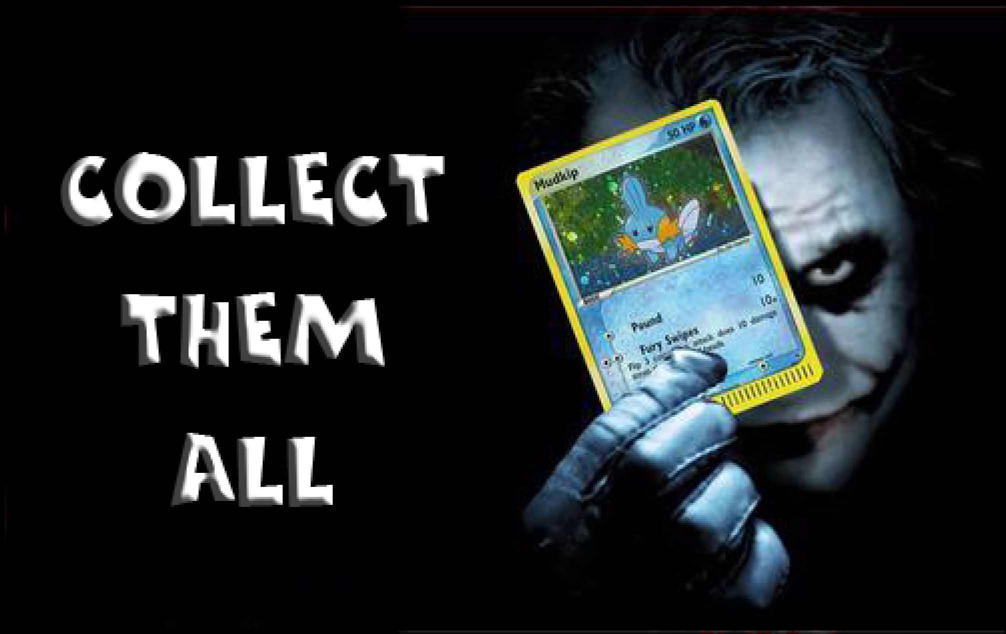
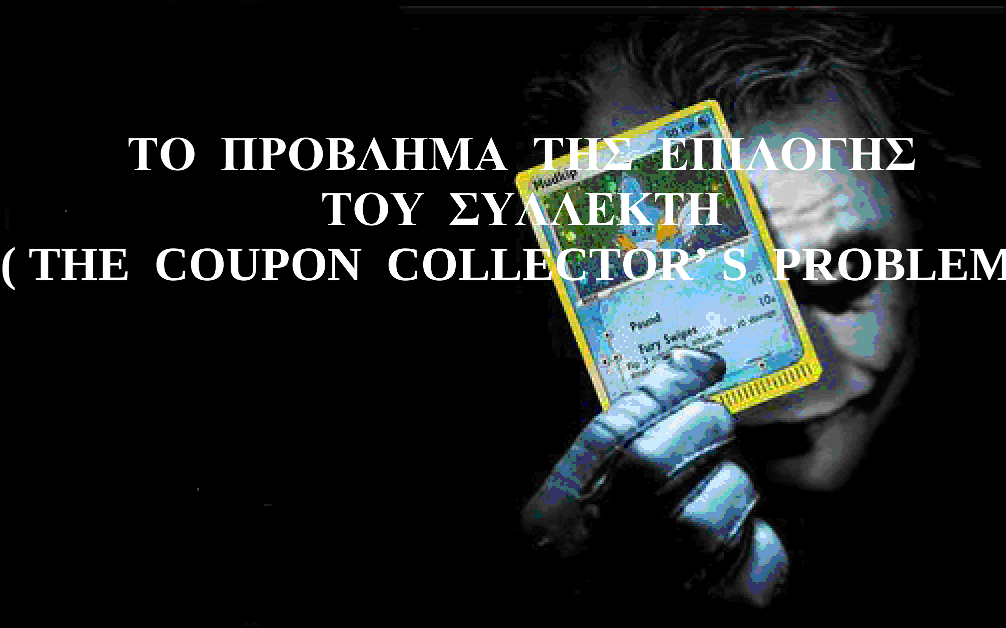
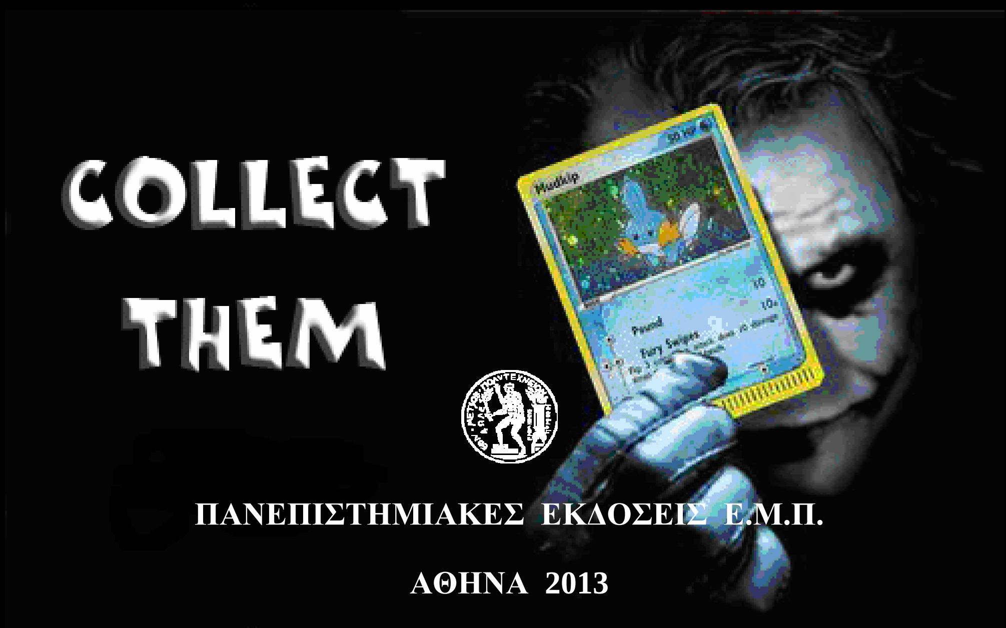
ΕΥΑΓΓΕΛΟΣ Κ. ΓΚΟΛΙΑΣ



# ΤΟ ΠΡΟΒΛΗΜΑ ΤΗΣ ΕΠΙΛΟΓΗΣ ΤΟΥ ΣΥΛΛΕΚΤΗ
# ( THE COUPON COLLECTOR’ S PROBLEM )




ΠΑΝΕΠΙΣΤΗΜΙΑΚΕΣ ΕΚΔΟΣΕΙΣ Ε.Μ.Π.

ΑΘΗΝΑ 2013

Το Πρόβλημα της Επιλογής του Συλλέκτη
( The Coupon Collector’ s Problem )

**ΕΥΑΓΓΕΛΟΣ Κ. ΓΚΟΛΙΑΣ**

# ΤΟ ΠΡΟΒΛΗΜΑ ΤΗΣ ΕΠΙΛΟΓΗΣ ΤΟΥ ΣΥΛΛΕΚΤΗ
# ( THE COUPON COLLECTOR' S PROBLEM )

**ΠΑΝΕΠΙΣΤΗΜΙΑΚΕΣ ΕΚΔΟΣΕΙΣ Ε.Μ.Π.**

**ΑΘΗΝΑ 2013**

**Το Πρόβλημα**
**της Επιλογής**
**του Συλλέκτη**
**( The Coupon Collector' s Problem)**

Διπλωματική Εργασία
**Ευάγγελου Κ. Γκόλια**

Επιβλέπων Καθηγητής
**Βασίλειος Παπανικολάου**
*Καθηγητής Ε.Μ.Π.*

Μέλη Εξεταστικής Επιτροπής

**Ιωάννης Σπηλιώτης**
*Αναπλ. Καθηγητής Ε.Μ.Π.*

**Χρυσηίς Καρώνη**
*Αναπλ. Καθηγήτρια Ε.Μ.Π.*

**Τομέας Μαθηματικών**

**Σχολή Εφαρμοσμένων Μαθηματικών και Φυσικών Επιστημών**

**Εθνικό Μετσόβιο Πολυτεχνείο**

**Αθήνα 2013**

*Στους γονείς μου*
*Κωνσταντίνο και Ζανέτα*

# Πρόλογος

Η διπλωματική εργασία με τίτλο «Το Πρόβλημα της Επιλογής του Συλλέκτη» εκπονήθηκε κατά τη διάρκεια του δέκατου εξαμήνου της Σχολής Εφαρμοσμένων Μαθηματικών και Φυσικών Επιστημών του Εθνικού Μετσόβιου Πολυτεχνείου.

Το πρόβλημα της επιλογής του συλλέκτη στην πιο απλή του μορφή έχει εμφανιστεί στο κλασσικό σύγγραμμα του Feller [8] και έχει ελκύσει την προσοχή διαφόρων ερευνητών λόγω του πλήθους των εφαρμογών του σε πολλές περιοχές της επιστήμης (επιστήμη υπολογιστών – αλγόριθμοι, μαθηματικός προγραμματισμός, βελτιστοποίηση, διαδικασίες μάθησης, μηχανική, οικολογία και γλωσσολογία).

Στο πρώτο κεφάλαιο παρουσιάζονται διάφορα μοντέλα κατανομής πιθανότητας για διακριτές ή συνεχείς τυχαίες μεταβλητές και δίνονται οι ορισμοί των ασυμπτωτικών σχέσεων. Επίσης περιγράφεται αναλυτικά η ανέλιξη Poisson, αποδεικνύεται η φόρμουλα αθροίσματος Euler – Maclaurin και αναφέρεται ο τρόπος με τον οποίο προκύπτει η σταθερά Euler.

Στις πρώτες ενότητες του δευτέρου κεφαλαίου παρουσιάζεται το πρόβλημα της επιλογής του συλλέκτη από τη σκοπιά των ανελίξεων Poisson και υπολογίζονται τα ασυμπτωτικά αναπτύγματα των ροπών πρώτης και δεύτερης τάξης στην περίπτωση των ίσων πιθανοτήτων. Στην τελευταία ενότητα αναλύεται το πρόβλημα με τη χρησιμοποίηση κατάλληλων οριακών θεωρημάτων για την εξαγωγή των αποτελεσμάτων.

Στο παράρτημα περιλαμβάνονται διάφορες αποδείξεις που παραλήφθηκαν από το κυρίως κείμενο για να μη διασπαστεί η ροή και η συνοχή του.

Στο σημείο αυτό θα ήθελα να εκφράσω τις ευχαριστίες μου προς τον καθηγητή και επιβλέποντα την παρούσα διπλωματική εργασία, κ.κ. Βασίλειο Παπανικολάου, για την αμέριστη βοήθεια που μου προσέφερε και για τις εποικοδομητικές παρατηρήσεις του καθ’ όλη τη διάρκεια της επεξεργασίας των διαφόρων ενοτήτων. Η ευρεία γνώση του επί της Θεωρίας Πιθανοτήτων συνέβαλε στην υπέρβαση όσων δυσχερειών ανέκυψαν. Ο επαγγελματισμός του αποτελεί πρότυπο για τη μελλοντική μου πορεία.

Επίσης εκφράζονται ιδιαίτερες ευχαριστίες στα άλλα δύο μέλη της εξεταστικής επιτροπής, καθηγητές κ.κ. Ιωάννη Σπηλιώτη και κ.κ. Χρυσηίδα Καρώνη, για το χρόνο που αφιέρωσαν στη μελέτη της εργασίας στην τελική της μορφή.

Τέλος, θα ήθελα να εκφράσω τις θερμές μου ευχαριστίες προς τους γονείς μου για την ακλόνητη στήριξη που μου προσέφεραν, την υπομονή και τη συμπαράστασή τους στις δύσκολες στιγμές.

Αθήνα, Φεβρουάριος 2013 Ευάγγελος Κ. Γκόλιας

# Περιεχόμενα

# Εισαγωγή

# 1

## 1.1. Γεωμετρική Κατανομή

Ας υποθέσουμε ότι εκτελούμε ανεξάρτητες δοκιμές μέχρι να έχουμε την πρώτη επιτυχημένη. Η πιθανότητα μία δοκιμή να είναι επιτυχημένη είναι p. Αν συμβολίσουμε με X τον αριθμό των δοκιμών που απαιτούνται μέχρι να έχουμε την πρώτη επιτυχημένη, τότε λέμε ότι η X ακολουθεί τη Γεωμετρική κατανομή με παράμετρο p. Η συνάρτηση μάζας πιθανότητας δίνεται από την εξίσωση

$$p_n = P\{X = n\} = (1-p)^{n-1} p, \quad n = 1,2,\ldots \tag{1.1}$$

Η εξίσωση αυτή προκύπτει από το γεγονός ότι για να είναι η X ίση με n, είναι απαραίτητο οι πρώτες $n-1$ δοκιμές να είναι αποτυχημένες και η $n$ – οστή δοκιμή επιτυχημένη.

Για να ελέγξουμε ότι η $p_n$ είναι συνάρτηση μάζας πιθανότητας κάνουμε τον υπολογισμό

$$\sum_{n=1}^{\infty} p_n = p \sum_{n=1}^{\infty} (1-p)^{n-1} = 1$$

Από την εξίσωση (1.1) έχουμε

$$E[X] = \sum_{n=1}^{\infty} np(1-p)^{n-1} = p \sum_{n=1}^{\infty} nq^{n-1} \qquad \text{όπου } q = 1-p$$

$$= p\sum_{n=1}^{\infty} \frac{d}{dq}\left(q^n\right) = p\frac{d}{dq}\left(\sum_{n=1}^{\infty} q^n\right)$$

$$= p\frac{d}{dq}\left(\frac{q}{1-q}\right) = \frac{p}{(1-q)^2}$$

$$= \frac{1}{p}$$

και

$$E\left[X(X-1)\right] = \sum_{n=1}^{\infty} n(n-1)p(1-p)^{n-1} = p\sum_{n=1}^{\infty} n(n-1)q^{n-1} \quad \text{όπου } q = 1-p$$

$$= pq\sum_{n=1}^{\infty} \frac{d^2}{dq^2}\left(q^n\right) = pq\frac{d^2}{dq^2}\left(\sum_{n=1}^{\infty} q^n\right)$$

$$= pq\frac{d^2}{dq^2}\left(\frac{q}{1-q}\right) = pq\frac{2}{(1-q)^3}$$

$$= \frac{2(1-p)}{p^2}$$

Συνεπώς,

$$V[X] = E[X - E[X]]^2 = E\left[X^2 - 2XE[X] + \{E[X]\}^2\right]$$

$$= E[X]^2 - \{E[X]\}^2 + E[X] - E[X]$$

$$= E[X(X-1)] + E[X] - \{E[X]\}^2$$

$$= \frac{2(1-p)}{p^2} + \frac{1}{p} - \frac{1}{p^2} = \frac{1-p}{p^2}$$

## 1.2. Γάμμα Κατανομή

Μία συνεχής τυχαία μεταβλητή ακολουθεί την κατανομή Γάμμα όταν η συνάρτηση πυκνότητας πιθανότητας είναι η

$$f(x)=\begin{cases}\dfrac{\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}}{\Gamma(\alpha)} & , \quad x\geq 0\\ 0 & , \quad x<0\end{cases}$$

για $\lambda > 0$ και $\alpha > 0$. Οι παράμετροι της κατανομής Γάμμα είναι οι αριθμοί $\alpha, \lambda$.

Η ποσότητα $\Gamma(\bullet)$ είναι η συνάρτηση Γάμμα η οποία ορίζεται από τη σχέση

$$\Gamma(\alpha)=\int_0^{\infty} e^{-x}x^{\alpha-1}\,dx$$

## 1.3. Εκθετική Κατανομή

Μία συνεχής τυχαία μεταβλητή X ακολουθεί την Εκθετική κατανομή με παράμετρο $\lambda$, $\lambda > 0$, αν η συνάρτηση πυκνότητας πιθανότητας δίνεται από τη σχέση

$$f(x)=\begin{cases}\lambda e^{-\lambda x} & , \quad x\geq 0\\ 0 & , \quad x<0\end{cases}$$

ή, ισοδύναμα, αν η συνάρτηση κατανομής πιθανότητας δίνεται από τη σχέση

$$F(x)=\int_{-\infty}^{x} f(y)\,dy=\begin{cases}1-e^{-\lambda x} & , \quad x\geq 0\\ 0 & , \quad x<0\end{cases}$$

Για τη μέση τιμή της Εκθετικής κατανομής, $E[X]$, έχουμε

$$E[X]=\int_{-\infty}^{\infty} xf(x)\,dx=\int_0^{\infty}\lambda x e^{-\lambda x}\,dx$$

$$=\left[-xe^{-\lambda x}\right]_0^{\infty}+\int_0^{\infty} e^{-\lambda x}\,dx$$

$$=\frac{1}{\lambda}$$

Η χαρακτηριστική συνάρτηση $\varphi(t)$ της Εκθετικής κατανομής είναι

$$\varphi(t) = E\left[e^{tX}\right] = \int_0^{\infty} e^{tx} \lambda e^{-\lambda x}\, dx$$

$$= \frac{\lambda}{\lambda - t} \quad \text{για } t < \lambda \tag{1.2}$$

Όλες οι ροπές της X μπορούν να υπολογιστούν παραγωγίζοντας την εξίσωση (1.2). Για παράδειγμα,

$$E\left[X^2\right] = \left.\frac{d^2}{dt^2}\varphi(t)\right|_{t=0} = \left.\frac{2\lambda}{(\lambda - t)^3}\right|_{t=0}$$

$$= \frac{2}{\lambda^2}$$

Συνεπώς,

$$\mathrm{Var}(X) = E\left[X^2\right] - (E[X])^2$$

$$= \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

### 1.3.1. Ιδιότητες της Εκθετικής Κατανομής

Μία τυχαία μεταβλητή X έχει την ιδιότητα της «απώλειας μνήμης» αν

$$P\{X > s + t \mid X > t\} = P\{X > s\} \quad \forall\, s, t \geq 0 \tag{1.3}$$

Ισοδύναμα μπορούμε να γράψουμε

$$\frac{P\{X > s + t\,,\, X > t\}}{P\{X > t\}} = P\{X > s\}$$

ή

$$P\{X > s + t\} = P\{X > s\}\, P\{X > t\} \tag{1.4}$$

Επειδή η εξίσωση (1.4) ικανοποιείται όταν η X είναι εκθετικά κατανεμημένη $\left(e^{-\lambda(s+t)} = e^{-\lambda s} e^{-\lambda t}\right)$, έπεται ότι οι τυχαίες μεταβλητές που ακολουθούν Εκθετική κατανομή έχουν την ιδιότητα της «απώλειας μνήμης».

Ας υποθέσουμε ότι οι τυχαίες μεταβλητές $X_1, \dots, X_n$ είναι ανεξάρτητες και όμοια κατανεμημένες και ακολουθούν Εκθετική κατανομή με μέση τιμή $\frac{1}{\lambda}$. Θα δείξουμε ότι η τυχαία μεταβλητή $X_1 + \dots + X_n$ ακολουθεί τη Γάμμα κατανομή με παραμέτρους n και λ, χρησιμοποιώντας τη μέθοδο της μαθηματικής επαγωγής.

