες ανισότητες

$$d_{TV}(n) > \frac{2^n \lambda^{n+1}}{(n+1)!}\left(1 - \frac{2\lambda}{n+2}\left(1 - \frac{1}{2^{n+1}}\right)\right),$$

$$d_{TV}(n) < \frac{2^n \lambda^{n+1}}{(n+1)!}\left(1 - \frac{2\lambda}{n+2}\left(1 - \frac{1}{2^{n+1}}\right) + \frac{4\lambda^2}{(n+2)(n+3)}\left(1 - \frac{n+3}{2^{n+2}}\right)\right). \tag{2.12}$$

Επομένως, καθώς $n \to \infty$,

$$d_{TV}(n) \sim \frac{2^n \lambda^{n+1}}{(n+1)!} \text{ και, ακριβέστερα,}$$

$$d_{TV}(n) = \frac{2^n \lambda^{n+1}}{(n+1)!}\left(1 - \frac{2\lambda}{n+2} + o\left(\frac{1}{n}\right)\right). \tag{2.13}$$

Απόδειξη

Η (2.13) είναι άμεση συνέπεια των ανισοτήτων (2.12). Οι ανισότητες (2.12) προκύπτουν από την (2.11) και το γεγονός ότι

$$1 - \lambda y < e^{-\lambda y} < 1 - \lambda y + \frac{1}{2}\lambda^2 y^2, \quad 0 < y < 1.$$

Μένει να δείξουμε την (2.11). Από τη σχέση $d_{TV}(\mathrm{X}_1, \mathrm{X}_2) = \sum_{j=0}^{\infty}(p_1(j) - p_2(j))^+$ με $p_1 = p_{n;\lambda}$ και p_2 η συνάρτηση πιθανότητας της *Poisson(λ)*, παίρνουμε

$$d_{TV}(n) = \sum_{j=0}^{n}\frac{\lambda^j}{j!}\left[\sum_{i=0}^{n-j}\frac{(-\lambda)^i}{i!} - e^{-\lambda}\right]^+$$

$$= \sum_{j=0}^{n}\frac{\lambda^j}{j!}\left[\frac{(-1)^{n-j}}{(n-j)!}\int_0^{\lambda}(\lambda - x)^{n-j} e^{-x} dx\right]^+$$

$$\overset{k=n-j}{=} \sum_{k=0}^{n}\frac{\lambda^{n-k}}{(n-k)!}\left[\frac{(-1)^k}{k!}\int_0^{\lambda}(\lambda - x)^k e^{-x} dx\right]^+$$

$$= \frac{1}{n!}\int_0^{\lambda} e^{-x}\left(\sum_{k\ \ \alpha\rho\tau\iota o\varsigma}\binom{n}{k}(\lambda - x)^k \lambda^{n-k}\right)dx.$$

Επειδή $\sum_{k\ \ \alpha\rho\tau\iota o\varsigma}\binom{n}{k}(\lambda - x)^k \lambda^{n-k} = \frac{1}{2}\left[x^n + (2\lambda - x)^n\right],$ έχουμε

$$d_{TV}(n) = \frac{1}{2n!}\int_0^{\lambda}[x^n + (2\lambda - x)^n]e^{-x}dx.$$

Αν θέσουμε $x = \lambda y$ παίρνουμε την (2.11).

Το Θεώρημα 2.3 αποδείχτηκε.

Παρατήρηση 2.1

Αν και η απόσταση παραγοντικών ροπών d_α είναι μεγαλύτερη από την απόσταση ολικής κύμανσης όταν $\alpha \geq 2,$ η κατάσταση για $\alpha \in (0,2)$ είναι τελείως διαφορετική.

Υποθέτουμε ότι για κάποιο $\alpha \in (0,2)$ και $t \geq 0$ μπορούμε να βρούμε μια πεπερασμένη σταθερά $C = C(\alpha, t) > 0$ έτσι ώστε

$$d_{TV}(\mathrm{X}_1, \mathrm{X}_2) \leq C d_\alpha(\mathrm{X}_1, \mathrm{X}_2) \text{ για όλες τις } \mathrm{X}_1, \mathrm{X}_2 \in \mathcal{X}(t). \tag{2.14}$$

Προφανώς, Z και Z_n, n=1,2,… είναι στο $\mathcal{X}(\infty) \subset \mathcal{X}(t).$

Από το Θεώρημα 2.3 ξέρουμε ότι $\lim_{n\to\infty}\frac{(n+1)!}{2^n}d_{TV}(Z_n, Z) = 1.$

Από την άλλη μεριά, από την (2.8) με $\lambda = 1,$

$d_\alpha(Z_n, Z) < \frac{\alpha^n}{(n+1)!}\left(1 + \frac{\alpha}{n+2} + \frac{\alpha^2 e^\alpha}{(n+2)(n+3)}\right),$ και, επειδή $\frac{\alpha}{2} < 1,$ έχουμε

$$1 = \lim_{n\to\infty}\frac{(n+1)!}{2^n}d_{TV}(Z_n, Z) \leq C\lim_{n\to\infty}\left(\left(\frac{\alpha}{2}\right)^n\left(1 + \frac{\alpha}{n+2} + \frac{\alpha^2 e^\alpha}{(n+2)(n+3)}\right)\right) = 0, \text{ άτοπο.}$$

Κ Ε Φ Α Λ Α Ι Ο 3

Προσέγγιση Poisson

3.1. Ένα Ενιαίο Φράγμα για την Προσέγγιση Poisson

Ας θεωρήσουμε μία τυχαία μεταβλητή Χ η οποία λαμβάνει τιμές στο $\{0,1,2,\ldots\}$ με συνάρτηση πιθανότητας $p,$ μέση τιμή μ και πεπερασμένη διασπορά σ^2. Ορίζουμε την συνάρτηση

$$h_{\mathrm{X}}(x)=\sum_{k=0}^{x}(\mu-k)\,p(k), \quad x=0,1,2,\ \ldots \tag{3.1}$$

Χρησιμοποιώντας την παραπάνω συνάρτηση και άλλα πολύτιμα εργαλεία, θα αποδείξουμε σταδιακά ότι

$$d_{TV}(\mathrm{X},\mathrm{Z})\leq\frac{1-e^{-\lambda}}{\lambda}\sum_{x=0}^{+\infty}\left|h_{\mathrm{X}}(x)-\lambda p(x)\right|+\min\left\{1,\lambda^{-1/2}\right\}\left|\lambda-\mu\right|, \tag{3.2}$$

όπου η Ζ ακολουθεί την $Poisson(\lambda)$.

Λήμμα 3.1

Για την συνάρτηση $h_{\mathrm{X}}(x)$ που ορίζεται στην (3.1), ισχύει ότι $h_{\mathrm{X}}(x)\geq 0$ για κάθε $x\in N.$

Απόδειξη

Από την σχέση $\sum_{k=0}^{+\infty}(\mu-k)\,p(k)=\sum_{k=0}^{\infty}\mu p(k)-\sum_{k=0}^{\infty}kp(k)=\mu\sum_{k=0}^{\infty}p(k)-\mu=\mu\cdot 1-\mu=\mu-\mu=0$ έπεται ότι

$$\sum_{k=0}^{x}(\mu-k)\,p(k)+\sum_{k=x+1}^{+\infty}(\mu-k)\,p(k)=0$$

ή $$\sum_{k=0}^{x}(\mu-k)\,p(k)=\sum_{k=x+1}^{\infty}(k-\mu)\,p(k).$$

Αν λοιπόν $x\leq\mu,$ τότε προφανώς $h_{\mathrm{X}}(x)=\sum_{k=0}^{x}(\mu-k)\,p(k)\geq 0$ και αν $x>\mu,$ τότε έχουμε $h_{\mathrm{X}}(x)=\sum_{k=x+1}^{+\infty}(k-\mu)\,p(k)\geq 0.$

Το επόμενο λήμμα εισήχθη από τους Cacoullos and Papathanasiou (1989).

Λήμμα 3.2 (Ταυτότητα Συνδιακύμασης)

Για κάθε συνάρτηση g με

$$E\left|(\mathrm{X}-\mu)g(\mathrm{X})\right| < +\infty \text{ και } \sum_{x=0}^{+\infty}\left|\Delta g(x)\right|h_{\mathrm{X}}(x) < +\infty,$$

ισχύει

$$Cov(\mathrm{X}, g(\mathrm{X})) = \sum_{x=0}^{+\infty}\Delta g(x)h_{\mathrm{X}}(x), \tag{3.3}$$

όπου $\Delta g(x) = g(x+1) - g(x)$ η προς τα εμπρός διαφορά.

Απόδειξη

Από τις υποθέσεις του λήμματος, παρατηρούμε αρχικά ότι και τα δύο μέλη της (3.3) είναι πεπερασμένα. Πράγματι,

$$\left|Cov(\mathrm{X}, g(\mathrm{X}))\right| = \left|E[(\mathrm{X}-\mu)g(\mathrm{X})]\right| \le E\left|(\mathrm{X}-\mu)g(\mathrm{X})\right| < +\infty$$

και

$$\left|\sum_{x=0}^{+\infty}\Delta g(x)h_{\mathrm{X}}(x)\right| \le \sum_{x=0}^{+\infty}\left|\Delta g(x)\right|h_{\mathrm{X}}(x) < +\infty.$$

Ξεκινώντας από το δεύτερο μέλος της (3.3),

$$\sum_{x=0}^{+\infty}\Delta g(x)h_{\mathrm{X}}(x) = \sum_{x=0}^{+\infty}\left(\Delta g(x)\sum_{k=0}^{x}(\mu-k)p(k)\right)$$

και εφαρμόζοντας το Θεώρημα Fubini έχουμε

$$\begin{aligned}\sum_{x=0}^{+\infty}\Delta g(x)h_{\mathrm{X}}(x) &= \sum_{k=0}^{+\infty}\left((\mu-k)p(k)\sum_{x=k}^{+\infty}\Delta g(x)\right)\\ &= \sum_{k=0}^{+\infty}(\mu-k)p(k)(-g(k))\\ &= E[(\mathrm{X}-\mu)g(\mathrm{X})]\\ &= Cov(\mathrm{X}, g(\mathrm{X})).\end{aligned}$$

Στο σημείο αυτό πρέπει να σημειώσουμε ότι τα παραπάνω λήμματα ισχύουν και για $\sigma^2 = 0$.

Επιπλέον, εύκολα μπορεί κανείς να διαπιστώσει το ακόλουθο.

Πόρισμα 3.1

(i) Γενικά ισχύει $Cov(\mathrm{X},\mathrm{X})=\sigma^2=\sum_{x=0}^{+\infty} h_{\mathrm{X}}(x).$

(ii) Αν $\sigma^2>0,$ τότε η συνάρτηση $p^*(x)=\sigma^{-2}h_{\mathrm{X}}(x),\ x=0,1,2,\dots$ είναι συνάρτηση πιθανότητας κάποιας τυχαίας μεταβλητής X^*. Στην περίπτωση αυτή η σχέση (3.3) (ταυτότητα συνδιακύμανσης) γράφεται

$$Cov(\mathrm{X},g(\mathrm{X}))=\sigma^2 E[\Delta g(\mathrm{X}^*)]$$

με την προυπόθεση ότι $E|(\mathrm{X}-\mu)g(\mathrm{X})|<+\infty$ και $E|\Delta g(\mathrm{X}^*)|<+\infty.$

Μπορούμε να έχουμε και μια τρίτη μορφή της ταυτότητας συνδιακύμανσης με την βοήθεια της συνάρτησης

$$w_{\mathrm{X}}(x)=\frac{p^*(x)}{p(x)}=\frac{1}{\sigma^2 p(x)}h_{\mathrm{X}}(x)=\frac{1}{\sigma^2 p(x)}\sum_{k=0}^{x}(\mu-k)p(k) \tag{3.4}$$

(όταν αυτή ορίζεται), γνωστής ως *w*-συνάρτηση.

Πόρισμα 3.2

Έστω $\sigma^2>0$ και η τυχαία μεταβλητή X που ικανοποιεί την $\{x: p(x)>0\}=\{0,1,\dots,b\}$ $(1\le b\le+\infty)$. Τότε για την συνάρτηση $w_{\mathrm{X}}(x)$ στην (3.4), η ταυτότητα συνδιακύμανσης γράφεται

$$Cov(\mathrm{X},g(\mathrm{X}))=\sigma^2 E[w_{\mathrm{X}}(\mathrm{X})\Delta g(\mathrm{X})],$$

με την προϋπόθεση ότι $E|(\mathrm{X}-\mu)g(\mathrm{X})|<+\infty$ και $E[w_{\mathrm{X}}(\mathrm{X})|\Delta g(\mathrm{X})|]<+\infty.$

Από το Λήμμα 3.1 είναι φανερό ότι για κάθε $x\in\{0,1,2,\dots,b\},\ w_{\mathrm{X}}(x)\ge 0.$ Επιπλέον η w-συνάρτηση είναι άμεσα συνδεδεμένη με την κατανομή $Poisson(\lambda)$, όπως φαίνεται και στην παρακάτω πρόταση.

Πρόταση 3.1

Η τυχαία μεταβλητή Z ακολουθεί την $Poisson(\lambda)$ αν και μόνο αν $w_Z(x)=1.$

Απόδειξη

Για το ευθύ, ας υποθέσουμε ότι η τυχαία μεταβλητή Z ακολουθεί την κατανομή $Poisson(\lambda)$ με μέση τιμή λ και διασπορά λ (>0). Τότε από την (3.4) έπεται

$$w_X(x)=\frac{1}{\lambda e^{-\lambda}\frac{\lambda^x}{x!}}\sum_{k=0}^{x}(\lambda-k)e^{-\lambda}\frac{\lambda^k}{k!}$$

$$=\frac{x!}{e^{-\lambda}\lambda^{x+1}}e^{-\lambda}\sum_{k=0}^{x}(\lambda-k)\frac{\lambda^k}{k!}$$

$$=\frac{x!}{\lambda^{x+1}}\sum_{k=0}^{x}(\lambda-k)\frac{\lambda^k}{k!}$$

$$=\frac{x!}{\lambda^{x+1}}\left(\sum_{k=0}^{x}\lambda\frac{\lambda^k}{k!}-\sum_{k=0}^{x}k\frac{\lambda^k}{k!}\right)$$

$$=\frac{x!}{\lambda^{x+1}}\left(\sum_{k=0}^{x}\frac{\lambda^{k+1}}{k!}-\sum_{k=0}^{x}\frac{\lambda^k}{(k-1)!}\right)$$

$$=\frac{x!}{\lambda^{x+1}}\left(\sum_{k=0}^{x}\frac{\lambda^{k+1}}{k!}-\sum_{k=0}^{x+1}\frac{\lambda^{k+1}}{k!}\right)$$

$$=\frac{x!}{\lambda^{x+1}}\frac{\lambda^{x+1}}{x!}=1.$$

Αντίστροφα, ας θεωρήσουμε μια τυχαία μεταβλητή Z με συνάρτηση πιθανότητας $q(x)$, μέση τιμή μ, διασπορά σ^2 (>0) και $w_Z(x)=1.$ Από τον ορισμό της w-συνάρτησης (βλέπε (3.4)), για την Z προκύπτει ότι

$$\sigma^2 q(x)=\sum_{k=0}^{x}(\mu-k)q(k),\qquad x=0,1,2,\ldots \tag{3.5}$$

Με επαγωγή (ως προς x) θα δείξουμε, χρησιμοποιώντας τη (3.5), ότι

$$q(x)=e^{-\lambda}\frac{\lambda^x}{x!},\qquad x=0,1,2,\ldots$$

όπου $\lambda=\mu=\sigma^2.$

Για $x=0$, έχουμε $\sigma^2 q(0)=\mu q(0)$, άρα $\sigma^2=\mu$ το οποίο ορίζουμε ως λ.

Για $x=1$, έχουμε $\lambda q(1)=\lambda q(0)+(\lambda-1)q(1)$, άρα $q(1)=\lambda q(0)=\lambda c$ $(c=q(0))$.

Υποθέτουμε ότι για τυχαίο x, ισχύει $q(x)=q(0)\dfrac{\lambda^x}{x!}$. Τότε για $x+1$, η (3.5) γράφεται

$$\lambda q(x+1)=\sum_{k=0}^{x+1}(\lambda-k)q(k)$$

ή
$$\lambda q(x+1)=\sum_{k=0}^{x}(\lambda-k)q(k)+(\lambda-x-1)q(x+1)$$

ή
$$\lambda q(x+1)=\lambda q(x)+(\lambda-x-1)q(x+1)$$

ή
$$(x+1)q(x+1)=\lambda q(x)$$

ή
$$q(x+1)=\frac{1}{x+1}\lambda q(0)\frac{\lambda^x}{x!}$$

ή
$$q(x+1)=q(0)\frac{\lambda^{x+1}}{(x+1)!}.$$

Άρα $q(x)=c\dfrac{\lambda^x}{x!}$ x=0,1,… και $c=e^{-\lambda}$.

Η απόδειξη ολοκληρώθηκε.

Από το Πόρισμα 3.2 και την Πρόταση 3.1 προκύπτουν άμεσα δύο ενδιαφέρουσες παρατηρήσεις.

Παρατήρηση 3.1 (M. Majsnerowska, 1998)

Έστω μια ακολουθία τ.μ. $(\mathrm{X}_n,\ n=1,2,\ldots)$ με τιμές στο Ν και με αντίστοιχες μέσες τιμές μ_n, διασπορές $\sigma_n{}^2$ και w_n-συναρτήσεις $(n=1,2,\ldots)$ έτσι ώστε $\lim_{n\to+\infty}\mu_n=\lim_{n\to+\infty}\sigma_n{}^2=\lambda$ $(\lambda>0)$. Ας υποθέσουμε επιπλέον ότι η τ.μ. Ζ ακολουθεί την κατανομή $Poisson(\lambda)$. Τότε

$$d_{TV}(\mathrm{X}_n,Z)\to 0 \quad \text{καθώς } n\to+\infty$$

αν και μόνο αν

$$w_n(\mathrm{X}_n)\xrightarrow{p}1 \quad \text{καθώς } n\to+\infty.$$

Παρατήρηση 3.2

Έστω Z μια τυχαία μεταβλητή που ακολουθεί την κατανομή $Poisson(\lambda)$ $(\lambda > 0)$ και μια τυχούσα συνάρτηση g. Τότε η ταυτότητα συνδιακύμανσης, όπως αυτή γράφεται στο Πόρισμα 3.2, είναι

$$Cov(Z, g(Z)) = \lambda E[\Delta g(Z)]$$

ή $$E[Zg(Z)] - E(Z)E[g(Z)] = \lambda E[g(Z+1) - g(Z)]$$

ή $$E[Zg(Z)] - \lambda E[g(Z)] = \lambda E[g(Z+1)] - \lambda E[g(Z)]$$

ή $$E[Zg(Z)] = \lambda E[g(Z+1)]$$

εφόσον $E\left|(\mathrm{X}-\mu)g(\mathrm{Z})\right| < +\infty$ και $E[\Delta g(Z)] < +\infty$.

