ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

Σχολή Εφαρμοσμένων Μαθηματικών και Φυσικών Επιστημών

Ασυμπτωτική συμπεριφορά των ροπών μιας κατηγορίας τυχαίων μεταβλητών

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

ΔΟΥΜΑΣ Β. ΑΡΙΣΤΕΙΔΗΣ

ΕΠΙΒΛΕΠΩΝ:

ΒΑΣΙΛΕΙΟΣ - ΓΕΩΡΓΙΟΣ ΠΑΠΑΝΙΚΟΛΑΟΥ

Καθηγητής Ε.Μ.Π.

ΑΘΗΝΑ, 14 Δεκεμβρίου 2012

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

Σχολή Εφαρμοσμένων Μαθηματικών και Φυσικών Επιστημών

Ασυμπτωτική συμπεριφορά των ροπών μιας κατηγορίας τυχαίων μεταβλητών

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

ΔΟΥΜΑΣ Β. ΑΡΙΣΤΕΙΔΗΣ

ΤΡΙΜΕΛΗΣ ΣΥΜΒΟΥΛΕΥΤΙΚΗ ΕΠΙΤΡΟΠΗ:

(i) Β. Γ. ΠΑΠΑΝΙΚΟΛΑΟΥ, Καθ. Ε.Μ.Π. (Επιβλέπων)

(ii) Κ. ΚΥΡΙΑΚΗ, Καθ. Ε.Μ.Π.

(iii) Ι. ΣΠΗΛΙΩΤΗΣ, Αναπλ. Καθ. Ε.Μ.Π.

ΕΠΤΑΜΕΛΗΣ ΕΞΕΤΑΣΤΙΚΗ Ε-ΠΙΤΡΟΠΗ:

(i) Β. Γ. ΠΑΠΑΝΙΚΟΛΑΟΥ, Καθ. Ε.Μ.Π. (Επιβλέπων)

(ii) Κ. ΚΥΡΙΑΚΗ, Καθ. Ε.Μ.Π.

(iii) Ι. ΣΠΗΛΙΩΤΗΣ, Αναπλ. Καθ. Ε.Μ.Π.

(iv) Δ. ΦΟΥΣΚΑΚΗΣ, Επικ. Καθ. Ε.Μ.Π.

(v) Μ. ΛΟΥΛΑΚΗΣ, Επικ. Καθ. Ε.Μ.Π.

(vi) Δ. ΧΕΛΙΩΤΗΣ, Επικ. Καθ. Ε.Κ.Π.Α.

(vii) Α. ΓΙΑΝΝΑΚΟΠΟΥΛΟΣ, Καθ. Ο.Π.Α.

ΑΘΗΝΑ, 14 Δεκεμβρίου 2012

Στον Πνευματικό μου Πατέρα και καθοδηγητή της ζωής μου
Θεοφιλέστατο Επίσκοπο Τανάγρας π. Πολύκαρπο

Στην πολυαγαπημένη σύζυγό μου Ισμήνη και στα παιδάκια μας
Ρωσσάννα-Θεοδώρα και Εμμέλεια-Αναστασία

Στην μνήμη του αδελφού, φίλου, ευεργέτου, και συνεργάτου μου
Παναγιώτη Θεοδωρίδη, συγγραφέα, Ηλ/γου Μηχ/κου και Μηχ/κου
Η/Υ, Ε.Μ.Π.

Περιεχόμενα

ΠΡΟΛΟΓΟΣ ix

SUMMARY xiii

ΕΥΧΑΡΙΣΤΙΕΣ xv

I Το πρόβλημα του συλλέκτη: Ασυμπτωτική συμπεριφορά της διασποράς 1

1 Στοιχεία από την Ασυμπτωτική Ανάλυση 3

2 Το πρόβλημα του συλλέκτη: Ιστορία και Εφαρμογές 9

3 Εισαγωγικά 15

4 Η περίπτωση των ισοπίθανων κουπονιών 21

4.1 Ένα Οριακό Θεώρημα στην περίπτωση των ισοπίθανων κουπονιών 26

5 Η αναγκαιότητα για την ασυμπτωτική μελέτη του προβλήματος. Μία προσέγγιση με ακολουθίες από μέτρα πιθανότητας. 29

6 Η Διχοτομία 35

7 Δεύτερη Ροπή και Διασπορά Ι. Η περίπτωση $L_i(\alpha) < \infty$ 39

8 Δεύτερη Ροπή και Διασπορά ΙΙ. Η περίπτωση $L_i(\alpha) = \infty$ 43

8.1 Ο πρώτος όρος στο ασυμπτωτικό ανάπτυγμα της Δεύτερης Ροπής 43
8.2 Βαθύτερη Ασυμπτωτική Συμπεριφορά της Μέσης Τιμής $E[T_N]$ 47
8.2α′ Το Ολοκλήρωμα $I_1(N)$ 50
8.2β′ Το Ολοκλήρωμα $I_2(N)$ 58
8.3 Βαθύτερη Ασυμπτωτική Συμπεριφορά της Δεύτερης Ροπής $E[T_N(T_N+1)]$ 62
8.3α′ Το Ολοκλήρωμα $I_3(N)$ 62
8.3β′ Το Ολοκλήρωμα $I_4(N)$ 70
8.4 Συμπέρασμα. Ασυμπτωτική Συμπεριφορά της Διασποράς $V[T_N]$ 72

9 Δύο Οριακά Θεωρήματα για το γενικό πρόβλημα 75

II Το πρόβλημα του συλλέκτη: Ασυμπτωτική συμπεριφορά όλων των Ροπών ανώτερης τάξης 81

10 Οι Ανοδικές Ροπές της τ.μ. T_N 83

10.1 Εισαγωγικά 83
10.2 Η Περίπτωση των Ισοπίθανων Κουπονιών 84
10.3 Η αναγκαιότητα για την ασυμπτωτική μελέτη του προβλήματος – Η Διχοτομία 86
10.4 Η περίπτωση $L_r(\alpha) < \infty$ 90
10.5 Η περίπτωση $L_r(\alpha) = \infty$ 90
10.5α′ Ο πρωτεύων όρος του ασυμπτωτικού αναπτύγματος των (ανοδικών) ροπών της τ.μ. T_N 90
10.5β′ Ασυμπτωτικές εκτιμήσεις των (ανοδικών) ροπών της τ.μ. T_N μέσω της σύγκρισης με ακολουθίες, για τις οποίες οι αντίστοιχες ροπές είναι γνωστές 92

III Παραδείγματα 97

11 Εφαρμογές επί γενικών ακολουθιών 99

Α′ Η συνάρτηση Γάμμα 113

Βιβλιογραφία 117

ΠΡΟΛΟΓΟΣ

A man is like a fraction whose numerator is what he is
and whose denominator is what he thinks of himself.
The larger the denominator the smaller the fraction.

Count Lev Nikolgevich Tolstoy (1828-1910)

Η παρούσα διατριβή ασχολείται με μερικές από τις πιο βασικές πτυχές του προβλήματος του συλλέκτη (**Coupon Collector's Problem (CCP)**), ενός εκ των πιο γνωστών προβλημάτων που προέρχονται από τον χώρο των πιθανοτήτων. Θεωρούμε έναν πληθυσμό που αποτελείται από N διαφορετικά αντικείμενα (ψάρια, λέξεις, κάρτες με παίκτες του baseball, ιοί, κτλ.), τα οποία συμβατικά θα αποκαλούμε κουπόνια. Έστω T_N ο αριθμός των δοκιμών, που ανεξάρτητα και με επανάθεση απαιτούνται, ώστε ένας συλλέκτης να έχει μία πλήρη συλλογή και από τα N διαφορετικά κουπόνια.[1] Για την τυχαία μεταβλητή T_N εγείρονται, φυσιολογικά, ερωτήματα ως προς την μέση τιμή, τις ροπές ανωτέρας τάξεως, την διασπορά και, φυσικά, την κατανομή της, καθώς $N \to \infty$. Τούτη η εργασία αποτελείται από τρία μέρη.
Στο πρώτο μέρος παρουσιάζουμε τεχνικές με τις οποίες υπολογίζουμε εις βάθος την ασυμπτωτική συμπεριφορά της πρώτης και δεύτερης ροπής της τ.μ. T_N καθώς $N \to \infty$. Στην συνέχεια προκύπτει ο πρωτεύων όρος επί του ασυμπτωτικού αναπτύγματος της **διασποράς** $V[T_N]$ που αποτελεί τον κύριο αποτέλεσμα του πρώτου μέρους. Συνδυάζοντας τα αποτελέσματα αυτά με δύο γνωστά –πλην όμως γενικά– θεωρήματα βρίσκουμε την **οριακή κατονομή** της ποσότητας T_N (καταλλήλως κανονικοποιημένη), η οποία για μία μεγάλη κλάση κατανομών αποδεικνύεται ότι είναι η συνήθης κατανομή Gumbel.
Στο δεύτερο μέρος και για μία μεγάλη κλάση κατανομών υπολογίζουμε τον πρωτεύοντα όρο επί του ασυμπτωτικού αναπτύγματος **όλων των ροπών** ανωτέρας τάξεως για την τ.μ. T_N.
Τέλος, στο τρίτο μέρος παραθέτουμε μερικά γενικά παραδείγματα, μέσω των οποίων παρουσιάζουμε τα αποτελέσματα του πρώτου και δεύτερου μέρους.

Προς τον αναγνώστη : Μία πιο αναλυτική περίληψη παραθέτουμε στο τέλος του Κεφαλαίου 5, αφού πρώτα έχουμε μελετήσει αναλυτικά την πιο απλή περίπτωση

[1] Ο αναγνώστης θα αναγνωρίσει, πιθανότατα, στο πρόβλημα αυτό την προσπάθεια ενός παιδιού να συμπληρώσει ένα album με 100 αριθμημένες θέσεις από ποδοσφαιριστές της πρώτης κατηγορίας, όταν αγοράζοντας ένα παγωτό βρίσκει κάθε φορά, και μία φωτογραφία ενός ποδοσφαιριστή που βρίσκεται ως δώρο στο κάτω μέρος της συσκευασίας.

του προβλήματος (δηλαδή την περίπτωση κατά την οποία όλα τα κουπόνια έχουν ίση πιθανότητα εκλογής), και ενώ έχουμε εισάγει τον απαιτούμενο συμβολισμό για την αντιμετώπιση του γενικού προβλήματος.
Σε όλη την έκταση της διατριβής υπάρχουν 19 συνολικά παρατηρήσεις και 7 παραδείγματα. Αυτές δεν ακολουθούν αρίθμηση σύμφωνη με την εκάστοτε παραγραφοποίηση, αλλά μία απλή αύξουσα αρίθμηση.

Σχετικά με τα παραπάνω έχουν γίνει οι ακόλουθες δημοσιεύσεις:

- A.V. Doumas and V.G. Papanicolaou, The Coupon Collector's problem revisited: Asymptotics of the Variance, *Adv. Appl. Prob.*, 44 (1) (2012) 166–195.

- A.V. Doumas and V.G. Papanicolaou, Asymptotics of the rising moments for the Coupon Collector's Problem, *Electr. J. Prob.*, accepted for publication.[2]

και οι ακόλουθες παρουσιάσεις στο 14ο Πανελλήνιο συνέδριο Μαθηματικής Ανάλυσης που πραγματοποιήθηκε στην Πάτρα το 2012:

- V.G. Papanicolaou (joint work with A.V. Doumas), Asymptotics and Limit Distributions for the General Collector's Problem.

- A. V. Doumas (joint work with V. G. Papanicolaou), Asymptotic Analysis of the Moments for the Coupon Collector's Problem.

Ο γράφων έδωσε ομιλία με θέμα: Asymptotics and Limit Distributions for the General Collector's Problem, στα πλαίσια του Σεμιναρίου Πιθανοτήτων και Στοχαστικής Ανάλυσης, Οικονομικό Παν/μιο Αθηνών, Μάιος 2012.

Στο περιθώριο του διαδακτορικού υπάρχουν τρεις ακόμη δημοσιευμένες εργασίες:

- V.G. Papanicolaou, A.V. Doumas, On the discrete one-dimensional inverse transmission eigenvalue problem, *Inverse Problems*, 27 (2011), no. 1, 015004, 13 pp,

- A.V. Doumas, An alternative approach to a problem by A. de Moivre, *Math. Gazette*, accepted for publication.[3]

- A.V. Doumas, On the analytic evaluation of a certain class of trigonometric sums, accepted for publication, *Scientia Series A: Mathematical Sciences,*[4]

ενώ έχει υποβληθεί μία ακόμη εργασία:

[2] A.V. Doumas and V.G. Papanicolaou, Asymptotics of the rising moments for the coupon collector's problem, *Electron. J. Probab.* **Vol. 18** (Article no. 41) (2012) 1–15 (doi: 10.1214/EJP.v18-1746).

[3] A. V. Doumas, An alternative approach to a problem by A. de Moivre, *Mathematical Gazette*, **97** (November, No. 540) 446-454 (2013).

[4] To appear in 2014 issue.

- A.V. Doumas and V.G. Papanicolaou, Some new aspects for the Random Coupon Collector's Problem.

Αθήνα, 14 Δεκεμβρίου 2012

Αριστείδης Β. Δούμας

SUMMARY

This Ph.D. thesis deals with a famous Urn problem. We consider a population whose members are of N different *types* (e.g. colors, fish, baseball cards, viruses, ants, words, etc). For $1 \leq j \leq N$ we denote by p_j the probability that a member of the population is of type j, where $p_j > 0$ and $\sum_{j=1}^{N} p_j = 1$. The members of the population are sampled independently with replacement and their types are recorded. The so-called **"coupon collector problem" (CCP)** deals with questions arising in the above procedure. Some key quantities are the moments of the number T_N of trials it takes until all N types are detected (at least once), the variance, and (of course), the distribution of T_N.
Our work here is divided in three parts.
In Part I we develop techniques of computing detailed asymptotics of the first and second moment of the random variable T_N of coupons that a collector has to buy in order to find all N existing different coupons as $N \to \infty$. The probabilities (occurring frequencies) of the coupons can be quite arbitrary. From these asymptotics we obtain the leading behavior of the **variance** $V[\,T_N\,]$ of T_N. Then, we combine our results with the general limit theorems of P. Neal in order to derive the **limit distribution** of T_N (appropriately normalized), which, for a large class of probabilities, turns out to be the standard Gumbel distribution.
In Part II, and for a large class of distributions, we arrive at the leading behavior of the rising **moments** of the random variable T_N as $N \to \infty$. We also present various illutrative examples.
Finally, in Part III we present several general examples that illustrate the results of the previous chapters.

To the reader:
An extensive summary for Part I is exhibited at the end of Section 5, following the case of equal sampling probabilities and introducing the necessary notation for the study of the general problem.
Throughout this thesis there are 19 Remarks and 5 Examples. They follow their own enumeration.

Results with respect to the present thesis have been published as follows:

- A.V. Doumas and V.G. Papanicolaou, The coupon collector's problem revisited: asymptotics of the variance, *Adv. Appl. Prob.*, 44 (1) (2012) 166–195.
- A.V. Doumas and V.G. Papanicolaou, Asymptotics of the rising moments

for the Coupon Collector's Problem, *Electr. J. Prob.*, accepted for publication.[5]

as well as, the following presentations in the 14th Greek Conference on Mathematical Analysis (Patra, 2012):

- V.G. Papanicolaou (joint work with A.V. Doumas), Asymptotics and Limit Distributions for the General Collector's Problem.

- A. V. Doumas (joint work with V. G. Papanicolaou), Asymptotic Analysis of the Moments for the Coupon Collector's Problem.

The author was an invited speaker, at Athens University of Economics and Business, Probability and Stochastic Analysis Seminar. He delivered a talk under the title: Asymptotics and Limit Distributions for the General Collector's Problem, May 2012.

Aside this Ph.D, the author has published the following papers:

- V.G. Papanicolaou, A.V. Doumas, On the discrete one-dimensional inverse transmission eigenvalue problem, *Inverse Problems*, 27 (2011), no. 1, 015004, 13 pp.

- A.V. Doumas, An alternative approach to a problem by A. de Moivre, *Math. Gazette*, accepted for publication.[6]

- A.V. Doumas, On the analytic evaluation of a certain class of trigonometric sums, accepted for publication, *Scientia Series A: Mathematical Sciences*.[7]

Moreover, the following article has been submitted:

- A.V. Doumas and V.G. Papanicolaou, Some new aspects for the Random Coupon Collector's Problem.

Athens 2012, December 14

Aristides V. Doumas

[5] A.V. Doumas and V.G. Papanicolaou, Asymptotics of the rising moments for the coupon collector's problem, *Electron. J. Probab.* **Vol. 18** (Article no. 41) (2012) 1–15 (doi: 10.1214/EJP.v18-1746).

[6] A. V. Doumas, An alternative approach to a problem by A. de Moivre, *Mathematical Gazette*, **97** (November, No. 540) 446-454 (2013).

[7] To appear in 2014 issue.

ΕΥΧΑΡΙΣΤΙΕΣ

Με την ολοκλήρωση της εργασίας αυτής οφείλω θερμές ευχαριστίες στην πολυαγαπημένη μου σύζυγο Ισμήνη για την υποστήριξη και την υπομονή της. Ισμήνη, αυτό το διδακτορικό είναι και δικό σου. Εκφράζω ακόμη την αγάπη μου στους καρπούς της αγάπης μας, τα παιδάκια μας Ρωσσάννα-Θεοδώρα και Εμμέλεια-Αναστασία που προσπάθησα να στερηθώ κατά το δυνατόν λιγότερο τα τελευταία χρόνια. Κορίτσια, είμαι σίγουρος ότι στο μέλλον θα τα καταφέρετε πολύ καλύτερα από εμένα.
Ο επιβλέπων την διατριβή τούτη, καθηγητής κ. Βασίλειος – Γεώργιος Παπανικολάου δεν έχει, ασφαλώς, ανάγκη από τα δικά μου καλά λόγια. Τον ευχαριστώ όμως ιδιαίτερα για την ποιότητα των προβλημάτων που μου εμπιστεύθηκε, την καθοδήγηση, και τον τρόπο με τον οποίο με διόρθωνε. Ακόμη, για τις ιδιαίτερα ευχάριστες συζητήσεις μας, οι οποίες πιστεύω ότι θα συνεχιστούν στο μέλλον.
Είμαι ευγνώμων στην καθηγήτρια (και μέλος της τριμελούς συμβουλευτικής μου επιτροπής) κα. Κυριακή Κυριάκη για τις συμβουλές της σε όλη την διάρκεια αυτού του πονήματος. Περισσότερο όμως, διότι με την δική της ενθάρρυνση αποφάσισα να κάνω, τελικά, αίτηση για διδακτορικές σπουδές.
Ευχαριστώ επίσης, θερμότατα τους καθηγητές κ. Ι. Σπηλιώτη, Δ. Φουσκάκη, Μ. Λουλάκη, Α. Γιαννακόπουλο, και Δ. Χελιώτη που μου έκαναν την τιμή να συμμετάσχουν στην επταμελή επιτροπή προς κρίση της διατριβής αυτής. Ειδικότερα, τον καθηγητή κ. Ι Σπηλιώτη που αποτέλεσε μέλος της τριμελούς συμβουλευτικής μου επιτροπής.
Έχω την αίσθηση ότι οι στιγμές επιτυχίας στην μαθηματική έρευνα αποτελούν, συχνά, ένα σύνολο μηδενικού μέτρου. Ωστόσο, καθίστανται συχνά τόσο ισχυρές που το ταξίδι αυτό να μην τελειώνει ποτέ. Είμαι ευγνώμων στον Θεό για τις ευλογίες Του και ελπίζω ότι τούτο το τέλος θα είναι η αρχή για κάτι νέο. Άλλωστε, όπως ο Billy Corgan (Smashing Pumpkins) έχει τραγουδήσει: *The End Is The Beginning Is The End...*

Αριστείδης Β. Δούμας[8]

[8] Η εργασία αυτή γράφτηκε με την βοήθεια των φίλων μου: **Johann Sebastian Bach, Silversun Pickups, Ramones, Pixies, Sonic Youth, Wipers, Galaxie 500, Feelies, Green on Red, Camper Van Beethoven, Go Betweens, Foo Fighters, Louis Armstrong, Billie Holiday, Fleshtones, Screaming Trees, Dramarama, John Mayall, Driven, Johnny Cash, Ella Fitzgerald, Nirvana, Dream Syndicate, and SUNNYBOYS.**

Μέρος I

Το πρόβλημα του συλλέκτη: Ασυμπτωτική συμπεριφορά της διασποράς

Κεφάλαιο 1

Στοιχεία από την Ασυμπτωτική Ανάλυση

Στο κεφάλαιο αυτό υπενθυμίζουμε στον αναγνώστη ορισμένους βασικούς ορισμούς από την ασυμπτωτική ανάλυση, τους οποίους θα χρησιμοποιήσουμε αρκετά συχνά στην παρούσα διατριβή. Για περισσότερες λεπτομέρειες παραπέμπουμε σχετικά στα [5], [9], [24], και, [60].

Ορισμός 1.0.1. *Εάν $g(x) > 0$ για κάθε $x \geq a$, γράφουμε*

(1.1)

$f(x) = O\left(g(x)\right)$ *(διαβάζουμε: η συνάρτηση $f(x)$ είναι μεγάλο O της $g(x)$)*[1]

εννοώντας ότι η συνάρτηση πηλίκο: $f(x)/g(x)$ είναι φραγμένη για $x \geq a$, δηλαδή ότι υπάρχει μία σταθερά $M > 0$ εις τρόπον ώστε

$$|f(x)| \leq M g(x), \quad \forall\, x \geq a.$$

Εάν η συνάρτηση $g(x)$ είναι μη αρνητική ο αντίστοιχος ορισμός είναι

$$\lim_{x\to\infty} \sup \frac{|f(x)|}{g(x)} < \infty.$$

Μία εξίσωση της μορφής

$$f(x) = h(x) + O\left(g(x)\right) \quad \text{σημαίνει ότι} \quad f(x) - h(x) = O\left(g(x)\right).$$

Σημειώνουμε ότι από την (1.1) δεν έπεται ότι $g(x) = O\left(f(x)\right)$, δηλαδή δεν πρόκειται για μία συμμετρική ισότητα. Ειδικότερα, η σχέση (1.1) δηλώνει ότι η συνάρτηση $f(x)$ ανήκει στο σύνολο $O\left(g(x)\right)$, δηλαδή το σύνολο των συναρτήσεων εις τρόπον ώστε, να υπάρχει μία θετική σταθερά M και ένα a, έτσι ώστε το πηλίκο $f(x)/g(x)$ να είναι φραγμένο, για $x \geq a$. Επομένως, η έκφραση

$$e^x = 1 + x + \frac{x^2}{2} + O\left(x^3\right) \quad \text{καθώς } x \to 0,$$

δηλώνει το γεγονός ότι η διαφορά $e^x - 1 - x - x^2/2$ είναι μικρότερη κατ᾽ απόλυτη τιμή από τον όρο x^3 επί κάποια σταθερά, όταν το x είναι κοντά στο 0.

[1] ο συμβολισμός αυτός συχνά αποκαλείται και Bachmann–Landau O–notation.

Ορισμός 1.0.2. *Εάν*

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1,$$

λέμε ότι η συνάρτηση $f(x)$ είναι ασυμπτωτικά ίση (asymptotic to) ως προς την συνάρτηση $g(x)$ καθώς $x \to \infty$ και γράφουμε

$$f(x) \sim g(x) \quad \text{καθώς } x \to \infty. \tag{1.2}$$

Ο οριμός αυτός ισχύει εάν αντικαταστήσουμε το ∞ με κάθε $x_0 \in \mathbb{R}$.

Επίσης από την

$$f(x) \sim g(x) \quad \text{καθώς } x \to \infty \quad \text{έπεται ότι} \quad g(x) \sim f(x) \quad \text{καθώς } x \to \infty.$$

Έτσι λοιπόν,

$$e^x + x \sim e^x, \quad x \to \infty$$

και παρατηρούμε ότι η διαφορά μεταξύ των δύο μελών της ασυμπτωτικής αυτής ισότητας x, απειρίζεται, καθώς $x \to \infty$. Συνεπώς, ακόμη και εάν δύο συναρτήσεις είναι ασυμπτωτικά ίσες η διαφορά τους μπορεί διαρκώς να μεγαλώνει.

Ορισμός 1.0.3. *Ο συμβολισμός*

$$f(x) = o\left(g(x)\right), \quad x \to \infty \quad \text{(διαβάζουμε: η συνάρτηση } f(x) \text{ είναι μικρό ο της } g(x)\text{)} \tag{1.3}$$

δηλώνει ότι

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0.$$

Μία εξίσωση της μορφής

$$f(x) = h(x) + o\left(g(x)\right) \quad \text{καθώς } x \to \infty, \text{ σημαίνει ότι} \quad f(x) - h(x) = o\left(g(x)\right).$$

Πιο γενικά, η ασυμπτωτική σχέση

$$f(x) \sim g(x), \quad x \to \infty,$$

σημαίνει ότι

$$f(x) = g(x) + o\left(g(x)\right) = g(x)\left[1 + o\left(1\right)\right], \quad x \to \infty.$$

Ας αναφερθεί επίσης ότι εάν

$$f(x) = O\left(1\right) \quad \text{τότε } f(x) = o\left(x\right), \quad x \to \infty.$$

Για παράδειγμα,

$$x = o\left(-10\right), \quad x \to 0^+, \text{ αν και τα πρόσημα των δύο μελών είναι αντίθετα.}$$

Ορισμός 1.0.4. *Η ασυμπτωτική συμπεριφορά της συνάρτησης $f(z)$ περιγράφεται από την δυναμοσειρά $\sum_{n=0}^{\infty} a_n \left(x - x_0\right)^n$ καθώς $x \to x_0$ και γράφουμε*[2]

$$f(x) \sim \sum_{n=0}^{\infty} a_n \left(x - x_0\right)^n \quad \left(x \to x_0\right), \quad \text{εάν}$$

[2] Σε κάποια συγγράμματα (π.χ. [9]) χρησιμοποιείται η ακόλουθη ορολογία: Η δυναμοσειρά $\sum_{n=0}^{\infty} a_n \left(x - x_0\right)^n$ λέγεται ότι είναι *ασυμπτωτική ως προς την συνάρτηση* $f(x)$ (asymptotic to the function $f(x)$), καθώς $x \to x_0$. Αυτός ο ορισμός είναι επίσης γνωστός ως Poincaré expansion.

$$f(x) - \sum_{n=0}^{N} a_n (x - x_0)^n = o(x - x_0)^N \quad \textit{για κάθε } N. \tag{1.4}$$

Συνεπώς, μία δυναμοσειρά είναι ασυμπτωτική ως προς μια συνάρτηση, εάν το υπόλοιπο μετά από N όρους είναι πολύ μικρότερο από τον επόμενο όρο, καθώς $x \to x_0$. Από τον ορισμό αυτό βλέπουμε ότι μία σειρά μπορεί να είναι ασυμπτωτική ως προς μία συνάρτηση, χωρίς αναγκαστικά να είναι συγκλίνουσα.
Συναντούμε επίσης την περίπτωση μη ακεραίων τιμών του $(x - x_0)$.
Η σειρά $\sum_{n=0}^{\infty} a_n (x - x_0)^{an}$ $(a > 0)$ είναι *ασυμπτωτική ως προς την συνάρτηση* $f(x)$ εάν

$$f(x) - \sum_{n=0}^{N} a_n (x - x_0)^{an} = o(x - x_0)^{aN} \, (x \to x_0), \quad \text{για κάθε } N.$$

Εάν $x_0 = \infty$ ο αντίστοιχος ορισμός είναι $f(x) \sim \sum_{n=0}^{\infty} a_n x^{-an}$ $(x \to \infty)$ εάν

$$f(x) - \sum_{n=0}^{N} a_n x^{-an} = o\left(x^{-aN}\right) \, (x \to \infty), \quad \text{για κάθε } N.$$

Ας σημειωθεί ότι δεν μπορούν όλες οι συναρτήσεις να γραφούν ως το ασυμπτωτικό ανάπτυγμα μιας δυναμοσειράς. Για την συνάρτηση e^x δεν μπορεί να βρεθεί δυναμοσειρά του τύπου $\sum_{n=0}^{\infty} a_n x^{-an}$ καθώς $x \to \infty$, διότι καθώς το $x \to +\infty$ το e^x μεγαλώνει.
Εάν η συνάρτηση $f(x)$ δύναται να αναπτυχθεί ασυμπτωτικά ως δυναμοσειρά $f(x) - \sum_{n=0}^{N} a_n x^{-aN} = o\left(x^{-aN}\right)$ $(x \to x_0)$, τότε οι συντελεστές του αναπτύγματος είναι μοναδικοί. Αυτό οφείλεται στο γεγονός ότι ο παραπάνω ορισμός παρέχει έναν ασφαλή τρόπο να υπολογιστούν οι συντελεστές a_n κατά μοναδικό τρόπο:

$$a_0 = \lim_{x \to x_0} f(x),$$

$$a_1 = \lim_{x \to x_0} \frac{f(x) - a_0}{(x - x_0)^a},$$

και γενικότερα,

$$a_N = \lim_{x \to x_0} \frac{f(x) - \sum_{n=0}^{N-1} a_n (x - x_0)^{an}}{(x - x_0)^{aN}}.$$

Η συνθήκη για να έχει μία συνάρτηση ασυμπτωτικά ανάπτυγμα δυναμοσειράς, είναι όλα τα παραπάνω όρια να υπάρχουν.

Η διαφορά μεταξύ συγκλίνουσας και ασυμπτωτικής σειράς είναι η εξής:

Εάν $f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n$ συγκλίνει για $|x - x_0| < R$, τότε το υπόλοιπο $\varepsilon_N(x) := f(x) - \sum_{n=0}^{N} a_n (x - x_0)^n$ τείνει στον μηδέν καθώς $N \to \infty$, για κάθε σταθεροποιημένο x (fixed), $|x - x_0| < R$.
Από την άλλη μεριά, εάν η σειρά είναι ασυμπτωτική ως προς την συνάρτηση $f(x)$ καθώς $x \to x_0$, τότε το υπόλοιπο τείνει στο μηδέν ταχύτερα από $(x - x_0)^N$ καθώς $x \to x_0$, αλλά δεν χρειάζεται να τείνει στο μηδέν καθώς $N \to \infty$, για κάθε x (fixed).
Η σύγκλιση είναι μία απόλυτη έννοια, μία εγγενής ιδιότητα των συντελεστών, του αναπτύγματος, a_n. Μπορεί κανείς, φυσικά, να αποδείξει ότι μία σειρά συγκλίνει

χωρίς να γνωρίζει την συνάρτηση στην οποία συγκλίνει. Ωστόσο, η ασυμπτωτικότητα είναι μία σχετική ιδιότητα των συντελεστών του αναπτύγματος *και* της συνάρτησης $f(x)$ ως προς την οποία είναι ασυμπτωτική η σειρά. Για να να αποδείξει κανείς ότι μία δυναμοσειρά είναι ασυμπτωτική ως προς μία συνάρτηση $f(x)$, πρέπει να θεωρήσει *από κοινού* την συνάρτηση $f(x)$ και τους συντελεστές του αναπτύγματος, βλ. [9] .
Οι περισσότερες λειτουργίες (πρόσθεση, πολλαπλασιασμός, διαίρεση, ολοκλήρωση, κτλ.) μπορούν γενικά να εκτελεστούν στις ασυμπτωτικές σειρές όρο προς όρο, ακριβώς σαν να ήταν συγκλίνουσες σειρές. Σημαντική εξαίρεση, για παράδειγμα, η παραγωγισιμότητα η οποία δεν ισχύει χωρίς επιπλέον περιορισμούς (Tauberian theorems–βλέπε, π.χ., [9]). Λόγου χάριν, οι συναρτήσεις $f(x)$ και

$$g(x) = f(x) + e^{-1/(x-x_0)^2} \sin\left(e^{-1/(x-x_0)^2}\right)$$

διαφέρουν κατά μία συνάρτηση (συνήθως την αποκαλούμε subdominant function) και επομένως έχουν το ίδιο ασυμπτωτικό ανάπτυγμα καθώς $x \to x_0$. Όμως, δεν είναι κατ΄ ανάγκην αληθές ότι συμβαίνει το ίδιο για τις $f'(x)$ και

$$g'(x) = f'(x) - 2\,(x-x_0)^{-3} \cos\left(e^{-1/(x-x_0)^2}\right)$$

$$+2\,(x-x_0)^{-3}\, e^{-1/(x-x_0)^2} \sin\left(e^{-1/(x-x_0)^2}\right)$$

Είναι σημαντικό να αναφέρουμε ότι αν και υπάρχουν πολλές διαφορετικές συναρτήσεις ασυμπτωτικά σε μία δοθείσα δυναμοσειρά, υπάρχει **μόνο μία** ασυμπτωτική δυναμοσειρά για κάθε συνάρτηση.
Συνεχίζουμε με ένα χρήσιμο αποτέλεσμα σχετικά με την ολοκλήρωση όρο προς όρο. Εάν
$f(t,x) \sim \sum_{n=0}^{\infty} f_n(t)\,(x-x_0)^{an}$, $x \to x_0$, $a > 0$, **ομοιόμορφα** για $a \le t \le b$, τότε

$$\int_a^b f(t,x)dt \sim \sum_{n=0}^{\infty} (x-x_0)^{an} \int_a^b f_n(t)dt, \quad x \to x_0,$$

εφόσον, όλοι οι όροι του δεξιού μέλους είναι πεπερασμένοι.

Θυμίζουμε στον αναγνώστη τον τύπο άθροισης των Euler–Maclaurin Euler–Maclaurin Summation Formula[3], ένα ισχυρό εργαλείο για τον υπολογισμό πλήρους μορφής ασυμπτωτικών αναπτυγμάτων σε αθροίσματα του τύπου

$$S(N) = \sum_{m=0}^{N} f(m) \quad \text{καθώς } N \to \infty.$$

Ισχύει,

$$S(N) \sim \int_0^N f(x)dx + C + \frac{1}{2} f(N) + \sum_{k=1}^{\infty} (-1)^{k+1} \frac{B_{k+1}}{(k+1)!} f^{(k)}(N), \tag{1.5}$$

[3] Ο τύπος αυτός ανακαλύφθηκε ανεξάρτητα από τους Leonhard Euler / Colin Maclaurin γύρω στα 1735 (και αργότερα γενικεύτηκε υπό τον τίτλο Darboux formula). Ο Euler την χρειάστηκε για να υπολογίσει άπειρες σειρές που συνέκλιναν αργά, ενώ ο Maclaurin για να υπολογίσει ολοκληρώματα.

όπου C είναι μία σταθερά που δίδεται από την ακόλουθη (συχνά περίπλοκη) έκφραση

$$C = \lim_{m\to\infty}\left[\sum_{j=1}^{m}\frac{(-1)^j B_{j+1}}{(j+1)!}f^{(j)}(0)+\frac{1}{2}f(0)+\frac{(-1)^m}{(m+1)!}\int_0^\infty B_{m+1}\left(t-\lfloor t\rfloor\right)f^{(m+1)}(t)dt\right]$$

και B_k είναι ο k–οστος αριθμός Bernoulli που ορίζεται ως:

$$\frac{z}{e^z-1}=\sum_{k=0}^{\infty}\frac{B_k}{k!}z^k \tag{1.6}$$

(π.χ., $B_0 = 0$, $B_1 = -1/2$, $B_2 = 1/6$, $B_4 = -1/30$, $B_k = 0$, για όλους τους περιττούς, με $k \geq 3$). Οι αριθμοί Bernoulli ικανοποιούν την ακόλουθη γνωστή αναδρομική σχέση (βλέπε, π.χ., [5])

$$B_n = \sum_{k=0}^{n}\binom{n}{k}B_k, \quad n \geq 2 \ \text{ και } B_0 = 1,\ B_1 = -1/2.$$

Για μία απόδειξη της (1.5), βλέπε, π.χ., [5], [6], [52].
Σημειώνουμε ότι ο πρώτος όρος επί της (1.5) είναι είτε το ολοκλήρωμα: $\int_0^N f(x)dx$, (τούτο συμβαίνει εάν $\sum_{n=1}^{\infty} 1/f(n) = \infty$), είτε η σταθερά C (εάν $\sum_{n=1}^{\infty} 1/f(n) < \infty$). Ενδεικτικά παραδείγματα αναφέρουμε στις (4.11), (4.12), αλλά και στο Κεφάλαιο 11.

Κεφάλαιο 2

Το πρόβλημα του συλλέκτη: Ιστορία και Εφαρμογές

Neo, sooner or later you're going to realize,
just as I did,
that there is a difference
between knowing the path, and walking the path.

Morpheus, The Matrix (1999)

Θεωρούμε έναν πληθυσμό από N διαφορετικά *αντικείμενα* (κάρτες με μουσικά συγκροτήματα της δεκαετίας του ΄60, όλες οι διαφορετικές λέξεις που ήξερε ο Shakespeare, ο αριθμός διαφορετικών ψαριών που ζουν στην λίμνη Michigan, κτλ). Τα διαφορετικά αυτά είδη θα τα αποκαλούμε –σε συμφωνία με την γνωστή βιβλιογραφία – "κουπόνια". Ένας συλλέκτης εκτελεί το ακόλουθο πείραμα. Εκλέγει ανεξάρτητα και με επανάθεση κάθε φορά, από ένα κουπόνι, προσπαθώντας να αποκτήσει μία πλήρη συλλογή και από τα N διαφορετικά είδη. Για $1 \leq j \leq N$ συμβολίζουμε με p_j την πιθανότητα ένα μέλος του πληθυσμού να είναι του τύπου j με $p_j > 0$, $\sum_{j=1}^{N} p_j = 1$. Ας σημειωθεί ότι το N θεωρείται ότι είναι γνωστό, καθώς επίσης και ότι το κουπόνι j έχει μονοσήμαντα πιθανότητα p_j να εκλεγεί, $j = 1, 2, \cdots, N$. Το πρόβλημα του συλλέκτη, **"Coupon Collector's Problem" (CCP)** ασχολείται με ερωτήματα που εγείρονται από την παραπάνω διαδικασία.
Συμβολίζουμε με T_N τον αριθμό δοκιμών που απαιτούνται, ώστε όλα τα N διαφορετικά είδη να επιλεγούν (τουλάχιστον μία φορά). Για την τυχαία μεταβλητή T_N, η μέση τιμή, οι ροπές ανώτερης τάξεως, η διασπορά, και φυσικά, η κατανομή είναι ιδιαίτερης σημασίας.
Το CCP ανήκει στην κατηγορία των προβλημάτων κάλπης (urn problems). Στην γενική περίπτωση μία κάλπη περιέχει r διαφορετικά σφαιρίδια. Τα σφαιρίδια εξάγονται διαδοχικά και με επανάθεση, έως ότου k από αυτά έχουν επιλεγεί τουλάχιστον m φορές το κάθε ένα. Πόσες δοκιμές χρειάζονται; Πόσα σφαιρίδια έχουν

επιλεγεί ακριβώς m φορές; Ειδικές περιπτώσεις τέτοιων προβλημάτων συχνά αποκαλούνται π.χ., πρόβλημα των γενεθλίων (birthday problem),[1] πρόβλημα του συλλέκτη (collector's problem), πρόβλημα της κατάληψης (occupancy problem), dixie cup problem. Το πρόβλημα της κατάληψης, για παράδειγμα, ερωτά: πόσα διαφορετικά σφαιρίδια έχουν εξαχθεί μετά από m δοκιμές; Αυτό είναι ουσιαστικά ισοδύναμο με τα Προβλήματα 18 και 19, στις πραγματείες του **A. De Moivre** *De Mensura Sortis,* (1712) και *Doctrine of Chances,* (1756) αντίστοιχα (βλέπε, π.χ., [51], [101]). Το αντίστροφο πρόβλημα που ρωτά: πόσες δοκιμές απιτούνται ώστε να έχουμε k διαφορετικά σφαιρίδια; είναι το πρόβλημα με το οποίο ασχολείται η διατριβή αυτή. Το πρόβλημα του συλλέκτη εισήγαγε ο **P. S. Laplace** στην πρωτοποριακή του πραγματεία *Theorie Analytique de Probabilites, 1812* (βλέπε π.χ., [27]).

Το πρόβλημα έγινε δημοφιλές στα 1930, όταν η αμερικανική εταιρεία Dixie Cup στην προσπάθειά της να αυξήσει τις πωλήσεις της στο παγωτό ατομικής συσκευασίας, και στοχεύοντας στο παιδικό κοινό, τοποθέτησε σε κάθε συσκευασία μία κάρτα, αρχικά με χαρακτήρες του τσίρκου, και αργότερα με αστέρες του Hollywood και παίκτες πρώτης κατηγορίας του baseball.[2] Ίσως, οι παραπάνω πληροφορίες εξηγούν το γεγονός ότι ο όρος *cartophily* αναφέρεται στον τίτλο δύο εκ των πρώτων εργασιών του περασμένου αιώνα σχετικά με το CCP. Οι εργασίες αυτές (και οι δύο δημοσιεύτηκαν στο *Math. Gazette* και έως σήμερα παραμένουν σχετικά άγνωστες) ασχολούνται με την πιο απλή περίπτωση του προβλήματος (βλέπε σχετικά [71], [50]).

Η απλούστερη περίπτωση στο πρόβλημα του συλλέκτη είναι όταν όλα τα κουπόνια έχουν την ίδια πιθανότητα να εκλεγούν, και φυσιολογικά είναι η πρώτη που μελετήθηκε. Τρεις είναι οι κυριότερες αναφορές για τούτη την περίπτωση:

(*i*) το κλασικό σύγγραμμα του **W. Feller**, *An Introduction to Probability Theory and Its Applications, Vol. I,* (1950),[3] [39]. Το κλασικό αποτέλεσμα εδώ είναι

$$E[T_N] = N\left(1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{N-1}+\frac{1}{N}\right), \quad N\to\infty.$$

και από την φόρμουλα άθροισης των Euler - Maclaurin (βλ. Κεφάλαιο 4) λαμβάνουμε άμεσα

$$E\left[T_N\right] = N\ln N + \gamma N + \frac{1}{2} + O\left(\frac{1}{N}\right), \tag{2.1}$$

όπου $\gamma = 0.5772\cdots$ είναι η σταθερά των Euler-Mascheroni.

(*ii*) η εργασία των **D. J. Newman, L. Shepp**, *Double Dixie Cup problem,* (1960), [78], όπου ένα πιο γενικό πρόβλημα μελετήθηκε. Συγκεκριμένα, υπολογίστηκε αναλυτικά το ασυμπτωτικό ανάπτυγμα της μέσης τιμής του αριθμού των δοκιμών που απαιτούνται, για την συμπλήρωση m πλήρων συλλογών και από τα N κουπόνια. Σε αυτό το πρόβλημα, και για κάθε σταθεροποιημένο θετικό ακέραιο m, οι συγγραφείς απέδειξαν ότι

$$E\left[T_m(N)\right] = N\ln N + (m-1)\,N\ln\ln N + NC_m + o(N) \tag{2.2}$$

[1] Το πρόβλημα αυτό εισήγαγε το 1932 ο R. von Mises (βλέπε, π.χ., [27]).

[2] Η κίνηση αυτή είχε πολύ θετικό αποτέλεσμα και την μιμήθηκαν πολλές άλλες εταιρείες. Για την ιστορία της εταιρείας Dixie Cup παραπέμπουμε σχετικά στον σύνδεσμο [104].

[3] Ο αναγνώστης θα παρατηρήσει ότι τα αποτελέσματα [71] και [50] για την κλασική μορφή του προβλήματος, εμφανίστηκαν πριν την πρώτη έκδοση του βιβλίου του W. Feller.

καθώς $N \to \infty$, και όπου C_m είναι μία σταθερά που εξαρτάται από τον αριθμό m. Η φόρμουλα (2.2) μας λέει ότι, κατά μέσο όρο, το πρώτο set κουπονιών 'κοστίζει' $N \ln N + O(N)$, ενώ κάθε επιπλέον set έχει ένα επιπρόσθετο κόστος: $N \ln\ln N + O(N)$, (σύγκρινε με την (2.1)), και

(*iii*) η πολύ γνωστή εργασία των **P. Erdős, A. Rényi**, *On a classical problem of Probability theory*, (1961), [37], όπου παρουσιάστηκε η οριακή κατανομή της τ.μ. που μελετήθηκε στην εργασία των D. J. Newman, L. Shepp (καταλλήλως κανονικοποιημένη). Επίσης, υπολογίστηκε η τιμή της σταθεράς της σχέσης (2.2). Συγκεκριμένα, οι συγγραφείς απέδειξαν ότι

$$C_m = \gamma - \ln (m-1)!. \tag{2.3}$$

καθώς επίσης ότι για κάθε πραγματικό αριθμό x το ακόλουθο οριακό αποτέλεσμα ισχύει:

$$\lim_{N\to\infty} P\left\{ \frac{T_m(N)}{N} < \log N + (m-1)\ln\ln N + x \right\} = \exp\left(-\frac{e^{-x}}{(m-1)!} \right). \tag{2.4}$$

Σημειώνουμε ότι για $m = 1, 2$ δηλαδή μία αλλά και δύο πλήρεις συλλογές κουπονιών, προκύπτει ότι η κατανομή της τ.μ. $T_m(N)$ είναι η συνήθης κατανομή Gumbel (βλέπε επίσης, στην Ενότητα 4.1).
Η τελευταία πρόταση αυτής της εργασίας είναι αξιοσημείωτη, καθώς αποτελεί σύνδεση με την θεωρία γράφων: '*Finally we mention that the problem treated above is analogous to a problem concerning random graphs which we considered recently*', βλέπε και [38].

Έκτοτε, το πρόβλημα μελετήθηκε από μία πληθώρα επιστημόνων και σε αρκετές παραλλαγές, καθώς βρήκε εφαρμογές σε διάφορες επιστημονικές περιοχές (μαθηματικός προγραμματισμός, βελτιστοποίηση, οικολογία, αναζήτηση αλγορίθμων, γλωσσολογία, βλέπε π.χ., [11]).
Ασυμπτωτικά αποτελέσματα στην περίπτωση που οι πιθανότητες εκλογής των κουπονιών είναι ίσες, ή περίπου ίσες, έχουν παρουσιαστεί από διάφορους ερευνητές (βλέπε, π.χ., [8], [59], και [56]). Σημειώνουμε ότι ο Pintacuda υπολόγισε την μέση τιμή $E[T_N]$ (στην κλασική περίπτωση του προβλήματος) χρησιμοποιώντας το martingale stopping theorem, [81]. Πέραν της απλότητάς της αυτή η περίπτωση έχει, επιπλέον, την ιδιότητα της ελάχιστης μέσης τιμής $E[T_N]$, βλέπε [55].

