

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

**ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ
ΚΑΙ ΑΣΦΑΛΙΣΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ**

**ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ ΣΤΗΝ
ΕΦΑΡΜΟΣΜΕΝΗ ΣΤΑΤΙΣΤΙΚΗ**

**ΣΤΟΧΑΣΤΙΚΑ ΥΠΟΔΕΙΓΜΑΤΑ
ΡΟΩΝ ΣΕ ΘΕΜΑΤΑ
ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΗΣ
ΑΝΑΛΥΣΗΣ**

Κωνσταντίνα Δ. Γεωργαντώνη

Διπλωματική Εργασία

που υποβλήθηκε στο Τμήμα Στατιστικής και Ασφαλιστικής
Επιστήμης του Πανεπιστημίου Πειραιώς ως μέρος των απαιτήσεων
για την απόκτηση του Μεταπτυχιακού Διπλώματος Ειδίκευσης στην
Εφαρμοσμένη Στατιστική

Πειραιάς

2007

Η παρούσα Διπλωματική Εργασία εγκρίθηκε ομόφωνα από την Τριμελή Εξεταστική Επιτροπή που ορίσθηκε από τη ΓΣΕΣ του Τμήματος Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς στην υπ' αριθμ. συνεδρίασή του σύμφωνα με τον Εσωτερικό Κανονισμό Λειτουργίας του Προγράμματος Μεταπτυχιακών Σπουδών στην Εφαρμοσμένη Στατιστική Τα μέλη της Επιτροπής ήταν:

- (Επιβλέπων)
-
-

Η έγκριση της Διπλωματικής Εργασίας από το Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς δεν υποδηλώνει αποδοχή των γνώμων του συγγραφέα.

UNIVERSITY OF PIRAEUS

**DEPARTMENT OF STATISTICS
AND INSURANCE SCIENCE**

**POSTGRADUATE PROGRAM IN
APPLIED STATISTICS**

**STOCHASTIC RUNS MODELS
APPLIED TO
FINANCIAL ANALYSIS TOPICS**

By

Konstantina D. Georgantoni

MSc Dissertation

**submitted to the Department of Statistics and Insurance Science of the
University of Piraeus in partial fulfillment of the requirements for the
degree of Master of Science in Applied Statistics**

Piraeus, Greece

2007

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

Στην οικογένεια μου
και στο Γιώργο

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

Ευχαριστίες

Θέλω να ευχαριστήσω τον επιβλέποντα καθηγητή μου κύριο Μάρκο Κούτρα για την πολύτιμη καθοδήγησή του και για τις γνώσεις που μου πρόσφερε, καθώς και τον κύριο Φώτη Μιλιένο για την χρήσιμη βοήθεια του με τον προγραμματισμό. Επίσης ευχαριστώ τους γονείς μου για την ηθική και οικονομική τους βοήθεια όλα τα χρόνια των σπουδών μου, το Γιώργο Κατσούδα και όλους τους φίλους μου για την στήριξη τους όλο αυτό τον καιρό.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

Περίληψη

Στην εργασία αυτή περιγράφεται αρχικά η έννοια των ροών καθώς και πού αυτές χρησιμοποιούνται. Στη συνέχεια αναλύεται η τυχαία μεταβλητή T_k , η οποία εκφράζει το χρόνο αναμονής μέχρι την πρώτη εμφάνιση ροής επιτυχιών τάξης k . Η μελέτη της T_k παρουσιάζεται για την περίπτωση ανεξάρτητων και ισόνομων (i.i.d.) δοκιμών και για την περίπτωση εξαρτημένων δοκιμών που εμφανίζουν Μαρκοβιανή εξάρτηση πρώτης τάξης. Δίνονται, και για τις δύο περιπτώσεις, διάφοροι εναλλακτικοί τρόποι εύρεσης της κατανομής της T_k , δηλαδή τύποι για τη συνάρτηση πυκνότητας πιθανότητας και την αθροιστική συνάρτηση κατανομής της, καθώς και τύποι για τη μέση τιμή και τη διακύμανσή της.

Στην συνέχεια περιγράφεται ένα μοντέλο δεσμευμένων πιθανοτήτων, που σχετίζεται με χρόνους ανομονής για την εμφάνιση μιας ροής μήκους k , το οποίο χρησιμοποιείται για την πρόβλεψη της κατάλληλης χρονικής στιγμής για επένδυση σε μια μετοχή. Το μοντέλο αυτό περιγράφεται και για τις δύο περιπτώσεις δοκιμών που αναφέραμε παραπάνω. Το πρόβλημα που μας απασχολεί είναι να βρούμε την χρονική στιγμή (ημέρα), έστω x , στην οποία θα έχουμε την μεγαλύτερη πιθανότητα στις επόμενες από αυτή y δοκιμές (ημέρες), να παρουσιαστεί ροή ανόδων (επιτυχιών) τάξης k , δεδομένου ότι μέχρι αυτή τη χρονική στιγμή x δεν έχει παρουσιαστεί μια τέτοια ροή. Δίνονται διαγράμματα της δεσμευμένης πιθανότητας $f_k(x/y)$ για διάφορες τιμές του χρονικού ορίζοντα y , του μήκους της ροής ανόδων k και της πιθανότητας ανόδου p .