Για να αποδείξουμε την (3.2), χρειαζόμαστε και άλλα εργαλεία εκτός της συνάρτησης $h_{\mathrm{X}}(x)$. Έστω λοιπόν μια τυχαία μεταβλητή Z που ακολουθεί την $Poisson(\lambda)$ $(\lambda > 0)$ με συνάρτηση πιθανότητας q, μέση τιμή $\mu = \lambda$ και διασπορά $\sigma^2 = \lambda$. Αν A υποσύνολο του N με αντίστοιχη δείκτρια

$$I(k \in \mathrm{A}) = \begin{cases} 1, & k \in \mathrm{A} \\ 0, & k \notin \mathrm{A} \end{cases},$$

τότε ορίζουμε την συνάρτηση

$$\begin{aligned} l_{\lambda,\mathrm{A}}(x) &= \frac{1}{q(x-1)} \sum_{k=0}^{x-1} (I(k \in \mathrm{A}) - P(\mathrm{Z} \in \mathrm{A}))q(k) \\ &= \frac{1}{e^{-\lambda} \dfrac{\lambda^{x-1}}{(x-1)!}} \sum_{k=0}^{x-1} (I(k \in \mathrm{A}) - P(\mathrm{Z} \in \mathrm{A}))e^{-\lambda} \frac{\lambda^k}{k!} \\ &= \frac{(x-1)!}{\lambda^{x-1}} \sum_{k=0}^{x-1} (I(k \in \mathrm{A}) - P(\mathrm{Z} \in \mathrm{A})) \frac{\lambda^k}{k!} \end{aligned} \tag{3.6}$$

(για $x = 1,2,\ldots$). Κατά σύμβαση δεχόμαστε ότι $l_{\lambda,\mathrm{A}}(0) = 0$. Παρατηρούμε ότι $l_{\lambda,\mathrm{A}}(x) = \lambda g_{\lambda,\mathrm{A}}(x)$ όπου $g_{\lambda,\mathrm{A}}$ είναι η λύση της εξίσωσης Chen-Stein (C-S).

Η εξίσωση Chen-Stein είναι της μορφής

$$\lambda g(x+1) - xg(x) = I(\mathrm{X} \in \mathrm{A}) - P(\mathrm{Z} \in \mathrm{A}), \qquad x = 0,1,2,\ldots$$

Δεν είναι δύσκολο να λύσει κανείς την εξίσωση C-S για μικρές τιμές του x:

Για $x=0$, είναι φανερό ότι

$$\lambda g(1) = I(0 \in \mathrm{A}) - P(\mathrm{Z} \in \mathrm{A}).$$

Για $x=1$, είναι φανερό ότι

$$\lambda g(2) = g(1) + I(1 \in \mathrm{A}) - P(\mathrm{Z} \in \mathrm{A})$$

$$= \frac{1}{\lambda}[I(0 \in \mathrm{A}) - P(\mathrm{Z} \in \mathrm{A}) + \lambda(I(1 \in \mathrm{A}) - P(\mathrm{Z} \in \mathrm{A}))].$$

Για $x=2$, έχουμε

$$\lambda g(3) = 2g(2) + I(2 \in \mathrm{A}) - P(\mathrm{Z} \in \mathrm{A})$$

$$= \frac{2}{\lambda!}\left[I(0 \in \mathrm{A}) - P(\mathrm{Z} \in \mathrm{A}) + \lambda(I(1 \in \mathrm{A}) - P(\mathrm{Z} \in \mathrm{A})) + \frac{\lambda^2}{2!}(I(2 \in \mathrm{A}) - P(\mathrm{Z} \in \mathrm{A})) \right].$$

Επαγωγικά αποδεικνύεται ότι η λύση της εξίσωσης (C-S) δίνεται από τον τύπο

$$g_{\lambda,\mathrm{A}}(x) = \frac{(x-1)!}{\lambda^x}\sum_{k=0}^{x-1}(I(k \in \mathrm{A}) - P(Z \in \mathrm{A}))\frac{\lambda^k}{k!}$$

ή
$$g_{\lambda,\mathrm{A}}(x) = \frac{1}{\lambda q(x-1)}\sum_{k=0}^{x-1}(I(k \in \mathrm{A}) - P(\mathrm{Z} \in \mathrm{A}))q(k),$$

όπου $q(k) = e^{-\lambda}\dfrac{\lambda^k}{k!}$ $(k = 0,1,2,\ldots)$ είναι η συνάρτηση πιθανότητας της $Poisson(\lambda)$.

Όρίζοντας τις νόρμες

$$\|g\| = \|g_{\lambda,\mathrm{A}}\| = \sup_{\mathrm{A},x}|g_{\lambda,\mathrm{A}}(x)| \quad \text{και} \quad \|\Delta g\| = \|\Delta g_{\lambda,\mathrm{A}}\| = \sup_{\mathrm{A},x}|g_{\lambda,\mathrm{A}}(x+1) - g_{\lambda,\mathrm{A}}(x)|,$$

οι Barbour and Eagleson (1983) αποδεικνύουν το παρακάτω ενδιαφέρον αποτέλεσμα: Αν $g(x) = g_{\lambda,\mathrm{A}}(x)$ είναι η λύση της εξίσωσης (C-S), τότε

$$\|g\| \le \min\left\{1, \frac{1}{\sqrt{\lambda}}\right\} \quad \text{και} \quad \|\Delta g\| \le \frac{1-e^{-\lambda}}{\lambda} \le \min\left\{1, \frac{1}{\lambda}\right\}.$$

Λήμμα 3.3

Η συνάρτηση $l_{\lambda,\mathrm{A}}(x)$ ικανοποεί τις σχέσεις

$$\sup_{x,\mathrm{A}}|l_{\lambda,\mathrm{A}}(x)| \le \min\{\lambda, \sqrt{\lambda}\}$$

και
$$\sup_{x,\mathrm{A}}|\Delta l_{\lambda,\mathrm{A}}(x)| \le 1 - e^{-\lambda} \le \min\{\lambda, 1\}.$$

Θεώρημα 3.1

Ας θεωρήσουμε την τυχαία μεταβλητή Χ με συνάρτηση πιθανότητας p, μέση τιμή μ και διασπορά $\sigma^2 < +\infty$, καθώς και την τυχαία μεταβλητή Ζ η οποία ακολουθεί την $Poisson(\lambda)$ με $\lambda > 0$. Τότε

$$d_{TV}(\mathrm{X}, \mathrm{Z}) \leq \frac{1 - e^{-\lambda}}{\lambda} \sum_{x=0}^{+\infty} \left| h_{\mathrm{X}}(x) - \lambda p(x) \right| + \min\{1, \lambda^{-1/2}\} |\lambda - \mu|$$

(δηλαδή ισχύει η σχέση (3.2)).

Απόδειξη

Αρχικά υπολογίζουμε την $\Delta l_{\lambda,\mathrm{A}}(x)$.

$$\begin{aligned}
\Delta l_{\lambda,\mathrm{A}}(x) &= l_{\lambda,\mathrm{A}}(x+1) - l_{\lambda,\mathrm{A}}(x) \\
&= \frac{x!}{\lambda^x} \sum_{k=0}^{x} (I(k \in \mathrm{A}) - P(Z \in \mathrm{A})) \frac{\lambda^k}{k!} - \frac{(x-1)!}{\lambda^{x-1}} \sum_{k=0}^{x-1} (I(k \in \mathrm{A}) - P(Z \in \mathrm{A})) \frac{\lambda^k}{k!} \\
&= \frac{x!}{\lambda^x} (I(x \in \mathrm{A}) - P(Z \in \mathrm{A})) \frac{\lambda^x}{x!} + \left(\frac{x!}{\lambda^x} - \frac{(x-1)!}{\lambda^{x-1}} \right) \sum_{k=0}^{x-1} (I(k \in \mathrm{A}) - P(Z \in \mathrm{A})) \frac{\lambda^k}{k!} \\
&= I(x \in \mathrm{A}) - P(Z \in \mathrm{A}) + \frac{(x-1)!}{\lambda^{x-1}} \left(\frac{x}{\lambda} - 1 \right) \sum_{k=0}^{x-1} (I(k \in \mathrm{A}) - P(Z \in \mathrm{A})) \frac{\lambda^k}{k!} \\
&= I(x \in \mathrm{A}) - P(Z \in \mathrm{A}) + \frac{x - \lambda}{\lambda} l_{\lambda,\mathrm{A}}(x).
\end{aligned}$$

Παίρνοντας τη μέση τιμή κάθε μέλους της σχέσης

$$\Delta l_{\lambda,\mathrm{A}}(x) = I(x \in \mathrm{A}) - P(Z \in \mathrm{A}) + \frac{x - \lambda}{\lambda} l_{\lambda,\mathrm{A}}(x),$$

έχουμε

$$P(\mathrm{X} \in \mathrm{A}) - P(Z \in \mathrm{A}) = E[\Delta l_{\lambda,\mathrm{A}}(\mathrm{X})] - \frac{1}{\lambda} E[(\mathrm{X} - \lambda) l_{\lambda,\mathrm{A}}(\mathrm{X})]. \tag{3.7}$$

Όμως η ποσότητα $E[(\mathrm{X} - \lambda) l_{\lambda,\mathrm{A}}(\mathrm{X})]$ είναι

$$\begin{aligned}
E[(\mathrm{X} - \lambda) l_{\lambda,\mathrm{A}}(\mathrm{X})] &= E[\mathrm{X} l_{\lambda,\mathrm{A}}(\mathrm{X})] - \lambda E[l_{\lambda,\mathrm{A}}(\mathrm{X})] \\
&= E[\mathrm{X} l_{\lambda,\mathrm{A}}(\mathrm{X})] - \mu E[l_{\lambda,\mathrm{A}}(\mathrm{X})] + \mu E[l_{\lambda,\mathrm{A}}(\mathrm{X})] - \lambda E[l_{\lambda,\mathrm{A}}(\mathrm{X})] \\
&= Cov(\mathrm{X}, l_{\lambda,\mathrm{A}}(\mathrm{X})) + (\mu - \lambda) E[l_{\lambda,\mathrm{A}}(\mathrm{X})].
\end{aligned}$$

Συνεπώς, η ισότητα (3.7) γράφεται

$$P(\mathrm{X} \in \mathrm{A}) - P(Z \in \mathrm{A}) = E[\Delta l_{\lambda,\mathrm{A}}(\mathrm{X})] - \frac{1}{\lambda} Cov(\mathrm{X}, l_{\lambda,\mathrm{A}}(\mathrm{X})) + \frac{\lambda - \mu}{\lambda} E[l_{\lambda,\mathrm{A}}(\mathrm{X})]. \tag{3.8}$$

Για να εφαρμόσουμε την ταυτότητα συνδιακύμανσης (βλέπε (3.3)) στην σχέση (3.8), πρέπει να ικανοποιείται η συνθήκη $\sum_{x=0}^{+\infty} |\Delta l_{\lambda,\mathrm{A}}(x)| h_{\mathrm{X}}(x) < +\infty$. Πράγματι, ισχύει

$$\sum_{x=0}^{+\infty} |\Delta l_{\lambda,\mathrm{A}}(x)| h_{\mathrm{X}}(x) \le (1 - e^{-\lambda}) \sum_{x=0}^{+\infty} h_{\mathrm{X}}(x) = (1 - e^{-\lambda})\sigma^2 < +\infty.$$

Αντικαθιστώντας λοιπόν την ταυτότητα συνδιακύμανσης στην (3.8), έχουμε

$$\begin{aligned}
P(\mathrm{X} \in \mathrm{A}) - P(Z \in \mathrm{A}) &= E[\Delta l_{\lambda,\mathrm{A}}(\mathrm{X})] - \frac{1}{\lambda} \sum_{x=0}^{+\infty} \Delta l_{\lambda,\mathrm{A}}(x) h_{\mathrm{X}}(x) + \frac{\lambda - \mu}{\lambda} E[l_{\lambda,\mathrm{A}}(\mathrm{X})] \\
&= \sum_{x=0}^{+\infty} \Delta l_{\lambda,\mathrm{A}}(x) p(x) - \frac{1}{\lambda} \sum_{x=0}^{+\infty} \Delta l_{\lambda,\mathrm{A}}(x) h_{\mathrm{X}}(x) + \frac{\lambda - \mu}{\lambda} E[l_{\lambda,\mathrm{A}}(\mathrm{X})] \\
&= \sum_{x=0}^{+\infty} \Delta l_{\lambda,\mathrm{A}}(x) [p(x) - \lambda^{-1} h_{\mathrm{X}}(x)] + \frac{\lambda - \mu}{\lambda} E[l_{\lambda,\mathrm{A}}(\mathrm{X})].
\end{aligned}$$

Λαμβάνοντας τις απόλυτες τιμές, έπεται

$$\begin{aligned}
|P(\mathrm{X} \in \mathrm{A}) - P(\mathrm{Z} \in \mathrm{A})| &= \left| \sum_{x=0}^{+\infty} \Delta l_{\lambda,\mathrm{A}}(x) [p(x) - \lambda^{-1} h_{\mathrm{X}}(x)] + \frac{\lambda - \mu}{\lambda} E[l_{\lambda,\mathrm{A}}(\mathrm{X})] \right| \\
&\le \sum_{x=0}^{+\infty} |\Delta l_{\lambda,\mathrm{A}}(x)| |p(x) - \lambda^{-1} h_{\mathrm{X}}(x)| + \frac{|\lambda - \mu|}{\lambda} E|l_{\lambda,\mathrm{A}}(\mathrm{X})| \\
&\le (1 - e^{-\lambda}) \sum_{x=0}^{+\infty} |p(x) - \lambda^{-1} h_{\mathrm{X}}(x)| + \frac{|\lambda - \mu|}{\lambda} \min\{\lambda, \sqrt{\lambda}\},
\end{aligned}$$

όπου η τελευταία ανισότητα προκύπτει από το Λήμμα 3.3 και είναι ανεξάρτητη του συνόλου A. Για το λόγο αυτό, θεωρώντας το sup ως προς A, καταλήγουμε στη ζητούμενη σχέση.

Πόρισμα 3.3

Έστω ότι ισχύουν οι υποθέσεις του Θεωρήματος 3.1 και επιπλέον ορίζεται η w-συνάρτηση της τυχαίας μεταβλητής X. Τότε

$$d_{TV}(\mathrm{X}, \mathrm{Z}) \le \frac{1 - e^{-\lambda}}{\lambda} \sum_{x=0}^{+\infty} |\sigma^2 w_{\mathrm{X}}(x) - \lambda| + \min\{1, \lambda^{-1/2}\} |\lambda - \mu|.$$

3.2. Παραδείγματα και Εφαρμογές

Τέλος παραθέτουμε τρία χαρακτηριστικά παραδείγματα τα οποία θα μας βοηθήσουν στην κατανήση των εννοιών.

Παράδειγμα 3.1 (M.Majsnerowska, 1998)

Έστω μια τυχαία μεταβλητή Χ~ $\Delta\iota\omega\nu\upsilon\mu\iota\kappa\acute{\eta}(n,p)$ $(0<p<1)$ με συνάρτηση πιθανότητας

$$p(x)=\binom{n}{x}p^x(1-p)^{n-x}, \qquad x=0,1,2,\dots,n,$$

μέση τιμή $\mu=np$ και διασπορά $\sigma^2=np(1-p)$. Με επαγωγή (ως προς x) αποδεικνύεται ότι

$$h_{\mathrm{X}}(x)=\frac{n!}{x!(n-x-1)!}p^{x+1}(1-p)^{n-x}, \qquad x=0,1,2,\dots,\ n-1.$$

Συνεπώς,

$$p^*(x)=\frac{h_{\mathrm{X}}(x)}{\sigma^2}=\frac{\dfrac{n!}{x!(n-x-1)!}p^{x+1}(1-p)^{n-x}}{np(1-p)}=\frac{(n-1)!}{x!(n-x-1)!}p^x(1-p)^{n-x-1}=\binom{n-1}{x}p^x(1-p)^{n-x-1}$$

για $x=0,1,2,\dots,n-1$, δηλαδή η τυχαία μεταβλητή $\mathrm{X}^*\sim\Delta\iota\omega\nu\upsilon\mu\iota\kappa\acute{\eta}(n-1,p)$. Επιπλέον ισχύει

$$w_{\mathrm{X}}(x)=\frac{p^*(x)}{p(x)}=\frac{\binom{n-1}{x}p^x(1-p)^{n-x-1}}{\binom{n}{x}p^x(1-p)^{n-x}}=\frac{\dfrac{(n-1)!}{x!(n-x-1)!}p^x(1-p)^{n-x}}{\dfrac{n!}{x!(n-x)!}p^x(1-p)^{n-x}}=\frac{n-x}{n(1-p)}, \qquad x=0,1,2,\dots,n.$$

Ας υποθέσουμε τώρα ότι η παράμετρος p μεταβάλλεται με το n, δηλαδή $p=p_n$.