Για την γενική και ασφαλώς πιο ενδιαφέρουσα περίπτωση όπου οι πιθανότητες εκλογής των κουπονιών δεν είναι όλες ίσες, η πρώτη εργασία δημοσιεύτηκε από τον Herman Von Schelling αρχικά το 1934 (πριν (!) την πρώτη έκδοση του W. Feller) σε Γερμανικό περιοδικό περιορισμένης κυκλοφορίας [93]. Το άρθρο αυτό παρέμεινε πρακτικά άγνωστο. Επιπλέον, τα αποτελέσματά του είχαν παρουσιαστεί χωρίς αποδείξεις. Αργότερα, (1954) ο ίδιος ερευνητής δημοσίευσε ολοκληρωμένη εργασία στο περιοδικό *American Mathematical Monthly*, [94]. Ειδικότερα παρουσιάστηκαν γενικές εκφράσεις για την πρώτη και δεύτερη ροπή της τ.μ. T_N:

$$E[T_N] = \sum_{\substack{J \subset \{1,\dots,N\} \\ J \neq \emptyset}} \frac{(-1)^{|J|-1}}{\sum_{j\in J} p_j} = \sum_{m=1}^{N} (-1)^{m-1} \sum_{1 \le j_1 < \cdots < j_m \le N} \frac{1}{p_{j_1} + \cdots + p_{j_m}}$$

και

$$\begin{aligned} E\left[T_N\left(T_N+1\right)\right] &= 2 \sum_{\substack{J\subset\{1,\dots,N\}\\ J\neq\emptyset}} \frac{(-1)^{|J|-1}}{\left(\sum_{j\in J} p_j\right)^2} \\ &= 2\sum_{m=1}^{N}(-1)^{m-1} \sum_{1\leq j_1<\cdots<j_m\leq N} \frac{1}{\left(p_{j_1}+\cdots+p_{j_m}\right)^2}, \end{aligned}$$

βλέπε σχετικά τις σχέσεις (3.15), (3.18) του επομένου κεφαλαίου. Ωστόσο, είναι πολύ δύσκολο να διαχειριστεί κανείς τις παραπάνω εκφράσεις για μεγάλα N. Ακριβώς για τον λόγο αυτό πρέπει να αναζητήσει κανείς την ασυμπτωτική συμπεριφορά, αρχικά της μέσης τιμής $E[\,T_N\,]$. Παρόλα αυτά, τέτοια ασυμπτωτικά αποτελέσματα είχαν βρεθεί μόνο για ειδικές περιπτώσεις των p_j, (τουλάχιστον ώσπου η εργασία [12] δημοσιεύτηκε). Οι κυριότερες είναι:

Η περίπτωση

$$p_j = \frac{2j}{N\,(N+1)}, \quad j=1,2,\cdots,N,$$

γνωστή και ως *γραμμική* (***linear case***). Τότε,

$$E[\,T_N\,] \sim \left[\left(2\pi/\sqrt{3}\right)-3\right] N\,(N+1)\,, \quad N\to\infty,$$

βλέπε [22].

Η περίπτωση

$$p_j = \frac{1}{jH_N}, \quad j=1,2,\cdots,N,$$

όπου με H_N συμβολίζουμε τον N–οστό αρμονικό αριθμό: $H_N := \sum_{k=1}^{N} 1/k$. Η κατανομή αυτή είναι γνωστή ως ***Zipf distribution***. Τότε,

$$E[\,T_N\,] \sim NH_N \ln N,$$

βλέπε [41].

Νωρίτερα (1963), και σε μία πολύ ενδιαφέρουσα εργασία (βλέπε [16]), ο R.K. Brayton, (διδακτορική διατριβή υπό την επίβλεψη του **N. Levinson**) το γενικότερο πρόβλημα των D.J. Newman - L.Shepp (δηλαδή, όταν ο στόχος του συλλέκτη είναι m sets κουπονιών) μελετήθηκε. Για το πρόβλημα αυτό ο Brayton παρουσίασε την ασυμπτωτική συμπεριφορά (πρωτεύων όρος) της διασποράς $V[\,T_m(N)\,]$ υπό ισχυρές υποθέσεις. Συγκεκριμένα οι πιθανότητες p_j θεωρήθηκαν περίπου ίσες. Στο Κεφάλαιο 5 (βλέπε τις σχέσεις (5.15), (5.16)), συζητάμε εκτενώς την εργασία αυτή καθώς και τα αποτελέσματά της σε σχέση με τα αντίστοιχα της παρούσης διατριβής.

Αναμφισβήτητα, τα βαθύτερα αποτελέσματα για την μέση τιμή παρουσιάστηκαν στην εργασία των V.G. Papanicolaou, S. Boneh, [12]. Αυτό το άρθρο αποτέλεσε μία αληθινή έμπνευση για αυτήν την διατριβή. Εδώ οι συγγραφείς υπολόγισαν τον πρώτο όρο στο ασυμπτωτικό ανάπτυγμα της μέσης τιμής $E[\,T_N\,]$, για μία μεγάλη κλάση κατανομών. Ειδικότερα, διαπιστώνουν μία κομβικής σημασίας **διχοτομία** (την οποία υιοθετούμε και εμείς εδώ) και υπό ασθενείς συνθήκες, που ικανοποιούνται από μία πληθώρα κατανομών, καταλήγουν στο ζητούμενο. Μία ακόμη σημαντική συνεισφορά της εργασίας αυτή είναι η φόρμουλα (8.1), την οποία επίσης ασπαζόμαστε στην διατριβή μας με σκοπό τον υπολογισμό της διασποράς $V[T_N]$.

Η κατανομή της τ.μ. T_N έχει, ασφαλώς, ιδιαίτερη σημασία στην γενική περίπτωση του προβλήματος. Ο P. Neal, [76] παρουσίασε το (2008) δύο οριακά–πλην όμως **πολύ γενικά**– θεωρήματα για την τ.μ. T_N. Τα αποτελέσματα αυτά συνδυάστηκαν με τα αντίστοιχα αποτελέσματα από την διατριβή αυτή (βλέπε [29]), και προέκυψε ότι η οριακή κατανομή της T_N (καταλλήλως κανονικοποιημένη), και για μία μεγάλη κλάση κατανομών είναι η συνήθης κατανομή Gumbel.

Όλες οι ροπές ανώτερης τάξης για την τ.μ. T_N έχουν μελετηθεί αναλυτικά στην εργασία [30], για μία μεγάλη κλάση κατανομών. Βλέπε σχετικώς στο Κεφάλαιο 10).

Το παραδοσιακό CCP επεκτείνεται στο ακόλουθο πολύ ενδιαφέρον πρόβλημα. Υποθέτουμε ότι η εταιρεία Dixie Cup πουλά ατομικής συσκευασίας παγωτά τοποθετώντας σε κάθε συσκευασία και μία κάρτα (κουπόνι) με ένα μουσικό συγκρότημα της δεκαετίας του ΄60. Συνολικά υπάρχουν N διαφορετικά κουπόνια, και το κάθε ένα έχει πιθανότητα p_j να εκλεγεί. Ο κύριος και η κυρία Smith έχουν μία κόρη και $(r-1)$ αγόρια, όλα τους εθισμένα στο παγωτό και την μουσική των sixties. Το κορίτσι (είναι η μεγαλύτερη) είναι το μόνο παιδί της οικογένειας που αγοράζει παγωτό. Προσπαθεί να συμπληρώσει μία πλήρη συλλογή με τα N μουσικά συγκροτήματα. Όταν βρίσκει μία νέα κάρτα την τοποθετεί στο album της, ενώ όταν βρει μία που ήδη έχει την δίνει (μαζί με το παγωτό) στον αμέσως μικρότερο αδελφό της, και όταν εκείνος πάρει μία κάρτα που ήδη έχει την δίνει στον αμέσως μικρότερο αδελφό του κ.ο.κ. Όταν το κορίτσι έχει αγοράσει T_N παγωτά, έχει συμπληρώσει το album της, ενώ ταυτόχρονα εκκρεμούν U_k^N κενές θέσεις στο album του k συλλέκτη $k = 2, 3, \cdots, r$, (δηλαδή του $(k-1)$ αδελφού). Η παραλλαγή αυτή συχνά αποκαλείται Γενικευμένο Πρόβλημα του Συλλέκτη (**Generalised CCP (GCCP)**). Ενδιαφέρον παρουσιάζει, φυσιολογικά, η μέση τιμή της τυχαίας ποσότητας U_k^N. Για την περίπτωση $k = 2$ ο Pintacuda, [81] έδωσε απάντηση χρησιμοποιώντας το martingale stopping theorem. Για γενικές – πλην όμως σταθεροποιημένες – τιμές του k το πρόβλημα έχει λυθεί για την περίπτωση των ισοπίθανων κουπονιών, βλέπε σχετικά τις εργασίες [46], [45], [47]. Εάν επιπλέον οι πιθανότητες p_j είναι άνισες έχουν δοθεί πολύ γενικά φράγματα για την $E[U_k^N]$, βλέπε [2]. Απάντηση στο πρόβλημα αυτό για μία μεγάλη κλάση κατανομών μπορεί να βρεθεί στην εργασία [31].

Μία διαφοροποιημένη μορφή του κλασικού προβλήματος παρουσιάστηκε στην εργασία [80], όπου οι πιθανότητες p_n θεωρήθηκαν εις τρόπον ώστε $p_n = a_n/A_N$, όπου a_n είναι ανεξάρτητες και ισόνομες τυχαίες μεταβλητές και $A_N = \sum_{n=1}^{N} a_n$. Εδώ, παρουσιάστηκε η ασυμπτωτική εκτίμηση για την ποσότητα $E\left[\overline{T_N}\right]$ όπου $\overline{T_N}$ δηλώνει την μέση τιμή της T_N, δοθέντων των p_n, $n = 1, 2 \cdots, N$. Οι συγγραφείς αναφέρθηκαν σε αυτήν την μορφή του προβλήματος με τον όρο **"Random Coupon Collector's Problem" (RCCP)**. Το αποτέλεσμα της εργασίας [80] βελτιώθηκε στην [57], όπου βρέθηκε η οριακή κατανομή της T_N για μία μεγάλη κλάση πρότερων (prior) κατανομών. Ασυμπτωτικά αποτελέσματα σχετικά με την μέση τιμή της διασποράς $\overline{T_N}$ μπορούν να βρεθούν στην εργασία [32].

Αρκετές ακόμη ενδιαφέρουσες παραλλαγές του προβλήματος έχουν μελετηθεί. Ενδεικτικά αναφέρουμε τις εργασίες [3], [74], [75], [89], [20], [73], [17], [102], και [69], οι οποίες καλύπτουν την χρονική περίοδο 2001-2012.

Σημαντικά οριακά θεωρήματα έχουν παρουσιαστεί για διάφορες μορφές του κλασικού προβλήματος και για διάφορες τυχαίες μεταβλητές που μπορεί κανείς να μελετήσει. Παραπέμπουμε σχετικά στις εργασίες [85], [36], [86], [63], [96], και, [44].

Από την φύση του, το CCP θεωρείται σημαντικό πρόβλημα και για την Στατιστική. Σχετικά παραπέμπουμε στις εργασίες [18] [27], [26], [42], [84], και [61]. Ειδικότερα, στις εργασίες [27][4] και [26], οι P. Diaconis και S. Holmes μελετούν το κλασικό, αλλά και το γενικό πρόβλημα, θεωρώντας μία αρχική (prior) κατανομή επί των $(p_1, p_1, \cdots, p_N)$ και στην συνέχεια μελετώντας την εκ των υστέρων (posterior) κατανομή.

Αξιοσημείωτη είναι και η σύνδεση του CCP τόσο με τυχαίους περίπατους, Μαρκοβιανές αλυσίδες και martingales, όσο και με την Θεωρία Γράφων. Αρκετές εργασίες έχουν δημοσιευτεί σχετικά. Ο αναγνώστης μπορεί να ανατρέξει στις [33], [7], [66], [98], [90], [4], [23], [95], και [62].

Εφαρμογές του προβλήματος: Όπως προαναφέραμε το CCP έχει βρει αρκετές εφαρμογές σε διάφορες επιστημονικές περιοχές. Μία πληθώρα εργασιών έχουν γραφεί σχετικά. Ενδεικτικά αναφέρουμε εφαρμογές στην βιολογία (βλ. [82]), στην γλωσσολογία (βλέπε, π.χ., τις εργασίες [77], [35] και [83]), αλλά και την πληροφορική (βλέπε, π.χ., [64], [100], [25], [97], [48], [67], [79], [1], [49], [68], [10], και [13].) Ειδικότερα, η εργασία [13] αποτέλεσε το ερέθισμα για το άρθρο **Computer passwords Speak, friend, and enter**, που δημοσιεύτηκε στον *Economist, 24 March 2012*.

Κλείνοντας αυτήν την σύντομη παρουσίαση της ιστορίας του CCP αξίζει να αναφερθεί ότι το πρόβλημα τούτο, ως ένα εκ των κλασικών προβλημάτων της βασικής θεωρίας πιθανοτήτων, αναφέρεται και σε κάθε σύγχρονο σύγγραμμα της περιοχής αυτής. Ενδεικτικά, παραπέμπουμε στα [43], [99], [88], και [70].

[4] Την εργασία αυτή έφερε υπ' όψιν μας ο καθηγητής του ΕΚΠΑ κος Δ. Χελιώτης.

Κεφάλαιο 3

Εισαγωγικά

Στο παρόν κεφάλαιο εισάγουμε τον αναγνώστη στο πρόβλημα και τις ποσότητες που θα μελετήσουμε στην συνέχεια.
Είναι βολικό να θεωρήσουμε τα ενδεχόμενα A_j^k, $1 \leq j \leq N$: το κουπόνι τύπου j **δεν** έχει επιλεγεί ως και την k δοκιμή (συμπεριλαμβάνεται). Υπενθυμίζουμε ότι με T_N συμβολίζουμε τον αριθμό δοκιμών που απαιτείται, ώστε όλα τα N κουπόνια να επιλεγούν (τουλάχιστον μία φορά), ενώ με p_j την πιθανότητα ένα κουπόνι να είναι του τύπου j (όπου $p_j > 0$, $\sum_{j=1}^{N} p_j = 1$). Τότε

$$P\{T_N \geq k\} = P\left(A_1^{k-1} \cup \cdots \cup A_N^{k-1}\right), \qquad k = 1, 2, \ldots .$$

Από την αρχή Εγκλεισμού–Αποκλεισμού έχουμε

$$\begin{aligned} P\{T_N \geq k\} &= P\left(A_1^{k-1}\right) + \cdots + P\left(A_N^{k-1}\right) \\ &\quad - P\left(A_1^{k-1} A_2^{k-1}\right) - \cdots - P\left(A_{N-1}^{k-1} A_N^{k-1}\right) \\ &\quad \vdots \\ &\quad + (-1)^{N-1} P\left(A_1^{k-1} \cdots A_N^{k-1}\right), \end{aligned}$$

και σε μία πιο συμπαγή μορφή

$$P\{T_N \geq k\} = \sum_{\substack{J \subset \{1,\ldots,N\} \\ J \neq \emptyset}} (-1)^{|J|-1} P\left(\bigcap_{j \in J} A_j^{k-1}\right), \tag{3.1}$$

όπου το άθροισμα εκτείνεται επί όλων των $2^N - 1$ μη κενών υποσυνόλων J του $\{1, ..., N\}$, ενώ με $|J|$ δηλώνουμε τον πληθικό αριθμό του συνόλου J[1].

[1] Μπορεί κανείς να αποδείξει αυστηρά την σχέση (3.1) χρησιμοποιώντας τις γνωστές και ως **δείκτριες συναρτήσεις**. Εξ ορισμού, η δείκτρια συνάρτηση ενός υποσυνόλου A του συνόλου X είναι η συνάρτηση $\mathbf{1}_A : X \to \{0, 1\}$ ορισμένη ως

$$\mathbf{1}_A(x) := \begin{cases} 1 & \text{εάν } x \in A, \\ 0 & \text{εάν } x \notin A. \end{cases}$$

Θυμίζουμε στον αναγνώστη τις βασικές ιδιότητες των δεικτριών συναρτήσεων. Εάν A και B είναι δύο υποσύνολα του X, τότε

$$\mathbf{1}_{A\cap B} = \min\{\mathbf{1}_A, \mathbf{1}_B\} = \mathbf{1}_A \cdot \mathbf{1}_B, \quad \mathbf{1}_{A\cup B} = \max\{\mathbf{1}_A, \mathbf{1}_B\} = \mathbf{1}_A + \mathbf{1}_B - \mathbf{1}_A \cdot \mathbf{1}_B,$$

Φυσικά,

$$P\{T_N \geq k\} = 1, \qquad \text{εάν} \quad 1 \leq k \leq N. \tag{3.4}$$

Είναι εύκολο να διαπιστώσει κανείς ότι

$$P\left(A_j^{k-1}\right) = (1-p_j)^{k-1}, \qquad P\left(A_j^{k-1}A_i^{k-1}\right) = [1-(p_j+p_i)]^{k-1},$$

και ούτω καθεξής. Γενικεύοντας, εάν $J \subset \{1,...,N\}$, τότε

$$P\left(\bigcap_{j\in J} A_j^{k-1}\right) = \left[1-\left(\sum_{j\in J} p_j\right)\right]^{k-1}. \tag{3.5}$$

Σημειώνουμε επίσης ότι $P(A_j^0) = 1$ για κάθε j. Ώστε,

$$P\left(\bigcap_{j\in J} A_j^0\right) = 1,$$

για όλα τα υποσύνολα J του $\{1,...,N\}$, συμπεριλαμβανομένου και του $\{1,...,N\}$ (ώστε, για $k=1$ και $J=\{1,...,N\}$, η σχέση(3.5) παραμένει αληθής υπό την σύμβαση: $0^0=1$). Συνδυάζοντας τις σχέσεις (3.5) και (3.1) έχουμε

$$P\{T_N \geq k\} = \sum_{\substack{J\subset\{1,...,N\}\\ J\neq\emptyset}} (-1)^{|J|-1}\left[1-\left(\sum_{j\in J} p_j\right)\right]^{k-1}, \qquad k=1,2,... \tag{3.6}$$

και η δείκτρια συνάρτηση του συμπληρώματος του A, δηλαδή του A^C είναι:

$$\mathbf{1}_{A^C} = 1 - \mathbf{1}_A.$$

Γενικότερα, έστω $A_1,\ldots,A_N$ είναι μία συλλογή υποσυνόλων του X. Για κάθε $x \in X$, η ποσότητα:

$$\prod_{j\in I}(1-\mathbf{1}_{A_j}(x))$$

είναι προφανώς ένα γινόμενο από 0 και 1. Το γινόμενο αυτό έχει τιμή 1 σε ακριβώς αυτά τα $x \in X$ που δεν ανήκουν σε κανένα από τα σύνολα A_j, και την τιμή 0 σε κάθε διαφορετική περίπτωση. Έτσι,

$$\prod_{j\in I}(1-\mathbf{1}_{A_j}) = \mathbf{1}_{X-\bigcup_j A_j} = 1-\mathbf{1}_{\bigcup_k A_j}.$$

Αναπτύσσοντας το γινόμενο που εμφανίζεται στο αριστερό μέλος έχουμε

$$\mathbf{1}_{\bigcup_j A_j} = 1 - \sum_{J\subseteq\{1,2,\ldots,N\}} (-1)^{|J|}\mathbf{1}_{\bigcap_J A_j} = \sum_{\emptyset\neq J\subseteq\{1,2,\ldots,N\}} (-1)^{|J|+1}\mathbf{1}_{\bigcap_J A_j} \tag{3.2}$$

όπου $|J|$ είναι ο πληθικός αριθμός του συνόλου J. Αυτή είναι μία μορφή της αρχής εγκλεισμού–αποκλεισμού.
Ειδικότερα, εάν X είναι χώρος πιθανότητας με μέτρο πιθανότητας $\mathbb{P}$ και A είναι μετρήσιμο σύνολο, τότε η 1_A είναι μία τυχαία μεταβλητή με μέση τιμή ίση με την πιθανότητα του συνόλου A. Δηλαδή, δοθέντος του χώρου πιθανότητας $(\Omega,\mathcal{F},\mathbb{P})$ με $A\in\mathcal{F}$, η (δείκτρια) τυχαία μεταβλητή $\mathbf{1}_A : \Omega\to\mathbb{R}$ ορίζεται ως: $\mathbf{1}_A(\omega)=1$ εάν $\omega\in A$, και $\mathbf{1}_A(\omega)=0$ διαφορετικά. Έτσι,

$$\mathrm{E}(\mathbf{1}_A) = \int_X \mathbf{1}_A(x)\,d\mathbb{P} = \int_A d\mathbb{P} = \mathrm{P}(A). \tag{3.3}$$

Το επιθυμητό αποτέλεσμα έπεται άμεσα παίρνοντας μέσες τιμές στην (3.2) και χρησιμοποιώντας την (3.3).

(η τελευταία σχέση, πάλι, είναι αληθής για την τετριμμένη περίπτωση όπου $k = 1$ παίρνοντας απλώς $0^0 = 1$).

Παρατήρηση 1. Ένα παράπλευρο αποτέλεσμα που προκύπτει από την ανωτέρω ανάλυση, είναι ο (κατά έναν τρόπο) μη τετριμμένος τύπος

$$\sum_{J \subset \{1,...,N\}} (-1)^{|J|} \left[1 - \left(\sum_{j \in J} p_j \right) \right]^n = 0, \quad n = 0, 1, ..., (N-1),$$

στον οποίο καταλήγει κανείς άμεσα χρησιμοποιώντας την σχέση (3.4) στην (3.6).

Λήμμα 3.0.5. *Έστω X μία τ.μ. με τιμές στο $\mathbb{N} = \{0, 1, 2, ...\}$. Εάν $g : \mathbb{N} \to \mathbb{R}$ είναι μία συνάρτηση τέτοια ώστε η $E[\,g(X)\,]$ υπάρχει (δηλαδή, $E[\,|\,g(X)\,|\,] < \infty$), τότε*

$$E[\,g(X)\,] = g(0) + \sum_{k=1}^{\infty} [\,g(k) - g(k-1)\,]\, P\{X \geq k\}\,. \tag{3.7}$$

Ειδικότερα, στην περίπτωση κατά την οποία $g(k) = k$, η σχέση (3.7) έχει την μορφή

$$E[\,X\,] = \sum_{k=1}^{\infty} P\{X \geq k\}\,. \tag{3.8}$$

Απόδειξη. Είναι

$$E[\,g(X)\,] = \sum_{k=0}^{\infty} g(k) P\{X = k\}$$

$$= g(0)P\{X=0\} + g(1)P\{X=1\} + g(2)P\{X=2\} + g(3)P\{X=3\} + \cdots.$$

Το παραπάνω άθροισμα δύναται να γραφεί ως:[2]

$$\begin{aligned} &g(0)\,[\,P\{X=0\} + P\{X=1\} + P\{X=2\} + P\{X=3\} + \cdots] \\ &+ [\,g(1) - g(0)\,]\,[\,P\{X=1\} + P\{X=2\} + P\{X=3\} + \cdots] \\ &+ [\,g(2) - g(1)\,]\,[\,P\{X=2\} + P\{X=3\} + \cdots] \\ &+ [\,g(3) - g(2)\,]\,[\,P\{X=3\} + \cdots] \\ &\vdots \end{aligned}$$

που είναι το δεξιό μέλος της σχέσης (3.7). ∎

Παρατήρηση 2. Εάν $g : \mathbb{N} \to \mathbb{R}$ είναι αύξουσα, τότε η (3.7) ισχύει ακόμη και εάν έχουμε $E[\,g(X)\,] = \infty$ (ομοίως και εάν η συνάρτηση g είναι φθίνουσα).

Για $z \in \mathbb{C}, |z| \geq 1$, εισάγουμε την ακόλουθη πιθανογεννήτρια συνάρτηση της τ.μ. T_N,

$$G(z) := E\left[z^{-T_N}\right] = 1 + \left(z^{-1} - 1\right) \sum_{k=1}^{\infty} z^{-(k-1)} P\{T_N \geq k\} \tag{3.9}$$

[2] Εφαρμόζουμε την μέθοδο άθροισης κατά παράγοντες του Abel.

(η δεύτερη ισότητα προκύπτει εύκολα με άθροιση κατά παράγοντες, δηλαδή με εφαρμογή του Λήμματος 3.0.5 για την συνάρτηση $g(k) = z^{-k}$).
Συνεπώς, από την (3.6) έχουμε

$$G(z) = 1 + \left(z^{-1} - 1\right) \sum_{\substack{J \subset \{1,\ldots,N\} \\ J \neq \emptyset}} (-1)^{|J|-1} \sum_{k=1}^{\infty} z^{-(k-1)} \left[1 - \left(\sum_{j \in J} p_j\right)\right]^{k-1}$$

και αθροίζοντας την γεωμετρική σειρά

$$G(z) = 1 + \left(z^{-1} - 1\right) \sum_{\substack{J \subset \{1,\ldots,N\} \\ J \neq \emptyset}} (-1)^{|J|-1} \frac{1}{1 - z^{-1}\left[1 - \left(\sum_{j \in J} p_j\right)\right]},$$

δηλαδή

$$G(z) = 1 - (z - 1) \sum_{\substack{J \subset \{1,\ldots,N\} \\ J \neq \emptyset}} \frac{(-1)^{|J|-1}}{z - 1 + \left(\sum_{j \in J} p_j\right)}. \tag{3.10}$$

Συνεχίζουμε παρατηρώντας ότι

$$\prod_{j=1}^{N} \left(1 - e^{-p_j t}\right) = \sum_{J \subset \{1,\ldots,N\}} (-1)^{|J|} \exp\left(-t \sum_{j \in J} p_j\right), \tag{3.11}$$

από την οποία συνάγουμε άμεσα ότι

$$\left[1 - \prod_{j=1}^{N} \left(1 - e^{-p_j t}\right)\right] e^{-(z-1)t} = \sum_{\substack{J \subset \{1,\ldots,N\} \\ J \neq \emptyset}} (-1)^{|J|-1} \exp\left[-t\left(z - 1 + \sum_{j \in J} p_j\right)\right].$$

Ώστε, τουλάχιστον για $\Re\{z\} \geq 1$[3],

$$\int_0^{\infty} \left[1 - \prod_{j=1}^{N} \left(1 - e^{-p_j t}\right)\right] e^{-(z-1)t} dt = \sum_{\substack{J \subset \{1,\ldots,N\} \\ J \neq \emptyset}} \frac{(-1)^{|J|-1}}{z - 1 + \left(\sum_{j \in J} p_j\right)}. \tag{3.12}$$

Τέλος, συνδυάζοντας τις σχέσεις (3.10) και (3.12) έχουμε

$$G(z) = 1 - (z - 1) \int_0^{\infty} \left[1 - \prod_{j=1}^{N} \left(1 - e^{-p_j t}\right)\right] e^{-(z-1)t} dt, \tag{3.13}$$

ή ισοδύναμα με την αντικατάσταση $x = e^{-t}$ στο ολοκλήρωμα

$$G(z) = 1 - (z - 1) \int_0^{1} \left[1 - \prod_{j=1}^{N} \left(1 - x^{p_j}\right)\right] x^{z-2} dx. \tag{3.14}$$

[3] για λόγους πληρότητας τούτη η ισότητα ισχύει εάν $\Re\{z - 1\} > -\min_{1 \leq j \leq N} \{p_j\}$, δηλαδή καταλήγει κανείς στην ισότητα (3.12), μόνο εάν $\Re\{z\} > 1 - \min_{1 \leq j \leq N} \{p_j\}$.

Παρατηρούμε ότι
$$E\,[\,T_N\,] = -\lim_{z\to 1^+} G'(z).$$
Έτσι, φθάνει κανείς στις γνωστές εκφράσεις (βλέπε, π.χ., [93], [94], [22], [87])
$$E\,[\,T_N\,] = \sum_{\substack{J\subset\{1,\dots,N\}\\ J\neq\emptyset}} \frac{(-1)^{|J|-1}}{\sum_{j\in J} p_j}$$
$$= \sum_{m=1}^{N}(-1)^{m-1}\sum_{1\le j_1<\cdots<j_m\le N}\frac{1}{p_{j_1}+\cdots+p_{j_m}}, \tag{3.15}$$
$$E\,[\,T_N\,] = \int_0^\infty \left[1-\prod_{j=1}^{N}\left(1-e^{-p_j t}\right)\right]dt, \tag{3.16}$$
και
$$E\,[\,T_N\,] = \int_0^1 \left[1-\prod_{j=1}^{N}\left(1-x^{p_j}\right)\right]\frac{dx}{x}. \tag{3.17}$$
Εναλλακτικά μπορεί κανείς να φθάσει στις παραπάνω σχέσεις για την μέση τιμή $E\,[\,T_N\,]$ απ΄ ευθείας από τις (3.8) και (3.6), δηλαδή χωρίς την χρήση της $G(z)$.

Συνεχίζουμε παρατηρώντας ότι, επίσης από την (3.9) είναι
$$E\,[\,T_N\,(T_N+1)\,] = \lim_{z\to 1^+} G''(z).$$
Για ακόμη μία φορά, από την (3.6), έχει κανείς
$$E\,[\,T_N\,(T_N+1)\,] = 2\sum_{\substack{J\subset\{1,\dots,N\}\\ J\neq\emptyset}} \frac{(-1)^{|J|-1}}{\left(\sum_{j\in J} p_j\right)^2}$$
$$= 2\sum_{m=1}^{N}(-1)^{m-1}\sum_{1\le j_1<\cdots<j_m\le N}\frac{1}{\left(p_{j_1}+\cdots+p_{j_m}\right)^2}, \tag{3.18}$$
$$E\,[\,T_N\,(T_N+1)\,] = 2\int_0^\infty \left[1-\prod_{j=1}^{N}\left(1-e^{-p_j t}\right)\right]t\,dt, \tag{3.19}$$
και
$$E\,[\,T_N\,(T_N+1)\,] = -2\int_0^1 \left[1-\prod_{j=1}^{N}\left(1-x^{p_j}\right)\right]\frac{\ln x}{x}dx. \tag{3.20}$$
Όπως για την μέση τιμή έτσι και οι εκφράσεις για την δεύτερη ροπή $E\,[\,T_N\,(T_N+1)\,]$ μπορούν να προκύψουν χωρίς την $G(z)$, αλλά με εφαρμογή την (3.7) για την συνάρτηση $g(k)=k^2$ και κατόπιν χρησιμοποιώντας την (3.6). Υπενθυμίζουμε ότι οι σχέσεις (3.15) και (3.18) εμφανίστηκαν για πρώτη φορά με ολοκληρωμένη απόδειξη το 1954 στην εργασία του H.V. Schelling, [94].
Φυσικά,
$$V\,[\,T_N\,] = E\,[\,T_N(T_N+1)\,] - E\,[\,T_N\,] - E\,[\,T_N\,]^2\,. \tag{3.21}$$

Κεφάλαιο 4

Η περίπτωση των ισοπίθανων κουπονιών

If I feel unhappy, I do mathematics to become happy.
If I am happy, I do mathematics to keep happy.

Alfréd Rényi

Φυσιολογικά, η απλούστερη περίπτωση σχετικά τις προαναφερθείσες εκφράσεις εμφανίζεται όταν έχουμε

$$p_1 = \cdots = p_N = \frac{1}{N}. \tag{4.1}$$

Είναι εύκολο να αποδειχθεί ότι δοθέντος $t > 0$, η ποσότητα

$$\prod_{j=1}^{N} \left(1 - e^{-p_j t}\right),$$

υπό την συνθήκη

$$p_1 + \cdots + p_N = 1,$$

μεγιστοποιείται όταν όλα τα p_j είναι ίσα μεταξύ τους. Πράγματι, από την ανισότητα αριθμητικού–γεωμετρικού μέσου (AM–GM inequality) έχουμε [1]

$$\prod_{j=1}^{N} \left(1 - e^{-p_j t}\right) \leq \left(1 - \frac{1}{N} \sum_{j=1}^{N} e^{-p_j t}\right)^N. \tag{4.2}$$

Από την ανισότητα Jensen έχουμε επίσης

$$\frac{1}{N} \sum_{j=1}^{N} e^{-p_j t} \geq e^{-t(p_1+p_2+\cdots p_N)/N} = e^{-t/N}, \tag{4.3}$$

[1] εναλλακτικά, μπορεί κανείς να εργαστεί με λογαρίθμους.

με την ισότητα τόσο στην (4.2) όσο και την (4.3) να ισχύει, εάν και μόνον εάν $p_1 = p_2 = \cdots = p_N = 1/N$. Ώστε,

$$\prod_{j=1}^{N} \left(1 - e^{-p_j t}\right) \leq \left(1 - e^{-t/N}\right)^N .$$

Επομένως, από τις εκφράσεις (3.13), (3.16), και (3.19) έχουμε άμεσα ότι η

$$G(z) = E\left[z^{-T_N}\right], \qquad z \geq 1,$$

επιτυγχάνει μέγιστη τιμή, ενώ οι

$$E\left[T_N\right] \qquad \text{και} \qquad E\left[T_N\left(T_N+1\right)\right]$$

επιτυγχάνουν ελάχιστη τιμή, όταν όλες οι πιθανότητες p_j είναι ίσες. Για να βρούμε αυτές τις ακραίες τιμές παρατηρούμε ότι υπό την συνθήκη (4.1), οι σχέσεις (3.10) και (3.14) έχουν την μορφή

$$
\begin{aligned}
G(z) &= 1 - (z-1)\int_0^1 \left[1 - \left(1 - x^{1/N}\right)^N\right] x^{z-2} dx \qquad (4.4)\\
&= 1 - (z-1)\sum_{m=1}^{N} \binom{N}{m} \frac{(-1)^{m-1}}{z-1+(m/N)},
\end{aligned}
$$

ενώ οι σχέσεις (3.15), (3.17) και (3.18), (3.20) αποδίδουν αντίστοιχα

$$E\left[T_N\right] = \int_0^1 \left[1 - \left(1 - x^{1/N}\right)^N\right] \frac{dx}{x} = N \sum_{m=1}^{N} \binom{N}{m} \frac{(-1)^{m-1}}{m}, \qquad (4.5)$$

$$
\begin{aligned}
E\left[T_N\left(T_N+1\right)\right] &= -2\int_0^1 \left[1 - \left(1 - x^{1/N}\right)^N\right] \frac{\ln x}{x} dx\\
&= 2N^2 \sum_{m=1}^{N} \binom{N}{m} \frac{(-1)^{m-1}}{m^2}. \qquad (4.6)
\end{aligned}
$$

Οι παραπάνω εκφράσεις απλοποιούνται σημαντικά με την βοήθεια του μετασχηματισμού, $u = 1 - x^{1/N}$. Από το ολοκλήρωμα της σχέσης (4.4) έχουμε

$$
\begin{aligned}
G(z) &= 1 - N(z-1)\int_0^1 \left(1 - u^N\right)(1-u)^{N(z-1)-1} du\\
&= N(z-1)\int_0^1 u^N (1-u)^{N(z-1)-1} du.
\end{aligned}
$$

Χρησιμοποιώντας τον ορισμό της συνάρτησης Βήτα, αλλά και την γνωστή σχέση η οποία συνδέει τις συναρτήσεις Βήτα και Γάμμα, δηλαδή

$$B(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}, \quad x, y > 0,$$

έχουμε

$$\begin{aligned} G(z) &= N(z-1)\,B\Big((z-1)N, N+1\Big) \\ &= N(z-1)\,\frac{\Gamma\Big((z-1)N\Big)\,\Gamma(N+1)}{\Gamma(zN+1)}. \end{aligned}$$

Μέσω των γνωστών ταυτοτήτων $x\Gamma(x) = \Gamma(x+1),\ x > 0$, και $\Gamma(N+1) = N!$, που ισχύει για όλους τους θετικούς ακεραίους, έχουμε

$$G(z) = N!\,\frac{\Gamma\Big((z-1)N+1\Big)}{\Gamma(zN+1)}. \tag{4.7}$$

Τέλος, από την αναπαράσταση γινομένου της συνάρτησης Γάμμα (βλέπε $(A.6)$), καταλήγουμε

$$G(z) = N!\,e^{\gamma N}\prod_{k=1}^{\infty}\frac{1+(zN/k)}{1+((z-1)N/k)}\,e^{-N/k}.$$

Ομοίως, αντικαθιστώντας $u = 1 - x^{1/N}$ στο ολοκλήρωμα της (4.5) έχουμε

$$E[T_N] = N\int_0^1 \frac{1-u^N}{1-u}du = N\int_0^1\left(\sum_{m=1}^{N}u^{m-1}\right)du = NH_N, \tag{4.8}$$

όπου έχουμε εισάγει τον συμβολισμό

$$H_N := \sum_{m=1}^{N}\frac{1}{m}, \tag{4.9}$$

(δηλαδή τον γνωστό και ως N–οστό αρμονικό αριθμό (N–th harmonic number)). Με την ίδια αντικατάσταση, το ολοκλήρωμα της σχέσεως (4.6) είναι

$$\begin{aligned} E[T_N(T_N+1)] &= -2N^2\int_0^1\frac{1-u^N}{1-u}\ln(1-u)du \\ &= -2N^2\sum_{m=1}^{N}\int_0^1 u^{m-1}\ln(1-u)du. \end{aligned}$$

Με ολοκλήρωση κατά παράγοντες λαμβάνουμε

$$\begin{aligned} E[T_N(T_N+1)] &= 2N^2\sum_{m=1}^{N}\frac{1}{m}\int_0^1\frac{1-u^m}{1-u}du \\ &= 2N^2\sum_{m=1}^{N}\frac{1}{m}\int_0^1\left(\sum_{l=1}^{m}u^{l-1}\right)du, \end{aligned}$$

από όπου βρίσκουμε

$$E[T_N(T_N+1)] = 2N^2\sum_{m=1}^{N}\frac{H_m}{m} = N^2\left(H_N^2 + \sum_{m=1}^{N}\frac{1}{m^2}\right). \tag{4.10}$$

Για να αποδείξουμε την τελευταία ισότητα της σχέσης (4.10) εναλλάσουμε την σειρά άθροισης και χρησιμοποιούμε την (4.9). Είναι

$$\begin{aligned} S_N := & \sum_{m=1}^{N} \frac{H_m}{m} \\ = & \sum_{m=1}^{N} \frac{1}{m} \sum_{k=1}^{m} \frac{1}{k} \\ = & \sum_{k=1}^{N} \frac{1}{k} \sum_{m=k}^{N} \frac{1}{m} \\ = & \sum_{k=1}^{N} \frac{1}{k} (H_N - H_{k-1}) \\ = & H_N \left(\sum_{k=1}^{N} \frac{1}{k} \right) - \sum_{k=1}^{N} \frac{1}{k} H_{k-1} \\ = & H_N^2 - \sum_{k=1}^{N} \frac{1}{k} H_{k-1} \\ = & H_N^2 - \sum_{k=1}^{N} \frac{1}{k} \left(H_k - \frac{1}{k} \right) = H_N^2 - S_N + \sum_{k=1}^{N} \frac{1}{k^2}, \end{aligned}$$

και έχουμε το ζητούμενο.

Τέλος, μπορεί κανείς εύκολα να υπολογίσει το **πλήρες ασυμπτωτικό ανάπτυγμα** των αθροισμάτων H_N και $\sum_{m=1}^{N} m^{-2}$ (καθώς, φυσικά, $N \to \infty$), με εφαρμογή του αθροιστικού τύπου Euler–Maclaurin (βλέπε (1.5)), για τις συναρτήσεις $f(t) = (t+1)^{-1}$ και $g(t) = (t+1)^{-2}$ αντίστοιχα:

$$H_N \sim \ln N + \gamma + \frac{1}{2N} - \sum_{k=2}^{\infty} \frac{B_k}{kN^k} \tag{4.11}$$

και

$$\sum_{m=1}^{N} \frac{1}{m^2} \sim \frac{\pi^2}{6} - \frac{1}{N} + \frac{1}{2N^2} - \sum_{k=2}^{\infty} \frac{B_k}{N^{k+1}}, \tag{4.12}$$

όπου $\gamma = 0.5772...$ είναι η σταθερά του Euler, ενώ B_k είναι ο $k-$οστός αριθμός Bernoulli που ορίζεται από τον τύπο,

$$\frac{z}{e^z - 1} = \sum_{k=0}^{\infty} \frac{B_k}{k!} z^k.$$

Από την (4.11) η σχέση (4.8) δίδει άμεσα

$$E\,[\,T_N\,] \sim N \ln N + \gamma N + \frac{1}{2} - \sum_{k=2}^{\infty} \frac{B_k}{kN^{k-1}}. \tag{4.13}$$

Ομοίως, από τις (4.11),(4.12), και (4.10), μπορεί κανείς να γράψει την πλήρη ασυμπτωτική συμπεριφορά της δεύτερης ροπής $E\left[T_N\left(T_N+1\right)\right]$. Συγκεκριμένα,

$$E\left[\,T_N\left(T_N+1\right)\,\right]=N^2\left[(\ln N)^2+2\gamma\ln N+\gamma^2+\frac{\pi^2}{6}+O\left(\frac{\ln N}{N}\right)\right]. \tag{4.14}$$

Έχοντας τις $E\left[\,T_N\,\right]$ και $E\left[\,T_N\left(T_N+1\right)\,\right]$, μπορούμε να υπολογίσουμε το πλήρες ασυμπτωτικό ανάπτυγμα της διασποράς $V\left[\,T_N\,\right]$ της τ.μ. T_N, από την σχέση (3.21). Έχουμε:

$$V\left[\,T_N\,\right]=\frac{\pi^2}{6}N^2-N\ln N-(\gamma+1)N+O\left(\frac{\ln N}{N}\right). \tag{4.15}$$

Ο συντελεστής $\pi^2/6$ επί του ασυμπτωτικού αναπτύγματος της διασποράς $V\left[\,T_N\,\right]$ σχετίζεται με την συνήθη κατανομή Gumbel (standard Gumbel distribution) και μάλιστα εμφανίζεται σε μία **πληθώρα** κατανομών (βλέπε σχετικά, το Θεώρημα 8.4.1 και την (9.10)). Επίσης, εμφανίστηκε στην περίπτωση όπου οι πιθανότητες p_j είναι περίπου ίσες (βλέπε [16]).
Συνοψίζοντας, αυτή η περίπτωση πέραν της απλότητός της έχει την ιδιότητα ανάμεσα σε όλες τις ακολουθίες, να είναι αυτή με τις ελάχιστες δυνατές ροπές της τ.μ. T_N (βλέπε, επίσης, [55][2] και [87][3]).

♣ Μία διασκεδαστική προσομοίωση του CCP για την περίπτωση των ισοπίθανων κουπονιών είναι διαθέσιμη σε κάθε υπολογιστή (με εγκατεστημένη Java) στον σύνδεσμο [105].

Εικασία 1. Η διασπορά $V\left[\,T_N\,\right]$ γίνεται ελάχιστη όταν όλες οι πιθανότητες p_j είναι ίσες μεταξύ τους.

Σχόλιο. Τα αποτελέσματα της παρούσης διατριβής (βλέπε τα Θεωρήματα 7.0.3 και 8.4.1) συνιστούν ότι ότι για μία μεγάλη κλάση κατανομών, η διασπορά $V\left[\,T_N\,\right]$ πράγματι ελαχιστοποιείται όταν όλα τα κουπόνια έχουν ίση πιθανότητα εκλογής, για αρκούντως μεγάλες τιμές του N.

Εικασία 2. Η διασπορά $V\left[\,T_N\,\right]$ είναι μία συνάρτηση κυρτή κατά Schur (Schur convex function).

Σχόλιο. Εάν η Εικασία 2 είναι αληθής, τότε από μία γνωστή ιδιότητα των κυρτών κατά Schur συναρτήσεων, η συνάρτηση $V\left[\,T_N\,\right]=V\left(p_1,p_2,\cdots,p_N\right)$ ελαχιστοποιείται όταν όλες τις οι μεταβλητές είναι ίσες (βλέπε π.χ., [72], σελ. 305), επομένως η Εικασία 1 είναι αληθής. Για τον λόγο αυτό η Εικασία 2 είναι ισχυρότερη από την Εικασία 1.
Προφανώς η $V\left[\,T_N\,\right]=V\left(p_1,p_2,\cdots,p_N\right)$ παραμένει αναλλοίωτη υπό οποιαδήποτε μετάθεση των μεταβλητών της, άρα είναι συμμετρική. Για να αποδείξει κανείς την

[2] Εδώ αποδείχθη ότι η μέση τιμή $E\left[\,T_N\,\right]$ είναι κυρτή συνάρτηση κατά Schur (Schur convex function). Επομένως λαμβάνει ελάχιστη τιμή όταν όλες οι μεταβλητές της, p_j, είναι ίσες μεταξύ τους. Τούτο το αποτέλεσμα ισχύει επίσης για όλες τις ροπές της τ.μ. T_N. Για μία αναλυτική συζήτηση των Schur συναρτήσεων παραπέμπουμε σχετικά στο σύγγραμμα [72].

[3] Σε μία κατ' ιδίαν επικοινωνία με τον Arnon Boneh, (1990) ο S. Ross επεσήμανε ότι ένας ισχυρότερος ισχυρισμός θα μπορούσε να αποδειχθεί: T_N ελαχιστοποιείται *στοχαστικά* (*minimized stochastically*) όταν όλα τα p_j είναι ίσα, (βλέπε [11]).

Εικασία 2, πρέπει επιπλέον να ισχύει:
(4.16)
$$(p_i - p_j)\left(\frac{\partial V(p_1, p_2, \cdots, p_N)}{\partial p_i} - \frac{\partial V(p_1, p_2, \cdots, p_N)}{\partial p_j}\right) \geq 0, \;\; 1 \leq i \neq j \leq N.$$

Στην περίπτωση αυτή από το Θεώρημα Schur–Osrowski μπορεί να προκύψει το ζητούμενο (βλέπε, ξανά, [72]).