Για $n = 1$ η τυχαία μεταβλητή $X_1$ ακολουθεί τη Γάμμα κατανομή με παραμέτρους 1 και λ. Υποθέτουμε ότι η $X_1 + \dots + X_{n-1}$ έχει συνάρτηση πυκνότητας πιθανότητας η οποία δίδεται από τη σχέση

$$f_{X_1+\dots+X_{n-1}}(t) = \lambda e^{-\lambda t}\ \frac{(\lambda t)^{n-2}}{(n-2)!}$$

Θα δείξουμε ότι η $X_1 + \dots + X_n$ έχει συνάρτηση πυκνότητας πιθανότητας την

$$f_{X_1+\dots+X_n}(t) = \lambda e^{-\lambda t}\ \frac{(\lambda t)^{n-1}}{(n-1)!}$$

Πράγματι έχουμε ότι

$$f_{X_1+\dots+X_{n-1}+X_n}(t) = \frac{d}{dt} F_{X_1+\dots+X_{n-1}+X_n}(t)$$

$$= \frac{d}{dt} P\left\{X_1 + \dots + X_{n-1} + X_n \le t\right\}$$

Θέτουμε

$$Y = X_1 + \dots + X_{n-1}$$

Κατά συνέπεια,

$$f_{Y+X_n}(t) = \frac{d}{dt} P\left\{Y + X_n \le t\right\} = \frac{d}{dt} P\left\{Y \in (-\infty, +\infty), X_n \in (-\infty, t-y]\right\}$$

$$= \frac{d}{dt}\int_{-\infty}^{+\infty}\int_{-\infty}^{t-y} f(x, y)\,dxdy = \frac{d}{dt}\int_{-\infty}^{+\infty}\int_{-\infty}^{t-y} f_{X_n}(x) f_Y(y)\,dxdy$$

$$= \frac{d}{dt}\int_{-\infty}^{+\infty}\left(\int_{-\infty}^{t-y} f_{X_n}(x)\,dx\right) f_Y(y)\,dy = \frac{d}{dt}\int_{-\infty}^{+\infty} F_{X_n}(t-y) f_Y(y)\,dy$$

$$= \int_{-\infty}^{+\infty} \frac{d}{dt}\left(F_{X_n}(t-y)\right) f_Y(y)\,dy = \int_{-\infty}^{+\infty} f_{X_n}(t-y) f_Y(y)\,dy$$

$$= \int_{-\infty}^{+\infty} f_{X_n}(t-y) f_{X_1+\ldots+X_{n-1}}(y)\, dy = \int_0^t \lambda e^{-\lambda(t-y)} \lambda e^{-\lambda y} \frac{(\lambda y)^{n-2}}{(n-2)!} dy$$

$$= \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!}$$

## 1.4. Poisson Κατανομή

Η τυχαία μεταβλητή X, που παίρνει μία από τις τιμές 0,1,2, … , ακολουθεί Poisson κατανομή με παράμετρο λ, αν για $\lambda > 0$,

$$p(i) = P\{X = i\} = e^{-\lambda} \frac{\lambda^i}{i!}, \quad i = 0,1, \ldots$$

Η παραπάνω εξίσωση αποτελεί πράγματι συνάρτηση μάζας πιθανότητας αφού

$$\sum_{i=0}^{\infty} p(i) = e^{-\lambda} \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = e^{-\lambda} e^{\lambda} = 1$$

Χρησιμοποιώντας τη συνάρτηση μάζας πιθανότητας παίρνουμε

$$E[X] = \sum_{i=0}^{\infty} \frac{i e^{-\lambda} \lambda^i}{i!} = e^{-\lambda} \sum_{i=1}^{\infty} \frac{\lambda^i}{(i-1)!}$$

$$= \lambda e^{-\lambda} \sum_{i=1}^{\infty} \frac{\lambda^{i-1}}{(i-1)!} = \lambda e^{-\lambda} \sum_{\kappa=0}^{\infty} \frac{\lambda^{\kappa}}{\kappa!}$$

$$= \lambda e^{-\lambda} e^{\lambda} = \lambda$$

Η χαρακτηριστική συνάρτηση φ(t) της Poisson κατανομής είναι

$$\varphi(t) = E\left[e^{tX}\right] = \sum_{i=0}^{\infty} \frac{e^{ti} e^{-\lambda} \lambda^i}{i!}$$

$$= e^{-\lambda} \sum_{i=0}^{\infty} \frac{\left(\lambda e^t\right)^i}{i!} = e^{-\lambda} e^{\lambda e^t}$$

$$= \exp\{\lambda(e^t - 1)\}$$

Παραγωγίζοντας έχουμε

$$\varphi'(t) = \lambda e^{t} \exp\{\lambda(e^{t} - 1)\}$$

και

$$\varphi''(t) = \lambda e^{t} \exp\{\lambda(e^{t} - 1)\} + (\lambda e^{t})^{2} \exp\{\lambda(e^{t} - 1)\}$$

Οπότε

$$E[X^2] = \frac{d^2}{dt^2}\varphi(t)\Big|_{t=0} = \lambda + \lambda^2$$

Συνεπώς,

$$V[X] = E[X^2] - \{E[X]\}^2 = \lambda$$

Βλέπουμε λοιπόν ότι τόσο η μέση τιμή όσο και η διασπορά της τυχαίας μεταβλητής, που ακολουθεί Poisson κατανομή με παράμετρο λ, είναι ίσες με λ.

## 1.5. Ασυμπτωτικές Σχέσεις

Στην ενότητα αυτή θα εξετάσουμε τις έννοιες τριών συμβόλων που θα χρησιμοποιήσουμε στο υπόλοιπο τμήμα της εργασίας.

### 1.5.1. Ασυμπτωτική Ισοδυναμία

Όταν γράφουμε

$$f(x) \sim g(x), \quad x \to x_0$$

το οποίο διαβάζεται ״η f(x) είναι ασυμπτωτικά ισοδύναμη στη g(x) καθώς το x τείνει στο $x_0$״, εννοούμε ότι το σχετικό σφάλμα μεταξύ της f και της g ״πηγαίνει״ στο μηδέν καθώς το $x \to x_0$. Δηλαδή

$$\lim_{x \to x_0} \frac{f(x) - g(x)}{g(x)} = 0$$

Ισοδύναμα έχουμε ότι

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = 1$$

Παρατηρούμε ότι αν $f(x) \sim g(x)$ $(x \to x_0)$ τότε και $g(x) \sim f(x)$ $(x \to x_0)$.

### 1.5.2. «Μεγάλο ο»

Ένα άλλο σύμβολο το οποίο θα χρησιμοποιήσουμε παρακάτω είναι το 'O'. Ορίζουμε

$$f(x) = O[g(x)], \quad x \to x_0$$

και διαβάζουμε ″η f(x) είναι το πολύ τάξης g(x) καθώς το $x \to x_0$″ ή ″η f(x) είναι 'O' της g(x) καθώς το $x \to x_0$″ αν το πηλίκο $\frac{f(x)}{g(x)}$ είναι φραγμένο για x ″κοντά″ στο $x_0$. Δηλαδή

$$\left|\frac{f(x)}{g(x)}\right| < M$$

για μία σταθερά M αν το x είναι αρκετά ″κοντά″ στο $x_0$. Παρατηρούμε ότι αν $f(x) \sim g(x)$ καθώς το $x \to x_0$, τότε $f(x) = O[g(x)]$ καθώς $x \to x_0$. Το σύμβολο 'O' είναι ιδιαίτερα χρήσιμο γιατί εκφράζει την τάξη μεγέθους του πρώτου όρου που παραλείπεται όταν αυτός ο όρος δεν έχει υπολογιστεί με ακρίβεια.

### 1.5.3. «Μικρό ο»

Το τελευταίο σύμβολο, αλλά όχι το λιγότερο σημαντικό, που θα παρουσιάσουμε, είναι το 'o'. Ορίζουμε

$$f(x) = o[g(x)], \quad x \to x_0$$

και διαβάζουμε ″η f(x) είναι 'o' της g(x) καθώς το $x \to x_0$″ αν το πηλίκο $\frac{f(x)}{g(x)}$ ″πηγαίνει″ στο μηδέν καθώς το x τείνει στο $x_0$. Δηλαδή

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = 0$$

Στην ειδική περίπτωση όπου $g(x) = x$ και $x_0 = 0$ έχουμε

$$f(x) = o(x), \quad x \to 0$$

Συνεπώς,

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

Για να συμβαίνει κάτι τέτοιο, θα πρέπει το f(x) να ˝πηγαίνει˝ στο μηδέν γρηγορότερα από το x. Δηλαδή για μικρό x, το f(x) πρέπει να είναι μικρό όταν συγκρίνεται με το x.

***Παρατήρηση:*** *Οι τρεις έννοιες: ασυμπτωτική ισοδυναμία, «μεγάλο o» και «μικρό o» μπορούν επίσης να οριστούν στις περιπτώσεις* $x \to x_0^+$, $x \to x_0^-$, $x \to +\infty$ *και* $x \to -\infty$.

## 1.6. Ανέλιξη Poisson

### 1.6.1. Απαριθμήτρια Ανέλιξη (Counting Process)

Μία στοχαστική ανέλιξη $\{N(t)\,,\, t \geq 0\}$ ονομάζεται απαριθμήτρια ανέλιξη αν το N(t) αντιπροσωπεύει το συνολικό αριθμό των ˝γεγονότων˝ που συμβαίνουν μέχρι τη χρονική στιγμή t.

Από τον ορισμό βλέπουμε ότι μία απαριθμήτρια ανέλιξη N(t) ικανοποιεί τα παρακάτω:

i) $N(t) \geq 0$.

ii) Το N(t) είναι ακέραιος αριθμός.

iii) Αν $s < t$, τότε $N(s) \leq N(t)$.

iv) Για $s < t$, το $N(t) - N(s)$ είναι ίσο με τον αριθμό των γεγονότων που συμβαίνουν στο χρονικό διάστημα $(s\,,\, t]$.

Μία απαριθμήτρια ανέλιξη λέγεται ότι έχει ανεξάρτητες προσαυξήσεις αν οι αριθμοί των γεγονότων που συμβαίνουν σε μη αλληλοεπικαλυπτόμενα χρονικά διαστήματα είναι ανεξάρτητοι. Δηλαδή για κάθε n και κάθε $t_0 < t_1 < \ldots < t_n$ οι διαφορές

$$Y(t_j) = N(t_j) - N(t_{j-1}), \quad j = 1,2, \ldots ,n$$

είναι ανεξάρτητες τυχαίες μεταβλητές.

Μία απαριθμήτρια ανέλιξη λέγεται ότι έχει στάσιμες προσαυξήσεις αν η κατανομή του αριθμού των γεγονότων που συμβαίνουν μεταξύ δύο χρονικών στιγμών εξαρτάται αποκλειστικά από την απόσταση μεταξύ των χρονικών στιγμών. Με άλλα λόγια, η ανέλιξη έχει στάσιμες προσαυξήσεις αν οι τυχαίες μεταβλητές $N(s + t) - N(s)$ έχουν την ίδια κατανομή για όλα τα s.

### 1.6.2. Ορισμός της Ανέλιξης Poisson

Μία από τις σημαντικότερες απαριθμήτριες ανελίξεις είναι η ανέλιξη Poisson η οποία ορίζεται ως εξής:

**Ορισμός 1.1**: Η απαριθμήτρια ανέλιξη $\{N(t)\,,\, t \geq 0\}$ ονομάζεται ανέλιξη Poisson με ρυθμό λ, $\lambda > 0$, αν

i) $N(0) = 0$.

ii) Η ανέλιξη έχει ανεξάρτητες προσαυξήσεις.

iii) Ο αριθμός των γεγονότων μεταξύ δύο χρονικών στιγμών, που βρίσκονται σε απόσταση t μεταξύ τους, ακολουθεί κατανομή Poisson με παράμετρο λt. Δηλαδή, για όλα τα $s, t \geq 0$

$$P\{N(s+t) - N(s) = n\} = e^{-\lambda t}\frac{(\lambda t)^n}{n!}, \quad n = 0,1, \ldots$$

Από τη συνθήκη (iii) προκύπτει ότι η ανέλιξη Poisson έχει στάσιμες προσαυξήσεις και ότι

$$E\,[N(t)] = \lambda t$$

Ένας διαφορετικός ορισμός της ανέλιξης Poisson είναι ο ακόλουθος:

**Ορισμός 1.2**: Η απαριθμήτρια ανέλιξη $\{N(t)\,,\, t \geq 0\}$ ονομάζεται ανέλιξη Poisson με ρυθμό λ, $\lambda > 0$, αν

i) $N(0) = 0$.

ii) Η ανέλιξη είναι στάσιμων και ανεξάρτητων προσαυξήσεων.

iii) $P\{N(h) = 1\} = \lambda h + o(h)$.

iv) $P\{N(h) \geq 2\} = o(h)$.

**Θεώρημα 1.1**: Ο Ορισμός 1.1 είναι ισοδύναμος με τον Ορισμό 1.2.

**Απόδειξη**: Πρώτα θα δείξουμε ότι από τον Ορισμό 1.1 συνεπάγεται ο Ορισμός 1.2. Οι πρώτες δύο συνθήκες του Ορισμού 1.2 προκύπτουν άμεσα από τις συνθήκες του Ορισμού 1.1. Οπότε αρκεί να αποδείξουμε τις συνθήκες (iii) και (iv) του Ορισμού 1.2. Από τη συνθήκη (iii) του Ορισμού 1.1 προκύπτει ότι

$$\begin{aligned} P\{N(h) = 1\} &= P\{N(s+h) - N(s) = 1\} \\ &= \lambda h e^{-\lambda h} \end{aligned} \tag{1.5}$$

Επίσης έχουμε ότι

$$\begin{aligned} P\{N(h) \geq 2\} &= 1 - P\{N(h) = 0\} - P\{N(h) = 1\} \\ &= 1 - e^{-\lambda h} - \lambda h e^{-\lambda h} \end{aligned} \tag{1.6}$$

Από την εξίσωση (1.5) και από τον ορισμό του συμβόλου 'ο', η συνθήκη (iii) του Ορισμού 1.2 είναι ισοδύναμη με

$$\lim_{h \to 0} \frac{\lambda h e^{-\lambda h} - \lambda h}{h} = 0 \tag{1.7}$$

Η εξίσωση (1.7) αποδεικνύεται κάνοντας χρήση των ιδιοτήτων του ορίου.

Με αντίστοιχα επιχειρήματα αποδεικνύεται και η συνθήκη (iv) του Ορισμού 1.2 χρησιμοποιώντας την εξίσωση (1.6).

Τώρα θα δείξουμε ότι από τον Ορισμό 1.2 συνεπάγεται ο Ορισμός 1.1. Αρκεί να αποδείξουμε την ισχύ της συνθήκης (iii) του Ορισμού 1.1. Θεωρούμε $u \geq 0$ και θέτουμε

$$g(t) = E\,[\exp\{-uN(t)\}]$$

Μία διαφορική εξίσωση για το g(t) προκύπτει ως εξής:

$$\begin{aligned} g(t+h) &= E\left[\exp\{-uN(t+h)\}\right] \\ &= E\left[\exp\{-uN(t)\}\exp\{-u(N(t+h)-N(t))\}\right] \\ &= E\left[\exp\{-uN(t)\}\right]E\left[\exp\{-u(N(t+h)-N(t))\}\right] \\ &\qquad \text{λόγω ανεξάρτητων προσαυξήσεων} \\ &= g(t)\,E\left[\exp\{-uN(h)\}\right] \qquad (1.8) \\ &\qquad \text{λόγω στάσιμων προσαυξήσεων} \end{aligned}$$

Από τις συνθήκες (iii) και (iv) του Ορισμού 1.2 συνεπάγεται ότι

$$P\{N(h) = 0\} = 1 - \lambda h + o(h) \tag{1.9}$$

Οπότε,

$$\begin{aligned} E\left[\exp\{-uN(h)\}\right] &= \sum_{\kappa=0}^{\infty} e^{-u\kappa} P\{N(h)=\kappa\} \\ &= P\{N(h)=0\} + e^{-u} P\{N(h)=1\} + \\ &\quad + \sum_{\kappa=2}^{\infty} e^{-u\kappa} P\{N(h)=\kappa\} \qquad (1.10) \end{aligned}$$

Επειδή ισχύει ότι

$$\sum_{\kappa=2}^{\infty} e^{-u\kappa} P\{N(h)=\kappa\} = e^{-u} P\{N(h) \geq 2\} + \sum_{\kappa=2}^{\infty} \left(e^{-u\kappa} - e^{-u(\kappa-1)}\right) P\{N(h) \geq \kappa\}$$

(δες και Παράρτημα Π.4.)

έχουμε,

$$\sum_{\kappa=2}^{\infty} e^{-u\kappa} P\{N(h)=\kappa\} = e^{-u} o(h) + (1-e^{u}) o(h) \sum_{\kappa=2}^{\infty} e^{-u\kappa}$$

$$= o(h) \tag{1.11}$$

Χρησιμοποιώντας τις εξισώσεις (1.9), (1.11) και τη συνθήκη (iii) του Ορισμού 1.2, η εξίσωση (1.10) γράφεται

$$E[\exp\{-uN(h)\}] = 1 - \lambda h + o(h) + e^{-u}(\lambda h + o(h)) + o(h)$$

$$= 1 - \lambda h + e^{-u}\lambda h + o(h) \tag{1.12}$$

Από τις εξισώσεις (1.8) και (1.12) λαμβάνουμε ότι

$$g(t+h) = g(t)(1 - \lambda h + e^{-u}\lambda h) + o(h)$$

από το οποίο συνεπάγεται ότι

$$\frac{g(t+h)-g(t)}{h} = g(t)\lambda(e^{-u}-1) + \frac{o(h)}{h}$$

Παίρνοντας το όριο για $h \to 0$ έχουμε

$$g'(t) = g(t)\lambda(e^{-u}-1)$$

Ισοδύναμα,

$$\frac{g'(t)}{g(t)} = \lambda(e^{-u}-1)$$

Ολοκληρώνοντας και χρησιμοποιώντας ότι $g(0) = 1$, παίρνουμε ότι

$$\ln(g(t)) = \lambda t(e^{-u}-1)$$

ή

$$g(t) = \exp\{\lambda t(e^{-u}-1)\}$$

Δηλαδή η ροπογεννήτρια συνάρτηση της τυχαίας μεταβλητής N(t) στο $-u$ είναι ίση με $\exp\{\lambda t(e^{-u} - 1)\}$. Ακριβώς το ίδιο αποτέλεσμα παίρνουμε για τη ροπογεννήτρια συνάρτηση μιας τυχαίας μεταβλητής που ακολουθεί κατανομή Poisson με παράμετρο λt. Επειδή ισχύει ότι αν δύο τυχαίες μεταβλητές έχουν την ίδια ροπογεννήτρια συνάρτηση, τότε έχουν την ίδια κατανομή, συμπεραίνουμε ότι η τυχαία μεταβλητή N(t) ακολουθεί κατανομή Poisson με παράμετρο λt.

### 1.6.3. Κατανομή του Ενδιάμεσου Χρόνου και του Χρόνου Αναμονής

Ας θεωρήσουμε μία ανέλιξη Poisson και ας συμβολίσουμε με $T_1$ τη χρονική στιγμή του πρώτου γεγονότος. Για $n > 1$, συμβολίζουμε με $T_n$ το χρόνο μεταξύ των γεγονότων $n - 1$ και $n$. Η ακολουθία $\{T_n, n = 1,2, \ldots\}$ ονομάζεται ακολουθία των ενδιάμεσων χρόνων.

Τώρα θα προσδιορίσουμε την κατανομή της τυχαίας μεταβλητής $T_n$. Παρατηρούμε ότι το ενδεχόμενο $\{T_1 > t\}$ πραγματοποιείται αν και μόνο αν δε συμβαίνουν γεγονότα της ανέλιξης Poisson το χρονικό διάστημα $[0,t]$ και έτσι,

$$P\{T_1 > t\} = P\{N(t) = 0\} = e^{-\lambda t}$$

Οπότε η τυχαία μεταβλητή $T_1$ ακολουθεί Εκθετική κατανομή με μέση τιμή $\frac{1}{\lambda}$. Τώρα,

$$\begin{aligned} P\{T_2 > t\} &= E\left[P\{T_2 > t \mid T_1\}\right] \\ &= \int_0^\infty P\{T_2 > t \mid T_1 = s\} f_{T_1}(s)\,ds \qquad (1.13) \end{aligned}$$

Εν τούτοις,

$$\begin{aligned} P\{T_2 > t \mid T_1 = s\} &= P\{N(s+t) - N(s) = 0 \mid N(s) = 1\} \\ &= P\{N(s+t) - N(s) = 0\} \\ &\quad \text{λόγω ανεξάρτητων προσαυξήσεων} \\ &= e^{-\lambda t} \qquad (1.14) \\ &\quad \text{από τη συνθήκη (iii) του Ορισμού 1.1} \end{aligned}$$

Η εξίσωση (1.13) μέσω της εξίσωσης (1.14) γράφεται

$$P\{T_2 > t\} = \int_0^\infty e^{-\lambda t} \lambda e^{-\lambda s}\,ds$$

$$= \lambda e^{-\lambda t} \left[ -\frac{1}{\lambda} e^{-\lambda s} \right]_0^{\infty} = e^{-\lambda t} \qquad (1.15)$$

Από τις εξισώσεις (1.14) και (1.15) συμπεραίνουμε ότι η τυχαία μεταβλητή $T_2$ είναι ανεξάρτητη από την $T_1$ και ακολουθεί και αυτή Εκθετική κατανομή με μέση τιμή $\frac{1}{\lambda}$. Επαναλαμβάνοντας τα ίδια επιχειρήματα οδηγούμαστε στην ακόλουθη πρόταση.