Επιπλέον, ας δεχθούμε ότι $np_n\to\lambda$ $(\lambda>0)$ καθώς $n\to+\infty$, δηλαδή $p_n\to 0$ με ταχύτητα σύγκλισης $\frac{1}{n}$. Τότε $w_{\mathrm{X},n}\to 1$ καθώς $n\to+\infty$ και από την Παρατήρηση 3.1 έπεται ότι

$$d_{TV}(\Delta\iota\omega\nu(n,p_n),P(\lambda))\to 0 \ \text{ καθώς } \ n\to+\infty.$$

Πιο συγκεκριμένα, από το Πόρισμα 3.3 προκύπτει

$$d_{TV}(\Delta\iota\omega\nu(n,p),P(\lambda))\le\frac{1-e^{-\lambda}}{\lambda}E\left|np(1-p)\frac{n-\mathrm{X}}{n(1-p)}-\lambda\right|+\min\{1,\lambda^{-1/2}\}|np-\lambda|.$$

Θέτοντας όπου λ το np στην παραπάνω σχέση, έχουμε

$$E\left|np(1-p)\frac{n-\mathrm{X}}{n(1-p)}-np\right| = E\left|p(n-\mathrm{X})-np\right| = E\left|pn-p\mathrm{X}-np\right| = pE(\mathrm{X}) = pnp = np^2$$

Άρα $\quad d_{TV}(\Delta\iota\omega\nu(n,p),P(np)) \le \frac{1-e^{-np}}{np}np^2,$ οπότε

$$d_{TV}(\Delta\iota\omega\nu(n,p),P(np)) \le (1-e^{-np})p.$$

Το τελευταίο φράγμα τείνει στο 0 και η τάξη σύγκλισης είναι $\frac{1}{n}$.

Παράδειγμα 3.2

Ας θεωρήσουμε μια τυχαία μεταβλητή $\mathrm{X}\sim Pascal(n,p)$ $(0<p<1)$ (Αρνητική Διωνυμική) με συνάρτηση πιθανότητας

$$p(x) = \binom{n+x-1}{x}p^n(1-p)^x, \qquad x=0,1,2,\ldots,$$

μέση τιμή $\mu = \frac{n(1-p)}{p}$ και διασπορά $\sigma^2 = \frac{n(1-p)}{p^2}$. Με επαγωγή (ως προς x) αποδεικνύεται ότι

$$h_{\mathrm{X}}(x) = \frac{(n+x)!}{x!(n-1)!}p^{n-1}(1-p)^{x+1}, \qquad x=0,1,2,\ldots$$

Συνεπώς, $p^*(x) = \frac{h_{\mathrm{X}}(x)}{\sigma^2} = \binom{n+x}{x}p^{n+1}(1-p)^x$ για $x=0,1,2,\ldots,$ δηλαδή η τυχαία μεταβλητή $\mathrm{X}^*\sim Pascal(n+1,p)$. Επιπλέον ισχύει $w_{\mathrm{X}}(x) = \frac{p^*(x)}{p(x)} = \frac{p}{n}(n+x) \qquad x=0,1,2,\ldots$

Ας υποθέσουμε τώρα ότι η παράμετρος p μεταβάλλεται με το n, δηλαδή $p=p_n$.

Επιπλέον, ας δεχθούμε ότι $n(1-p_n)\to\lambda$ $(\lambda>0)$ καθώς $n\to+\infty$, δηλαδή $p_n\to 1$ με ταχύτητα σύγκλισης $\frac{1}{n}$. Τότε $w_{\mathrm{X},n}\to 1$ καθώς $n\to+\infty$ και από την Παρατήρηση 3.1 έπεται ότι

$$d_{TV}(Pascal(n,p_n),P(\lambda))\to 0 \text{ καθώς } n\to+\infty.$$

Πιο συγκεκριμένα, από το Πόρισμα 3.3 προκύπτει

$$d_{TV}(Pascal(n,p),P(\lambda)) \le \frac{1-e^{-\lambda}}{\lambda}E\left|\frac{n(1-p)}{p^2}\frac{p}{n}(n+\mathrm{X})-\lambda\right| + \min\{1,\lambda^{-1/2}\}\left|\frac{n(1-p)}{p}-\lambda\right|.$$

Θέτοντας όπου λ το $\dfrac{n(1-p)}{p}$ στην παραπάνω σχέση, έχουμε

$$E\left|\frac{n(1-p)}{p^2}\frac{p}{n}(n+\mathrm{X})-\frac{n(1-p)}{p}\right|=E\left|\frac{1-p}{p}(n+\mathrm{X})-\frac{n(1-p)}{p}\right|=\frac{1-p}{p}E(\mathrm{X})=\frac{n(1-p)^2}{p^2}.$$

Άρα
$$d_{TV}(Pascal(n,p),P((n-np)/p))\le(1-e^{-(n-np)/p})\frac{1-p}{p}.$$

Το τελευταίο φράγμα τείνει στο 0 και η τάξη σύγκλισης είναι $\dfrac{1}{n}$.

Παράδειγμα 3.3

Ας θεωρήσουμε μια τυχαία μεταβλητή $\mathrm{X}\sim \Upsilon\pi\varepsilon\rho\gamma\varepsilon\omega\mu\varepsilon\tau\rho\iota\kappa\acute{\eta}(n,m,r)$ με συνάρτηση πιθανότητας

$$p(x)=\frac{\binom{m}{x}\binom{n-m}{r-x}}{\binom{n}{r}},\qquad \max\{0,r-n+m\}\le x\le\min\{r,m\},$$

μέση τιμή $\mu=\dfrac{mr}{n}$ και διασπορά $\sigma^2=\dfrac{mr(n-r)(n-m)}{n^2(n-1)}$. Με επαγωγή (ως προς x) αποδεικνύεται ότι

$$h_{\mathrm{X}}(x)=\frac{m!r!(n-m)!(n-r)!}{nn!x!(m-x-1)!(r-x-1)!(n-m-r+x)!},$$

για όλα τα κατάλληλα x.

Συνεπώς, $p^*(x)=\dfrac{h_{\mathrm{X}}(x)}{\sigma^2}=\dfrac{\binom{m-1}{x}\binom{n-m-1}{r-x-1}}{\binom{n-2}{r-1}}$, δηλαδή η τυχαία μεταβλητή

$\mathrm{X}^*\sim\Upsilon\pi\varepsilon\rho\gamma\varepsilon\omega\mu\varepsilon\tau\rho\iota\kappa\acute{\eta}(n-2,m-1,r-1)$. Επιπλέον ισχύει $w_{\mathrm{X}}(x)=\dfrac{p^*(x)}{p(x)}=\dfrac{n(n-1)(m-x)(r-x)}{mr(n-m)(n-r)}$.

Ας υποθέσουμε τώρα ότι οι παράμετροι m και r εξαρτώνται από το n, δηλαδή $m=m_n$

και $r=r_n$.

Επιπλέον, ας δεχθούμε ότι $\dfrac{m_n r_n}{n}\to\lambda$ $(\lambda>0)$ καθώς $n,m_n,r_n\to+\infty$. Τότε $w_{\mathrm{X},n}\to 1$ καθώς $n\to+\infty$ και από την Παρατήρηση 3.1 έπεται ότι

$$d_{TV}(\Upsilon\pi\varepsilon\rho\gamma(n,m_n,r_n),P(\lambda)) \to 0 \quad \text{καθώς} \quad n,m,r_n \to +\infty.$$

Πιο συγκεκριμένα, από το Πόρισμα 3.3 προκύπτει

$$d_{TV}(\Upsilon\pi\varepsilon\rho\gamma(n,m,r),P(\lambda)) \le \frac{1-e^{-\lambda}}{\lambda} E\left|\frac{mr(n-r)(n-m)}{n^2(n-1)}\frac{n(n-1)(m-\mathrm{X})(r-\mathrm{X})}{mr(n-m)(n-r)} - \lambda\right| + \min\{1,\lambda^{-1/2}\}\left|\frac{mr}{n} - \lambda\right|.$$

Θέτοντας όπου λ το $\dfrac{mr}{n}$ στην παραπάνω σχέση, έχουμε

$$E\left|\frac{mr(n-r)(n-m)}{n^2(n-1)}\frac{n(n-1)(m-\mathrm{X})(r-\mathrm{X})}{mr(n-m)(n-r)} - \frac{mr}{n}\right| = \frac{mr}{n}\left[\frac{m+r}{n} - \frac{1}{n}\left(\frac{mr}{n} + \frac{(n-r)(n-m)}{n(n-1)}\right)\right].$$

Άρα $\quad d_{TV}(\Upsilon\pi\varepsilon\rho\gamma(n,m,r),P(mr/n)) \le (1-e^{-mr/n})\left[\frac{m+r}{n} - \frac{1}{n}\left(\frac{mr}{n} + \frac{(n-r)(n-m)}{n(n-1)}\right)\right].$

Κ Ε Φ Α Λ Α Ι Ο 4

Προσέγγιση Κανονικής Κατανομής

4.1. Αποδείξεις Κεντρικού Οριακού Θεωρήματος

Εδώ αποδεικνύεται μια συνελιξιακή ανισότητα για το τυποποιημένο άθροισμα ανεξάρτητων απόλυτα συνεχών τυχαίων μεταβλητών. Με βάση την ανισότητα αυτή, που ισχύει με την προυπόθεση ότι το στήριγμα –support των τυχαίων μεταβλητών είναι διάστημα, αποδεικνύεται ότι η τάξη σύγκλισης κατά ολική κύμανση των τυποποιημένων μερικών αθροισμάτων $\frac{S_n - n\mu}{\sigma\sqrt{n}}$ των ανεξάρτητων και ισόνομων τυχαίων μεταβλητών $\mathrm{X}, \mathrm{X}_1, \mathrm{X}_2, \ldots$ προς την τυποποιημένη κανονική Ζ είναι τουλάχιστον $O\left(n^{-\frac{1}{2}}\right)$. Επίσης δίνεται τιμή της σταθεράς c_X για την οποία $d_{TV}\left(\frac{S_n - n\mu}{\sigma\sqrt{n}}, Z\right) \le \frac{c_X}{\sqrt{n}}$.

Έστω μια κανονικοποιημένη απόλυτα συνεχής τ.μ. X με μια συνάρτηση κατανομής F και πυκνότητα f και έστω Z μια τυποποιημένη κανονική τ.μ. με συνάρτηση κατανομής Φ.
Οι Cacoullos, Papathanasiou and Utev (CPU, 1994) βρήκαν το ακόλουθο φράγμα για την απόσταση ολικής κύμανσης $d_{TV}(F,\Phi) = \sup\{|F(A) - \Phi(A)|, \text{A Borel}\}$ μεταξύ F και Φ
(ή X και Z)

$$d_{TV}(F,\Phi) \le 2E\,|\,w(X) - 1| \le 2\sqrt{Var(w(X))} \tag{4.1}$$

όπου η συνάρτηση $w(\cdot)$ ορίζεται για κάθε x στο στήριγμα της X από τη σχέση
$w(x)f(x) = -\int_{-\infty}^{x} tf(t)dt = \int_{x}^{\infty} tf(t)dt.$

Σημειώνεται ότι η δεύτερη ανισότητα προκύπτει από τη σχέση $E[w(X)] = 1,$ από τη βασική ταυτότητα συνδιακύμανσης $Cov(\mathrm{X}, g(\mathrm{X})) = E(\mathrm{X}g(\mathrm{X})) = E[w(\mathrm{X})g'(\mathrm{X})]$ και από τη σχέση $(E\,|\,w-1|)^2 \le E(w-1)^2$. Επομένως έχουμε

$$(E\,|\,w-1|)^2 \le E(w-1)^2$$

$$\Rightarrow E\,|\,w-1| \leq \sqrt{E(w-1)^2}$$

$$\Rightarrow E\,|\,w-1| \leq \sqrt{Var[w(X)]}$$

Ο παράγοντας 2 μπορεί να αντικατασταθεί από $\frac{3}{2}$, όπως δείχνουμε στο ακόλουθο Λήμμα.

Λήμμα 4.1

Αν X έχει μέση τιμή 0, διασπορά 1, πυκνότητα f και συνάρτηση κατανομής F, τότε

$$d_{TV}(F,\Phi) \leq \frac{3}{2}E|w(\mathrm{X})-1| \leq \frac{3}{2}\sqrt{Var[w(\mathrm{X})]}.$$

Απόδειξη

Για ένα αυθαίρετο Borel σύνολο A, θεωρούμε τη συνάρτηση $\psi_{\mathrm{A}}(x),\ \ x \in R$, ορισμένη ως

$$\psi_{\mathrm{A}}(x) = \exp\left(\frac{x^2}{2}\right)\int_{-\infty}^{x}\big(I(t \in \mathrm{A}) - \Phi(\mathrm{A})\big)\exp\left(-\frac{t^2}{2}\right)dt.$$

Εφαρμόζοντας τη βασική ταυτότητα συνδιακύμανσης με $g = \psi_{\mathrm{A}}$ και χρησιμοποιώντας την

(1.3) του CPU (1994) $\psi'(x) = x\psi(x) + (I_{\mathrm{A}}(x) - \Phi(\mathrm{A}))$ έχουμε

$$F(\mathrm{A}) - \Phi(\mathrm{A}) = E\{\psi'_{\mathrm{A}}(\mathrm{X}) - \mathrm{X}\psi_{\mathrm{A}}(\mathrm{X})\}$$

$$= E\{\psi'_{\mathrm{A}}(\mathrm{X})[1-w(\mathrm{X})]\}$$

$$= E\{I(\mathrm{X} \in \mathrm{A})[1-w(\mathrm{X})]\} + E\{\mathrm{X}\psi_{\mathrm{A}}(\mathrm{X})[1-w(\mathrm{X})]\}.$$

Παρατηρούμε ότι

$$E\{I(\mathrm{X} \in \mathrm{A})[1-w(\mathrm{X})]\} = \int_{\mathrm{A}}[1-w(\mathrm{X})]dF(x)$$

$$\leq \int_{\mathrm{A}_0}[1-w(\mathrm{X})]dF(x) = \frac{1}{2}E|w(\mathrm{X})-1|, \qquad (4.2)$$

όπου $\mathrm{A}_0 = \{x \in R : w(x) \leq 1\}$.

Από την άλλη μεριά,

$$E\{\mathrm{X}\psi_{\mathrm{A}}(\mathrm{X})[1-w(\mathrm{X})]\} \leq E|\mathrm{X}\psi_{\mathrm{A}}(\mathrm{X})||1-w(\mathrm{X})| \leq E|1-w(\mathrm{X})|, \qquad (4.3)$$

επειδή $|x\psi_{\mathrm{A}}(x)| \leq (|x|/\varphi(x))\min\{\Phi(x), 1-\Phi(x)\} \leq 1$ για όλα τα x και A (φ είναι η

τυποποιημένη κανονική πυκνότητα)

Τώρα η πρώτη ανισότητα προκύπτει από τις σχέσεις (4.2) και (4.3) και από το γεγονός ότι $\sup_{\mathrm{A}}|F(\mathrm{A})-\Phi(\mathrm{A})|=\sup_{\mathrm{A}}[F(\mathrm{A})-\Phi(\mathrm{A})],$ και η δεύτερη ανισότητα είναι προφανής επειδή $E[w(\mathrm{X})]=1.$

Το Κ.Ο.Θ. προκύπτει από την (4.1) και το εξής.

Θεώρημα 4.1 (CPU 1992)

Έστω $\mathrm{X}_1, \ldots, \mathrm{X}_n, \ldots$ ανεξάρτητες, ισόνομες, απόλυτα συνεχείς τυχαίες μεταβλητές με

$$E[\mathrm{X}_1]=0, \ E[\mathrm{X}_1^{\,2}]=1 \ \text{ και } \ E[w^2(\mathrm{X}_1)]<\infty,$$

όπου w η w-συνάρτηση της X_1.

Τότε $Var[w_n(S_n)] \to 0$ καθώς $n \to \infty$,

όπου $S_n=(\mathrm{X}_1+\ \ldots+\mathrm{X}_n)/\sqrt{n}$ και w_n είναι η w-συνάρτηση του S_n.

Η απόδειξη δίνεται στο CPU, 1992 κάτω από την υπόθεση ότι $Var[w(\mathrm{X})]<\infty$, εφαρμόζοντας ιδιότητες των w-συναρτήσεων. Μια απλούστερη απόδειξη δίνεται παρακάτω.

Για την απόδειξη, χρησιμοποιούμε το παρακάτω Λήμμα.

Λήμμα 4.2

Κάτω από τις υποθέσεις του Θεωρήματος, για $j \le n$ έχουμε

$$E[w_j(S_j)w_n(S_n)]=E[w_n^{\,2}(S_n)].$$

Απόδειξη

Έστω $\mathrm{A}_j=E[w_j(S_j)w_n(S_n)].$ Από τη βασική ταυτότητα συνδιακύμανσης

$Cov[\mathrm{X}, g(\mathrm{X})]=\sigma^2 E[w(\mathrm{X})g'(\mathrm{X})],$ παίρνουμε ότι για αυθαίρετη συνάρτηση g

$$E[S_j g(S_j)]=E[w_j(S_j)g'(S_j)].$$

Από την άλλη μεριά, χρησιμοποιώντας την ίδια ταυτότητα,

$$E[S_j g(S_j)]=E\left[\left(\sqrt{\frac{j-1}{j}}S_{j-1}+\sqrt{\frac{1}{j}}\mathrm{X}_j\right)g(S_j)\right]$$

$$=\sqrt{\frac{j-1}{j}}E\{E[S_{j-1}g(S_j)\,|\,\mathrm{X}_j]\}+\sqrt{\frac{1}{j}}E\{E[\mathrm{X}_j g(S_j)\,|\,S_{j-1}]\}$$

$$= \frac{j-1}{j} E[w_{j-1}(S_{j-1}) g'(S_j)] + \frac{1}{j} E[w(\mathrm{X}_j) g'(S_j)].$$

Επομένως, $E[w_j(S_j) g'(S_j)] = \frac{j-1}{j} E[w_{j-1}(S_{j-1}) g'(S_j)] + \frac{1}{j} E[w(\mathrm{X}_j) g'(S_j)].$

Εφαρμόζοντας την παραπάνω σχέση για $g'(x) = w_n\left(\sqrt{\frac{j}{n}} x + \sqrt{\frac{n-j}{n}} S_{n-j}{}^*\right)$, όπου

$S_{n-j}{}^* = (\mathrm{X}_{j+1} + \ldots + \mathrm{X}_n)/\sqrt{(n-j)}$ για $j < n$ και $S_0{}^* = 0$, παίρνουμε

$E[w_j(S_j) w_n(S_n) \mid S_{n-j}{}^*] = \frac{j-1}{j} E[w_{j-1}(S_{j-1}) w_n(S_n) \mid S_{n-j}{}^*] + \frac{1}{j} E[w(\mathrm{X}_j) w_n(S_n) \mid S_{n-j}{}^*].$

Επομένως, χρησιμοποιώντας το γεγονός ότι $E[w(\mathrm{X}_j) w_n(S_n)] = E[w_1(S_1) w_n(S_n)]$ για κάθε j, συμπεραίνουμε ότι $j\mathrm{A}_j = (j-1)\mathrm{A}_{j-1} + \mathrm{A}_1$. Αυτό σημαίνει ότι $\mathrm{A}_1 = \ldots = \mathrm{A}_n$ και η απόδειξη ολοκληρώνεται.