4.1 Ένα Οριακό Θεώρημα στην περίπτωση των ισοπίθανων κουπονιών

Όπως προαναφέραμε ένα από τα πιο σημαντικά αποτελέσματα του κλασικού CCP δηλαδή, όταν $p_j = 1/N$, $j = 1, 2, \cdots, N$, δημοσιεύτηκε το 1961 σε μία από κοινού εργασία των Erdős, Rényi, [37]. Συγκεκριμένα, απέδειξαν ότι (βλέπε επίσης [34])

$$P\left\{\frac{T_N - N \ln N}{N} \leq x\right\} \to \exp\left(-e^{-x}\right), \qquad N \to \infty. \tag{4.17}$$

Ένα πιο γενικό οριακό αποτέλεσμα το οποίο ισχύει για *περίπου ίσες* πιθανότητες, και πιο συγκεκριμένα για αυτές που ικανοποιούν την συνθήκη (5.15), έχει βρεθεί στην εργασία [58]. Παρόμοια οριακά θεωρήματα για πιο γενικές οικογένειες πιθανοτήτων είναι **ιδιαίτερης** σημασίας (βλέπε, την Παρατήρηση 12 του Κεφαλαίου 8, καθώς και το Κεφάλαιο 9).
Για μία απτή εφαρμογή του παραπάνω οριακού θεωρήματος θεωρούμε το πρόβλημα:

Παράδειγμα -1. Ποιά η πιθανότητα σε ένα χωριό 2190 ($= 6 \cdot 365$) κατοίκων, να αντιπροσωπεύονται όλα τα γενέθλια του έτους; Πόσο διαφοροποιείται η απάντηση όταν ο πληθυσμός του χωριού είναι 1825 ($= 5 \cdot 365$) κάτοικοι;

► Έχουμε $N = 365$ και $N \ln N = 2153$. Επομένως,

$$\begin{aligned} P(T_{365} \leq 2190) &= P\left((T_{365} - 2153)/365 \leq 37/365\right) \\ &\approx \exp(-e^{-0.1014}) = \exp(-0.9036) = 0.4051. \end{aligned}$$

Στην δεύτερη περίπτωση:

$$\begin{aligned} P(T_{365} \leq 1825) &= P\left((T_{365} - 2153)/365 \leq -328/365\right) \\ &\approx \exp(-e^{0.8986}) = \exp(-2.4562) = 0.085. \end{aligned}$$

Το ακόλουθο είναι ένα διασκεδαστικό παράδειγμα:

Παράδειγμα 0. Ένας διδακτορικός φοιτητής του Ε.Μ.Π. είναι ο πρώτος που επισκέπτεται κάθε ημέρα την βιβλιοθήκη του ιδρύματος για μελέτη. Η βιβλιοθήκη έχει 200 αριθμημένες θυρίδες από το 1 έως το 200. Κάθε μέρα ο υπάλληλος της βιβλιοθήκης εκλέγει με πιθανότητα 1/200, ένα αριθμημένο κλειδί το οποίο παραλαμβάνει ο φοιτητής μας για την φύλαξη των προσωπικών του αντικειμένων ενώ μελετά. Ποιά πρέπει να είναι η διάρκεια του διδακτορικού, ώστε με πιθανότητα 0.90, ο φοιτητής να έχει δανειστεί και τα 200 κλειδιά (τουλάχιστον μία φορά); Υποθέτουμε ότι η βιβλιοθήκη είναι 365 ημέρες τον χρόνο ανοικτή και ότι ο φοιτητής είναι πρόθυμος να μελετά κάθε ημέρα.

► Εδώ έχουμε $N = 200$ και $N \ln N = 1059.66$. Υποθέτουμε ότι η απάντηση είναι λ χρόνια. Έχουμε

$$\begin{aligned} P\left(T_{200} \leq \lambda \cdot 365\right) &= P\left(\left(T_{200} - 1059.66\right)/200 \leq \left(\lambda \cdot 365 - 1059.66\right)/200\right) \\ &\approx \exp(-e^{-\alpha}) = 0.90, \end{aligned}$$

όπου $\alpha = \left(\lambda \cdot 365 - 1059.66\right)/200$.
Ώστε, $\alpha = -\ln\left[-\ln\left(0.90\right)\right] = 2.25037$, και κατά συνέπεια με πιθανότητα 0.90, το διδακτορικό αυτό θα πρέπει να διαρκέσει όχι λιγότερο από 4.14 χρόνια.

Κεφάλαιο 5

Η αναγκαιότητα για την ασυμπτωτική μελέτη του προβλήματος. Μία προσέγγιση με ακολουθίες από μέτρα πιθανότητας.

Για μεγάλες τιμές του N δεν παίρνουμε ξεκάθαρη πληροφορία από τις σχέσεις (3.15)–(3.17) και (3.18)–(3.20)[1] του Κεφαλαίου 3, για τις ροπές πρώτης και δεύτερης τάξεως, $E[T_N]$, $E[T_N(T_N+1)]$. Κατά συνέπεια, το ίδιο ισχύει και για την διασπορά $V[T_N]$. Για τον λόγο αυτόν είναι ανάγκη να αναπτύξουμε αποδοτικές τεχνικές κατάλληλες, ώστε να περιγράψουμε την ασυμπτωτική συμπεριφορά των ποσοτήτων αυτών, καθώς $N \to \infty$. (Έχουμε ήδη αναλύσει στο προηγούμενο κεφάλαιο την ειδική περίπτωση όπου όλες οι πιθανότητες εκλογής των κουπονιών είναι ίσες βλέπε τις σχέσεις (4.13), (4.14), και (4.15)).
Έστω $\alpha = \{a_j\}_{j=1}^{\infty}$ ακολουθία **αυστηρά** θετικών όρων. Τότε για κάθε ακέραιο $N > 0$, μπορεί κανείς να ορίσει ένα μέτρο πιθανότητας $\pi_N = \{p_1, ..., p_N\}$ επί του συνόλου $\{1, ..., N\}$ ως ακολούθως

$$p_j = \frac{a_j}{A_N}, \qquad \text{όπου} \quad A_N = \sum_{j=1}^{N} a_j. \tag{5.1}$$

Σημειώνουμε ότι οι πιθανότητες p_j **εξαρτώνται, τόσο από την ακολουθία** α, **όσο και από το** N. Συνεπώς, δοθείσης της ακολουθίας α, έχει έννοια να αναζητηθεί η ασυμπτωτική συμπεριφορά των $E[T_N]$, $E[T_N(T_N+1)]$, και $V[T_N]$, καθώς $N \to \infty$. Τούτη η μέθοδος εισαγωγής ακολουθιών από μέτρα πιθανότητας εμφανίστηκε για πρώτη φορά στην εργασία των V. G. Papanicolaou,

[1] Είναι προφανές ότι οι υπολογισμοί για δοθέντα, πλην όμως μεγάλα N, δεν είναι πρακτικώς εφικτοί.

S. Boneh, [12].
Η ακολουθία των μέτρων πιθανότητας π_N, $N = 2, 3, ...$, όπως κατασκευάστηκε από την ακολουθία α μέσω της σχέσης (5.1) έχει μία ενδιαφέρουσα ιδιότητα:
Για κάθε $N = 2, 3, ...$ έστω

$$\Omega_N := \left\{\omega = \{\omega_j\}_{j=1}^{\infty} : \ \omega_j = 1 \text{ ή } 2 \text{ ή} \cdots \text{ή } N\right\}, \tag{5.2}$$

δηλαδή, $\Omega_N = \{1, \dots, N\}^{\mathbb{N}}$, όπου $\mathbb{N} = \{1, 2, \dots\}$. Ο δειγματικός χώρος Ω_N περιγράφει το πείραμα επιλογής N κουπονιών με επανάθεση, εις το διηνεκές.
Σε κάθε m-άδα $(k_1, \dots, k_m) \in \{1, \dots, N\}$ $(m \in \mathbb{N})$ συσχετίζουμε ένα κυλινδρικό υποσύνολο του Ω_N:

$$A_{k_1,\dots,k_m} := \{k_1\} \times \cdots \times \{k_m\} \times \prod_{j=m+1}^{\infty} \{1, \dots, N\}. \tag{5.3}$$

Τότε, η ακολουθία μέτρων $\pi_N = \{p_1, ..., p_N\}$, όπως ορίστηκε στην (5.1), επάγει μία πιθανότητα (a set function) P_N επί αυτών των κυλινδρικών συνόλων:

$$P_N \left\{A_{k_1,\dots,k_m}\right\} := p_{k_1}^N \cdots p_{k_m}^N \tag{5.4}$$

(εδώ, N είναι ένας δείκτης (superscript) που δηλώνει την εξάρτηση των p_k από το N). Από το Θεώρημα Επέκτασης του Καραθεοδωρή (Extension Theorem of Caratheodory), το P_N επεκτείνεται σε ένα πλήρες μέτρο πιθανότητας (complete probability measure), το οποίο επίσης συμβολίζουμε με P_N, επί των $(\Omega_N, \mathcal{F}_N)$, όπου $\mathcal{F}_N$ είναι η πλήρωση της σ-άλγεβρας που παράγεται από τα κυλινδρικά σύνολα.
Εάν L_{N+1} είναι το υποσύνολο του Ω_{N+1} που ορίζουμε ως

$$L_{N+1} := \left\{\omega = \{\omega_j\}_{j=1}^{\infty} : \ \omega_j = N+1 \text{ για όλα, πλην πεπερασμένα το πλήθος } j\right\},$$

τότε είναι προφανές ότι

$$P_{N+1} \left\{L_{N+1}\right\} = 0. \tag{5.5}$$

Για την ακολουθία $\omega = \{\omega_j\}_{j=1}^{\infty} \in \Omega_{N+1} \setminus L_{N+1}$ συμβολίζουμε με $\iota_{N+1}(\omega)$ την ακολουθία (επί του Ω_N) που παράγεται από την ω με διαγραφή όλων των όρων ω_j, έτσι ώστε $\omega_j = N + 1$ (σημειώνουμε ότι η $\iota_{N+1}(\omega)$ είναι μία ακολουθία, έχει δηλαδή άπειρους όρους, καθόσον $\omega \notin L_{N+1}$).
Για να είμαστε πιο ακριβείς, δοθείσης της ακολουθίας $\omega = \{\omega_j\}_{j=1}^{\infty}$ ας θεωρήσουμε το σύνολο δεικτών

$$\{j_n \in \mathbb{N} : \ \omega_{j_n} \neq N + 1, \quad j_n < j_{n+1}\}.$$

Τότε, $\iota_{N+1}(\omega) = \{\omega_{j_n}\}_{n=1}^{\infty}$. Αποκαλούμε: $\iota_{N+1}(\omega)$ την $(N+1)$- *συστολή* ($(N+1)$-*contraction* of) $\omega \in \Omega_{N+1} \setminus L_{N+1}$. Για παράδειγμα, εάν

$$\omega = (3, 1, 2, 3, 3, 2, 1, 2, 2, 3, 1, 1, \dots) \in \Omega_3 \setminus L_3,$$

τότε

$$\iota_3(\omega) = (1, 2, 2, 1, 2, 2, 1, 1, \dots) \in \Omega_2.$$

Καθώς $\Omega_N \subset (\Omega_{N+1} \setminus L_{N+1})$, έχουμε ότι $\iota_{N+1} : (\Omega_{N+1} \setminus L_{N+1}) \to \Omega_N$ είναι μία 'προβολή' ("projection") και, μέσω της ι_{N+1}, το P_{N+1} επάγει ένα μέτρο $\tilde{P}_N$ επί της $\mathcal{F}_N$ που ορίζεται ως

$$\tilde{P}_N\{A\} := P_{N+1}\{\iota_{N+1}^{-1}(A)\}. \tag{5.6}$$

Πρόταση. Τα μέτρα P_N και $\tilde{P}_N$ συμπίπτουν, δηλαδή

$$\tilde{P}_N\{A\} = P_N\{A\} \qquad \text{για κάθε } A \in \mathcal{F}_N.$$

Η απόδειξη είναι άμεση.

Σχόλιο. Για την ακολουθία $\omega = \{\omega_j\}_{j=1}^{\infty} \in \Omega_{N+1} \setminus L_{N+1}$ έχουμε ότι $T_{N+1}(\omega) =$ είναι το μικρότερο k για το οποίο η k–άδα $(\omega_1, \dots, \omega_k)$ περιέχει όλα τα στοιχεία του συνόλου $\{1, \dots, N+1\}$. Εάν θέσουμε $\tilde{T}_N(\omega) := T_N(\iota_{N+1}(\omega))$, τότε, προφανώς $\tilde{T}_N(\omega) \le T_{N+1}(\omega)$. Επίσης, είναι εύκολο να δει κανείς ότι οι τ.μ. $\tilde{T}_N$ και T_N έχουν την ίδια κατανομή. Επομένως η ακολουθία $\{T_N\}_{N=2}^{\infty}$ είναι *στοχαστικά αύξουσα* (*stochastically increasing*), δηλαδή η ανισότητα $P\{T_{N+1} \ge k\} \ge P\{T_N \ge k\}$ ικανοποιείται για όλα τα k και για όλα τα $N \ge 2$.

Μετά από αυτήν την σύντομη παρέκκλιση, επιστρέφουμε στην ασυμπτωτική μελέτη των $E[\,T_N\,]$ και $E[\,T_N(T_N+1)\,]$.
Κινητοποιούμενοι από την σχέση (3.15) εισάγουμε τον συμβολισμό (βλέπε [12])
(5.7)

$$E_N(\alpha) := \sum_{\substack{J \subset \{1,\dots,N\} \\ J \ne \emptyset}} \frac{(-1)^{|J|-1}}{\sum_{j \in J} a_j} = \sum_{k=1}^{N} (-1)^{k-1} \sum_{1 \le j_1 < \cdots < j_k \le N} \frac{1}{a_{j_1} + \cdots + a_{j_k}}.$$

Τότε, όπως στις (3.16) και (3.17), έχουμε

$$E_N(\alpha) = \int_0^{\infty} \left[1 - \prod_{j=1}^{N} \left(1 - e^{-a_j t}\right)\right] dt = \int_0^1 \left[1 - \prod_{j=1}^{N} \left(1 - x^{a_j}\right)\right] \frac{dx}{x}. \tag{5.8}$$

Εάν $s\alpha := \{s a_j\}_{j=1}^{\infty}$, από την (5.7) προκύπτει άμεσα ότι

$$E_N(s\alpha) = \frac{1}{s} E_N(\alpha),$$

επομένως υπό τις (3.15) και (5.1)

$$E[\,T_N\,] = E_N(A_N^{-1}\alpha) = A_N E_N(\alpha). \tag{5.9}$$

Παρομοίως, με εκκίνηση την σχέση (3.18), και για να αναλύσουμε την $E[\,T_N\,(T_N+1)\,]$ εισάγουμε τον συμβολισμό

$$\begin{aligned} Q_N(\alpha) :&= 2 \sum_{\substack{J \subset \{1,\dots,N\} \\ J \ne \emptyset}} \frac{(-1)^{|J|-1}}{\left(\sum_{j \in J} a_j\right)^2} \\ &= 2 \sum_{k=1}^{N} (-1)^{k-1} \sum_{1 \le j_1 < \cdots < j_k \le N} \frac{1}{\left(a_{j_1} + \cdots + a_{j_k}\right)^2}. \end{aligned} \tag{5.10}$$

Όπως στις σχέσεις (3.19) και (3.20), είναι

$$Q_N(\alpha) = 2 \int_0^{\infty} \left[1 - \prod_{j=1}^{N} \left(1 - e^{-a_j t}\right)\right] t\, dt \tag{5.11}$$

και

$$Q_N(\alpha) = -2\int_0^1 \left[1 - \prod_{j=1}^{N}(1 - x^{a_j})\right]\frac{\ln x}{x}dx. \tag{5.12}$$

Από την σχέση (5.10) έχουμε άμεσα

$$Q_N(s\alpha) = \frac{1}{s^2}\,Q_N(\alpha), \tag{5.13}$$

ώστε

$$E\left[T_N(T_N+1)\right] = Q_N(A_N^{-1}\alpha) = A_N^2 Q_N(\alpha). \tag{5.14}$$

Γενική Παρατήρηση. Όπως είχε σημειωθεί στην εργασία των V.G. Papanicolaou, S. Boneh [12], για την μέση τιμή $E\,[\,T_N\,]$, έτσι και το πρόβλημα της δεύτερης ροπής $E\,[T_N(T_N+1)]$, $N \to \infty$, δύναται να αντιμετωπιστεί ως δύο ξεχωριστά προβλήματα, και πιο συγκεκριμένα της εκτίμησης, πρωτίστως, του $Q_N(\alpha)$, και εν συνεχεία του A_N^2. Η ανάλυσή μας εστιάζει στον πρώτο στόχο. Η εκτίμηση του A_N^2 είναι ένα διαφορετικό πρόβλημα το οποίο, μάλιστα, μπορεί κανείς να χειριστεί με γνωστά και ισχυρά εργαλεία, όπως για παράδειγμα η φόρμουλα άθροισης των Euler–Maclaurin **(Euler-Maclaurin Summation formula)**, η μέθοδος Laplace για αθροίσματα **(Laplace method for sums)** ή μέθοδος αθροίσεως κατά παράγοντες **(Abel's summation by parts)**, (βλέπε, π.χ., [9]).

Είμαστε πλέον σε θέση να δώσουμε μία αναλυτικότερη περίληψη των επομένων κεφαλαίων, ως ακολούθως:

Στο Κεφάλαιο 6 προβάλλουμε μία κομβικής σημασίας **Διχοτομία**, σύμφωνα με την οποία αποδεικνύουμε ότι η ακολουθία α που παράγει τις πιθανότητες p_j μπορεί να είναι δύο ειδών.
Το Κεφάλαιο 7 ασχολείται με τις ακολουθίες του πρώτου τύπου. Σε αυτήν την περίπτωση, το κύριο αποτέλεσμα είναι ο πρωτεύων όρος επί του ασυμπτωτικού αναπτύγματος της διασποράς $V[\,T_N\,]$ καθώς $N \to \infty$ (βλέπε το Θεώρημα 7.0.3 και την Παρατήρηση 5).
Στο Κεφάλαιο 8 θεωρούμε μία μεγάλη κλάση των ακολουθιών του δεύτερου είδους. Εδώ, οι υπολογισμοί είναι πολύ πιο περίπλοκοι. Αφού παρουσιάζουμε αναλυτικά το ασυμπτωτικό ανάπτυγμα **έως τον πέμπτο όρο της μέσης τιμής** $E\,[\,T_N\,]$ (στο Θεωρήματα 8.2.5 και την Παρατήρηση 8), καθώς επίσης και **έως τον έκτο όρο** της δεύτερης ροπής $E\,[\,T_N(T_N+1)\,]$ (στο Θεώρημα 8.3.3 και την Παρατήρηση 9), φθάνουμε, τελικά, στον πρωτεύοντα όρο επί του ασυμπτωτικού αναπτύγματος της διασποράς $V[\,T_N\,]$ καθώς $N \to \infty$. Το αποτέλεσμα αυτό παρουσιάζουμε στο στο Θεώρημα 8.4.1. Όπως έχουμε αναφέρει, στο Κεφάλαιο 2, σε μία προηγούμενη εργασία (βλέπε [16]) του R.K. Brayton (Διδακτορική διατριβή υπό την επίβλεψη του **Norman Levinson**) ο πρωτεύων όρος επί του ασυμπτωτικού αναπτύγματος της διασποράς $V[\,T_N\,]$ είχε παρουσιαστεί υπό πολύ περιοριστικούς όρους για τις ακολουθίες α. Ειδικότερα, οι πιθανότητες p_j που θεωρήθηκαν στην εργασία [16], έπρεπε να ικανοποιούν την συνθήκη:

$$\lambda(N) := \frac{\max_{1\le j\le N}\{p_j\}}{\min_{1\le j\le N}\{p_j\}} \le M, \qquad \text{ανεξαρτήτως του } N, \tag{5.15}$$

δηλαδή θεωρήθηκαν "περίπου" ίσες. Το αποτέλεσμα του R.K. Brayton ήταν ότι για κάθε σταθεροποιημένο θετικό ακέραιο m είναι:

$$V\left[T_m(N)\right] = N^2\left[\frac{\pi^2}{6} + O\left(\frac{\ln\ln\ln N}{\ln\ln N}\right)\right] \quad \text{καθώς} \quad N \to \infty. \tag{5.16}$$

Τα αποτελέσματα της παρούσης διατριβής συμπληρώνουν και εν γένει ολοκληρώνουν τα αντίστοιχα της εργασίας [16], αφού καλύπτουν αρκετά γενικές ακολουθίες για τις οποίες το παραπάνω πηλίκο $\lambda(N)$ δεν είναι φραγμένο. Εξόχως, καλύπτουμε σημαντικές οικογένειες κατανομών που έχουν αποσχολήσει επιστήμονες από διάφορες περιχές (για παράδειγμα τις κατανομές Zipf, linear) βλέπε σχετικά το Κεφάλαιο 2.

Στην συνέχεια, και στο Κεφάλαιο 9, χρησιμοποιούμε τα ασυμπτωτικά αναπτύγματα για την μέση τιμή $E[T_N]$ και την διασπορά $V[T_N]$ στα οριακά–πλην γενικά–θεωρήματα του P. Neal, (βλέπε [76]), και παρουσιάζουμε συγκεκριμένα οριακά θεωρήματα για την τ.μ. T_N (βλέπε τις (9.10) και (9.11)).

Στο Δεύτερο μέρος (Κεφάλαιο 10) υπολογίζουμε τον πρώτο όρο του ασυμπτωτικού αναπτύγματος όλων των ροπών της τ.μ. T_N. Η φιλοσοφία είναι πάνω–κάτω η ίδια με αυτήν του Πρώτου μέρους της διατριβής. Αφού αναλύουμε την περίπτωση των ισοπίθανων κουπονιών τα κύρια αποτελέσματα παρουσιάζονται στα Θεωρήματα 10.4.1 και 10.5.1 Επιπλέον, στην Υποενότητα 10.5.2 επεκτείνουμε την κλάση κατανομών για τις οποίες το Θεώρημα 10.5.1 δεν γίνεται να εφαρμοστεί.

Το Τρίτο μέρος της διατριβής εξαντλείται στο Κεφάλαιο 11, όπου και αναφέρουμε, αναλυτικά, μερικά γενικά παραδείγματα μέσω των οποίων απεικονίζονται τα αποτελέσματα των προηγούμενων κεφαλαίων.

Κεφάλαιο 6

Η Διχοτομία

Χάριν ευκολίας συμβολίζουμε

$$f_N^{\alpha}(x) := \prod_{j=1}^{N}(1 - x^{a_j}), \quad 0 \le x \le 1. \tag{6.1}$$

Οι ακόλουθες ιδιότητες των συναρτήσεων f_N^{α} είναι άμεσες:
 (i) $f_N^{\alpha}(0) = 1$ και $f_N^{\alpha}(1) = 0$,
 (ii) $f_N^{\alpha}(x)$ ειναι (μονότονα) φθίνουσα ως προς x,
 (iii) $f_{N+1}^{\alpha}(x) \le f_N^{\alpha}(x)$, ειδικότερα το όριο

$$\lim_N f_N^{\alpha}(x) = \prod_{j=1}^{\infty}(1 - x^{a_j})$$

υπάρχει.
Επομένως από το Θεώρημα μονότονης σύγκλισης (Monotone Convergence Theorem) έχουμε από τις (5.8) και (5.12), αντίστοιχα

$$L_1(\alpha) := \lim_N E_N(\alpha) = \int_0^1 \left[1 - \prod_{j=1}^{\infty}(1 - x^{a_j})\right] \frac{dx}{x} \tag{6.2}$$

και

$$L_2(\alpha) := \lim_N Q_N(\alpha) = -2\int_0^1 \left[1 - \prod_{j=1}^{\infty}(1 - x^{a_j})\right] \frac{\ln x}{x} dx. \tag{6.3}$$

Ας σημειώσουμε ότι $L_1(\alpha)$, $L_2(\alpha) > 0$, για κάθε α (αφού για κάθε $x \in (0,1)$, $f_N^{\alpha}(x) < 1$ με τιμές που φθίνουν καθώς το N αυξάνει). Ωστόσο, μπορούμε να έχουμε $L_1(\alpha) = \infty$ και/ή $L_2(\alpha) = \infty$. Μάλιστα, όπως θα δούμε στην συνέχεια (βλέπε Παρατήρηση 3), $L_1(\alpha) = \infty$, εάν και μόνο εάν, $L_2(\alpha) = \infty$.

Θεώρημα 6.0.1. *$L_2(\alpha) < \infty$, εάν και μόνο εάν υπάρχει $\xi \in (0,1)$, έτσι ώστε*

$$\sum_{j=1}^{\infty} \xi^{a_j} < \infty. \tag{6.4}$$

Πριν την απόδειξη του Θεωρήματος ανακαλούμε το ακόλουθο Λήμμα (βλέπε [12]):

Λήμμα 6.0.2. *Έστω* $\{b_j\}_{j=1}^{\infty}$ *μία ακολουθία πραγματικών αριθμών με* $0 \le b_j \le 1$, *για κάθε* j. *Εάν* $\sum_{j=1}^{\infty} b_j < \infty$, *τότε*

$$\sum_{j=1}^{\infty} b_j - \sum_{1\le l<j} b_l b_j \le 1 - \prod_{j=1}^{\infty}(1-b_j) \le \sum_{j=1}^{\infty} b_j .$$

(Αρχικά αποδεικνύει κανείς το δεξιό σκέλος της διπλής αυτής ανισότητας με επαγωγή και παίρνοντας όριο. Χρησιμοποιώντας αυτό το αποτέλεσμα και με παρόμοιο τρόπο ολοκληρώνεται η απόδειξη.

Απόδειξη του Θεωρήματος 6.0.1. Υποθέτουμε ότι υπάρχει $\xi \in (0,1)$, τέτοιο ώστε η (6.4) είναι αληθής. Τότε από την (6.3) και το Λήμμα 6.0.2 έχουμε

$$\begin{aligned} L_2(\alpha) &\le -2\int_0^{\xi}\left[\sum_{j=1}^{\infty} x^{a_j}\right]\ln x\frac{dx}{x} - 2\int_{\xi}^{1}\left[1-\prod_{j=1}^{\infty}(1-x^{a_j})\right]\ln x\frac{dx}{x} \\ &\le -2\int_0^{\xi}\left[\sum_{j=1}^{\infty} x^{a_j-1}\right]\ln(x)\,dx + \ln^2\xi. \end{aligned} \tag{6.5}$$

Με ολοκλήρωση κατά παράγοντες υπολογίζουμε το ολοκλήρωμα

$$\begin{aligned} I_j(\xi) := \int_0^{\xi} x^{a_j-1}\ln x\,dx &= \left[\frac{x^{a_j}\ln x}{a_j}\right]_{x=0}^{\xi} - \int_0^{\xi}\frac{x^{a_j-1}}{a_j}\,dx \\ &= \frac{1}{a_j}\,\xi^{a_j}\ln\xi - \frac{1}{a_j^2}\,\xi^{a_j}. \end{aligned} \tag{6.6}$$

Από το θεώρημα Tonelli (βλέπε, π.χ., [91]) και την (6.6), η (6.5) αποδίδει

$$L_2(\alpha) \le -2\sum_{j=1}^{\infty}\left(\frac{1}{a_j}\,\xi^{a_j}\ln\xi - \frac{1}{a_j^2}\,\xi^{a_j}\right) + \ln^2\xi. \tag{6.7}$$

Επίσης, από την (6.4) έχουμε $\xi^{a_j} \to 0$[1], επομένως $a_j \to \infty$. Κατά συνέπεια,

$$\min_j \{a_j\} = a_{j_0} > 0.$$

Έχουμε λοιπόν,

$$L_2(\alpha) \le 2\,\frac{1}{a_{j_0}^2}\sum_{j=1}^{\infty}\xi^{a_j} - 2\,\frac{\ln\xi}{a_{j_0}}\sum_{j=1}^{\infty}\xi^{a_j} + \ln^2\xi.$$

Εφόσον $\xi \in (0,1)$, είναι $L_2(\alpha) < \infty$.

Αντιστρόφως, εάν

$$\sum_{j=1}^{\infty}\xi^{a_j} = \infty, \quad \text{για κάθε } \ \xi \in (0,1)$$

[1] Το αντίστροφο, φυσικά, δεν ισχύει πάντοτε.

τότε, από ένα γνωστό θεώρημα των απειρογινομένων[2] (βλέπε, π.χ., [92], σελ. 300)

$$\prod_{j=1}^{\infty}(1-x^{a_j})=0, \qquad \text{για κάθε } x\in(0,1)$$

οπότε από την (6.3) έχουμε

$$L_2(\alpha)=-2\int_0^1(\ln x/x)\,dx=\infty. \qquad \blacksquare$$

Μία εναλλακτική απόδειξη για το αντίστροφο του Θεωρήματος 6.0.1 είναι και η ακόλουθη:
Θέτουμε χάριν ευκολίας $\xi^{a_j}:=q_j$. Προφανώς, $q_j\in(0,1)$ για κάθε $\xi\in(0,1)$. Είναι μία εύκολη άσκηση για κάποιον να αποδείξει ότι ισχύει [3]

$$\exp\left(-\frac{q_j}{1-q_j}\right)\le 1-q_j\le\exp\left(-q_j\right) \tag{6.9}$$

ώστε,

$$\exp\left(-\sum_{j=1}^{\infty}\frac{q_j}{1-q_j}\right)\le\prod_{j=1}^{\infty}(1-q_j)\le\exp\left(-\sum_{j=1}^{\infty}q_j\right).$$

Επομένως, εάν

$$\sum_{j=1}^{\infty}q_j=\infty$$

έχουμε

$$\prod_{j=1}^{\infty}(1-q_j)=0, \qquad \text{για κάθε } \; q_j\in(0,1).$$

Παρατήρηση 3. Έχει αποδειχθεί (βλέπε [12]), ότι $L_1(\alpha)<\infty$, εάν και μόνο εάν υπάρχει $\xi\in(0,1)$ τέτοιο ώστε $\sum_{j=1}^{\infty}\xi^{a_j}<\infty$. Επομένως $L_2(\alpha)<\infty$, εάν και μόνο εάν $L_1(\alpha)<\infty$.

Συνοψίζοντας έχουμε τις ακόλουθες **διχοτομίες**:

$$(i)\quad 0<L_1(\alpha),\ L_2(\alpha)<\infty \qquad \text{ή} \qquad (ii)\quad L_1(\alpha)=L_2(\alpha)=\infty. \tag{6.10}$$

Παρατήρηση 4. Θεωρούμε το ακόλουθο σφάλμα (error term), εκπεφρασμένο ως την διαφορά

$$\Delta_2(N):=L_2(\alpha)-Q_N(\alpha).$$

[2]**Θεώρημα:** Έστω $0\le u_n<1$. Τότε

$$\prod_{n=1}^{\infty}(1-u_n)>0 \;\; \text{εάν και μόνο εάν} \;\; \sum_{n=1}^{\infty}u_n<\infty.$$

[3]Επειδή η εκθετική συνάρτηση είναι κυρτή, έχουμε

$$e^x\ge x+1,\ x\in\mathbb{R}, \tag{6.8}$$

ώστε το δεξιό σκέλος της (6.9) είναι αληθές. Το αριστερό σκέλος είναι ισοδύναμο με $\exp\left(\frac{q_j}{1-q_j}\right)\ge\frac{1}{1-q_j}$. Και αυτή η ανισότητα έπεται από την κυρτότητα της εκθετικής συνάρτησης εάν θέσουμε στην (6.8), $x=q_j/\left(1-q_j\right),\ q_j\in(0,1)$.

Από τις σχέσεις (5.12) και (6.3) με εφαρμογή του Θεωρήματος Tonelli και της ολοκλήρωσης κατά παράγοντες παίρνουμε

$$\Delta_2(N) \leq 2 \sum_{j=N+1}^{\infty} \frac{1}{a_j^2}. \tag{6.11}$$

Επομένως, εάν

$$\sum_{j=N+1}^{\infty} \frac{1}{a_j^2} < \infty, \tag{6.12}$$

η σχέση (6.12) είναι ένα άνω φράγμα (μερικές φορές αρκετά ικανοποιητικό) για το σφάλμα $\Delta_2(N)$. Η ίδια σχέση δείχνει επίσης ότι όσο ταχύτερα η a_j^2 τείνει στο ∞, τόσο ταχύτερη είναι η σύγκλιση του $L_2(\alpha)$ στο $Q_N(\alpha)$.

Κεφάλαιο 7

Δεύτερη Ροπή και Διασπορά I. Η περίπτωση $L_i(\alpha) < \infty$

Έστω A_N και $L_i(\alpha)$ όπως έχουν οριστεί από τις (5.1), (6.2), και (6.3) αντίστοιχα. Σημειώνουμε ότι από το Θεώρημα 6.0.1, $L_i(\alpha) < \infty$ συνεπάγεται ότι $\lim_j a_j = \infty$, (επομένως $\lim_N A_N = \infty$).

Θεώρημα 7.0.3. *Εάν $L_1(\alpha) < \infty$, $L_2(\alpha) < \infty$, τότε καθώς $N \to \infty$,*

$$E\left[\, T_N(T_N+1) \,\right] = A_N^2 L_2(\alpha) \left[1 + o(1)\right], \tag{7.1}$$

και

$$V\left[\, T_N \,\right] = A_N^2 \left[L_2(\alpha) - L_1(\alpha)^2\right] + o\left(A_N^2\right). \tag{7.2}$$

Απόδειξη. Είναι γνωστό (βλέπε [12]) ότι

$$E\left[\, T_N \,\right] = A_N L_1(\alpha) \left[1 + o(1)\right], \qquad N \to \infty. \tag{7.3}$$

Συνεπώς, οι ζητούμενες από το θεώρημα σχέσεις επάγονται άμεσα από τις (5.14), (6.3), και (3.21). ∎

Καθώς $V\left[\, T_N \,\right] > 0$ έχουμε από την (7.2)

$$L_2(\alpha) - L_1(\alpha)^2 \geq 0.$$

Ωστόσο, για να είναι η (7.0.4) ακριβής, θα πρέπει να εξαιρέσουμε την περίπτωση κατά την οποία (πιθανώς) ισχύει

$$L_2(\alpha) = L_1^2(\alpha).$$

Στην περίπτωση αυτή ο πρώτος όρος στο ασυμπτωτικό ανάπτυγμα της διασποράς $V\left[\, T_N \,\right]$, θα προέκυπτε υποχρεωτικά από τον όρο $o\left(A_N^2\right)$.

Θεώρημα 7.0.4.

$$L_2(\alpha) - L_1(\alpha)^2 > 0.$$

Απόδειξη. Θέτουμε

$$F(x) := 1 - \prod_{j=1}^{\infty} \left(1 - x^{a_j}\right), \quad x \in [0,1].$$

Προφανώς η συνάρτηση F είναι αύξουσα στο διάστημα $[0,1]$, δεξιά συνεχής, με $F(0) = 0$ και $F(1) = 1$. Κατά συνέπεια, η $F(\cdot)$ είναι συνάρτηση κατανομής πιθανότητας μίας τυχαίας μεταβλητής X (μη τετριμμένης, καθόσον $L_1(\alpha)$, $L_2(\alpha) < \infty$) με τιμές στο διάστημα $[0,1]$. Υπό τις (6.2) και (6.3), θέλουμε να αποδείξουμε ότι

$$-2\int_0^1 \frac{\ln x}{x}\, F(x)\, dx > \left[\int_0^1 \frac{F(x)}{x}\, dx\right]^2. \tag{7.4}$$

Με ολοκλήρωση κατά παράγοντες μένει μόνο να πιστοποιήσουμε ότι[1]

$$E_F\left[\ln^2 X\right] = \int_0^1 \ln^2 x\, dF(x) > \left[\int_0^1 \ln x\, dF(x)\right]^2 = E_F\left[\ln X\right]^2. \tag{7.5}$$

Η συνάρτηση $\ln^2(x)$ είναι κυρτή, $\forall\, x < e$. Επομένως, η (7.5) είναι αληθής από την ανισότητα Jensen. Ώστε, η (7.4) έπεται. Σημειώνουμε ότι η ανισότητα είναι αυστηρή, επειδή η συνάρτηση $\ln x$ δεν είναι σταθερή συνάρτηση στο διάστημα $[0,1]$. ■

Συνοψίζοντας, ένα μέρος του Θεωρήματος 7.0.3 (συγκεκριμένα η σχέση (7.2)), λαμβάνει τελικά την μορφή που περιγράφεται στην ακόλουθη

Παρατήρηση 5. Εάν $L_1(\alpha) < \infty,\ L_2(\alpha) < \infty$, τότε καθώς $N \to \infty$,

$$V\left[\,T_N\,\right] = A_N^2\left[L_2(\alpha) - L_1(\alpha)^2\right]\,[1 + o(1)]. \tag{7.6}$$

Συνεχίζουμε με ένα αρκετά ενδιαφέρον σχόλιο. Έστω

$$x_s := \sup\{x \geq 0 \,:\, \sum_{j=1}^{\infty} x^{a_j} < \infty\}.$$

Φυσικά, $x_s \leq 1$. Από το Θεώρημα 6.0.1 έχουμε επίσης, $x_s > 0$. Η συνάρτηση $F(x) = 1 - \prod_{j=1}^{\infty}\left(1 - x^{a_j}\right)$ που εμφανίζεται στην απόδειξη του Θεωρήματος 7.0.4 είναι συνεχής στο διάστημα $[0, x_s)$. Άρα, το απειρογινόμενο συγκλίνει ομοιόμορφα στα συμπαγή υποσύνολα του $[0, x_s)$ (π.χ., από το M–test του Weierstrass επί της σειράς $\sum_{j=1}^{\infty} \ln(1 - x^{a_j})$). Επιπλέον, $F(x) \equiv 1$, για $x > x_s$, ενώ $F(x)$ είναι αριστερά συνεχής στο $x = x_s$, από το Θεώρημα μονότονης σύγκλισης επί της σειράς $\sum_{j=1}^{\infty} \ln(1 - x^{a_j})$.
Εάν $x_s = 1$, τότε έχουμε ότι η συνάρτηση $F(x)$ είναι συνεχής στο διάστημα $[0,1]$. Ωστόσο, εάν $x_s < 1$, τότε η $F(x)$ δύναται να έχει άλμα ασυνέχειας στο $x = x_s$. Για παράδειγμα:
(i) εάν $a_j = \ln j$, τότε $x_s = e^{-1}$ και η $F(x)$ είναι συνεχής στο $x = x_s$
(ii) όμως εάν $a_j = \ln j + 2\ln\ln j$, τότε $x_s = e^{-1}$ και η $F(x)$ δεν είναι συνεχής στο $x = x_s$.

[1] Η συνάρτηση $F(\cdot)$ είναι αύξουσα και επομένως φραγμένης κύμανσης. Τα ολοκληρώματα Riemann–Stieltjes της ανισότητας (7.5) είναι καλώς ορισμένα.

Γενική Παρατήρηση. Το Θεώρημα 7.0.3 κελεύει ότι εάν $L_1(\alpha) < \infty$, $L_2(\alpha) < \infty$ τότε η ασυμπτωτική συμπεριφορά των $E[\,T_N\,]$, $E\,[\,T_N(T_N+1)\,]$ και (ασφαλώς) της διασποράς $V[\,T_N\,]$ καθορίζεται (διαισθητικά) από την αντίστοιχη της ποσότητας A_N. Όπως έχουμε ήδη αναφέρει υπάρχουν αρκετές μέθοδοι για τον υπολογισμό αυτό οι οποίες μπορούν να εφαρμοστούν σε μία πληθώρα περιπτώσεων. Εναλλακτικά, μπορεί κανείς να προσφύγει σε συγκεκριμένα χαρακτηριστικά της ακολουθίας. Για παράδειγμα, εάν η ακολουθία α είναι του τύπου

$$a_j = e^{j\,c_j}, \quad \text{όπου} \quad c_j \nearrow \infty, \quad \text{τότε καθώς} \quad N \to \infty, \tag{7.7}$$

$$A_N = \sum_{j=1}^{N} a_j \sim a_N, \tag{7.8}$$

Για να αποδείξουμε την ασυμπτωτική ισότητα της (7.8) χρησιμοποιούμε την μονοτονία της a_j και αθροίζουμε την ακολούθως προκύπτουσα γεωμετρική σειρά. Αναλυτικότερα,

$$\begin{aligned} A_N &= \sum_{j=1}^{N} e^{j\,c_j} \le \sum_{j=1}^{N} e^{j\,c_N} \\ &= \frac{e^{c_N(N+1)}-1}{e^{c_N}-1} \le M e^{c_N(N+1)}, \end{aligned}$$

όπου $M = 1/(e^{c_1}-1)$. Επειδή,

$$\lim_{N\to\infty} M\frac{e^{c_N(N+1)}}{e^{c_{N+1}(N+1)}} = \lim_{N\to\infty} M e^{(c_N - c_{N+1})\,(N+1)} = 0,$$

έχουμε το ζητούμενο, δηλαδή την (7.8). Με άλλα λόγια, εάν η ακολουθία ικανοποιεί τις (7.7), τότε σε ένα άθροισμα του τύπου της σχέσης (7.8), ο τελευταίος όρος **ακυρώνει** όλους τους προγενέστερους όρους.

Παραδείγματα τέτοιων ακολουθιών είναι: $a_j = e^{j^r}$ με $r > 1$, $a_j = j^j$ και $a_j = j!$ (βλέπε, το Παράδειγμα 5 του Κεφαλαίου 11).

Κεφάλαιο 8

Δεύτερη Ροπή και Διασπορά ΙΙ. Η περίπτωση $L_i(\alpha) = \infty$

Όπως θα δούμε αυτή η περίπτωση είναι πολύ πιο απαιτητική.

8.1 Ο πρώτος όρος στο ασυμπτωτικό ανάπτυγμα της Δεύτερης Ροπής

Από το Θεώρημα 6.0.1 και την Παρατήρηση 3, εάν $L_2(\alpha) = \infty$ έχουμε ισοδύναμα $L_1(\alpha) = \infty$ και επίσης (ισοδύναμα),

$$\sum_{j=1}^{\infty} x^{a_j} = \infty, \qquad \text{για κάθε } x \in (0,1).$$

Για την μετέπειτα ανάλυσή μας ακολουθούμε τους V. Papanicolaou, S. Boneh, (βλέπε [12]), και γράφουμε την ακολουθία a_j στην μορφή

$$a_j = \frac{1}{f(j)}, \tag{8.1}$$

όπου

$$f(x) > 0 \qquad \text{και} \qquad f'(x) > 0. \tag{8.2}$$

Υποθέτουμε ότι η συνάρτηση $f(x)$ ικανοποιεί τις ακόλουθες (ασθενείς) υποθέσεις, καθώς $x \to \infty$:

$$\begin{aligned}
&(i) && f(x) \to \infty, \\
&(ii) && \frac{f'(x)}{f(x)} \to 0, \\
&(iii) && \frac{f''(x)/f'(x)}{[f'(x)/f(x)]} = O(1), \\
&(iv) && \frac{f'''(x)\ f(x)^2}{f'(x)^3} = O(1).
\end{aligned} \tag{8.3}$$

Για τον αναγνώστη θυμίζουμε ότι στην εργασία [12], ο πρώτος όρος στο ασυμπτωτικό ανάπτυγμα της μέσης τιμής είχε βρεθεί υπό τις ελαφρώς ασθενέστερες συνθήκες:
(8.4)
$$(i^{'})\ f(x) \nearrow \infty,\ (ii^{'})\ \frac{f'(x)}{f(x)} \searrow 0, \text{ και } (iii^{'})\ \frac{f''(x)/f'(x)}{[f'(x)/f(x)]\ln[f'(x)/f(x)]} \to 0.$$

Για να δώσουμε στον αναγνώστη μία ιδέα για την κλάση (ας τις καλέσουμε $\mathcal{S}$), στην οποία ανήκουν οι συναρτήσεις $f(\cdot)$ που ικανοποιούν τις συνθήκες (8.3), έχουμε να παρατηρήσουμε τα εξής. Η κλάση $\mathcal{S}$ περιέχει τις θετικές και αυστηρώς αύξουσες $C^3(0,\infty)$ συναρτήσεις $f(x)$, **που τείνουν στο ∞ (καθώς $x\to\infty$) πιο αργά από τις εκθετικές συναρτήσεις, αλλά ταχύτερα από αυτές που είναι δυνάμεις λογαρίθμων**.
Για παράδειγμα, οι συναρτήσεις

$$f(x)=x^p(\ln x)^q, \quad p>0,\ q\in\mathbb{R}, \qquad\qquad f(x)=\exp(x^r), \quad 0<r<1,$$

όπως επίσης και οι κυρτοί συνδυασμοί γινομένων τέτοιων συναρτήσεων, ανήκουν στην κλάση $\mathcal{S}$. Από την άλλη μεριά, η συνάρτηση $f(x)=e^{bx}$, $b>0$ (και, γενικότερα η $f(x)=e^{\phi(x)}$ με $\phi'(x)\geq b>0$ για αρκούντως μεγάλες τιμές του x), δεν ανήκει στην $\mathcal{S}$ καθώς παραβιάζει την συνθήκη (ii) της (8.3), και η $f(x)=(\ln x)^\gamma$, $\gamma>0$ επίσης, δεν ανήκει στην κλάση $\mathcal{S}$ διότι παραβιάζει την συνθήκη (iii) της (8.3). Οι συνθήκες (8.3) (iii) και (iv) είναι περισσότερο τεχνικές συνθήκες. Ειδικότερα, η (iii) είναι καθαρά τεχνική και μας λέει ότι η λογαριθμική παράγωγος της $f'(x)$ κυριαρχείται από την λογαριθμική παράγωγο της συνάρτησης $f(x)$ – διαισθητικά, εάν ο ρυθμός αύξησης της $f(x)$ είναι πολύ αργός, τότε η $f(\cdot)$ παραβιάζει την (8.3) (iii). Συνοψίζοντας, η κλάση $\mathcal{S}$ είναι μία *πολύ πλούσια κλάση* και σε αυτήν θα επικεντρωθούμε στην συνέχεια.

Παρατήρηση 6.
α. Από τις συνθήκες (i) και (ii) της (8.3), προκύπτει άμεσα ότι

$$\lim_{x\to+\infty}\frac{f(x+1)}{f(x)}=1. \tag{8.5}$$

Πράγματι, αρκεί κανείς να θεωρήσει την συνάρτηση $g(x):=\ln f(x)$ και να εφαρμόσει το Θεώρημα της Μέσης Τιμής του Διαφορικού Λογισμού.

β. Από την συνθήκη (iii) σε συνδυασμό με τις (i) και (ii) έχουμε

$$\frac{\ln f'(x)}{\ln f(x)}=O(1). \tag{8.6}$$

Αποδεικνύει κανείς την (8.6) ξεκινώντας από την (iii). Υπάρχει μία θετική σταθερά M, τέτοια ώστε για αρκούντως μεγάλα x

$$\left|(\ln f'(x))'\right| \leq M\left|(\ln f(x))'\right|.$$

Επειδή $f(x),\ f'(x)>0$ η παραπάνω σχέση γίνεται

$$\left|(\ln f'(x))'\right| \leq M\,(\ln f(x))'.$$

Για αρκούντως μεγάλες τιμές των x_0, x έχουμε

$$\int_{x_0}^{x} \left|(\ln f'(x))'\right| dx \leq M \int_{x_0}^{x} (\ln f(x))' .$$

Το ολοκλήρωμα της απόλυτης τιμής μίας συνάρτησης είναι μεγαλύτερο, ή ίσο, της απόλυτης τιμής του ολοκληρώματός της. Επομένως,

$$\left|\int_{x_0}^{x} (\ln f'(x))' \, dx\right| \leq M \int_{x_0}^{x} (\ln f(x))'$$

ώστε,

$$|\ln f'(x) - \ln f'(x_0)| \leq M \left(\ln f(x) - \ln f(x_0)\right).$$

Εάν διαιρέσουμε και τα δύο μέλη της τελευταίας ανισότητας με την θετική συνάρτηση $\ln f(x)$ και χρησιμοποιήσουμε την (i) έχουμε το ζητούμενο.