**Πρόταση 1.1**: Οι τυχαίες μεταβλητές $T_n$, n = 1,2, … είναι ανεξάρτητες και όμοια κατανεμημένες και ακολουθούν Εκθετική κατανομή με μέση τιμή $\frac{1}{\lambda}$.

Μία άλλη τυχαία μεταβλητή για την οποία ενδιαφερόμαστε είναι η $S_n$, δηλαδή ο χρόνος άφιξης του n – οστού γεγονότος, που επίσης ονομάζεται χρόνος αναμονής μέχρι το n – οστό γεγονός. Προφανώς ισχύει ότι

$$S_n = \sum_{i=1}^{n} T_i, \quad n \geq 1$$

Από την Πρόταση 1.1 και τα αποτελέσματα της Παραγράφου 1.3.1. προκύπτει ότι η τυχαία μεταβλητή $S_n$ ακολουθεί την κατανομή Γάμμα με παραμέτρους n και λ. Δηλαδή η συνάρτηση πυκνότητας πιθανότητας της $S_n$ δίνεται από τη σχέση

$$f_{S_n}(t) = \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!}, \quad t \geq 0 \qquad (1.16)$$

Στην εξίσωση (1.16) μπορούμε επίσης να καταλήξουμε παρατηρώντας ότι το n – οστό γεγονός θα συμβεί πριν από ή στη χρονική στιγμή t αν και μόνο αν ο αριθμός των γεγονότων που συμβαίνουν μέχρι τη χρονική στιγμή t είναι τουλάχιστον ίσος με n.

$$N(t) \geq n \quad \Leftrightarrow \quad S_n \leq t$$

Οπότε,

$$F_{S_n}(t) = P\{S_n \leq t\} = P\{N(t) \geq n\} = \sum_{j=n}^{\infty} e^{-\lambda t} \frac{(\lambda t)^j}{j!}$$

που κατόπιν παραγώγισης ως προς t δίνει

$$f_{S_n}(t) = -\sum_{j=n}^{\infty} \lambda e^{-\lambda t} \frac{(\lambda t)^j}{j!} + \sum_{j=n}^{\infty} \lambda e^{-\lambda t} \frac{(\lambda t)^{j-1}}{(j-1)!}$$

$$= \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!} + \sum_{j=n+1}^{\infty} \lambda e^{-\lambda t} \frac{(\lambda t)^{j-1}}{(j-1)!} - \sum_{j=n}^{\infty} \lambda e^{-\lambda t} \frac{(\lambda t)^j}{j!}$$

$$= \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!}$$

### 1.6.4. Ιδιότητες της Ανέλιξης Poisson

Ας θεωρήσουμε μία ανέλιξη Poisson $\{N(t)\,,\, t \geq 0\}$ με ρυθμό λ και ας υποθέσουμε ότι κάθε φορά που συμβαίνει ένα γεγονός, αυτό χαρακτηρίζεται είτε ως τύπου Ι είτε ως τύπου ΙΙ. Ας υποθέσουμε περαιτέρω ότι κάθε γεγονός χαρακτηρίζεται ως τύπου Ι με πιθανότητα p ή ως τύπου ΙΙ με πιθανότητα 1 – p, ανεξάρτητα από όλα τα άλλα γεγονότα. Με $N_1(t)$ και $N_2(t)$ δηλώνουμε τον αριθμό των γεγονότων τύπου Ι, και αντίστοιχα τύπου ΙΙ, που συμβαίνουν στο χρονικό διάστημα [0,t]. Παρατηρούμε ότι $N(t) = N_1(t) + N_2(t)$.

**Πρόταση 1.2**: Οι στοχαστικές ανελίξεις $\{N_1(t)\,,\, t \geq 0\}$ και $\{N_2(t)\,,\, t \geq 0\}$ είναι ανελίξεις Poisson με αντίστοιχους ρυθμούς λp και λ(1 – p). Επίσης, οι δύο ανελίξεις είναι ανεξάρτητες.

**Απόδειξη**: Θα δείξουμε ότι η $\{N_1(t)\,,\, t \geq 0\}$ ικανοποιεί τις συνθήκες του Ορισμού 1.2, οπότε είναι ανέλιξη Poisson με ρυθμό λp. Ακριβώς με τον ίδιο τρόπο προκύπτει ότι και η $\{N_2(t)\,,\, t \geq 0\}$ είναι ανέλιξη Poisson με ρυθμό λ(1 – p).

- $N_1(0) = 0$ το οποίο προκύπτει άμεσα από το γεγονός ότι $N(0) = 0$.
- Η $\{N_1(t)\,,\, t \geq 0\}$ κληρονομεί τις ιδιότητες των στάσιμων και ανεξάρτητων προσαυξήσεων της ανέλιξης $\{N(t)\,,\, t \geq 0\}$. Αυτό ισχύει γιατί την κατανομή του αριθμού των γεγονότων τύπου Ι μπορούμε να την πάρουμε χρησιμοποιώντας δέσμευση ως προς τον αριθμό των γεγονότων σε εκείνο το διάστημα, και η κατανομή αυτής της τελευταίας ποσότητας εξαρτάται μόνο από το μήκος του διαστήματος και είναι ανεξάρτητη από ότι έχει συμβεί σε οποιοδήποτε χρονικό διάστημα που δεν το επικαλύπτει.
- $$P\{N_1(h)=1\} = E\left[P\{N_1(h)=1 \mid N(h)\}\right]$$

$$= P\{N_1(h)=1 \mid N(h)=1\}\,P\{N(h)=1\}$$

$$+ P\{N_1(h)=1 \mid N(h)\geq 2\}\,P\{N(h)\geq 2\}$$

$$= p(\lambda h + o(h)) + o(h)$$

$$=\lambda ph + o(h)$$

- $$P\{N_1(h) \geq 2\} \leq P\{N(h) \geq 2\} = o(h)$$

Οπότε η $\{N_1(t)\,,\, t \geq 0\}$ είναι μία ανέλιξη Poisson με ρυθμό λp. Επειδή η πιθανότητα ενός γεγονότος τύπου I στο διάστημα από t μέχρι t + h είναι ανεξάρτητη από ότι έχει συμβεί σε διαστήματα που δεν επικαλύπτουν το (t,t + h), είναι ανεξάρτητη από τη γνώση του πότε συμβαίνουν γεγονότα τύπου II. Αυτό δείχνει ότι οι δύο ανελίξεις Poisson είναι ανεξάρτητες.

## 1.7. Η Φόρμουλα Αθροίσματος Euler – Maclaurin

Πριν αποδείξουμε τη φόρμουλα αθροίσματος Euler – Maclaurin θα αναφέρουμε κάποια γενικά στοιχεία που την αφορούν. Η φόρμουλα αυτή αποτελεί μία κομψή και γενική έκφραση για το ασυμπτωτικό ανάπτυγμα των αθροισμάτων της μορφής

$$F(n) = \sum_{\kappa=0}^{n} f(\kappa) \tag{1.17}$$

καθώς το $n \to \infty$.

Στη φόρμουλα αθροίσματος Euler – Maclaurin υπάρχουν τα πολυώνυμα Bernoulli $B_n(x)$ τα οποία ορίζονται από τη σχέση

$$\frac{te^{xt}}{e^t - 1} = \sum_{n=0}^{\infty} B_n(x) \frac{t^n}{n!} \tag{1.18}$$

Ισοδύναμα μπορούμε να γράψουμε ότι

$$B_n(x) = \frac{d^n}{dt^n} \left( \frac{te^{xt}}{e^t - 1} \right) \Bigg|_{t=0} \tag{1.19}$$

Λίγα από τα πρώτα πολυώνυμα Bernoulli είναι: $B_0(x) = 1$, $B_1(x) = x - \frac{1}{2}$, $B_2(x) = x^2 - x + \frac{1}{6}$. Το $B_n(x)$ είναι ένα πολυώνυμο βαθμού n. Οι αριθμοί Bernoulli ορίζονται με τη βοήθεια των πολυωνύμων Bernoulli καθώς $B_n = B_n(0)$. Μερικοί από τους πρώτους αριθμούς Bernoulli είναι: $B_0 = 1$, $B_1 = -\frac{1}{2}$, $B_2 = \frac{1}{6}$, $B_3 = 0$, $B_4 = -\frac{1}{30}$.

**Θεώρημα 1.2**: Το πλήρες ασυμπτωτικό ανάπτυγμα της F(n) στην εξίσωση (1.17) είναι

$$F(n) \sim \frac{1}{2} f(n) + \int_0^n f(t)\,dt + c + \sum_{j=1}^{\infty} (-1)^{j+1} \frac{B_{j+1}}{(j+1)!} f^{(j)}(n), \quad n \to \infty, \quad (1.20)$$

όπου c είναι μία σταθερά που δίνεται από τη σχέση

$$c = \lim_{m\to\infty} \left[ \sum_{j=1}^{m} \frac{(-1)^j B_{j+1}}{(j+1)!} f^{(j)}(0) + \frac{1}{2} f(0) + \right.$$

$$\left. + \frac{(-1)^m}{(m+1)!} \int_0^{\infty} B_{m+1}(t-[t]) f^{(m+1)}(t)\,dt \right] \qquad (1.21)$$

και [t] είναι ο μεγαλύτερος ακέραιος που είναι μικρότερος από το t.

**Απόδειξη**: Οι τύποι (1.20) και (1.21), που αποτελούν τη φόρμουλα αθροίσματος Euler – Maclaurin, αποδεικνύονται με μία διαδικασία πέντε βημάτων:

Βήμα 1.: Θα αποδείξουμε ότι

$$\frac{1}{2}[f(\kappa) + f(\kappa+1)] - \int_{\kappa}^{\kappa+1} f(t)\,dt = \int_{\kappa}^{\kappa+1} \left(t - \kappa - \frac{1}{2}\right) f'(t)\,dt \qquad (1.22)$$

Από την εξίσωση (1.22) ισοδύναμα έχουμε

$$\frac{1}{2}[f(\kappa) + f(\kappa+1)] - \int_{\kappa}^{\kappa+1} f(t)\,dt = \int_{\kappa}^{\kappa+1} d\left[\left(t - \kappa - \frac{1}{2}\right) f(t)\right] - \int_{\kappa}^{\kappa+1} f(t)\,dt \quad \Leftrightarrow$$

$$\frac{1}{2}[f(\kappa) + f(\kappa+1)] = \left[\left(t - \kappa - \frac{1}{2}\right) f(t)\right]_{\kappa}^{\kappa+1} \quad \Leftrightarrow$$

$$\frac{1}{2}[f(\kappa) + f(\kappa+1)] = \frac{1}{2}[f(\kappa) + f(\kappa+1)]$$

Βήμα 2.: Η F(n) στην εξίσωση (1.17) ικανοποιεί τη σχέση

$$F(n) = \frac{1}{2}[f(0) + f(n)] + \int_0^n f(t)\,dt + \int_0^n B_1(t-[t]) f'(t)\,dt \qquad (1.23)$$

Πράγματι, αν στην εξίσωση (1.22) αθροίσουμε ως προς κ, παίρνουμε ότι

$$\frac{1}{2}\sum_{\kappa=0}^{n}\left[f(\kappa)+f(\kappa+1)\right]-\sum_{\kappa=0}^{n}\int_{\kappa}^{\kappa+1}f(t)\,dt=\sum_{\kappa=0}^{n}\int_{\kappa}^{\kappa+1}\left(t-\kappa-\frac{1}{2}\right)f'(t)\,dt$$

Ισοδύναμα,

$$\frac{1}{2}\left[\sum_{\kappa=0}^{n}f(\kappa)+\sum_{\kappa=0}^{n}f(\kappa)-f(0)+f(n+1)\right]-\int_{0}^{n+1}f(t)\,dt=$$

$$=\sum_{\kappa=0}^{n}\int_{\kappa}^{\kappa+1}\left(t-\kappa-\frac{1}{2}\right)f'(t)\,dt \quad \Leftrightarrow$$

$$F(n)-\frac{1}{2}f(0)+\frac{1}{2}f(n+1)-\int_{0}^{n+1}f(t)\,dt=\sum_{\kappa=0}^{n}\int_{\kappa}^{\kappa+1}\left(t-\kappa-\frac{1}{2}\right)f'(t)\,dt \qquad (1.24)$$

Από το δεξιό μέλος της εξίσωσης (1.24) έχουμε

$$\sum_{\kappa=0}^{n}\int_{\kappa}^{\kappa+1}\left(t-\kappa-\frac{1}{2}\right)f'(t)\,dt=\int_{0}^{1}B_1\left(t-\left[\frac{1}{2}\right]\right)f'(t)\,dt+\int_{1}^{2}B_1\left(t-\left[\frac{3}{2}\right]\right)f'(t)\,dt+$$

$$+\ldots+\int_{n}^{n+1}B_1\left(t-\left[\frac{2n+1}{2}\right]\right)f'(t)\,dt$$

$$=\int_{0}^{1}B_1(t-[t])f'(t)\,dt+\int_{1}^{2}B_1(t-[t])f'(t)\,dt+$$

$$+\ldots+\int_{n}^{n+1}B_1(t-[t])f'(t)\,dt$$

$$=\int_{0}^{n+1}B_1(t-[t])f'(t)\,dt \qquad (1.25)$$

Από τις εξισώσεις (1.24), (1.25) μπορούμε να γράψουμε

$$F(n)=\frac{1}{2}f(0)+\left\{-\frac{1}{2}f(n+1)+\int_{n}^{n+1}f(t)\,dt+\int_{n}^{n+1}\left(t-n-\frac{1}{2}\right)f'(t)\,dt\right\}+$$

$$+\int_{0}^{n}f(t)\,dt+\int_{0}^{n}B_1(t-[t])f'(t)\,dt$$

Ισοδύναμα,

$$F(n)=\frac{1}{2}[f(0)+f(n)]+\int_0^n f(t)\,dt+\int_0^n B_1(t-[t])\,f'(t)\,dt$$

<u>Βήμα 3.</u>: Τα πολυώνυμα Bernoulli $B_n(x)$, τα οποία ορίζονται από την εξίσωση (1.18), ικανοποιούν τις σχέσεις

$$B_n'(x)=n\,B_{n-1}(x) \tag{1.26}$$

και

$$B_n(0)=B_n(1), \quad n\geq 2 \tag{1.27}$$

Πρώτα θα αποδείξουμε την (1.26). Παραγωγίζοντας ως προς x την εξίσωση (1.19) παίρνουμε

$$B_n'(x)=\frac{d^n}{dt^n}\left(\frac{t^2e^{xt}}{e^t-1}\right)\Bigg|_{t=0}$$

Ισοδύναμα,

$$\begin{aligned}
B_n'(x)&=\frac{d^{n-1}}{dt^{n-1}}\left\{\frac{te^{xt}}{e^t-1}+t\,\frac{d}{dt}\left(\frac{te^{xt}}{e^t-1}\right)\right\}\Bigg|_{t=0}\\
&=\frac{d^{n-2}}{dt^{n-2}}\left\{2\,\frac{d}{dt}\left(\frac{te^{xt}}{e^t-1}\right)+t\,\frac{d^2}{dt^2}\left(\frac{te^{xt}}{e^t-1}\right)\right\}\Bigg|_{t=0}\\
&=\frac{d^{n-3}}{dt^{n-3}}\left\{3\,\frac{d^2}{dt^2}\left(\frac{te^{xt}}{e^t-1}\right)+t\,\frac{d^3}{dt^3}\left(\frac{te^{xt}}{e^t-1}\right)\right\}\Bigg|_{t=0}\\
&=\left\{n\,\frac{d^{n-1}}{dt^{n-1}}\left(\frac{te^{xt}}{e^t-1}\right)+t\,\frac{d^n}{dt^n}\left(\frac{te^{xt}}{e^t-1}\right)\right\}\Bigg|_{t=0}\\
&=n\,B_{n-1}(x)
\end{aligned}$$

Τώρα θα αποδείξουμε την εξίσωση (1.27). Από την (1.18) για $x=1$ έχουμε

$$\sum_{n=0}^{\infty}B_n(1)\frac{t^n}{n!}=\frac{te^t}{e^t-1}=t+\frac{t}{e^t-1} \tag{1.28}$$

Η (1.18) για $x=0$ δίνει

$$\frac{t}{e^{t}-1}=\sum_{n=0}^{\infty} B_n(0)\frac{t^n}{n!} \tag{1.29}$$

Με τη χρησιμοποίηση της εξίσωσης (1.29), η (1.28) γράφεται

$$\sum_{n=0}^{\infty} B_n(1)\frac{t^n}{n!}=t+\sum_{n=0}^{\infty} B_n(0)\frac{t^n}{n!}$$

Από την τελευταία εξίσωση προκύπτει

$$B_n(0)=B_n(1), \quad n \neq 1 \tag{1.30}$$

και

$$B_1(1)=B_1(0)+1, \quad n = 1$$

Από την (1.30) έχουμε την (1.27).

<u>Βήμα 4.</u>: Θα δείξουμε ότι

$$F(n)=\frac{1}{2}\left[f(0)+f(n)\right]+\int_0^n f(t)\,dt+\sum_{j=1}^{m}(-1)^{j+1}\frac{B_{j+1}}{(j+1)!}\left[f^{(j)}(n)-f^{(j)}(0)\right]+$$

$$+\frac{(-1)^m}{(m+1)!}\int_0^n B_{m+1}(t-[t])\,f^{(m+1)}(t)\,dt \tag{1.31}$$

Στο δεύτερο ολοκλήρωμα του δεξιού μέλους της εξίσωσης (1.23) εφαρμόζουμε επανειλημμένα ολοκλήρωση κατά παράγοντες και χρησιμοποιώντας τις εξισώσεις (1.26), (1.27) παίρνουμε

$$\int_0^n B_1(t-[t])\,f'(t)\,dt=\frac{1}{2}\int_0^n B_2'(t-[t])\,f'(t)\,dt$$

$$=\frac{1}{2}\left[B_2(t-[t])\,f'(t)\right]_0^n-\frac{1}{2}\int_0^n B_2(t-[t])\,f''(t)\,dt$$

$$=\frac{1}{2}\left[B_2(1)f'(n)-B_2(1)f'(0)\right]-\frac{1}{2}\cdot\frac{1}{3}\int_0^n B_3'(t-[t])\,f''(t)\,dt$$

$$=\frac{1}{2!}B_2\left[f'(n)-f'(0)\right]-\frac{1}{3!}\left\{\left[B_3(t-[t])\,f''(t)\right]_0^n-\right.$$

$$- \int_0^n B_3(t-[t])\, f'''(t)\, dt \Big\}$$

$$= \frac{1}{2!} B_2 [f'(n) - f'(0)] - \frac{1}{3!} B_3 [f''(n) - f''(0)] + \frac{1}{3!} \int_0^n B_3(t-[t])\, f'''(t)\, dt$$

$$= \sum_{j=1}^{m} (-1)^{j+1} \frac{B_{j+1}}{(j+1)!} \left[ f^{(j)}(n) - f^{(j)}(0) \right] +$$

$$+ \frac{(-1)^m}{(m+1)!} \int_0^n B_{m+1}(t-[t])\, f^{(m+1)}(t)\, dt \qquad (1.32)$$

Αντικαθιστώντας την εξίσωση (1.32) στην (1.23) παίρνουμε την εξίσωση (1.31).

Βήμα 5.: Η (1.31) με αναδιάταξη των όρων της γράφεται

$$F(n) = \frac{1}{2} f(n) + \int_0^n f(t)\, dt + \left\{ \sum_{j=1}^{m} (-1)^j \frac{B_{j+1}}{(j+1)!} f^{(j)}(0) + \frac{1}{2} f(0) + \right.$$

$$\left. + \frac{(-1)^m}{(m+1)!} \int_0^n B_{m+1}(t-[t])\, f^{(m+1)}(t)\, dt \right\} + \sum_{j=1}^{m} (-1)^{j+1} \frac{B_{j+1}}{(j+1)!} f^{(j)}(n) \quad (1.33)$$

Από την εξίσωση (1.33), παίρνοντας το όριο καθώς το $m \to \infty$, προκύπτει

$$F(n) = \frac{1}{2} f(n) + \int_0^n f(t)\, dt + \lim_{m\to\infty} \left\{ \sum_{j=1}^{m} (-1)^j \frac{B_{j+1}}{(j+1)!} f^{(j)}(0) + \frac{1}{2} f(0) + \right.$$

$$\left. + \frac{(-1)^m}{(m+1)!} \int_0^n B_{m+1}(t-[t])\, f^{(m+1)}(t)\, dt \right\} + \sum_{j=1}^{\infty} (-1)^{j+1} \frac{B_{j+1}}{(j+1)!} f^{(j)}(n) \quad (1.34)$$

Αν στην εξίσωση (1.34) πάρουμε το όριο καθώς το $n \to \infty$, προκύπτουν οι εξισώσεις (1.20), (1.21), που αποτελούν τη φόρμουλα αθροίσματος Euler – Maclaurin, και ολοκληρώνεται η απόδειξη του θεωρήματος.