Απόδειξη Θεωρήματος 4.1

Έστω $S_n = Var[w_n(S_n)]$. Προκύπτει από το Λήμμα 4.2 ότι

$$\sigma_n - \sigma_{n+1} = E[w_n{}^2(S_n)] - E[w_{n+1}{}^2(S_{n+1})]$$

$$= E[w_n(S_n) - w_{n+1}(S_{n+1})]^2.$$

Επομένως, σ_n είναι φθίνουσα και συγκλίνει (επειδή $\sigma_1 < \infty$ από τις υποθέσεις).

Έτσι, $\sigma_n - \sigma_{2n} = E[w_{2n}(S_n) - w_n(S_n)]^2 \to 0.$

Συνεπώς, $E[w_{2n}(S_{2n}) - w_n(S_n)]^2 \geq E[w_{2n}(S_{2n}) - E[w_{2n}(S_{2n}) \mid S_n]]^2$

$$= E\{Var[w_{2n}(S_{2n}) \mid S_n]\} \to 0.$$

Επίσης, αν $S_n{}^* = (\mathrm{X}_{n+1} + \ldots + \mathrm{X}_{2n})/\sqrt{n}$, παίρνουμε

$$E[w_{2n}(S_{2n}) - E[w_{2n}(S_{2n}) \mid S_n]]^2$$

$$= E\{E([w_{2n}(S_{2n}) - E[w_{2n}(S_{2n}) \mid S_n]]^2 \mid S_n{}^*)\}$$

$$\geq E\{E^2([w_{2n}(S_{2n}) - E[w_{2n}(S_{2n}) \mid S_n]] \mid S_n{}^*)\}$$

$$= E\left[E[w_{2n}(S_{2n}) \mid S_n{}^*] - 1\right]^2$$

$$= E\left[E[w_{2n}(S_{2n}) \mid S_n] - 1\right]^2$$

$$= Var\{E[w_{2n}(S_{2n}) \mid S_n]\} \to 0.$$

Επομένως, $\sigma_{2n} = E\{Var[w_{2n}(S_{2n}) \mid S_n]\} + Var\{E[w_{2n}(S_{2n}) \mid S_n]\} \to 0$ από τα παραπάνω επιχειρήματα, που ολοκληρώνει την απόδειξη.

Στο παρόν κεφάλαιο θα αποδείξουμε ένα αποτέλεσμα που σχετίζεται με τον ρυθμό σύγκλισης στο κεντρικό οριακό θεώρημα.

Ορισμός 4.1:

Έστω η οικογένεια C των τυχαίων μεταβλητών Χ που ικανοποιούν τις ακόλουθες συνθήκες:

Α) $E(\mathrm{X}) = 0,\ E(\mathrm{X}^2) = 1$ (χωρίς βλάβη της γενικότητας)

Β) Χ απόλυτα συνεχής με πυκνότητα f και το στήριγμα της είναι είναι διάστημα, όχι απαραίτητα πεπερασμένο.

Γ) Η w-συνάρτηση της Χ έχει πεπερασμένη δεύτερη ροπή, $E[w^2(\mathrm{X})] < \infty$.

Θεώρημα 4.2

Υποθέτουμε ότι $\mathrm{X}_1, \ldots, \mathrm{X}_n, \ldots$ είναι ανεξάρτητες ισόνομες τυχαίες μεταβλητές με $\mathrm{X}_1 \in C$.

Τότε
$$d_{TV}\left(\frac{S_n}{\sqrt{n}}, \mathrm{Z}\right) \le \frac{c}{\sqrt{n}},$$

όπου η σταθερά c μπορεί να εκλεγεί ως

$$c = 2\sqrt{Var[w_1(\mathrm{X}_1)]}$$

και w_1 η w-συνάρτηση που αντιστοιχεί στη X_1.

Η απόδειξη του παραπάνω θεωρήματος βασίζεται στη συνελιξιακή ανισότητα που δίνεται από το παρακάτω Λήμμα.

Λήμμα 4.3

Έστω Χ,Υ ανεξάρτητες τυχαίες μεταβλητές με $\mathrm{X} \in C,\ \mathrm{Y} \in C$ και θεωρούμε την τυχαία μεταβλητή $S = a\mathrm{X} + b\mathrm{Y}$ όπου α,b είναι πραγματικές σταθερές τέτοιες ώστε $\alpha^2 + b^2 = 1$.

Τότε $S \in C$ και

$$Var[w_S(S)] \le \alpha^4 Var[w_{\mathrm{X}}(\mathrm{X})] + b^4 Var[w_{\mathrm{Y}}(\mathrm{Y})],$$

όπου $w_{\mathrm{X}}, w_{\mathrm{Y}}, w_S$ είναι οι w-συναρτήσεις των Χ, Υ, S αντίστοιχα.

Απόδειξη

Πρώτα παρατηρούμε ότι η S έχει πυκνότητα με στήριγμα διάστημα, μέση τιμή 0 και διασπορά 1. Έτσι η w-συνάρτηση w_S ορίζεται από τη σχέση $w(x)f(x) = -\int_{-\infty}^{x} tf(t)dt = \int_{x}^{\infty} tf(t)dt.$

Για κάθε απόλυτα συνεχής συνάρτηση g ισχύει η ακόλουθη ταυτότητα συνδιακύμανσης

$$E[Sg(S)] = E[w_S(S)g'(S)],$$

όπου g' η παράγωγος της g.

Από την άλλη μεριά

$$\begin{aligned} E[Sg(S)] &= E[(\alpha \mathrm{X} + b\mathrm{Y})g(a\mathrm{X} + b\mathrm{Y})] \\ &= \alpha E\{E[\mathrm{X}g(\alpha \mathrm{X} + b\mathrm{Y}) \mid \mathrm{Y}]\} + bE\{E[\mathrm{Y}g(\alpha \mathrm{X} + b\mathrm{Y}) \mid \mathrm{X}]\} \\ &= \alpha^2 E[w_{\mathrm{X}}(\mathrm{X})g'(S)] + b^2 E[w_{\mathrm{Y}}(\mathrm{Y})g'(S)]. \end{aligned}$$

Άρα έχουμε

$$E[w_S(S)g'(S)] = E[(\alpha^2 w_{\mathrm{X}}(\mathrm{X}) + b^2 w_{\mathrm{Y}}(\mathrm{Y}))g'(S)].$$

Για $g'(x) = w_S(x)\mathrm{I}\{w_S(x) \le N\}$ έχουμε

$$\begin{aligned} \left(E[w_S{}^2(S)I\{w_S(S) \le N\}]\right)^2 &= \left(E[(\alpha^2 w_{\mathrm{X}}(\mathrm{X}) + b^2 w_{\mathrm{Y}}(\mathrm{Y}))w_S(S)I\{w_S(S) \le N\}]\right)^2 \\ &\le E[(\alpha^2 w_{\mathrm{X}}(\mathrm{X}) + b^2 w_{\mathrm{Y}}(\mathrm{Y}))^2]E[w_S{}^2(S)I\{w_S(S) \le N\}] \end{aligned}$$

από την ανισότητα Cauchy-Schwarz.

Άρα $E[w_S{}^2(S)I\{w_S(S) \le N\}] \le E[(\alpha^2 w_{\mathrm{X}}(\mathrm{X}) + b^2 w_{\mathrm{Y}}(\mathrm{Y}))^2].$

Παίρνοντας το όριο καθώς $N \to \infty$ έχουμε ότι

$$E[w_S{}^2(S)] < \infty \text{ και τότε } S \in C.$$

Επομένως $E[w_S{}^2(S)] \le E[(\alpha^2 w_{\mathrm{X}}(\mathrm{X}) + b^2 w_{\mathrm{Y}}(\mathrm{Y}))^2]$ το οποίο είναι ισοδύναμο με $Var[w_S(S)] \le \alpha^4 Var[w_{\mathrm{X}}(\mathrm{X})] + b^4 Var[w_{\mathrm{Y}}(\mathrm{Y})],$ επειδή $E[w_{\mathrm{X}}(\mathrm{X})] = 1$ για κάθε τυχαία μεταβλητή X.

Παρατήρηση 4.1

Προφανώς το Λήμμα 4.3, όταν $S=\sum_{i=1}^{n}\alpha_i X_i$ με $\sum_{i=1}^{n}\alpha_i^{\,2}=1$, δίνει την ανισότητα

$$Var[w_S(S)]\le\sum_{i=1}^{n}\alpha_i^{\,4}Var[w_{X_i}(X_i)].$$

Πόρισμα 4.1

Κάτω από τις συνθήκες του Λήμματος 4.3 έχουμε

$$E[w_S^{\,2}(S)]\le\alpha^2E[w_X^{\,2}(X)]+b^2E[w_Y^{\,2}(Y)],$$

με ισότητα αν και μόνο αν X,Y και S τυποποιημένες κανονικές, δοθέντος $ab\ne 0$.

Απόδειξη

Εύκολα βλέπουμε ότι

$$\alpha^2E[w_X^{\,2}(X)]+b^2E[w_Y^{\,2}(Y)]-E[(\alpha^2w_X(X)+b^2w_Y(Y))^2]$$

$$=\alpha^2E[w_X^{\,2}(X)]+b^2E[w_Y^{\,2}(Y)]-E[a^4w_X^{\,2}(X)+b^4w_Y^{\,2}(Y)+2\alpha^2b^2w_X(X)w_Y(Y)]$$

$$=\alpha^2E[w_X^{\,2}(X)]+b^2E[w_Y^{\,2}(Y)]-\alpha^4E[w_X^{\,2}(X)]-b^4E[w_Y^{\,2}(Y)]-2\alpha^2b^2E[w_X(X)w_Y(Y)]$$

$$=\alpha^2(1-\alpha^2)E[w_X^{\,2}(X)]+b^2(1-b^2)E[w_Y^{\,2}(Y)]-2\alpha^2b^2E[w_X(X)]E[w_Y(Y)]$$

$$=\alpha^2b^2E[w_X^{\,2}(X)]+b^2\alpha^2E[w_Y^{\,2}(Y)]-2\alpha^2b^2\cdot1\cdot1$$

$$=\alpha^2b^2\left\{E[w_X^{\,2}(X)]+E[w_Y^{\,2}(Y)]-2\right\}$$

$$=\alpha^2b^2\left\{E[w_X^{\,2}(X)]-1+E[w_Y^{\,2}(Y)]-1\right\}$$

$$=\alpha^2b^2\left\{Var[w_X(X)]+Var[w_Y(Y)]\right\}\ge 0.$$

Άρα $E[(a^2w_X(X)+b^2w_Y(Y))^2]\le\alpha^2E[w_X^{\,2}(X)]+b^2E[w_Y^{\,2}(Y)]$,

οπότε $E[w_S^{\,2}(S)]\le\alpha^2E[w_X^{\,2}(X)]+b^2E[w_Y^{\,2}(Y)]$.

Πόρισμα 4.2

Αν $X_1, \ldots, X_n, \ldots$ είναι ανεξάρτητες ισόνομες τυχαίες μεταβλητές τέτοιες ώστε $X_1 \in C$, τότε $S_n = (X_1 + \ldots + X_n)/\sqrt{n} \in C$. Επιπλέον, για την ακολουθία $\sigma_n = Var[w_n(S_n)]$, έχουμε

$$\sigma_n \le \frac{\sigma_1}{n},$$

όπου w_n είναι η w-συνάρτηση της S_n.

Απόδειξη

Ο ισχυρισμός προκύπτει από την παρατήρηση με $\alpha_1 = \alpha_2 = \ldots = \alpha_n = \dfrac{1}{\sqrt{n}}$.

Απόδειξη Θεωρήματος 4.2

Έχουμε $d_{TV}\left(\dfrac{S_n}{\sqrt{n}}, Z\right) \le 2\sqrt{\sigma_n}$ και $\sigma_n \le \dfrac{\sigma_1}{n}$, άρα $d_{TV}\left(\dfrac{S_n}{\sqrt{n}}, Z\right) \le 2\dfrac{\sqrt{\sigma_1}}{\sqrt{n}}$.

4.2. Εφαρμογές σε Τυχαία Αθροίσματα

Τώρα θα δώσουμε κάποιες εφαρμογές των παραπάνω αποτελεσμάτων σε τυχαία αθροίσματα. Αποδεικνύεται ότι κάτω από κατάλληλες συνθήκες (στην πραγματικότητα, όταν το N είναι πολύ μεγάλο), το Κεντρικό Οριακό Θεώρημα συνεχίζει να ισχύει για τα τυποποιημένα αθροίσματα N ανεξάρτητων και ισόνομων τυχαίων μεταβλητών. Σημειώνεται, παρόλα αυτά, ότι εδώ ο όρος ¨τυποποιημένα¨ αθροίσματα έχει ένα κάπως διαφορετικό νόημα με την έννοια ότι

$$S_N = \frac{1}{\sqrt{N}}(X_1 + \ldots + X_N)$$

(όπου $S_0 = 0$ από ορισμό) δεν χρειάζεται να έχει διασπορά 1.

Πρώτα αποδεικνύουμε το παρακάτω βοηθητικό αποτέλεσμα.

Λήμμα 4.4

Υποθέτουμε ότι X, $X_1, \ldots, X_i, \ldots$ είναι ανεξάρτητες ισόνομες τυχαίες μεταβλητές όπως στο Θεώρημα 4.1. Τότε, για κάθε μη αρνητική ακέραια τυχαία μεταβλητή N, ανεξάρτητη των $\{X_i\}$,

$$d_{TV}(F_{\mathrm{N}},\Phi) \le P(\mathrm{N} \le m) + c/\sqrt{m+1}, \;\; m=0,1,\dots ,$$

όπου F_{N} η συνάρτηση κατανομής του S_{N} και η σταθερά c μπορεί να εκλεγεί όπως στο Θεώρημα 4.2.

Απόδειξη

Για ένα αυθαίρετο Borel σύνολο A έχουμε, από τη σχέση $d_{TV}(F_n,\Phi) \le c/\sqrt{n}$,

$$\left|P(S_{\mathrm{N}} \in \mathrm{A}) - \Phi(\mathrm{A})\right| \le \sum_{n=0}^{\infty} P(\mathrm{N}=n)\left|P[S_{\mathrm{N}} \in \mathrm{A} \mid \mathrm{N}=n] - \Phi(\mathrm{A})\right|$$

$$\le \sum_{n=0}^{m} P(\mathrm{N}=n) + \sum_{n=m+1}^{\infty} cP(\mathrm{N}=n)/\sqrt{n},$$

και η απόδειξη ολοκληρώνεται.

Χρησιμοποιώντας το παραπάνω Λήμμα, έχουμε το ακόλουθο Θεώρημα.

Θεώρημα 4.3

Έστω η ακολουθία $\{\mathrm{X}_i\}$ ικανοποιεί τις υποθέσεις του Λήμματος 4.4 και υποθέτουμε ότι οι μη αρνητικές ακέραιες τυχαίες μεταβλητές N_n τείνουν κατά πιθανότητα στο άπειρο (με την έννοια ότι για κάθε $m > 0$, $P(\mathrm{N}_n > m) \to 1$ καθώς $n \to \infty$). Τότε

$$d_{TV}(F_n,\Phi) \to 0, \;\; \text{καθώς} \;\; n \to \infty,$$

όπου F_n η συνάρτηση κατανομής του S_{N_n}.

Απόδειξη

Προκύπτει από το Λήμμα 4.4 ότι για κάθε $\varepsilon > 0$ (αυθαίρετα μικρό) και $m > 0$ (αυθαίρετα μεγάλο),

$$d_{TV}(F_n,\Phi) \le \varepsilon + c/\sqrt{m+1} \quad \text{όταν} \quad n > n_0(\varepsilon, m)$$

και ο ισχυρισμός προκύπτει από την αυθαιρεσία των m και ε.

Πόρισμα 4.3

Αν $\mathrm{N}_n/\alpha_n \to \Theta$ κατά πιθανότητα, όπου Θ μια θετική τυχαία μεταβλητή και $\alpha_n \to \infty$.

Απόδειξη

Είναι εύκολο να επαληθεύσουμε τις συνθήκες του Θεωρήματος 4.3.

Το ακόλουθο Θεώρημα δίνει εκτιμήσεις για τον ρυθμό σύγκλισης, κάτω από ισχυρότερες συνθήκες για την $\{\mathrm{N}_n\}$.

Θεώρημα 4.4

Έστω οι ακολουθίες $\{\mathrm{X}_i\}$ και $\{\alpha_n\}$ όπως στο Πόρισμα 4.3 και υποθέτουμε ότι

$$\sup_n\left\{\sqrt{\alpha_n}\,E\left|\frac{\mathrm{N}_n}{\alpha_n}-\Theta\right|\right\}=C<\infty,$$

όπου Θ είναι μια θετική τυχαία μεταβλητή τέτοια ώστε $P(\Theta\geq\delta)=1$ για κάποιο $\delta>0$. Τότε

$$d_{TV}(F_n,\Phi)\leq O(\alpha_n^{-1/2}),\ \text{ καθώς } n\to\infty.$$

Απόδειξη

Ανισότητα Markov:

Για τυχούσα τυχαία μεταβλητή $\mathrm{X}\geq 0$, και για κάθε $\alpha>0$,

$$P(\mathrm{X}\geq\alpha)\leq\frac{1}{\alpha}E(\mathrm{X}).$$

Χρησιμοποιώντας την ανισότητα του Markov, εύκολα προκύπτει ότι

$$P(2\mathrm{N}_n\leq\delta\alpha_n)\leq P\left[\left|\frac{\mathrm{N}_n}{\alpha_n}-\Theta\right|\geq\frac{\delta}{2}\right]\leq\frac{2C}{\delta\sqrt{\alpha_n}}.$$

Εφαρμόζοντας το Λήμμα 4.4 με $m=[\delta\alpha_n/2]$, παίρνουμε

$$d_{TV}(F_n,\Phi)\leq\frac{2C/\delta+c\sqrt{2}/\sqrt{\delta}}{\sqrt{\alpha_n}}$$

και η απόδειξη ολοκληρώνεται.