Θεώρημα 8.1.1. *Εάν $\alpha = \{1/f(j)\}_{j=1}^{\infty}$, και η συνάρτηση f ικανοποιεί τις συνθήκες (8.2) και (8.3), τότε*

$$Q_N(\alpha) \sim f(N)^2 \ln \left(\frac{f(N)}{f'(N)}\right)^2 = F(N)^2, \qquad N \to \infty \tag{8.7}$$

Απόδειξη. (Υιοθετούμε την απόδειξη που δόθηκε στην εργασία των V. Papanicolaou, S. Boneh, (βλέπε [12]), για την εύρεση του πρώτου όρου στο ασυμπτωτικό ανάπτυγμα της ποσότητας $E_N(\alpha)$).
Θέτουμε

$$F(x) := -f(x) \ln \left[\frac{f'(x)}{f(x)}\right]. \tag{8.8}$$

Ας σημειωθεί ότι από τις (8.2) και (ii) της (8.3), έχουμε $F(x) > 0$, για αρκούντως μεγάλα x. Επομένως, μπορεί κανείς να γράψει την (5.11) ως:

$$\begin{aligned} Q_N(\alpha) &= F(N)^2 \, Q_N\left[\, F(N)\, \alpha \,\right] \\ &= 2F(N)^2 \int_0^1 \left[1 - \exp\left(\sum_{j=1}^{N} \ln\left(1 - e^{-\frac{F(N)}{f(j)} s}\right)\right)\right] s \, ds \\ &+ 2F(N)^2 \int_1^{\infty} \left[1 - \exp\left(\sum_{j=1}^{N} \ln\left(1 - e^{-\frac{F(N)}{f(j)} s}\right)\right)\right] s \, ds. \end{aligned} \tag{8.9}$$

Έχει αποδειχθεί ότι (βλέπε [12]),

$$\lim_{N} \sum_{j=1}^{N} \ln\left(1 - e^{-\frac{F(N)}{f(j)} s}\right) = \begin{cases} -\infty, & \text{εάν } s < 1; \\ 0, & \text{εάν } s \geq 1, \end{cases} \tag{8.10}$$

καθώς επίσης

$$\int_1^N e^{-\frac{F(N)}{f(x)} s} dx \sim \frac{1}{s \ln\left[f(N)/f'(N)\right]} \left[\frac{f(N)}{f'(N)}\right]^{1-s}. \tag{8.11}$$

Αυτά τα δύο αποτελέσματα προέκυψαν υπό τις συνθήκες (8.4).[1]
Εφαρμόζουμε το Θεώρημα φραγμένης σύγκλισης (Bounded Convergence Theorem) στο πρώτο ολοκλήρωμα της σχέσης (8.9) και με την βοήθεια της (8.10) έχουμε

$$\begin{aligned} Q_N(\alpha) &= 2F(N)^2\left[\frac{1}{2}+o(1)\right] \\ &+ 2F(N)^2\int_1^\infty\left[1-\exp\left(\sum_{j=1}^N \ln\left(1-e^{-\frac{F(N)}{f(j)}s}\right)\right)\right] s ds. \end{aligned} \tag{8.12}$$

Ακολούθως, θέλουμε να εκτιμήσουμε το ολοκλήρωμα της τελευταίας σχέσης. Ξεκινάμε με την παρατήρηση ότι από το Θεώρημα κυριαρχημένης σύγκλισης (Dominated Convergence Theorem), έχουμε (καθώς $f(N)/f'(N)\to\infty$)

$$\lim_N \int_1^\infty\left[1-\exp\left(-\frac{(f(N)/f'(N))^{1-s}}{s\ln(f(N)/f'(N))}\right)\right] s\, ds = 0.$$

Αξιοποιώντας την (8.11) λαμβάνουμε

$$\lim_N \int_1^\infty\left[1-\exp\left(-\int_1^N e^{-\frac{F(N)}{f(x)}s}dx\right)\right] s\, ds = 0. \tag{8.13}$$

Η συνάρτηση f είναι (γνησίως) αύξουσα. Επομένως και η $(-1)/f$ είναι (γνησίως) αύξουσα. Άρα,

$$\begin{aligned} \int_1^N e^{-\frac{F(N)}{f(x)}s}dx &\le \sum_{j=1}^N e^{-\frac{F(N)}{f(j)}s} \\ &\le \int_1^{N+1} e^{-\frac{F(N)}{f(x)}s}dx \\ &\le \int_1^N e^{-\frac{F(N)}{f(x)}s}dx + e^{-\frac{F(N)}{f(N+1)}s}. \end{aligned} \tag{8.14}$$

Από τις ανισότητες αυτές έπεται ότι

$$\begin{aligned} 1-\exp\left(-\int_1^N e^{-\frac{F(N)}{f(x)}s}dx\right) &\le 1-\exp\left(-\sum_{j=1}^N e^{-\frac{F(N)}{f(j)}s}\right) \\ &\le 1-\exp\left(-\int_1^N e^{-\frac{F(N)}{f(x)}s}dx + e^{-\frac{F(N)}{f(N+1)}s}\right). \end{aligned} \tag{8.15}$$

[1] Η ιδέα ήταν να αντικατασταθεί το άθροισμα

$$\sum_{j=1}^N \ln\left[1-e^{-\frac{F(N)}{f(j)}s}\right] \quad \text{με το} \quad -\sum_{j=1}^N e^{-\frac{F(N)}{f(j)}s} \quad \text{και εν συνεχεία,}$$

$$\text{με το ολοκλήρωμα} \quad \int_1^N e^{-\frac{F(N)}{f(x)}s}dx.$$

Το τελευταίο βήμα ήταν να βρεθεί η ασυμπτωτική συμπεριφορά του ολοκληρώματος, υπό τις συνθήκες (8.4).

Ωστόσο, από την (8.11) είναι

$$\lim_N \int_1^N e^{-\frac{F(N)}{f(x)}s} dx = \begin{cases} \infty, & \text{εάν } s < 1; \\ 0, & \text{εάν } s \geq 1. \end{cases} \tag{8.16}$$

Παίροντας τώρα όρια στην (8.15) και χρησιμοποιώντας τις (8.13) και (8.5), έχουμε

$$\lim_N \int_1^\infty \left[1 - \exp\left(\sum_{j=1}^N \ln\left(1 - e^{-\frac{F(N)}{f(j)}s}\right)\right)\right] s\, ds = 0. \tag{8.17}$$

Τέλος, από τον ορισμό της $F(\cdot)$, και το ανάπτυγμα Taylor του λογαρίθμου, δηλαδή, $\ln(1-x) \sim -x$ καθώς $x \to 0$, η (8.12) δίδει

$$Q_N(\alpha) \sim F(N)^2 = f(N)^2 \ln\left(\frac{f(N)}{f'(N)}\right)^2 \quad \text{καθώς } N \to \infty, \tag{8.18}$$

ολοκληρώνοντας την απόδειξη. ■

Παρατήρηση 7. Από το Θεώρημα 8.1.1 και την (5.14) έχουμε

$$E[T_N(T_N+1)] \sim A_N^2 f(N)^2 \ln\left(\frac{f(N)}{f'(N)}\right)^2, \quad N \to \infty. \tag{8.19}$$

Σημειώνουμε ότι από την (5.1) έχουμε

$$A_N f(N) = \frac{1}{p_N} = \frac{1}{\min_{1\leq j\leq N}\{p_j\}}. \tag{8.20}$$

8.2 Βαθύτερη Ασυμπτωτική Συμπεριφορά της Μέσης Τιμής $E[T_N]$

Sometimes it is useful to know how large is your zero.

Έχει αποδειχθεί (βλέπε [12]), ότι

$$E[T_N] \sim A_N f(N) \ln\left(\frac{f(N)}{f'(N)}\right), \qquad N \to \infty. \tag{8.21}$$

Από την (3.21) είναι φανερό ότι οι εκφράσεις (8.19) και (8.21) **ΔΕΝ παρέχουν αρκετή πληροφορία**, ώστε να προκύψει ο πρώτος όρος στο ασυμπτωτικό α-νάπτυγμα της διασποράς $V[T_N]$. Συνεπώς, καθίσταται αναγκαία η βαθύτερη (α-συμπτωτικά) μελέτη των $E[T_N]$ και $E[T_N(T_N+1)]$.
Ξεκινώντας κανείς από την (5.8), γράφει την $E_N(\alpha)$ στην μορφή

$$\begin{aligned} E_N(\alpha) = & F(N)\left\{ 1 - \int_0^1 \exp\left[\sum_{j=1}^N \ln\left(1 - e^{-\frac{F(N)}{f(j)}s}\right)\right] ds \right. \\ & \left. + \int_1^\infty \left[1 - \exp\left\{ \sum_{j=1}^N \ln\left(1 - e^{-\frac{F(N)}{f(j)}s}\right)\right\}\right] ds \right\}. \end{aligned} \tag{8.22}$$

Θέτουμε

$$I_1(N) := \int_0^1 \exp\left[\sum_{j=1}^{N} \ln\left(1 - e^{-\frac{F(N)}{f(j)}s}\right)\right] ds, \tag{8.23}$$

και

$$I_2(N) := \int_1^\infty \left\{1 - \exp\left[\sum_{j=1}^{N} \ln\left(1 - e^{-\frac{F(N)}{f(j)}s}\right)\right]\right\} ds. \tag{8.24}$$

Από τα Θεωρήματα Φραγμένης και Κυριαρχημένης Σύγκλισης (Bounded/Dominated Convergence Theorem) αντίστοιχα, έχουμε (βλέπε [12]),

$$I_1(N) = o(1) \qquad \text{και} \qquad I_2(N) = o(1). \tag{8.25}$$

Για την βαθύτερη ανάλυση των δύο αυτών ολοκληρωμάτων θα χρειαστούμε κατ' αρχάς το ακόλουθο:

Λήμμα 8.2.1. *Θέτουμε*

$$J_m(N) := \int_1^N f(x)^m e^{-\frac{F(N)}{f(x)}s} dx, \quad m \geq 0. \tag{8.26}$$

Τότε, υπό τις (8.3) και (8.8), έχουμε

$$J_m(N) = \frac{f(N)^{m+2}}{sF(N)f'(N)} e^{-\frac{F(N)}{f(N)}s} \left[1 + O\left(\frac{f(N)}{F(N)}\right)\right], \quad N \to \infty \tag{8.27}$$

ομοιόμορφα ως προς $s \in [s_0, \infty)$, για κάθε $s_0 > 0$.

Απόδειξη. Με ολοκλήρωση κατά πράγοντες παίρνουμε

$$\begin{aligned} \int_1^N f(x)^m e^{-\frac{F(N)}{f(x)}s} dx &= \left[\frac{f(x)^{m+2} e^{-\frac{F(N)}{f(x)}s}}{sF(N)f'(x)}\right]_{x=1}^N \\ &- \int_1^N \frac{e^{-\frac{F(N)}{f(x)}s}}{sF(N)} \left[\frac{f(x)^{m+2}}{f'(x)}\right]' dx. \end{aligned}$$

Είναι

$$\begin{aligned} \int_1^N \frac{e^{-\frac{F(N)}{f(x)}s}}{sF(N)} \left[\frac{f(x)^{m+2}}{f'(x)}\right]' dx = \frac{m+2}{s} \int_1^N \frac{f(x)^{m+1}}{F(N)} e^{-\frac{F(N)}{f(x)}s} dx \\ - \frac{1}{s} \int_1^N \frac{f''(x)/f'(x)}{f'(x)/f(x)} \frac{f(x)^{m+1}}{F(N)} e^{-\frac{F(N)}{f(x)}s} dx. \end{aligned} \tag{8.28}$$

Εφόσον η συνάρτηση $f(\cdot)$ είναι (γνησίως) αύξουσα έχουμε

$$\begin{aligned} \int_1^N \frac{f(x)^{m+1}}{F(N)} e^{-\frac{F(N)}{f(x)}s} dx &\leq \frac{f(N)}{F(N)} \int_1^N f(x)^m e^{-\frac{F(N)}{f(x)}s} dx \\ &= \frac{f(N)}{F(N)} J_m(N) \\ &= o\left(J_m(N)\right). \end{aligned}$$

Από τις (8.3) και (8.8) έχουμε επίσης

$$\int_1^N \frac{f''(x)/f'(x)}{f'(x)/f(x)} \frac{f(x)^{m+1}}{F(N)} e^{-\frac{F(N)}{f(x)}s} dx = \left[\int_1^N \frac{f(x)^{m+1}}{F(N)} e^{-\frac{F(N)}{f(x)}s} dx\right] \cdot O(1),$$

το οποίο ολοκληρώνει την απόδειξη. ■

Θα χρειαστούμε επίσης και τον δεύτερο όρο στο ασυμπτωτικό ανάπτυγμα του ολοκληρώματος $J_m(N)$. Έχουμε σχετικά το ακόλουθο:

Λήμμα 8.2.2. *Έστω $J_m(N)$ όπως ορίστηκε στο ως στο Λήμμα 8.2.1. Τότε καθώς $N \to \infty$*

$$J_m(N) = \frac{f(N)^{m+2}}{sF(N)f'(N)} e^{-\frac{F(N)}{f(N)}s} + \omega(N)\, \frac{f(N)^{m+3}}{s^2F(N)^2f'(N)} e^{-\frac{F(N)}{f(N)}s} \left[1 + O\left(\frac{f(N)}{F(N)}\right)\right], \tag{8.29}$$

όπου

$$\omega(N) := -2 + \frac{f''(N)/f'(N)}{f'(N)/f(N)}. \tag{8.30}$$

Η ασυμπτωτική αυτή συμπεριφορά είναι, όπως και στο Λήμμα 8.2.1, ομοιόμορφη ως προς $s \in [s_0, \infty)$, για κάθε $s_0 > 0$.

Απόδειξη. Εφαρμόζουμε ολοκλήρωση κατά παράγοντες στην (8.28) και έχουμε

$$\frac{m+2}{s} \int_1^N \frac{f(x)^{m+1}}{F(N)} e^{-\frac{F(N)}{f(x)}s} dx - \frac{1}{s}\int_1^N \frac{f''(x)/f'(x)}{f'(x)/f(x)} \frac{f(x)^{m+1}}{F(N)} e^{-\frac{F(N)}{f(x)}s} dx$$

$$= (m+2)\left[\frac{f(x)^{m+3} e^{-\frac{F(N)}{f(x)}s}}{s^2F(N)^2f'(x)}\right]_{x=1}^N - (m+2)\int_1^N \frac{e^{-\frac{F(N)}{f(x)}s}}{s^2F(N)^2} \left[\frac{f(x)^{m+3}}{f'(x)}\right]' dx$$

$$-\left[\frac{f(x)^{m+3} e^{-\frac{F(N)}{f(x)}s}}{s^2F(N)^2f'(x)} \frac{f''(x)/f'(x)}{f'(x)/f(x)}\right]_{x=1}^N + \int_1^N \frac{e^{-\frac{F(N)}{f(x)}s}}{s^2F(N)^2}\left[\frac{f(x)^{m+3}}{f'(x)} \frac{f''(x)/f'(x)}{f'(x)/f(x)}\right]' dx.$$

Επιπλέον,

$$\begin{aligned}\int_1^N \frac{e^{-\frac{F(N)}{f(x)}s}}{s^2F(N)^2}\left[\frac{f(x)^{m+3}}{f'(x)}\right]' dx &= (m+3)\int_1^N \frac{f(x)^{m+2}}{s^2F(N)^2} e^{-\frac{F(N)}{f(x)}s} dx \\ &- \int_1^N \frac{f''(x)/f'(x)}{f'(x)/f(x)} \frac{f(x)^{m+2}}{s^2F(N)^2} e^{-\frac{F(N)}{f(x)}s} dx.\end{aligned}$$

Από την μονοτονία της συνάρτησης $f(\cdot)$ και την (iii) της (8.3) λαμβάνουμε

$$\int_1^N \frac{e^{-\frac{F(N)}{f(x)}s}}{s^2F(N)^2}\left[\frac{f(x)^{m+3}}{f'(x)}\right]' dx = O\left(\frac{f(N)^2}{F(N)^2}\, J_m(N)\right) = o\left(J_m(N)\right).$$

Επίσης,

$$\int_1^N \frac{e^{-\frac{F(N)}{f(x)}s}}{s^2F(N)^2}\left[\frac{f(x)^{m+3}}{f'(x)}\,\frac{f''(x)/f'(x)}{f'(x)/f(x)}\right]'dx = \int_1^N \frac{e^{-\frac{F(N)}{f(x)}s}}{s^2F(N)^2}\frac{f''(x)/f'(x)}{f'(x)/f(x)}\left[\frac{f(x)^{m+3}}{f'(x)}\right]'dx$$
$$+\int_1^N \frac{e^{-\frac{F(N)}{f(x)}s}}{s^2F(N)^2}\frac{f(x)^{m+3}}{f'(x)}\left[\frac{f''(x)/f'(x)}{f'(x)/f(x)}\right]'dx,$$

και

$$\int_1^N \frac{e^{-\frac{F(N)}{f(x)}s}}{s^2F(N)^2}\frac{f(x)^{m+3}}{f'(x)}\left[\frac{f''(x)/f'(x)}{f'(x)/f(x)}\right]'dx = \int_1^N \frac{f(x)^2f'''(x)}{\left(f'(x)\right)^3}f(x)^{m+2}\frac{e^{-\frac{F(N)}{f(x)}s}}{s^2F(N)^2}dx$$
$$+\int_1^N \frac{f''(x)/f'(x)}{f'(x)/f(x)}f(x)^{m+2}\frac{e^{-\frac{F(N)}{f(x)}s}}{s^2F(N)^2}dx$$
$$-2\int_1^N \left(\frac{f''(x)/f'(x)}{f'(x)/f(x)}\right)^2 f(x)^{m+2}\frac{e^{-\frac{F(N)}{f(x)}s}}{s^2F(N)^2}dx.$$

Ώστε από την (8.3) είναι

$$\int_1^N \frac{e^{-\frac{F(N)}{f(x)}s}}{s^2F(N)^2}\frac{f(x)^{m+3}}{f'(x)}\left[\frac{f''(x)/f'(x)}{f'(x)/f(x)}\right]'dx = O\left(\frac{f(N)^2}{F(N)^2}\,J_m(N)\right) = o\left(J_m(N)\right).$$

Από την (8.30), δηλαδή:

$$\omega(N) = -2+\frac{f''(N)/f'(N)}{f'(N)/f(N)},$$

λαμβάνουμε

$$J_m(N) = \frac{f(N)^{m+2}}{sF(N)f'(N)}e^{-\frac{F(N)}{f(N)}s}+\omega(N)\,\frac{f(N)^{m+3}}{s^2F(N)^2f'(N)}e^{-\frac{F(N)}{f(N)}s}+o\left(J_m(N)\right),$$

και η απόδειξη ολοκληρώνεται από το Λήμμα 8.2.1, (δηλαδή από τον πρώτο όρο του ασυμπτωτικού αναπτύγματος του ολοκληρώματος $J_m(N)$). ■

Σχόλιο. Ας σημειωθεί ότι από τις συνθήκες (8.3) προκύπτει άμεσα ότι η ποσότητα $\omega(N)$ είναι φραγμένη, δηλαδή, $\omega(N)=O(1)$ καθώς $N\to\infty$.

8.2α′ Το Ολοκλήρωμα $I_1(N)$

Αναφορικά με το ολοκλήρωμα που ορίστηκε στην (8.23), δοθέντος $\varepsilon\in(0,1)$ έχουμε
(8.31)
$$I_1(N) = \int_0^{1-\varepsilon}\exp\left(\sum_{j=1}^N \ln\left(1-e^{-\frac{F(N)}{f(j)}s}\right)\right)ds+\int_{1-\varepsilon}^1\exp\left(\sum_{j=1}^N \ln\left(1-e^{-\frac{F(N)}{f(j)}s}\right)\right)ds.$$

Για το πρώτο ολοκλήρωμα της (8.31) ισχύουν διαδοχικά

$$
\begin{aligned}
I_{11}(N) &:= \int_0^{1-\varepsilon} \exp\left(\sum_{j=1}^{N} \ln\left(1 - e^{-\frac{F(N)}{f(j)}s}\right)\right) ds \\
&< (1-\varepsilon)\exp\left(\sum_{j=1}^{N} \ln\left(1 - e^{-\frac{F(N)}{f(j)}(1-\varepsilon)}\right)\right) \\
&< \exp\left(\sum_{j=1}^{N} \ln\left(1 - e^{-\frac{F(N)}{f(j)}(1-\varepsilon)}\right)\right) \\
&< \exp\left(-\sum_{j=1}^{N} e^{-\frac{F(N)}{f(j)}(1-\varepsilon)}\right),
\end{aligned}
$$

καθώς, $\ln(1-x) < -x$, για $0 < x < 1$.
Από την (8.14) (δηλαδή την σύγκριση αθροίσματος – ολοκληρώματος), έχουμε

$$
I_{11}(N) < \exp\left(-\int_1^N e^{-\frac{F(N)}{f(x)}(1-\varepsilon)} dx\right).
$$

Εφαρμόζουμε τώρα το Λήμμα 8.2.1, για $m = 0$, και παίρνουμε

$$
I_{11}(N) < \exp\left\{-\frac{f(N)^2}{(1-\varepsilon)F(N)f'(N)} e^{-\frac{F(N)}{f(N)}(1-\varepsilon)} \left(1 + M_1 \frac{f(N)}{F(N)}\right)\right\}, \tag{8.32}
$$

όπου M_1 είναι μία θετική σταθερά. Από την (8.8), δηλαδή τον ορισμό της $F(\cdot)$, έχουμε

$$
I_{11}(N) < \exp\left\{-\frac{1}{1-\varepsilon} \frac{\left(\frac{f(N)}{f'(N)}\right)^{\varepsilon}}{\ln\left(\frac{f(N)}{f'(N)}\right)} \left(1 + M_1 \frac{1}{\ln\left(\frac{f(N)}{f'(N)}\right)}\right)\right\}. \tag{8.33}
$$

Εφόσον, $\frac{f'(x)}{f(x)} \to 0$ και $\varepsilon \in (0,1)$ καταλήγουμε στην εκτίμηση

$$
I_{11}(N) < \left[\frac{1}{\ln\left(\frac{f(N)}{f'(N)}\right)}\right]^3, \tag{8.34}
$$

για αρκούντως μεγάλες τιμές του N.

Το επόμενο μας μέλημα είναι να υπολογίσουμε μερικούς όρους του ασυμπτωτικού αναπτύγματος του δεύτερου ολοκληρωματος της σχέσης (8.31).

Βήμα 1. *Γράφουμε το ολοκλήρωμα I_{12} σε μία μορφή πιο εύκολα διαχειρίσιμη.*

Θέτουμε

$$
B(N;s) := \sum_{j=1}^{N} \ln\left(1 - e^{-\frac{F(N)}{f(j)}s}\right). \tag{8.35}
$$

Επειδή

$$\frac{F(N)}{f(j)} \to \infty \qquad \text{καθώς} \quad N\to\infty, \tag{8.36}$$

και

$$\ln(1-x) = -x + O(x^2) \qquad \text{καθώς} \quad x \to 0,$$

έχουμε (για όσο τουλάχιστον είναι, $s \geq s_0 > 0$)

$$B(N;s) = \sum_{j=1}^{N}\left[-e^{-\frac{F(N)}{f(j)}s} + O\left(e^{-\frac{2F(N)}{f(j)}s}\right)\right]. \tag{8.37}$$

Από την σύγκριση σειράς και ολοκληρώματος, δηλαδή την (8.14), η σχέση (8.37) παίρνει την μορφή

$$B(N;s) = -\left[\int_1^N e^{-\frac{F(N)}{f(x)}s}dx + O\left(e^{-\frac{F(N)}{f(N+1)}s}\right)\right] + \sum_{j=1}^{N} O\left(e^{-\frac{2F(N)}{f(j)}s}\right).$$

Η τελευταία σε συνδυασμό με το Λήμμα 8.2.2, για $m=0$, γίνεται

$$\begin{aligned} B(N;s) = &- \frac{f(N)^2}{sF(N)f'(N)}e^{-\frac{F(N)}{f(N)}s} \\ &- \omega(N)\,\frac{f(N)^3}{s^2F(N)^2f'(N)}e^{-\frac{F(N)}{f(N)}s}\left[1+O\left(\frac{f(N)}{F(N)}\right)\right] \\ &+ O\left(e^{-\frac{F(N)}{f(N+1)}s}\right) + O\left(Ne^{-\frac{2F(N)}{f(N)}s}\right). \end{aligned} \tag{8.38}$$

Χρησιμοποιώντας την (8.5) (βλέπε Παρατήρηση 6), από την (8.38) λαμβάνουμε

$$\begin{aligned} B(N;s) = &- \frac{f(N)^2}{sF(N)f'(N)}e^{-\frac{F(N)}{f(N)}s} \\ &- \omega(N)\,\frac{f(N)^3}{s^2F(N)^2f'(N)}e^{-\frac{F(N)}{f(N)}s}\left[1+O\left(\frac{f(N)}{F(N)}\right)\right]. \end{aligned} \tag{8.39}$$

Επομένως,

$$\begin{aligned} I_{12}(N) :=& \int_{1-\varepsilon}^{1} e^{B(N;s)}ds \\ =& \int_{1-\varepsilon}^{1} \exp\left[-\frac{f(N)^2}{sF(N)f'(N)}e^{-\frac{F(N)}{f(N)}s}\right. \\ &\left. -\omega(N)\,\frac{f(N)^3}{s^2F(N)^2f'(N)}e^{-\frac{F(N)}{f(N)}s}\left[1+O\left(\frac{f(N)}{F(N)}\right)\right]\right] ds \end{aligned} \tag{8.40}$$

καθώς $N\to\infty$. Από τον ορισμό της $F(\cdot)$ και την αντικατάσταση $s=1-t$, έχουμε

$$I_{12}(N) = \int_0^{\varepsilon} \exp\left[-\frac{1}{1-t}\,\frac{\left(\frac{f(N)}{f'(N)}\right)^t}{\ln\left(\frac{f(N)}{f'(N)}\right)} - \omega(N)\frac{1}{(1-t)^2}\,\frac{\left(\frac{f(N)}{f'(N)}\right)^t}{\ln^2\left(\frac{f(N)}{f'(N)}\right)}\right.$$

$$\times \left[1 + O\left(\frac{1}{\ln\left(\frac{f(N)}{f'(N)}\right)}\right)\right]\Bigg] dt.$$

Χάριν ευκολίας εισάγουμε τον ακόλουθο συμβολισμό

$$A := \frac{f(N)}{f'(N)} \tag{8.41}$$

(σημειώνουμε ότι $A \to \infty$ καθώς $N \to \infty$). Επομένως,

$$I_{12} = \int_0^{\varepsilon} \exp\left[-\frac{1}{1-t}\,\frac{A^t}{\ln A} - \omega(N)\frac{1}{(1-t)^2}\,\frac{A^t}{\ln^2 A}\left[1 + O\left(\frac{1}{\ln A}\right)\right]\right] dt \tag{8.42}$$

ή σε μία πιο συμπαγή μορφή:

(8.43)

$$I_{12} = \int_0^{\varepsilon} \exp\left[-\,\frac{A^t}{\ln A}\left(\sum_{n=0}^{\infty} t^n\right) - \omega(N)\frac{A^t}{\ln^2 A}\,\left(\sum_{n=1}^{\infty} n\, t^{n-1}\right)\left[1 + O\left(\frac{1}{\ln A}\right)\right]\right] dt.$$

Η αντικατάσταση

$$u = \frac{A^t}{\ln A}$$

στο ανωτέρω ολοκλήρωμα, αποδίδει

$$I_{12} = \int_{1/\ln A}^{A^{\varepsilon}/\ln A} \exp\left[-\frac{u}{1 - \frac{\ln u}{\ln A} - \frac{\ln(\ln A)}{\ln A}} - \frac{\omega(N)}{\ln A}\,\frac{u}{\left(1 - \frac{\ln u}{\ln A} - \frac{\ln(\ln A)}{\ln A}\right)^2}\right.$$
$$\left.\times \left[1 + O\left(\frac{1}{\ln A}\right)\right]\right] \frac{du}{u \ln A}.$$

Εάν,

$$\delta := \frac{1}{\ln A} = \frac{1}{\ln\left(\frac{f(N)}{f'(N)}\right)} = \frac{f(N)}{F(N)} \tag{8.44}$$

(επομένως $A \to \infty$ συνεπάγεται ότι $\delta \to 0^+$) το ανωτέρω ολοκλήρωμα λαμβάνει την ακόλουθη πιο διαχειρίσιμη μορφή

$$I_{12} = \delta \int_{\delta}^{\delta \exp(\varepsilon/\delta)} \exp\left[-\frac{u}{1 - \delta \ln u + \delta \ln \delta} - \frac{\omega(N) u \delta}{(1 - \delta \ln u + \delta \ln \delta)^2}\,(1 + O(\delta))\right] \frac{du}{u}.$$

Βήμα 2. *Αναλύουμε το ολοκλήρωμα I_{12} παρεμβάλλοντας τον όρο $1/\sqrt{\delta}$.*

Συνεχίζουμε με την παρατήρηση ότι

$$1/\sqrt{\delta} = o\left(\delta e^{\varepsilon/\delta}\right), \quad \delta \to 0^+.$$

Έχουμε

$$I_{12} = \delta \int_{\delta}^{1/\sqrt{\delta}} \exp\left[-\frac{u}{1-\delta\ln u + \delta \ln\delta} - \frac{\omega(N)u\delta}{(1-\delta\ln u+\delta\ln\delta)^2}\left(1+O\left(\delta\right)\right)\right]\frac{du}{u} +$$

(8.45)

$$+\delta \int_{1/\sqrt{\delta}}^{\delta\exp(\varepsilon/\delta)} \exp\left[-\frac{u}{1-\delta\ln u + \delta \ln\delta} - \frac{\omega(N)u\delta}{(1-\delta\ln u+\delta\ln\delta)^2}\left(1+O\left(\delta\right)\right)\right]\frac{du}{u}.$$

Βήμα 3. *Εκτιμούμε το δεύτερο ολοκλήρωμα του Βήματος 2.*

Αρχικά εργαζόμαστε με το δεύτερο ολοκλήρωμα της σχέσης (8.45) και λαμβάνουμε ένα άνω φράγμα ως ακολούθως:

$$\int_{1/\sqrt{\delta}}^{\delta\exp(\varepsilon/\delta)} \exp\left[-\frac{u}{1-\delta\ln u + \delta \ln\delta} - \frac{\omega(N)u\delta}{(1-\delta\ln u+\delta\ln\delta)^2}\ \left(1+O\left(\delta\right)\right)\right]\frac{du}{u}$$

$$=\int_{1/\sqrt{\delta}}^{\delta\exp(\varepsilon/\delta)} \exp\left[-\frac{u}{1-\delta\ln u + \delta \ln\delta}\left[1+ \frac{\omega(N)}{1-\delta\ln u+\delta\ln\delta}\,\delta\left(1+O\left(\delta\right)\right)\right]\right]\frac{du}{u}$$

$$\leq \int_{1/\sqrt{\delta}}^{\delta\exp(\varepsilon/\delta)} \exp\left[-\frac{u\left(1+O\left(\delta\right)\right)}{1-\delta\ln(1/\sqrt{\delta}) + \delta \ln\delta}\right]\frac{du}{1/\sqrt{\delta}}$$

$$=\sqrt{\delta}\int_{1/\sqrt{\delta}}^{\delta\exp(\varepsilon/\delta)} \exp\left(-\frac{u\left(1+O\left(\delta\right)\right)}{1+\frac{3}{2}\delta\ln\delta}\right)du$$

$$=O\left(\sqrt{\delta}\int_{1/\sqrt{\delta}}^{\infty} e^{-u\ (1+O(\delta))}du\right)$$

(8.46)

$$=O\left(\sqrt{\delta}\ e^{-1/\sqrt{\delta}}\right).$$

Βήμα 4. *Εργαζόμαστε με το πρώτο ολοκλήρωμα του Βήματος 2, $K_1(\delta)$. Εφαρμόζουμε ανάπτυγμα κατά Taylor.*

Το πρώτο ολοκλήρωμα της σχέσης (8.45) είναι

(8.47)

$$K_1(\delta) := \int_{\delta}^{1/\sqrt{\delta}} \exp\left[-\frac{u}{1-\delta\ln u + \delta \ln\delta} - \frac{\omega(N)u\delta}{(1-\delta\ln u+\delta\ln\delta)^2}\left(1+O\left(\delta\right)\right)\right]\frac{du}{u}.$$

Καθόσον,

$$(1-x)^{-2} = \sum_{n=1}^{\infty} n\, x^{n-1} \qquad \text{για} \ \ |x|<1,$$

η σχέση (8.47) γίνεται

$$K_1(\delta) = \int_{\delta}^{1/\sqrt{\delta}} \exp\left[-u\sum_{n=0}^{\infty}\left[\delta\ln\left(\frac{u}{\delta}\right)\right]^n - u\omega(N)\delta\left(1+O\left(\delta\right)\right)\sum_{n=1}^{\infty} n\left[\delta\ln\left(\frac{u}{\delta}\right)\right]^{n-1}\right]\frac{du}{u}$$

$$= \int_{\delta}^{1/\sqrt{\delta}} \frac{e^{-u}}{u}\ \exp\left(-u\sum_{n=1}^{\infty}\left[\delta\ln\left(\frac{u}{\delta}\right)\right]^n\right)$$

$$\times\exp\left(-u\,\omega(N)\,\delta\left(1+O\left(\delta\right)\right)\sum_{n=1}^{\infty} n\left[\delta\ln\left(\frac{u}{\delta}\right)\right]^{n-1}\right) du. \tag{8.48}$$

Εν συνεχεία αναπτύσουμε τους εκθετικούς όρους και παίρνουμε από την (8.48)

$$K_1(\delta) = \int_{\delta}^{1/\sqrt{\delta}} \frac{e^{-u}}{u}\ \left\{1 - u\sum_{n=1}^{\infty}\left(\delta\ln\frac{u}{\delta}\right)^n + O\left(u\sum_{n=1}^{\infty}\left(\delta\ln\frac{u}{\delta}\right)^n\right)^2\right\}$$

$$\times\left\{1-\omega(N)\ u\ \delta\left(1+O\left(\delta\right)\right)\sum_{n=1}^{\infty}\left(\delta\ln\frac{u}{\delta}\right)^{n-1}\right.$$

$$\left.+O\left(\omega(N)\ u\ \delta\left(1+O\left(\delta\right)\right)\sum_{n=1}^{\infty}\left(\delta\ln\frac{u}{\delta}\right)^{n-1}\right)^2\right\} du,$$

(διότι $e^x = 1 + x + O(x^2)$ καθώς $x \to 0$). Ώστε,[2]

Βήμα 5. *Αναλύουμε το ολοκλήρωμα $K_1(\delta)$ παρεμβάλλοντας το ∞.*

$$K_1(\delta) = \int_{\delta}^{1/\sqrt{\delta}} \frac{e^{-u}}{u}\left[1 - u\left(\delta\ln\frac{u}{\delta} + \omega(N)\delta\left(1+O\left(\delta\right)\right)\right) + u^2 O\left(\delta^2\ln^2\frac{u}{\delta}\right)\right] du$$

$$= \int_{\delta}^{\infty} \frac{e^{-u}}{u}\left[1 - u\left(\delta\ln\frac{u}{\delta} + \omega(N)\delta\left(1+O\left(\delta\right)\right)\right) + u^2 O\left(\delta^2\ln^2\frac{u}{\delta}\right)\right] du$$

$$- \int_{1/\sqrt{\delta}}^{\infty} \frac{e^{-u}}{u}\left[1 - u\left(\delta\ln\frac{u}{\delta} + \omega(N)\delta\left(1+O\left(\delta\right)\right)\right) + u^2 O\left(\delta^2\ln^2\frac{u}{\delta}\right)\right] du. \tag{8.49}$$

Βήμα 6. *Εκτιμούμε το δεύτερο ολοκλήρωμα του Βήματος 5.*

[2] Από την πρώτη αγκύλη χρειαζόμαστε μόνο τον πρώτο όρο της σειράς $\sum_{n=1}^{\infty}\left(\delta\ln\frac{u}{\delta}\right)^n$. Από την δεύτερη αγκύλη, επίσης, μόνον ο πρώτος όρος της σειράς $\sum_{n=1}^{\infty}\left(\delta\ln\frac{u}{\delta}\right)^{n-1}$ αρκεί για τους υπολογισμούς μας.

Ωστόσο,

$$\int_{1/\sqrt{\delta}}^{\infty}\frac{e^{-u}}{u}\left[1-u\left(\delta\ln\frac{u}{\delta}+\omega(N)\delta\left(1+O\left(\delta\right)\right)\right)+u^2O\left(\delta^2\ln^2\frac{u}{\delta}\right)\right]du$$

$$\leq\int_{1/\sqrt{\delta}}^{\infty}\frac{e^{-u}}{1/\sqrt{\delta}}\left[1-\frac{1}{\sqrt{\delta}}\,\delta\ln\left(\frac{1/\sqrt{\delta}}{\delta}\right)-\ \omega(N)\delta\left(1+O\left(\delta\right)\right)\right]du$$

$$+\int_{1/\sqrt{\delta}}^{\infty}ue^{-u}O\left(\delta^2\ln^2\frac{u}{\delta}\right)du$$

$$=\int_{1/\sqrt{\delta}}^{\infty}\sqrt{\delta}\,e^{-u}\left(1+\frac{3}{2}\sqrt{\delta}\ln\delta-\omega(N)\delta\left(1+O\left(\delta\right)\right)\right)du+O\left(e^{-1/\sqrt{\delta}}\delta^{\frac{3}{2}}\ln^2\delta\right)$$

$$=\sqrt{\delta}\,e^{-1/\sqrt{\delta}}\left(1+\frac{3}{2}\sqrt{\delta}\ln\delta-\omega(N)\delta\left(1+O\left(\delta\right)\right)\right)+O\left(e^{-1/\sqrt{\delta}}\delta^{\frac{3}{2}}\ln^2\delta\right)$$

$$=O\left(\sqrt{\delta}e^{-1/\sqrt{\delta}}\right)+O\left(e^{-1/\sqrt{\delta}}\delta^{\frac{3}{2}}\ln^2\delta\right)=O\left(\sqrt{\delta}e^{-1/\sqrt{\delta}}\right)\tag{8.50}$$

καθώς $\delta\to 0^+$.

Βήμα 7. *Goal ! Γράφουμε το ολοκήρωμα $K_1(\delta)$ με άνω άκρο ολοκλήρωσης το ∞.*

Συνεπώς, στην ολοκληρωτική έκφραση του $K_1(\delta)$ μπορούμε να αντικαταστήσουμε το άνω άκρο ολοκλήρωσης με το ∞. Έχουμε λοιπόν,
(8.51)

$$K_1(\delta)=\int_{\delta}^{\infty}\frac{e^{-u}}{u}\left[1-u\left(\delta\ln\frac{u}{\delta}+\omega(N)\delta\left(1+O\left(\delta\right)\right)\right)+u^2O\left(\delta^2\ln^2\frac{u}{\delta}\right)\right]du.$$

Βήμα 8. *Ασυμπτωτική συμπεριφορά του ολοκληρώματος I_{12} καθώς $\delta\to 0^+$.*

Από την (8.51), το ολοκλήρωμα $I_{12}(N)$ της (8.45) γίνεται
(8.52)

$$I_{12}(N)=\delta\int_{\delta}^{\infty}\frac{e^{-u}}{u}\left[1-u\left(\delta\ln\frac{u}{\delta}+\omega(N)\delta\left(1+O\left(\delta\right)\right)\right)+u^2O\left(\delta^2\ln^2\frac{u}{\delta}\right)\right]du,$$

με πολύ μικρά λάθη, καθώς $\delta\to 0^+$ (βλέπε (8.46) και (8.50)). Σημειώνουμε ότι η σχέση (8.52) μας λέει ότι μπορούμε να αντικαταστήσουμε την πρώτη από τις σειρές που εμφανίζονται στην (8.43), με τους δύο πρώτους όρους του αναπτύγματος Taylor γύρω από το $t=0$. Ομοίως, η δεύτερη σειρά που εμφανίζεται στην (8.43), μπορεί τελικά να αντικατασταθεί από τον πρώτο όρο του αντίστοιχου αναπτύγματος Taylor.

Για την συνέχεια της ανάλυσής μας θα χρειαστούμε τα ακόλουθα λήμματα :

Λήμμα 8.2.3. *Το εκθετικό ολοκλήρωμα (exponential integral),*

$$E(x):=\int_x^{\infty}\frac{e^{-t}}{t}dt,$$

έχει το ασυμπτωτικό ανάπτυγμα

$$E(x)\sim-\ln x-\gamma+x-\frac{1}{4}x^2+\frac{1}{18}x^3-\cdots\tag{8.53}$$

καθώς $x \to 0^+$, *($\gamma = 0.5772...$, είναι η σταθερά του Euler).*

Απόδειξη. (Βλέπε [9], σελ. 252). ■

Λήμμα 8.2.4. *Για το ολοκλήρωμα,*

$$G(x) := \int_x^\infty \ln t \, e^{-t} dt,$$

έχουμε το ασυμπτωτικό ανάπτυγμα

$$G(x) \sim -\gamma - x\ln x + x + \frac{1}{2}x^2 \ln x - \frac{1}{2}x^2 - \frac{1}{6}x^3 \ln x + \frac{1}{6}x^3 + \frac{1}{24}x^4 \ln x - \frac{1}{24}x^4 + \cdots \tag{8.54}$$

καθώς $x \to 0^+$.

Απόδειξη. Επειδή

$$\frac{dG(x)}{dx} = -\ln x \, e^{-x} = -\ln x \left(1 - x + \frac{1}{2}x^2 - \frac{1}{6}x^3 + \cdots\right),$$

έχουμε

$$G(x) \sim C - x\ln x + x + \frac{1}{2}x^2 \ln x - \frac{1}{2}x^2 - \frac{1}{6}x^3 \ln x + \frac{1}{6}x^3 + \frac{1}{24}x^4 \ln x - \frac{1}{24}x^4 + \cdots, \tag{8.55}$$

όπου C είναι μία σταθερά. Για να ολοκληρωθεί η απόδειξη αρκεί να υπολογίσουμε την τιμή της C. Από την σχέση (8.55) έχουμε

$$C = \lim_{x\to 0^+} \left(\int_x^\infty \ln t \, e^{-t} dt\right) = \int_0^\infty \ln t \, e^{-t} dt. \tag{8.56}$$

Το τελευταίο ολοκλήρωμα ισούται με $\Gamma'(1)$. Επομένως, $C = -\gamma$ και έχουμε το ζητούμενο (βλέπε επίσης την $(A.3)$). ■

Από τα Λήμματα 8.2.3, 8.2.4, και την (8.52,) έχουμε

$$I_{12} = \delta\left\{\left[-\ln\delta - \gamma + \delta + O(\delta^2)\right] - \delta\left[-\gamma + O(\delta\ln\delta)\right] \right.$$
$$\left. + \delta\ln\delta \, e^{-\delta} - \omega(N)\delta\left(1 + O(\delta)\right)e^{-\delta}\right\}$$
$$+ O\left\{\delta^3\left(\int_\delta^\infty ue^{-u}\ln^2 u \, du + \ln^2\delta \int_\delta^\infty u \, e^{-u} du - 2\ln\delta \int_\delta^\infty ue^{-u} \, du\right)\right\},$$

ή

$$\begin{aligned} I_{12} = &-\delta\ln\delta - \gamma\delta + \delta^2\ln\delta \, e^{-\delta} + (1+\gamma)\delta^2 - \delta^2\omega(N) \, e^{-\delta}\left(1 + O(\delta)\right) \\ &+ O(\delta^3) + O\left(\delta^2\ln\delta\right) \\ &+ O\left(\delta^3 \int_\delta^\infty ue^{-u}\ln^2 u du + \delta^3\ln^2\delta \int_\delta^\infty ue^{-u} du - 2\delta^3\ln\delta \int_\delta^\infty ue^{-u} du\right). \end{aligned} \tag{8.57}$$

Με παρόμοιους ισχυρισμούς όπως στα Λήμματα 8.2.3 και 8.2.4 έχουμε ότι καθώς $x \to 0^+$,

$$\int_x^\infty t\, e^{-t} \ln^2 t\, dt \sim \; C_0 = \lim_{x\to 0^+} \left(\int_x^\infty t\, e^{-t} \ln^2 t\, dt \right)$$
$$= \int_0^\infty t\, e^{-t} \ln^2 t\, dt = \left(\frac{\pi^2}{6} + \gamma^2 - 2\gamma \right). \tag{8.58}$$

Πράγματι, με ολοκλήρωση κατά παράγοντες έχει κανείς

$$\int_0^\infty t \ln^2 t\, e^{-t} dt = \int_0^\infty \ln^2 t\, e^{-t} dt + 2 \int_0^\infty \ln t\, e^{-t} dt. \tag{8.59}$$

Παρατηρούμε τώρα ότι $\int_0^\infty \ln^2 t\, e^{-t} dt = \Gamma''(1)$. Από την (Α΄.4) είναι

$$\int_0^\infty \ln^2 t\, e^{-t} dt = \frac{\pi^2}{6} + \gamma^2. \tag{8.60}$$

Συνδυάζοντας τις (8.56) και (8.60) καταλήγουμε στην ακριβή τιμή της σταθεράς C_0. Επίσης από το ανάπτυγμα Taylor της εκθετικής συνάρτησης έχουμε

$$e^{-\delta} = 1 + O(\delta), \qquad \delta \to 0^+.$$

Επιπλέον

$$\int_\delta^\infty u\, e^{-u}\, du = (1+\delta)\, e^{-\delta} = (1+\delta)\,(1+O(\delta))$$
$$= 1 + \delta + O(\delta) + O(\delta^2) = 1 + O(\delta). \tag{8.61}$$

Από τις (8.58) και (8.61) η σχέση (8.57) γίνεται τελικά

$$I_{12} = -\delta \ln \delta - \gamma\, \delta + \delta^2 \ln \delta + [\, 1 + \gamma - \omega(N)\,]\, \delta^2 + O\left(\delta^3 \ln^2 \delta\right). \tag{8.62}$$

Στο σημείο αυτό σημειώνουμε ότι το εκτιμώμενο σφάλμα στην (8.62) είναι ισχυρότερο από τα αντίστοιχα σφάλματα των σχέσεων (8.46) και (8.50).