Τέλος καταφεύγουμε σε πραγματικά δεδομένα του χρηματιστηρίου, δηλαδή χρονοσειρές της εξέλιξης κάποιων μετοχών, για να συγκρίνουμε τα αποτελέσματα του πιθανοθεωρητικού μοντέλου με τα αντίστοιχα εμπειρικά. Περιγράφεται και σχολιάζεται η φύση των χρονοσειρών αυτών, επιλέγονται κατάλληλες χρονικές περίοδοι που ταιριάζουν στην περίπτωση των i.i.d. δοκιμών και γίνεται η εφαρμογή του μοντέλου. Τέλος, γίνεται η σύγκριση μεταξύ θεωρητικών και πειραματικών αποτελεσμάτων για την επιτυχία του προτεινόμενου μοντέλου.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

Abstract

In the present dissertation the notion of success runs is presented firstly along with the fields of their application. Next, the random variable T_k , which denotes the waiting time until the occurrence, for the first time, of a sequence of k consecutive successes is analyzed. This random variable, is studied for the case of independent identically distributed (i.i.d.) trials as well as for first order Markov-dependent trials. For both cases, we give provide various alternative techniques for establishing the associated distribution, that is to say formulae for the probability mass function, cumulative distribution function and their moments (mean and variance).

Then we describe a model of conditional probabilities, which is related to the waiting times of runs of length k , which is subsequently used for identifying the most appropriate time instance for investment action in a stock market share. The model is developed for the two cases of trials (i.i.d. and Markov dependent) mentioned above. The problem we are dealing with is to spot out the time (day) x , which maximizes the probability that in the next y trials (days), a run of rises (successes) of order k will occur, given that, up to time x no such run has been observed. Several graphs of the conditional probability $f_k(x/y)$ are given for a variety of choices for the time horizon y , length of run k and success (rise) probability p .

Finally, we use real data from the Athens Stock Exchange market, i.e. time series of specific stocks, to compare the theoretical results to the empirical ones. The form of these time series is analysed by the aid of quality control tools (control charts), and appropriate time periods are selected so that the assumption of i.i.d. trials be justified. Then a comparison between the theoretical and experimental results is performed in order to assess the efficiency of our conditional probability model.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

Περιεχόμενα

Κατάλογος Πινάκων

Κατάλογος Σχημάτων

Κατάλογος Συντομογραφιών

1. Εισαγωγή: Έννοια και χρήση των ροών	1
2. Τυχαία μεταβλητή T_k: χρόνος αναμονής μέχρι την πρώτη εμφάνιση ροής τάξης k	4
2.1 Εισαγωγή της τυχαίας μεταβλητής T_k	4
2.2 Εύρεση της κατανομής της τυχαίας μεταβλητής T_k	5
2.3 Μέση τιμή και διακύμανση της τυχαίας μεταβλητής T_k	15
2.4 Χρόνος αναμονής μέχρι την πρώτη εμφάνιση ροής επιτυχιών τάξης k για Μαρκοβιανά εξαρτημένες δοκιμές	17
3. Δεσμευμένες κατανομές που σχετίζονται με χρόνους αναμονής ροής επιτυχιών μήκους k	20
3.1 Εισαγωγή	20
3.2 Ένα μοντέλο πιθανοτήτων για ανεξάρτητες και ισόνομες δοκιμές	21
3.3 Ένα μοντέλο για Μαρκοβιανά εξαρτημένες δοκιμές	33
4. Εφαρμογή του μοντέλου σε πραγματικά δεδομένα	44
4.1 Εισαγωγή	44
4.2 Φύση των δεδομένων	45
4.3 Επιλογή δεδομένων και σύγκριση θεωρητικών και πειραματικών πιθανοτήτων	51
Βιβλιογραφία	61

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

Κατάλογος πινάκων

4.3.1	Επιλεγμένες από τις χρονοσειρά του ΟΤΕ περιοχές με πιθανότητες ανόδου $p=0,40$, $p=0,60$ και $p=0,80$	51
4.3.2	Ακριβείς τιμές της $f_k(x/y)$ για $y=2$ και $k=3$	53
4.3.3	Ακριβείς τιμές της $f_k(x/y)$ για $y=2$ και $k=5$	54
4.3.4	Ακριβείς τιμές της $f_k(x/y)$ για $y=3$ και $k=5$	55
4.3.5	Πειραματικές τιμές της $f_k(x/y)$ για την μετοχή του ΟΤΕ σε περίοδο με $p=0,80$	56
4.3.6	Θεωρητικές και πειραματικές τιμές της $f_k(x/y)$ για $y=2$, $k=3$ και $p=0,80$	57
4.3.7	Θεωρητικές και πειραματικές τιμές της $f_k(x/y)$ για $y=2$, $k=5$ και $p=0,80$	58
4.3.8	Θεωρητικές και πειραματικές τιμές της $f_k(x/y)$ για $y=3$, $k=5$ και $p=0,80$	59
4.3.9	Αποκλίσεις θεωρητικών και πειραματικών τιμών της $f_k(x/y)$ με $p=0,80$	60