ΚΕΦΑΛΑΙΟ 5

Μετασχηματισμός Τυχαίας Μεταβλητής Χ

Έστω Χ μια απόλυτα συνεχής τυχαία μεταβλητή με πεπερασμένη διασπορά σ^2.

Τότε υπάρχει μια νέα τυχαία μεταβλητή X^*, η οποία θεωρείται ως μετασχηματισμός της Χ με μονοκόρυφη πυκνότητα, που ικανοποιεί την ταυτότητα συνδιακύμανσης

$$Cov[\mathrm{X}, g(\mathrm{X})] = \sigma^2 E[g'(\mathrm{X}^*)]$$

για κάθε απόλυτα συνεχής συνάρτηση g με παράγωγο g', δοθέντος ότι $E[g'(\mathrm{X}^*)] < \infty$.

Θα αναφέρουμε τις ιδιότητες της τυχαίας μεταβλητής X^* και μια εφαρμογή σε φράγματα διασποράς.

5.1. Ιδιότητες της Τυχαίας Μεταβλητής X^*

Έστω Χ μια απόλυτα συνεχής τυχαία μεταβλητή με πυκνότητα f, μέση τιμή μ, διασπορά σ^2 και στήριγμα $S(\mathrm{X})$. Για λόγους απλότητας, θεωρούμε το στήριγμα μιας απόλυτα συνεχούς τυχαίας μεταβλητής Χ με πυκνότητα f να είναι το σύνολο $S(\mathrm{X}) = \{x : f(x) > 0\}$.

Ορίζουμε X^* μια τυχαία μεταβλητή με πυκνότητα f^*, που δίνεται από τη σχέση

$$f^*(x) = \frac{1}{\sigma^2}\int_{-\infty}^{x} (\mu - t) f(t)dt = \frac{1}{\sigma^2}\int_{x}^{\infty} (t-\mu) f(t)dt.$$

Προφανώς, η f^* είναι μη αρνητική.

Μια εφαρμογή του Θεωρήματος Tonelli δείχνει ότι $\int_{-\infty}^{\infty} f^* = 1$, άρα η f^* είναι πυκνότητα πιθανότητας. Το ακόλουθο Λήμμα συνοψίζει τις ιδιότητες της X^*.

Λήμμα 5.1

i) f^* είναι μια μονοκόρυφη απόλυτα συνεχής πυκνότητα με κορυφή μ και μέγιστη τιμή

$$f^*(\mu) = \frac{E|\mathrm{X}-\mu|}{2\sigma^2}. \tag{5.1}$$

Επιπλέον, η συνάρτηση $(f^*(x))'/(\mu - x)$ (ορισμένη σχεδόν παντού) είναι μη αρνητική και ολοκληρώσιμη.

ii) Για κάθε απόλυτα συνεχή συνάρτηση g με $E\left|g'(\mathrm{X}^*)\right| < \infty$,

$$Cov[\mathrm{X}, g(\mathrm{X})] = \sigma^2 E[g'(\mathrm{X}^*)]. \tag{5.2}$$

iii) Αν η τυχαία μεταβλητή Y ικανοποιεί την ταυτότητα

$$Cov[\mathrm{X}, g(\mathrm{X})] = \sigma^2 E[g'(\mathrm{Y})] \tag{5.3}$$

για κάθε απόλυτα συνεχή συνάρτηση g με $E|g'(\mathrm{Y})| < \infty$, τότε $\mathrm{Y} \overset{d}{=} \mathrm{X}^*$, δηλαδή Y και X^* έχουν την ίδια κατανομή.

iv) $S(\mathrm{X}^*) = (\text{essinf } \mathrm{S(X)}, \text{esssup } \mathrm{S(X)})$, όπου
essinf S(X)= ουσιώδες infinum=$\sup\{x \in R : F(x) = 0\}$,
esssup S(X)=ουσιώδες suprenum=$\inf\{x \in R : F(x) = 1\}$
και F η συνάρτηση κατανομής της X.

v) Για οποιουσδήποτε πραγματικούς αριθμούς $\alpha \neq 0$ και b,

$$(\alpha\mathrm{X} + b)^* \overset{d}{=} a\mathrm{X}^* + b.$$

vi) Για ανεξάρτητες απόλυτα συνεχείς τυχαίες μεταβλητές $\mathrm{X}_1, \mathrm{X}_2$ με μέσες τιμές μ_1, μ_2, διασπορές $\sigma_1{}^2, \sigma_2{}^2$ και οποιουσδήποτε αριθμούς α_1 και α_2 με $\alpha_1\alpha_2 \neq 0$

$$(\alpha_1\mathrm{X}_1 + \alpha_2\mathrm{X}_2)^* \overset{d}{=} B(\alpha_1\mathrm{X}_1{}^* + \alpha_2\mathrm{X}_2) + (1-B)(\alpha_1\mathrm{X}_1 + \alpha_2\mathrm{X}_2{}^*), \tag{5.4}$$

όπου X_1, X_2, $\mathrm{X}_1{}^*$, $\mathrm{X}_2{}^*$ και B είναι ανεξάρτητες με

$$P(B=1) = \frac{\alpha_1{}^2\sigma_1{}^2}{\alpha_1{}^2\sigma_1{}^2 + a_2{}^2\sigma_2{}^2} = 1 - P(B=0).$$

Απόδειξη

i), iv) και v) είναι προφανή. Το ii) προκύπτει από τον ορισμό της X^*, χρησιμοπιώντας το Θεώρημα Fubini επειδή από την υπόθεση ότι $E\left|g'(\mathrm{X}^*)\right| < \infty$, οι μη αρνητικές συναρτήσεις $g_1(x,t) = |g'(x)|(\mu - t) f(t) I(t < x)$ και $g_2(x,t) = |g'(x)|(t-\mu) f(t) I(t > x)$ είναι ολοκληρώσιμες στα σύνολα $(-\infty, \mu] \times (-\infty, \mu]$ και $[\mu, \infty) \times [\mu, \infty)$, αντίστοιχα. iii) προκύπτει από το ii) και από το γεγονός ότι $E[G(\mathrm{X}^*)] = E[G(\mathrm{Y})]$ για κάθε φραγμένη (μετρική) συνάρτηση G. Θεωρώντας vi), παρατηρούμε ότι για $S = \alpha_1\mathrm{X}_1 + \alpha_2\mathrm{X}_2$ και g' φραγμένη, εμείς έχουμε από ii):

$$Cov[S, g(S)] = \sigma^2 E[g'(S^*)],$$

όπου $\sigma^2 = Var(S) = Var(\alpha_1\mathrm{X}_1 + \alpha_2\mathrm{X}_2) = \alpha_1{}^2 Var(\mathrm{X}_1) + \alpha_2{}^2 Var(\mathrm{X}_2) = \alpha_1{}^2\sigma_1{}^2 + \alpha_2{}^2\sigma_2{}^2.$

Από την άλλη μεριά

$$Cov[S, g(S)] = Cov[\alpha_1\mathrm{X}_1 + \alpha_2\mathrm{X}_2, g(\alpha_1\mathrm{X}_1 + \alpha_2\mathrm{X}_2)]$$

$$=\alpha_1 Cov[X_1, g(\alpha_1 X_1+\alpha_2 X_2)]+\alpha_2 Cov[X_2, g(\alpha_1 X_1+\alpha_2 X_2)]$$

$$=\alpha_1^2\sigma_1^2 E[g'(\alpha_1 X_1^*+\alpha_2 X_2)]+\alpha_2^2\sigma_2^2 E[g'(\alpha_1 X_1+\alpha_2 X_2^*)].$$

Προκύπτει ότι για κάθε φραγμένη συνάρτηση G,

$$E\left[G(S^*)\right]=\frac{\alpha_1^2\sigma_1^2}{\sigma^2}E\left[G(\alpha_1 X_1^*+\alpha_2 X_2)\right]+\frac{\alpha_2^2\sigma_2^2}{\sigma^2}E\left[G(\alpha_1 X_1+\alpha_2 X_2^*)\right]$$

$$=E[G\left(B(\alpha_1 X_1^*+\alpha_2 X_2)+(1-B)(\alpha_1 X_1+\alpha_2 X_2^*)\right)],$$

και ολοκληρώνεται η απόδειξη.

Παρατήρηση 5.1

Σημειώνεται ότι ο όρος μονοκόρυφη χρησιμοποιείται για να ορίσουμε μια συνάρτηση $h: R\to R$ με την ιδιότητα ότι υπάρχει κάποιο $m\in R$ τέτοιο ώστε $h(x)$ να μην είναι φθίνουσα για $x\le m$ και να μην είναι αύξουσα για $x\ge m$. Κάθε m με αυτή την ιδιότητα λέγεται κορυφή της h.

Η ταυτότητα (5.2) ισχύει για κάθε μη-φθίνουσα ή μη-αύξουσα απόλυτα συνεχή συνάρτηση g, ακόμα και στην περίπτωση όπου $E[g'(X^*)]=\infty$ (ή $-\infty$ αν g είναι μη-αύξουσα) όπως προκύπτει από το Θεώρημα Tonelli. Σε αυτήν την περίπτωση $E[(X-\mu)(g(X)-g(\mu))]=\infty$ (ή $-\infty$). Αν, όμως, $E[g'(X^*)]=\infty$ και g αυθαίρετη, ίσως $E\,|(X-\mu)(g(X)-g(\mu))|<\infty$, όπως δείχνει το ακόλουθο παράδειγμα.

Παράδειγμα 5.1

Υποθέτουμε ότι η X έχει πυκνότητα $f(x)=\frac{3}{8}\min\{1, x^{-4}\}$, $-\infty<x<\infty$ και

$$g(x)=\sum_{n=1}^{\infty}(2n-1)(|x|-\alpha_{2n-2})I(\alpha_{2n-2}\le|x|<\alpha_{2n-1})+\sum_{n=1}^{\infty}2n(\alpha_{2n}-|x|)I(\alpha_{2n-1}\le|x|<\alpha_{2n}),$$

όπου $\alpha_o=0$ και $\alpha_n=1+\frac{1}{2}+\ldots+\frac{1}{n}$ για $n\ge 1$. Προκύπτει ότι $0\le g(x)\le 1$ και $g(-x)=g(x)$ για κάθε x. Επίσης, η g είναι απόλυτα συνεχής με παράγωγο $g'(x)$ που ικανοποιεί την σχέση $|g'(x)|\ge|x|\ge|g(x)|=g(x)$ για σχεδόν όλα τα x. Επειδή $E[X]=0$, $E|X|=\frac{3}{4}$, $Var(X)=1$ και $S(X)=(-\infty,+\infty)$, συμπεραίνουμε ότι η X^* έχει πυκνότητα

$$f^*(x)=\frac{3}{16}\left[(2-x^2)I(|x|\le 1)+x^{-2}I(|x|>1)\right] \text{ και } E\left|g'(X^*)\right|\ge E\left|X^*\right|=\infty, \text{ ενώ } E|Xg(X)|\le E|X|=\frac{3}{4}.$$

Θεώρημα 5.1

Υποθέτουμε ότι η τυχαία μεταβλητή Y έχει μια μονοκόρυφη απόλυτα συνεχής πυκνότητα *h*.

i) Αν η κορυφή *m* της *h* είναι μοναδική ($h(x) < h(m)$ για όλα τα $x \neq m$), τότε υπάρχει μια τυχαία μεταβλητή X_m έτσι ώστε $X_m{}^* \overset{d}{=} Y$ αν και μόνο αν η συνάρτηση $\dfrac{h'(x)}{(m-x)}$ είναι ολοκληρώσιμη. Επίσης, αν $X_1{}^* \overset{d}{=} Y$ και $X_2{}^* \overset{d}{=} Y$, τότε $X_1 \overset{d}{=} X_2$ (και τότε X_m είναι μοναδική).

ii) Αν $\{x$: x είναι μια κορυφή της $h\}=[\alpha, b]$ με $\alpha < b$, τότε για κάθε $\mu \in (\alpha, b)$, υπάρχει πάντα μια μοναδική τυχαία μεταβλητή X_μ τέτοια ώστε $E[X_\mu] = \mu$ και $X_\mu{}^* \overset{d}{=} Y$. Επίσης, για $\mu = \alpha$ ή $\mu = b$, υπάρχει μια τυχαία μεταβλητή X_μ τέτοια ώστε $E[X_\mu] = \mu$ και $X_\mu{}^* \overset{d}{=} Y$ αν και μόνο αν η συνάρτηση $\dfrac{h'(x)}{(\mu - x)}$ είναι ολοκληρώσιμη. Τελικά, αν $X_1{}^* \overset{d}{=} Y$, $X_2{}^* \overset{d}{=} Y$ και $E[X_1] = E[X_2]$, τότε $X_1 \overset{d}{=} X_2$.

Απόδειξη

i) Αν X είναι μια τυχαία μεταβλητή με μέση τιμή μ, διασπορά σ^2 και πυκνότητα f τέτοια ώστε $X^* \overset{d}{=} Y$. Προκύπτει από Λήμμα 5.1 i) ότι μ πρέπει να είναι μια κορυφή της $f^* = h$ και ότι $\dfrac{h'(x)}{\mu - x}$ είναι ολοκληρώσιμη. Υποθέτουμε τώρα ότι η συνάρτηση $\dfrac{h'(x)}{m-x}$ (ορισμένη σχεδόν παντού) είναι ολοκληρώσιμη. Παρατηρούμε ότι είναι επίσης μη αρνητική (επειδή *m* είναι η κορυφή της *h*) και ορίζουμε τη τυχαία μεταβλητή X_m με πυκνότητα $f_m(x) = \dfrac{h'(x)}{c(m-x)}$, όπου $c = \displaystyle\int_{-\infty}^{\infty} \frac{h'(x)}{m-x} dx > 0.$

Επειδή $\lim\limits_{x\to\infty} h(x) = 0$, $\lim\limits_{x\to-\infty} h(x) = 0$ (επειδή *h* είναι μια μονοκόρυφη πυκνότητα), έχουμε $E[m - X_m] = 0$. Εφαρμόζοντας το θεώρημα του Tonelli έχουμε

$$\int_{-\infty}^{m} (m-x)h'(x)dx = \int_{-\infty}^{m}\int_{x}^{m} h'(x)du dx = \int_{-\infty}^{m} h(u)du,$$

και ομοίως $\displaystyle\int_{m}^{\infty} (x-m)(-h'(x))dx = \int_{m}^{\infty} h(u)du,$

και τότε $Var[\mathrm{X}_m]=\dfrac{1}{c}$. Επομένως, $\mathrm{X}_m{}^* \overset{d}{=} \mathrm{Y}$.

Θα δείξουμε τώρα ότι η X_m είναι μοναδική. Πράγματι, αν $\mathrm{X}^* \overset{d}{=} \mathrm{Y}$ για μια τυχαία μεταβλητή X με μέση τιμή μ, διασπορά σ^2 και πυκνότητα f, προκύπτει από Λήμμα 5.1 i) ότι μ πρέπει να είναι μια κορυφή της $f^*=h$ και επομένως $\mu=m$. Τότε $f(x)=\dfrac{\sigma^2 h'(x)}{m-x}=c\sigma^2 f_m(x)$ για σχεδόν όλα τα x, και επομένως $\mathrm{X}\overset{d}{=}\mathrm{X}_m$.

ii) Επειδή h είναι σταθερά στο $[\alpha,b]$, $h'\equiv 0$ στο (a,b).

Τότε, για όλα τα $\mu\in(\alpha,b)$,

$$\int_{-\infty}^{\infty}\frac{h'(x)}{\mu-x}dx=\int_{-\infty}^{a}\frac{h'(x)}{\mu-x}dx+\int_{b}^{\infty}\frac{-h'(x)}{x-\mu}dx\le\frac{h(a)}{\mu-a}+\frac{h(b)}{b-\mu}<\infty.$$

Θεωρούμε τη τυχαία μεταβλητή X_μ με πυκνότητα

$$f_\mu(x)=\frac{h'(x)}{c_\mu(\mu-x)},\ \text{όπου}\ \ c_\mu=\int_{-\infty}^{\infty}\frac{h'(x)}{\mu-x}dx.$$

Τότε $E[\mathrm{X}_\mu]=\mu$, $Var[\mathrm{X}_\mu]=\dfrac{1}{c_\mu}$ και επομένως $\mathrm{X}_\mu{}^* \overset{d}{=} \mathrm{Y}$.

Εύκολα βλέπουμε ότι αν $\mathrm{X}^* \overset{d}{=} \mathrm{Y}$ και $E[\mathrm{X}]=\mu$ για κάποια τυχαία μεταβλητή X με πυκνότητα f και διασπορά σ^2, τότε $\mathrm{X}\overset{d}{=}\mathrm{X}_\mu$. Πράγματι, προκύπτει από το Θεώρημα 5.1 και Λήμμα 5.1 i) ότι $f(x)=\dfrac{\sigma^2 h'(x)}{\mu-x}=c_\mu\sigma^2 f_\mu(x)$ για σχεδόν όλα τα x. Επομένως, επειδή η μέση τιμή κάθε τυχαίας μεταβλητής X που ικανοποιεί $\mathrm{X}^* \overset{d}{=} \mathrm{Y}$ πρέπει να είναι μια κορυφή της h, είτε $E[\mathrm{X}]=\mu\in(a,b)$ (και επομένως $\mathrm{X}\overset{d}{=}\mathrm{X}_\mu$) ή $E[\mathrm{X}]=a$ (και $\dfrac{h'(x)}{a-x}$ είναι ολοκληρώσιμη) ή $E[\mathrm{X}]=b$ (και $\dfrac{h'(x)}{b-x}$ είναι ολοκληρώσιμη).

Το ακόλουθο παράδειγμα δείχνει ότι όλες οι περιπτώσεις που περιγράφονται στο Θεώρημα 5.1 είναι πιθανές.