8.2β΄ Το Ολοκλήρωμα $I_2(N)$

Το τελευταίο βήμα για να επιτευχθεί ο στόχος της βαθύτερης μελέτης επί της ασυμπτωτικής συμπεριφοράς της μέσης τιμής $E[\,T_N\,]$, είναι να υπολογίσουμε τον πρώτο όρο στο ασυμπτωτικό ανάπτυγμα του ολοκληρώματος $I_2(N)$. (Όπως θα καταστεί ξεκάθαρο παρακάτω, ο πρώτος όρος του $I_2(N)$, θα είναι αρκετός για την εύρεση του πρώτου όρου της διασποράς $V[\,T_N\,]$). Προς αυτήν την κατεύθυνση θα ακολουθήσουμε διαφορετική προσέγγιση. Ξέρουμε ότι

$$\frac{\ln(1-x)}{x} \to -1, \qquad x \to 0.$$

Επομένως, δοθέντος $\vartheta \in (0,1)$, υπάρχει $\eta = \eta(\vartheta)$ τέτοιο ώστε για $|x| < \eta$, να ισχύει

$$-1 - \vartheta < \frac{\ln(1-x)}{x} < -1 + \vartheta.$$

Εάν περιοριστούμε σε θετικές τιμές του x, έχουμε

$$-(1+\vartheta)\,x < \ln(1-x) < -(1-\vartheta)\,x. \tag{8.63}$$

Ξέρουμε επίσης ότι

$$\frac{1-e^{-x}}{x} \to 1, \qquad x \to 0.$$

Ομοίως, για θετικά x και δοθέντος $\vartheta \in (0,1)$, υπάρχει $\eta = \eta(\vartheta)$ εις τρόπον ώστε, για $0 < x < \eta$, να ισχύει

$$(1-\vartheta)\,x < 1-e^{-x} < (1+\vartheta)\,x. \tag{8.64}$$

Για $j = 1, \dots, N$, $s \geq 1$, χρησιμοποιούμε τον ορισμό της $F(\cdot)$, καθώς επίσης και το γεγονός ότι η $f(\cdot)$ είναι αύξουσα, και παίρνουμε

$$0 < x = e^{-\frac{F(N)}{f(j)}s} \leq e^{-\frac{F(N)}{f(N)}s} \leq e^{-\frac{F(N)}{f(N)}} = \frac{f'(N)}{f(N)} \to 0, \qquad N \to \infty.$$

Επομένως, δοθέντος $\vartheta \in (0,1)$, υπάρχει $N_0 = N_0(\vartheta)$ τέτοιο ώστε, για $N \geq N_0$, από την σχέση (8.63) προκύπτει

$$-(1+\vartheta)\,e^{-\frac{F(N)}{f(j)}s} < \ln\left(1 - e^{-\frac{F(N)}{f(j)}s}\right) < -(1-\vartheta)\,e^{-\frac{F(N)}{f(j)}s}, \quad j = 1, \dots, N.$$

Αθροίζουμε επί των τιμών του j και εισάγοντας την (8.35) έχουμε

$$-(1+\vartheta)\sum_{j=1}^{N} e^{-\frac{F(N)}{f(j)}s} < B(s;N) < -(1-\vartheta)\sum_{j=1}^{N} e^{-\frac{F(N)}{f(j)}s}. \tag{8.65}$$

Από την σύγκριση σειράς και ολοκληρώματος (δηλαδή την σχέση (8.14)), και για δοθέν $\vartheta \in (0,1)$, είναι

$$\vartheta\sum_{j=1}^{N} e^{-\frac{F(N)}{f(j)}s} \leq \vartheta\, e^{-\frac{F(N)}{f(N+1)}s} + \vartheta\int_1^N e^{-\frac{F(N)}{f(x)}s}dx, \tag{8.66}$$

$$-\sum_{j=1}^{N} e^{-\frac{F(N)}{f(j)}s} \leq -\int_1^N e^{-\frac{F(N)}{f(x)}s}dx. \tag{8.67}$$

Προσθέτουμε τις ανισότητες (8.66) και (8.67) λαμβάνοντας

$$-(1-\vartheta)\sum_{j=1}^{N} e^{-\frac{F(N)}{f(j)}s} \leq \vartheta\, e^{-\frac{F(N)}{f(N+1)}s} - (1-\vartheta)\int_1^N e^{-\frac{F(N)}{f(x)}s}dx. \tag{8.68}$$

Από την (8.14) έχουμε επίσης

$$-(1+\vartheta)\left[e^{-\frac{F(N)}{f(N+1)}s} + \int_1^N e^{-\frac{F(N)}{f(x)}s}dx\right] \leq -(1+\vartheta)\sum_{j=1}^{N} e^{-\frac{F(N)}{f(j)}s}. \tag{8.69}$$

Χρησιμοποιούμε τις ανισότητες (8.68) και (8.69) στην (8.65) και έχουμε

$$\begin{aligned} & -(1+\vartheta)\left[e^{-\frac{F(N)}{f(N+1)}s} + \int_1^N e^{-\frac{F(N)}{f(x)}s}dx\right] \\ & < B(s;N) \\ & < \vartheta\, e^{-\frac{F(N)}{f(N+1)}s} - (1-\vartheta)\int_1^N e^{-\frac{F(N)}{f(x)}s}dx. \end{aligned} \tag{8.70}$$

Από την (8.25) έχουμε $B(s;N) \to 0$, ομοιόμορφα ως προς $s \in [1,\infty)$ καθώς $N \to \infty$. Επομένως, δοθέντος $\vartheta > 0$, υπάρχει $N_1 = N_1(\vartheta)$, εις τρόπον ώστε, για $N \geq N_1$, η ανισότητα (8.64) αποδίδει

$$-(1-\vartheta)\,B(s;N) < 1 - e^{B(s;N)} < -(1+\vartheta)\,B(s;N).$$

Ώστε, (βλέπε τις σχέσεις (8.24) και (8.35)),

$$-(1-\vartheta)\int_1^\infty B(s;N)ds < I_2(N) < -(1+\vartheta)\int_1^\infty B(s;N)ds. \tag{8.71}$$

Χρησιμοποιώντας τα φράγματα που δόθηκαν στην (8.70) για την ποσότητα $B(s;N)$, και εισαγάγοντάς τα στην (8.71), καταλήγουμε ότι για όλα τα $N \geq N_2 = \max\{N_0, N_1\}$ είναι

$$\begin{aligned} & (1-\vartheta)^2\int_1^\infty\int_1^N e^{-\frac{F(N)}{f(x)}s}dx\,ds - \vartheta\,(1-\vartheta)\int_1^\infty e^{-\frac{F(N)}{f(N+1)}s}ds \\ & < I_2(N) \\ & < (1+\vartheta)^2\int_1^\infty\int_1^N e^{-\frac{F(N)}{f(x)}s}dx\,ds + (1+\vartheta)^2\int_1^\infty e^{-\frac{F(N)}{f(N+1)}s}ds. \end{aligned} \tag{8.72}$$

Εν συνεχεία, υπολογίζουμε το

$$\begin{aligned} \int_1^\infty\int_1^N e^{-\frac{F(N)}{f(x)}s}dx\,ds &= \frac{1}{F(N)}\int_1^N f(x)e^{-\frac{F(N)}{f(x)}}dx \\ &= \left(\frac{f(N)}{F(N)}\right)^2\left[1 + O\left(\frac{f(N)}{F(N)}\right)\right], \end{aligned}$$

όπου εφαρμόσαμε εναλλαγή της σειράς ολοκλήρωσης και ακολούθως, το Λήμμα 8.2.1, για $m = 1$.
Επιπλέον, προκύπτει άμεσα ότι

$$\int_1^\infty e^{-\frac{F(N)}{f(N+1)}s}\,ds = \frac{f(N+1)}{F(N)}\,e^{-\frac{F(N)}{f(N+1)}} = o\left(\frac{f(N)^2}{F(N)^2}\right). \tag{8.73}$$

Πράγματι, χρησιμοποιώντας τον ορισμό της $F(\cdot)$, (βλέπε (8.8)), έχουμε

$$\begin{aligned} & \frac{f(N+1)}{F(N)}\,e^{-\frac{F(N)}{f(N+1)}}\left(\frac{F(N)}{f(N)}\right)^2 \\ & = -\frac{f(N+1)}{f(N)}\,\ln\left(\frac{f'(N)}{f(N)}\right)e^{\frac{f(N)}{f(N+1)}\ln\left(\frac{f'(N)}{f(N)}\right)}. \end{aligned} \tag{8.74}$$

Το ζητούμενο έπεται τώρα από τις (8.3) και (8.5).
Εφόσον το $\vartheta \in (0,1)$ είναι αυθαίρετο, η (8.72) αποδίδει[3]

$$I_2(N) = \left(\frac{f(N)}{F(N)}\right)^2 \left[1 + O\left(\frac{f(N)}{F(N)}\right)\right] \quad \text{καθώς} \quad N \to \infty. \tag{8.75}$$

Από τον ορισμό της $F(\cdot)$, αλλά και την (8.44) καταλήγουμε

$$I_2(N) = \delta^2 \left(1 + O(\delta)\right) \quad \text{καθώς} \quad \delta \to 0^+. \tag{8.76}$$

Είμαστε πλέον σε θέση να παρουσιάσουμε το ακόλουθο αποτέλεσμα.

Θεώρημα 8.2.5. *Έστω δ όπως έχει οριστεί στην (8.44) (επομένως, $\delta \to 0^+$ καθώς $N \to \infty$) και $\omega(N)$ ως έχει τεθεί στην (8.30). Τότε (γ είναι, ως συνήθως, η σταθερά του Euler)*

$$E_N(\alpha) = f(N)\left[\frac{1}{\delta} + \ln\delta + \gamma - \delta\ln\delta + (\omega(N) - \gamma)\,\delta + O\left(\delta^2 \ln^2 \delta\right)\right] \tag{8.77}$$

Απόδειξη. Το αποτέλεσμα αυτό προκύπτει άμεσα από τον συνδυασμό των σχέσεων (8.23), (8.24), (8.34), (8.62), και (8.76), με την σχέση (8.22). ∎

Παρατήρηση 8. Χρησιμοποιώντας το Θεώρημα 8.2.5 στην σχέση (5.9), έχουμε καθώς $N \to \infty$ (και από την (8.44) καθώς $\delta \to 0^+$)

$$E[T_N] = A_N f(N)\left[\frac{1}{\delta} + \ln\delta + \gamma - \delta\ln\delta + (\omega(N) - \gamma)\,\delta + O\left(\delta^2 \ln^2 \delta\right)\right], \tag{8.78}$$

όπου ω ως έχει τεθεί στην σχέση (8.30).
Σημειώνουμε ότι στην εργασία των V. G. Papanicolaou, S. Boneh, (βλέπε [12]), είχε προσδιοριστεί ο πρώτος όρος επί του ασυμπτωτικού αναπτύγματος της μέσης τιμής $E[T_N]$. Η σχέση (8.78) είναι μία σαφής βελτίωση, υπό τις συνθήκες (8.3)).

⋆⋆⋆ **Σκιαγραφούμε τώρα την προσέγγισή μας επί της παρουσίασης των πέντε (!) πρώτων όρων επί του ασυμπτωτικού αναπτύγματος της μέσης τιμής $E[T_N]$.**

Εφόσον, $I_1(N) = o(1)$ και $I_2(N) = o(1)$, (βλέπε (8.25)), είναι καταρχάς σημαντικό να βρει κανείς ποιό από τα ολοκληρώματα $I_1(N)$, $I_2(N)$ είναι πιο σημαντικό. Η διαίσθηση προκρίνει το $I_1(N)$. Μία λοιπόν λογική στρατηγική είναι να ξεκινήσει κανείς από το $I_2(N)$, για το οποίο ο πρώτος όρος φαντάζει αρκετός. Ως προς το $I_1(N)$ είναι καθοριστικό να δει κανείς ότι η δύναμις του ολοκληρώματος αυτού προέρχεται από το άνω άκρο ολοκλήρωσης, δηλαδή το 1. Συνεπώς, αναλύουμε το $I_1(N)$, ως άθροισμα των $I_{11}(N)$ και $I_{12}(N)$ (βλέπε (8.23) και (8.24)). Είναι εύκολο να αποδείξει κάποιος ότι $I_{11}(N) = o\left(I_2(N)\right)$. Συνεπώς, πρέπει να εστιάσουμε στο $I_{11}(N)$, το οποίο πρέπει να αναλυθεί βαθύτερα. Στην πραγματικότητα – για να επιτευχθεί ο στόχος της εύρεσης του πρώτου όρου του ασυμπτωτικού αναπτύγματος της διασποράς – θα πρέπει κανείς να βρει τόσους όρους στο ασυμπτωτικο ανάπτυγμα του $I_{11}(N)$, ώστε να φτάσει σε όρο **της αυτής τάξεως** με

[3] Η παραπάνω αυστηρή προσέγγιση ουσιαστικά μας λέει ότι μπορούμε να αναπτύξουμε κατά Taylor, ως προς s, την ολοκληρωτέα συνάρτηση του I_2.

τον πρωτεύοντα όρο του $I_2(N)$. Μία παρόμοια προσέγγιση θα αποδειχθεί απαραίτητη και για την δεύτερη ροπή. Αυτή θα παρουσιάσουμε στην επόμενη ενότητα, και όπως θα δούμε είναι αναγκαίο να εμβαθύνει κανείς έως **και τον έκτο όρο** της δεύτερης ροπής.

8.3 Βαθύτερη Ασυμπτωτική Συμπεριφορά της Δεύτερης Ροπής $E[\,T_N(T_N+1)\,]$

Ακολουθώντας την προσέγγισή της Ενότητας 8.2 εκφράζουμε την ποσότητα $Q_N(\alpha)$ ως,

$$Q_N(\alpha) = 2F(N)^2\left[\frac{1}{2} - \int_0^1 \exp\left(\sum_{j=1}^N \ln\left(1 - e^{-\frac{F(N)}{f(j)}s}\right)\right) s\,ds\right]$$

$$+2F(N)^2\int_1^\infty \left[1 - \exp\left(\sum_{j=1}^N \ln\left(1 - e^{-\frac{F(N)}{f(j)}s}\right)\right)\right] s\,ds. \tag{8.79}$$

Θέτουμε

$$I_3(N) := \int_0^1 \left[\exp\left(\sum_{j=1}^N \ln\left(1 - e^{-\frac{F(N)}{f(j)}s}\right)\right)\right] s\,ds, \tag{8.80}$$

και

$$I_4(N) := \int_1^\infty \left[1 - \exp\left(\sum_{j=1}^N \ln\left(1 - e^{-\frac{F(N)}{f(j)}s}\right)\right)\right] s\,ds. \tag{8.81}$$

Από τις (8.12) και (8.17), γνωρίζουμε ότι

$$I_3(N) = o(1) \qquad \text{και} \qquad I_4(N) = o(1). \tag{8.82}$$

8.3α′ Το Ολοκλήρωμα $I_3(N)$

Ως προς το ολοκλήρωμα $I_3(N)$ της σχέσης (8.80), και για δοθέν $\varepsilon \in (0,1)$ έχουμε

$$I_3(N) = I_{31}(N) + I_{32}(N), \tag{8.83}$$

όπου

$$I_{31}(N) := \int_0^{1-\varepsilon} \left[\exp\left(\sum_{j=1}^N \ln\left(1 - e^{-\frac{F(N)}{f(j)}s}\right)\right)\right] s\,ds \tag{8.84}$$

και

$$I_{32}(N) := \int_{1-\varepsilon}^1 \left[\exp\left(\sum_{j=1}^N \ln\left(1 - e^{-\frac{F(N)}{f(j)}s}\right)\right)\right] s\,ds \tag{8.85}$$

Ως προς το ολοκλήρωμα $I_{31}(N)$, όπως ορίστηκε στην σχέση (8.84), ισχύουν διαδοχικά

$$\begin{aligned} I_{31}(N) &:= \int_0^{1-\varepsilon} \left[\exp\left(\sum_{j=1}^{N} \ln\left(1-e^{-\frac{F(N)}{f(j)}s}\right)\right)\right] s\,ds \\ &< (1-\varepsilon)^2 \exp\left(\sum_{j=1}^{N} \ln\left(1-e^{-\frac{F(N)}{f(j)}(1-\varepsilon)}\right)\right) \\ &< \exp\left(\sum_{j=1}^{N} \ln\left(1-e^{-\frac{F(N)}{f(j)}(1-\varepsilon)}\right)\right) \\ &< \exp\left(-\sum_{j=1}^{N} e^{-\frac{F(N)}{f(j)}(1-\varepsilon)}\right). \end{aligned}$$

Επομένως από την σύγκριση ολοκληρώματος και αθροίσματος (8.14) λαμβάνουμε

$$I_{31}(N) < \exp\left(-\int_1^N e^{-\frac{F(N)}{f(x)}(1-\varepsilon)}dx\right).$$

Εφαρμόζοντας το Λήμμα 8.2.1, για $m=0$, έχουμε

$$I_{31}(N) < \exp\left\{-\frac{f(N)^2}{(1-\varepsilon)F(N)f'(N)} e^{-\frac{F(N)}{f(N)}(1-\varepsilon)} \left(1+M_2\frac{f(N)}{F(N)}\right)\right\}. \tag{8.86}$$

όπου M_2 είναι μία θετική σταθερά. Από τον ορισμό της $F(\cdot)$, (δηλαδή την σχέση (8.8)), η τελευταία ανισότητα παίρνει την μορφή

$$I_{31}(N) < \exp\left\{-\frac{1}{1-\varepsilon}\ \frac{\left(\frac{f(N)}{f'(N)}\right)^{\varepsilon}}{\ln\left(\frac{f(N)}{f'(N)}\right)} \left(1+M_2\frac{1}{\ln\left(\frac{f(N)}{f'(N)}\right)}\right)\right\}. \tag{8.87}$$

Επειδή $\frac{f'(x)}{f(x)} \to 0$ και $\varepsilon \in (0,1)$ έχουμε

$$I_{31}(N) < \left[\frac{1}{\ln\left(\frac{f(N)}{f'(N)}\right)}\right]^3, \tag{8.88}$$

για αρκούντως μεγάλες τιμές του N.
Επόμενός μας στόχος είναι να υπολογίσουμε μερικούς όρους έπί του ασυμπτωτικού αναπτύγματος του ολοκληρώματος της σχέσης (8.85).

Βήμα 1. *Γράφουμε το ολοκλήρωμα I_{32} σε μία πιο εύκολα διαχειρίσιμη μορφή.*

Μπορούμε να διαχειριστούμε το $I_{32}(N)$, όπως αυτό ορίστηκε στην (8.85), κατά τρόπο ανάλογο με αυτόν που διαχειριστήκαμε το $I_{12}(N)$ στην Υποενότητα

8.2.1. Φθάνει λοιπόν κανείς στην ακόλουθη έκφραση

$$
\begin{aligned}
I_{32}(N) &:= \int_{1-\varepsilon}^{1} s\, e^{B(N;s)} ds \\
&= \int_{1-\varepsilon}^{1} s\ \exp\Bigg[-\frac{f(N)^2}{sF(N)f'(N)} e^{-\frac{F(N)}{f(N)}s} \\
&\qquad -\omega(N)\ \frac{f(N)^3}{s^2F(N)^2f'(N)} e^{-\frac{F(N)}{f(N)}s} \left[1 + O\left(\frac{f(N)}{F(N)}\right)\right]\Bigg] ds
\end{aligned}
\tag{8.89}
$$

καθώς $N \to \infty$. Από τον ορισμό της συνάρτησης $F(\cdot)$, αλλά και την αντικατάσταση $s = 1 - t$, η σχέση (8.89) λαμβάνει την μορφή

$$
\begin{aligned}
I_{32}(N) = \int_0^{\varepsilon} (1-t)\exp\Bigg[&-\frac{1}{1-t}\ \frac{\left(\frac{f(N)}{f'(N)}\right)^t}{\ln\left(\frac{f(N)}{f'(N)}\right)} \\
&-\omega(N)\frac{1}{(1-t)^2}\ \frac{\left(\frac{f(N)}{f'(N)}\right)^t}{\ln^2\left(\frac{f(N)}{f'(N)}\right)} \left[1 + O\left(\frac{1}{\ln\left(\frac{f(N)}{f'(N)}\right)}\right)\right]\Bigg] dt,
\end{aligned}
$$

ή σε μία πιο συμπαγή μορφή

$$
\begin{aligned}
I_{32}(N) = \int_0^{\varepsilon} (1-t)\exp\Bigg[&-\frac{\left(\frac{f(N)}{f'(N)}\right)^t}{\ln\left(\frac{f(N)}{f'(N)}\right)}\ \left(\sum_{n=0}^{\infty} t^n\right) \\
&-\omega(N)\frac{\left(\frac{f(N)}{f'(N)}\right)^t}{\ln^2\left(\frac{f(N)}{f'(N)}\right)}\ \sum_{n=1}^{\infty} n t^{n-1}\ \left[1 + O\left(\frac{1}{\ln\left(\frac{f(N)}{f'(N)}\right)}\right)\right]\Bigg] dt.
\end{aligned}
\tag{8.90}
$$

Με τρόπο ανάλογο με αυτόν της Υποενότητας 8.2.1, έχουμε

$$
I_{32} = \int_0^{\varepsilon} (1-t)\exp\left[-\frac{1}{1-t}\ \frac{A^t}{\ln A} - \omega(N)\frac{1}{(1-t)^2}\ \frac{A^t}{\ln^2 A}\left[1 + O\left(\frac{1}{\ln A}\right)\right]\right] dt,
\tag{8.91}
$$

όπου A όπως ορίστηκε στην (8.41), (δηλαδή, $A = f(N)/f'(N)$). Από τις συνθήκες (8.3) παίρνουμε, όπως πριν, $A \to \infty$. Αντικαθιστώντας, και πάλι,

$$
u = \frac{A^t}{\ln A}
$$

και χρησιμοποιώντας την σχέση (8.42), η (8.91) γίνεται

$$
\begin{aligned}
I_{32} = I_{12} - \int_{1/\ln A}^{A^{\varepsilon}/\ln A} &\frac{\ln u + \ln(\ln A)}{\ln A} \exp\left[-\frac{u}{1 - \frac{\ln u}{\ln A} - \frac{\ln(\ln A)}{\ln A}}\right] \\
&\times \exp\left[-\frac{\omega(N)}{\ln A}\ \frac{u}{\left(1 - \frac{\ln u}{\ln A} - \frac{\ln(\ln A)}{\ln A}\right)^2}\left[1 + O\left(\frac{1}{\ln A}\right)\right]\right] \frac{du}{u \ln A}.
\end{aligned}
\tag{8.92}
$$

Εμπλέκοντας την (8.44), δηλαδή την σχέση,

$$\delta = \frac{1}{\ln A} = \frac{1}{\ln\left(\frac{f(N)}{f'(N)}\right)} = \frac{f(N)}{F(N)},$$

η (8.92) έχει την μορφή

$$I_{32}(N) = (1+\delta\ln\delta)\, I_{12} - \delta^2 \int_{\delta}^{\delta\,\exp(\varepsilon/\delta)} \frac{\ln u}{u}\, \exp\left[-\frac{u}{1-\delta\ln u+\delta\ln\delta} - \frac{\omega(N)u\,\delta\,\left(1+O\left(\delta\right)\right)}{\left(1-\delta\ln u+\delta\ln\delta\right)^2}\right] du. \tag{8.93}$$

(φυσικά, $A\to\infty$ συνεπάγεται ότι, $\delta\to 0^+$).

Βήμα 2. *Αναλύουμε το ολοκλήρωμα I_{32} παρεμβάλλοντας τον όρο $1/\sqrt{\delta}$.*

Ως στην Υποενότητα 8.2.1, σημειώνουμε ότι

$$1/\sqrt{\delta} << \delta\exp\left(\varepsilon/\delta\right),$$

συνεπώς, η σχέση (8.93) δύναται να γραφεί ως

$$\begin{aligned} I_{32}(N) = {} & (1+\delta\ln\delta)\, I_{12} \\ & - \delta^2 \int_{\delta}^{1/\sqrt{\delta}} \frac{\ln u}{u}\, \exp\left[-\frac{u}{1-\delta\ln u+\delta\ln\delta} - \frac{\omega(N)\,u\,\delta\,\left(1+O\left(\delta\right)\right)}{\left(1-\delta\ln u+\delta\ln\delta\right)^2}\right] du \\ & - \delta^2 \int_{1/\sqrt{\delta}}^{\delta\,\exp(\varepsilon/\delta)} \frac{\ln u}{u}\, \exp\left[-\frac{u}{1-\delta\ln u+\delta\ln\delta} - \frac{\omega(N)\,u\,\delta\,\left(1+O\left(\delta\right)\right)}{\left(1-\delta\ln u+\delta\ln\delta\right)^2}\right] du. \end{aligned} \tag{8.94}$$

Βήμα 3. *Εκτιμούμε το δεύτερο ολοκλήρωμα του Βήματος 2.*

Αρχικά εργαζόμαστε με το δεύτερο ολοκλήρωμα της σχέσης (8.94) και βρίσκουμε ένα άνω φράγμα ως ακολούθως:

$$\left|\int_{1/\sqrt{\delta}}^{\delta\,\exp(\varepsilon/\delta)} \frac{\ln u}{u}\, \exp\left[-\frac{u}{1-\delta\ln u+\delta\ln\delta} - \frac{\omega(N)\,u\,\delta}{\left(1-\delta\ln u+\delta\ln\delta\right)^2}\,\left(1+O\left(\delta\right)\right)\right] du\right|$$

$$=\int_{1/\sqrt{\delta}}^{\delta\exp(\varepsilon/\delta)}\frac{\ln u}{u}\exp\left[-\frac{u}{1-\delta\ln u+\delta\ln\delta}\left(1+\frac{\omega(N)}{1-\delta\ln u+\delta\ln\delta}\,\delta\left(1+O\left(\delta\right)\right)\right)\right]du$$
$$\leq\int_{1/\sqrt{\delta}}^{\delta\exp(\varepsilon/\delta)}\exp\left(-\frac{u\ \left(1+O\left(\delta\right)\right)}{1-\delta\ln(1/\sqrt{\delta})+\delta\ln\delta}\right)\frac{\ln\left(1/\sqrt{\delta}\right)du}{1/\sqrt{\delta}}$$
$$=-\frac{\sqrt{\delta}\ln\delta}{2}\int_{1/\sqrt{\delta}}^{\delta\exp(\varepsilon/\delta)}\exp\left(-\frac{u\left(1+O\left(\delta\right)\right)}{1+\frac{3}{2}\delta\ln\delta}\right)du$$
$$=O\left(\sqrt{\delta}\ln\delta\int_{1/\sqrt{\delta}}^{\infty}e^{-u\ (1+O(\delta))}du\right)$$
(8.95)
$$=O\left(\sqrt{\delta}\ln\delta\ e^{-1/\sqrt{\delta}}\right).$$

Βήμα 4. *Εργαζόμαστε με το πρώτο ολοκλήρωμα του Βήματος 2, $K_2(\delta)$. Εφαρμόζουμε ανάπτυγμα κατά Taylor.*

Ως προς το πρώτο ολοκλήρωμα της σχέσης (8.94), θέτουμε
(8.96)
$$K_2(\delta):=\int_{\delta}^{1/\sqrt{\delta}}\exp\left[-\frac{u}{1-\delta\ln u+\delta\ln\delta}-\frac{\omega(N)\ u\ \delta}{\left(1-\delta\ln u+\delta\ln\delta\right)^2}\left(1+O\left(\delta\right)\right)\right]\frac{\ln u}{u}du,$$
και εφόσον, $(1-x)^{-2}=\sum_{n=1}^{\infty}nx^{n-1}$, για $|x|<1$, παίρνουμε
$$\begin{aligned}K_2(\delta)=&\int_{\delta}^{1/\sqrt{\delta}}\exp\Bigg[-u\sum_{n=0}^{\infty}\left(\delta\ln\left(\frac{u}{\delta}\right)\right)^n-\omega(N)\ u\ \delta\left(1+O\left(\delta\right)\right)\\&\times\sum_{n=1}^{\infty}n\left(\delta\ln\left(\frac{u}{\delta}\right)\right)^{n-1}\Bigg]\frac{\ln u}{u}du\\=&\int_{\delta}^{1/\sqrt{\delta}}e^{-u}\exp\left[-u\sum_{n=1}^{\infty}\left(\delta\ln\left(\frac{u}{\delta}\right)\right)^n\right]\\&\times\exp\left[-\omega(N)u\delta\left(1+O\left(\delta\right)\right)\sum_{n=1}^{\infty}n\left(\delta\ln\left(\frac{u}{\delta}\right)\right)^{n-1}\right]\frac{\ln u}{u}du.\end{aligned}$$
Εν συνεχεία, αναπτύσουμε τους εκθετικούς όρους έχοντας
$$\begin{aligned}K_2(\delta)=&\int_{\delta}^{1/\sqrt{\delta}}\frac{e^{-u}\ln u}{u}\left\{1-u\sum_{n=1}^{\infty}\left(\delta\ln\frac{u}{\delta}\right)^n+O\left(u\sum_{n=1}^{\infty}\left(\delta\ln\frac{u}{\delta}\right)^n\right)^2\right\}\\&\times\Bigg\{1-\omega(N)\ u\ \delta\left(1+O\left(\delta\right)\right)\sum_{n=1}^{\infty}\left(\delta\ln\frac{u}{\delta}\right)^{n-1}\\&+O\left(\omega(N)\ u\ \delta\left(1+O\left(\delta\right)\right)\sum_{n=1}^{\infty}\left(\delta\ln\frac{u}{\delta}\right)^{n-1}\right)^2\Bigg\}du.\end{aligned}$$ [4]

[4] Όπως στην Υποενότητα 8.2.1, από την πρώτη αγκύλη χρειαζόμαστε μόνον τον πρώτο όρο της σειράς $\sum_{n=1}^{\infty}\left(\delta\ln\frac{u}{\delta}\right)^n$. Επίσης, από την δεύτερη αγκύλη ο πρώτος όρος της σειράς $\sum_{n=1}^{\infty}\left(\delta\ln\frac{u}{\delta}\right)^{n-1}$ θα αποδειχθεί ότι τελικά είναι αρκετός.

Βήμα 5. *Αναλύουμε το ολοκλήρωμα $K_1(\delta)$ παρεμβάλλοντας το ∞.*

Καταλήγει λοιπόν κανείς στην έκφραση

$$\begin{aligned}
K_2(\delta) &= \int_\delta^{1/\sqrt{\delta}} \frac{e^{-u}\ln u}{u}\left[1-u\left(\delta\ln\frac{u}{\delta}+\omega(N)\delta\left(1+O\left(\delta\right)\right)\right)+u^2O\left(\delta^2\ln^2\frac{u}{\delta}\right)\right]du\\
&= \int_\delta^{\infty} \frac{e^{-u}\ln u}{u}\left[1-u\left(\delta\ln\frac{u}{\delta}+\omega(N)\delta\left(1+O\left(\delta\right)\right)\right)+u^2O\left(\delta^2\ln^2\frac{u}{\delta}\right)\right]du\\
&\quad -\int_{1/\sqrt{\delta}}^{\infty} \frac{e^{-u}\ln u}{u}\left[1-u\left(\delta\ln\frac{u}{\delta}+\omega(N)\delta\left(1+O\left(\delta\right)\right)\right)+u^2O\left(\delta^2\ln^2\frac{u}{\delta}\right)\right]du.
\end{aligned}\tag{8.97}$$

Βήμα 6. *Εκτιμούμε το δεύτερο ολοκλήρωμα του Βήματος 5.*

Ωστόσο, καθώς $\delta\to 0^+$

$$\begin{aligned}
&\int_{1/\sqrt{\delta}}^{\infty} e^{-u}\,\frac{\ln u}{u}\left[1-u\left(\delta\ln\frac{u}{\delta}+\omega(N)\delta\left(1+O\left(\delta\right)\right)\right)+u^2O\left(\delta^2\ln^2\frac{u}{\delta}\right)\right]du\\
&\le \int_{1/\sqrt{\delta}}^{\infty} e^{-u}\,\frac{\ln\left(1/\sqrt{\delta}\right)}{1/\sqrt{\delta}}\left(1-\frac{1}{\sqrt{\delta}}\,\delta\ln\frac{1/\sqrt{\delta}}{\delta}-\ \omega(N)\delta\left(1+O\left(\delta\right)\right)\right)du\\
&+\int_{1/\sqrt{\delta}}^{\infty} u\ \ln u\; e^{-u}O\left(\delta^2\ln^2\frac{u}{\delta}\right)du\\
&\qquad =\int_{1/\sqrt{\delta}}^{\infty} -\frac{\sqrt{\delta}\ln\delta}{2}\;e^{-u}\left(1+\frac{3}{2}\sqrt{\delta}\ln\delta-\omega(N)\delta\left(1+O\left(\delta\right)\right)\right)du\\
&\qquad +O\left(e^{-1/\sqrt{\delta}}\delta^{\frac{3}{2}}\ln^3\delta\right)\\
&\qquad =-\frac{\sqrt{\delta}\ln\delta}{2}\;e^{-1/\sqrt{\delta}}\left(1+\frac{3}{2}\sqrt{\delta}\ln\delta-\omega(N)\delta\left(1+O\left(\delta\right)\right)\right)\\
&\qquad +O\left(e^{-1/\sqrt{\delta}}\delta^{\frac{3}{2}}\ln^3\delta\right)\\
&\qquad =O\left(\sqrt{\delta}\ln\delta\;e^{-1/\sqrt{\delta}}\right)+O\left(e^{-1/\sqrt{\delta}}\delta^{\frac{3}{2}}\ln^3\delta\right)\\
&\qquad =O\left(\sqrt{\delta}\ln\delta\;e^{-1/\sqrt{\delta}}\right).
\end{aligned}\tag{8.98}$$

Βήμα 7. *Goal !! Γράφουμε το ολοκλήρωμα $K_2(\delta)$ με άνω άκρο ολοκλήρωσης το ∞.*

Έπεται λοιπόν ότι στην έκφραση του ολοκληρώματος $K_2(\delta)$ μπορούμε να αντικαταστήσουμε το άνω άκρο ολοκλήρωσης με το ∞, δηλαδή

$$K_2(\delta)=\int_\delta^{\infty}\ln u\;\frac{e^{-u}}{u}\left[1-u\left(\delta\ln\frac{u}{\delta}+\omega(N)\delta\left(1+O\left(\delta\right)\right)\right)+u^2O\left(\delta^2\ln^2\frac{u}{\delta}\right)\right]du.\tag{8.99}$$

Βήμα 8. *Ασυμπτωτική συμπεριφορά του ολοκληρώματος I_{32} καθώς $\delta \to 0^+$.*

Από την σχέση (8.99), το ολοκλήρωμα $I_{32}(N)$ της σχέσης (8.94) γίνεται

(8.100)
$$I_{32}(N) = (1 + \delta \ln \delta)\, I_{12} - \delta^2 \int_\delta^\infty \frac{e^{-u} \ln u}{u} \left[1 - u \left(\delta \ln \frac{u}{\delta} + \omega(N)\delta \left(1 + O\left(\delta\right)\right)\right) + u^2 O\left(\delta^2 \ln^2 \frac{u}{\delta}\right)\right] du,$$

με λάθος πολύ μικρής τάξεως, καθώς $\delta \to 0^+$ (βλέπε (8.95), (8.98)).
Σημειώνουμε ακόμη ότι η σχέση (8.100) μας λέει ότι μπορούμε –τελικά– να αντικαταστήσουμε την πρώτη από τις σειρές που εμφανίζονται στην σχέση (8.90), με τους δύο πρώτους όρορυς από το αντίστοιχο ανάπτυγμα Taylor γύρω από το $t = 0$, και την δεύτερη σειρά της ίδιας σχέσης με τον πρώτο, αντίστοιχα, όρο.
Για την συνέχεια της ανάλυσής μας θα χρειαστούμε τα ακόλουθα λήμματα:

Λήμμα 8.3.1. *Για το ολοκλήρωμα,*

$$L(x) := \int_x^\infty \ln t \, \frac{e^{-t}}{t} dt,$$

έχουμε το ακόλουθο ασυμπτωτικό ανάπτυγμα

(8.101)
$$L(x) \sim -\frac{1}{2} \ln^2 x + \frac{1}{2} \left(\gamma^2 + \frac{\pi^2}{6}\right) + x \ln x - x - \frac{1}{4} x^2 \ln x + \frac{1}{8} x^2 + \frac{1}{18} x^3 \ln x - \frac{1}{54} x^3 + \cdots .$$

καθώς $x \to 0^+$.

Απόδειξη. Επειδή

$$\frac{dL(x)}{dx} = -\ln x \, \frac{e^{-x}}{x} = \ln x \left(-\frac{1}{x} + 1 - \frac{1}{2} x + \frac{1}{6} x^2 - \cdots\right),$$

έχουμε

(8.102)
$$L(x) \sim -\frac{1}{2} \ln^2 x + C^* + x \ln x - x - \frac{1}{4} x^2 \ln x + \frac{1}{8} x^2 + \frac{1}{18} x^3 \ln x - \frac{1}{54} x^3 + \cdots ,$$

όπου C^*, είναι μία πραγματική σταθερά, της οποίας και θα υπολογίσουμε την τιμή.
Από την σχέση (8.102) είναι

$$L(x) + \frac{1}{2} \ln^2 x \sim C^*, \quad x \to 0^+$$

άρα

$$C^* = \lim_{x \to 0^+} \left(\int_x^\infty \ln t \, \frac{e^{-t}}{t} dt + \frac{1}{2} \ln^2 x\right).$$

Με ολοκλήρωση κατά παράγοντες έχουμε

$$2C^* = \int_0^\infty e^{-t} \ln^2 t dt. \tag{8.103}$$

Από την σχέση (Α′.3), ή την σχέση (8.60), είναι

$$2C^* = \gamma^2 + \frac{\pi^2}{6} = \Gamma''(1), \tag{8.104}$$

και η απόδειξη ολοκληρώθηκε. ■

Λήμμα 8.3.2. *Έστω το ολοκλήρωμα*

$$M(x) := \int_x^\infty \ln^2 t\; e^{-t} dt.$$

Καθώς $x \to 0^+$, έχουμε την ακόλουθη ασυμπτωτική συμπεριφορά

(8.105)

$$M(x) \sim \left(\gamma^2 + \frac{\pi^2}{6}\right) - x\ln^2 x + 2x\ln x - 2x - \frac{1}{2}x^2\ln x + \frac{1}{4}x^2 + \frac{1}{9}x^3\ln x - \frac{1}{27}x^3 + \cdots .$$

Απόδειξη. Καθώς

$$\frac{dM(x)}{dx} = -\ln^2 x\; e^{-x} = -\ln^2 x \left(1 - x + \frac{x^2}{2} - \frac{x^3}{6} + \cdots\right),$$

έχουμε

(8.106)

$$M(x) \sim C^* - x\ln^2 x + 2x\ln x - 2x - \frac{1}{2}x^2\ln x + \frac{1}{4}x^2 + \frac{1}{9}x^3\ln x - \frac{1}{27}x^3 + \cdots ,$$

όπου C^*, είναι μία σταθερά. Υπολογίζουμε τώρα την τιμή της C^*. Από την σχέση (8.105) είναι

$$M(x) \sim C^*, \quad (x \to 0^+),$$

επομένως

$$C^* = \lim_{x\to 0^+}\left(\int_x^\infty \ln^2 t\; e^{-t} dt\right).$$

Από την σχέση (8.103) είναι

$$2C^* = \gamma^2 + \frac{\pi^2}{6}, \tag{8.107}$$

και έχουμε το ζητούμενο. ■

Από τα Λήμματα 8.2.3, 8.2.4, 8.3.1, και 8.3.2, το ολοκλήρωμα $I_{32}(N)$ της σχέσης (8.100) γίνεται

$$\begin{aligned}
I_{32}(N) = {} & (1+\delta\ln\delta)\, I_{12} \\
& - \delta^2\left[-\frac{1}{2}\ln^2\delta + \frac{1}{2}\left(\gamma^2 + \frac{\pi^2}{6}\right) + \delta\ln\delta - \delta + O\left(\delta^2\ln\delta\right)\right] \\
& + \delta^3\left[\gamma^2 + \frac{\pi^2}{6} + O\left(\delta\ln^2\delta\right)\right] \\
& - \delta^3\ln\delta\left[-\gamma + O\left(\delta\ln\delta\right)\right] + \delta^3\left(1 + O(\delta)\right)\omega(N)\left(-\gamma + O\left(\delta\ln\delta\right)\right) \\
& + O\left(\delta^4\int_\delta^\infty u\, e^{-u}\ln^3 u\, du + \delta^4\ln^2\delta\int_\delta^\infty u\, e^{-u}\ln u du\right. \\
& \left. -2\delta^4\ln\delta\int_\delta^\infty u\, e^{-u}\ln^2 u\, du\right).
\end{aligned} \tag{8.108}$$

Με ισχυρισμούς παρόμοιους με αυτούς που χρησιμοποιήθηκαν στα προηγούμενα λήμματα (για παράδειγμα όπως στο Λήμμα 8.3.2) είναι εύκολο να δει κανείς ότι ο πρωτεύων όρος επί των ασυμπτωτικών αναπτυγμάτων και των τριών ολοκληρωμάτων που αναφέρονται στην εκτίμηση σφάλματος της σχέσης (8.108), είναι κάθε φορά και μία σταθερά, καθώς $\delta \to 0^+$ (σύγκρινε με την σχέση (8.58)). Για να υπολογίσουμε τις σταθερές αυτές ολοκληρώνουμε κατά παράγοντες και έχουμε

$$\int_0^\infty t\, e^{-t} \ln t \, dt = \int_0^\infty \; e^{-t} \left(\ln t + 1\right) \, dt = 1 - \gamma$$

και

$$\int_0^\infty t\, e^{-t} \ln^3 t \, dt = \int_0^\infty \; e^{-t} \ln^3 t \, dt + 3 \int_0^\infty \; e^{-t} \ln^2 t \, dt = \Gamma'''(1) + 3\Gamma''(1)$$

$$= -2\zeta(3) - \gamma\zeta(2) - \gamma^3 + 3\gamma^2 + \zeta(2),$$

όπου χρησιμοποιήσαμε τις (Α′.5) και (Α′.4). Κατά συνέπεια, από την σχέση (8.62), η (8.108) παίρνει την μορφή

$$\begin{aligned} I_{32}(N) = &- \delta \ln \delta - \gamma\delta - \frac{1}{2}\delta^2 \ln^2 \delta + (1-\gamma)\, \delta^2 \ln \delta \\ &+ \left[1 + \gamma - \omega(N) - \frac{1}{2}\left(\gamma^2 + \frac{\pi^2}{6}\right)\right] \delta^2 + O\left(\delta^3 \ln^2 \delta\right). \end{aligned} \tag{8.109}$$

Σημειώνουμε ότι η εκτίμηση σφάλματος στην σχέση (8.109) είναι ισχυροτέρα από τις αντίστοιχες των σχέσεων (8.95) και (8.98).

8.3β′ Το Ολοκλήρωμα $I_4(N)$

Με παρόμοιο τρόπο όπως αυτόν της Υποενότητας 8.2.2 (σύγκρινε με την σχέση (8.72)), δοθέντος $\vartheta \in (0,1)$ έχουμε

$$\begin{aligned} &(1-\vartheta)^2 \int_1^\infty \left(\int_1^N e^{-\frac{F(N)}{f(x)}s} dx\right) s ds - \vartheta\,(1-\vartheta) \int_1^\infty s\; e^{-\frac{F(N)}{f(N+1)}s} ds \\ &< I_4(N) \\ &< (1+\vartheta)^2 \left[\int_1^\infty \left(\int_1^N e^{-\frac{F(N)}{f(x)}s} dx\right) s ds + \int_1^\infty s\; e^{-\frac{F(N)}{f(N+1)}s} ds\right]. \end{aligned} \tag{8.110}$$

Εναλλάσουμε την σειρά ολοκλήρωσης, χρησιμοποιούμε τον ορισμό της συνάρτησης $F(\cdot)$, και εφαρμόζουμε το Λήμμα 8.2.1, για $m=1$ και $m=2$, και παίρνουμε

$$\begin{aligned} &\int_1^\infty \left(\int_1^N e^{-\frac{F(N)}{f(x)}s} dx\right) s\, ds \\ &= \frac{1}{F(N)} \int_1^N f(x)\, e^{-\frac{F(N)}{f(x)}} \, dx + \frac{1}{F(N)^2} \int_1^N f^2(x)\, e^{-\frac{F(N)}{f(x)}} \, dx \\ &= \left\{\left(\frac{f(N)}{F(N)}\right)^2 + \left(\frac{f(N)}{F(N)}\right)^3\right\} \left[1 + O\left(\frac{f(N)}{F(N)}\right)\right]. \end{aligned}$$

Επιπλέον, χρησιμοποιώντας τις συνθήκες (8.3), αλλά και την (8.5), και, ξανά, τον ορισμό της $F(\cdot)$, εύκολα έχει κανείς

$$\begin{aligned}&\int_1^{\infty} s\, e^{-\frac{F(N)}{f(N+1)}s}\, ds \\ &= \frac{f(N+1)}{F(N)}\, e^{-\frac{F(N)}{f(N+1)}} + \left[\frac{f(N+1)}{F(N)}\right]^2 e^{-\frac{F(N)}{f(N+1)}} \\ &= o\left(\frac{f(N)^2}{F(N)^2}\right).\end{aligned}$$

Ειδικότερα, μέσω της σχέσης (8.73) μένει μόνο να δείξουμε ότι ισχύει

$$\left[\frac{f(N+1)}{F(N)}\right]^2 e^{-\frac{F(N)}{f(N+1)}} = o\left(\frac{f(N)^2}{F(N)^2}\right).$$

Τούτο όμως είναι άμεσο και μάλιστα έπεται από τις σχέσεις (8.8), (8.3), και (8.5). Καθώς το $\vartheta \in (0,1)$ είναι αυθαίρετο, η διπλή ανισότητα (8.110) αποδίδει

(8.111)
$$I_4(N) = \left\{\left(\frac{f(N)}{F(N)}\right)^2 + \left(\frac{f(N)}{F(N)}\right)^3\right\}\left[1 + O\left(\frac{f(N)}{F(N)}\right)\right] \quad \text{καθώς} \quad N \to \infty.$$

Χρησιμοποιώντας τον ορισμό της $F(\cdot)$, αλλά και την σχέση (8.44) έχουμε

(8.112)
$$I_4(N) = \delta^2 + \delta^3 + O\left(\delta^4\right) \quad \text{καθώς} \quad \delta \to 0^+.$$

Είμαστε, πλέον, σε θέση να παρουσιάσουμε το ακόλουθο

Θεώρημα 8.3.3. *Έστω δ όπως έχει οριστεί στην σχέση (8.44) (επομένως $\delta \to 0^+$ καθώς $N \to \infty$) και $\omega(N)$ ως ορίστηκε στην (8.30). Τότε*

$$\begin{aligned}Q_N(\alpha) = f(N)^2 \Bigg\{&\frac{1}{\delta^2} + \frac{2\ln\delta}{\delta} + \frac{2\gamma}{\delta} + \ln^2\delta + 2\,(\gamma-1)\ln\delta \\ &+ \left(2\omega(N) + \gamma^2 + \frac{\pi^2}{6} - 2\gamma\right) + O\left(\delta\ln^2\delta\right)\Bigg\}.\end{aligned} \tag{8.113}$$

Απόδειξη. Το ζητούμενο έπεται άμεσα, εάν συνδυάσουμε τις σχέσεις (8.80), (8.81) (8.88), (8.109), και (8.112), με την (8.79). ∎

Παρατήρηση 9. Χρησιμοποιώντας το Θεώρημα 8.3.3 επί της (5.14), έχουμε καθώς $N \to \infty$ (και από την (8.44) καθώς $\delta \to 0^+$)

$$\begin{aligned}E[\,T_N(T_N+1)\,] = A_N^2 f(N)^2 \Bigg\{&\frac{1}{\delta^2} + \frac{2\ln\delta}{\delta} + \frac{2\gamma}{\delta} + \ln^2\delta + 2\,(\gamma-1)\ln\delta \\ &+ \left(2\omega(N) + \gamma^2 + \frac{\pi^2}{6} - 2\gamma\right) + O\left(\delta\ln^2\delta\right)\Bigg\},\end{aligned} \tag{8.114}$$

όπου $\omega(N)$ είναι, ως έχει οριστεί στην σχέση (8.30).