## 1.8. Η Σταθερά Euler

Η ακολουθία $(\alpha_n)$ με $\alpha_n = \left(1+\frac{1}{n}\right)^n$ είναι γνησίως αύξουσα και άνω φραγμένη, οπότε συγκλίνει στο $\Re$. Το όριό της είναι ο αριθμός e. Η ακολουθία ($\beta_n$) με $\beta_n = \left(1+\frac{1}{n}\right)^{n+1}$ είναι γνησίως φθίνουσα και κάτω φραγμένη, οπότε συγκλίνει και αυτή στο $\Re$. Το όριο της ($\beta_n$) είναι ο αριθμός e. Επομένως μπορούμε να γράψουμε την ανισότητα

$$\left(1+\frac{1}{n}\right)^{n} < e < \left(1+\frac{1}{n}\right)^{n+1}$$

Κάνοντας χρήση της ιδιότητας ότι η συνάρτηση ln είναι γνησίως αύξουσα έχουμε

$$\ln\left(1+\frac{1}{n}\right)^{n} < \ln e < \ln\left(1+\frac{1}{n}\right)^{n+1}$$

Ισοδύναμα,

$$n\ln\left(1+\frac{1}{n}\right) < 1 < (n+1)\ln\left(1+\frac{1}{n}\right)$$

Λαμβάνουμε

$$n\ln\left(1+\frac{1}{n}\right) < 1 \quad \text{και} \quad (n+1)\ln\left(1+\frac{1}{n}\right) > 1$$

Συνεπώς,

$$\frac{1}{n+1} < \ln\left(1+\frac{1}{n}\right) < \frac{1}{n} \tag{1.35}$$

Τώρα θα αποδείξουμε ότι η ακολουθία ($\gamma_n$) με $\gamma_n = \sum_{\kappa=1}^{n}\frac{1}{\kappa} - \ln n$ είναι γνησίως φθίνουσα και κάτω φραγμένη.

- Έχουμε

$$\begin{aligned}\gamma_n - \gamma_{n+1} &= \left(\sum_{\kappa=1}^{n}\frac{1}{\kappa} - \ln n\right) - \left(\sum_{\kappa=1}^{n+1}\frac{1}{\kappa} - \ln(n+1)\right)\\ &= \ln\left(1+\frac{1}{n}\right) - \frac{1}{n+1} \qquad (1.36)\end{aligned}$$

Από την (1.36), λόγω της (1.35), συνεπάγεται

$$\gamma_n - \gamma_{n+1} > 0 \ \text{ για κάθε } \ n \in \aleph$$

Οπότε η ακολουθία $(\gamma_n)$ είναι γνησίως φθίνουσα.

- Η ακολουθία $(\gamma_n)$ είναι κάτω φραγμένη, γιατί

$$\gamma_n = \sum_{\kappa=1}^{n} \frac{1}{\kappa} - \ln n > \sum_{\kappa=1}^{n} \ln\left(1+\frac{1}{\kappa}\right) - \ln n \text{ (με τη βοήθεια της (1.35))}$$

$$= \ln\left(\frac{2}{1}\cdot\frac{3}{2}\cdot\frac{4}{3}\cdot\ldots\cdot\frac{n+1}{n}\cdot\frac{1}{n}\right)$$

$$= \ln\left(\frac{n+1}{n}\right) > 0$$

Άρα $\gamma_n > 0$ για κάθε $n \in \aleph$.

Επομένως η ακολουθία $(\gamma_n)$ συγκλίνει σε πραγματικό αριθμό. Το όριο της ακολουθίας $(\gamma_n)$ συμβολίζεται με το γράμμα γ. Ο αριθμός γ ονομάζεται σταθερά Euler και έχει βρεθεί ότι

$$\gamma = 0{,}57721566490\ldots$$

# The Coupon Collector's Problem

# 2

## 2.1. Το Πρόβλημα της Επιλογής του Συλλέκτη Κουπονιών από τη Σκοπιά των Ανελίξεων Poisson

Έστω ότι υπάρχουν m κουπόνια διαφορετικά μεταξύ τους. Κάθε φορά επιλέγουμε ένα κουπόνι, ανεξάρτητα από αυτά που έχουμε ήδη επιλέξει. Το j – οστό κουπόνι, $j = 1, \ldots ,m$, επιλέγεται με πιθανότητα $p_j$ και προφανώς $\sum_{j=1}^{m} p_j = 1$. Συμβολίζουμε με N τον αριθμό των κουπονιών που πρέπει να επιλεγούν, ώστε να έχουμε τουλάχιστον ένα από όλα τα κουπόνια.

### 2.1.1. Υπολογισμός της Μέσης Τιμής του N

Συμβολίζουμε με $N_j$ τον αριθμό των κουπονιών που πρέπει να επιλεγούν, ώστε να αποκτήσουμε το j – οστό κουπόνι. Οπότε μπορούμε να εκφράσουμε το N ως

$$N = \max_{1 \le j \le m} N_j$$

Από τον ορισμό της τυχαίας μεταβλητής $N_j$, προκύπτει ότι ακολουθεί τη Γεωμετρική κατανομή με παράμετρο $p_j$.

**Πρόταση 2.1**: Οι τυχαίες μεταβλητές $N_j$ , $j = 1, \ldots ,m$, δεν είναι ανεξάρτητες.

**Απόδειξη**: Έστω ότι οι $N_j$ είναι ανά δύο ανεξάρτητες. Οπότε από τον ορισμό της ανεξαρτησίας τυχαίων μεταβλητών έχουμε

$$P\left\{N_j=\kappa \,\middle|\, N_i=\ell\right\}=P\left\{N_j=\kappa\right\} \tag{2.1}$$

για $i \neq j \in \{1, \ldots ,m\}$ και $\kappa < \ell$ χωρίς βλάβη της γενικότητας. Για τον υπολογισμό της πιθανότητας του $\{N_j = \kappa \mid N_i = \ell\}$ κάνουμε το ακόλουθο σχήμα.

| $1$ | $5$ | $\ldots$ | $\kappa-1$ | $\kappa$ | $\ldots$ | $\ell$ |
|---|---|---|---|---|---|---|
| $1-p_j-p_i$ | $(1-p_j-p_i)^5$ | | $(1-p_j-p_i)^{\kappa-1}$ | $(1-p_j-p_i)^{\kappa-1}\times$ <br> $\times p_j$ | | $(1-p_j-p_i)^{\kappa-1}\times$ <br> $\times p_j(1-p_i)^{\ell-1-\kappa}p_i$ |

Με τη βοήθεια του σχήματος και από τον ορισμό της έννοιας της δεσμευμένης πιθανότητας

$$\begin{aligned} P\left\{N_j=\kappa \,\middle|\, N_i=\ell\right\} &= \frac{P\left\{(N_j=\kappa)\cap(N_i=\ell)\right\}}{P\left\{N_i=\ell\right\}} \\ &= \frac{(1-p_j-p_i)^{\kappa-1}p_j\,(1-p_i)^{\ell-\kappa-1}p_i}{(1-p_i)^{\ell-1}p_i} \\ &= (1-p_j-p_i)^{\kappa-1}p_j\,(1-p_i)^{-\kappa} \end{aligned} \tag{2.2}$$

Επίσης,

$$P\left\{N_j=\kappa\right\}=(1-p_j)^{\kappa-1}p_j \tag{2.3}$$

Από τις (2.2) και (2.3) οδηγούμαστε στο συμπέρασμα ότι δεν ισχύει η (2.1). Οπότε οι τυχαίες μεταβλητές δεν είναι ανά δύο ανεξάρτητες και κατά συνέπεια δεν είναι ανεξάρτητες.

Για αυτόν ακριβώς το λόγο θα πρέπει να τροποποιήσουμε την έκφραση του N, ώστε να υπολογίσουμε τη μέση τιμή του μέγιστου ανεξάρτητων τυχαίων μεταβλητών. Για να το κάνουμε αυτό, υποθέτουμε ότι τα κουπόνια επιλέγονται σε χρονικές στιγμές σύμφωνα με μία ανέλιξη Poisson με ρυθμό $\lambda = 1$. Ένα γεγονός της ανέλιξης αυτής είναι τύπου $j$, $1 \leq j \leq m$, αν το κουπόνι που επιλέγεται εκείνη τη χρονική στιγμή είναι το $j$. Συμβολίζουμε με $N_j(t)$ τον αριθμό των $j$ κουπονιών που επιλέγονται μέχρι τη χρονική στιγμή $t$. Από την Πρόταση 1.2 προκύπτει ότι οι στοχαστικές ανελίξεις $\{ N_j(t) , t \geq 0\}$ , $j = 1, \ldots ,m$, είναι ανεξάρτητες ανελίξεις Poisson με αντίστοιχους ρυθμούς $\lambda p_j = p_j$. Δηλώνουμε με $X_j$ το χρόνο του πρώτου γεγονότος της $j$ – οστής ανέλιξης, δηλαδή τη χρονική στιγμή που επιλέχθηκε το πρώτο $j$ κουπόνι.

**Πρόταση 2.2**: Οι τυχαίες μεταβλητές $X_j$ , j = 1, … ,m, ακολουθούν Εκθετική κατανομή με παράμετρο $p_j$.

**Απόδειξη**: Θα δείξουμε ότι η $X_j$ έχει τη συνάρτηση κατανομής πιθανότητας μιας τυχαίας μεταβλητής που ακολουθεί Εκθετική κατανομή με παράμετρο $p_j$. Πράγματι,

$$\begin{aligned} F(x) &= P\left\{X_j \le x\right\} = 1 - P\left\{X_j > x\right\} \\ &= 1 - P\left\{\bigcap_{t \le x}\left(N_j(t) = 0\right)\right\} \\ &= 1 - P\left\{N_j(x) = 0\right\} \\ &= 1 - e^{-p_j x}, \quad x \ge 0 \end{aligned}$$

**Πρόταση 2.3**: Οι τυχαίες μεταβλητές $X_j$ , j = 1, … ,m, είναι ανεξάρτητες.

**Απόδειξη**: Αρκεί να δείξουμε ότι

$$P\left\{\bigcap_{j=1}^{m}\left(X_j \le x_j\right)\right\} = \prod_{j=1}^{m} P\left\{X_j \le x_j\right\}$$

Πράγματι,

$$\begin{aligned} P\left\{\bigcap_{j=1}^{m}\left(X_j \le x_j\right)\right\} &= P\left\{\left[\bigcup_{j=1}^{m}\left(X_j > x_j\right)\right]^c\right\} = 1 - P\left\{\bigcup_{j=1}^{m}\left(X_j > x_j\right)\right\} \\ &= 1 - P\left\{\bigcup_{j=1}^{m}\left(N_j(x_j) = 0\right)\right\} = 1 - P\left\{\left[\bigcap_{j=1}^{m}\left(N_j(x_j) = 0\right)^c\right]^c\right\} \\ &= 1 - 1 + P\left\{\bigcap_{j=1}^{m}\left(N_j(x_j) = 0\right)^c\right\} = \prod_{j=1}^{m} P\left\{\left(N_j(x_j) = 0\right)^c\right\} \\ &= \prod_{j=1}^{m}\left[1 - P\left\{N_j(x_j) = 0\right\}\right] = \prod_{j=1}^{m}\left[1 - P\left\{X_j > x_j\right\}\right] \\ &= \prod_{j=1}^{m} P\left\{X_j \le x_j\right\} \end{aligned}$$

Συμβολίζουμε με X τη χρονική στιγμή κατά την οποία συμπληρώνεται μία ολόκληρη συλλογή από κουπόνια. Με άλλα λόγια είναι το χρονικό διάστημα που απαιτείται ώστε στη συλλογή να βρίσκονται για πρώτη φορά m διαφορετικά κουπόνια. Οπότε μπορούμε να εκφράσουμε το X ως

$$X = \max_{1 \le j \le m} X_j$$

Χρησιμοποιώντας τις προτάσεις 2.2 και 2.3 έχουμε

$$P\{X < t\} = P\left\{\max_{1 \le j \le m} X_j < t\right\} = P\left\{\bigcap_{j=1}^{m}\left(X_j < t\right)\right\}$$

$$= \prod_{j=1}^{m} P\{X_j < t\} = \prod_{j=1}^{m} P\{X_j \le t\}$$

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

Έτσι,

$$E[X] = \int_0^{\infty} t f(t)\,dt = [tF(t)]_0^{\infty} - \int_0^{\infty} F(t)\,dt$$

$$= \int_0^{\infty} dt - \int_0^{\infty} F(t)\,dt = \int_0^{\infty}\left[1 - P\{X \le t\}\right]dt$$

$$= \int_0^{\infty}\left[1 - \prod_{j=1}^{m}\left(1 - e^{-p_j t}\right)\right]dt \tag{2.4}$$

Τώρα θα μελετήσουμε τη σχέση που συνδέει το E [X], δηλαδή τη μέση τιμή του χρόνου που απαιτείται για να συμπληρώσουμε μία ολόκληρη συλλογή, με το E [N], που είναι η μέση τιμή του αριθμού των κουπονιών που χρειάζονται ώστε να έχουμε μία πλήρη συλλογή. Συμβολίζουμε με $T_i$ το $i$ – οστό χρονικό διάστημα μιας ανέλιξης Poisson που μετρά τον αριθμό των κουπονιών που έχουν ληφθεί. Δηλαδή η $T_i$ αναπαριστά το χρόνο που μεσολαβεί από τη λήψη του κουπονιού $i-1$ μέχρι τη λήψη του $i$ – οστού.

**Πρόταση 2.4**: Οι τυχαίες μεταβλητές $T_i$ , $1 \le i \le N$, ακολουθούν Εκθετική κατανομή με παράμετρο 1 και είναι ανεξάρτητες.

**Απόδειξη**: Από όσα ειπώθηκαν στην Παράγραφο 1.6.3. και από την Πρόταση 1.1 έχουμε το ζητούμενο.

Έχει ειπωθεί ότι ο αριθμός των κουπονιών που χρειάζονται για να συμπληρωθεί μία ολόκληρη συλλογή είναι N. Το άθροισμα των χρονικών διαστημάτων, μεταξύ των διαδοχικών λήψεων των N κουπονιών, δίνει το συνολικό χρόνο, που απαιτείται για την ολοκλήρωση της συλλογής, δηλαδή το X. Συνεπώς,

$$X = \sum_{i=1}^{N} T_i$$

Από την Πρόταση 2.4 και από το γεγονός ότι η N είναι ανεξάρτητη από τις $T_i$, έχουμε

$$E[X] = E\left[E[X \mid N]\right] = \sum_{n=1}^{\infty}\left(E[X \mid N=n]\; P\{N=n\}\right)$$

$$= \sum_{n=1}^{\infty}\left(E\left[\sum_{i=1}^{N} T_i \,\middle|\, N=n\right] P\{N=n\}\right)$$

$$= \sum_{n=1}^{\infty}\left(E\left[\sum_{i=1}^{n} T_i\right] P\{N=n\}\right)$$

$$= \sum_{n=1}^{\infty}\left(nE[T_i]\,P\{N=n\}\right)$$

$$= E[T_i]\sum_{n=1}^{\infty} nP\{N=n\}$$

$$= E[T_i]\,E[N] \tag{2.5}$$

Οι τυχαίες μεταβλητές $T_i$ έχουμε αποδείξει ότι ακολουθούν Εκθετική κατανομή με παράμετρο 1, οπότε

$$E[T_i] = 1 \tag{2.6}$$

Από τη (2.5), λόγω της (2.6), προκύπτει

$$E[N] = E[X]$$

Κατά συνέπεια το E [N] δίνεται από την Εξίσωση (2.4).

## 2.1.2. Υπολογισμός της Μέσης Τιμής του Αριθμού των Κουπονιών Που Εμφανίζονται Μόνο Μία Φορά σε Ολόκληρη τη Συλλογή

Ορίζουμε η τυχαία μεταβλητή $I_i$ να είναι ίση με 1, αν υπάρχει ένα μόνο i κουπόνι σε ολόκληρη τη συλλογή, ή να είναι ίση με 0 σε οποιαδήποτε άλλη περίπτωση. Οπότε θα πρέπει να υπολογίσουμε το $E\left[\sum_{i=1}^{m} I_i\right]$. Έχουμε

$$E\left[\sum_{i=1}^{m} I_i\right] = \sum_{i=1}^{m} E\left[I_i\right]$$

$$= \sum_{i=1}^{m} P\left\{I_i = 1\right\} \qquad (2.7)$$

Στην τελική συλλογή θα υπάρχει ένα μόνο i κουπόνι αν την έχουμε ολοκληρώσει πριν την έλευση του δεύτερου i κουπονιού. Με άλλα λόγια, η ύπαρξη μοναδικού i κουπονιού στη συλλογή οφείλεται στην κατοχή τουλάχιστον ενός από όλα τα κουπόνια πριν την άφιξη του δεύτερου i κουπονιού.

Συμβολίζουμε με $S_i$ τη χρονική στιγμή κατά την οποία λαμβάνεται το δεύτερο i κουπόνι. Έστω τώρα η απαριθμήτρια ανέλιξη $\{N_i(t)$, $t \geq 0\}$ με την οποία δηλώνεται, όπως έχουμε πει, ο αριθμός των i κουπονιών που επιλέγονται μέχρι τη χρονική στιγμή t. Σύμφωνα με την Πρόταση 1.2 η ανέλιξη αυτή είναι ανέλιξη Poisson με ρυθμό $\lambda p_i = p_i$. Ορίζουμε με $T_{i,1}$ τη χρονική στιγμή κατά την οποία λαμβάνεται το πρώτο i κουπόνι. Για $n > 1$, ορίζουμε με $T_{i,n}$ το χρονικό διάστημα που μεσολαβεί από τη λήψη του κουπονιού $n-1$ τύπου i μέχρι τη λήψη του $n-$οστού του ιδίου τύπου. Η ακολουθία $\{T_{i,n}, n = 1,2, \ldots\}$ ονομάζεται ακολουθία των ενδιάμεσων χρόνων του $i-$οστού κουπονιού. Επαναλαμβάνοντας την επιχειρηματολογία της Παραγράφου 1.6.3. καταλήγουμε σε μία πρόταση αντίστοιχη της Πρότασης 1.1.

**Πρόταση 2.5**: Οι τυχαίες μεταβλητές $T_{i,n}$, $n = 1,2, \ldots$ είναι ανεξάρτητες και όμοια κατανεμημένες και ακολουθούν Εκθετική κατανομή με παράμετρο $p_i$.

Από τον ορισμό της τυχαίας μεταβλητής $S_i$ προκύπτει ότι

$$S_i = \sum_{\kappa=1}^{2} T_{i,\kappa}$$

Από την Πρόταση 2.5 και τις ιδιότητες της Εκθετικής κατανομής, όπως αυτές αναλύονται στην Παράγραφο 1.3.1., έχουμε ότι η $S_i$ ακολουθεί την κατανομή Γάμμα με παραμέτρους 2 και $p_i$.