Παράδειγμα 5.2

i) Υποθέτουμε ότι η Y έχει τη μονοκόρυφη απόλυτα συνεχής πυκνότητα $h(x)=\frac{1}{3}(\min\{2-|x|,1\})^{+}$ με παράγωγο $h'(x)=-\frac{1}{3}sign(x)I(1<|x|<2)$ για σχεδόν όλα τα x. Τότε, για όλα τα $\mu\in(-1,1)$, η τυχαία μεταβλητή X_μ με πυκνότητα $f_\mu(x)=(3c_\mu|x-\mu|)^{-1}I(1<|x|<2)$, όπου $c_\mu=\frac{1}{3}\log\left[\frac{4-\mu^2}{1-\mu^2}\right]$, ικανοποιεί $E[X]=\mu$ και $X_\mu^{*}\stackrel{d}{=}Y$. Επίσης, $\frac{h'(x)}{\mu-x}$ δεν είναι ολοκληρώσιμη για $\mu=\pm1$.

ii) Υποθέτουμε ότι η Y έχει την μονοκόρυφη απόλυτα συνεχής πυκνότητα

$$h(x)=\frac{6}{19}\left(\min\{x+2,1\}I(-2<x<1)+x(2-x)I(1\le x<2)\right).$$

Τότε, για κάθε $\mu\in(-1,1]$, η τυχαία μεταβλητή X_μ με πυκνότητα

$$f_\mu(x)=\frac{6}{12c_\mu|x-\mu|}(I(-2<x<-1)+2(x-1)I(1<x<2)),$$

όπου $c_\mu=\frac{6}{19}\left(2+\log\frac{\mu+2}{\mu+1}-2(1-\mu)\log\frac{2-\mu}{1-\mu}\right)$, ικανοποιεί την $E[X_\mu]=\mu$ και $X_\mu^{*}\stackrel{d}{=}Y$, ενώ $\frac{h'(x)}{-1-x}$ δεν είναι ολοκληρώσιμη. Παρομοιώς, για τη τυχαία μεταβλητή $W=-Y$ με πυκνότητα $h(-x)$, τότε υπάρχει μια τυχαία μεταβλητή R_μ τέτοια ώστε $E[R_\mu]=\mu$ και $R_\mu^{*}\stackrel{d}{=}W$ αν και μόνο αν $\mu\in[-1,1)$.

iii) Για την τυχαία μεταβλητή Y με πυκνότητα

$$h(x)=\frac{3}{10}(|x|(2-|x|)I(1<|x|<2)+I(|x|\le1)),$$

υπάρχει μια τυχαία μεταβλητή X_μ τέτοια ώστε $E[X_\mu]=\mu$ και $X_\mu^{*}\stackrel{d}{=}Y$ για κάθε κορυφή $\mu\in[-1,1]$.

5.2. Εφαρμογή σε Φράγματα Διασποράς

Λήμμα 5.2

Έστω X μια απόλυτα συνεχής τυχαία μεταβλητή με μέση τιμή μ και διασπορά σ^2. Τότε, για κάθε απόλυτα συνεχή συνάρτηση g με παράγωγο g', έχουμε τα ακόλουθα φράγματα:

i)
$$E[(g(\mathrm{X})-g(\mu))^2] \le \sigma^2 E[(g'(\mathrm{X}^*))^2], \tag{5.5}$$

με ισότητα αν και μόνο αν $E[g^2(\mathrm{X})]=\infty$ ή

$$g(x)-g(\mu)=\begin{cases} a_1(x-\mu) & av \quad x\le\mu, \\ a_2(x-\mu) & av \quad x\ge\mu, \end{cases}$$

για κάποιες σταθερές α_1, α_2 και για όλα τα $x\in S(\mathrm{X}^*)$.

ii) Αν $E\left|g'(\mathrm{X}^*)\right|<\infty$,

$$Var[g(\mathrm{X})]\ge \sigma^2 E^2[g'(\mathrm{X}^*)], \tag{5.6}$$

με ισότητα αν και μόνο αν $P[g(\mathrm{X})=a\mathrm{X}+b]=1$ για κάποιες σταθερές α και b.

Απόδειξη

i) Έστω f μια πυκνότητα της X. Τότε έχουμε

$$\sigma^2 E[(g'(\mathrm{X}^*))^2] = \int_{-\infty}^{\mu}(g'(x))^2\int_{-\infty}^{x}(\mu-t)f(t)dtdx+\int_{\mu}^{\infty}(g'(x))^2\int_{x}^{\infty}(t-\mu)f(t)dtdx$$

$$=\int_{-\infty}^{\mu}f(t)(\mu-t)\int_{t}^{\mu}(g'(x))^2dxdt+\int_{\mu}^{\infty}f(t)(t-\mu)\int_{\mu}^{t}(g'(x))^2dxdt$$

$$\ge\int_{-\infty}^{\mu}f(t)(g(\mu)-g(t))^2dt+\int_{\mu}^{\infty}f(t)(g(t)-g(\mu))^2dt$$

$$=E[(g(\mathrm{X})-g(\mu))^2],$$

από Θεώρημα Tonelli και την ανισότητα Cauchy-Schwarz για ολοκληρώματα. Παρατηρούμε ότι αν $E[g^2(\mathrm{X})]=\infty$, η ισότητα ισχύει με ένα προφανή τρόπο ($\infty=\infty$).

Αλλιώς, η ισότητα ισχύει αν και μόνο αν υπάρχουν σταθερές α_1 και α_2 τέτοιες ώστε $g'(x)=\alpha_1+(\alpha_2-\alpha_1)I(x\ge\mu)$ για σχεδόν όλα τα $x\in S(\mathrm{X}^*)$, που ολοκληρώνει την απόδειξη.

ii) Έχουμε από (5.2)

$$\sigma^4 E^2[g'(X^*)] = Cov^2[X, g(X)] \le \sigma^2 Var[g(X)]$$

Από την ανισότητα Cauchy-Schwarz για τυχαίες μεταβλητές και η απόδειξη είναι πλήρης.

Πόρισμα 5.1

Για κάθε απόλυτα συνεχή συνάρτηση g

$$Var[g(X)] \le \sigma^2 E[(g'(X^*))^2], \tag{5.7}$$

με ισότητα αν και μόνο αν $E[g^2(X)] = \infty$ ή g είναι γραμμική στο $S(X^*)$.

Σημειώνουμε ότι η (5.6) συνεχίζει να ισχύει για κάθε μη φθίνουσα (μη αύξουσα) απόλυτα συνεχή συνάρτηση g, ακόμα και στην περίπτωση που $E[g'(X^*)] = \pm\infty$ (σε αυτήν την περίπτωση $E[g^2(X)] = \infty$). Επίσης, επειδή $P[X \in S(X^*)] = 1$, η ισότητα στην (5.7) συνεπάγεται την ισότητα στην (5.6)

Το αντίστροφο δεν ισχύει πάντα, όπως δείχνει το ακόλουθο παράδειγμα.

Παράδειγμα 5.3

Έστω X ομοιόμορφα κατανεμημένη στο $(-2,-1) \cup (1,2)$ και $g(x) = xI(|x|>1) + x^3 I(|x| \le 1)$.

Τότε $\mu = 0$, $\sigma^2 = \dfrac{7}{3}$ και X^* έχει πυκνότητα $f^*(x) = \dfrac{3}{28}(\min\{4-x^2, 3\})^+$. Προκύπτει ότι

$$Var[g(X)] = E[(g(X) - g(\mu))^2] = \sigma^2 \text{ και } \sigma^2 E^2[g'(X^*)] = Var[g(X)] < \sigma^2 E[(g'(X^*))^2] = \frac{53}{15}.$$

Αυτό δείχνει ότι η g είναι γραμμική σε σχέση με το επαγόμενο μέτρο της X ενώ δεν είναι γραμμική σε σχέση με το επαγόμενο μέτρο της X^*.

Αυτό το παράδειγμα βασίζεται στο γεγονός ότι το $S(X)$ δεν είναι διάστημα. Αν παρόλα αυτά το $S(X)$ είναι πεπερασμένο ή μη πεπερασμένο διάστημα, τα γνωστά άνω και κάτω φράγματα για τη διασπορά της $g(X)$ παίρνουν την παρακάτω μορφή:

$$\sigma^2 E^2[w(X) g'(X)] \le Var[g(X)] \le \sigma^2 E[w(X)(g'(X))^2],$$

για κάποια μη αρνητική συνάρτηση w ορισμένη στο S(X), όπου και οι δύο ανισότητες ισχύουν αν και μόνο αν η g είναι γραμμική στο $S(X)$, δοθέντος ότι $E[w(X)(g'(X))^2] < \infty$.

Το ακόλουθο αποτέλεσμα δείχνει την ισοδυναμία μεταξύ των φραγμάτων διασποράς και της ταυτότητας συνδιακύμανσης.

Θεώρημα 5.2

Υποθέτουμε ότι για την απόλυτα συνεχή τυχαία μεταβλητή X με πεπερασμένη διασπορά σ^2, η τυχαία μεταβλητή Y ικανοποιεί ένα από τα ακόλουθα:

i) Για κάθε απόλυτα συνεχή συνάρτηση g με παράγωγο g',

$$Var[g(\mathrm{X})] \le \sigma^2 E[(g'(\mathrm{Y}))^2].$$

ii) Για κάθε απόλυτα συνεχή συνάρτηση g με παράγωγο g', έτσι ώστε $E|g'(\mathrm{Y})| < \infty$

$$Var[g(\mathrm{X})] \ge \sigma^2 E^2[g'(\mathrm{Y})].$$

Τότε $\mathrm{Y} \overset{d}{=} \mathrm{X}^*$.

Απόδειξη

Υποθέτουμε ότι το i) ισχύει. Έστω h' μια μετρική φραγμένη συνάρτηση και θεωρούμε την απόλυτα συνεχή συνάρτηση $g(x) = x + \lambda h(x)$, όπου λ είναι μια αυθαίρετη σταθερά και h ένα αόριστο ολοκλήρωμα της h'. Προκύπτει ότι g είναι απόλυτα συνεχής με φραγμένη παράγωγο $g' = 1 + \lambda h'$. Τότε $Var[g(\mathrm{X})] < \infty$, $Var[h(\mathrm{X})] < \infty$ και

$$Var[g(\mathrm{X})] = \sigma^2 + \lambda^2 Var[h(\mathrm{X})] + 2\lambda Cov[\mathrm{X}, h(\mathrm{X})]$$

$$\le \sigma^2 \left(1 + \lambda^2 E[(h'(\mathrm{Y}))^2] + 2\lambda E[h'(\mathrm{Y})]\right).$$

Επομένως, η τετραγωνική μορφή $\theta\lambda^2 + 2\delta\lambda$, όπου $\theta = \sigma^2 E[(h'(\mathrm{Y}))^2] - Var[h(\mathrm{X})] \ge 0$ και $\delta = \sigma^2 E[h'(\mathrm{Y})] - Cov[\mathrm{X}, h(\mathrm{X})]$, είναι μη αρνητική για κάθε λ και τότε $\delta = 0$.

Συμπεραίνουμε ότι για κάθε φραγμένη συνάρτηση H $E[\mathrm{H}(\mathrm{Y})] = E[\mathrm{H}(\mathrm{X}^*)]$ και το αποτέλεσμα προκύπτει.

Όμοια για το ii).

Θεώρημα 5.3

Έστω Y μια αυθαίρετη τυχαία μεταβλητή. Τότε η Y έχει μια μονοκόρυφη απόλυτα συνεχή πυκνότητα h τέτοια ώστε η συνάρτηση $\dfrac{h'(x)}{m-x}$ να είναι ολοκληρώσιμη για κάποια κορυφή m της h αν και μόνο αν υπάρχει κάποια απόλυτα συνεχής τυχαία μεταβλητή X με πεπερασμένη ροπή δεύτερης τάξης, έτσι ώστε να ισχύει κάποιο από τα ακόλουθα:

i) Για κάθε απόλυτα συνεχή συνάρτηση g με παράγωγο g' τέτοια ώστε $E|g'(\mathrm{Y})| < \infty$,

$$Cov[\mathrm{X}, g(\mathrm{X})] = Var[\mathrm{X}]E[g'(\mathrm{Y})].$$

ii) Για κάθε απόλυτα συνεχή συνάρτηση g με παράγωγο g',

$$Var[g(\mathrm{X})] \le Var[\mathrm{X}]E[(g'(\mathrm{Y}))^2].$$

iii) Για κάθε απόλυτα συνεχή συνάρτηση g με παράγωγο g' τέτοια ώστε $E|g'(\mathrm{Y})| < \infty$,

$$Var[g(\mathrm{X})] \ge Var[\mathrm{X}]E^2[g'(\mathrm{Y})].$$

iv) $\mathrm{X}^* \overset{d}{=} \mathrm{Y}.$

Επίσης, X είναι μοναδική αν και μόνο αν η κορυφή της h είναι μοναδική.

Στα παραπάνω θεωρήματα η σταθερά είναι ίση με $Var[\mathrm{X}]$. Όμως υπάρχει περίπτωση η σταθερά να μην είναι ίση με $Var[\mathrm{X}]$. Τότε η Y δεν είναι αναγκαστικά η X^*.
Αυτό δείχνει το ακόλουθο παράδειγμα.

Παράδειγμα 5.4

Έστω X ομοιόμορφα κατανεμημένη στο (0,1) και η Y έχει πυκνότητα $h(x) = \dfrac{3}{4}(1-x^2)I(|x|<1)$. Προκύπτει ότι για κάθε απόλυτα συνεχή συνάρτηση g,

$$Var[g(\mathrm{X})] \le \frac{1}{12}E[(g'(\mathrm{X}^*))^2] = \frac{1}{2}\int_0^1 x(1-x)(g'(x))^2 dx \le cE[(g'(\mathrm{Y}))^2],$$

για κάποια σταθερά $c \le \dfrac{1}{3}$, επειδή $\dfrac{1}{2}x(1-x)I(0<x<1) \le \dfrac{1}{3}h(x)$ για $x \in (-1,1)$.

Από την άλλη μεριά, μια εφαρμογή στη συνάρτηση $g(x)=x^2$ δείχνει ότι $c \geq \frac{1}{9} > \frac{1}{12} = Var[\mathrm{X}]$.

ΚΕΦΑΛΑΙΟ 6

Ένα Γενικό Φράγμα για την Απόσταση Ολικής Κύμανσης

6.1. Ένα Ενιαίο Φράγμα για την Απόσταση Ολικής Κύμανσης μεταξύ δύο Κατανομών

Θα παρουσιάσουμε ένα γενικό φράγμα για την απόσταση ολικής κύμανσης

$$d_{TV}(\mathrm{X},\mathrm{Y}) = \sup_{\mathrm{A}}\left\{\left|P\{\mathrm{X}\in\mathrm{A}\} - P\{\mathrm{Y}\in\mathrm{A}\}\right|,\ \mathrm{A}\ \ Borel\right\} = \frac{1}{2}\int_{-\infty}^{+\infty}\left|f(x)-h(x)\right|dx$$

μεταξύ δύο τυχαίων μεταβλητών X, Y με πυκνότητες $f, h,$ μέσες τιμές μ, m και διασπορές $\sigma^2,\ s^2$ αντίστοιχα.

Θεωρούμε ότι X είναι μια τυχαία μεταβλητή και Y η οριακή τυχαία μεταβλητή.

Τα αποτελέσματα εφαρμόζονται όταν $S(\mathrm{Y})$ είναι ανοιχτό, πεπερασμένο ή μη πεπερασμένο διάστημα και $S(\mathrm{X}^*) \subset S(\mathrm{Y}),$ δηλαδή όταν το επαγόμενο μέτρο της X είναι απόλυτα συνεχές ως προς το επαγόμενο μέτρο της Y.

Τα αποτελέσματα γενικεύονται για μια τυχαία μεταβλητή X που δεν έχει απαραίτητα στήριγμα διάστημα.

Υποθέτουμε ότι $S(\mathrm{Y}) = (a,b),$ όπου $-\infty \le \alpha < b \le \infty.$

Έστω A Borel σύνολο και θεωρούμε την απόλυτα συνεχής συνάρτηση

$$\psi_{\mathrm{A}}(x) = \frac{1}{h^*(x)}\int_a^x \left(I(t\in\mathrm{A}) - P\{\mathrm{Y}\in\mathrm{A}\}\right)h(t)dt, \quad \alpha < x < b, \tag{6.1}$$

όπου h^* η πυκνότητα της Y^*.

Λήμμα 6.1

Οι συναρτήσεις $\psi_{\mathrm{A}}(x)$ και $\dfrac{h^*(x)}{h(x)}\psi'_{\mathrm{A}}(x)$ (η δεύτερη ορίζεται για σχεδόν όλα τα $x\in(a,b)$) είναι απόλυτα φραγμένες από κάποιες πεπερασμένες σταθερές c και $c',$ αντίστοιχα που δεν εξαρτώνται από το $x\in(a,b)$ και το A. Μια δυνατή επιλογή είναι $c'=2$ και $c = s^2 c_{\mathrm{Y}},$

όπου $c_{\mathrm{Y}} = \dfrac{2\max\{P\{\mathrm{Y}\le m\}, P\{\mathrm{Y}\ge m\}\}}{E|\mathrm{Y}-m|} < \dfrac{2}{E|\mathrm{Y}-m|}.$

Απόδειξη

Έστω $\alpha < x < x+y \le m.$ Προκύπτει ότι

$$\int_\alpha^x h(t)dt\int_x^{x+y}(m-t)h(t)dt \le (m-x)\int_a^x h(t)dt\int_x^{x+y}h(t)dt \le \int_a^x (m-t)h(t)dt\int_x^{x+y}h(t)dt.$$

Προσθέτοντας και στα δύο μέλη την ποσότητα $\int_a^x h(t)dt \int_a^x (m-t)h(t)dt$ συμπεραίνουμε ότι

$$s^2 h^*(x)H(x+y) \geq s^2 h^*(x+y)H(x),$$

όπου H η συνάρτηση κατανομής της Y.

Επειδή $x > \alpha,\ h^*(x) > 0$ και $H(x) > 0,$ προκύπτει ότι η συνάρτηση $\dfrac{H(x)}{h^*(x)}$ είναι μη φθίνουσα

για $x \in (a,m]$. Ομοίως, $\dfrac{(1-H(x))}{h^*(x)}$ είναι μη αύξουσα για $x \in [m,b)$.

Επειδή $|\psi_{\mathrm{A}}(x)| \leq \dfrac{1}{h^*(x)} \min\{H(x), 1-H(x)\}$ για κάθε $x \in (a,b)$ έχουμε ότι

$|\psi_{\mathrm{A}}(x)| \leq \dfrac{1}{h^*(m)} \max\{H(m), 1-H(m)\} = \dfrac{2s^2 \max\{H(m), 1-H(m)\}}{E|\mathrm{Y}-m|}$ από Λήμμα 5.1 i).