8.4 Συμπέρασμα. Ασυμπτωτική Συμπεριφορά της Διασποράς $V[T_N]$

Είμαστε πλέον έτοιμοι για το κύριο αποτέλεσμά μας ως προς την ασυμπτωτική συμπεριφορά της διασποράς $V[T_N]$.

Θεώρημα 8.4.1. *Έστω $\alpha = \{a_j\}_{j=1}^{\infty} = 1/f(j)$, όπου η συνάρτηση $f(\cdot)$ ικανοποιεί τις συνθήκες (8.3) (συνεπώς, $L_i(\alpha) = \infty$). Τότε καθώς $N \to \infty$, έχουμε*

$$V[T_N] \sim \frac{\pi^2}{6}\, A_N^2\, f(N)^2 = \frac{\pi^2}{6} \cdot \frac{1}{p_N^2} = \frac{\pi^2}{6} \cdot \frac{1}{\min_{1 \le j \le N}\{p_j\}^2}, \tag{8.115}$$

όπου $A_N = \sum_{j=1}^{N} a_j$ ($p_j = a_j/A_N$ είναι οι πιθανότητες εκλογής των κουπονιών (coupon probabilities)).

Απόδειξη. Από τις σχέσεις (8.78) και (8.114), και με λίγη υπομονή, μετά από μερικούς απλούς υπολογισμούς έχουμε

$$E[T_N(T_N+1)] - E[T_N]^2 \sim \frac{\pi^2}{6}\, A_N^2\, f(N)^2 \quad \text{καθώς } N \to \infty.$$

Υπό το πρίσμα της σχέσης (3.21), και για να ολοκληρωθεί η απόδειξη μας, απομένει να δείξουμε ότι

$$\frac{E[T_N]}{A_N^2\, f(N)^2} \to 0, \quad N \to \infty. \tag{8.116}$$

Από το αποτέλεσμα για την μέση τιμή, δηλαδή την σχέση (8.78), έχουμε

$$E[T_N] \sim A_N f(N) \frac{1}{\delta}.$$

Εμπλέκοντας την (8.44) παίρνουμε

$$E[T_N] \sim A_N F(N),$$

και από την (8.8), φθάνουμε στην

$$E[T_N] \sim A_N f(N) \ln\left(\frac{f(N)}{f'(N)}\right).$$

Συνεπώς, η προς απόδειξη σχέση (8.116), είναι ισοδύναμη με την

$$\frac{\ln f(N) - \ln f'(N)}{A_N\, f(N)} = \frac{1}{A_N}\left[\frac{\ln f(N)}{f(N)} - \frac{\ln f'(N)}{f(N)}\right] \to 0, \quad N \to \infty. \tag{8.117}$$

Για να επιβεβαιώσουμε το όριο αυτό ας παρατηρήσουμε πρώτα ότι από την συνθήκη (i) της (8.3) έχουμε άμεσα

$$\frac{\ln f(N)}{f(N)} \to 0.$$

Επιπλέον,

$$\frac{\ln f'(N)}{f(N)} = \frac{\ln f'(N)}{\ln f(N)} \cdot \frac{\ln f(N)}{f(N)}.$$

Από την (8.6) συνάγουμε εύκολα την ορθότητα της σχέσης (8.117), **είτε η ποσότητα A_N είναι φραγμένη είτε όχι**. Η απόδειξη του Θεωρήματος 8.4.1 έχει πλέον ολοκληρωθεί. ■

Ας σημειωθεί ότι ο πρωτεύων όρος επί του ασυμπτωτικού αναπτύγματος της διασποράς $V[T_N]$, της τ.μ. T_N, είναι **ανεξάρτητος** των ποσοτήτων δ της σχέσης (8.44), αλλά και της ποσότητας $\omega(N)$ της σχέσης (8.30).

Παρατήρηση 10. Είναι αξιοσημείωτο ότι η σταθερά $\pi^2/6$, η οποία πρωτοεμφανίστηκε στην σχέση (4.15), αποκτά διαστάσεις *καθολικότητας*. Όπως θα δούμε στο Κεφάλαιο 9 αυτό το αποτέλεσμα δεν είναι τελικά, τόσο αναπάντεχο.

Παρατήρηση 11. Εάν $C_f := \sum_{n=1}^{\infty} 1/f(n) < \infty$, τότε

$$A_N = C_f\Big(1 + o(1)\Big).$$

Εξάλλου, εάν $C_f = \infty$, τότε καθώς $N \to \infty$ είναι

$$A_N \sim \int_1^N \frac{dx}{f(x)}.$$

Παρατήρηση 12. Η οριακή συνάρτηση κατανομής της σχέσεως (4.17), (βλέπε το Κεφάλαιο 4 όπου μελετήσαμε την περίπτωση των ισοπίθανων κουπονιών), είναι η αθροιστική συνάρτηση της κατανομής Gumbel (*standard Gumbel distribution*):[5]

$$F(x) = \exp\left(-e^{-x}\right),$$

Είναι γνωστό, αλλά και ταυτόχρονα εύκολο να δει κανείς, ότι η ροπογεννήτρια συνάρτηση της αντίστοιχης τ.μ. είναι

$$G(s) := \int_{-\infty}^{\infty} e^{sx} dF(x) = \Gamma(1-s), \qquad \Re\{s\} < 1^6.$$

Κατά συνέπεια για την μέση τιμή έχουμε $G'(0) = -\Gamma'(1) \overset{(A.3)}{=} \gamma$, ενώ για την διασπορά ισχύει $G''(0) - G'(0)^2 = \Gamma''(1) - \gamma^2 \overset{(A.4)}{=} \pi^2/6$.
Η παρατήρηση αυτή από κοινού με το αποτέλεσμά μας για την $E[\,T_N\,]$ (δηλαδή την σχέση (8.78)), αλλά και το Θεώρημα 8.4.1, **υπαινίσονται** ότι μπορεί κανείς να περιμένει ένα οριακό θεώρημα της μορφής
(8.118)

$$P\left\{\frac{T_N - A_N f(N)[\delta(N)^{-1} + \ln\delta(N)]}{A_N f(N)} \le x\right\} \to F(x) = \exp\left(-e^{-x}\right), \quad N \to \infty,$$

όπου

$$\delta(N) = 1/\ln\left(\frac{f(N)}{f'(N)}\right)$$

είναι η ποσότητα που εισήχθη στην σχέση (8.44) και όπως έχουμε δει, $\delta(N) \to 0^+$ καθώς $N \to \infty$. Σε περίπτωση αληθείας της, η (8.118) μας λέει ότι για την κανονικοποίηση της τ.μ. T_N, **απαιτούνται όλοι οι όροι του ασυμπτωτικού**

[5] Η κατανομή Gumbel έχει πολλές εφαρμογές στην υδρολογία.
[6] κάνουμε την αντικατάσταση $e^{-x} = t$.

αναπτύγματος της μέσης τιμής, οι οποίοι είναι πολύ ισχυρότεροι την τυπική απόκλιση. Η διαίσθηση αυτή θα επαληθευτεί στο ακόλουθο κεφάλαιο.

Κεφάλαιο 9

Δύο Οριακά Θεωρήματα για το γενικό πρόβλημα

Ο P. Neal [76], παρουσίασε δύο **γενικά** οριακά θεωρήματα σχετικά με την τ.μ. T_N, όπου $\pi_N = \{p_1^N, p_2^N, ..., p_N^N\}$, $N = 1, 2, ...$, είναι τυχαία μέτρα πιθανότητας (ακόμη και (sub)probability measures), όχι κατ΄ ανάγκην της μορφής (5.1):

Θεώρημα 9.0.2. *(N1) Υποθέτουμε ότι υπάρχουν ακολουθίες $\{b_N\}$ και $\{k_N\}$ εις τρόπον ώστε $k_N/b_N \to 0$, καθώς $N \to \infty$ και ότι για $y \in \mathbb{R}$,*

$$S_N(y) := \sum_{j=1}^{N} \exp\left(-p_j^N(b_N + yk_N)\right) \to g(y), \qquad N \to \infty, \tag{9.1}$$

για μία φθίνουσα (nonincreasing) συνάρτηση $g(\cdot)$ με $g(y) \to \infty$ καθώς $y \to -\infty$, και $g(y) \to 0$ καθώς $y \to \infty$. Τότε

$$\frac{T_N - b_N}{k_N} \xrightarrow{D} Y, \qquad N \to \infty, \tag{9.2}$$

όπου η τ.μ. Y έχει συνάρτηση κατανομής $F(y) = P\{Y \leq y\} = e^{-g(y)}$, $y \in \mathbb{R}$.

Θεώρημα 9.0.3. *(N2) Υποθέτουμε ότι υπάρχει ακολουθία $\{k_N\}$ εις τρόπον ώστε, για $y \in \mathbb{R}^+$,*

$$\sum_{j=1}^{N} \exp\left(-p_j^N yk_N\right) \to \hat{g}(y), \qquad N \to \infty, \tag{9.3}$$

για μία φθίνουσα (nonincreasing) συνάρτηση $\hat{g}(\cdot)$ με $\hat{g}(y) \to \infty$ καθώς $y \to 0$, και $\hat{g}(y) \to 0$ καθώς $y \to \infty$. Επιπλέον, υποθέτουμε ότι υπάρχει μία συνάρτηση $h(\cdot)$ εις τρόπον ώστε, για κάθε $y \in \mathbb{R}^+$,

$$\prod_{j=1}^{N} \left[1 - \exp\left(-p_j^N yk_N\right)\right] \to h(y), \qquad N \to \infty,$$

Τότε η (9.3) εξασφαλίζει ότι $h(y) \to 0$ καθώς $y \to 0$ και $h(y) \to 1$ καθώς $y \to \infty$, και επιπλέον ότι

$$\frac{T_N}{k_N} \xrightarrow{D} \hat{Y}, \qquad N \to \infty,$$

όπου η τ.μ. $\hat{Y}$ έχει συνάρτηση κατανομής $\hat{F}(y) = P\{\hat{Y} \leq y\} = h(y)$, $y \in \mathbb{R}^+$.

Τα Θεωρήματα 9.0.2 και 9.0.3 **δεν υποδεικνύουν την επιλογή των ακολουθιών** $\{b_N\}$ **και** $\{k_N\}$. Εδώ τα ασυμπτωτικά μας αποτελέσματα μπορούν να βοηθήσουν.

Περίπτωση 1. Το συμπέρασμα του Θεωρήματος 9.0.2 (βλέπε την (9.2)), υπονοεί ότι καθώς $N \to \infty$

$$b_N \sim E[T_N] \qquad \text{και} \qquad k_N \sim c\sqrt{V[T_N]}, \quad \text{για κάποιο } c \neq 0.$$

Θυμίζουμε στον αναγνώστη, ότι στην πορεία της παρούσης διατριβής οι πιθανότητες των κουπονιών (coupon probabilities) p_j^N, $1 \leq j \leq N$, $N = 1, 2, ...$, ικανοποιούν τις

$$p_j^N = \frac{a_j}{A_N} \qquad \text{με} \qquad A_N = \sum_{j=1}^{N} a_j.$$

Εάν $a_j = 1/f(j)$, όπου η συνάρτηση $f(x)$ ικανοποιεί τις συνθήκες (8.3), τότε, υπό το πρίσμα των (8.3), η ασυμπτωτική έκφραση (8.78) από κοινού με το Θεώρημα 8.4.1 **προτείνουν** τις επιλογές

$$b_N = A_N f(N) \left[\rho(N) - \ln \rho(N)\right] \qquad \text{και} \qquad k_N = A_N f(N), \tag{9.4}$$

όπου

$$\rho(N) := 1/\delta = \ln\left(f(N)/f'(N)\right).$$

Σημειώνουμε, ότι καθώς $N \to \infty$, $\rho(N) \to \infty$, και συνεπώς $k_N/b_N \to 0$, ως απαιτείται. Επίσης, όπως αναφέραμε στην Παρατήρηση 12 (βλέπε Κεφάλαιο 9), χρησιμοποιούμε τους όρους από το ασυμπτωτικό ανάπτυγμα της μέσης τιμής, οι οποίοι είναι ισχυρότεροι της τυπικής απόκλισης της τ.μ. T_N.
Τότε, $S_N(y)$ (βλέπε (9.1)) έχει την μορφή

$$S_N(y) := \sum_{j=1}^{N} \exp\left(-\frac{f(N)}{f(j)}\left[\rho(N) - \ln \rho(N) + y\right]\right).$$

Επειδή,

$$S_N(y) - S_{N-1}(y) = \exp\left[-\left(\rho(N) - \ln \rho(N) + y\right)\right] \to 0$$

και η συνάρτηση $f(\cdot)$ είναι αύξουσα, έχουμε[1]
(9.8)
$$S_N(y) \sim I_N(y) := \int_1^N \exp\left(-\frac{f(N)}{f(x)}\left[\rho(N) - \ln\rho(N) + y\right]\right) dx, \qquad N \to \infty.$$
Με ολοκλήρωση κατά παράγοντες λαμβάνουμε
(9.9)
$$I_N(y) = \left[\frac{1}{M} \cdot \frac{f(N)^2}{f'(N)} \exp\left(-\frac{M}{f(x)}\right)\right]_{x=1}^N - \frac{1}{M}\int_1^N \left[\frac{f(x)^2}{f'(x)}\right]' \exp\left(-\frac{M}{f(x)}\right) dx,$$
όπου, χάριν ευκολίας έχουμε θέσει
$$M := f(N)[\rho(N) - \ln\rho(N) + y].$$
Αξιοποιώντας τις συνθήκες (8.3) για την συνάρτηση $f(\cdot)$ και για τιμές του N αρκούντως μεγάλες, υπάρχει σταθερά $K > 0$ έτσι ώστε καθώς $L \to \infty$,
$$\left|\frac{1}{M}\left[\frac{f(N)^2}{f'(N)}\right]'\right| = \left|\frac{f(N)}{M}\left[2 - \frac{1}{f(N)} \cdot \frac{f''(N)/f'(N)}{f'(N)/f(N)}\right]\right| \le \left|\frac{2f(N)}{M} + \frac{K}{M}\right| \to 0.$$
Συνεπώς, το ολοκλήρωμα που εμφανίζεται στο δεξιό μέλος της (9.9) είναι $o(I_N(y))$. Επομένως,
$$I_N(y) \sim \frac{f(N)}{f'(N)} \cdot \frac{\exp\left(-\rho(N) + \ln\rho(N) - y\right)}{\rho(N) - \ln\rho(N) + y} \sim e^{-y}.$$
Έπεται λοιπόν ότι
$$S_N(y) \to e^{-y}.$$
Η συνάρτηση $g(y) = e^{-y}$ ικανοποιεί τις απαιτήσεις του Θεωρήματος 9.0.2, δηλαδή είναι φθίνουσα, και $g(y) \to \infty$ καθώς $y \to -\infty$, ενώ $g(y) \to 0$ καθώς $y \to \infty$. Άρα το Θεώρημα 9.0.2 αποδίδει πως $\forall\, y \in \mathbb{R}$,

(9.10)
$$P\left\{\frac{T_N - b_N}{k_N} \le y\right\} \to \exp(-e^{-y}), \qquad N \to \infty,$$

όπου $\{b_N\}$ και $\{k_N\}$ δίδονται από την (9.4). Είναι αξιοσημείωτο ότι η οριακή κατανομή της (9.10) είναι η συνήθης Gumbel κατανομή (standard Gumbel distribution), **ανεξάρτητα** από την επιλογή της συνάρτησης $f(x)$! Μάλιστα, η ίδια

[1] Ένας πιο αυστηρά διατυπωμένος δρόμος προς την σχέση (9.8) ακολουθεί. Επειδή η συνάρτηση $f(\cdot)$ είναι αύξουσα και για αρκούντως μεγάλες τιμές του N η ποσότητα: $f(N)[\rho(N) - \ln\rho(N) + y] > 0$, έχουμε
(9.5)
$$S_{N-1}(y) = \sum_{j=1}^{N-1}\int_j^{j+1} e^{-\frac{f(N)}{f(j)}[\rho(N)-\ln\rho(N)+y]}dx \le \int_1^N e^{-\frac{f(N)}{f(x)}[\rho(N)-\ln\rho(N)+y]}dx := I_N(y)$$
και

(9.6)
$$S_N(y) = \sum_{j=1}^{N}\int_{j-1}^{j} e^{-\frac{f(N)}{f(j)}[\rho(N)-\ln\rho(N)+y]}dx \ge I_N(y).$$

Από τις (9.5), (9.6), την μονοτονία της $f(\cdot)$, και το γεγονός ότι $e^{-[\rho(N)-\ln\rho(N)+y]} \to 0$ (καθώς $N \to \infty$) λαμβάνουμε

(9.7)
$$0 \le S_N(y) - I_N(y) \le S_N(y) - S_{N-1}(y) = o(1), \qquad N \to \infty,$$

και το επιθυμητό αποτέλεσμα έπεται.

οριακή κατανομή εμφανίζεται για διάφορες ακόμη επιλογές των πιθανοτήτων των κουπονιών, συμπεριλαμβανομένης και της περίπτωσης που τα κουπόνια μας είναι ισοπίθανα ή "περίπου" ισοπίθανα – βλέπε, π.χ., (5.15), [58], [11].

Περίπτωση 2. Σχετικά με το Θεώρημα 9.0.3, διαπιστώνουμε πως οι προτάσεις του υπονοούν ότι καθώς $N \to \infty$

$$\frac{E[T_N]}{\sqrt{V[T_N]}} \to c_1 \in \mathbb{R} \qquad \text{και} \qquad k_N \sim c_2\sqrt{V[T_N]}, \quad \text{για κάποια σταθερά } c_2 > 0.$$

Για $p_j^N = a_j/A_N$, με τα $\{a_j\}$ να ικανοποιούν την (6.4) για κάποιο $\xi \in (0,1)$, το Θεώρημα 7.0.3 υποδεικνύει ότι η σωστή επιλογή για την ακολουθία k_N είναι

$$k_N = A_N.$$

Πράγματι, για $y \in \mathbb{R}^+$ έχουμε

$$\sum_{j=1}^{N} \exp\left(-p_j^N y k_N\right) = \sum_{j=1}^{N} \exp\left(-\frac{a_j}{A_N} y A_N\right) = \sum_{j=1}^{N} \exp\left(-a_j y\right) \overset{(6.4)}{\to} \hat{g}(y), \quad N \to \infty.$$

Επιπλέον,

$$\prod_{j=1}^{N}\left(1 - e^{-a_j y}\right) \to \prod_{j=1}^{\infty}\left(1 - e^{-a_j y}\right) := h(y), \qquad N \to \infty.$$

Τελικά, το Θεώρημα 9.0.3 καθώς $N \to \infty$, αποδίδει

$$P\left\{\frac{T_N}{A_N} \le y\right\} \to \prod_{j=1}^{\infty}\left(1 - e^{-a_j y}\right). \tag{9.11}$$

Παρατηρούμε ότι στην περίπτωση αυτή η οριακή κατανομή **εξαρτάται** από την επιλογή της ακολουθίας $\{a_j\}$.
Τέλος, αξίζει να αναφέρουμε ότι η διχοτομία του Κεφαλαίου 6 εμφανίζεται, πάλι και εδώ: Περίπτωση 1 / Περίπτωση 2. Στην πρώτη περίπτωση έχουμε

$$E[T_N]/\sqrt{V[T_N]} \to \infty,$$

ενώ στην δεύτερη

$$E[T_N]/\sqrt{V[T_N]} \to c_1 \in \mathbb{R}.$$

Προς πίστωσιν των παραπάνω προτάσεών μας, εξετάζουμε το ακόλουθο παράδειγμα.

★ ★ ★ Είναι γνωστό ότι στην λίμνη Michigan των Η.Π.Α. ζουν 100 είδη ψαριών, βλέπε [106]. Υποθέτουμε ότι ένας ψαράς είναι τόσο ικανός, ώστε σε κάθε δοκιμή να βγάζει ψάρι από την λίμνη. Στην συνέχεια το ελευθερώνει και πάλι στο νερό. Ο στόχος του είναι να βρει και τα 100 είδη ψαριών (τουλάχιστον μία φορά). Ποιός είναι ο ελάχιστος αριθμός δοκιμών που απαιτούνται, ώστε με πιθανότητα 0.90 να επιτύχει τον στόχο του, εάν

(i) $a_j = 1/j,\ j = 1, 2, \cdots, 100$. Τούτη η περίπτωση είναι γνωστή ως **Zipf distribution** (βλέπε και το Παράδειγμα 4 του Κεφαλαίου 9).
(ii) $a_j = j,\ j = 1, 2, \cdots, 100$. Η περίπτωση αυτή είναι γνωστή ως **Linear case** (βλέπε και το Παράδειγμα 2 του Κεφαλαίου 9).
(iii) $a_j = 1$, δηλαδή $p_j = 1/100$. Τούτη η περίπτωση έχει μελετηθεί αναλυτικά στο Κεφάλαιο 4 (**Equal probabilities case**).

► (i) Στην περίπτωση αυτή ισχύουν:
$N = 100,\ f(N) = N,\ \rho(N) = \ln(f(N)/f'(N)) = \ln N, A_N = H_{100} = 5.18738$.
Συνεπώς, $b_{100} = 1596.67,\ k_{100} = 518.738$. Υποθέτουμε ότι η απάντηση είναι a δοκιμές. Έχουμε

$$\begin{aligned} P\left(T_{100} \leq a\right) &= P\left(\left(T_{100} - 1596.67\right)/518.738 \leq \left(a - 1596.67\right)/518.738\right) \\ &\approx \exp(-e^{-\lambda}) = 0.90, \end{aligned}$$

όπου $\lambda = (a - 1596.67)/518.738$.
Ώστε, $\lambda = -\ln[-\ln(0.90)] = 2.25037$, και κατά συνέπεια με πιθανότητα 0.90, κατ΄ ελάχιστον απαιτούνται **2,765 δοκιμές** για μία πλήρη συλλογή και των 100 ψαριών της λίμνης. (Ο ακριβής υπολογισμός δίδει $\alpha = 2,764.07$).

► (ii). Όπως θα δούμε απαιτούνται **περισσότερες από 12,000 δοκιμές**. Έχουμε, $N = 100$ και $A_N = \sum_{j=1}^{100} j = 5,050$.

$$\begin{aligned} P\left(T_{100} \leq 12,000\right) &= P\left(T_{100}/5,050\right) \leq \left(12,000/5,050\right) \\ &\approx \prod_{j=1}^{\infty}\left(1 - e^{-2.37624\,j}\right) = 0.898477 \approx 0.90. \end{aligned}$$

► (iii). Ένα παρόμοιο παράδειγμα εξετάσαμε στην Ενότητα 4.1. Είναι $N = 100$ και $N \ln N = 460.515$. Υποθέτουμε ότι η απάντηση είναι b δοκιμές. Έχουμε από το αντίστοιχο οριακό Θεώρημα, βλέπε (4.17)

$$\begin{aligned} P\left(T_{100} \leq b\right) &= P\left(\left(T_{100} - 460.517\right)/100 \leq \left(b - 460.517\right)/100\right) \\ &\approx \exp(-e^{-u}) = 0.90, \end{aligned}$$

όπου $u = (b - 460.517)/100$.
Ώστε, $u = -\ln[-\ln(0.90)] = 2.25037$, και κατά συνέπεια με πιθανότητα 0.90, **απαιτούνται τουλάχιστον 686 δοκιμές**. (Ο ακριβής υπολογισμός δίδει $b = 685.554$).
Όσο αφορά στην τυπική απόκλιση έχουμε την ακόλουθη προσέγγιση, για κάθε περίπτωση αντίστοιχα:

$$\sqrt{V[\,T_{N_{\text{Zipf}}}\,]} \approx \sqrt{\frac{\pi^2}{6} N^2 \left(\ln N\right)^2} \stackrel{(11.39)}{=} \sqrt{348,851} \approx 590.64$$

$$\sqrt{V[\,T_{N_{\text{Linear}}}\,]} \approx \sqrt{\left(L_2(\alpha) - L_1(\alpha)^2\right) A_N^2} \stackrel{(11.20)}{=} \sqrt{20,945,800} \approx 4,576.66$$

$$\sqrt{V[\,T_{N_{\text{Equal}}}\,]} \approx \sqrt{\frac{\pi^2}{6} N^2} \stackrel{(4.15)}{=} \sqrt{1,6449.3} \approx 128.26.$$

Μέρος II

Το πρόβλημα του συλλέκτη: Ασυμπτωτική συμπεριφορά όλων των Ροπών ανώτερης τάξης

Κεφάλαιο 10

Οι Ανοδικές Ροπές της τ.μ. T_N

A method is a trick applied twice.

Στο δεύτερο μέρος της διατριβής θα αναπτύξουμε τεχνικές για τον υπολογισμό του πρωτεύοντος όρου επί του ασυμπτωτικού αναπτύγματος των ανοδικών ροπών της τ.μ. T_N, δηλαδή του αριθμού κουπονιών τα οποία πρέπει ένας συλλέκτης να αγοράσει, έτσι ώστε να έχει ένα πλήρες σετ από τα N διαφορετικά κουπόνια, καθώς $N \to \infty$. Αφού αρχικά αναφερθούμε στην περίπτωση κατά την οποία οι πιθανότητες των κουπονιών είναι ίσες θεωρούμε την γενική περίπτωση κατά την οποία είναι άνισες. Για μία μεγάλη κλάση κατανομών –και ενώ υιοθετούμε μία διχοτομία– καταλήγουμε στον ζητούμενο όρο καθώς $N \to \infty$. Επιπλέον παρουσιάζουμε ένα θεώρημα με την βοήθεα του οποίου εκτιμούμε ασυμπτωτικώς τις (ανοδικές) ροπές της τ.μ. T_N συγκρίνοντάς τες με τις αντίστοιχες ροπές που έχουν γνωστές ακολουθίες (βλέπε Θεώρημα 10.5.2 έναντι του Θεωρήματος 10.5.1.)

10.1 Εισαγωγικά

Ας ανακαλέσουμε από το Κεφάλαιο 3 ότι για $z \in \mathbb{C}, |z| \geq 1$, είχαμε εισάγει την ακόλουθη πιθανογεννήτρια συνάρτηση της τ.μ. T_N,

$$G(z) := E\left[z^{-T_N}\right] = 1 + \left(z^{-1} - 1\right) \sum_{k=1}^{\infty} z^{-(k-1)} P\left\{T_N \geq k\right\},$$

καθώς και την επαγόμενη έκφραση:

$$G(z) = 1 - (z-1) \int_0^{\infty} \left[1 - \prod_{j=1}^{N} \left(1 - e^{-p_j t}\right)\right] e^{-(z-1)t} dt,$$

ή ισοδύναμα με χρήση της αντικατάστασης $x = e^{-t}$ υπό την μορφή,

$$G(z) = 1 - (z-1)\int_0^1 \left[1 - \prod_{j=1}^{N}\left(1 - x^{p_j}\right)\right] x^{z-2}dx.$$

Έστω ακέραιος $r \geq 1$. Θέτουμε

$$T_N^{(r)} := T_N\left(T_N+1\right)\left(T_N+2\right)\cdots\left(T_N+r-1\right). \tag{10.1}$$

Θεωρούμε την ακόλουθη r-τάξεως ροπή της τ.μ. T_N (για την ακρίβεια θα πρέπει κανείς να την αποκαλέσει ανοδική r–τάξεως ροπή)

$$E\left[T_N^{(r)}\right] = E\left[T_N\left(T_N+1\right)\left(T_N+2\right)\cdots\left(T_N+r-1\right)\right]. \tag{10.2}$$

Παρατηρούμε ότι,

$$E\left[T_N^{(r)}\right] = (-1)^r \lim_{z\to 1^+} G^{(r)}(z).$$

Με παρόμοιους συλλογισμούς όπως αυτούς του Κεφαλαίου 3 καταλήγουμε στις ακόλουθες εκφράσεις

$$\begin{aligned}
E\left[T_N^{(r)}\right] &= r \sum_{\substack{J\subset\{1,\ldots,N\}\\ J\neq\emptyset}} \frac{(-1)^{|J|-1}}{\left(\sum_{j\in J} p_j\right)^r} \\
&= r\sum_{m=1}^{N}(-1)^{m-1}\sum_{1\leq j_1<\cdots<j_m\leq N}\frac{1}{\left(p_{j_1}+\cdots+p_{j_m}\right)^r} \\
&= r\int_0^\infty \left[1-\prod_{j=1}^{N}\left(1-e^{-p_j t}\right)\right] t^{r-1}\,dt \\
&= (-1)^{r-1}\, r\int_0^1 \left[1-\prod_{j=1}^{N}\left(1-x^{p_j}\right)\right](\ln x)^{r-1}\,\frac{dx}{x}. \qquad (10.3)
\end{aligned}$$

10.2 Η Περίπτωση των Ισοπίθανων Κουπονιών

Όπως αναφέραμε στο Κεφάλαιο 4, η απλούστερη περίπτωση σχετικά με τις προαναφερθείσες εκφράσεις εμφανίζεται όταν έχουμε $p_1 = \cdots = p_N = 1/N$. Ανακαλούμε, επίσης, ότι τούτη η περίπτωση πέραν από την απλότητά της έχει επιπλέον την ιδιότητα ανάμεσα από όλες τις ακολουθίες να είναι αυτή κατά την οποία οι ροπές κάθε τάξης της τ.μ. T_N ελαχιστοποιούνται (βλέπε [55], [11]). Για παράδειγμα, από τις σχέσεις (3.14) και (10.3) έχουμε άμεσα ότι

$$G(z) = E\left[z^{-T_N}\right], \qquad \text{όπου} \quad z \geq 1,$$

μεγιστοποιείται (βλέπε την (4.7)), ενώ οι (ανοδικές) ροπές $E\left[T_N^{(r)}\right]$ ελαχιστοποιούνται, όταν όλες οι πιθανότητες p_j είναι ίσες. Επομένως από την (10.3) έχουμε

$$E\left[T_N^{(r)}\right] = (-1)^{r-1}\, r\int_0^1 \left[1-\left(1-x^{1/N}\right)^N\right]\frac{\ln x}{x}dx. \tag{10.4}$$

Με την αντικατάσταση $u = 1 - x^{1/N}$, το ολοκλήρωμα της σχέσης (10.4) γίνεται

$$\begin{aligned} E\left[T_N^{(r)}\right] &= (-1)^{r-1}\, r\, N^r \int_0^1 \frac{1-u^N}{1-u} \ln^{r-1}(1-u)du \\ &= (-1)^{r-1}\, r\, N^r \sum_{m=1}^{N} \int_0^1 u^{m-1} \ln^{r-1}(1-u)du. \end{aligned}$$

Με επαναλαμβανόμενη εφαρμογή της ολοκλήρωσης κατά παράγοντες έχουμε

$$E\left[T_N^{(r)}\right] = r!\, N^r \sum_{m=1}^{N} \left(\frac{1}{m} \sum_{m_1=1}^{m} \left(\frac{1}{m_1} \sum_{m_2=1}^{m_1} \frac{1}{m_2} \cdots \right) \sum_{m_{r-1}=1}^{m_{r-2}} \frac{1}{m_{r-1}} \right). \tag{10.5}$$

Εξ΄ όσων γνωρίζουμε, εκφράσεις για ροπές της τ.μ. T_N πρωτοεμφανίστηκαν στην εργασία του H. J. Goodwin [50], το έτος 1949.
Για να συνεχίσουμε θεωρούμε τον αναδρομικό τύπο

$$\alpha_1^{(N)} = \sum_{m=1}^{N} \frac{1}{m}, \quad \alpha_r^{(N)} = \sum_{m=1}^{N} \frac{1}{m} \alpha_{r-1}^{(m)}.$$

Οι D. Foata, H. Guo-Niu, και B. Lass (βλέπε [46]), ονόμασαν τους αριθμούς αυτούς *υπεραρμονικούς (hyperharmonic)* και υπολόγισαν την ασυμπτωτική τους συμπεριφορά χρησιμοποιώντας πολυμεταβλητές γεννήτριες συναρτήσεις (multivariate generating fuctions). Λίγο αργότερα, οι I. Adler, S. Oren, και S. Ross (βλέπε [2]), παρουσίασαν μία ακριβή έκφραση των υπεραρμονικών αριθμών χρησιμοποιώντας βασικά εργαλεία της θεωρίας πιθανοτήτων. Είναι άξιοσημείωτο ότι η μελέτη αυτών των ποσοτήτων εμφανίστηκε σε μία πολύ πιο γενική εκδοχή του κλασικού CCP, την οποία ονομάζουμε **"Generalised Coupon Collector's Problem" (GCCP)**, βλέπε σχετικά στο Κεφάλαιο 2. Συγκεκριμένα έχουμε, βλέπε [46]

$$\alpha_r^{(N)} \sim \frac{(\ln N)^r}{r!}, \qquad N \to \infty,$$

επομένως από την (10.5) έχουμε

$$E\left[T_N^{(r)}\right] \sim N^r (\ln N)^r, \qquad N \to \infty. \tag{10.6}$$

Για κάθε r, έχουμε αναλυτική ασυμπτωτική συμπεριφορά για τις αντίστοιχες ροπές $E\left[T_N^{(r)}\right]$ είτε από τα αποτελέσματα της εργασίας [46], είτε με εφαρμογή άλλων γνωστών και αρκετά ισχυρών μεθόδων, όπως για παράδειγμα η μέθοδος μερικής αθροίσεως του Abel (**Abel's partial summation method**).[1] Για παράδειγμα, θεωρούμε την περίπτωση $r = 3$. Θέτουμε

$$S_3(N) := \sum_{m=1}^{N} \left(\frac{1}{m} \sum_{j=1}^{m} \left(\frac{1}{j} \sum_{i=1}^{j} \frac{1}{i} \right) \right).$$

Είναι

$$S_3(N) \overset{(4.10)}{=} \frac{1}{2} \sum_{m=1}^{N} \frac{1}{m} \left(H_m^2 + \sum_{k=1}^{m} \frac{1}{k^2} \right) = \frac{1}{2} \sum_{m=1}^{N} \frac{1}{m} H_m^2 + \frac{1}{2} \sum_{m=1}^{N} \frac{1}{m} \sum_{k=1}^{m} \frac{1}{k^2}.$$

[1] Abel has left mathematicians enough to keep them busy for 500 years, *Charles Hermite (1822–1901), Calculus Gems.*

Εισάγουμε τον συμβολισμό

$$H_m^{(2)} := \sum_{k=1}^{m} 1/k^2.$$

Λαμβάνουμε

$$S_3(N) = \frac{1}{2}\sum_{m=1}^{N}\frac{H_m^2}{m} + \frac{1}{2}\sum_{m=1}^{N}\frac{H_m^{(2)}}{m}. \tag{10.7}$$

Με επαναλαμβανόμενη εφαρμογή της μεθόδου μερικής αθροίσεως του Abel και διαθέτοντας λίγη υπομονή και αρκετό χαρτί για τις απαιτούμενες πράξεις, καταλήγουμε στις ακόλουθες εκφράσεις

$$\sum_{m=1}^{N}\frac{H_m^2}{m} = \frac{5}{3}\sum_{m=1}^{\infty}\frac{1}{m^3} + \frac{1}{3}(\gamma + \ln N)^3 + O\left(\frac{\ln N}{N}\right), \tag{10.8}$$

$$\sum_{m=1}^{N}\frac{H_m^{(2)}}{m} = \frac{\pi^2}{6}(\gamma + \ln N) - \sum_{m=1}^{\infty}\frac{1}{m^3} + O\left(\frac{\ln N}{N}\right). \tag{10.9}$$

Αξίζει να σημειώσουμε ότι και οι δύο σχέσεις (10.8), (10.9) απετέλεσαν το πρόβλημα υπ′ αριθμόν [4582], προταθέν από τον **Murray S. Klamkin** το οποίο δημοσιεύτηκε στο περιοδικό American Mathematical Monthly, problem section, (1954). Η λύση του προβλήματος δημοσιεύτηκε το (1955), από τον **Leonard Carlitz**, (βλέπε [65], [19]).
Από τις (10.8) και (10.9), η σχέση (10.7) έχει την μορφή

$$\begin{aligned} S_3(N) = {} & \frac{1}{6}\ln^3 N + \frac{1}{2}\gamma\ln^2 N + \frac{1}{2}\left(\gamma^2 + \frac{\pi^2}{6}\right)\ln N \\ & + \left(\frac{1}{3}\zeta(3) + \frac{1}{6}\gamma^3 + \frac{\gamma\pi^2}{12}\right) + O\left(\frac{\ln N}{N}\right). \end{aligned} \tag{10.10}$$

Τέλος, από τις σχέσεις (10.5) και (10.10) έχουμε

$$\begin{aligned} E\left[T_N^{(3)}\right] = N^3 \Bigg[& \ln^3 N + 3\gamma\ln^2 N + \left(3\gamma^2 + \frac{\pi^2}{2}\right)\ln N \\ & + \left(2\zeta(3) + \gamma^3 + \frac{\gamma\pi^2}{2}\right) + O\left(\frac{\ln N}{N}\right)\Bigg]. \end{aligned} \tag{10.11}$$

10.3 Η αναγκαιότητα για την ασυμπτωτική μελέτη του προβλήματος – Η Διχοτομία

Όταν οι τιμές του N είναι μεγάλες, δεν είναι ξεκάθαρη η πληροφορία που παίρνουμε από τις εκφράσεις (10.3) σχετικά με τις (ανοδικές) ροπές $E\left[T_N^{(r)}\right]$. Όπως στο Κεφάλαιο 5, θεωρούμε την ακολουθία **αυστηρά** θετικών όρων, $\alpha = \{a_j\}_{j=1}^{\infty}$. Τότε, για κάθε ακέραιο $N > 0$, μπορεί κανείς να ορίσει ένα μέτρο πιθανότητας $\pi_N = \{p_1, ..., p_N\}$ επί του συνόλου $\{1, ..., N\}$ ως εξής:

$$p_j = \frac{a_j}{A_N}, \qquad \text{όπου} \quad A_N = \sum_{j=1}^{N} a_j.$$

Παρατήρηση 13. Προφανώς, για δοθέν N οι πιθανότητες p_j μπορούν να ληφθούν ώστε ως προς j να είναι μονότονες χωρίς βλάβη της γενικότητας. Όσο για την ακολουθία $\{a_j\}_{j=1}^{\infty}$,
(i) εάν $a_j \to \infty$, τότε για κάθε $k \in \mathbb{N}$ υπάρχει $j = j(k) \geq k$, εις τρόπον ώστε, $a_j \geq a_i$, $\forall\ i \leq j$. Τούτο μας λέει πως με μία αναδιάταξη των όρων a_i, όπου $j(k) \leq i \leq j(k+1)$, η ακολουθία $\{a_j\}_{j=1}^{\infty}$ μπορεί να ληφθεί ότι είναι αύξουσα (nondecreasing) χωρίς βλάβη της γενικότητας.
(ii) Ομοίως, εάν $a_j \to 0$, τότε $\{a_j\}_{j=1}^{\infty}$ μπορεί να ληφθεί φθίνουσα (nonicreasing) χωρίς βλάβη της γενικότητας.

Κινητοποιούμενοι από την (10.3) εισάγουμε τον συμβολισμό, (βλέπε [12])

$$\begin{aligned} H_N(\alpha; r) :&= r \sum_{\substack{J \subset \{1,\ldots,N\} \\ J \neq \emptyset}} \frac{(-1)^{|J|-1}}{\left(\sum_{j\in J} a_j\right)^r} \\ &= r \sum_{k=1}^{N} (-1)^{k-1} \sum_{1 \leq j_1 < \cdots < j_k \leq N} \frac{1}{(a_{j_1} + \cdots + a_{j_k})^r}. \end{aligned} \tag{10.12}$$

Έχουμε ήδη μελετήσει τις περιπτώσεις $r = \{1, 2\}$, όπου $H_N(\alpha; 1) = E_N(\alpha)$, και $H_N(\alpha; 2) = Q_N(\alpha)$, (βλέπε τις (5.10) και (5.7) αντίστοιχα). Είναι

$$\begin{aligned} H_N(\alpha; r) &= r \int_0^{\infty} \left[1 - \prod_{j=1}^{N} \left(1 - e^{-a_j t}\right)\right] t^{r-1}\, dt \\ &= (-1)^{r-1}\, r \int_0^1 \left[1 - \prod_{j=1}^{N} (1 - x^{a_j})\right] (\ln x)^{r-1} \frac{dx}{x}. \end{aligned} \tag{10.13}$$

Εάν $s\alpha := \{sa_j\}_{j=1}^{\infty}$, τότε από την (10.12) προκύπτει άμεσα

$$H_N(s\alpha; r) = s^{-r} H_N(\alpha; r) \tag{10.14}$$

ώστε από την (10.3) και τον ορισμό των ακολουθιών από μέτρα πιθανότητας, έχουμε

$$E\left[T_N^{(r)}\right] = A_N^r H_N(\alpha; r). \tag{10.15}$$

Όπως είχε σημειωθεί στην εργασία των V.G. Papanicolaou, S. Boneh [12], για την μέση τιμή $E\,[\,T_N\,]$, αλλά και στο Κεφάλαιο 5 για την δεύτερη (ανοδική) ροπή, έτσι και το γενικό πρόβλημα των ροπών $E\left[T_N^{(r)}\right]$, $N \to \infty$, δύναται να αντιμετωπιστεί ως δύο ξεχωριστά προβλήματα, και πιο συγκεκριμένα της εκτίμησης, πρωτίστως, του $H_N(\alpha; r)$, και εν συνεχεία του A_N^r. Η ανάλυσή μας εστιάζει στον πρώτο στόχο. Η εκτίμηση του A_N^r είναι ένα διαφορετικό πρόβλημα το οποίο, μάλιστα, μπορεί κανείς να χειριστεί με γνωστά και ισχυρά εργαλεία, όπως για παράδειγμα η φόρμουλα άθροισης των Euler–Maclaurin **(Euler-Maclaurin Summation formula)**, η μέθοδος Laplace για αθροίσματα **(Laplace method for sums)** ή μέθοδος αθροίσεως κατά παράγοντες **(Abel's summation by parts)**, (βλέπε, π.χ., [9]).

Υιοθετούμε τώρα την **Διχοτομία** του Κεφαλαίου 6. Με εφαρμογή του Θεωρήματος μονότονης σύγκλισης επί της σχέσεως (10.13) έχουμε

$$\begin{aligned} L_r(\alpha) :&= \lim_N H_N(\alpha; r) \\ &= (-1)^{r-1}\, r \int_0^1 \left[1 - \prod_{j=1}^{\infty}(1 - x^{a_j})\right] (\ln x)^{r-1}\, \frac{dx}{x}. \end{aligned} \tag{10.16}$$

Σημειώνουμε ότι $L_r(\alpha) > 0$, για κάθε ακολουθία α (εφόσον, για κάθε $x \in (0,1)$, είναι $f_N^{\alpha}(x) < 1$ με τιμές που φθίνουν καθώς το N αυξάνει). Θυμίζουμε ότι από την σχέση (6.1) έχουμε θέσει

$$f_N^{\alpha}(x) := \prod_{j=1}^{N}(1 - x^{a_j}), \quad 0 \le x \le 1.$$

Αν και η ποσότητα $L_r(\alpha)$ λαμβάνει θετικές τιμές μπορεί να έχουμε $L_r(\alpha) = \infty$. Στην πραγματικότητα όπως θα αποδείξουμε στην συνέχεια (βλέπε Παρατήρηση 14), $L_r(\alpha) = \infty$, εάν και μόνον εάν, $L_1(\alpha) = \infty$.