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

Κατάλογος Σχημάτων

2.2.1	Γεωμετρική κατανομή τάξης $k=3$	9
2.2.2	Γεωμετρική κατανομή τάξης $k=6$	10
2.2.3	Γεωμετρική κατανομή τάξης $k=8$	11
2.2.4	Γεωμετρική κατανομή τάξης k και πιθανότητα επιτυχίας $p=0.3$	12
2.2.5	Γεωμετρική κατανομή τάξης k και πιθανότητα επιτυχίας $p=0.5$	13
2.2.6	Γεωμετρική κατανομή τάξης k και πιθανότητα επιτυχίας $p=0.8$	14
3.2.1	Δεσμευμένη πιθανότητα $f_k(x/y)$ για $k>1$	23
3.2.2	Δεσμευμένη πιθανότητα $f_k(x/y)$ για $y=2$	24
3.2.3	Δεσμευμένη πιθανότητα $f_k(x/y)$ για $y=3$	27
3.2.4	Δεσμευμένη πιθανότητα $f_k(x/y)$ για $y=5$	29
3.2.5	Δεσμευμένη πιθανότητα $f_k(x/y)$ για $y=10$	30
3.3.1	Δεσμευμένη πιθανότητα $f_k(x/y)$ για $y=2$ σε Μαρκοβιανά εξαρτημένες δοκιμές	37
3.3.2	Δεσμευμένη πιθανότητα $f_k(x/y)$ για $y=3$ σε Μαρκοβιανά εξαρτημένες δοκιμές	39
3.3.3	Δεσμευμένη πιθανότητα $f_k(x/y)$ για $y=5$ σε Μαρκοβιανά εξαρτημένες δοκιμές	40
3.3.4	Δεσμευμένη πιθανότητα $f_k(x/y)$ για $y=10$ σε Μαρκοβιανά εξαρτημένες δοκιμές	42
4.2.1	Διάγραμμα χρονοσειράς του γενικού δείκτη του χρηματιστηρίου	45
4.2.2	p διάγραμμα ελέγχου για τον γενικό δείκτη με μέγεθος δείγματος $n=100$	46
4.2.3	Διάγραμμα χρονοσειράς της μετοχής του ΟΤΕ	47
4.2.4	p διάγραμμα ελέγχου για την μετοχή του ΟΤΕ με μέγεθος δείγματος $n=100$	48
4.2.5	Διάγραμμα χρονοσειράς της μετοχής της Εθνικής	49
4.2.6	p διάγραμμα ελέγχου για την μετοχή της Εθνικής με μέγεθος δείγματος $n=100$	50

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

4.3.1	Δεσμευμένη πιθανότητα $f_k(x/y)$ για $y=2$ και $k=3$	53
4.3.2	Δεσμευμένη πιθανότητα $f_k(x/y)$ για $y=2$ και $k=5$	54
4.3.3	Δεσμευμένη πιθανότητα $f_k(x/y)$ για $y=3$ και $k=5$	55

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

Κατάλογος Συντομογραφιών

τ.μ.	τυχαία μεταβλητή
σ.π.π.	συνάρτηση πυκνότητας πιθανότητας
σ.κ.	συνάρτηση κατανομής
i.i.d.	independent and identically distributed (ανεξάρτητες και ομοιόμορφα κατανεμημένες)

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

ΚΕΦΑΛΑΙΟ 1

Εισαγωγή

Έννοια και χρήση των ροών

Το αντικείμενο στο οποίο στηρίζεται η εργασία αυτή είναι οι ροές. Σκοπός της εργασίας είναι, χρησιμοποιώντας την θεωρία των ροών, να εφαρμόσουμε κάποιο μοντέλο δεσμευμένων πιθανοτήτων, το οποίο σχετίζεται με χρόνους ανομονής ροής μήκους k , και να δούμε κατά πόσο μπορούμε με αυτήν την εφαρμογή, να προβλέψουμε με επιτυχία την κατάλληλη χρονική στιγμή για επένδυση σε μια μετοχή. Η ροή σαν έννοια είναι μια συνεχόμενη ακολουθία στοιχείων, δηλαδή είναι μια ακολουθία όμοιων στοιχείων που δεν διακόπτεται. Για να γίνει η έννοια της ροής κατανοητή μπορούμε να δώσουμε απλά, καθημερινά παραδείγματα. Έστω ότι κάνουμε έλεγχο ποιότητας σε ένα προϊόν. Επιλέγουμε ένα δείγμα από μονάδες του προϊόντος αυτού και ελέγχουμε αν η κάθε μονάδα είναι ελαττωματική ή όχι. Κωδικοποιώντας τις ελαττωματικές με τον αριθμό 1 και τις μη ελαττωματικές με τον αριθμό 0, έστω ότι έχουμε τα εξής αποτελέσματα: 0000110011111. Έτσι έχουμε μια ροή ελαττωματικών μονάδων τάξης 2 και άλλη μία τάξης 5. Αν ο παραγωγός ή ο καταναλωτής έχει ως όριο αποδοχής του δείγματος τις 4 συνεχόμενες ελαττωματικές μονάδες, τότε το δείγμα αυτό θα απορριφθεί γιατί έχουμε ροή ελαττωματικών μονάδων τάξης 5. Έστω τώρα ότι έχουμε ένα σύστημα το οποίο αποτελείται από 5 πομπούς μετάδοσης κάποιου σήματος. Το σήμα μεταδίδεται από τον έναν πομπό στον άλλο και κάθε πομπός έχει εμβέλεια μετάδοσης μέχρι και 3 πομπούς μετά από αυτόν. Αν σε αυτό το σύστημα έχουμε 3 ή περισσότερους συνεχόμενους πομπούς χαλασμένους, δηλαδή έχουμε μια ροή χαλασμένων πομπών τάξης 3 ή μεγαλύτερης, το σύστημα δεν λειτουργεί και έτσι το σήμα δεν μεταδίδεται.