Στο σημείο αυτό θα πρέπει να θυμίσουμε ότι έχουμε συμβολίσει με $X_j$ τη χρονική στιγμή που επιλέχθηκε το πρώτο j κουπόνι και ότι για να υπάρχει στην τελική συλλογή ένα μόνο i κουπόνι θα πρέπει να έχουν εμφανιστεί όλα τα κουπόνια πριν την έλευση του δεύτερου i κουπονιού. Οπότε, χρησιμοποιώντας και το ότι η $S_i$ ακολουθεί την κατανομή Γάμμα με παραμέτρους 2 και $p_i$, μπορούμε να γράψουμε

$$\begin{aligned}
P\{I_i = 1\} &= P\left\{X_j < S_i,\ \text{για κάθε}\ j \neq i\right\} \\
&= E\left[P\left\{X_j < S_i,\ \text{για κάθε}\ j \neq i \,\middle|\, S_i\right\}\right] \\
&= \int_0^\infty P\left\{X_j < S_i,\ \text{για κάθε}\ j \neq i \,\middle|\, S_i = x\right\} f_{S_i}(x)\,dx \\
&= \int_0^\infty P\left\{X_j < x,\ \text{για κάθε}\ j \neq i\right\} \frac{p_i e^{-p_i x} p_i x}{\Gamma(2)}\,dx \\
&= \int_0^\infty P\left\{\bigcap_{j \neq i}(X_j < x)\right\} p_i^2 x e^{-p_i x}\,dx \\
&= \int_0^\infty \prod_{j \neq i} P\left\{X_j < x\right\} p_i^2 x e^{-p_i x}\,dx \\
&= \int_0^\infty \prod_{j \neq i}\left(1 - e^{-p_j x}\right) p_i^2 x e^{-p_i x}\,dx
\end{aligned} \tag{2.8}$$

Αν στην Εξίσωση (2.7) αντικαταστήσουμε τη (2.8) παίρνουμε

$$\begin{aligned}
E\left[\sum_{i=1}^m I_i\right] &= \sum_{i=1}^m \left(\int_0^\infty \prod_{j \neq i}\left(1 - e^{-p_j x}\right) p_i^2 x e^{-p_i x}\,dx\right) \\
&= \sum_{i=1}^m \left(\int_0^\infty \prod_{j=1}^m \left(1 - e^{-p_j x}\right) x p_i^2 \frac{e^{-p_i x}}{1 - e^{-p_i x}}\,dx\right) \\
&= \int_0^\infty x \prod_{j=1}^m \left(1 - e^{-p_j x}\right) \sum_{i=1}^m p_i^2 \frac{e^{-p_i x}}{1 - e^{-p_i x}}\,dx
\end{aligned}$$

## 2.2. Υπολογισμός των Ασυμπτωτικών Αναπτυγμάτων των Ροπών Πρώτης και Δεύτερης Τάξης στο Πρόβλημα της Επιλογής του Συλλέκτη Κουπονιών

Στη συγκεκριμένη ενότητα, και στις υποενότητές της, θα θεωρήσουμε την ύπαρξη N κουπονιών, που είναι διαφορετικά μεταξύ τους. Για $1 \le j \le N$, θα συμβολίσουμε με $p_j$ την πιθανότητα το κουπόνι που επιλέγεται να είναι τύπου j. Θα πρέπει εδώ να σημειώσουμε ότι τα κουπόνια επιλέγονται ανεξάρτητα το ένα από το άλλο με επανατοποθέτηση και ο τύπος τους καταγράφεται. Ο αριθμός των κουπονιών, που πρέπει να επιλεγούν ώστε να έχουμε τουλάχιστον ένα κουπόνι από όλους τους τύπους, συμβολίζεται με $T_N$.

### 2.2.1. Υπολογισμός των Ροπών Πρώτης και Δεύτερης Τάξης

Συμβολίζουμε με $A_j^\kappa$, $1 \le j \le N$, το ενδεχόμενο το $j-$οστό κουπόνι να μην υπάρχει μέσα στα κ τον αριθμό κουπόνια που έχουμε επιλέξει. Τότε

$$P\{T_N \ge \kappa\} = P\{A_1^{\kappa-1} \cup \ldots \cup A_N^{\kappa-1}\}, \quad \kappa = 1,2,\ \ldots$$

Χρησιμοποιώντας το Γενικευμένο αθροιστικό τύπο παίρνουμε

$$\begin{aligned}
P\{T_N \ge \kappa\} &= P\{A_1^{\kappa-1}\} + \ldots + P\{A_N^{\kappa-1}\} - P\{A_1^{\kappa-1} \cap A_2^{\kappa-1}\} - \ldots - P\{A_{N-1}^{\kappa-1} \cap A_N^{\kappa-1}\} + \\
&\quad + P\{A_1^{\kappa-1} \cap A_2^{\kappa-1} \cap A_3^{\kappa-1}\} + \ldots + P\{A_{N-2}^{\kappa-1} \cap A_{N-1}^{\kappa-1} \cap A_N^{\kappa-1}\} - \\
&\quad - \ldots + (-1)^{N-1} P\{A_1^{\kappa-1} \cap \ldots \cap A_N^{\kappa-1}\} \\
&= (1-p_1)^{\kappa-1} + \ldots + (1-p_N)^{\kappa-1} - (1-p_1-p_2)^{\kappa-1} - \ldots - (1-p_{N-1} - \\
&\quad - p_N)^{\kappa-1} + (1-p_1-p_2-p_3)^{\kappa-1} + \ldots + (1-p_{N-2}-p_{N-1}-p_N)^{\kappa-1} - \\
&\quad - \ldots + (-1)^{N-1}(1-p_1-\ldots-p_N)^{\kappa-1} \\
&= \left(1 - \sum_{j\in\{1\}} p_j\right)^{\kappa-1} + \ldots + \left(1 - \sum_{j\in\{N\}} p_j\right)^{\kappa-1} - \left(1 - \sum_{j\in\{1,2\}} p_j\right)^{\kappa-1} - \\
&\quad - \ldots - \left(1 - \sum_{j\in\{N-1,N\}} p_j\right)^{\kappa-1} + \left(1 - \sum_{j\in\{1,2,3\}} p_j\right)^{\kappa-1} +
\end{aligned}$$

$$+\ldots+\left(1-\sum_{j\in\{N-2,N-1,N\}} p_j\right)^{\kappa-1}-\ldots+(-1)^{N-1}\left(1-\sum_{j\in\{1,\ldots,N\}} p_j\right)^{\kappa-1}$$

$$=\sum_{\substack{J\subseteq\{1,\ldots,N\}\\ J\neq\emptyset}}(-1)^{|J|-1}\left(1-\sum_{j\in J} p_j\right)^{\kappa-1},\quad \kappa=1,2,\ldots \tag{2.9}$$

όπου αθροίζουμε ως προς τα $2^N-1$ μη – κενά υποσύνολα του $\{1,\ldots$ $\ldots,N\}$. Το $|J|$ συμβολίζει τον πληθάριθμο του συνόλου J. Για $z\in\mathbb{C}$, $|z|\geq$ $\geq 1$, ορίζουμε

$$G(z)=E\left[z^{-T_N}\right] \tag{2.10}$$

Για τον υπολογισμό του δεξιού μέλους της (2.10) έχουμε

$$E\left[z^{-T_N}\right]=\sum_{\kappa=0}^{\infty} z^{-\kappa}P\{T_N=\kappa\}$$

$$=1+\sum_{\kappa=1}^{\infty}\left(z^{-\kappa}-z^{-(\kappa-1)}\right)P\{T_N\geq\kappa\}$$

(δες και Παράρτημα Π.4.)

$$=1+\left(z^{-1}-1\right)\sum_{\kappa=1}^{\infty} z^{-(\kappa-1)}P\{T_N\geq\kappa\} \tag{2.11}$$

Η (2.10) λόγω των εξισώσεων (2.11) και (2.9) γράφεται ως εξής:

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

$$=1+\left(z^{-1}-1\right)\sum_{\substack{J\subseteq\{1,\ldots,N\}\\ J\neq\emptyset}}(-1)^{|J|-1}\sum_{\kappa=1}^{\infty} z^{-(\kappa-1)}\left(1-\sum_{j\in J} p_j\right)^{\kappa-1}$$

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

$$=1+\left(z^{-1}-1\right)\sum_{\substack{J\subseteq\{1,\dots,N\}\\ J\neq\emptyset}}(-1)^{|J|-1}\frac{1}{1-z^{-1}\left(1-\sum_{j\in J}p_j\right)}$$

$$=1+\left(z^{-1}-1\right)\sum_{\substack{J\subseteq\{1,\dots,N\}\\ J\neq\emptyset}}(-1)^{|J|}\frac{-z}{z-1+\sum_{j\in J}p_j}$$

$$=1+(z-1)\sum_{\substack{J\subseteq\{1,\dots,N\}\\ J\neq\emptyset}}\frac{(-1)^{|J|}}{z-1+\sum_{j\in J}p_j} \tag{2.12}$$

Παραγωγίζοντας ως προς z τη σχέση (2.10) έχουμε

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

Παίρνοντας το όριο καθώς το $z\to 1^+$ προκύπτει

$$E\left[T_N\right]=-\lim_{z\to 1^+}G'(z) \tag{2.13}$$

Από τη σχέση (2.12) συνεπάγεται

$$G'(z)=\sum_{\substack{J\subseteq\{1,\dots,N\}\\ J\neq\emptyset}}\frac{(-1)^{|J|}}{z-1+\sum_{j\in J}p_j}+(z-1)\sum_{\substack{J\subseteq\{1,\dots,N\}\\ J\neq\emptyset}}\frac{(-1)^{|J|-1}}{\left(z-1+\sum_{j\in J}p_j\right)^2} \tag{2.14}$$

Αντικαθιστώντας τη (2.14) στη (2.13) και υπολογίζοντας το όριο παίρνουμε

$$E\left[T_N\right]=\sum_{\substack{J\subseteq\{1,\dots,N\}\\ J\neq\emptyset}}\frac{(-1)^{|J|-1}}{\sum_{j\in J}p_j}$$

$$=\sum_{m=1}^{N}(-1)^{m-1}\sum_{1\le j_1<\dots<j_m\le N}\frac{1}{p_{j_1}+\dots+p_{j_m}} \tag{2.15}$$

Η σχέση (2.15) μπορεί να γραφεί και ως ακολούθως:

$$E\left[T_N\right]=\sum_{m=1}^{N}(-1)^{m-1}\sum_{1\le j_1<\ldots<j_m\le N}\int_0^{\infty}e^{-\left(p_{j_1}+\ldots+p_{j_m}\right)t}\,dt$$

$$=\int_0^{\infty}\left(\sum_{m=1}^{N}(-1)^{m-1}\sum_{1\le j_1<\ldots<j_m\le N}e^{-\left(p_{j_1}+\ldots+p_{j_m}\right)t}\right)dt$$

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

Κάνοντας αλλαγή μεταβλητής ($e^{-t}=x$ , $dt=-\frac{dx}{x}$) στην παραπάνω σχέση,

$$E\left[T_N\right]=\int_0^{1}\left[1-\prod_{j=1}^{N}\left(1-x^{p_j}\right)\right]\frac{dx}{x} \tag{2.16}$$

Τώρα θα υπολογίσουμε το $E\left[T_N(T_N+1)\right]$. Παραγωγίζοντας δύο φορές ως προς z τη σχέση (2.10) έχουμε

$$G''(z)=E\left[T_N\left(T_N+1\right)z^{-T_N-2}\right]$$

Παίρνοντας στην παραπάνω σχέση το όριο καθώς το $z\to 1^+$ προκύπτει

$$E\left[T_N\left(T_N+1\right)\right]=\lim_{z\to 1^+}G''(z) \tag{2.17}$$

Από τη σχέση (2.14) συνεπάγεται

$$G''(z)=2\sum_{\substack{J\subseteq\{1,\ldots,N\}\\ J\neq\emptyset}}\frac{(-1)^{|J|-1}}{\left(z-1+\sum_{j\in J}p_j\right)^2}+(z-1)\sum_{\substack{J\subseteq\{1,\ldots,N\}\\ J\neq\emptyset}}\frac{2(-1)^{|J|}}{\left(z-1+\sum_{j\in J}p_j\right)^3} \tag{2.18}$$

Αντικαθιστώντας τη (2.18) στη (2.17) και υπολογίζοντας το όριο παίρνουμε

$$E\left[T_N\left(T_N+1\right)\right]=2\sum_{\substack{J\subseteq\{1,\ldots,N\}\\ J\neq\emptyset}}\frac{(-1)^{|J|-1}}{\left(\sum_{j\in J}p_j\right)^2}$$

$$= 2\sum_{m=1}^{N}(-1)^{m-1}\sum_{1\le j_1<\ldots<j_m\le N}\frac{1}{\left(p_{j_1}+\ldots+p_{j_m}\right)^2} \tag{2.19}$$

Η σχέση (2.19) μπορεί να γραφεί και ως ακολούθως:

$$E\left[T_N\left(T_N+1\right)\right]=2\sum_{m=1}^{N}(-1)^{m-1}\sum_{1\le j_1<\ldots<j_m\le N}\int_0^{\infty}e^{-\left(p_{j_1}+\ldots+p_{j_m}\right)t}\,t\,dt$$

$$=2\int_0^{\infty}\left(\sum_{m=1}^{N}(-1)^{m-1}\sum_{1\le j_1<\ldots<j_m\le N}e^{-\left(p_{j_1}+\ldots+p_{j_m}\right)t}\right)t\,dt$$

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

Κάνοντας αλλαγή μεταβλητής ($e^{-t} = x$ , $t = -\ln x$ , $dt = -\frac{dx}{x}$) στην παραπάνω σχέση,

$$E\left[T_N\left(T_N+1\right)\right]=-2\int_0^{1}\left[1-\prod_{j=1}^{N}\left(1-x^{p_j}\right)\right]\frac{\ln x}{x}\,dx \tag{2.20}$$

Από τις εξισώσεις (2.16) και (2.20) έχουμε τις τιμές των $E\,[T_N]$ και $E\,[T_N(T_N + 1)]$ αντίστοιχα, οπότε μπορούμε να υπολογίσουμε τη $V\,[T_N]$. Έχουμε

$$V\left[T_N\right]=E\left[T_N-E\left[T_N\right]\right]^2=E\left[T_N^{\,2}-2T_N E\left[T_N\right]+\left\{E\left[T_N\right]\right\}^2\right]$$

$$=E\left[T_N\right]^2-\left\{E\left[T_N\right]\right\}^2+E\left[T_N\right]-E\left[T_N\right]$$

$$=E\left[T_N\left(T_N+1\right)\right]-E\left[T_N\right]-\left\{E\left[T_N\right]\right\}^2 \tag{2.21}$$

### 2.2.2. Η Περίπτωση των Ίσων Πιθανοτήτων

Η απλούστερη περίπτωση, όσον αφορά τους προηγούμενους τύπους, είναι εκείνη κατά την οποία

$$p_1=\ldots=p_N=\frac{1}{N} \tag{2.22}$$

Η (2.15), λόγω της (2.22), γράφεται

$$E\left[T_N\right] = \sum_{m=1}^{N} (-1)^{m-1} \binom{N}{m} \frac{N}{m}$$

$$= N \sum_{m=1}^{N} (-1)^{m-1} \binom{N}{m} \frac{1}{m}$$

$$= N \sum_{m=1}^{N} \frac{1}{m} \quad \text{(δες και Παράρτημα Π.2.)}$$

$$= N H_N, \quad \text{όπου } H_N = \sum_{m=1}^{N} \frac{1}{m} \tag{2.23}$$

Οι (2.19) και (2.20), λόγω της (2.22), γράφονται

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

Στο ολοκλήρωμα της παραπάνω σχέσης κάνοντας αλλαγή μεταβλητής ($u = 1 - x^{1/N}$ , $x = \left(1 - u\right)^N$ , $\ln x = N \ln\left(1 - u\right)$ , $dx = -N\left(1 - u\right)^{N-1} du$) έχουμε

$$E\left[T_N\left(T_N + 1\right)\right] = -2N^2 \int_0^1 \left(1 - u^N\right) \frac{\ln\left(1 - u\right)}{1 - u} du$$

$$= -2N^2 \int_0^1 \left(1 - u\right)\left(\sum_{m=0}^{N-1} u^m\right) \frac{\ln\left(1 - u\right)}{1 - u} du$$

$$= -2N^2 \sum_{m=1}^{N} \int_0^1 \left(\frac{u^m}{m}\right)' \ln\left(1 - u\right) du$$

$$= -2N^2 \sum_{m=1}^{N} \left\{\left[\frac{u^m}{m} \ln\left(1 - u\right)\right]_0^1 + \int_0^1 \frac{u^m}{m} \cdot \frac{1}{1 - u} du\right\} \tag{2.24}$$

Το ολοκλήρωμα του δεξιού μέλους της σχέσης (2.24) μπορεί να υπολογιστεί με αλλαγή μεταβλητής ($\ell = 1 - u$ , $du = -d\ell$). Συνεπώς

$$\int_0^1 \frac{u^m}{1 - u} du = \int_0^1 \frac{\left(1 - \ell\right)^m}{\ell} d\ell \tag{2.25}$$

Χρησιμοποιώντας το θεώρημα του διωνύμου,

$$\left(1-\ell\right)^{m}=\sum_{j=0}^{m}(-1)^{j}\binom{m}{j}\ell^{j} \tag{2.26}$$

Αντικαθιστώντας τη (2.26) στη (2.25),

$$\begin{aligned}\int_0^1 \frac{u^m}{1-u}\,du &= \int_0^1\left[\frac{1}{\ell}\sum_{j=0}^{m}(-1)^{j}\binom{m}{j}\ell^{j}\right]d\ell \\ &= \int_0^1 \frac{1}{\ell}\,d\ell+\sum_{j=1}^{m}\left[(-1)^{j}\binom{m}{j}\int_0^1 \ell^{j-1}\,d\ell\right] \\ &= -\lim_{\ell\to 0^+}\ln\ell+\sum_{j=1}^{m}(-1)^{j}\binom{m}{j}\frac{1}{j} \end{aligned} \tag{2.27}$$

Η (2.24), λόγω της (2.27), γίνεται

$$\begin{aligned} E\left[T_N\left(T_N+1\right)\right] &= -2N^2\sum_{m=1}^{N}\left\{\frac{1}{m}\left[\lim_{u\to 1^-}\ln\left(1-u\right)-\lim_{\ell\to 0^+}\ln\ell+\sum_{j=1}^{m}(-1)^{j}\binom{m}{j}\frac{1}{j}\right]\right\} \\ &= 2N^2\sum_{m=1}^{N}\left\{\frac{1}{m}\left[\sum_{j=1}^{m}(-1)^{j-1}\binom{m}{j}\frac{1}{j}\right]\right\} \\ &= 2N^2\sum_{m=1}^{N}\left(\frac{1}{m}\sum_{j=1}^{m}\frac{1}{j}\right) \quad \text{(δες και Παράρτημα Π.2.)} \\ &= 2N^2\sum_{m=1}^{N}\frac{H_m}{m}, \quad \text{όπου } H_m=\sum_{j=1}^{m}\frac{1}{j} \end{aligned} \tag{2.28}$$

Η σχέση (2.28) μπορεί να γραφεί και ως εξής:

$$E\left[T_N\left(T_N+1\right)\right]=2N^2\left[\sum_{m=1}^{N}\frac{1}{m^2}+\sum_{m=1}^{N-1}\left(\frac{1}{m}\sum_{j=m+1}^{N}\frac{1}{j}\right)\right]$$

$$= N^2 \left[ \left\{ \sum_{m=1}^{N} \frac{1}{m^2} + 2 \sum_{m=1}^{N-1} \left( \frac{1}{m} \sum_{j=m+1}^{N} \frac{1}{j} \right) \right\} + \sum_{m=1}^{N} \frac{1}{m^2} \right]$$

$$= N^2 \left[ \left( \sum_{m=1}^{N} \frac{1}{m} \right)^2 + \sum_{m=1}^{N} \frac{1}{m^2} \right]$$

$$= N^2 \left( H_N^2 + \sum_{m=1}^{N} \frac{1}{m^2} \right) \tag{2.29}$$

Τώρα θα υπολογίσουμε το ασυμπτωτικό ανάπτυγμα του $E\,[T_N]$ με τη βοήθεια των σχέσεων (1.20) και (1.21). Για $f(t) = \frac{1}{t+1}$ και $F(n) = \sum_{t=0}^{n-1} \frac{1}{t+1}$ από την (1.20), έχοντας αντικαταστήσει στο δεξιό μέλος της όπου n το $n-1$, παίρνουμε

$$F(n) \sim \frac{1}{2n} + \int_0^{n-1} \frac{1}{t+1}\,dt + c + \frac{1/6}{2!}\left( -\frac{1}{n^2} \right) + \frac{-1/30}{4!}\left( -\frac{6}{n^4} \right) + \ldots\,, \quad n \to \infty$$

Ισοδύναμα,

$$F(n) \sim \frac{1}{2n} + \ln n + c - \frac{1}{12n^2} + \frac{1}{120n^4} + \ldots\,, \quad n \to \infty \tag{2.30}$$