Θεώρημα 10.3.1. *$L_r(\alpha) < \infty$, εάν και μόνον εάν υπάρχει $\xi \in (0,1)$, εις τρόπον ώστε*

$$\sum_{j=1}^{\infty} \xi^{a_j} < \infty. \tag{10.17}$$

Απόδειξη. Υποθέτουμε ότι υπάρχει ένα $\xi \in (0,1)$ τέτοιο ώστε να αληθεύει η (10.17). Τότε για κάθε θετικό ακέραιο r, από την (10.16) και το Λήμμα 6.0.2 είναι

$$\begin{aligned} L_r(\alpha) \le& (-1)^{r-1} r \int_0^{\xi} \left[\sum_{j=1}^{\infty} x^{a_j}\right] (\ln x)^{r-1}\, \frac{dx}{x} \\ &+ (-1)^{r-1} r \int_{\xi}^{1} \left[1 - \prod_{j=1}^{\infty}(1 - x^{a_j})\right] (\ln x)^{r-1}\, \frac{dx}{x} \\ <& (-1)^{r-1} r \int_0^{\xi} \left[\sum_{j=1}^{\infty} x^{a_j - 1}\right] \ln(x)^{r-1}\, dx + (-1)^r (\ln \xi)^r. \end{aligned} \tag{10.18}$$

Με ολοκλήρωση κατά παράγοντες υπολογίζουμε το ολοκλήρωμα,

$$\begin{aligned} I_j(\xi; r) &:= \int_0^{\xi} x^{a_j - 1} (\ln x)^{r-1}\, dx \\ &= \left[\frac{x^{a_j} (\ln x)^{r-1}}{a_j}\right]_{x=0}^{\xi} - (r-1) \int_0^{\xi} \frac{x^{a_j - 1}}{a_j} (\ln x)^{r-2}\, dx. \end{aligned}$$

Με επαναλαμβανόμενη εφαρμογή της ίδιας μεθόδου έχουμε

$$I_j(\xi; r) = \frac{1}{a_j}\, \xi^{a_j} \sum_{k=0}^{r-1} (-1)^k (r-1)_k\, \frac{1}{a_j^k} (\ln \xi)^{r-1-k}, \tag{10.19}$$

όπου $(r-1)_k = (r-1)!/\,(r-1-k)!$ είναι το καθοδικό παραγοντικό (falling Pochhammer symbol). Από το θεώρημα Tonelli και την (10.19), η ανισότητα (10.18) έχει την μορφή

$$\begin{aligned} L_r(\alpha) \leq & (-1)^{r-1}\, r \left(\sum_{j=1}^{\infty} \frac{1}{a_j}\, \xi^{a_j} \left(\sum_{k=0}^{r-1} (-1)^k\, (r-1)_k\, \frac{1}{a_j^k}\, (\ln \xi)^{r-1-k} \right) \right. \\ & + (-1)^r r \ln^r \xi . \end{aligned}$$

Παρατηρούμε τώρα ότι από την συνθήκη (10.17) έχουμε ότι

$$\xi^{a_j} \to 0^2$$

επομένως, $a_j \to \infty$. Συνεπώς,

$$\min_j \{a_j\} = a_{j_0} > 0.$$

Υπό την παρατήρηση αυτή η παραπάνω ανισότητα συνεπάγεται ότι,

$$\begin{aligned} L_r(\alpha) \leq & (-1)^{r-1}\, r\, \frac{1}{a_{j_0}} \left(\sum_{j=1}^{\infty} \xi^{a_j} \right) \left(\sum_{k=0}^{r-1} (-1)^k\, (r-1)_k\, \frac{1}{a_{j_0}^k}\, (\ln \xi)^{r-1-k} \right) \\ & + (-1)^r \ln^r \xi . \end{aligned}$$

Επειδή $\xi \in (0,1)$, r είναι ένας θετικός ακέραιος, και η συνθήκη (10.17) είναι σε ισχύ, έχουμε $L_r(\alpha) < \infty$.

Αντιστρόφως, εάν $\sum_{j=1}^{\infty} \xi^{a_j} = \infty$, για κάθε $\xi \in (0,1)$, τότε από γνωστή ιδιότητα των απειρογινομένων είναι (βλέπε, π.χ., [92], σελ. 300)

$$\prod_{j=1}^{\infty} (1 - x^{a_j}) = 0, \qquad \text{για κάθε } x \in (0,1)$$

οπότε η (10.16) γίνεται

$$L_r(\alpha) = (-1)^{r-1}\, r \int_0^1 \left[(\ln x)^{r-1} / x \right] dx = \infty. \qquad \blacksquare$$

Παρατήρηση 14. Έχει αποδειχθεί (βλέπε [12]), ότι $L_1(\alpha) < \infty$, εάν και μόνον εάν υπάρχει $\xi \in (0,1)$ έτσι ώστε $\sum_{j=1}^{\infty} \xi^{a_j} < \infty$. Επομένως, $L_r(\alpha) < \infty$ εάν και μόνον εάν $L_1(\alpha) < \infty$. Συνοψίζοντας, έχουμε τις ακόλουθες **διχοτομίες, ταυτόχρονα για όλους** τους θετικούς ακεραίους r:

(10.20) $\qquad (i) \quad 0 < L_r(\alpha) < \infty \qquad$ ή $\qquad (ii) \quad L_r(\alpha) = \infty.$

Παρατήρηση 15. Θεωρούμε το ακόλουθο σφάλμα (error term), εκπεφρασμένο ως την διαφορά

$$\Delta_r(N) := L_r(\alpha) - H_N(\alpha; r).$$

[2] Το αντίστροφο, φυσικά, δεν ισχύει πάντοτε.

Τότε για κάθε θετικό ακέραιο r από τις (10.13), (10.16), το Θεώρημα Tonelli, και επαναλαμβανομένης εφαρμογής της ολοκλήρωσης κατά παράγοντες, έχουμε

$$\Delta_r(N) = (-1)^{r-1} r \int_0^1 \prod_{j=1}^{N} (1 - x^{a_j}) \left[1 - \prod_{j=N+1}^{\infty} (1 - x^{a_j}) \right] (\ln x)^{r-1} \frac{dx}{x}$$

$$\leq (-1)^{r-1} r \int_0^1 \left(\sum_{j=N+1}^{\infty} x^{a_j} \right) (\ln x)^{r-1} \frac{dx}{x} = r! \sum_{j=N+1}^{\infty} a_j^{-r}. \tag{10.21}$$

Ώστε, εάν

$$\sum_{j=1}^{\infty} a_j^{-r} < \infty, \tag{10.22}$$

η (10.22) είναι ένα άνω φράγμα (μερικές φορές αρκετά ικανοποιητικό) για το σφάλμα $\Delta_r(N)$. Η ίδια σχέση δείχνει επίσης ότι όσο ταχύτερα η a_j^r τείνει στο ∞, τόσο ταχύτερη είναι η σύγκλιση του $L_r(\alpha)$ στο $H_N(\alpha; r)$.

10.4 Η περίπτωση $L_r(\alpha) < \infty$

Έστω A_N και $L_r(\alpha)$ ως έχουν οριστεί στις (5.1) και (10.16) αντίστοιχα. Υπενθυμίζουμε ότι, από το Θεώρημα 10.3.1, $L_r(\alpha) < \infty$ συνεπάγεται ότι: $\lim_j a_j = \infty$ (επομένως, $\lim_N A_N = \infty$).

Θεώρημα 10.4.1. *Εάν $L_r(\alpha) < \infty$, τότε καθώς $N \to \infty$,*

$$E\left[T_N^{(r)}\right] = A_N^r L_r(\alpha) \left[1 + o(1)\right], \tag{10.23}$$

Απόδειξη. Το ζητούμενο προκύπτει άμεσα από τις (10.15) και (10.16). ■

Το Θεώρημα 10.4.1 μας λέει ότι εάν $L_r(\alpha) < \infty$, τότε ο πρωτεύων όρος επί του ασυμπτωτικού αναπτύγματος των (ανοδικών) ροπών, $E\left[T_N^{(r)}\right]$, καθορίζεται, διαισθητικά τουλάχιστον, από την συμπεριφορά της ποσότητας A_N.
Συνεχίζουμε με την σαφώς πιο δύσκολη περίπτωση.

10.5 Η περίπτωση $L_r(\alpha) = \infty$

10.5α΄ Ο πρωτεύων όρος του ασυμπτωτικού αναπτύγματος των (ανοδικών) ροπών της τ.μ. T_N

Από το Θεώρημα 10.3.1 και την Παρατήρηση 14, $L_r(\alpha) = \infty$ είναι ισοδύναμο με τις ισότητες $L_j(\alpha) = \infty$, για όλα τα $j = 1, 2, \cdots, r-1$, και επίσης ισοδύναμο με $\sum_{j=1}^{\infty} x^{a_j} = \infty$, για κάθε $x \in (0,1)$. Για την συνέχεια της ανάλυσής μας ακολουθούμε την εργασία βλέπε [12], και γράφουμε την ακολουθία a_j υπό την μορφή

$$a_j = \frac{1}{f(j)},$$

όπου

$$f(x) > 0, \tag{10.24}$$

και υποθέτουμε επιπλέον, ότι η συνάρτηση $f(x)$ είναι δύο φορές παραγωγίσιμη και ικανοποιεί τις ακόλουθες συνθήκες καθώς $x \to \infty$:

$$\begin{aligned} &(i) \quad f(x) \nearrow \infty, \\ &(ii) \quad \frac{f'(x)}{f(x)} \searrow 0, \\ &(iii) \quad \frac{f''(x)/f'(x)}{[f'(x)/f(x)]\ln[f'(x)/f(x)]} \to 0. \end{aligned} \tag{10.25}$$

Θεώρημα 10.5.1. *Εάν $\alpha = \{1/f(j)\}_{j=1}^{\infty}$, όπου $f(\cdot)$ ικανοποιεί τις (10.24) και (10.25), τότε*

$$H_N(\alpha; r) \sim f(N)^r \ln\left(\frac{f(N)}{f'(N)}\right)^r, \quad N \to \infty. \tag{10.26}$$

Απόδειξη. (Υιοθετούμε την απόδειξη των V.G. Papanicolaou, S. Boneh, [12] για τον πρωτεύοντα όρο επί του ασυμπτωτικού αναπτύγματος της πρώτης ροπής, δηλαδή όταν $H_N(\alpha; 1)$). Θέτουμε

$$F(x) := -f(x)\ln\left[\frac{f'(x)}{f(x)}\right].$$

Προφανώς $F(x) > 0$, τουλάχιστον για τιμές του x αρκούντως μεγάλες. Επομένως, μπορεί κανείς να γράψει την (10.13) ως:

$$\begin{aligned} H_N(\alpha; r) =& F(N)^r H_N\left[\, F(N)\, \alpha; r\, \right] \\ =& \, rF(N)^r \int_0^1 \left[1 - \exp\left(\sum_{j=1}^N \ln\left(1 - e^{-\frac{F(N)}{f(j)}s}\right)\right)\right] s^{r-1} ds \\ &+ rF(N)^r \int_1^\infty \left[1 - \exp\left(\sum_{j=1}^N \ln\left(1 - e^{-\frac{F(N)}{f(j)}s}\right)\right)\right] s^{r-1} ds. \end{aligned} \tag{10.27}$$

Από το Θεώρημα Φραγμένης Σύγκλισης για το πρώτο ολοκλήρωμα της σχέσης (10.27) δίδει (υπό το πρίσμα της σχέσης (8.10))

$$\begin{aligned} H_N(\alpha; r) =& \, rF(N)^r \left[\frac{1}{r} + o(1)\right] \\ &+ rF(N)^r \int_1^\infty \left[1 - \exp\left(\sum_{j=1}^N \ln\left(1 - e^{-\frac{F(N)}{f(j)}s}\right)\right)\right] s^{r-1} ds. \end{aligned} \tag{10.28}$$

Για την συνέχεια θέλουμε να εκτιμήσουμε το ολοκλήρωμα της ανωτέρω σχέσης. Από το Θεώρημα Κυριαρχημένης Σύγκλισης (καθώς $f(N)/f'(N) \to \infty$) έχουμε

$$\lim_N \int_1^\infty \left[1 - \exp\left(-\frac{(f(N)/f'(N))^{1-s}}{s\ln(f(N)/f'(N))}\right)\right] s^{r-1} ds = 0.$$

Από την σχέση (8.11) έπεται ότι

$$\lim_N \int_1^\infty \left[1 - \exp\left(-\int_1^N e^{-\frac{F(N)}{f(x)}s}dx\right)\right] s^{r-1}ds = 0. \tag{10.29}$$

Επειδή η συνάρτηση $f(\cdot)$ είναι αύξουσα έχουμε από την σύγκριση σειράς και ολοκληρώματος (ανακαλούμε την διπλή ανισότητα (8.15)) δηλαδή,

$$\begin{aligned}1 - \exp\left(-\int_1^N e^{-\frac{F(N)}{f(x)}s}dx\right) \leq & 1 - \exp\left(-\sum_{j=1}^N e^{-\frac{F(N)}{f(j)}s}\right)\\ \leq & 1 - \exp\left(-\int_1^N e^{-\frac{F(N)}{f(x)}s}dx + e^{-\frac{F(N)}{f(N+1)}s}\right).\end{aligned}$$

Επομένως παίρνοντας όρια στην (8.15) και χρησιμοποιώντας τις (10.29), (8.16), και (8.5) (βλέπε την Παρατήρηση 6), έχουμε

$$\lim_N \int_1^\infty \left[1 - \exp\left(\sum_{j=1}^N \ln\left(1 - e^{-\frac{F(N)}{f(j)}s}\right)\right)\right] s^{r-1}ds = 0. \tag{10.30}$$

Τέλος, από τον ορισμό της συνάρτησης $F(\cdot)$, και το ανάπτυγμα Taylor του λογαρίθμου, δηλαδή $\ln(1-x) \sim -x$ καθώς $x \to 0$, η (10.28) αποδίδει

$$H_N(\alpha; r) \sim F(N)^r = f(N)^r \ln\left(\frac{f(N)}{f'(N)}\right)^r, \quad N \to \infty \tag{10.31}$$

και η απόδειξη ολοκληρώθηκε. ■

Παρατήρηση 16. Χρησιμοποιώντας το Θεώρημα 10.5.1 στην βασική σχέση (10.15), έχουμε ότι καθώς $N \to \infty$

$$\begin{aligned}E\left[T_N^{(r)}\right] \sim & A_N^r f(N)^r \ln\left(\frac{f(N)}{f'(N)}\right)^r\\ = & \frac{1}{p_N^r} \ln\left(\frac{f(N)}{f'(N)}\right)^r\\ = & \frac{1}{\min_{1\leq j\leq N}\{p_j\}^r} \ln\left(\frac{f(N)}{f'(N)}\right)^r,\end{aligned} \tag{10.32}$$

όπου η τελευταία ισότητα έπεται από την σχέση (5.1).

10.5β′ Ασυμπτωτικές εκτιμήσεις των (ανοδικών) ροπών της τ.μ. T_N μέσω της σύγκρισης με ακολουθίες, για τις οποίες οι αντίστοιχες ροπές είναι γνωστές

Στην υποενότητα αυτή θα παρουσιάσουμε ένα θεώρημα το οποίο μας βοηθά να έχουμε ασυμπτωτικές εκτιμήσεις με την βοήθεια της σύγκρισης με ακολουθίες α για τις οποίες ασυμπτωτικές εκτιμήσεις για τις ποσότητες $H_N(\alpha; r)$ είναι γνωστές (για παράδειγμα, μέσω του Θεωρήματος 10.5.1). Έτσι, το Θεώρημα 10.5.2 που θα δούμε στην συνέχεια, επεκτείνει την κλάση των κατανομών για τις οποίες το

Θεώρημα 10.5.1 δεν μπορεί να εφαρμοστεί. Ένα παρόμοιο θεώρημα, που αφορά στην περίπτωση όπου $r = 1$, μπορεί να βρεθεί στην εργασία [12]. Ανακαλούμε έναν γνωστό συμβολισμό.
Έστω $\{s_j\}_{j=1}^{\infty}$ και $\{t_j\}_{j=1}^{\infty}$ είναι δύο ακολουθίες μη αρνητικών όρων. Με το σύμβολο $s_j \asymp t_j$ εννούμε ότι υπάρχουν δύο σταθερές $C_1 > C_2 > 0$ και ένας ακέραιος $j_0 > 0$, εις τρόπον ώστε

$$C_2 t_j \leq s_j \leq C_1 t_j, \quad \text{για κάθε} \quad j \geq j_0, \tag{10.33}$$

δηλαδή, $s_j = O(t_j)$ και $t_j = O(s_j)$.

Θεώρημα 10.5.2. *Έστω $\alpha = \{a_j\}_{j=1}^{\infty}$ και $\beta = \{b_j\}_{j=1}^{\infty}$ ακολουθίες αυστηρά θετικών όρων, έτσι ώστε $\lim_N H_N(\alpha; r) = \lim_N H_N(\beta; r) = \infty$.*

(i) Εάν υπάρχει j_0 έτσι ώστε $a_j = b_j$, $\forall\, j \geq j_0$, τότε η διαφορά $H_N(\beta; r) - H_N(\alpha; r)$ είναι φραγμένη,

(ii) εάν $a_j = O(b_j)$, τότε $H_N(\beta; r) = O\left(H_N(\alpha; r)\right)$ καθώς $N \to \infty$,

(iii) εάν $a_j = o(b_j)$, τότε $H_N(\beta, r) = o\left(H_N(\alpha; r)\right)$ καθώς $N \to \infty$,

(iv) εάν $a_j \asymp b_j$, τότε $H_N(\beta; r) \asymp H_N(\alpha; r)$ καθώς $N \to \infty$,

(v) εάν $a_j \sim b_j$, τότε $H_N(\beta; r) \sim H_N(\alpha; r)$ καθώς $N \to \infty$.

Απόδειξη. Η περίπτωση (i) έπεται εύκολα από την (5.8):

$$\begin{aligned}
&|H_N(\beta; r) - H_N(\alpha; r)| \\
=& r \left| \int_0^{\infty} \prod_{j=j_0}^{N} \left(1 - e^{-a_j t}\right) \left[\prod_{j=1}^{j_0-1} \left(1 - e^{-a_j t}\right) - \prod_{j=1}^{j_0-1} \left(1 - e^{-b_j t}\right) \right] t^{r-1}\, dt \right| \\
\leq& \int_0^{\infty} \left| \left[\prod_{j=1}^{j_0-1} \left(1 - e^{-a_j t}\right) - \prod_{j=1}^{j_0-1} \left(1 - e^{-b_j t}\right) \right] \right| t^{r-1}\, dt \\
=& \int_0^{\infty} \left| \sum_{J \subset \{1,\dots,j_0-1\}} (-1)^{|J|} \left\{ \exp\left(-t \sum_{j \in J} a_j\right) - \exp\left(-t \sum_{j \in J} b_j\right) \right\} t^{r-1} \right| dt \; < \infty,
\end{aligned}$$

όπου έχουμε χρησιμοποιήσει την (3.11). Το άθροισμα εκτείνεται επί όλων των 2^{j_0-1} υποσυνόλων J του $\{1, ..., j_0 - 1\}$, ενώ με $|J|$ δηλώνουμε τον πληθικό αριθμό του συνόλου J.

(ii) Εφόσον $a_j = O(b_j)$, υπάρχει μία θετική σταθερά M και ένας ακέραιος j_0, έτσι ώστε $a_j \leq M b_j$, $\forall\, j \geq j_0$. Από την πρόταση (i) του παρόντος Θεωρήματος έχουμε

$$|H_N(M\beta; r) - H_N(\alpha; r)| \leq C, \quad \text{για κάποια θετική σταθερά } C \text{ καθώς } N \to \infty.$$

Με την βοήθεια της (10.14) λαμβάνουμε

$$\left| \frac{1}{M^2} H_N(\beta; r) - H_N(\alpha; r) \right| \leq C, \quad \text{ι.ε.} \quad H_N(\beta; r) \leq M^2 H_N(\alpha; r) + CM^2,$$

και το ζητούμενο έπεται άμεσα από τον ορισμό του O (big O notation).

(iii) Σταθεροποιούμε πρώτα ένα $\epsilon > 0$. Τότε $a_j \leq \epsilon b_j$, $\forall\ j \geq j_0(\epsilon)$. Επομένως, από την πρόταση (i) υπάρχει $M = M(\epsilon)$ έτσι ώστε

$$H_N(\epsilon\beta; r) - H_N(\alpha; r) \leq M.$$

Από την (10.14) έχουμε

$$\frac{1}{\epsilon^r} H_N(\beta; r) \leq H_N(\alpha; r) + M, \quad \forall\ N \geq N_0(\epsilon).$$

Εάν διαιρέσουμε την προηγούμενη ανισότητα με $H_N(\alpha)$ και κατόπιν επιτρέψουμε: $N \to \infty$, η (iv) έπεται άμεσα, καθώς το ϵ είναι αυθαίρετο και $\lim_N H_N(\alpha) = \infty$.

(iv) Επειδή $a_j \asymp b_j$, από την (10.33) είναι $a_j = O(b_j)$ και $b_j = O(a_j)$. Από την πρόταση (ii) έχουμε καθώς $N \to \infty$, $H_N(\beta) = O\left(H_N(\alpha)\right)$ και $H_N(\alpha) = O\left(H_N(\beta)\right)$. Το ζητούμενο έπεται τώρα, πάλι από την (10.33).

Για να αποδείξουμε την πρόταση (v) σταθεροποιούμε πρώτα ένα $\epsilon > 0$. Τότε,

$$(1-\epsilon)b_j \leq a_j \leq (1+\epsilon)b_j, \quad \forall\ j \geq j_0(\epsilon).$$

Τότε, από την πρόταση (i) υπάρχει $M = M(\epsilon)$, έτσι ώστε

$$H_N((1+\epsilon)\,\beta; r) - M \leq H_N(\alpha; r) \leq H_N((1+\epsilon)\,\beta; r) + M, \quad \forall\ N \geq N_0(\epsilon),$$

και εμπλέκοντας την (10.14) έχουμε άμεσα

$$\left(\frac{1}{1+\epsilon}\right)^r H_N(\beta; r) - M \leq H_N(\alpha; r) \leq \left(\frac{1}{1-\epsilon}\right)^r H_N(\beta; r) + M,$$

$\forall\ \ N \geq N_0(\epsilon)$.
Διαιρούμε τώρα την τελευταία ανισότητα με $H_N(\beta; r)$ και αφήνουμε το $N \to \infty$, καταλήγοντας στην (v), εφόσον το ϵ είναι αυθαίρετο, και $\lim_N H_N(\beta; r) = \infty$. ■

Ένα παράδειγμα του Θεωρήματος 10.5.2 για την περίπτωση όπου $a_j = 1, \forall\ j$, δίδεται στο Παράδειγμα 1 του ακόλουθου κεφαλαίου.

Παρατήρηση 17. Ο σκοπός αυτού του κεφαλαίου ήταν να αναπτύξει τεχνικές με σκοπό τον υπολογισμό του πρωτεύοντος όρου επί του ασυμπτωτικού αναπτύγματος όλων των (ανοδικών) ροπών $E[\,T_N^{(r)}\,]$ της τυχαίας μεταβλητής T_N καθώς $N \to \infty$. Έχουμε ήδη αναφερθεί στην εργασία του H.J. Godwin [50], στην περίπτωση που οι πιθανότητες εκλογής των κουπονιών ακολουθούν ομοιόμορφη κατανομή. Εξ όσων γνωρίζουμε, δεν υπάρχουν (άλλες) δημοσιευμένες εργασίες σχετικά με την ασυμπτωτική συμπεριφορά των ανοδικών ροπών για τις τιμές $r \geq 3$, και στην γενική περίπτωση των άνισων πιθανοτήτων. Φυσικά, στην υπάρχουσα βιβλιογραφία υπάρχει πληθώρα εργασιών σχετικά με την ασυμπτωτική συμπεριφορά της μέσης τιμής $E[\,T_N\,]$ και επίσης, μερικές εργασίες σχετικά με την συμπεριφορά της δεύτερης ροπής $E[\,T_N^2\,]$ και της διασποράς $V[\,T_N\,]$. Για μερικές αντιπροσωπευτικές εργασίες βλέπε στο Κεφάλαιο 2.
Η περίπτωση $r = 1, 2$ εξετάστηκε στην εργασία [29] και συνιστά ένα εκ των κυρίων αποτελεσμάτων της παρούσης διατριβής (Κεφάλαια 7–9).

Πρόσφατα, οι J. Du Boisberranger, D. Gardy, and Y. Ponty, [10] θεώρησαν το word collector problem, i.e. the expected number of calls to a random weighted generator before all the words of a given length in a language are generated. Όπως χαρακτηριστικά αναφέρουν στην εργασία αυτή the main ingredient of this instance of the non-uniform coupon collector lies in the, potentially large, multiplicity of the words (coupons) of a given probability (composition). Ειδικότερα, παρουσίασαν ένα γενικό θεώρημα που δίδει ένα αποτέλεσμα ασυμπτωτικά ισοδύναμο με την μέση τιμή του γενικού προβλήματος Coupon Collector στην περίπτωση $r = 1$. Το θεώρημα αυτό είναι ιδιαιτέρως κατάλληλο για κλάσεις κουπονιών που εμφανίζουν υψηλό αριθμό πολλαπλότητας στην εμφάνισή τους. Τα αποτελέσματα αυτά είναι συμπληρωματικά ως προς αυτά της εργασίας [12].

Μία ενδιαφέρουσα περίπτωση για τις εφαρμογές[3] είναι όταν έχουμε ένα υποσύνολο από κουπόνια (subcollection of coupons) το οποίο μεγαλώνει (καθώς το N μεγαλώνει) με την ιδιότητα όλα τα κουπόνια αυτού του υποσυνόλου έχουν την ίδια πιθανότητα. Μπορούμε να μοντελοποιήσουμε την περίπτωση αυτή με μία ακολουθία $\{a_j\}_{j=1}^{\infty}$ η οποία είναι μία *ένωση* δύο υπακολουθιών, μία από τις οποίες είναι σταθερή (αυτή αντιστοιχεί στην ομοιόμορφη υπο-συλλογή κουπονιών, ενώ η δεύτερη είναι του τύπου που συζητήθηκε στην υποενότητα 10.5.1. Στην περίπτωση αυτή, πιστεύουμε, ότι το Θεώρημα 10.5.1 ισχύει (ίσως υπό μία κατάλληλη μετονομασία του δείκτη) δεδομένου ότι η *πυκνότητα* της σταθερής υπακολουθίας είναι επαρκώς μικρή. Με άλλα λόγια η ομοιοόμορφη υπο-συλλογή δεν επηρεάζει την ασυμπτωτική κατανομή.

Επιπρόσθετα, εάν η ακολουθία $\{a_j\}_{j=1}^{\infty}$ είναι η *ένωση* δύο φθινουσών υπακολουθιών με όριο το μηδέν, αλλά η μία φθίνει πολύ γρηγορότερα από την άλλη τότε, πιστεύουμε, ότι η πιο γρήγορη υπερισχύει στην ασυμπτωτική συμπεριφορά, δεδομένου ότι η *πυκνότητά* της δεν είναι πολύ μικρή. Σημειώνουμε ότι στο Παράδειγμα 3 του Κεφαλαίου 11 συζητούμε την περίπτωση όπου η ακολουθία $\{a_j\}_{j=1}^{\infty}$ φθίνει εκθετικά γρήγορα. Για την περίπτωση αυτή η ασυμπτωτική συμπεριφορά είναι δραστικά διαφορετική από την αντίστοιχη κατά την οποία η ακολουθία μεν φθίνει, πλην όμως αργά.

[3] Αυτή η περίπτωση μας υπεδείχθη κατά την διάρκεια της κρίσης του άρθρου μας με τίτλο, Asymptotics of the rising moments for the Coupon Collector's Problem, που υπεβλήθη στο περιοδικό: Electronic Journal of Probability, (2012).

Μέρος III

Παραδείγματα

Κεφάλαιο 11

Εφαρμογές επί γενικών ακολουθιών

Στο παρόν Κεφάλαιο δίδουμε διάφορα παραδείγματα μέσω των οποίων παρουσιάζονται τα αποτελέσματα των προηγούμενων κεφαλαίων και ενοτήτων.

Παράδειγμα 1. Η περίπτωση $a_j = 1$, $\forall j$, έχει ήδη αναλυθεί στο Κεφάλαιο 4 και την Ενότητα 10.2. Επιπρόσθετα, τούτη η περίπτωση παρέχει μία εφαρμογή του Θεωρήματος 10.5.2:
Εάν $\beta = \{b_j\}_{j=1}^{\infty}$ είναι μία ακολουθία έτσι ώστε $0 < \underline{\lim} b_j \leq \overline{\lim} b_j < \infty$ τότε, υπάρχουν σταθερές $C > c > 0$ και ακέραιος $j_0 > 0$ ώστε

$$cb_j \leq 1 \leq Cb_j, \qquad \forall\, j \geq j_0, \quad \text{δηλαδή, } 1 \asymp b_j.$$

Επομένως, από την πρόταση (iv) του Θεωρήματος 10.5.2,

$$H_N(\beta; r) \asymp \ln^r N.$$

Εάν επιπλέον το όριο $\lim b_j = b$ υπάρχει, τότε

$$\lim_{j\to\infty} \frac{ba_j}{b_j} = 1, \quad \text{δηλαδή, } ba_j \sim b_j.$$

Τότε, από την πρόταση (v) του Θεωρήματος 10.5.2,

$$H_N(\beta; r) \sim H_N(b\alpha; r).$$

Από την σχέση (10.14), παίρνουμε ότι για κάθε θετικό ακέραιο r

$$H_N(\beta; r) \sim b^{-r} \ln^r N.$$

Ειδικότερα, εάν $r = 2$, έχουμε

$$H_N(\beta; 2) = Q_N(\beta) \sim b^{-2} \ln^2 N.$$

Παράδειγμα 2. $a_j = j^p$, όπου $p > 0$. Στην περίπτωση αυτή

$$L_{1,p} := L_1(\alpha) = \int_0^1 \left[1 - \prod_{j=1}^{\infty} (1 - x^{j^p}) \right] \frac{dx}{x},$$

$$L_{2,p} :=L_2(\alpha) = (-2) \int_0^1 \left[1 - \prod_{j=1}^{\infty}(1 - x^{j^p}) \right] \frac{\ln x}{x}\, dx,$$

και γενικότερα,

$$L_{r,p} :=L_r(\alpha) = (-1)^{r-1}\, r\, \int_0^1 \left[1 - \prod_{j=1}^{\infty}(1 - x^{j^p}) \right] (\ln x)^{r-1} \frac{dx}{x},$$

(σημειώνουμε ότι η ποσότητα $L_{r,p}$ φθίνει ως προς p, για κάθε δοθέντα θετικό ακέραιο r).
Είναι εύκολο να δει κανείς ότι

$$L_{1,p} < \infty,[1] \tag{11.1}$$

ώστε από την Παρατήρηση 14 θα είναι

$$0 < L_r(\alpha) < \infty,$$

για όλους τους θετικούς ακεραίους r.
Σύμφωνα με τις σχέσεις (7.3), (7.1), και (7.2) πρέπει επιπλέον να υπολογίσουμε την ποσότητα A_N. Ανακαλούμε από το Κεφάλαιο 1 τον τύπο άθροισης των Euler-Maclaurin:

$$S(N) = \sum_{m=0}^{N} f(m) \sim \int_0^N f(x)dx + C + \frac{1}{2}f(N) + \sum_{k=1}^{\infty}(-1)^{k+1} \frac{B_{k+1}}{(k+1)!} f^{(k)}(N).$$

Για την συνάρτηση $f(t) = (t+1)^p$ έχουμε $f^{(k)}(t) = 0$, για κάθε θετικό ακέραιο $k \geq p+1$. Συνεπώς, εάν p είναι ένας **θετικός ακέραιος**, προκύπτει εύκολα το **πλήρες** ασυμπτωτικό ανάπτυγμα της ποσότητας $A_N = \sum_{n=1}^{N} n^p$, με την μορφή ενός πολυωνύμου του N, βαθμού $p+1$ δηλαδή

$$A_N = \sum_{n=1}^{N} n^p = \frac{N^{p+1}}{p+1} \left[1 + O\left(\frac{1}{N}\right) \right].[2]$$

Επομένως,

$$E\left[T_N^{(r)} \right] = \frac{N^{r(p+1)}}{(p+1)^r} L_{r,p} \left[1 + o(1) \right].$$

[1] Ένας τρόπος να αποδειχθεί η (11.1), είναι με το ολοκληρωτικό κριτήριο. Θεωρούμε την συνάρτηση: $f(x) = e^{x^p \ln \xi}$, $\xi \in (0,1)$, $p > 0$, η οποία είναι συνεχής, μη αρνητική και φθίνουσα για κάθε x. Με την αντικατάσταση $x^p \ln \xi = -y$ έχουμε

$$\int_1^{\infty} e^{x^p \ln \xi} dx = \frac{1}{p} \left(-\ln \xi\right)^{-1/p} \int_{-\ln \xi}^{\infty} e^{-y} y^{\frac{1}{p}-1} dy = O(1).$$

(Σημειώνουμε ότι εάν $p \geq 1$, το κριτήριο λόγου δίδει άμεσα το ζητούμενο.)

[2] **Παρατήρηση 18.** Το αποτέλεσμα αυτό είναι σε συμφωνία με την γνωστή ταυτότητα:

$$\sum_{n=1}^{N} n^p = 1^p + 2^p + \cdots + N^p = \sum_{i=1}^{p} S\left(p, i\right) \binom{N+1}{i+1}\, i!\,,$$

όπου $S\left(p,i\right)$ είναι ο αριθμός των τρόπων διαμέρισης ενός συνόλου $p-$ αντικειμένων σε i μη κενά υποσύνολα (blocks), δηλαδή οι αριθμοί Stirling δευτέρου είδους.

Ειδικότερα,

$$E\left[\,T_N\,\right] = A_N E_N(\alpha) = \frac{N^{p+1}}{p+1} L_{1,p}\left[1+o(1)\right],$$

$$E\left[T_N(T_N+1)\right] = \frac{N^{2(p+1)}}{(p+1)^2} L_{2,p}\left[1+o(1)\right].$$

Ώστε

$$V\left[\,T_N\,\right] = \frac{N^{2(p+1)}}{(p+1)^2}\left(L_{2,p} - L_{1,p}^2\right)\left[1+o(1)\right].$$

★ Η περίπτωση $p=1$ είναι γνωστή ως η *γραμμική περίπτωση* (*linear case*), και παρουσιάζει ιδιαίτερο ενδιφέρον.
Από τον τύπο του **Euler** για τους πεντάγωνους αριθμούς (**Euler's pentagonal-number formula**),

$$\prod_{j=1}^{\infty}\left(1-x^j\right) = 1 + \sum_{k=1}^{\infty}(-1)^k\left[x^{\omega(k)} + x^{\omega(-k)}\right], \tag{11.2}$$

όπου [3]

$$\omega(k) = \frac{3k^2-k}{2}, \qquad k = 0, \pm 1, \pm 2, ...$$

Μία αξιοθαύμαστη απόδειξη η οποία, μάλιστα, χρησιμοποιεί συνδυαστικά επιχειρήματα οφείλεται στον F. Franklin[4] και δημοσιεύτηκε το 1881. Βλέπε, π.χ.,[5].
• Για γενικές τιμές του r, ο αριθμός L_r της (10.16) είναι

$$L_r = (-1)^r\, r \sum_{k=1}^{\infty}(-1)^k\left[\int_0^1 x^{\omega(k)-1}(\ln x)^{r-1}\,dx + \int_0^1 x^{\omega(-k)-1}(\ln x)^{r-1}\,dx\right].$$

Με επανάληψη της ολοκλήρωσης κατά παράγοντες έχουμε

$$\begin{aligned} L_r &= r!\,\sum_{k=1}^{\infty}(-1)^{k+1}\left\{\frac{1}{\omega(k)^r} + \frac{1}{\omega(-k)^r}\right\} \\ &= 2^r\, r!\,\sum_{k=1}^{\infty}(-1)^{k+1}\left[\frac{1}{(3k^2-k)^r} + \frac{1}{(3k^2+k)^r}\right]. \end{aligned} \tag{11.3}$$

Για κάθε θετικό ακέραιο r μπορεί κανείς να αναλύσει την ποσότητα της παραπάνω αγκύλης σε κλάσματα που συνδέονται με γνωστά αθροίσματα. Οι περιπτώσεις $r=1,2$ υπολογίζονται αναλυτικά στην συνέχεια του παραδείγματος.
Επιπλέον, μπορούμε να έχουμε ένα άνω φράγμα για το σφάλμα

$$\Delta_r(N) = L_r(\alpha) - H_N(\alpha; r)$$

[3] Οι αριθμοί $\omega(n) = (3n^2-n)/2$ και $\omega(-n) = (3n^2+n)/2$, $n = 1,2,3,...$ είναι γνωστοί ως πεντάγωνοι αριθμοί (*pentagonal numbers*) και σχετίζονται με τα πεντάγωνα, βλέπε π.χ., [5]. Οι αριθμοί αυτοί είναι επίσης τα μερικά αθροίσματα των όρων της αριθμητικής προόδου $1, 4, 7, 10, 13, \cdots, 3n+1, \cdots$. Η ταυτότητα (11.2) έπεται άμεσα και από την ταυτότητα τριπλού γινομένου (Jacobi's triple product identity): $\prod_{m=1}^{\infty}\left(1-q^{2m}\right)\left(1-q^{2m-1}z^2\right)\left(1-q^{2m-1}z^{-2}\right) = \sum_{n=-\infty}^{\infty} q^{n^2} z^{2n}$, για $q = x^{3/2}$ και $z^2 = -x^{1/2}$.

[4] Ο Hans Rademacher αναφερόμενος στην απόδειξη αυτή την αποκάλεσε: *the first major achievement of American mathematics.*

ως ακολούθως. Για κάθε $\lambda \in (0,1)$, έχουμε

$$\Delta_r(N) = (-1)^{r-1}\, r \int_0^{1-N^{-\lambda}} \left[-\prod_{j=1}^{N}(1-x^j) + \prod_{j=1}^{\infty}(1-x^j) \right] (\ln x)^{r-1}\, \frac{dx}{x}$$

$$+ (-1)^{r-1}\, r \int_{1-N^{-\lambda}}^{1} \prod_{j=1}^{N}\left(1-x^j\right) \left[1 - \prod_{j=N+1}^{\infty}\left(1-x^j\right) \right] (\ln x)^{r-1}\, \frac{dx}{x}$$

$$\leq (-1)^{r}\, r \int_0^{1-N^{-\lambda}} \left[-\prod_{j=1}^{N}\left(1-x^j\right) + 1 + \sum_{k=1}^{N}(-1)^k \left[x^{\omega(k)} + x^{\omega(-k)}\right] \right] (\ln x)^{r-1}\, \frac{dx}{x}$$

$$+ (-1)^{r}\, r \int_0^{1-N^{-\lambda}} \left[\sum_{k=N+1}^{\infty} (-1)^k \left[x^{\omega(k)} + x^{\omega(-k)}\right] \right] (\ln x)^{r-1}\, \frac{dx}{x}$$

$$+ (-1)^{r-1}\, r \int_{1-N^{-\lambda}}^{1} 1 \cdot 1 \cdot (\ln x)^{r-1}\, \frac{dx}{x}$$

$$= (-1)^{r-1}\, r \int_0^{1-N^{-\lambda}} \left[\prod_{j=1}^{N}\left(1-x^j\right) - 1 - \sum_{k=1}^{N}(-1)^k \left[x^{\omega(k)} + x^{\omega(-k)}\right] \right] (\ln x)^{r-1}\, \frac{dx}{x}$$

$$+ O\left(N^{-r\lambda}\right),$$

όπου έχουμε χρησιμοποιήσει ότι $\ln(1-N^{-\lambda}) \sim N^{-\lambda},\ N \to \infty$. Είναι γνωστό ότι η ποσότητα εντός της αγκύλης στο τελευταίο ολοκλήρωμα είναι φραγμένη από τον όρο Nx^{N+1} (βλέπε την απόδειξη του Θεωρήματος 14.3 στο σύγγραμμα [5]). Συνεπώς,

$$\Delta_r(N) \leq (-1)^{r-1}\ rN \int_0^{1-N^{-\lambda}} x^N (\ln x)^{r-1}\, dx + O\left(N^{-r\lambda}\right).$$

Με επαναλαμβανόμενη εφαρμογή της ολοκλήρωσης κατά παράγοντες (δηλαδή με την βοήθεια της (10.19)) έχουμε

$$\Delta_r(N) \leq \frac{(-1)^{r-1}\, rN}{N+1} \left(1 - \frac{1}{N^\lambda}\right)^{N+1}$$

$$\times \sum_{k=0}^{r-1} (-1)^k\, (r-1)_k\, \frac{1}{(N+1)^k} \left[\ln\left(1 - \frac{1}{N^\lambda}\right)\right]^{r-1-k} + O\left(N^{-r\lambda}\right).$$

Επομένως,

$$\Delta_r(N) = O\left(N^{-r\lambda}\right), \quad \forall\ \lambda \in (0,1). \tag{11.4}$$

Σημειώνουμε ότι το γενικό φράγμα σφάλματος που δόθηκε στην (10.21) δίδει ($\forall$ $r \neq 1$),

$$\Delta_r(N) \leq r! \sum_{j=N+1}^{\infty} \frac{1}{j^r} = O\left(\frac{1}{N^{r-1}}\right).$$

Επομένως, για $\lambda > (r-1)/r$, η (11.4) παρέχει ένα καλύτερο φράγμα, από αυτό της (10.21).

• Για την συνέχεια θα υπολογίσουμε τα $L_1,\ L_2$. Συγκεκριμένα θα δείξουμε ότι

$$L_1 = \sum_{k=1}^{\infty} \frac{12(-1)^{k+1}}{9k^2-1} = \frac{4\pi\sqrt{3}}{3} - 6 \cong 1.2552, \tag{11.5}$$

$$L_2 = 4(54 - 8\pi\sqrt{3} - \pi^2) \cong 2.39684. \tag{11.6}$$

Για $r = 1$, η (11.3) έχει την μορφή

$$\begin{aligned} L_1 &= 2\sum_{k=1}^{\infty}(-1)^{k+1}\left[\frac{1}{3k^2-k} + \frac{1}{3k^2+k}\right] \\ &= 12\sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{9k^2-1}. \end{aligned}$$

Αναλύουμε σε απλά κλάσματα την ανωτέρω ποσότητα και έχουμε

$$L_1 = -6\left(S_1 - S_2\right), \tag{11.7}$$

όπου

$$S_1 := \sum_{k=1}^{\infty}(-1)^{k+1}\frac{1}{3k+1} = \frac{1}{4} - \frac{1}{7} + \frac{1}{10} - \cdots \tag{11.8}$$

και

$$S_2 := \sum_{k=1}^{\infty}(-1)^{k+1}\frac{1}{3k-1} = \frac{1}{2} - \frac{1}{5} + \frac{1}{8} - \cdots . \tag{11.9}$$

Υπολογίζουμε την σειρά S_1. Θεωρούμε την συνάρτηση

$$f : (-1,1) \to \mathbb{R}, \quad f(x) := \frac{1}{1+x^3}.$$

Είναι

$$\frac{1}{1+x^3} = 1 - x^3 + x^6 - x^9 + \cdots, \qquad -1 < x < 1.$$

Επομένως,

$$\int_0^x \frac{dt}{1+t^3} = x - \frac{x^4}{4} + \frac{x^7}{7} - \frac{x^{10}}{10} + \cdots, \qquad -1 < x < 1.$$

Από το κριτήριο Leibniz για εναλλάσουσες σειρές, η δυναμοσειρά

$$x - \frac{x^4}{4} + \frac{x^7}{7} - \frac{x^{10}}{10} + \cdots, \quad \text{συγκλίνει, για } x = 1.$$

Από το Θεώρημα του Abel για δυναμοσειρές, έχουμε

$$\lim_{x\to 1^-}\int_0^x \frac{dt}{1+t^3} = 1 - \frac{1}{4} + \frac{1}{7} - \frac{1}{10} + \cdots .$$

Άρα,

$$\lim_{x\to 1^-}\frac{1}{3}\left[\ln\frac{x+1}{\sqrt{x^2-x+1}}+\sqrt{3}\left(\arctan\frac{2x-1}{\sqrt{3}}-\arctan\frac{-1}{\sqrt{3}}\right)\right]=1-\frac{1}{4}+\frac{1}{7}-\frac{1}{10}+\cdots$$

ώστε,

$$1-\frac{1}{4}+\frac{1}{7}-\frac{1}{10}+\cdots=\frac{1}{3}\left(\ln 2+\frac{\pi}{\sqrt{3}}\right). \tag{11.10}$$

Συνδυάζοντας τις (11.10) και (11.8) έχουμε

$$S_1=1-\frac{1}{3}\left(\ln 2+\frac{\pi}{\sqrt{3}}\right). \tag{11.11}$$

Παρατηρούμε τώρα ότι

$$\left(\frac{1}{3}-\frac{1}{6}+\frac{1}{9}-\frac{1}{12}+\frac{1}{15}+\cdots\right)-(S_1+S_2)=-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\frac{1}{5}-\frac{1}{6}+\frac{1}{7}-\frac{1}{8}+\frac{1}{9}-\frac{1}{10}-\cdots$$

και επιπλέον, ότι

$$\frac{1}{3}-\frac{1}{6}+\frac{1}{9}-\frac{1}{12}+\frac{1}{15}+\cdots=\frac{1}{3}\left(1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\frac{1}{5}-\cdots\right)=\frac{1}{3}\ln 2.$$

Συνεπώς,

$$\frac{1}{3}\ln 2-(S_1+S_2)=\ln 2-1. \tag{11.12}$$

Από τις (11.11) και (11.12) είναι

$$S_2=\frac{1}{3}\left(\frac{\pi}{\sqrt{3}}-\ln 2\right). \tag{11.13}$$

Από τις σχέσεις (11.11), (11.13), και (11.7), συνάγουμε την (11.5).
Με παρόμοιο τρόπο θα υπολογίσουμε τον αριθμό L_2, επιβεβαιώνοντας αφενός μεν ότι $L_2<\infty$, αφετέρου ότι $L_2-L_1^2>0$. Για $r=2$, έχουμε από την (11.3) αναλύοντας σε απλά κλάσματα

$$\begin{aligned}L_2&=8\sum_{k=1}^{\infty}(-1)^{k+1}\left[\frac{2}{k^2}+\frac{9}{(3k-1)^2}+\frac{9}{(3k+1)^2}+\frac{18}{3k+1}-\frac{18}{3k-1}\right]\\&=72\sum_{k=1}^{\infty}(-1)^{k+1}\left[\frac{1}{(3k-1)^2}+\frac{1}{(3k+1)^2}\right]\\&+16\sum_{k=1}^{\infty}(-1)^{k+1}\frac{1}{k^2}-24\sum_{k=1}^{\infty}\frac{12(-1)^{k+1}}{9k^2-1}.\end{aligned} \tag{11.14}$$

Θέτουμε

$$S_3:=\sum_{k=1}^{\infty}(-1)^{k+1}\frac{1}{(3k-1)^2}=\frac{1}{2^2}-\frac{1}{5^2}+\frac{1}{8^2}-\frac{1}{11^2}+\frac{1}{14^2}-\cdots \tag{11.15}$$

και

$$S_4 := \sum_{k=1}^{\infty} (-1)^{k+1} \frac{1}{(3k+1)^2} = \frac{1}{4^2} - \frac{1}{7^2} + \frac{1}{10^2} - \frac{1}{13^2} + \frac{1}{16^2} - \cdots . \tag{11.16}$$

Παρατηρούμε ότι

$$S_3 + S_4 = 1 - \sum_{k=1}^{\infty} (-1)^{k+1} \frac{1}{k^2} + \left(\frac{1}{3^2} - \frac{1}{6^2} + \frac{1}{9^2} - \frac{1}{12^2} + \frac{1}{15^2} - \cdots \right)$$

(11.17)

$$= 1 - \sum_{k=1}^{\infty} (-1)^{k+1} \frac{1}{k^2} + \frac{1}{9} \sum_{k=1}^{\infty} (-1)^{k+1} \frac{1}{k^2} = 1 - \frac{8}{9} \sum_{k=1}^{\infty} (-1)^{k+1} \frac{1}{k^2} = 1 - \frac{2\pi^2}{27},$$ [5]

όπου χρησιμοποιήσαμε το γνωστό αποτέλεσμα

$$\sum_{k=1}^{\infty} (-1)^{k+1} \frac{1}{k^2} = \frac{\pi^2}{12}. \tag{11.18}$$

Από τις σχέσεις (11.17) και (11.5), η (11.14) αποδίδει

$$L_2 = \sum_{k \in \mathbb{Z} \setminus \{0\}} \frac{2(-1)^{k+1}}{\omega(k)^2} = 16\, \frac{\pi^2}{12} + 72 \left(1 - \frac{2\pi^2}{27} \right) - 24 \left(\frac{4\pi\sqrt{3}}{3} - 6 \right)$$

$$= 4(54 - 8\pi\sqrt{3} - \pi^2) \cong 2.39684. \tag{11.19}$$

Συνεπώς,

$$L_2 - L_1^2 = 180 - 16\pi\sqrt{3} - 28\pi^2/3 \cong 0.821323 > 0, \tag{11.20}$$

ένα αποτέλεσμα που είναι σε συμφωνία με το Θεώρημα 7.0.4.