Οι ροές έχουν χρησιμοποιηθεί σε πολλά στατιστικά προβλήματα και προβλήματα πιθανοτήτων. Πρώτος ο de Moivre (1738) χρησιμοποίησε τις ροές στο πρόβλημα: ποιά είναι η πιθανότητα να πάρουμε μια ροή μήκους r σε n δοκιμές. Με το ίδιο πρόβλημα ασχολήθηκαν αργότερα και ο Simson (1740), ο Laplace (1812), και ο Tobhunter (1865). Ο de Forest (1876)

χρησιμοποίησε τις ροές στις ακολουθίες σφαλμάτων (residuals) για να προσδιορίσει αν κάποια παρατηρούμενα δεδομένα ακολουθούν μια συγκεκριμένη κατανομή.

Αργότερα οι Wald and Wolfowitz (1940) χρησιμοποίησαν τις ροές για να πραγματοποιήσουν τον έλεγχο αν δύο τυχαία δείγματα από δύο πληθυσμούς προέρχονται από την ίδια κατανομή (two sample test). Δηλαδή έχουμε ένα τυχαίο δείγμα X_1, X_2, \dots, X_m από έναν πληθυσμό με αθροιστική συνάρτηση κατανομής $F_X(x)$ και ένα άλλο ανεξάρτητο τυχαίο δείγμα Y_1, Y_2, \dots, Y_n με αθροιστική συνάρτηση κατανομής $F_Y(x)$. Θέλουμε να ελέγξουμε την μηδενική υπόθεση $H_0: F_X(x) = F_Y(x)$ έναντι της εναλλακτικής $H_1: F_X(x) \neq F_Y(x)$. Συνδιάζουμε τα δύο δείγματα και τακτοποιούμε τις $m + n$ παρατηρήσεις κατά αύξουσα σειρά μεγέθους, αντικαθιστούμε τις τιμές με 0 ή 1 ανάλογα από ποιο δείγμα προέρχονται, δείγμα X ή Y αντίστοιχα και χρησιμοποιούμε τον συνολικό αριθμό των ροών σε αυτή την ακολουθία. Οι Wald and Wolfowitz (1940) πρότειναν την απόρριψη της μηδενικής υπόθεσης αν έχουμε μικρό αριθμό συνολικών ροών. Οι κριτικές τιμές του ελέγχου μπορούν να προσδιορισθούν για τα δοθέντα μεγέθη m και n των δειγμάτων και το επιθυμητό επίπεδο σημαντικότητας. Οι Wald and Wolfowitz (1940) δείξαν ότι η ισχύς αυτού του τεστ ροών δύο δειγμάτων τείνει στο 1 όταν τα μεγέθη των δειγμάτων m και n τείνουν και τα δύο στο ∞ .

Ακόμα οι ροές μπορούν να χρησιμοποιηθούν και για τον έλεγχο τυχειότητας ενός δείγματος με διαφορετική κρίσιμη περιοχή για κάθε εναλλακτική υπόθεση. Πιο συγκεκριμένα, έστω X_1, X_2, \dots, X_n n τυχαίες μεταβλητές με από κοινού αθροιστική συνάρτηση κατανομής $F(x_1, x_2, \dots, x_n)$. Θέλουμε να ελέγξουμε την μηδενική υπόθεση H_0 : οι X_i είναι αναξάρτητες και ομοιόμορφα κατανεμημένες (i.i.d). Θεωρούμε τις διαφορές $X_2 - X_1, X_3 - X_2, \dots, X_n - X_{n-1}$ και ορίζουμε ως *run up* μία ροή θετικών διαφορών και ως *run down* μία ροή αρνητικών διαφορών. Ανάλογα με την εναλλακτική υπόθεση ορίζουμε την κατάλληλη κριτική περιοχή για τον αριθμό των *runs up* και των *runs down* ή για το μήκος των *runs up* και των *runs down*. Αν για παράδειγμα η εναλλακτική υπόθεση είναι ότι υπάρχει κάποια τάση στις X_1, X_2, \dots, X_n τότε απορρίπτουμε την H_0 αν οι αριθμοί των *runs up* και των *runs down* είναι πολύ μικροί. Αν η εναλλακτική υπόθεση είναι ότι παρασιάζονται πολλοί μικροί κύκλοι στις X_1, X_2, \dots, X_n , τότε απορρίπτουμε την H_0 αν οι αριθμοί των *runs up* και των *runs down* είναι πολύ μεγάλοι.