Για τον υπολογισμό της σταθεράς c, η οποία ορίζεται από τη σχέση (1.21), από τη (2.30) έχουμε

$$\lim_{n \to \infty} F(n) = \lim_{n \to \infty} \ln n + c$$

Οπότε

$$c = \lim_{n \to \infty} \left( \sum_{t=0}^{n-1} \frac{1}{t+1} - \ln n \right) = \lim_{n \to \infty} \left( \sum_{t=1}^{n} \frac{1}{t} - \ln n \right) \tag{2.31}$$

Από τη (2.31) συμπεραίνουμε ότι η σταθερά c είναι ίση με τη σταθερά του Euler. Δηλαδή $c = \gamma$. Συνεπώς η σχέση (2.30), λαμβάνοντας υπόψη ότι $c = \gamma$, αντικαθιστώντας όπου n το N και αναδιατάσσοντας τους όρους του δεξιού μέλους της από το μεγαλύτερο προς το μικρότερο καθώς το $N \to \infty$, γίνεται

$$F(N) \sim \ln N + \gamma + \frac{1}{2N} - \frac{1}{12N^2} + \frac{1}{120N^4} + \ldots\,, \quad N \to \infty \tag{2.32}$$

Επειδή

$$F(N)=\sum_{t=0}^{N-1}\frac{1}{t+1}=\sum_{t=1}^{N}\frac{1}{t}=H_N$$

μπορούμε να γράψουμε ότι

$$E\left[T_N\right]=NH_N\sim N\ln N+\gamma N+\frac{1}{2}-\frac{1}{12N}+\frac{1}{120N^3}+\dots\,,\quad N\to\infty \qquad (2.33)$$

Χρησιμοποιώντας το σύμβολο 'O', η (2.33) λαμβάνει τη μορφή

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

Ο υπολογισμός του ασυμπτωτικού αναπτύγματος του $E\,[T_N(T_N+1)]$ γίνεται με ανάλογο τρόπο. Για $f(t)=\frac{1}{(t+1)^2}$ και $F(n)=\sum_{t=0}^{n-1}\frac{1}{(t+1)^2}$, από την (1.20), έχοντας και σε αυτή την περίπτωση αντικαταστήσει στο δεξιό μέλος της όπου n το n – 1, παίρνουμε

$$F(n)\sim\frac{1}{2n^2}+\int_0^{n-1}\frac{1}{(t+1)^2}dt+c+\frac{1/6}{2!}\left(-\frac{2}{n^3}\right)+\frac{-1/30}{4!}\left(-\frac{24}{n^5}\right)+\dots\,,\quad n\to\infty$$

Ισοδύναμα,

$$F(n)\sim\frac{1}{2n^2}-\frac{1}{n}+1+c-\frac{1}{6n^3}+\frac{1}{30n^5}+\dots\,,\quad n\to\infty \qquad (2.34)$$

Για τον υπολογισμό της σταθεράς c, από τη (2.34) έχουμε

$$\lim_{n\to\infty}F(n)=1+c$$

Οπότε

$$c=\lim_{n\to\infty}\left[\sum_{t=0}^{n-1}\frac{1}{(t+1)^2}\right]-1$$

$$=\lim_{n\to\infty}\left[\sum_{t=1}^{n}\frac{1}{t^2}\right]-1=\frac{\pi^2}{6}-1 \qquad (2.35)$$

Συνεπώς η σχέση (2.34), λαμβάνοντας υπόψη τη (2.35), αντικαθιστώντας όπου n το N και αναδιατάσσοντας τους όρους του δεξιού μέλους της από το μεγαλύτερο προς το μικρότερο, καθώς το $N \to \infty$, γίνεται

$$F(N) \sim \frac{\pi^2}{6} - \frac{1}{N} + \frac{1}{2N^2} - \frac{1}{6N^3} + \frac{1}{30N^5} + \dots , \quad N \to \infty$$

Επειδή

$$F(N) = \sum_{t=0}^{N-1} \frac{1}{(t+1)^2} = \sum_{m=1}^{N} \frac{1}{m^2}$$

μπορούμε να γράψουμε

$$\sum_{m=1}^{N} \frac{1}{m^2} \sim \frac{\pi^2}{6} - \frac{1}{N} + \frac{1}{2N^2} - \frac{1}{6N^3} + \frac{1}{30N^5} + \dots , \quad N \to \infty \quad (2.36)$$

Αφού

$$H_N \sim \ln N + \gamma + \frac{1}{2N} - \frac{1}{12N^2} + \dots , \quad N \to \infty$$

προκύπτει το συμπέρασμα ότι

$$H_N^2 \sim \left( \ln N + \gamma + \frac{1}{2N} - \frac{1}{12N^2} + \dots \right)^2 , \quad N \to \infty$$

Αναπτύσσοντας το τετράγωνο του δεξιού μέλους της παραπάνω σχέσης βρίσκουμε

$$H_N^2 \sim (\ln N)^2 + \gamma^2 + \frac{1}{4N^2} + \frac{1}{144N^4} + 2\gamma \ln N +$$
$$+ \frac{\ln N}{N} - \frac{\ln N}{6N^2} + \frac{\gamma}{N} - \frac{\gamma}{6N^2} - \frac{1}{12N^3} + \dots \quad (2.37)$$

Από την Εξίσωση (2.29), χρησιμοποιώντας τις σχέσεις (2.36), (2.37), εκτελώντας τις πράξεις και αναδιατάσσοντας τους όρους του δεξιού μέλους της από το μεγαλύτερο προς το μικρότερο, καθώς το $N \to \infty$, συνεπάγεται

$$E\left[T_N(T_N + 1)\right] \sim N^2 \left[ (\ln N)^2 + 2\gamma \ln N + \frac{\pi^2}{6} + \gamma^2 + \frac{\ln N}{N} - \right.$$

$$-\frac{1}{N}+\frac{\gamma}{N}-\frac{\ln N}{6N^2}+\frac{3}{4N^2}-\frac{\gamma}{6N^2}-$$

$$\left.-\frac{1}{4N^3}+\frac{1}{144N^4}+\frac{1}{30N^5}+\ldots\right], \quad N\to\infty \qquad (2.38)$$

Χρησιμοποιώντας το σύμβολο 'O', η (2.38) λαμβάνει τη μορφή

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

Για τον υπολογισμό του ασυμπτωτικού αναπτύγματος του V $[T_N]$ θα χρησιμοποιήσουμε τις σχέσεις

$$E\left[T_N\right]\sim N\ln N+\gamma N+\frac{1}{2}-\frac{1}{12N}+\ldots, \quad N\to\infty \qquad (2.39)$$

και

$$E\left[T_N\left(T_N+1\right)\right]\sim N^2\left[\left(\ln N\right)^2+2\gamma\ln N+\frac{\pi^2}{6}+\gamma^2+\frac{\ln N}{N}-\frac{1}{N}+\frac{\gamma}{N}-\frac{\ln N}{6N^2}+\right.$$

$$\left.+\frac{3}{4N^2}-\frac{\gamma}{6N^2}-\frac{1}{4N^3}+\frac{1}{144N^4}+\ldots\right], \quad N\to\infty \qquad (2.40)$$

οι οποίες προκύπτουν από τις (2.33) και (2.38) αντίστοιχα αν παραλείψουμε τον υπολογισθέντα τελευταίο όρο από τα δεξιά μέλη τους. Από την Εξίσωση (2.21), μετά από την αντικατάσταση των ασυμπτωτικών σχέσεων (2.39) και (2.40), έχουμε

$$V\left[T_N\right]\sim\left(N\ln N\right)^2+2\gamma N^2\ln N+\frac{\pi^2}{6}N^2+\left(\gamma N\right)^2+N\ln N-N+$$

$$+\gamma N-\frac{\ln N}{6}+\frac{3}{4}-\frac{\gamma}{6}-\frac{1}{4N}+\frac{1}{144N^2}-N\ln N-\gamma N-\frac{1}{2}+$$

$$+\frac{1}{12N}-\left(N\ln N\right)^2-\left(\gamma N\right)^2-\frac{1}{4}-\frac{1}{144N^2}-2\gamma N^2\ln N-$$

$$-N\ln N+\frac{\ln N}{6}-\gamma N+\frac{\gamma}{6}+\frac{1}{12N}+\ldots, \quad N\to\infty \qquad (2.41)$$

Από τη (2.41), εκτελώντας τις πράξεις και αναδιατάσσοντας τους όρους του δεξιού μέλους της από το μεγαλύτερο προς το μικρότερο, καθώς το $N\to\infty$, συνεπάγεται

$$V\left[T_N\right] \sim \frac{\pi^2}{6}N^2 - N\ln N - N - \gamma N - \frac{1}{12N} + \dots, \quad N \to \infty \quad (2.42)$$

Χρησιμοποιώντας το σύμβολο 'O', η (2.42) λαμβάνει τη μορφή

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

## 2.3. Το Πρόβλημα της Επιλογής του Συλλέκτη Κουπονιών από τη Σκοπιά των Οριακών Θεωρημάτων

Έστω ότι υπάρχουν n κουπόνια διαφορετικά μεταξύ τους. Υποθέτουμε ότι το i – οστό κουπόνι επιλέγεται τυχαία και ανεξάρτητα από αυτά που έχουν ήδη επιλεγεί. Ορίζουμε οι τυχαίες μεταβλητές $X_1, X_2, \dots$ να είναι ανεξάρτητες και όμοια κατανεμημένες και να λαμβάνουν τιμές από το σύνολο $\{1,2, \dots ,n\}$.

### 2.3.1. Η Ασυμπτωτική Συμπεριφορά του Χρόνου Συλλογής των n Κουπονιών

Ορίζουμε με $\tau_\kappa^n$ τη χρονική στιγμή κατά την οποία έχουμε για πρώτη φορά κ διαφορετικά μεταξύ τους κουπόνια. Συμβολικά μπορούμε να γράψουμε $\tau_\kappa^n = \inf\left\{ m : \left| \left\{X_1, \dots, X_m\right\} \right| = \kappa \right\}$. Οπότε ο χρόνος, που απαιτείται για τη συλλογή όλων των τύπων των κουπονιών, είναι το $\tau_n^n$, το οποίο αναπαρίσταται και ως $T_n$. Από τον ορισμό του $\tau_\kappa^n$ προκύπτει ότι όταν $\kappa = 1$ τότε $\tau_1^n = 1$. Για $\kappa = 0$ ορίζουμε $\tau_0^n = 0$. Το χρονικό διάστημα, που μεσολαβεί από τη χρονική στιγμή που υπάρχουν στη συλλογή μας $\kappa - 1$ διαφορετικά μεταξύ τους κουπόνια, μέχρι τη χρονική στιγμή που θα επιλέξουμε ένα κουπόνι διαφορετικού τύπου από αυτούς που προηγήθηκαν, συμβολίζεται με $X_{n,\kappa}$. Λαμβάνοντας υπόψη τον ορισμό του $\tau_\kappa^n$, έχουμε $X_{n,\kappa} \equiv \tau_\kappa^n - \tau_{\kappa-1}^n$ για $1 \le \kappa \le n$. Από την Παράγραφο 1.1. και από την έννοια της τυχαίας μεταβλητής $X_{n,\kappa}$ συμπεραίνουμε ότι ακολουθεί τη Γεωμετρική κατανομή. Η πιθανότητα να επιλέξουμε ένα κουπόνι διαφορετικού τύπου από εκείνους των κουπονιών που προηγήθηκαν υπολογίζεται ως εξής:

$$p = \frac{\binom{n-(\kappa-1)}{1}}{\binom{n}{1}} = \frac{n-(\kappa-1)}{n} = 1 - \frac{\kappa-1}{n} \quad (2.43)$$

Η (2.43) δίνει την παράμετρο της Γεωμετρικής κατανομής που ακολουθεί η $X_{n,\kappa}$. Στο σημείο αυτό θα πρέπει να παρατηρήσουμε ότι η $X_{n,\kappa}$, $1 \le \kappa \le n$, είναι ανεξάρτητη από τις τυχαίες μεταβλητές $X_{n,j}$, $1 \le j < \kappa$, διότι τα χρονικά διαστήματα δεν αλληλεπικαλύπτονται, όπως πολύ εύκολα μπορούμε να συμπεράνουμε από τον ορισμό τους, και η χρονική στιγμή που επιλέγεται ένα κουπόνι διαφορετικού τύπου από τους τύπους των κουπονιών που προηγήθηκαν δεν εξαρτάται από τις χρονικές στιγμές που επιλέγονται τα προηγούμενα διαφορετικού τύπου κουπόνια.

Από την Παράγραφο 1.1. έχουμε ότι όταν μία τυχαία μεταβλητή X ακολουθεί τη Γεωμετρική κατανομή με παράμετρο p τότε $E[X] = \frac{1}{p}$ και $V[X] = \frac{1-p}{p^2}$. Από την τελευταία σχέση συνεπάγεται ότι $V[X] \le \frac{1}{p^2}$. Χρησιμοποιώντας την ιδιότητα της γραμμικότητας της μέσης τιμής,

$$E[T_n] = E\left[\sum_{\kappa=1}^{n} X_{n,\kappa}\right] = \sum_{\kappa=1}^{n} E[X_{n,\kappa}]$$

$$= \sum_{\kappa=1}^{n}\left(1 - \frac{\kappa-1}{n}\right)^{-1} = n\sum_{\kappa=1}^{n}\frac{1}{n-\kappa+1}$$

$$= n\sum_{m=1}^{n}\frac{1}{m} \tag{2.44}$$

Από την ανισοτική σχέση (Π.15) συνεπάγεται

$$n\ln n \le n\sum_{m=1}^{n}\frac{1}{m} \le n + n\ln n \tag{2.45}$$

Επειδή η (2.45) ισχύει για κάθε $n \in \aleph$, θα ισχύει και καθώς το $n \to \infty$. Οπότε έχουμε

$$\lim_{n\to\infty}(n\ln n) \le \lim_{n\to\infty}\left(n\sum_{m=1}^{n}\frac{1}{m}\right) \le \lim_{n\to\infty}(n + n\ln n) = \lim_{n\to\infty}(n\ln n) \tag{2.46}$$

Από τη (2.46) προκύπτει

$$\lim_{n\to\infty}\left(n\sum_{m=1}^{n}\frac{1}{m}\right) = \lim_{n\to\infty}(n\ln n) \tag{2.47}$$

Συνεπώς από τις σχέσεις (2.44) και (2.47) παίρνουμε

$$E[T_n] = n\sum_{m=1}^{n}\frac{1}{m} \sim n\ln n, \quad n\to\infty \tag{2.48}$$

Από τον ορισμό της τυχαίας μεταβλητής $T_n$ και από το Λήμμα Π.1,

$$\begin{aligned} V[T_n] &= V\left[\sum_{\kappa=1}^{n} X_{n,\kappa}\right] = \sum_{\kappa=1}^{n} V[X_{n,\kappa}] \\ &\le \sum_{\kappa=1}^{n}\left(1-\frac{\kappa-1}{n}\right)^{-2} = n^2\sum_{\kappa=1}^{n}\frac{1}{(n-\kappa+1)^2} \\ &= n^2\sum_{m=1}^{n}\frac{1}{m^2} \le n^2\sum_{m=1}^{\infty}\frac{1}{m^2} \\ &= n^2\frac{\pi^2}{6} \end{aligned} \tag{2.49}$$

Ορίζουμε $\mu_n = E[T_n]$, $\sigma_n^2 = V[T_n]$ και $b_n = n\ln n$. Θα δείξουμε ότι $\frac{\sigma_n^2}{b_n^2} \to$
$\to 0$. Πράγματι

$$0 \le \frac{\sigma_n^2}{b_n^2} = \frac{V[T_n]}{n^2(\ln n)^2} \le \frac{n^2\frac{\pi^2}{6}}{n^2(\ln n)^2} = \frac{\frac{\pi^2}{6}}{(\ln n)^2} \to 0, \quad n\to\infty$$

Επειδή ισχύει η υπόθεση του Θεωρήματος Π.2 έχουμε

$$\frac{T_n-\mu_n}{b_n} = \frac{T_n - n\sum_{m=1}^{n} m^{-1}}{n\ln n} \to 0 \text{ κατά πιθανότητα} \tag{2.50}$$

Από τη σχέση (2.48)

$$\lim_{n\to\infty}\frac{n\sum_{m=1}^{n} m^{-1}}{n\ln n} = 1 \tag{2.51}$$

Η (2.50), λόγω της (2.51), συνεπάγεται ότι $\frac{T_n}{n\ln n} \to 1$ κατά πιθανότητα.

### 2.3.2. Η Κατανομή Poisson και το Πρόβλημα της Επιλογής του Συλλέκτη

Ας υποθέσουμε ότι έχουμε r μπάλες τις οποίες τοποθετούμε τυχαία μέσα σε n κουτιά.

**Θεώρημα 2.1**: Αν $ne^{-r/n} \to \lambda \in [0\,,\infty)$ τότε ο αριθμός των άδειων κουτιών προσεγγίζει την κατανομή Poisson με μέση τιμή λ.

**Απόδειξη**: Αρχικά παρατηρούμε ότι

$$P\{\text{τα κουτιά } i_1, i_2, \dots, i_\kappa \text{ είναι άδεια}\} = \frac{\binom{n-\kappa}{1}}{\binom{n}{1}} \cdot \frac{\binom{n-\kappa}{1}}{\binom{n}{1}} \cdot \;\dots\; \cdot \frac{\binom{n-\kappa}{1}}{\binom{n}{1}}$$

$1^{\eta}$ μπάλα, $2^{\eta}$ μπάλα, ... , r − οστή μπάλα

$$= \left(\frac{n-\kappa}{n}\right)^r = \left(1 - \frac{\kappa}{n}\right)^r \tag{2.52}$$

Συμβολίζουμε με $p_m(r,n)$ την πιθανότητα ακριβώς m κουτιά να είναι άδεια όταν r μπάλες τοποθετούνται μέσα σε n κουτιά. Τότε

$$P\left(\{\text{0 άδεια κουτιά}\} \cup \{\text{τουλάχιστον 1 άδειο κουτί}\}\right) = 1$$

Ισοδύναμα,

$$P\{\text{0 άδεια κουτιά}\} = 1 - P\{\text{τουλάχιστον 1 άδειο κουτί}\} \tag{2.53}$$

Χρησιμοποιώντας το Γενικευμένο αθροιστικό τύπο και την Εξίσωση (2.52) παίρνουμε

$$\begin{aligned}
P\{\text{τουλ. 1 κ. άδειο}\} &= P\left\{1^{o} \text{ κ. άδειο}\right\} + \dots + P\left\{n^{o} \text{ κ. άδειο}\right\} - P\left\{1^{o}, 2^{o} \text{ κ. άδεια}\right\} \\
&- \dots - P\left\{(n-1)^{o}, n^{o} \text{ κ. άδεια}\right\} + P\left\{1^{o}, 2^{o}, 3^{o} \text{ κ. άδεια}\right\} + \dots \\
&\dots + P\left\{(n-2)^{o}, (n-1)^{o}, n^{o} \text{ κ. άδεια}\right\} - \dots + \\
&+ (-1)^{n-1} P\left\{1^{o}, \dots, n^{o} \text{ κ. άδεια}\right\} \\
&= \binom{n}{1} P\{\text{το } i_1 \text{ κ. άδειο}\} - \binom{n}{2} P\{\text{τα } i_1, i_2 \text{ κ. άδεια}\} + \\
&+ \binom{n}{3} P\{\text{τα } i_1, i_2, i_3 \text{ κ. άδεια}\} - \dots +
\end{aligned}$$

$$+(-1)^{n-1}\binom{n}{n}P\{\text{τα } i_1,\ldots,i_n \text{ κ. άδεια}\}$$

$$=\sum_{\kappa=1}^{n}(-1)^{\kappa-1}\binom{n}{\kappa}P\{\text{τα } i_1,\ldots,i_\kappa \text{ κ. άδεια}\}$$

$$=\sum_{\kappa=1}^{n}(-1)^{\kappa-1}\binom{n}{\kappa}\left(1-\frac{\kappa}{n}\right)^r \qquad (2.54)$$

Από τη (2.53), λόγω της (2.54), έχουμε

$$p_0(r,n)=1-\sum_{\kappa=1}^{n}(-1)^{\kappa-1}\binom{n}{\kappa}\left(1-\frac{\kappa}{n}\right)^r$$

$$=\sum_{\kappa=0}^{n}(-1)^{\kappa}\binom{n}{\kappa}\left(1-\frac{\kappa}{n}\right)^r \qquad (2.55)$$