• Υπό την (11.4), η (5.14) αποδίδει

$$E\,[\, T_N \,(T_N + 1)\,] = \left(54 - 8\pi\sqrt{3} - \pi^2 \right) N^2 \,(N+1)^2 \left[1 + O(N^{-2\lambda}) \right], \ \forall \lambda \in (0,1).$$

Επίσης, (βλέπε [12],

$$E[\, T_N \,] = \left(\frac{2\pi\sqrt{3}}{3} - 3 \right) N(N+1) \left[1 + O(N^{-\lambda}) \right], \quad \forall \lambda \in (0,1).$$

Συμπεραίνουμε λοιπόν ότι για $p = 1$ είναι

(11.21)

$$V\,[T_N\,] = \left(45 - 4\pi\sqrt{3} - \frac{7\pi^2}{3} \right) N^2 \,(N+1)^2 \, \left[1 + O(N^{-\lambda}) \right], \quad \forall \lambda \in (0,1).$$

[5] Γενικά, ισχύει η ταυτότητα (r είναι πάντα ένας θετικός ακέραιος, βλ. π.χ., [53])

$$\sum_{k=1}^{\infty} (-1)^{k+1} \left[\frac{1}{(3k-1)^r} + \frac{1}{(3k+1)^r} \right] = 1 - \left(1 - \frac{1}{3^r} \right) \sum_{k=1}^{\infty} (-1)^{k+1} \frac{1}{k^r}.$$

Παρατήρηση 19.
a. Μπορεί κανείς να εκφράσει τον αριθμό L_r ως συνάρτηση του L_{r-1}. Για παράδειγμα, εάν $r=3$, είναι

$$\begin{aligned}\frac{1}{2^3 3!}\,L_3 = &-18\sum_{k=1}^{\infty}(-1)^{k+1}\frac{1}{k^2} + 27\sum_{k=1}^{\infty}(-1)^{k+1}\left[\frac{1}{(3k-1)^3} - \frac{1}{(3k+1)^3}\right] \\ &- 81\sum_{k=1}^{\infty}(-1)^{k+1}\left[\frac{1}{(3k-1)^2} + \frac{1}{(3k+1)^2}\right] \\ &+ 162\sum_{k=1}^{\infty}(-1)^{k+1}\left[\frac{1}{(3k-1)} - \frac{1}{(3k+1)}\right].\end{aligned}$$

Από την (11.14) προκύπτει εύκολα ότι

$$\frac{1}{2^3 3!}L_3 = 27\sum_{k=1}^{\infty}(-1)^{k+1}\left[\frac{1}{(3k-1)^3} - \frac{1}{(3k+1)^3}\right] - \frac{9}{8}\,L_2.$$

Στα ανωτέρω αθροίσματα εμφανίζεται η συνάρτηση ζήτα του Hurwitz:

$$\zeta(j,\alpha) = \sum_{k=0}^{\infty}(k+\alpha)^{-j}.$$

Τέλος, από την σχέση (11.19) έχουμε

$$L_3 \cong 6.68903.$$

b. Γενικότερα, για κάθε θετικό ακέραιο $r \geq 2$ υπάρχουν σταθερές A_j, B_j εις τρόπον ώστε

$$L_r = (-1)^r\, 2^r\, r! \sum_{k=1}^{\infty}(-1)^{k+1}\left[\sum_{j=1}^{\lfloor r/2 \rfloor}\frac{A_j}{k^{2j}} + \sum_{j=1}^{r}(-1)^{-j}\frac{B_j}{(3k-1)^j} + \sum_{j=1}^{r}\frac{B_j}{(3k+1)^j}\right].$$

Τώρα η ποσότητα L_r μπορεί να υπολογιστεί αναλυτικά με την βοήθεια των γνωστών ταυτοτήτων:

$$\sum_{k=1}^{\infty}(-1)^{k+1}k^{-2j} = \left(1 - \frac{1}{2^{2j-1}}\right)\zeta(2j), \tag{11.22}$$

$$\sum_{k=1}^{\infty}(-1)^{k+1}\frac{1}{(3k-1)^j} = 6^{-j}\left[\zeta\left(j,\frac{1}{3}\right) - \zeta\left(j,\frac{5}{6}\right)\right], \tag{11.23}$$

$$\sum_{k=1}^{\infty}(-1)^{k+1}\frac{1}{(3k+1)^j} = 6^{-j}\left[\zeta\left(j,\frac{2}{3}\right) - \zeta\left(j,\frac{7}{6}\right)\right], \tag{11.24}$$

όπου $\zeta(r,\alpha)$ είναι, πάντα, η συνάρτηση ζήτα του Hurwitz.

Παράδειγμα 3. Έστω, $b_j = e^{pj}$ και $a_j = e^{-pj}$, $p > 0$. Για την ακολουθία $\beta = \{b_j\}_{j=0}^{\infty}$ έχουμε $L_r(\beta) < \infty$[6]. Επιπλέον,

$$\Delta_1(N) = E_N(\beta) - L_1(\beta) \leq \sum_{j=N+1}^{\infty} e^{-pj} = \frac{e^{-p(N+1)}}{1-e^{-p}},$$

$$\Delta_2(N) = L_2(\beta) - Q_N(\beta) \leq 2 \sum_{j=N+1}^{\infty} e^{-2pj} = \frac{2e^{-2p(N+1)}}{1-e^{-2p}},$$

και γενικότερα,,

$$\Delta_r(N) = L_r(\beta) - H_N(\beta; r) \leq r! \sum_{j=N+1}^{\infty} e^{-rpj} = \frac{r!\, e^{-rp(N+1)}}{1-e^{-rp}}. \tag{11.25}$$

Επίσης

$$B_N := \sum_{j=0}^{\infty} b_j = \frac{e^{p(N+1)}-1}{e^p - 1}.$$

Συνεπώς, από τις σχέσεις (7.3), (7.1), (7.2), και (10.23), έχουμε

$$E\,[\,T_N\,] = \frac{e^{p\,(N+1)}}{e^p-1} L_1(\beta) + O(1), \tag{11.26}$$

$$E\,[T_N\,(T_N+1)] = \left(\frac{e^{p(N+1)}}{e^p-1}\right)^2 L_2(\beta) + O(1), \tag{11.27}$$

$$V\,[\,T_N\,] = \left(\frac{e^{p\,(N+1)}}{e^p-1}\right)^2 \left(L_2(\beta) - L_1(\beta)^2\right) + e^{pN} O(1), \tag{11.28}$$

και για κάθε θετικό ακέραιο r

$$E\left[\,T_N^{(r)}\,\right] = \left(\frac{e^{p(N+1)}}{e^p-1}\right)^r L_r(\beta) + O(1). \tag{11.29}$$

★ Στην ειδική περίπτωση όπου

$$b_j = 2^j$$

(δηλαδή, $p = \ln 2$), έχουμε

$$\phi(x) := \prod_{j=0}^{\infty}(1 - x^{2^j}) = \sum_{k=0}^{\infty}(-1)^{\delta(k)} x^k, \tag{11.30}$$

όπου $\delta(k)$ είναι το πλήθος του αριθμού 1 επί του δυαδικού αναπτύγματος του k. Η ισότητα της σχέσεως (11.30) είναι άμεση. Παραπέμπουμε σχετικά και στην εργασία [15]. Κατά συνέπεια από την (6.2)

$$L_1(\beta) = \int_0^1 \left[1 - \sum_{k=0}^{\infty}(-1)^{\delta(k)} x^k\right] \frac{dx}{x} = \sum_{k=1}^{\infty} \frac{(-1)^{\delta(k)-1}}{k}.$$

[6] Επειδή $e^{pj} \geq 1 + pj$ είναι $\sum_{j=1}^{\infty} \xi^{e^{pj}} \leq \xi \sum_{j=1}^{\infty} (\xi^p)^j < \infty$.

Με την βοήθεια του λογισμικού Mathematica βρίσκουμε

$$L_1(\beta) = \int_0^1 [1-\phi(x)]\,\frac{dx}{x} \cong 1.19628.$$

Ομοίως, από την (6.3)

$$L_2(\beta) = (-2)\int_0^1 \left[1-1-\sum_{k=1}^{\infty}(-1)^{\delta(k)}x^k\right]\frac{\ln x}{x}\,dx = 2\int_0^1 \frac{\ln x}{x}\sum_{k=1}^{\infty}(-1)^{\delta(k)}x^k\,dx$$

$$= 2\sum_{k=1}^{\infty}(-1)^{\delta(k)}\left[\int_0^1 x^{k-1}\,\ln x\,dx\right] = 2\sum_{k=1}^{\infty}\frac{(-1)^{\delta(k)-1}}{k^2}.$$

Με την βοήθεια του λογισμικού Mathematica βρίσκουμε[7]

$$L_2(\beta) = (-2)\int_0^1 [1-\phi(x)]\,\frac{\ln x}{x}\,dx \cong 2.31051.$$

Μάλιστα, και σε συμφωνία με το Θεώτημα 7.0.4

$$L_2(\beta) - L_1(\beta)^2 \cong 0.879418 > 0. \tag{11.31}$$

Η διασπορά $V\,[\,T_N\,]$ εκτιμάται τώρα εύκολα από την σχέση (7.6) τηε Παρατήρησης 5.

Στην γενική περίπτωση και για κάθε θετικό ακέραιο r είναι
(11.32)

$$L_r(\alpha) = (-1)^r\,r\sum_{k=1}^{\infty}(-1)^{\delta(k)}\left[\int_0^1 x^{k-1}\,(\ln x)^{r-1}\,dx\right] = r!\,\sum_{k=1}^{\infty}\frac{(-1)^{\delta(k)-1}}{k^r}.$$

★ Τώρα, όσο αφορά στην ακολουθία $\alpha = \{a_j\}_{j=0}^{\infty}$ έχουμε ηαε $L(\alpha) = \infty$[8]. Επιπροσθέτως, η συνάρτηση $f(x) = e^{px}$ **ΔΕΝ** ικανοποιεί την συνθήκη (ii) της (8.3). Συνεπώς τα Θεωρήματα 8.1.1, 8.2.5, 8.3.3, 10.5.1 και, (φυσικά), 8.4.1 δεν δύνανται να εφαρμοστούν. Ωστόσο,

$$A_N := \sum_{j=0}^{N} a_j = \frac{1-e^{-p(N+1)}}{1-e^{-p}} \tag{11.33}$$

και

$$H_N(\alpha;r) \overset{(11.34)}{=} H_N\left(e^{-pN}\beta\right)$$

$$\overset{(10.14)}{=} e^{rpN}H_N(\beta;r) \overset{(11.25)}{=} e^{rpN}\left[L_r(\beta)+O(e^{-rpN})\right] = e^{rpN}L_r(\beta)+O(1).$$

Επομένως,

$$E\left[\,T_N^{(r)}\,\right] = \left(\frac{1-e^{-p(N+1)}}{1-e^{-p}}\right)^r\left(e^{rpN}L_r(\beta)+O(1)\right) = \left(\frac{e^{p(N+1)}}{e^p-1}\right)^r L_r(\beta)+O(1).$$

[7] Το Mathematica χρειάζεται δέκα μόνον όρους για να δώσει σταθεροποιημένες απαντήσεις και για τα δύο αυτά αποτελέσματα.

[8] $\sum_{j=1}^{\infty}\xi^{e^{-pj}} > \sum_{j=1}^{\infty}\xi = \infty$.

Ως εκ τούτου, για την ακολουθία α, η εκτίμηση για τις ροπές κάθε τάξεως $E\left[T_N^{(r)}\right]$ θα είναι η ίδια με την αντίστοιχη που περιγράφεται στην σχέση (11.29). Μέ άλλα λόγια, για την ακολουθία α, η εκτιμήσεις για την μέση τιμή $E[\,T_N\,]$, αλλά και την δεύτερη (ανοδική) ροπή $E[\,T_N(T_N+1)\,]$ θα είναι οι ίδιες με τις αντίστοιχες της ακολουθίας β (βλέπε την (11.26)).
Στην πραγματικότητα, οι ακολουθίες β και α παράγουν τις ίδιες (!) πιθανότητες κουπονιών. Τούτο έπεται από το γεγονός ότι θέτοντας $c_N := e^{pN}$, για κάθε N, έχουμε
(11.34)
$$\{b_j\,:\,0\le j\le N\}=\left\{1,e^p,e^{2p},\cdots,e^{p(N-1)},e^{pN}\right\}=\{c_Na_j\,:\,0\le j\le N\}\,,$$

δηλαδή, οι ακολουθίες είναι, τελικά, ίδιες. Άρα, η εκτίμηση της διασποράς $V\,[\,T_N\,]$ θα είναι η ίδια για τις β και (11.28).
Σημειώνουμε ότι το Θεώρημα 8.4.1, αν και δεν εφαρμόζεται στην περίπτωση της ακολουθίας α, διατηρεί την τάξη μεγέθους της διασποράς $V\,[\,T_N\,]$ πολλαπλασιασμένη επί μία σταθερά. Προς πίστωσιν του επιχειρήματος τούτου, στην υποθετική περίπτωση όπου το θεώρημα ίσχυε, θα είχαμε

$$V\,[\,T_N\,]\sim\frac{\pi^2}{6}\left(\frac{e^{p(N+1)}-1}{e^p-1}\right)^2,\qquad N\to\infty, \tag{11.35}$$

απάντηση που συμφωνεί με την (11.28), ως προς την τάξη μεγέθους, αλλά όχι ως προς την σταθερά (σύγκρινε με την (11.31)).

Παράδειγμα 4. Έστω $a_j=1/j^p$, $p>0$. Αυτή είναι η γνωστή και ως γενικευμένη κατανομή Zipf (**generalized Zipf law**). Σημειώνουμε ότι στο Κεφάλαιο 2 αναφερθήκαμε στην περίπτωση $p=1$ - που αντιστοιχεί στην συνήθη **Zipf** κατανομή, η οποία μεταξύ άλλων έχει βρει εφαρμογές και στην γλωσσολογία. Χαρακτηριστική είναι η εργασία του D. R. Mc Neil με τίτλο: Estimating an Author's Vocabulary, [77]. (Για προγενέστερα αποτελέσματα στην κατανομή αυτή ως προς την $E\,[\,T_N\,]$ βλέπε, π.χ., [41] και [12]).
Στην περίπτωση αυτή από το Θεώρημα 10.3.1 και την Παρατήρηση 14 έχουμε

$$L_r(\alpha)=\infty,\qquad r=1,2,\cdots.$$

Εάν $f(x)=x^p$, τότε η συνάρτηση $f(\cdot)$ ικανοποιεί τις υποθέσεις $(i)-(iv)$ της (8.3), και επομένως τα Θεωρήματα 8.1.1, 8.2.5, 8.3.3,και 8.4.1 μπορούν να εφαρμοστούν. Παράλληλα,

$$A_N=\frac{N^{1-p}}{1-p}+C+\frac{1}{2N^p}+O\left(\frac{1}{N^{p+1}}\right),\qquad 0<p<1, \tag{11.36}$$

$$A_N=H_N=\ln N+\gamma+\frac{1}{2N}+O\left(\frac{1}{N^2}\right),\qquad p=1 \tag{11.37}$$

και

$$A_N=\zeta(p)-\frac{1}{p-1}\frac{1}{N^{p-1}}+\frac{1}{2N^p}+O\left(\frac{1}{N^{p+1}}\right),\ \ p>1, \tag{11.38}$$

όπου $\zeta(p)$ η συνάρτηση ζήτα του Riemann.
Οι σχέσεις αυτές απρρέουν άμεσα από τον τύπο άθροισης των Euler–Maclaurin,

(βλέπε (1.5)), για την συνάρτηση $f(t) = (t+1)^{-p}$. Η σταθερά C της (1.5) είναι ο πρωτεύων όρος στο ανάπτυγμα, μόνον εάν $p > 1$. Τούτο οφείλεται στο γεγονός ότι η σειρά $\sum_{j=1}^{\infty} j^{-p}$ συγκλίνει μόνον εάν $p > 1$.
Είναι τώρα εύκολο να υπολογίσουμε την ποσότητα A_N^2. Είναι

$$A_N^2 = \left(\frac{1}{1-p}\right)^2 N^{2-2p} + \frac{2C}{1-p} N^{1-p} + \frac{1}{1-p} N^{1-2p} + O\left(1\right), \quad 0 < p < 1,$$

$$A_N^2 = H_N^2 = \ln^2 N + 2\gamma \ln N + \gamma^2 + O\left(\frac{\ln N}{N}\right), \qquad p = 1$$

και

$$A_N^2 = \zeta(p)^2 - \frac{2\zeta(p)}{p-1}\frac{1}{N^{p-1}} + \frac{\zeta(p)}{N^p} + O\left(\frac{1}{N^{p+1}}\right), \qquad p > 1,$$

Επιπλέον, από τις σχέσεις (8.44) και (8.30) έχουμε αντίστοιχα

$$\delta = \frac{1}{\ln\left(N/p\right)} \quad \text{και} \quad \omega(N) = -\left(\frac{p+1}{p}\right).$$

Συνεπώς τα Θεωρήματα 8.2.5 και 8.3.3 αποδίδουν

(*i*) Για $0 < p < 1$,

$$\begin{aligned} E\left[\,T_N\,\right] &= \frac{N \ln\left(N/p\right)}{1-p} - \frac{N \ln\left(\ln\left(N/p\right)\right)}{1-p} + \frac{\gamma N}{1-p} \\ &\quad + \frac{N \ln\left(\ln\left(N/p\right)\right)}{\left(1-p\right)\ln\left(N/p\right)} - \frac{\gamma p + p + 1}{p\left(1-p\right)}\frac{N}{\ln\left(N/p\right)} + O\left(N\frac{\ln^2\left(\ln N\right)}{\ln^2 N}\right), \end{aligned}$$

και

$$\begin{aligned} E\left[\,T_N\left(T_N+1\right)\right] &= \frac{N^2 \ln^2\left(N/p\right)}{\left(1-p\right)^2} - \frac{2N^2 \ln\left(\ln\left(N/p\right)\ln\left(N/p\right)\right)}{\left(1-p\right)^2} + \frac{2\gamma N^2 \ln\left(N/p\right)}{\left(1-p\right)^2} \\ &\quad + \frac{N^2 \ln^2\left(\ln\left(N/p\right)\right)}{\left(1-p\right)^2} - \frac{2\left(\gamma-1\right)N^2 \ln\left(\ln\left(N/p\right)\right)}{\left(1-p\right)^2} \\ &\quad + \left(\frac{6\gamma^2 p - 12\gamma p + \pi^2 p - 12p - 12}{6p\left(1-p\right)^2}\right) N^2 + O\left(N^2\frac{\ln^2\left(\ln N\right)}{\ln^2 N}\right), \end{aligned}$$

(*ii*) για $p = 1$,

$$\begin{aligned} E\left[\,T_N\,\right] &= N \ln^2 N - N \ln N \ln\left(\ln N\right) + 2\gamma N \ln N + \left(1-\gamma\right) N \ln\left(\ln N\right) \\ &\quad + \left(\gamma^2 - \gamma - 2\right) N + O\left(N\frac{\ln^2\left(\ln N\right)}{\ln N}\right), \end{aligned}$$

και

$$\begin{aligned} E\left[\,T_N\left(T_N+1\right)\right] &= N^2 \ln^4 N - 2N^2 \ln^3 N \ln\left(\ln N\right) + 4\gamma N^2 \ln^3 N + N^2 \ln^2 N \ln^2\left(\ln N\right) \\ &\quad + N^2 \ln^2 N \ln^2\left(\ln N\right) + 2\left(1-3\gamma\right) N^2 \ln^2 N \ln\left(\ln N\right) \\ &\quad + \left(6\gamma^2 - 2\gamma + \frac{\pi^2}{6} - 4\right) N^2 \ln^2 N + O\left(N^2 \ln N \ln^2\left(\ln N\right)\right), \end{aligned}$$

(iii) για $p > 1$,

$$\begin{aligned} E\left[\, T_N \,\right] &= \zeta(p) N^p \ln\left(N/p\right) - \zeta(p) N^p \ln\left(\ln\left(N/p\right)\right) \\ &\quad + \zeta(p)\gamma N^p + \zeta(p) N^p \frac{\ln\left(\ln\left(N/p\right)\right)}{\ln\left(N/p\right)} \\ &\quad - \frac{\left(\gamma p + p + 1\right)\zeta(p)}{p} \frac{N^p}{\ln\left(N/p\right)} + O\left(N^p \frac{\ln^2\left(\ln N\right)}{\ln^2 N}\right), \end{aligned}$$

και

$$\begin{aligned} E\left[\, T_N\left(T_N + 1\right) \right] &= \zeta(p)^2 N^{2p} \ln^2\left(N/p\right) - 2\zeta(p)^2 N^{2p} \ln\left(\ln\left(N/p\right)\right) \ln\left(N/p\right) \\ &\quad + 2\gamma\zeta(p)^2 N^{2p} \ln\left(N/p\right) + \zeta(p)^2 N^{2p} \ln^2\left(\ln\left(N/p\right)\right) \\ &\quad - 2\left(\gamma - 1\right)\zeta(p)^2 N^{2p} \ln\left(\ln\left(N/p\right)\right) \\ &\quad + \zeta(p)^2 \left(\gamma^2 - 2\gamma + \frac{\pi^2}{6} - 2 - \frac{2}{p}\right) N^{2p} + O\left(N^{2p} \frac{\ln^2\left(\ln N\right)}{\ln^2 N}\right). \end{aligned}$$

Από τις ανωτέρω σχέσεις η ταυτότητα (3.21) δίδει το αποτέλεσμα του Θεωρήματος 8.4.1, δηλαδή

$$V\left[\, T_N \,\right] \sim \frac{\pi^2}{6}\, \frac{N^2}{\left(1-p\right)^2}, \qquad 0 < p < 1,$$

$$V\left[\, T_N \,\right] \sim \frac{\pi^2}{6}\, N^2 \left(\ln N\right)^2, \qquad p = 1, \tag{11.39}$$

και

$$V\left[\, T_N \,\right] \sim \frac{\pi^2}{6}\, \zeta(p)^2 N^{2p}, \qquad p > 1.$$

Γενικότερα, για τις ροπές ανωτέρας τάξεως έχουμε

$$A_N^r \sim \left(\frac{1}{1-p}\right)^r N^{r(1-p)}, \qquad 0 < p < 1,$$

$$A_N^r = H_N^r \sim \left(\ln N\right)^r, \qquad p = 1,$$

και

$$A_N^r \sim \zeta(p)^r, \qquad p > 1.$$

Το Θεώρημα 10.5.1 τώρα, δίδει

$$E\left[T_N^{(r)}\right] \sim \frac{N^r \left(\ln N\right)^r}{\left(1-p\right)^r} \qquad 0 < p < 1,$$

$$E\left[T_N^{(r)}\right] \sim N^r \left(\ln N\right)^{2r} \qquad p = 1,$$

$$E\left[T_N^{(r)}\right] \sim \zeta(p)^r N^{rp} \left(\ln N\right)^r \qquad p > 1.$$

Παράδειγμα 5. $a_j = j!$. Από το κριτήριο λόγου και το Θεώρημα 10.3.1 έχουμε $L_r(\alpha) < \infty$, για κάθε θετικό ακέραιο r. Εδώ, ο τύπος άθροισης Euler–Maclaurin

δεν είναι αποδοτικός. Ωστόσο, από τον τύπο του Stirling: $n! \sim (n/e)^n \sqrt{2\pi n}$ και την (7.8) έχουμε

$$A_N \sim N!.$$

Από το Θεώρημα 10.4.1 έχουμε

$$E\left[T_N^{(r)}\right] \sim L_r(\alpha)\,(N!)^r\,, \qquad N \to \infty.$$

Ειδικότερα,

$$E[\,T_N\,(T_N+1)\,] \sim L_2(\alpha)\,(N!)^2\,,$$

και (βλέπε επίσης [12]),

$$E\,[\,T_N\,] \sim L_1(\alpha) N!.$$

Συνεπώς, το Θεώρημα 7.0.3 αποδίδει

$$V\,[\,T_N\,] \sim \left(L_2(\alpha) - L_1(\alpha)^2\right)(N!)^2 \quad \text{καθώς} \quad N \to \infty.$$

Παράρτημα Α′

Η συνάρτηση Γάμμα

The world is full of obvious thing which nobody by any chance ever observes.

Sherlock Holms, *The Hound of the Baskervilles*,

Sir Arthur Conan Doyle.

Παρουσιάζουμε εδώ μερικά γνωστά αποτελέσματα σχετικά με την συνάρτηση *Γάμμα*, τα οποία χρησιμοποιήσαμε στις προηγούμενες ενότητες. Σε πολλά συγγράμματα τα αποτελέσματα αυτά αναφέρονται απλά ως ασκήσεις, (βλέπε π.χ., [54], [14]). Ο μοντέρνος ορισμός της συνάρτησης *Γάμμα* οφείλεται στον Legendre, (1809):

$$\Gamma(x) := \int_0^{\infty} e^{-t}\, t^{x-1}\, dt, \quad \Re\{x\} \geq 0. \tag{Α′.1}$$

Από την (Α′.1) παίρνουμε

$$\Gamma'(x) = \int_0^{\infty} e^{-t}\, t^{x-1} \ln t\, dt,$$

$$\Gamma''(x) = \int_0^{\infty} e^{-t}\, t^{x-1} (\ln t)^2\, dt,$$

$$\Gamma'''(x) = \int_0^{\infty} e^{-t}\, t^{x-1} (\ln t)^3\, dt.$$

Επομένως,

$$\Gamma'(1) = \int_0^\infty e^{-t} \, \ln t \, dt,$$

$$\Gamma''(1) = \int_0^\infty e^{-t} \, (\ln t)^2 \, dt,$$

$$\Gamma'''(1) = \int_0^\infty e^{-t} \, (\ln t)^3 \, dt.$$

Ειδικότερα,

$$\Gamma(1) = 1, \tag{A'.2}$$

$$\Gamma'(1) = -\gamma, \tag{A'.3}$$

$$\Gamma''(1) = \zeta(2) + \gamma^2, \quad \text{και}, \tag{A'.4}$$

$$\Gamma'''(1) = -\left(2\zeta(3) + 3\gamma\zeta(2) + \gamma^3\right), \tag{A'.5}$$

όπου $\gamma = 0.5772...$ είναι η σταθερά του Euler, ενώ με $\zeta(\cdot)$ συμβολίζουμε την συνάρτηση ζήτα του Riemann. Από την (A′.1), λαμβάνουμε άμεσα την (A′.2). Σκιαγραφούμε την απόδειξη των σχέσεων (A′.3)–(A′.5). Από την γνωστή αναπαράσταση γινομένου για την συνάρτηση *Γάμμα*, είναι:

$$\frac{1}{\Gamma(x)} = x \, e^{\gamma x} \prod_{k=1}^{\infty} \left(1 + \frac{x}{k}\right) e^{-\frac{x}{k}}. \tag{A'.6}$$

Παίρνοντας λογαρίθμους έχουμε

$$\ln \Gamma(x) = -\ln x - \gamma x - \sum_{n=1}^{\infty} \left\{ \ln\left(1 + \frac{x}{n}\right) - \frac{x}{n} \right\}. \tag{A'.7}$$

Παραγωγίζουμε δύο φορές και τα δύο μέλη της (A′.7) και από τον ορισμό της συνάρτησης *ψι* (psi function), δηλαδή της παραγώγου του λογαρίθμου της *Γάμμα* συνάρτησης, παίρνουμε

$$\psi(x) := \frac{\Gamma'(x)}{\Gamma(x)} = -\gamma - \frac{1}{x} - \sum_{n=1}^{\infty} \left\{ \left(\frac{1}{1 + \frac{x}{n}} \right) \frac{1}{n} - \frac{1}{n} \right\}, \quad x > 0 \quad \text{ανδ} \tag{A'.8}$$

$$\psi'(x) = \frac{1}{x^2} + \sum_{n=1}^{\infty} \left\{ \left(\frac{1}{\left(1 + \frac{x}{n}\right)^2} \right) \frac{1}{n^2} \right\}, \quad x > 0. \tag{A'.9}$$

Θέτοντας $x = 1$ στην (A′.9) έχουμε

$$\psi'(1) = \zeta(2). \tag{A'.10}$$

Επίσης, για $x = 1$ έχουμε από την (Α′.8)

$$\begin{aligned}
\psi(1) &= -\gamma - 1 - \sum_{n=1}^{\infty}\left(\frac{1}{n+1} - \frac{1}{n}\right) \\
&= -\gamma - 1 - \lim_{k\to\infty}\sum_{n=1}^{k}\left(\frac{1}{n+1} - \frac{1}{n}\right) \\
&= -\gamma = \Gamma'(1). \qquad \text{(Α}'\text{.11)}
\end{aligned}$$

Για την απόδειξη της (Α′.4) χρησιμοποιούμε τον ορισμό της συνάρτησης *ψι*, και τις (Α′.10), (Α′.3). Με παρόμοιο τρόπο αποδεικνύουμε και την (Α′.5).
Εναλλακτικά, μπορεί κανείς να χρησιμοποιήσει την αναδρομική σχέση (βλέπε, π.χ., [14])

$$\Gamma^{(n+1)}(1) = -\gamma\Gamma^{(n)}(1) + n!\sum_{k=1}^{n}\frac{(-1)^{k+1}}{(n-k)!}\,\zeta\,(k+1)\,\Gamma^{(n-k)}(1).$$

Βιβλιογραφία

[1] S. Abraham, G. Brockman, S. Sapp, and A. P. Godbole, Omnibus sequences, Coupon collection, and missing word counts, *arXiv:0905.4517v1 [math.PR]*, (2009).

[2] I. Adler, S. Oren and S. Ross, The coupon collector's problem revisited, *J. Appl. Prob.* **40** (2003) 513–518.

[3] I. Adler and S. Ross, The coupon subset collection, *J. Appl. Prob.*, **38** (2001) 737–746.

[4] D. Aldous, An Introduction to covering problems for random walks on graphs, *J. Theor. Prob.*, **2** (1) (1989) 87–89.

[5] T. M. Apostol, *Introduction to Analytic Number Theory*, Springer-Verlag, New York, 1976.

[6] T. M. Apostol, An Elementary View of Euler's Summation Formula, The American Mathematical Monthly **106** (5) (1999) 409–418.

[7] A. Astashkevich and I. Pak, Random walks on nilpotent groups, *preprint*, (2001).

[8] L. E. Baum and P. Billingsley, Asymptotic distributions for the Coupon Collector's problem, *Annals of Mathematical Statistics* **36** (1965) 1835–1839.

[9] C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers I: Asymptotic Methods and Perturbation Theory*, Springer-Verlag, New York, 1999.

[10] J. Du Boisberranger, D. Gardy, and Y. Ponty, The weighted words collector, *DMTCS Proc. AofA 23rd Intern. Meeting on Probabilistic, Combinatorial, and Asymptotic Methods for the Analysis of Algorithms* Montreal, Canada, 2012.

[11] A. Boneh and M. Hofri, The Coupon Collector Problem Revisited–A Survey of Engineering Problems and Computational Methods, *Comm. Statist. Stochastic Models* **13** (1) (1997) 39–66.

[12] S. Boneh and V.G. Papanicolaou, General Asymptotic Estimates for the Coupon Collector Problem, *J. Comp. Appl. Math.***67** (2) (1996) 277–289.

[13] J. Bonneau and E. Shutova, Linguistic properties of multi-word passphrases, *USEC'12: Workshop on Usable Security*, (2012) Kralendijk, Bonaire, The Netherlands, Featured in The Economist, 24 March 2012 under the title: Computer passwords Speak, friend, and enter.

[14] G. Boros, V. H. Moll, *Irresistible Integrals: symbolics, analysis and experiments in the Evaluation og Integrals*, Cambridge University Press, Cambridhe, U.K., 2004.

[15] J. M. Borwein and P. B. Borwein, Strange series and high precision fraud, *Amer. Math. Monthly* **99** (7) (1992) 622–640.

[16] R. K. Brayton, On the asymptotic behavior of the number of trials necessary to complete a set with random selection, *Journal of Mathematical Analysis and Applications* **7** (1963) 31–61.

[17] M. Brown, E. Peköz and S. Ross, Coupon collecting, *Prob. in the Engineering and Informational Sciences*, **22** (2008) 221–229.

[18] J. Bunge and M. Fitzpatrick, Estimating the number of species: A review, *J. Amer. Stat. Assoc.* **88** (421) (1993) 364–373.

[19] L. Carlitz, Solution of Problem 4582, *Amer. Math. Monthly* **62** (1955) 447–450.

[20] K. C. Chang and S. Ross, The multiple subset collecting problem, *Prob. in the Engineering and Informational Sciences*, **21** (2007) 435–440.

[21] K. L. Chung, *A Course in Probability Theory*, Third Edition, Academic Press, San Diego, CA, USA, 2001.

[22] F.N. David and D. E. Barton, *Combinatorial Chance* , Charles Griffin & Co., London, 1962.

[23] N. B. Dimitrov and C. G. Plaxton, Optimal Cover Time for a Graph-Based Coupon Collector Process, *32nd International Colloquium on Automata, Languages, and Programming ICALP*, **2** (2005) 702–716.

[24] N. G. De Bruijn, *Asymptotic Methods in Analysis*, North–Holland Publishing & Co., Amsterdam, 1961.

[25] B. Dawkins, Siobhan's problem: The Coupon collector revisited, *American Statistician*, **45** (1) (1991) 76–82.

[26] P. Diaconis and S. Holmes, Are there still things to do in Bayesian statistics?, *Erkenntnis*, **45** Numbers (2–3) (1996) 145–158.

[27] P. Diaconis and S. Holmes, A Bayesian peek into Feller volume I, *Sankhyā*, Special issue in memory of D. Basu, **64** Ser. A (3, part 2) (2002) 820–841.

[28] P. Diaconis and L. Saloff–Coste, Comparison techniques for random walk on finite groups, *Annals of Prob.*, **21** (4) (1996) 2131–2156.

[29] A. V. Doumas and V. G. Papanicolaou, The Coupon Collector's Problem Revisited: Asymptotics of the Variance, *Adv. Appl. Prob.* **44** (1) (2012) 166–195.

[30] A. V. Doumas and V. G. Papanicolaou, Asymptotics of the rising moments for the coupon collector's problem, *Electron. J. Probab.* **Vol. 18** (Article no. 41) (2012) 1–15 (doi: 10.1214/EJP.v18-1746).

[31] A. V. Doumas and V. G. Papanicolaou, The Siblings of the Coupon Collector, 2012, preprint.

[32] A. V. Doumas and V. G. Papanicolaou, Some new aspects for the Random Coupon Collector's Problem, 2012, submitted.

[33] P. Dupuis, C. Nuzman and P. Whiting, Large deviation asymptotics for occupancy problems, *Annals of Prob.*, **32** (3B) (2004) 2765–2818.

[34] R. Durrett, *Probability: Theory and Examples*, Third Edition, Duxbury Advanced Series, Brooks/Cole—Thomson Learning. Belmont, CA, USA, 2005.

[35] B. Efron and R. Thisted, Estimating the number of unsen species: How many words did Shakespeare know?, *Biometrika* **63** (3) (1976) 435–447.

[36] G. Englund, On the Coupon Collector's Remainder Term, *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **60** (1982) 381–391.

[37] P. Erdős and A. Rényi, On a classical problem of probability theory, *Magyar. Tud. Akad. Mat. Kutató Int. Kőzl.*, **6** (1961), 215–220.

[38] P. Erdős and A. Rényi, On the strength of connectedness of a random graph, *Acta Math. Acad. Sci. Hung.* **12** (1961), 261–267.

[39] W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. I, First Edition, John Wiley & Sons, Inc., New York, 1950.

[40] W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. II, Second Edition, John Wiley & Sons, Inc., New York, 1971.

[41] P. Flajolet, D. Gardy and L. Thimonier, Birthday paradox, coupon collectors, caching algorithms and self-organizing search, *Discrete Applied Mathematics* **39** (1992) 207–229.

[42] P. Flajolet and A.M. Odlyzko, Random mapping statistics, *Advances in Cryptology*, Proc. Eurocrypt"89, LNCS 434, J.-J. Quisquater and J. Vandewalle, Eds., Springer-Verlag, (1990) 329–354.

[43] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge University Press, U.K., 2002.

[44] L. Flatto, Limit Theorems for Some Random Variables Associated with Urn Models, *Annals of Prob.* **10** (4) (1982) 927–934.

[45] D. Foata, Hyperharmonic Numbers and the Phatry of the Coupon Collector, *2000-2001 Algorithms Seminal, INRIA* (2002) 19–22.

[46] D. Foata, H. Guo-Niu and B. Lass, Les nombres hyperharmonique et la fratrie du collectionneur de vignettes, *Sem. Lothar Combinatoire* **47** B47a (2001) (electronic).

[47] D. Foata and D. Zeilberger, The Collector's Brotherhood Problem using the Newman–Sheep symbolic method, *Algebra univers.* **49** (2003) 387–395.

[48] T. Gadrich and R. Ravid, The Coupon collector problem in Statistical quality control , *Proc. World Congress on Engineering, WCE* (2008), London, U.K..

[49] C. M. Goldie, R. Cornish, and C. L. Robinson, Computer Assisted Assessment: How many questions are enough?, *arXiv:math.PR/1002.2114* (2010).

[50] H. J. Goodwin, On cartophily and motor cars, *Math. Gazette* **33** (305) (1949) 169–171.

[51] Hald, A., A. de Moivre: De Mensura Sortis or On the Measurment of Chance *International Statistical Review* **52** (3) (1984) 229-262.

[52] G. H. Hardy, *Divergent Series*, Oxford University Press, London, U.K., 1949.

[53] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, Fourth Edition, Oxford University Press, London, U.K., 1960.

[54] J. Havil, *Gamma, Exploring Euler's Constant*, Princeton University Press, New Jersey, USA, 2003.

[55] M. V. Hildebrand, The Birthday problem, a notice appeared in the *Amer. Math. Monthly* **100** (7) (1993) p. 643.

[56] L. Holst, On Birthday, Collectors', Occupancy and other classical Urn problems, *International Statistical Review* **54** (1986) 15–27.

[57] L. Holst, Extreme value distributions for random coupon collector and birthday problems, *Extremes* **4** (2) (2001) 129–145.

[58] L. Holst, J. E. Kennedy, and M. P. Quine, Rates of Poisson Convergence for some Coverage and Urn Problems Using Coupling, *J. Appl. Prob.* **25** (1988) 717–724.

[59] S. Janson, Limit theorems for some sequential occupancy problems, *J. Appl. Prob.***20** (1983) 545–553.

[60] S. Janson, Probability asymptotics: notes on notation, *arXiv:1108.3924*.

[61] J. Jonasson, Biased random – to – top shuffling, *Annals Appl. Prob.*, **45** Numbers (2–3) (1996) 145–158.

[62] N. D. Kan, Martingale approach to the Coupon collection problem, *J. Math. Sci.* **127** (1) (2005) 1737–1744.

[63] N. Kaplan, A Generalization of a result of Erdős and Rényi, *J. Appl. Prob.* **14** (1) (1977) 212–216.

[64] M. Karwan, V. Lotfi, J. Telgen, and S. Zionts, *Redundancy in Mathematical Programming*, Springer-Verlag, Berlin, 1983.

[65] M. S. Klamkin, Problem 4582, *Amer. Math. Monthly* **61** (1954) p. 199.

[66] J. E. Kobza, S. H. Jacobson, and D. E. Vaughan, A survey of the Coupon Collector"s Problem with random sample sizes, *Method. Comput. Appl. Probab.*, **9** (2007) 573–584.

[67] A. Korolova, R. Motwani, S. U. Nabar, and Y. Xu, Link Privacy in Social Networks, *Proc. CIKM'08*, (2008) Napa Valley, California, USA.

[68] Y. Li, E. Soljanin, and P. Spasojević, Collecting coded coupons over generations, *arXiv:1002.1406v3 [cs.IT]*, (2010).

[69] H. M. Mahmoud and R. T. Smythe, On the joint behavior of types of coupons in generalized coupon collection, *J. Adv. Appl. Prob.*, **44** (2) (2012) 1–23.

[70] H. M. Mahmoud, *Pólya Urn Models*, Charman & Hall Press, U.S.A., 2009.

[71] F. G. Maunsell, A problem in cartophily, *Math. Gazette* **22** (251) (1938) 328–331.

[72] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and its Applications*, Academic Press, New York, 1979.

[73] R. May, Coupon collecting with quotas, *Electron. J. Combinatorics*, **15** (N31) (2008) 435–440.

[74] A. Myers and H.S. Wilfp, Some new aspects of the Coupon Collector Problem, *Siam J. Discrete Math.*, **17** (1) (2003) 1–17.

[75] T. Nakata and I.Kubo, A coupon collector"s problem with bonuses, *DMTCS Proc. AG Fourth Colloquium on Mathematics and Computer Science*, France, (2006) 215–224. ΔΜΤ"Σ προς. ΑΓ,

[76] P. Neal, The Generalised Coupon Collector Problem, *J. Appl. Prob.* **45** (2008) 621–629.

[77] D. Mc Neil, Estimating an Author's Vocabulary, *J. American Statistical Assosiation* **68** (1973) 92–96.

[78] D. J. Newman and L. Shepp, The Double Dixie Cup problem, *Amer. Math. Monthly* **67** (1960) 58–61.

[79] Luke O'Connor, Entropy Bounds for Traffic Confrmation,, *IACR, Technical Report 2008/365*, (2008).

[80] V. G. Papanicolaou, G. Kokolakis and S. Boneh, Asymptotics for the Random Coupon Collector Problem, *J. Comp. Appl. Math.***93** (1998) 95–105.

[81] N. Pintacuda, Coupons collectors via the martingales, *Bolletino Un. Mat. It.* **17A** (5) (1980) 174 – 177.

[82] A. Poon, B. H. Davis and L. Chao, The Coupon collector and the suppressor mutation: Estimating the number of compensatory mutations by maximum likelihood, *Genetics*, **170** (3) (2005) 1323–1332.

[83] A. Rafferty, T. L. Griffiths, and D. Klein, Convergence bounds for language evolution by iterated learning, *Proceed. 31st Annual Conference of the Cognitive Science Society*, (2009).

[84] K. L. Q. Read, A lognormal approximation for the Collector's Problem, *American Statistician*, **52** (2) (1998) 175–180.

[85] B. Rósen, Asymptotic normality in a coupon collector's problem, *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **12** (1969) 256–279.

[86] B. Rósen, On the coupon collector's waiting time, *Annals of Mathematical Statistics* **41** (1970) 1952–1969.

[87] S. Ross, *A First Course in Probability*, Seventh Edition, Pearson Prentice Hall, 2006.

[88] S. Ross, *Introduction to Probability Models*, Ninth Edition, Academic Press, San Diego, U.S.A., 2007.

[89] S. Ross, The Hypergeometric Coupon Collection Problem and its Dual, *J. of Industrial and Systems Engineering* **1** (1) (2007) 1–7.

[90] R. Rovatti, C. Passerini, G. Mazzini, Markov modeling of cooperative multiplayer Coupon Collectors' problems, *arXiv:0912.2523* (2009).

[91] H. Royden, *Real Analysis*, Third Edition Pearson Prentice Hall, 1988.

[92] W. Rudin, *Real and Complex Analysis*, Third Edition McGraw-Hill, New York, 1987.

[93] H. V. Schelling, Auf der Spur des Zufalls, *Deutsches Statistisches Zentralblatt* **26** (1934) 137–146.

[94] H. V. Schelling, Coupon Collecting for Unequal Probabilities, *Amer. Math. Monthly* **61** (5) (1954) 306–311.

[95] P. K. Sen, Invariance principles for the coupon collector's problem: A martingale approach, *Ann. Statist.* **7** (1979) 372–380.

[96] P. K. Sen, Limit theorems for an extended coupon collector's problem and for successive subsampling with varying probabilities, *Calcutta Statist. Assoc. Bull.* **29** (1980) 115–116, (1981) 113–132.

[97] S. Shioda, Some upper and lower bounds on the coupon collector problem, *J. Comp. Appl. Math.* **200** (2007) 154–167.

[98] S. Sidenko, Kac's random walk and Coupon Collector"s process on posets, *Ph.D. Thesis, MIT*, (2008), advisor: I. Pak.

[99] D. Stirzaker, *Elementary Probability*, Second Edition, Cambridge University Press, U.K., 2003.

[100] A. Torn and A. Zallinskas, *Global Optimization*, Lecture Notes in Computer Science, No. 350, Springer-Verlag, New York, 1989.

[101] I. Todhunter, *A History of the Mathematical Theory of Probability from the time of Pascal to that of Laplace*, Macmillan, London, 1865. Reprinted Chelsea, New York, 1949.

[102] W. Xu and A. K. Tang, A generalized coupon collector problem, *J. Appl. Probab.*, **48** (4) (2011) 1081–1094.

[103] D. Zeilberger, How many singles, doubles, triples, edtc., should the coupon collector expect?, *Personal Journal of Ekhad and Zeilberger* (2001) *http://www.math.rutgers.edu/~zeilberg/.*

[104] The Dixie Cup Company History: *http://academicmuseum.lafayette.edu/special/dixie/company.html*

[105] A Coupon Collector Problem simulation in the case of equal probabilities: *www-stat.stanford.edu/∼susan/surpise/Collector.html.*

[106] Fish of Lake Michigan, University of Wisconsin Sea Grant Institute: *http://www.seagrant.wisc.edu/greatlakesfish/LakeMichFishIndex.html.*

[107] Γ. Κοκολάκης και Ι. Σπηλιώτης, *Θεωρία Πιθανοτήτων και Στατιστική με Εφαρμογές*, Εκδ. Συμεών, Αθήνα, 2010.

[108] Γ. Κοκολάκης και Δ. Φουσκάκης, *Στατιστική – Θεωρία και Εφαρμογές*, Εκδ. Συμεών, Αθήνα, 2009.