Επίσης οι ροές χρησιμοποιούνται στον στατιστικό έλεγχο ποιότητας για να ελέγξουμε αν μια διαδικασία παραγωγής είναι εντός ελέγχου. Ο αριθμός και το μήκος των ροών έχουν χρησιμοποιηθεί στην ανάπτυξη μη παραμετρικών τεστ για την συμμετρία μιας κατανομής δεδομένων, Cohen and Menjoge (1988), Mc Williams (1990), Henze (1993), Modarres and Gastwirth (1996, 1998). Οι ροές παίζουν βασικό ρόλο και στα τεστ εκκίνησης μιας μηχανής ή μιας συσκευής, όπου ανάλογα με τον αριθμό συνεχόμενων επιτυχημένων εκκινήσεων (δηλαδή το μήκος της ροής επιτυχιών) αποδεχόμαστε την συσκευή ή όχι. Τέλος μια άλλη εφαρμογή των ροών είναι στα συστήματα αξιοπιστίας όπου μέσω της μεγαλύτερης ροής αποτυχιών κρίνεται αν το σύστημα αποτυγχάνει, Chiang και Niu (1981), Chao, Fu and Koutras (1995).

ΚΕΦΑΛΑΙΟ 2

Τυχαία μεταβλητή T_k : χρόνος αναμονής μέχρι την πρώτη εμφάνιση ροής τάξης k

2.1 Εισαγωγή της τυχαίας μεταβλητής T_k

Έστω ότι X_1, X_2, \dots, X_n μια ακολουθία δοκιμών, κάθε μία από τις οποίες μπορεί να έχει δύο αποτελέσματα: επιτυχία (Success (S)) ή αποτυχία (Failure (F)). Συμβολίζουμε με T_k την τυχαία μεταβλητή που δηλώνει το χρόνο αναμονής μέχρι να παρατηρηθεί για πρώτη φορά μια ακολουθία k συνεχόμενων επιτυχιών. Δηλαδή ενδιαφερόμαστε μόνο για τις επιτυχίες και έτσι η τυχαία μεταβλητή T_k εκφράζει το χρόνο αναμονής μέχρι την πρώτη εμφάνιση ροής επιτυχιών τάξης k , δηλαδή μήκους (μεγέθους) k . Έτσι μπορούμε να εκφράσουμε την τυχαία μεταβλητή T_k με τους εξής ισοδύναμους τύπους:

$$T_k = \min\{n : X_{n-k+1} = \dots = X_n = 1\} = \min\left\{n : \prod_{j=n-k+1}^n X_j = 1\right\} = \min\left\{n : \sum_{j=n-k+1}^n X_j = k\right\}.$$

Ας δώσουμε τώρα ένα παραδείγμα για την τυχαία μεταβλητή T_k . Έστω ότι έχουμε τα παρακάτω 30 αποτελέσματα μιας ακολουθίας :

S S S S F F F S S F F F S S S S S F F S S S S S S F S S

Τότε ο χρόνος αναμονής μέχρι την πρώτη εμφάνιση ροής επιτυχιών τάξης $k=2$ είναι $T_2 = 2$, ενώ ο χρόνος αναμονής μέχρι την πρώτη εμφάνιση ροής επιτυχιών τάξης $k=3$ είναι $T_3 = 3$, τάξης $k=4$ είναι $T_4 = 4$, τάξης $k=5$ είναι $T_5 = 18$, τάξης $k=6$ είναι $T_6 = 26$, τάξης $k=7$ είναι $T_7 = 27$, τάξης $k=8$ είναι $T_8 > 30$.

2.2 Εύρεση της κατανομής της τυχαίας μεταβλητής T_k

Ας θεωρήσουμε ότι έχουμε μια ακολουθία από δοκιμές Bernoulli, δηλαδή ότι έχουμε μια ακολουθία X_1, X_2, \dots από ανεξάρτητες και ισόνομες δοκιμές με πιθανότητα επιτυχίας $p = P(X_i = 1)$ και πιθανότητα αποτυχίας $q = P(X_i = 0)$. Τότε η κατανομή της τυχαίας μεταβλητής T_k , που εκφράζει το χρόνο αναμονής μέχρι την πρώτη εμφάνιση ροής επιτυχιών τάξης k , είναι γνωστή με την ονομασία γεωμετρική κατανομή τάξης k .