Τώρα θα υπολογίσουμε την πιθανότητα $p_m(r,n)$. Από τον ορισμό της συνεπάγεται

$$p_m(r,n)=\binom{n}{m}P\left(\{\text{τα κ. } i_1,\ldots,i_m \text{ άδεια }\}\cap\{\text{τα κ. } i_{m+1},\ldots,i_n \text{ γεμάτα όταν r μπάλες τοποθετούνται σε n κουτιά }\}\right)$$

$$=\binom{n}{m}P\{\text{τα κ. } i_1,\ldots,i_m \text{ άδεια }\}P\left(\{\text{τα κ. } i_{m+1},\ldots,i_n \text{ γεμάτα όταν r μπάλες τοποθετούνται σε n κουτιά }\}\,\middle|\,\{\text{τα κ. } i_1,\ldots,i_m \text{ άδεια }\}\right)$$

$$=\binom{n}{m}P\{\text{τα κ. } i_1,\ldots,i_m \text{ άδεια }\}P\{\text{0 άδεια κουτιά όταν r μπάλες τοποθετούνται σε } (n-m) \text{ κουτιά }\}$$

$$=\binom{n}{m}\left(1-\frac{m}{n}\right)^r p_0(r,n-m) \qquad (2.56)$$

Θέλουμε να δείξουμε ότι $p_m(r,n)\to e^{-\lambda}\frac{\lambda^m}{m!}$ καθώς το $n\to\infty$. Θα ξεκινήσουμε πρώτα δείχνοντας ότι αν $ne^{-r/n}\to\lambda$ τότε

$$\binom{n}{m}\left(1-\frac{m}{n}\right)^r\to\frac{\lambda^m}{m!} \qquad (2.57)$$

Επειδή $1-x \le e^{-x}$ και $ne^{-r/n} \to \lambda$ έχουμε

$$\binom{n}{m}\left(1-\frac{m}{n}\right)^r = \frac{(n-m)!(n-m+1)\cdot\ldots\cdot n}{m!(n-m)!}\left(1-\frac{m}{n}\right)^r$$

$$\le \frac{n^m}{m!}\left(e^{-m/n}\right)^r = \frac{\left(ne^{-r/n}\right)^m}{m!} \to \frac{\lambda^m}{m!} \qquad (2.58)$$

Παρατηρούμε ότι

$$\binom{n}{m} = \frac{(n-m)!(n-m+1)\cdot\ldots\cdot n}{m!(n-m)!} \ge \frac{(n-m)^m}{m!}$$

οπότε

$$\binom{n}{m}\left(1-\frac{m}{n}\right)^r \ge \frac{(n-m)^m}{m!}\left(1-\frac{m}{n}\right)^r = \left(1-\frac{m}{n}\right)^{m+r}\frac{n^m}{m!} \qquad (2.59)$$

Αν $0 \le t \le \frac{1}{2}$ τότε

$$\ln(1-t) = -t - \frac{t^2}{2} - \frac{t^3}{3} - \frac{t^4}{4} - \ldots \ge -t - \frac{t^2}{2} - \frac{t^2}{4} - \frac{t^2}{8} - \ldots$$

$$= -t - \frac{t^2}{2}\left(1+\frac{1}{2}+\frac{1}{4}+\ldots\right) = -t - \frac{t^2}{2}\frac{1}{1-\frac{1}{2}} = -t - t^2$$

Συνεπώς

$$\ln\left(n^m\left(1-\frac{m}{n}\right)^r\right) = m\ln n + r\ln\left(1-\frac{m}{n}\right) \ge m\ln n + r\left(-\frac{m}{n} - \left(\frac{m}{n}\right)^2\right)$$

$$= m\ln n - r\left(\frac{m}{n}\right) - r\left(\frac{m}{n}\right)^2 \qquad (2.60)$$

Από υπόθεση

$$ne^{-r/n} \to \lambda \Leftrightarrow \ln\left(ne^{-r/n}\right) \to \ln\lambda \Leftrightarrow -\frac{r}{n} + \ln n \to \ln\lambda$$

$$\Leftrightarrow \frac{r - n\ln n + n\ln\lambda}{n} \to 0 \Leftrightarrow r - n\ln n + n\ln\lambda = o(n)$$

$$\Leftrightarrow r = n\ln n - n\ln\lambda + o(n) \tag{2.61}$$

Οπότε

$$r\left(\frac{m}{n}\right)^2 = \left(n\ln n - n\ln\lambda + o(n)\right)\frac{m^2}{n^2}$$

$$= m^2\frac{\ln n}{n} - m^2\frac{\ln\lambda}{n} + m^2\frac{f(n)}{n}\cdot\frac{1}{n} \to 0 \tag{2.62}$$

Από την Εξίσωση (2.61), πολλαπλασιάζοντας με $\frac{m}{n}$, παίρνουμε

$$r\left(\frac{m}{n}\right) = m\ln n - m\ln\lambda + \frac{m}{n}o(n) \quad \Leftrightarrow$$

$$-\frac{n}{m}\left(m\ln n - \frac{rm}{n} - m\ln\lambda\right) = o(n) \quad \Leftrightarrow$$

$$-\frac{\frac{n}{m}\left(m\ln n - \frac{rm}{n} - m\ln\lambda\right)}{n} \to 0 \quad \Leftrightarrow$$

$$m\ln n - \frac{rm}{n} \to m\ln\lambda \tag{2.63}$$

Από τις (2.60), (2.62) και (2.63) έχουμε

$$\ln\left(n^m\left(1-\frac{m}{n}\right)^r\right) \geq m\ln n - \frac{rm}{n} - r\left(\frac{m}{n}\right)^2 \to m\ln\lambda \tag{2.64}$$

Η (2.64) συνεπάγεται

$$\liminf_{n\to\infty}\left[\ln\left(n^m\left(1-\frac{m}{n}\right)^r\right)\right] \geq \liminf_{n\to\infty}\left[m\ln n - \frac{rm}{n} - r\left(\frac{m}{n}\right)^2\right]$$

$$= m\ln\lambda$$

Συνεπώς

$$\ln\left[\liminf_{n\to\infty}\left(n^m\left(1-\frac{m}{n}\right)^r\right)\right] \geq \ln\lambda^m \quad \Leftrightarrow$$

$$\liminf_{n\to\infty}\left(n^m\left(1-\frac{m}{n}\right)^r\right)\geq\lambda^m \tag{2.65}$$

Επειδή $\left(1-\frac{m}{n}\right)^m\to 1$ καθώς το $n\to\infty$, από τις (2.59) και (2.65) παίρνουμε

$$\begin{aligned}\liminf_{n\to\infty}\left[\binom{n}{m}\left(1-\frac{m}{n}\right)^r\right] &\geq \liminf_{n\to\infty}\left\{\frac{1}{m!}\left[\left(1-\frac{m}{n}\right)^m\right]\left[n^m\left(1-\frac{m}{n}\right)^r\right]\right\}\\ &\geq\frac{1}{m!}\cdot 1\cdot\liminf_{n\to\infty}\left[n^m\left(1-\frac{m}{n}\right)^r\right]\\ &\geq\frac{\lambda^m}{m!}\end{aligned} \tag{2.66}$$

Από τη σχέση (2.58) έχουμε

$$\begin{aligned}\limsup_{n\to\infty}\left[\binom{n}{m}\left(1-\frac{m}{n}\right)^r\right] &\leq \limsup_{n\to\infty}\left[\frac{\left(ne^{-r/n}\right)^m}{m!}\right]\\ &=\frac{\lambda^m}{m!}\end{aligned} \tag{2.67}$$

Συνδυάζοντας τις (2.66) και (2.67) προκύπτει

$$\frac{\lambda^m}{m!}\leq\liminf_{n\to\infty}\left[\binom{n}{m}\left(1-\frac{m}{n}\right)^r\right]\leq\limsup_{n\to\infty}\left[\binom{n}{m}\left(1-\frac{m}{n}\right)^r\right]\leq\frac{\lambda^m}{m!}$$

Από την τελευταία σχέση συνεπάγεται η (2.57).

Τώρα θα διατυπώσουμε το Θεώρημα Κυριαρχημένης Σύγκλισης και θα το χρησιμοποιήσουμε για να συνεχίσουμε την απόδειξη του θεωρήματος.

Θεώρημα Κυριαρχημένης Σύγκλισης

Έστω $(f_n)$ μία ακολουθία συναρτήσεων. Αν $f_n\to f$ σημειακά, $|f_n|\leq g$ για όλα τα n και $\sum_{\kappa=0}^{\infty} g(\kappa)<\infty$ τότε $\sum_{\kappa=0}^{n} f_n(\kappa)\to\sum_{\kappa=0}^{\infty} f(\kappa)$ καθώς το $n\to\infty$.

Θεωρούμε την ακολουθία των συναρτήσεων $(f_n)$, που ορίζεται με τον τύπο

$$f_n(\kappa)=(-1)^\kappa \binom{n}{\kappa}\left(1-\frac{\kappa}{n}\right)^r, \quad n\in\aleph, \ \kappa\in\aleph_0$$

και τη συνάρτηση f, που ορίζεται με

$$f(\kappa)=(-1)^\kappa \frac{\lambda^\kappa}{\kappa!}, \quad \kappa\in\aleph_0$$

Από τη σχέση (2.57) έχουμε

$$(-1)^\kappa \binom{n}{\kappa}\left(1-\frac{\kappa}{n}\right)^r \to (-1)^\kappa \frac{\lambda^\kappa}{\kappa!}, \quad n\to\infty$$

για τυχαίο κ. Οπότε $f_n(\kappa)\to f(\kappa)$ για κάθε $\kappa\in\aleph_0$. Συνεπώς $f_n\to f$ σημειακά.
Από τον ορισμό της ακολουθίας των συναρτήσεων $(f_n)$ και από τη σχέση (2.58) παίρνουμε

$$|f_n(\kappa)|=\left|(-1)^\kappa \binom{n}{\kappa}\left(1-\frac{\kappa}{n}\right)^r\right|=\binom{n}{\kappa}\left(1-\frac{\kappa}{n}\right)^r\le\frac{\left(ne^{-r/n}\right)^\kappa}{\kappa!} \tag{2.68}$$

Επειδή $ne^{-r/n}\to\lambda$ συνεπάγεται

$$ne^{-r/n}\le\lambda+1 \tag{2.69}$$

Από τη (2.68), λόγω της (2.69), προκύπτει

$$|f_n(\kappa)|\le\frac{(\lambda+1)^\kappa}{\kappa!} \tag{2.70}$$

Ορίζουμε τη συνάρτηση g με $g(\kappa)=\frac{(\lambda+1)^\kappa}{\kappa!}$. Από τη (2.70) έχουμε ότι $|f_n(\kappa)|\le g(\kappa)$ για τυχαίο κ. Οπότε $|f_n|\le g$ για όλα τα n.
Από τον ορισμό της συνάρτησης g,

$$\sum_{\kappa=0}^{\infty} g(\kappa)=\sum_{\kappa=0}^{\infty}\frac{(\lambda+1)^\kappa}{\kappa!}=e^{\lambda+1}<\infty$$

Επειδή ισχύουν οι υποθέσεις του Θεωρήματος Κυριαρχημένης Σύγκλισης,

$$\sum_{\kappa=0}^{n}(-1)^{\kappa}\binom{n}{\kappa}\left(1-\frac{\kappa}{n}\right)^{r}\to\sum_{\kappa=0}^{\infty}(-1)^{\kappa}\frac{\lambda^{\kappa}}{\kappa!}=e^{-\lambda} \tag{2.71}$$

Από τη σχέση (2.71), λόγω της (2.55), παίρνουμε

$$p_0(r,n)\to e^{-\lambda} \tag{2.72}$$

Η (2.72) ισχύει αν $ne^{-r/n}\to\lambda$. Για σταθερό m, $(n-m)e^{-r/(n-m)}\to\lambda$, οπότε από (2.72)

$$p_0(r,n-m)\to e^{-\lambda} \tag{2.73}$$

Από τις σχέσεις (2.56), (2.57) και (2.73) προκύπτει

$$p_m(r,n)\to e^{-\lambda}\frac{\lambda^m}{m!},\quad n\to\infty$$

και ολοκληρώνεται η απόδειξη του θεωρήματος.

Στο πρόβλημα της επιλογής του συλλέκτη κουπονιών υπενθυμίζουμε ότι οι τυχαίες μεταβλητές $X_1, X_2, \ldots$ είναι ανεξάρτητες και όμοια κατανεμημένες και λαμβάνουν τιμές από το σύνολο $\{1,2, \ldots ,n\}$. Επίσης έχουμε συμβολίσει με $T_n$ το χρόνο που απαιτείται για τη συλλογή όλων των τύπων των κουπονιών και μπορούμε να γράψουμε $T_n=\inf\{\,m:\{\,X_1,\ldots,X_m\}=\{1,2,\ldots,n\,\}\}$. Στην παράγραφο αυτή οι $X_1, X_2, \ldots$ αναπαριστούν τις μπάλες που τοποθετούνται τυχαία μέσα σε n κουτιά. Σύμφωνα με αυτή τη θεώρηση το $T_n\le m$ αν και μόνο αν m μπάλες γεμίζουν και τα n κουτιά. Έστω τώρα ότι έχουμε r μπάλες με $r=n\ln n+nx$. Τότε

$$ne^{-r/n}=ne^{-\ln n-x}=e^{-x}\to e^{-x},\quad n\to\infty$$

Από το Θεώρημα 2.1 προκύπτει

$$P\{T_n-n\ln n\le nx\}=P\{T_n\le r\}=p_0(r,n)\to\exp\left(-e^{-x}\right)$$

# Παράρτημα

## Αποδείξεις που παραλήφθηκαν από το κυρίως κείμενο

### Π.1.

Μία οικογένεια τυχαίων μεταβλητών $X_i$, $i \in I$ με $E\left(X_i^2\right) < \infty$ ονομάζεται ασυσχέτιστη αν

$$E\left(X_i X_j\right) = EX_i \, EX_j, \quad i \neq j \tag{Π.1}$$

**Λήμμα Π.1**: Αν οι τυχαίες μεταβλητές $X_1, \ldots, X_n$ είναι ασυσχέτιστες και $E\left(X_i^2\right) < \infty$ για κάθε $i = 1, \ldots, n$ τότε

$$\mathrm{Var}\left(X_1 + \ldots + X_n\right) = \mathrm{Var}\left(X_1\right) + \ldots + \mathrm{Var}\left(X_n\right) \tag{Π.2}$$

**Απόδειξη**: Θέτουμε $\mu_i = EX_i$ $(i = 1, \ldots, n)$ και $S_n = \sum_{i=1}^n X_i$. Αφού $ES_n = \sum_{i=1}^n \mu_i$, χρησιμοποιώντας τον ορισμό της διασποράς έχουμε

$$\mathrm{Var}\left(S_n\right) = E\left(S_n - ES_n\right)^2 = E\left[\sum_{i=1}^n \left(X_i - \mu_i\right)\right]^2$$

Γράφοντας το τετράγωνο του αθροίσματος ως το γινόμενο δύο αθροισμάτων έχουμε

$$\mathrm{Var}\left(S_n\right) = E\left[\sum_{i=1}^n \left(X_i - \mu_i\right) \sum_{j=1}^n \left(X_j - \mu_j\right)\right]$$

$$= E\left[\sum_{i=1}^{n}\sum_{j=1}^{n}(X_i - \mu_i)(X_j - \mu_j)\right]$$

Χρησιμοποιώντας την ιδιότητα της γραμμικότητας της αναμενόμενης τιμής και αναπτύσσοντας το διπλό άθροισμα παίρνουμε

$$\operatorname{Var}(S_n) = \sum_{i=1}^{n} E\,(X_i - \mu_i)^2 + 2\sum_{i=1}^{n}\sum_{j=1}^{i-1} E\left[(X_i - \mu_i)(X_j - \mu_j)\right]$$

$$= \sum_{i=1}^{n} \operatorname{Var}(X_i) + 2\sum_{i=1}^{n}\sum_{j=1}^{i-1} E\left[(X_i - \mu_i)(X_j - \mu_j)\right] \qquad (\Pi.3)$$

Τώρα θα δείξουμε ότι ο δεύτερος όρος του δεξιού μέλους της ισότητας (Π.3) είναι ίσος με μηδέν. Πράγματι,

$$E\left[(X_i - \mu_i)(X_j - \mu_j)\right] = E\,(X_i X_j) - \mu_j E X_i - \mu_i E X_j + \mu_i \mu_j$$

$$= E\,(X_i X_j) - \mu_i \mu_j$$

$$= 0 \qquad (\Pi.4)$$

Η ισότητα (Π.4) είναι αληθής γιατί οι τυχαίες μεταβλητές $X_i$, $X_j$ είναι ασυσχέτιστες. Από τις ισότητες (Π.3), (Π.4) προκύπτει η (Π.2) και ολοκληρώνεται η απόδειξη του λήμματος.

## Π.2.

Θα αποδείξουμε ότι

$$\sum_{m=1}^{N}(-1)^{m-1}\binom{N}{m}\frac{1}{m} = \sum_{m=1}^{N}\frac{1}{m} \qquad (\Pi.5)$$

Έχουμε,

$$\sum_{m=0}^{N-1}(1-t)^m = 1 + (1-t) + (1-t)^2 + \ldots + (1-t)^{N-1}$$

$$= 1\cdot\frac{(1-t)^N - 1}{1-t-1} = \frac{1-(1-t)^N}{t} \qquad (\Pi.6)$$

Ολοκληρώνοντας το αριστερό μέλος της ισότητας (Π.6) παίρνουμε

$$\sum_{m=0}^{N-1}\int(1-t)^m\,dt=-\sum_{m=0}^{N-1}\frac{(1-t)^{m+1}}{m+1}+c_1$$

$$=-\sum_{m=1}^{N}\frac{(1-t)^m}{m}+c_1 \qquad (\Pi.7)$$

Τώρα ολοκληρώνουμε το δεξιό μέλος της (Π.6), οπότε

$$\int\frac{1-(1-t)^N}{t}\,dt=\int\frac{1}{t}\,dt-\int\frac{(1-t)^N}{t}\,dt \qquad (\Pi.8)$$

Χρησιμοποιώντας το θεώρημα του διωνύμου,

$$(1-t)^N=\sum_{m=0}^{N}(-1)^m\binom{N}{m}t^m \qquad (\Pi.9)$$

Αντικαθιστώντας την ισότητα (Π.9) στην (Π.8),

$$\int\frac{1-(1-t)^N}{t}\,dt=\int\frac{1}{t}\,dt-\sum_{m=0}^{N}(-1)^m\binom{N}{m}\int t^{m-1}\,dt$$

$$=\int\frac{1}{t}\,dt-\int\frac{1}{t}\,dt+\sum_{m=1}^{N}(-1)^{m-1}\binom{N}{m}\frac{t^m}{m}+c_2$$

$$=\sum_{m=1}^{N}(-1)^{m-1}\binom{N}{m}\frac{t^m}{m}+c_2 \qquad (\Pi.10)$$

Εξισώνοντας τα δεξιά μέλη των σχέσεων (Π.7), (Π.10),

$$-\sum_{m=1}^{N}\frac{(1-t)^m}{m}=\sum_{m=1}^{N}(-1)^{m-1}\binom{N}{m}\frac{t^m}{m}+c_3 \qquad (\Pi.11)$$

Για $t=0$, από την εξίσωση (Π.11) προκύπτει

$$c_3=-\sum_{m=1}^{N}\frac{1}{m} \qquad (\Pi.12)$$

Η (Π.11), λόγω της εξίσωσης (Π.12), γράφεται

$$-\sum_{m=1}^{N}\frac{(1-t)^m}{m}=\sum_{m=1}^{N}(-1)^{m-1}\binom{N}{m}\frac{t^m}{m}-\sum_{m=1}^{N}\frac{1}{m} \qquad (\Pi.13)$$

Για $t = 1$, από την (Π.13) προκύπτει η (Π.5) και ολοκληρώνεται η απόδειξη.