Παραγωγίζοντας την (6.1) έχουμε ότι για σχεδόν όλα τα $x \in (a,b)$,

$$\frac{h^*(x)}{h(x)}\psi'_{\mathrm{A}}(x) = \frac{x-m}{s^2}\psi_{\mathrm{A}}(x) + (I(x \in \mathrm{A}) - P\{\mathrm{Y} \in \mathrm{A}\}). \tag{6.2}$$

Τότε $\left|\dfrac{h^*(x)}{h(x)}\psi'_{\mathrm{A}}(x)\right| \leq \dfrac{1}{s^2}|(x-m)\psi_{\mathrm{A}}(x)| + 1 \leq \dfrac{|x-m|\min\{H(x), 1-H(x)\}}{s^2 h^*(x)} + 1.$

Επομένως, για σχεδόν όλα τα $x \in (a,m]$,

$$\left|\frac{h^*(x)}{h(x)}\psi'_{\mathrm{A}}(x)\right| \leq \frac{(m-x)\int_a^x h(t)dt}{\int_a^x (m-t)h(t)dt} + 1 \leq 2. \tag{6.3}$$

Ομοίως, $\left|\dfrac{h^*(x)}{h(x)}\psi'_{\mathrm{A}}(x)\right| \leq 2$ για σχεδόν όλα τα $x \in [m,b)$ και η απόδειξη ολοκληρώνεται.

Το ακόλουθο αποτέλεσμα είναι τώρα μια άμεση συνέπεια του παραπάνω Λήμματος.

Θεώρημα 6.1

Για X,Y όπως παραπάνω έχουμε

$$d_{TV}(\mathrm{X},\mathrm{Y}) \leq 2\int_a^b \left| f(x) - \frac{\sigma^2 h(x)}{s^2 h^*(x)} f^*(x)\right| dx + c_{\mathrm{Y}}|\mu - m|, \tag{6.4}$$

όπου f^* η πυκνότητα της X^*.

Απόδειξη

Επειδή $f(x)=f(x)I(x\in S(\mathrm{X}^*))$ σχεδόν παντού, υποθέτουμε ότι $S(\mathrm{X})\subset S(\mathrm{X}^*)$.

Παίρνοντας μέση τιμή στην (6.2) και από το Λήμμα 5.1 ii) έχουμε:

$$P\{\mathrm{X}\in \mathrm{A}\}-P\{\mathrm{Y}\in \mathrm{A}\}=E\left[\frac{h^*(\mathrm{X})}{h(\mathrm{X})}\psi'_{\mathrm{A}}(\mathrm{X})\right]-\frac{1}{s^2}Cov[\mathrm{X},\psi_{\mathrm{A}}(\mathrm{X})]-\frac{\mu-m}{s^2}E[\psi_{\mathrm{A}}(\mathrm{X})]$$

$$=E\left[\frac{h^*(\mathrm{X})}{h(\mathrm{X})}\psi'_{\mathrm{A}}(\mathrm{X})\right]-\frac{\sigma^2}{s^2}E\left[\psi'_{\mathrm{A}}(\mathrm{X}^*)\right]-\frac{\mu-m}{s^2}E[\psi_{\mathrm{A}}(\mathrm{X})]$$

$$=\int_{S(\mathrm{X}^*)}\left[f(x)\frac{h^*(x)}{h(x)}-\frac{\sigma^2}{s^2}f^*(x)\right]\psi'_{\mathrm{A}}(x)dx-\frac{\mu-m}{s^2}E\left[\psi_{\mathrm{A}}(\mathrm{X})\right].$$

Επομένως, επειδή $S(\mathrm{X})\subset S(\mathrm{X}^*)\subset S(\mathrm{Y})=(a,b)$ έχουμε από Λήμμα 6.1 ότι

$$\left|P\{\mathrm{X}\in \mathrm{A}\}-P\{\mathrm{Y}\in \mathrm{A}\}\right|\le \operatorname*{esssup}_{x\in(a,b)}\left|\frac{h^*(x)}{h(x)}\psi'_{\mathrm{A}}(x)\right|\int_a^b\left|f(x)-\frac{\sigma^2h(x)}{s^2h^*(x)}f^*(x)\right|dx+\frac{|\mu-m|}{s^2}\sup_{x\in(a,b)}|\psi_{\mathrm{A}}(x)|$$

$$\le 2\int_a^b\left|f(x)-\frac{\sigma^2h(x)}{s^2h^*(x)}f^*(x)\right|dx+c_{\mathrm{Y}}|\mu-m|,$$

που είναι ανεξάρτητο του A.

Συγκεκριμένα, για την κανονική τυχαία μεταβλητή $\mathrm{Z}_{m,\sigma^2}=\sigma\mathrm{Z}+m,$ όπου Z είναι η τυποποιημένη κανονική, ισχύει το ακόλουθο αποτέλεσμα.

Πόρισμα 6.1

Για κάθε απόλυτα συνεχής τυχαία μεταβλητή X με μέση τιμή μ και διασπορά σ^2,

$$d_{TV}(\mathrm{X},Z_{\mu,\sigma^2})\le 3d_{TV}(\mathrm{X},\mathrm{X}^*)+\frac{\sqrt{\pi}}{\sigma\sqrt{2}}|\mu-m|. \tag{6.5}$$

Απόδειξη

Από (6.3) έχουμε ότι για όλα τα x και A, $|(x-m)\psi_{\mathrm{A}}(x)|\le\sigma^2$. Από την άλλη μεριά, η κανονική πυκνότητα $h(x)=\varphi(x;\mu,\sigma^2)=\frac{1}{\sigma}\varphi\left(\frac{x-m}{\sigma}\right),$ όπου φ είναι η τυποποιημένη κανονική πυκνότητα, ικανοποιεί $h^*\equiv h$.

$$P\{\mathrm{X}\in \mathrm{A}\}-P\left\{Z_{m,\sigma^2}\in \mathrm{A}\right\}=\int_{S(\mathrm{X}^*)}(f(x)-f^*(x))\psi'_{\mathrm{A}}(x)dx-\frac{\mu-m}{\sigma^2}E\left[\psi_{\mathrm{A}}(\mathrm{X})\right]$$

$$=\int_{S(\mathrm{X}^*)}\frac{x-m}{\sigma^2}\psi_{\mathrm{A}}(x)(f(x)-f^*(x))dx+\int_{S(\mathrm{X}^*)}I(x\in \mathrm{A})(f(x)-f^*(x))dx-\frac{\mu-m}{\sigma^2}E\left[\psi_{\mathrm{A}}(\mathrm{X})\right].$$

Παρατηρούμε ότι για κάθε Borel σύνολο Α

$$\int_{S(\mathrm{X}^*)} I(x \in \mathrm{A})(f(x) - f^*(x))dx \le d_{TV}(\mathrm{X}, \mathrm{X}^*)$$

και

$$\int_{S(\mathrm{X}^*)} \frac{x-m}{\sigma^2}\psi_{\mathrm{A}}(x)(f(x) - f^*(x))dx \le 2d_{TV}(\mathrm{X}, \mathrm{X}^*).$$

Το αποτέλεσμα προκύπτει από το γεγονός ότι $E\left|Z_{m,\sigma^2} - m\right| = \sigma\sqrt{\frac{2}{\pi}}$ και την παρατήρηση ότι για κάθε δύο τυχαίες μεταβλητές X_1 και X_2

$$d_{TV}(\mathrm{X}_1, \mathrm{X}_2) = \sup_{\mathrm{A}}\{P(\mathrm{X}_1 \in \mathrm{A}) - P(\mathrm{X}_2 \in \mathrm{A})\}.$$

Μια ικανή και αναγκαία συνθήκη για να ακολουθεί μια απόλυτα συνεχής τυχαία μεταβλητή Χ με πεπερασμένη δεύτερη ροπή κανονική κατανομή είναι $\mathrm{X} \overset{d}{=} \mathrm{X}^*$.

Θεώρημα 6.2

Έστω X_n μια ακολουθία απόλυτα συνεχών τυχαίων μεταβλητών με μέσους μ_n και διασπορές $\sigma_n^{\ 2}$ έτσι ώστε $\mu_n \to \mu$ και $\sigma_n^{\ 2} \to \sigma^2$ καθώς $n \to \infty$, όπου $0 < \sigma^2 < \infty$.

Τότε, καθώς $n \to \infty$

i) $d_{TV}(\mathrm{X}_n, \mathrm{Z}_{\mu,\sigma^2}) \to 0$ αν και μόνο αν $d_{TV}(\mathrm{X}_n, \mathrm{X}_n^{\ *}) \to 0$.

ii) Αν $d_{TV}(\mathrm{X}_n, \mathrm{Z}_{\mu,\sigma^2}) \to 0$ τότε $d_{TV}(\mathrm{X}_n^{\ *}, \mathrm{Z}_{\mu,\sigma^2}) \to 0$.

Απόδειξη

Υποθέτουμε πρώτα ότι $d_{TV}(\mathrm{X}_n, \mathrm{X}_n^{\ *}) \to 0$.

$$d_{TV}(\mathrm{X}_n, \mathrm{Z}_{\mu,\sigma^2}) \le d_{TV}(\mathrm{X}_n, \mathrm{Z}_{\mu,\sigma_n^{\ 2}}) + d_{TV}(\mathrm{Z}_{\mu,\sigma_n^{\ 2}}, \mathrm{Z}_{\mu,\sigma^2})$$

$$\le 3d_{TV}(\mathrm{X}_n, \mathrm{X}_n^{\ *}) + \frac{\sqrt{\pi}}{\sigma_n\sqrt{2}}\left|\mu_n - \mu\right| + 2\left|1 - \frac{\sigma_n^{\ 2}}{\sigma^2}\right| \to 0.$$

Υποθέτουμε τώρα ότι $d_{TV}(\mathrm{X}_n, \mathrm{Z}_{\mu,\sigma^2}) \to 0$.

Προκύπτει ότι κάθε υπακολουθία $\{m\} \subset \{n\}$ περιέχει μια περαιτέρω υπακολουθία $\{k\} \subset \{m\}$ τέτοια ώστε για σχεδόν όλα τα x, $f_k(x) \to \varphi(x;\mu,\sigma^2)$ καθώς $k \to \infty$, όπου f_k είναι η πυκνότητα της X_k. Επομένως, για σχεδόν όλα τα x

$$\frac{(\mu_k - x)^2}{\sigma_k^{\ 2}} f_k(x) \to \frac{(\mu - x)^2}{\sigma^2}\varphi(x;\mu,\sigma^2) \text{ καθώς } k \to \infty$$

και από Λήμμα του *Scheffe'*

$$\int_{-\infty}^{+\infty}\left|\frac{(\mu_k-x)^2}{\sigma_k^{\,2}}f_k(x)-\frac{(\mu-x)^2}{\sigma^2}\varphi(x;\mu,\sigma^2)\right|dx\to 0. \tag{6.6}$$

Χρησιμοποιώντας το Θεώρημα Tonelli έχουμε

$$2d_{TV}(\mathrm{X}_k^{\,*},\mathrm{Z}_{\mu_k,\sigma_k^{\,2}})$$

$$=\int_{-\infty}^{\mu_k}\left|\int_{-\infty}^{x}\frac{\mu_k-t}{\sigma_k^{\,2}}(f_k(t)-\varphi(t;\mu_k,\sigma_k^{\,2}))dt\right|dx+\int_{\mu_k}^{\infty}\left|\int_{x}^{\infty}\frac{t-\mu_k}{\sigma_k^{\,2}}(f_k(t)-\varphi(t;\mu_k,\sigma_k^{\,2}))dt\right|dx$$

$$\le\int_{-\infty}^{\mu_k}\int_{-\infty}^{x}\frac{\mu_k-t}{\sigma_k^{\,2}}\left|f_k(t)-\varphi(t;\mu_k,\sigma_k^{\,2})\right|dtdx+\int_{\mu_k}^{\infty}\int_{x}^{\infty}\frac{t-\mu_k}{\sigma_k^{\,2}}\left|f_k(t)-\varphi(t;\mu_k,\sigma_k^{\,2})\right|dtdx$$

$$=\int_{-\infty}^{+\infty}\frac{(\mu_k-t)^2}{\sigma_k^{\,2}}\left|f_k(t)-\varphi(t;\mu_k,\sigma_k^{\,2})\right|dt$$

$$\le\int_{-\infty}^{+\infty}\left|\frac{(\mu_k-t)^2}{\sigma_k^{\,2}}f_k(t)-\frac{(\mu-t)^2}{\sigma^2}\varphi(t;\mu,\sigma^2)\right|dt+\int_{-\infty}^{+\infty}\left|\frac{(\mu-t)^2}{\sigma^2}\varphi(t;\mu,\sigma^2)-\frac{(\mu_k-t)^2}{\sigma_k^{\,2}}\varphi(t;\mu_k,\sigma_k^{\,2})\right|dt\to 0$$

καθώς $k\to\infty$ από (6.6).

Επομένως, $d_{TV}(\mathrm{X}_k,\mathrm{X}_k^{\,*})\le d_{TV}(\mathrm{X}_k,\mathrm{Z}_{\mu,\sigma^2})+d_{TV}(\mathrm{Z}_{\mu,\sigma^2},\mathrm{Z}_{\mu_k,\sigma_k^{\,2}})+d_{TV}(\mathrm{X}_k^{\,*},\mathrm{Z}_{\mu_k,\sigma_k^{\,2}})\to 0$

καθώς $k\to\infty$. Για το ii) έχουμε, λόγω του i), ότι

$$d_{TV}(\mathrm{X}_n^{\,*},\mathrm{Z}_{\mu,\sigma^2})\le d_{TV}(\mathrm{X}_n^{\,*},\mathrm{X}_n)+d_{TV}(\mathrm{X}_n,\mathrm{Z}_{\mu,\sigma^2})\to 0$$

που ολοκληρώνει την απόδειξη.

Παρατήρηση 6.1: Αν X έχει στήριγμα διάστημα,

$$d_{TV}(\mathrm{X},\mathrm{Y})\le 2E\left|1-\frac{\sigma^2 w(\mathrm{X})}{s^2 w_{\mathrm{Y}}(\mathrm{X})}\right|+c_{\mathrm{Y}}|\mu-m|,$$

όπου $w=\dfrac{f^*}{f}$ και $w_{\mathrm{Y}}=\dfrac{h^*}{h}$ είναι οι w-συναρτήσεις των X και Y αντίστοιχα.

Παράδειγμα 6.1

Έστω $\mathrm{X}_1,\ldots,\mathrm{X}_n$ ανεξάρτητες, ισόνομες τυχαίες μεταβλητές που ακολουθούν την ομοιόμορφη κατανομή στο (0,1) και $\mathrm{X}=n\mathrm{X}_{1:n}$, όπου $\mathrm{X}_{1:n}=\min\{\mathrm{X}_1,\ldots,\mathrm{X}_n\}$.

Τότε $\mu=\dfrac{n}{n+1}$, $\sigma^2=\dfrac{n^3}{(n+1)^2(n+2)}$, $w(x)=\dfrac{(n+1)(n+2)}{n^2}x\left(1-\dfrac{x}{n}\right)$, $0<x<n.$

Αν η Y ακολουθεί την εκθετική κατανομή με παράμετρο 1, τότε

$w_Y(x) = x,\ m = 1,\ s^2 = 1$ και από την Παρατήρηση 6.1 έχουμε

$$d_{TV}(X,Y) \le \frac{1}{n+1}\left(2\frac{2n+1}{n+1} + c_Y\right).$$

Η τάξη σύγκλισης είναι τουλάχιστον Ο(1/*n*).

Παράδειγμα 6.2

Αν η X ακολουθεί την *t*-κατανομή με $n > 3$ βαθμούς ελευθερίας, τότε $\mu = 0,\ \sigma^2 = \dfrac{n}{n-2}$

και $w(x) = (n-2)/(n-1)(1 + x^2/n)$. Αν η Y ακολουθεί την τυποποιημένη κανονική κατανομή τότε, από την Παρατήρηση 6.1, έχουμε

$$d_{TV}(X,Y) \le \frac{4}{n-2} \to 0$$

και η τάξη σύγκλισης είναι τουλάχιστον $O(1/n)$.

Παράδειγμα 6.3

Αν X=*n*Z, όπου η Z ακολουθεί την *F*- κατανομή με *n* και *m* βαθμούς ελευθερίας και η Y ακολουθεί την X^2-κατανομή με *n* βαθμούς ελευθερίας, τότε από την Παρατήρηση 6.1 έχουμε

$$d_{TV}(X,Y) \le \frac{1}{m-2}\left[2nc_Y + 2\left(2 + \frac{mn}{m-2}\right)\right].$$

Η τάξη σύγκλισης είναι τουλάχιστον Ο(1/*m*).

Παράδειγμα 6.4

Υποθέτουμε ότι X_n ακολουθεί την Βήτα-κατανομή $B(\alpha_n, b_n)$ με παραμέτρους $\alpha_n > 0$, $b_n > 0$. Αν $X = m_n X_n$ και η ακολουθία $\{m_n > 0\}$ επιλέγεται έτσι ώστε

$$m_n \to \infty,\quad E(X) \to \mu > 0 \ \text{ και } \ Var(X) \to \sigma^2 > 0,$$

και η Y ακολουθεί την Γάμμα κατανομή $G(\mu^2/\sigma^2, \sigma^2/\mu)$, τότε από την Παρατήρηση 6.1 έχουμε

$$d_{TV}(X,Y) \le 2\left|1 - \frac{\mu}{\sigma^2}\frac{m_n}{\alpha_n + b_n}\right| + 2\frac{\mu}{\sigma^2}\frac{m_n \alpha_n}{(a_n + b_n)^2} + c_Y\left|\frac{m_n a_n}{a_n + b_n} - \mu\right|.$$

6.2. Τοπικό Οριακό Θεώρημα

Έστω Χ, $X_1, X_2, \ldots$ μια ακολουθία ανεξάρτητων, ισόνομων τυχαίων μεταβλητών με μέση τιμή 0, διασπορά 1 και μια κοινή πυκνότητα f και θεωρούμε τα μερικά αθροίσματα $S_n = X_1 + X_2 + \ldots + X_n$ με πυκνότητες f_n.

Θεώρημα 6.3

Αν Ζ είναι μια τυποποιημένη κανονική τυχαία μεταβλητή, τότε

$$d_{TV}\left(\frac{S_n}{\sqrt{n}}, Z\right) \to 0 \text{ καθώς } n \to \infty.$$

Για την απόδειξη του Θεωρήματος, χρειαζόμαστε τα παρακάτω Λήμματα.