Αν συμβολίσουμε με $f(x)$ την συνάρτηση πυκνότητας πιθανότητας της T_k , δηλαδή της γεωμετρικής κατανομής τάξης k , $f(x) = P(T_k = x)$, $x=0,1,\dots$ τότε η $f(x)$ για $0 \leq x \leq 2k$ θα δίνεται από τον εξής τύπο:

$$f(x) = \begin{cases} 0 & , \text{αν } 0 \leq x < k \\ p^k & , \text{αν } x = k \\ qp^k & , \text{αν } k < x \leq 2k \end{cases}$$

Είναι πολύ εύκολο να κατανοήσουμε τον παραπάνω τύπο αν σκεφτούμε τον ορισμό της τυχαίας μεταβλητής T_k . Αφού αναμένουμε να εμφανιστούν k συνεχόμενες επιτυχίες, όταν ο χρόνος x βρίσκεται στο διάστημα $0 \leq x < k$, δηλαδή είναι μικρότερος από k , η πιθανότητα η τ.μ. T_k να βρίσκεται μέσα σε αυτό το διάστημα είναι φυσικά μηδενική. Για $x = k$, έχουμε $T_k = k$, το οποίο σημαίνει ότι οι k πρώτες δοκιμές μας είναι επιτυχίες άρα έχουμε την πιθανότητα $P(T_k = k) = p^k$, όπου p είπαμε ότι είναι η πιθανότητα επιτυχίας της κάθε δοκιμής. Για να βρούμε την $f(x)$ για $k < x \leq 2k$ σκεφτόμαστε ως εξής: σε αυτό το διάστημα η πιθανότητα $P(T_k = x)$ μπορεί να βρεθεί ως το γινόμενο $qp^k P(\text{στις προηγούμενες από το } F \text{ δοκιμές, } x-k-1 \text{ το πλήθος, δεν εμφανίζεται ροή επιτυχιών τάξης } k) = qp^k(1 - P(\text{στις προηγούμενες από το } F \text{ δοκιμές, } x-k-1 \text{ το πλήθος, εμφανίζεται ροή επιτυχιών τάξης } k)) = qp^k(1-0) = qp^k$.

Για $x \geq 2k$ υπάρχουν τρεις αναδρομικοί τύποι που μπορούν να χρησιμοποιηθούν για τον υπολογισμό της συνάρτησης πυκνότητας πιθανότητας της T_k . Ο πρώτος υπολογίζει την σ.π.π. της T_k αθροίζοντας τις διαφορετικές περιπτώσεις που μπορούμε να έχουμε ανάλογα σε ποια δοκιμή εμφανίζεται η πρώτη αποτυχία και είναι ο εξής:

$$f(x) = \sum_{i=1}^k qp^{i-1} f(x-i) \tag{2.2.1}$$

Αν έχουμε την περίπτωση όπου το πρώτο F εμφανίζεται στην πρώτη δοκιμή έχουμε ότι $f(x) = qP(T_k = x - 1)$, αν εμφανίζεται στην δεύτερη δοκιμή έχουμε $f(x) = pqP(T_k = x - 2)$, αν εμφανίζεται στην τρίτη δοκιμή $f(x) = p^2qP(T_k = x - 3)$, ..., και αν εμφανίζεται στην k δοκιμή $f(x) = p^{k-1}qP(T_k = x - k)$. Για να βρούμε την σ.π.π. της T_k αθροίζουμε όλες τις παραπάνω περιπτώσεις.

Ο δεύτερος τύπος βασίζεται στην τομή δύο ανεξάρτητων γεγονότων. Τα γεγονότα αυτά είναι: οι τελευταίες $k+1$ δοκιμές είχαν αποτελέσματα F, S, \dots, S και δεν εμφανίστηκε ροή επιτυχιών μήκους k ή μεγαλύτερη, μέχρι την $(x-k-1)$ -ή δοκιμή. Δηλαδή έχουμε ότι η $f(x)$ υπολογίζεται ως εξής:

$$f(x) = pqP(\text{στις προηγούμενες από το } F \text{ δοκιμές δεν εμφανίζονται } k \text{ συνεχόμενες επιτυχίες}) = p^k q (1 - P(\text{στις προηγούμενες από το } F \text{ δοκιμές, } x-k-1 \text{ το πλήθος, εμφανίζονται } k \text{ συνεχόμενες επιτυχίες}))$$

Έτσι αυτός ο δεύτερος αναδρομικός τύπος είναι:

$$f(x) = qp^k \left\{ 1 - \sum_{i=0}^{x-k-1} f(i) \right\} \quad (2.2.2)$$

Ο τρίτος τύπος, ο οποίος χρησιμοποιείται περισσότερο, υπολογίζεται εφαρμόζοντας την αφαίρεση $f(x-1) - f(x)$ στον τύπο (2.2.1) και είναι ο εξής:

$$f(x) = f(x-1) - qp^k f(x-k-1). \quad (2.2.3)$$

Συγγραφείς που έχουν αναφέρει και έχουν αποδείξει τους τρεις αναδρομικούς τύπους (2.2.1-2.2.3) είναι: Hahn and Gage (1983), Laurencelle (1987), Aki, Kuboki and Hirano (1984), Philippou and Makri (1985), Barry and Lo Bello (1993), Johnson, Kotz and Kemp (1992), Kemp (1997).