## Π.3.

Θα αποδείξουμε την ανισοτική σχέση

$$\sum_{m=2}^{n}\frac{1}{m}\le\int_{1}^{n}\frac{dx}{x}\le\sum_{m=1}^{n}\frac{1}{m} \qquad (\Pi.14)$$

Επίσης θα αποδείξουμε ότι από την (Π.14) προκύπτει

$$\ln n\le\sum_{m=1}^{n}\frac{1}{m}\le 1+\ln n \qquad (\Pi.15)$$

και αντίστροφα.
Έστω η συνάρτηση f με $f(x) = \frac{1}{x}$, $x > 0$. Θεωρούμε μία υποδιαίρεση του διαστήματος $[1,n]$ σε $n - 1$ μοναδιαίου μήκους διαστήματα. Τα συνολικά εμβαδά των εγγεγραμμένων και περιγεγραμμένων ορθογωνίων δίνονται αντίστοιχα, από τους τύπους

$$\sum_{m=2}^{n}f(m)=f(2)+f(3)+\ldots+f(n)$$

$$\sum_{m=1}^{n-1}f(m)=f(1)+f(2)+\ldots+f(n-1)$$

Είναι φανερό ότι το εμβαδό του χωρίου, που ορίζεται κάτω από τη γραφική παράσταση της f και μεταξύ των γραμμών $x = 1$, $x = n$ και $y = 0$ ικανοποιεί την ανισοτική σχέση

$$\sum_{m=2}^{n}\frac{1}{m}\le\int_{1}^{n}\frac{dx}{x}\le\sum_{m=1}^{n-1}\frac{1}{m}\le\sum_{m=1}^{n}\frac{1}{m} \qquad (\Pi.16)$$

Από την (Π.16) προκύπτει άμεσα η (Π.14). Από την τελευταία μπορούμε να γράψουμε ισοδύναμα

$$\sum_{m=2}^{n}\frac{1}{m}\le\int_{1}^{n}\frac{dx}{x}\quad \text{και}\quad \int_{1}^{n}\frac{dx}{x}\le\sum_{m=1}^{n}\frac{1}{m}\quad\Leftrightarrow$$

$$\sum_{m=1}^{n}\frac{1}{m}-1\le\ln n\quad \text{και}\quad \ln n\le\sum_{m=1}^{n}\frac{1}{m}\quad\Leftrightarrow$$

$$\sum_{m=1}^{n}\frac{1}{m}\le\ln n+1\quad \text{και}\quad \ln n\le\sum_{m=1}^{n}\frac{1}{m}$$

Από την τελευταία ανισοτική σχέση έχουμε την (Π.15). Το αντίστροφο ισχύει λόγω των ισοδυναμιών.

## Π.4.

Έστω X μία διακριτή τυχαία μεταβλητή και g μία συνάρτηση αυτής της τυχαίας μεταβλητής. Θα δείξουμε ότι

$$\sum_{\kappa=0}^{\infty}g(\kappa)P\{X=\kappa\}=g(0)+\sum_{\kappa=1}^{\infty}[g(\kappa)-g(\kappa-1)]P\{X\ge\kappa\}$$

Θα ξεκινήσουμε από το δεξιό μέλος της παραπάνω ισότητας και θα καταλήξουμε στο αριστερό. Πράγματι,

$$g(0)+\sum_{\kappa=1}^{\infty}[g(\kappa)-g(\kappa-1)]P\{X\ge\kappa\}=$$

$$=g(0)+\sum_{\kappa=1}^{\infty}[g(\kappa)-g(\kappa-1)]P[\{X=\kappa\}\cup\{X>\kappa\}]$$

$$=g(0)+\sum_{\kappa=1}^{\infty}([g(\kappa)-g(\kappa-1)][P\{X=\kappa\}+P\{X>\kappa\}])$$

$$=g(0)+\sum_{\kappa=1}^{\infty}g(\kappa)P\{X=\kappa\}+\sum_{\kappa=1}^{\infty}g(\kappa)P\{X>\kappa\}-\sum_{\kappa=1}^{\infty}g(\kappa-1)P\{X=\kappa\}-$$

$$-\sum_{\kappa=1}^{\infty}g(\kappa-1)P\{X>\kappa\}$$

$$=g(0)+\sum_{\kappa=1}^{\infty}g(\kappa)P\{X=\kappa\}+\sum_{\kappa=1}^{\infty}\left[g(\kappa)\sum_{\ell=\kappa+1}^{\infty}P\{X=\ell\}\right]-\sum_{\kappa=0}^{\infty}g(\kappa)P\{X=\kappa+1\}$$

$$-\sum_{\kappa=0}^{\infty}\left[g(\kappa)\sum_{\ell=\kappa+2}^{\infty}P\{X=\ell\}\right]$$

$$=g(0)+\sum_{\kappa=1}^{\infty}g(\kappa)P\{X=\kappa\}+\sum_{\kappa=1}^{\infty}\left[g(\kappa)\left(\sum_{\ell=\kappa+2}^{\infty}(P\{X=\ell\})+P\{X=\kappa+1\}\right)\right]-$$

$$-\sum_{\kappa=0}^{\infty}g(\kappa)P\{X=\kappa+1\}-g(0)\sum_{\ell=2}^{\infty}P\{X=\ell\}-\sum_{\kappa=1}^{\infty}\left[g(\kappa)\sum_{\ell=\kappa+2}^{\infty}P\{X=\ell\}\right]$$

$$=g(0)+\sum_{\kappa=1}^{\infty}g(\kappa)P\{X=\kappa\}+\sum_{\kappa=1}^{\infty}\left[g(\kappa)\sum_{\ell=\kappa+2}^{\infty}P\{X=\ell\}\right]+\sum_{\kappa=1}^{\infty}g(\kappa)P\{X=\kappa+1\}$$

$$-g(0)P\{X=1\}-\sum_{\kappa=1}^{\infty}g(\kappa)P\{X=\kappa+1\}-g(0)\sum_{\ell=2}^{\infty}P\{X=\ell\}-$$

$$-\sum_{\kappa=1}^{\infty}\left[g(\kappa)\sum_{\ell=\kappa+2}^{\infty}P\{X=\ell\}\right]$$

$$=g(0)+\sum_{\kappa=1}^{\infty}g(\kappa)P\{X=\kappa\}-g(0)\sum_{\ell=1}^{\infty}P\{X=\ell\}$$

$$=\sum_{\kappa=1}^{\infty}g(\kappa)P\{X=\kappa\}+g(0)\left(1-\sum_{\ell=1}^{\infty}P\{X=\ell\}\right)$$

$$=\sum_{\kappa=1}^{\infty}g(\kappa)P\{X=\kappa\}+g(0)P\{X=0\}$$

$$=\sum_{\kappa=0}^{\infty}g(\kappa)P\{X=\kappa\}$$

## Π.5.

Μία σημαντική ανισοτική σχέση, η οποία μας είναι απαραίτητη παρακάτω, είναι η Ανισότητα Chebyshev. Αυτή έχει ως εξής:

**Θεώρημα Π.1**: Έστω X τυχαία μεταβλητή και φ: $\Re \to \Re$ με φ(x) = = φ(–x) για κάθε $x \in \Re$, φ(x) > 0 για κάθε $x \in (0,+\infty)$ και φ αύξουσα στο $(0,+\infty)$. Αν επιπλέον E [φ(X)] < ∞, τότε

$$P\{|X| \ge x\} \le \frac{E[\varphi(X)]}{\varphi(x)} \text{ για κάθε } x > 0 \qquad (\Pi.17)$$

**Απόδειξη**: Μπορεί να βρεθεί στο [1] στη σελίδα 140.

Στο σημείο αυτό θα πρέπει να παρατηρήσουμε ότι επειδή $\{|X| > x\} \subseteq$ $\subseteq \{|X| \ge x\}$ έχουμε $P\{|X| > x\} \le P\{|X| \ge x\}$. Η τελευταία σχέση σε συνδυασμό με την (Π.17) μας οδηγεί στο συμπέρασμα

$$P\{|X| > x\} \le \frac{E[\varphi(X)]}{\varphi(x)} \text{ για κάθε } x > 0 \qquad (\Pi.18)$$

Έστω τώρα $X_1, X_2, \ldots$ ασυσχέτιστες τυχαίες μεταβλητές με $EX_i = \mu$ και $Var(X_i) < \infty$ (i = 1, 2, … ). Θέτουμε $S_n = X_1 + \ldots + X_n$ και $Z_n = \frac{S_n}{n} - \mu$.

**Λήμμα Π.2**: Αν κ > 0 και $E|Z_n|^\kappa \to 0$ καθώς το $n \to \infty$, τότε $Z_n \to 0$ κατά πιθανότητα.

**Απόδειξη**: Από τον ορισμό της σύγκλισης κατά πιθανότητα έχουμε

$$Z_n \xrightarrow{P} 0 \quad \Leftrightarrow \quad P\{|Z_n| > \varepsilon\} \to 0 \text{ για κάθε } \varepsilon > 0$$

Άρα αρκεί να αποδείξουμε το δεξιό μέλος της ισοδυναμίας. Θεωρούμε συνάρτηση φ με $\varphi(x) = |x|^\kappa$ και τυχαία μεταβλητή $X = Z_n$. Βλέπουμε ότι οι υποθέσεις του θεωρήματος Π.1 ικανοποιούνται, οπότε από την (Π.18) συνεπάγεται

$$P\{|Z_n| > \varepsilon\} \le \frac{E[\varphi(Z_n)]}{\varphi(\varepsilon)} \text{ για κάθε } \varepsilon > 0$$

Ισοδύναμα,

$$P\{|Z_n| > \varepsilon\} \le \varepsilon^{-\kappa} E|Z_n|^\kappa \text{ για κάθε } \varepsilon > 0$$

Για $n \to \infty$ έχουμε $\lim_n P\{|Z_n| > \varepsilon\} = 0$ για κάθε ε > 0 και ολοκληρώνεται η απόδειξη.

**Θεώρημα Π.2**: Έστω $\mu_n = ES_n$ και $\sigma_n^2 = \operatorname{Var}(S_n)$. Αν $\frac{\sigma_n^2}{b_n^2} \to 0$ καθώς το $n \to \infty$, τότε $\frac{S_n - \mu_n}{b_n} \to 0$ κατά πιθανότητα.

**Απόδειξη**: Θέτουμε $Z_n = \frac{S_n - \mu_n}{b_n}$. Έχουμε

$$E|Z_n|^2 = E\left(\frac{S_n - \mu_n}{b_n}\right)^2 = b_n^{-2}\, E(S_n - \mu_n)^2$$

$$= b_n^{-2}\, E(S_n - ES_n)^2 = b_n^{-2}\, \operatorname{Var}(S_n) = \frac{\sigma_n^2}{b_n^2} \to 0$$

καθώς το $n \to \infty$. Ικανοποιούνται λοιπόν οι υποθέσεις του λήμματος Π.2 και παίρνουμε το ζητούμενο.

# Πίνακας Συμβόλων

| | | |
|---|---|---|
| $\aleph$ | : | Το σύνολο των φυσικών αριθμών $1,2,\ldots,n,\ldots$ |
| $\aleph_0$ | : | Το σύνολο των φυσικών αριθμών και του μηδενός. |
| $\mathbb{Z}$ | : | Το σύνολο των ακέραιων αριθμών. |
| $\Re$ | : | Το σύνολο των πραγματικών αριθμών. |
| $\mathbb{C}$ | : | Το σύνολο των μιγαδικών αριθμών. |
| $n!$ | = | $1\cdot 2\cdot 3\cdot \ldots \cdot n$, αν $n$ φυσικός αριθμός, $0! = 1$. |
| $\alpha \in A$ | : | Το $\alpha$ είναι στοιχείο του συνόλου Α. |
| $A \bigcup B$ | : | Ένωση των συνόλων Α και Β. |
| $A \bigcap B$ | : | Τομή των συνόλων Α και Β. |
| $[A]^c$ ή $(A)^c$ | : | Το συμπλήρωμα του συνόλου Α. |
| $\{\ \}$ | : | Άγγιστρα συνόλου. |
| $\{\ \}$ | : | Άγγιστρα ενδεχομένου. |
| $\subseteq$ | : | Είναι υποσύνολο του … |
| $(\alpha,\beta)$ | : | Ανοικτό διάστημα με άκρα $\alpha$ και $\beta$. Το σύνολο όλων των πραγματικών αριθμών $x$ τέτοιων, ώστε $\alpha < x < \beta$. |
| $[\alpha,\beta]$ | : | Κλειστό διάστημα με άκρα $\alpha$ και $\beta$. Το σύνολο όλων των πραγματικών αριθμών $x$ τέτοιων, ώστε $\alpha \le x \le \beta$. |
| $[\alpha,\beta)$ | : | Το σύνολο όλων των πραγματικών αριθμών $x$ τέτοιων, ώστε $\alpha \le x < \beta$. |
| $(\alpha,\beta]$ | : | Το σύνολο όλων των πραγματικών αριθμών $x$ τέτοιων, ώστε $\alpha < x \le \beta$. |
| $[\alpha,+\infty)$ | : | Το σύνολο όλων των πραγματικών αριθμών $x$ τέτοιων, ώστε $x \ge \alpha$. |
| $(\alpha,+\infty)$ | : | Το σύνολο όλων των πραγματικών αριθμών $x$ τέτοιων, ώστε $x > \alpha$. |
| $(-\infty,\beta)$ | : | Το σύνολο όλων των πραγματικών αριθμών $x$ τέτοιων, ώστε $x \le \beta$. |
| $(-\infty,\beta)$ | : | Το σύνολο όλων των πραγματικών αριθμών $x$ τέτοιων, ώστε $x < \beta$. |

$\exp\{\dots\} \; = \; e^{\{\dots\}}$

$|\cdot|$ : Απόλυτη τιμή πραγματικού αριθμού.

$|\cdot|$ : Μέτρο μιγαδικού αριθμού.

<=> ή $\Leftrightarrow$ : Σύμβολο ισοδυναμίας.

$\forall$ : Για κάθε.

$\sim$ : Είναι «ασυμπτωτικά ισοδύναμη» στη …

$O(\dots)$ : Είναι «μεγάλο ο» της …

$o(\dots)$ : Είναι «μικρό ο» της …

$\rightarrow$ : Συγκλίνει.

$\rightarrow$ : Τείνει.

$f: A \rightarrow B$ : Συνάρτηση f με πεδίο (σύνολο) ορισμού το A (ή υποσύνολο του A) και πεδίο (σύνολο) τιμών το B (ή υποσύνολο του B).

$f(x)$ : Η τιμή της συνάρτησης f στο στοιχείο x του πεδίου ορισμού της.

$df(x)$ : Διαφορικό της συνάρτησης f.

$\lim\limits_{x \to x_0} f(x)$ : Η οριακή τιμή (το όριο) της συνάρτησης f, όταν το x τείνει στο $x_0$.

$\max\limits_{x \in A} f(x)$ : Μέγιστο του συνόλου των τιμών της $f(x)$, όπου x ανήκει στο A.

$\sup\limits_{x \in A} f(x)$ : Άνω πέρας του συνόλου των τιμών της $f(x)$, όπου x ανήκει στο A.

$\min\limits_{x \in A} f(x)$ : Ελάχιστο του συνόλου των τιμών της $f(x)$, όπου x ανήκει στο A.

$\inf\limits_{x \in A} f(x)$ : Κάτω πέρας του συνόλου των τιμών της $f(x)$, όπου x ανήκει στο A.

$$\sum_{\kappa=1}^{n} \alpha_\kappa \text{ ή } \textstyle\sum_{\kappa=1}^{n} \alpha_\kappa \; = \; \alpha_1 + \alpha_2 + \dots + \alpha_n$$

$$\sum_{\kappa \in \{j_1, j_2, \dots, j_m\}} \alpha_\kappa \; = \; \alpha_{j_1} + \alpha_{j_2} + \dots + \alpha_{j_m}$$

$$\prod_{\kappa=1}^{n} \alpha_\kappa \; = \; \alpha_1 \cdot \alpha_2 \cdot \dots \cdot \alpha_n$$

$$\prod_{\kappa \in \{j_1, j_2, \dots, j_m\}} \alpha_\kappa \; = \; \alpha_{j_1} \cdot \alpha_{j_2} \cdot \dots \cdot \alpha_{j_m}$$

$\frac{df}{dx}$ ή $f'(x)$ : Η παράγωγος της συνάρτησης f ως προς τη μεταβλητή x.

$\frac{d^2f}{dx^2}$ ή $f''(x)$ : Η παράγωγος δεύτερης τάξης της συνάρτησης f ως προς τη μεταβλητή x.

$\frac{d^nf}{dx^n}$ ή $f^{(n)}(x)$ : Η παράγωγος $n-$οστής τάξης της συνάρτησης f ως προς τη μεταβλητή x.

$\binom{\kappa}{n} = \frac{\kappa!}{n!(\kappa-n)!} \quad (\kappa \in \aleph,\ 0 \le n \le \kappa)$

$\int f(x)dx$ : Το αόριστο ολοκλήρωμα της συνάρτησης f ως προς τη μεταβλητή x.

$\int_\alpha^\beta f(x)dx$ : Το ορισμένο ολοκλήρωμα της συνάρτησης f ως προς τη μεταβλητή x στο κλειστό διάστημα [α,β].

$[f(x)]_\alpha^\beta = f(x)\big|_\alpha^\beta = f(\beta) - f(\alpha)$

$B_n(x)$ : Πολυώνυμο Bernoulli βαθμού n.

$B_n$ : Αριθμός Bernoulli, $B_n = B_n(0)$.

[x] : Ο μεγαλύτερος ακέραιος που είναι μικρότερος από το x.

$\infty$ : Το άπειρο.

P{E} : Η πιθανότητα ενός ενδεχομένου Ε.

P{A|B} ή P(A|B) : Η δεσμευμένη πιθανότητα ενός ενδεχομένου Α όταν γνωρίζουμε ότι έχει πραγματοποιηθεί το ενδεχόμενο Β.

E [X] ή E (X) : Η μέση τιμή της τυχαίας μεταβλητής X.

E [X|Y] : Η δεσμευμένη μέση τιμή της τυχαίας μεταβλητής X δοθείσης της τυχαίας μεταβλητής Y.

Var (X) ή V [X] : Η διασπορά της τυχαίας μεταβλητής X.

$F_X(x)$ ή $F(x)$ : Η συνάρτηση κατανομής πιθανότητας της τυχαίας μεταβλητής X, $F(x) = P\{X \le x\}\ \forall x \in \Re$.

$\{p_\kappa : \kappa \in \aleph\}$ : Η συνάρτηση μάζας πιθανότητας της τυχαίας μεταβλητής X, $p_\kappa = P\{X = x_\kappa\}\ \forall \kappa \in \aleph$.

$f_X(x)$ ή $f(x)$ : Η συνάρτηση πυκνότητας πιθανότητας της τυχαίας μεταβλητής X, $f(x) = F'(x)\ \forall\, x \in \Re$ εφόσον η παράγωγος υπάρχει.

$\xrightarrow{P}$ : Συγκλίνει κατά πιθανότητα.

# Βιβλιογραφία


**α) Ελληνική**

[1] ΚΟΚΟΛΑΚΗΣ, Γ.Ε. και ΣΠΗΛΙΩΤΗΣ, Ι. *Εισαγωγή στις Πιθανότητες*. Έκδοση 1η, Εκδόσεις Συμεών, Αθήνα, 2002.

[2] ΚΟΚΟΛΑΚΗΣ, Γ.Ε. *Σημειώσεις Στοχαστικών Ανελίξεων*. Αθήνα, 2005.

[3] ΡΑΣΣΙΑΣ, Θ.Μ. *Μαθηματική Ανάλυση Ι*. Τεύχος Α΄, Εκδόσεις Σαββάλας, Αθήνα, 2004.

[4] ΡΑΣΣΙΑΣ, Θ.Μ. *Μαθηματική Ανάλυση Ι*. Τεύχος Β΄, Εκδόσεις Σαββάλας, Αθήνα, 2005.

**β) Ξενόγλωσση**

[5] BENDER, C.M. and ORSZAG, S.A. *Advanced Mathematical Methods for Scientists and Engineers*. Springer, New York, 1999.

[6] DOUMAS, A.V. and PAPANICOLAOU, V.G. *The coupon collector's problem revisited: asymptotics of the variance*. Adv. Appl. Prob. 44, 166 – 195, 2012.

[7] DURRETT, R. *Probability: Theory and Examples*. 3rd ed., Cambridge University Press, 2005.

[8] FELLER, W. *An Introduction to Probability Theory and Its Applications*. Vol. 1, 3rd ed., John Wiley, New York, 1968.

[9] FELLER, W. *An Introduction to Probability Theory and Its Applications*. Vol. 2, 2nd ed., John Wiley, New York, 1971.

[10] ROSS, S.M. *Introduction to Probability Models*. 9th ed., Academic Press, 2006.