Λήμμα 6.2

Έχουμε $S_n^* \overset{d}{=} S_{n-1}^* + X_n$, όπου S_{n-1}^* ανεξάρτητη του X_n.

Απόδειξη

Από Λήμμα 5.1 v) έχουμε $(X_1 + X_2)^* \overset{d}{=} X_1^* + X_2$, όπου αποδεικνύει το Λήμμα για n=2. Με επαγωγή ως προς n

$$S_n^* \overset{d}{=} X_1^* + (X_2 + \ldots + X_n).$$

Επομένως, $\quad S_{n-1}^* + X_n \overset{d}{=} X_1^* + (X_2 + \ldots + X_{n-1}) + X_n \overset{d}{=} S_n^*.$

Λήμμα 6.3

Η ακολουθία $d_{TV}(S_n, S_n^*)$ είναι φθίνουσα και συγκλίνει σε μια μη αρνητική σταθερά d.

Απόδειξη

Από το προηγούμενο Λήμμα έχουμε ότι

$$d_{TV}(S_n, S_n^*) = d_{TV}(S_{n-1} + X_n, S_{n-1}^* + X_n) \le d_{TV}(S_{n-1}, S_{n-1}^*).$$

Λήμμα 6.4

Αν $S(X) = \{x : f(x) > 0\}$ είναι το στήριγμα της Χ, τότε

$$d \le 1 - P\{X^* \in S(X)\}.$$

Απόδειξη

Έστω G_n, n=1,2,… μια ακολουθία μετρήσιμων συναρτήσεων έτσι ώστε $0 \le G_n(x) \le 1$ για κάθε n και x. Χρησιμοποιώντας την $S_n^* \overset{d}{=} \mathrm{X}_1^* + (\mathrm{X}_2 + \ldots + \mathrm{X}_n)$ έχουμε

$$
\begin{aligned}
E[G_n(S_n^*)] &= E[G_n(\mathrm{X}_1^* + (\mathrm{X}_2 + \ldots + \mathrm{X}_n))] \\
&\ge E[G_n(\mathrm{X}_1^* + (\mathrm{X}_2 + \ldots + \mathrm{X}_n)) I(\mathrm{X}_1^* \in S(\mathrm{X}))] \\
&= E\left[\frac{f^*(\mathrm{X}_1)}{f(\mathrm{X}_1)} G_n(S_n)\right] \\
&= \frac{1}{n}\sum_{j=1}^{n} E\left[\frac{f^*(\mathrm{X}_j)}{f(\mathrm{X}_j)} G_n(S_n)\right],
\end{aligned}
$$

όπου f^* είναι η κοινή πυκνότητα των X_j^*.

Επομένως,

$$
\begin{aligned}
& E\left[G_n(S_n)\right] - E\left[G_n(S_n^*)\right] \\
&\le E\left[G_n(S_n)\right] - \frac{1}{n}\sum_{j=1}^{n} E\left[\frac{f^*(\mathrm{X}_j)}{f(\mathrm{X}_j)} G_n(S_n)\right] \\
&= E\left[G_n(S_n) - \frac{1}{n}\sum_{j=1}^{n} \frac{f^*(\mathrm{X}_j)}{f(\mathrm{X}_j)} G_n(S_n)\right] \\
&= E\left[\left(1 - \frac{1}{n}\sum_{j=1}^{n} \frac{f^*(\mathrm{X}_j)}{f(\mathrm{X}_j)}\right) G_n(S_n)\right] \\
&= E\left[\left(1 - \frac{1}{n}\sum_{j=1}^{n} \frac{f^*(\mathrm{X}_j)}{f(\mathrm{X}_j)}\right) G_n(S_n)\right] + P\{\mathrm{X}^* \in S(\mathrm{X})\} E\left[G_n(S_n)\right] - P\{\mathrm{X}^* \in S(\mathrm{X})\} E\left[G_n(S_n)\right] \\
&= E\left[\left(P\{\mathrm{X}^* \in S(\mathrm{X})\} - \frac{1}{n}\sum_{j=1}^{n} \frac{f^*(\mathrm{X}_j)}{f(\mathrm{X}_j)}\right) G_n(S_n)\right] + \left(1 - P\{\mathrm{X}^* \in S(\mathrm{X})\}\right) E\left[G_n(S_n)\right] \\
&\le E\left|P\{\mathrm{X}^* \in S(\mathrm{X})\} - \frac{1}{n}\sum_{j=1}^{n} \frac{f^*(\mathrm{X}_j)}{f(\mathrm{X}_j)}\right| + 1 - P\{\mathrm{X}^* \in S(\mathrm{X})\}.
\end{aligned}
$$

Επειδή οι μη αρνητικές, ανεξάρτητες και ισόνομες τυχαίες μεταβλητές $\mathrm{Y}_j = \dfrac{f^*(\mathrm{X}_j)}{f(\mathrm{X}_j)}$ έχουν μέση τιμή $E\left[\mathrm{Y}_j\right] = P\{\mathrm{X}^* \in S(\mathrm{X})\}$, προκύπτει από τον ασθενή νόμο των μεγάλων αριθμών ότι

$$\frac{1}{n}\sum_{j=1}^{n}\mathrm{Y}_j \xrightarrow{p} P\{\mathrm{X}^* \in S(\mathrm{X})\} = E\left[\frac{1}{n}\sum_{j=1}^{n}\mathrm{Y}_j\right], \quad \text{καθώς} \quad n \to \infty.$$

Επομένως, συμπεραίνουμε ότι

$$E\left|P\{\mathrm{X}^* \in S(\mathrm{X})\} - \frac{1}{n}\sum_{j=1}^{n}\mathrm{Y}_j\right| \to 0, \quad \text{καθώς} \quad n \to \infty.$$

Η απόδειξη ολοκληρώνεται διαλέγοντας $G_n(x) = I(x \in \mathrm{A}_n)$, όπου f_n^* είναι η πυκνότητα του S_n^* και $\mathrm{A}_n = \{x : f_n(x) > f_n^*(x)\}$.

Λήμμα 6.5

Έστω Y, W ανεξάρτητες και απόλυτα συνεχείς τυχαίες μεταβλητές με πυκνότητες h, g αντίστοιχα και έστω $q = h * g$ η πυκνότητα του $\mathrm{Y}+W$.

Τότε έχουμε τα ακόλουθα:

i) Το σύνολο $S(\mathrm{Y}+W) = \{x : q(x) > 0\}$ περιέχει κάποιο διάστημα.

ii) Αν $0 < P\{\mathrm{Y} \le 0\} < 1$ και $0 < P\{W \le 0\} < 1$, τότε υπάρχουν τέσσερις θετικοί αριθμοί $\alpha_1 < \alpha_2$ και $b_1 < b_2$ τέτοιοι ώστε $(-\alpha_2, -\alpha_1) \cup (b_1, b_2) \subset S(\mathrm{Y}+W)$.

iii) Αν $(\alpha, b) \subset S(\mathrm{Y}) = \{x : h(x) > 0\}$ και $(c, d) \subset S(W) = \{x : g(x) > 0\}$, τότε $(\alpha, b) + (c, d) = (a + c, b + d) \subset S(\mathrm{Y}+W)$.

Απόδειξη

Έστω $h_k(x) = \min\{k, h(x)\}$ για k=1,2,…. Προφανώς, h_k είναι φραγμένη και επομένως η συνάρτηση $h_k * g$ είναι συνεχής. Άρα το σύνολο $\mathrm{A}_k = \{x : (h_k * g)(x) > 0\}$ είναι ανοιχτό. Επειδή $h_k \uparrow h$ καθώς $k \to \infty$ και $h_k * g \le q$ για κάθε k προκύπτει ότι $\mathrm{A}_k \subset S(\mathrm{Y}+W)$. Από το Θεώρημα μονότονης σύγκλισης, $h_k * g \uparrow q$ καθώς $k \to \infty$, που δείχνει ότι για αρκετά μεγάλο k, $(h_k * g)(x) > 0$ αν $q(x) > 0$. Τότε το A_k είναι μη κενό για αρκετά μεγάλο k και περιέχει διάστημα, που αποδεικνύει το i). Για το ii), παρατηρούμε ότι $S(\mathrm{Y}+W)$ περιέχει και θετικούς και αρνητικούς αριθμούς και το αποτέλεσμα προκύπτει όπως το i). Το iii) είναι προφανές.

Ως άμεση συνέπεια του Λήμματος ισχύει το ακόλουθο αποτέλεσμα.

Πόρισμα 6.2

Υπάρχει ακέραιος αριθμός m και θετικός αριθμός α έτσι ώστε $(-n\alpha, n\alpha) \subset S(S_{nm})$ για n=1,2,… όπου $S(S_{nm})$ το στήριγμα του S_{nm}.

Απόδειξη

Από το παραπάνω Λήμμα ii) έχουμε

$$(-\alpha_2, -\alpha_1) \cup (b_1, b_2) \subset S(\mathrm{X}_1 + \mathrm{X}_2) = S(S_2).$$

Χρησιμοποιώντας Λήμμα iii) και το γεγονός ότι $S_{2n} = (\mathrm{X}_1 + \mathrm{X}_2) + \ldots + (\mathrm{X}_{2n-1} + \mathrm{X}_{2n})$ συμπεραίνουμε ότι

$$\bigcup_{k=0}^{n} (-k\alpha_2 + (n-k)b_1, -k\alpha_1 + (n-k)b_2) \subset S(S_{2n}).$$

Τώρα διαλέγουμε δύο ακέραιους αριθμούς $k < m'$ έτσι ώστε $\dfrac{\alpha_1}{b_2} + 1 < \dfrac{m'}{k} < \dfrac{\alpha_2}{b_1} + 1.$

Τότε υπάρχει ένας θετικός αριθμός α τέτοιος ώστε $(-\alpha, \alpha) \subset S(S_{2m'})$ και επομένως $(-n\alpha, n\alpha) \subset S(S_{nm})$ για κάθε $n \geq 1$, όπου $m = 2m'$.

Τέλος, θα χρησιμοποιήσουμε το ακόλουθο Λήμμα.

Λήμμα 6.6

Έστω $R, R_1, R_2, \ldots$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με μέση τιμή 0 και πεπερασμένη διασπορά $\tau^2 > 0$. Τότε, για κάθε $\alpha > 0$,

$$\frac{1}{n} E\left[\mathrm{T}_n^{\ 2} I(|\mathrm{T}_n| > n\alpha)\right] \to 0 \quad \text{καθώς} \quad n \to \infty,$$

όπου $\mathrm{T}_n = R_1 + \ldots + R_n$.

Για την απόδειξη βλέπε [17,p. 19] ή [5, p. 176].

Απόδειξη Θεωρήματος 6.3

Έστω α και m όπως στο Πόρισμα 6.2.

Από το παραπάνω Λήμμα προκύπτει ότι

$$
\begin{aligned}
P\{S_{nm}{}^{*} \in S(S_{nm})\} &= \int_{S(S_{nm})} f_{nm}{}^{*}(x)dx \\
&\geq \int_{-n\alpha}^{n\alpha} f_{nm}{}^{*}(x)dx \\
&= \frac{1}{nm}\left\{ n\alpha \int_{|x|>n\alpha} |x| f_{nm}(x)dx + \int_{-n\alpha}^{n\alpha} x^2 f_{nm}(x)dx \right\} \\
&= \frac{\alpha}{m}\int_{|x|>n\alpha} |x| f_{nm}(x)dx + \frac{1}{nm}\int_{-n\alpha}^{n\alpha} x^2 f_{nm}(x)dx \\
&= \frac{a}{m}E\left[|S_{nm}| I(|S_{nm}|>n\alpha)\right] + \frac{1}{nm}E\left[S_{nm}{}^{2} I(|S_{nm}| \leq n\alpha)\right] \\
&= 1 + \frac{\alpha}{m}E\left[|S_{nm}| I(|S_{nm}|>n\alpha)\right] - \frac{1}{nm}E\left[S_{nm}{}^{2} I(|S_{nm}| > n\alpha)\right] \\
&\geq 1 - \frac{1}{nm}E\left[S_{nm}{}^{2} I(|S_{nm}| > n\alpha)\right] \to 1 \qquad \kappa\alpha\theta\acute{\omega}\varsigma \qquad n \to \infty.
\end{aligned}
$$

Έστω R_j ανεξάρτητες και ισόνομες τυχαίες μεταβλητές τέτοιες ώστε

$$R_j \overset{d}{=} \frac{S_{mk}}{\sqrt{mk}}, \quad j=1,2,\ldots$$

Από Λήμμα 6.3 και 6.4 προκύπτει ότι

$$d = \lim_{n\to\infty} d_{TV}(S_{nmk}, S_{nmk}{}^{*}) = \lim_{n\to\infty} d_{TV}\left(\frac{S_{nmk}}{\sqrt{nmk}}, \frac{S_{nmk}{}^{*}}{\sqrt{nmk}}\right) \leq 1 - P\{R_1^{*} \in S(R_1)\} = 1 - P\{S_{mk}{}^{*} \in S(S_{mk})\} \to 0$$

καθώς $k \to \infty$, που δείχνει ότι $d = 0$. Το αποτέλεσμα προκύπτει από το Πόρισμα 6.1

$$d_{TV}\left(\frac{S_n}{\sqrt{n}}, Z\right) \leq 3d_{TV}\left(\frac{S_n}{\sqrt{n}}, \frac{S_n{}^{*}}{\sqrt{n}}\right) = 3d_{TV}(S_n, S_n{}^{*}) \to 0.$$

Β Ι Β Λ Ι Ο Γ Ρ Α Φ Ι Α

1) R. Arratia, L. Goldstein and L. Gordon, Poisson approximation and the Chen-Stein method, *Statist. Science*, **5** (1990), pp. 403-434.

2) G. Afendras and N. Papadatos, A factorial moment distance and an application to the matching problem, Theory of Probability and its Applications, **62** (2017), pp. 617-628.

3) A. D. Barbour and G. K. Eagleson, Poisson approximation for some statistics based on exchangeable trials, *Adv. Appl. Probability*, **15** (1983), pp. 585-600.

4) A. D. Barbour, L. Holst and S. Janson, *Poisson Approximation*, Oxford Stydies in Probability **2**, Clarendon Press, Oxford, 1992.

5) P. Billingsley, Convergence of Probability Measures, Wiley, New York, 1968.

6) T. Cacoullos, V. Papathanasiou and S. Utev, Variational inequalities with examples and an application to the central limit theorem, *The Annals of Probability*, **22** (1994), pp. 1607-1618.

7) T. Cacoullos, V. Papathanasiou and S. Utev, Another characterization of the normal law and a proof of the central limit theorem connected with it, Theory of Probability and its Applications, **37** (1992), pp. 581-588.

8) T. Cacoullos, N. Papadatos, V. Papathanasiou, Variance inequalities for covariance kernels and applications to central limit theorems, Theory of Probability and its Applications, **42** (1997), pp. 195-201.

9) T. Cacoullos, N. Papadatos, V. Papathanasiou, Three elementary proofs of the central limit theorem with applications to random sums, Stochastic Processes and Related Topics, (I. Karatzas, B.S. Rajput and M.S. Taqqu, Eds.), 1998, Birkhauser, Boston, pp. 15-23.

10) T. Cacoullos, N. Papadatos, V. Papathanasiou, An application of a density transform and the local limit theorem, Theory of Probability and its Applications, **46** (2001), pp. 803-810.

11) T. Cacoullos and V. Papathanasiou, Characterizations of distributions by variance bounds, Statist. Probab. Lett., 7 (1989), pp. 351-356.

12) T. Cacoullos and V. Papathanasiou, Lower variance bounds and a new proof of the central limit theorem, J. Multivariate Anal., 43 (1992), pp. 173-184.

13) Chen, L.H.Y. (1988), The central limit theorem and $Poincare'$ type inequalities, Ann. Probab. 16, pp. 300-304.

14) A. DasGupta, The matching problem with random decks and the Poisson approximation, Technical report, #99-01, Purdue Univ., West Lafayette, IN, 1999, 18 pp.

15) A. DasGupta, “The matching, birthday and the strong birthday problem: a contemporary review”, J. Statist. Plann. Inference, 130:1-2 (2005), 377–389.

16) P. Diaconis, “Application of the method of moments in probability and statistics”, Moments in mathematics (San Antonio, Tex., 1987), Proc. Sympos. Appl. Math., **37**, Amer. Math. Soc., Providence, RI, 1987, 125–142.

17) A.Gut, Stopped Random Walks: Limit Theorems and Applications, Springer-Verlag, New York, 1988.

18) M. Majsnerowska, A note on Poisson approximation by w-functions, *Applicationes Mathematicae*, **25** (1998), pp. 387-392.

19) P. R. de Montmort, Essay d’analyse sur les jeux de hazard, Photographic reprint of the 2nd ed., Chelsea Publishing, New York, 1980, xlii+416 pp.; 1st ed., 1708; 2nd ed., Jacques Quillau, Paris, 1713.

20) S. Niermann, “A generalization of the matching distribution”, Statist. Papers, 40:2 (1999), 233–238.

21) N. Papadatos and V. Papathanasiou, Distance in variation between two arbitrary distributions via the associated w-functions, *Theory Probab. Appl. (Russian)*, **40** (1995), pp. 685-694.

22) N. Papathatos, V. Papathanasiou, Total variation distance and generalized covariance kernels, Mathematical Methods of Statistics **7** (1998), pp. 230-244.

23) N. Papathatos, V. Papathanasiou, Unified variance bounds and a Stein type identity, (Ch.A. Charalambides,M.V. Koutras and N. Balakrishnan Eds.) , 2001, Chapman & Hall/CRC, New York, pp. 87-100.

24) Y. H. Wang, “A compound Poisson convergence theorem”, Ann. Probab., **19**:1 (1991), 452–455.

25) V. M. Zolotarev, Probability metrics, *Theory Prob. Appl. (Russian)*, **28** (1983), pp. 278-302.

26) Χ. Χαραλαμπίδης, *Θεωρία Πιθανοτήτων και Εφαρμογές*, Εκδόσεις Συμμετρία, Αθήνα, 1993.

Ε Υ Ρ Ε Τ Η Ρ Ι Ο Ο Ρ Ω Ν