Για την συνάρτηση πυκνότητας πιθανότητας της T_k υπάρχουν και μη αναδρομικοί τύποι. Έτσι έχουμε τους παρακάτω:

$$f(x) = \sum \binom{x_1 + x_2 + \dots + x_k + K}{x_1, x_2, \dots, x_k} p^x \left(\frac{q}{p} \right)^{x_1 + K + x_k}, \quad x \geq k$$

με την συνθήκη $\sum_{i=1}^k ix_i = x - k$, όπου x_1, x_2, \dots, x_k είναι μη-αρνητικοί ακέραιοι, Philippou and Muwafi (1982). Ο τύπος αυτός περιλαμβάνει πολυωνυμικά αποτελέσματα και πολλαπλά

αθροίσματα. Ένας άλλος μη-αναδρομικός τύπος για την συνάρτηση πυκνότητας πιθανότητας της T_K ο οποίος περιλαμβάνει δυνωμικά αποτελέσματα και μονά αθροίσματα είναι ο εξής:

$$f(x) = p^k \sum_{j=0}^{\infty} (-1)^j \binom{x-k-jk}{j} (qp^k)^j - p^{k+1} \sum_{j=0}^{\infty} (-1)^j \binom{x-k-jk-1}{j} (qp^k)^j, \quad x \geq k$$

, Uppuluri and Patil (1983).

Ένας δεύτερος μη-αναδρομικός τύπος ο οποίος περιλαμβάνει δυνωμικά αποτελέσματα είναι:

$$f(x) = \sum_{i=1}^{x-k} q^i p^{x-i} \sum_{j=0}^{\lfloor \frac{x-i-k}{k} \rfloor} (-1)^j \binom{i}{j} \binom{x-k(j+1)-1}{i-1}, \quad x \geq k+1$$

Ακόμα ένας μη-αναδρομικός τύπος ο οποίος περιλαμβάνει μονό άθροισμα είναι:

$$f(x) = \sum_{j=1}^{\lfloor \frac{x+1}{k+1} \rfloor} (-1)^{j-1} p^{jk} q^{j-1} \left\{ \binom{x-jk-1}{j-2} + \binom{x-jk-1}{j-1} \right\},$$

Muselli (1996).

Ακόμα μπορούμε να καταλήξουμε στη συνάρτηση πυκνότητας πιθανότητας της T_K χρησιμοποιώντας μια κατάλληλη αλυσίδα Markov. Παρακολουθούμε την διαδικασία, η οποία ανεβαίνει στάδια μόνο όταν έχουμε συνεχόμενες επιτυχίες, μέχρι να φτάσουμε στο στάδιο k , δηλαδή να έχουμε k συνεχόμενες επιτυχίες έστω μία φορά. Φτιάχνουμε τον πίνακα μετάβασης ο οποίος είναι της μορφής:

$$\Lambda = \begin{bmatrix} q & p & 0 & K & 0 & 0 \\ q & 0 & p & K & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ q & 0 & 0 & K & 0 & p \\ 0 & 0 & 0 & K & 0 & 1 \end{bmatrix}_{(k+1) \times (k+1)}$$

και η συνάρτηση πυκνότητας πιθανότητας της T_K δίνεται από τον τύπο:

$$f(x) = p \mathbf{e}_1 \Lambda^{x-1} \mathbf{e}'_k, \quad x \geq 1,$$

όπου \mathbf{e}_i , $i=1,2,\dots, k+1$ είναι μοναδιαία διανύσματα στο χώρο \mathfrak{R}^{k+1} . Μια τέτοια προσέγγιση χρησιμοποιήθηκε από τους Chao and Fu (1989, 1991), Fu and Lou (1991) για την μελέτη συστημάτων αξιοπιστίας. Ακόμα χρησιμοποιήθηκε από τον Feder (1974) για την διαμόρφωση ορισμένων μεγάλων συναθροίσεων με σειριακούς συνδέσμους. Μια

ενοποιημένη παρουσίαση αυτών των μεθόδων μπορεί να βρεθεί στην πρόσφατη δημοσίευση του Koutras (1996a).

Αν συμβολίσουμε τώρα με $F(x)$ την αθροιστική συνάρτηση κατανομής της T_k , δηλαδή της γεωμετρικής κατανομής τάξης k , $F(x) = P(T_k \leq x) = \sum_{y \leq x} f(y)$, $x=0,1,\dots$ τότε η $F(x)$ θα

δίνεται από τον τύπο:

$$F(x) = F(x-1) + qp^k \{1 - F(x-k-1)\}, \quad x \geq k+1, \quad (2.2.4)$$

με αρχικές συνθήκες

$$F(x) = \begin{cases} 0 & \text{αν } 0 \leq x < k \\ p^k & \text{αν } k \leq x < k+1. \end{cases} \quad (2.2.5)$$

Ένας άλλος τύπος για την αθροιστική συνάρτηση κατανομής της T_k είναι ο παρακάτω:

$$F(x) = 1 - \frac{p^{x+1}}{q} \sum \binom{x_1 + x_2 + \dots + x_k}{x_1, x_2, \dots, x_k} (q/p)^{\sum_{i=1}^k x_i}, \quad x \geq k,$$

με την συνθήκη $\sum_{i=1}^k ix_i = x+1$ όπου x_1, x_2, \dots, x_k είναι μη-αρνητικοί ακέραιοι, Philippou and Makri (1985, 1986), Georgiou and Philippou (1985).

Ένας τρίτος τύπος για την αθροιστική συνάρτηση κατανομής της T_k είναι μέσω της αλυσίδας Markov και ο τύπος είναι ο εξής:

$$F(x) = \mathbf{e}_1 \mathbf{\Lambda}^x \mathbf{e}'_{k+1}.$$

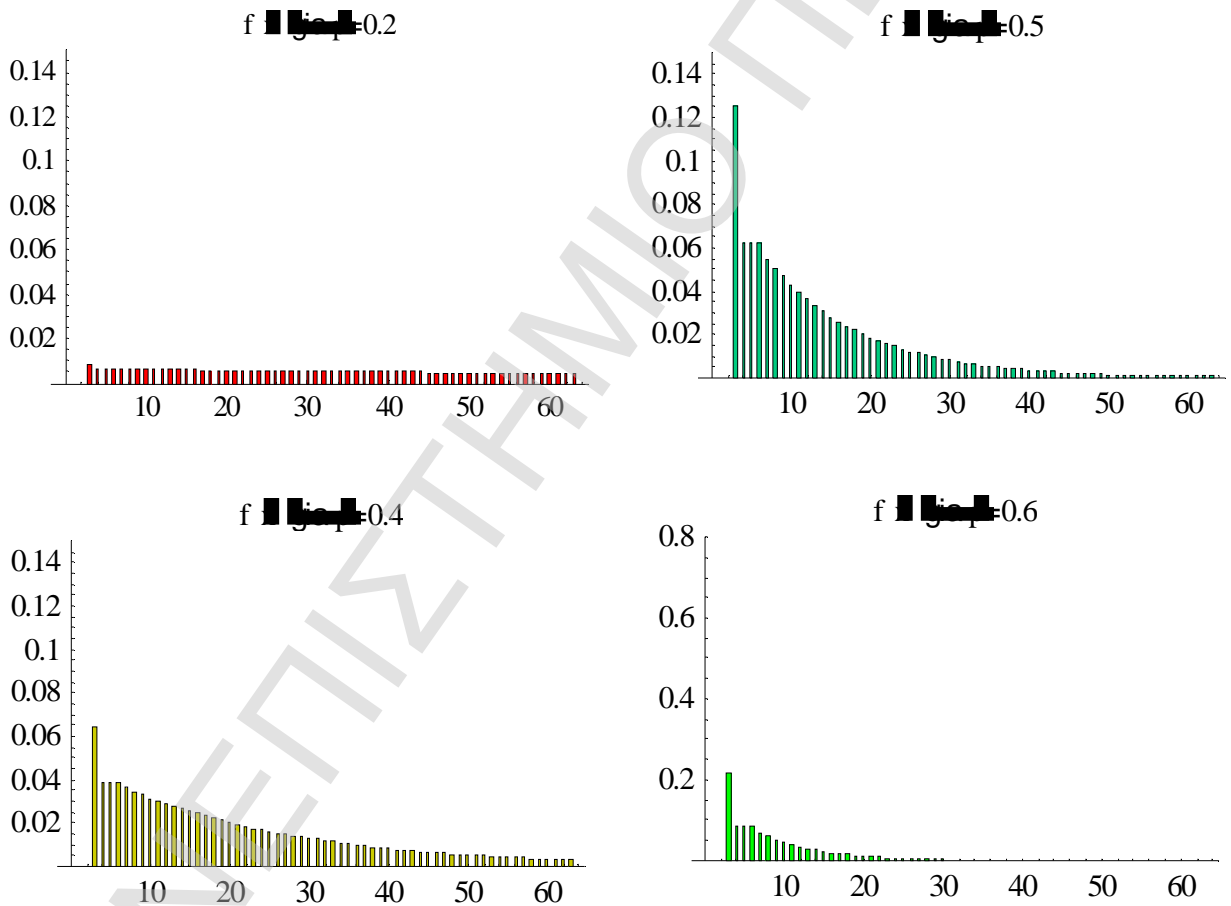
Παρακάτω θα δούμε κάποιες γραφικές παραστάσεις της συνάρτησης πυκνότητας πιθανότητας της T_k , για διάφορες τιμές της τάξης k της ροής επιτυχιών και για διάφορες τιμές της πιθανότητας επιτυχίας p στις δοκιμές. Για τον σχεδιασμό των παρακάτω γραφικών παραστάσεων χρησιμοποιήσαμε τον ακόλουθο αλγόριθμο στο Mathematica, αλλάζοντας τις τιμές του k και του p για την κάθε περίπτωση:

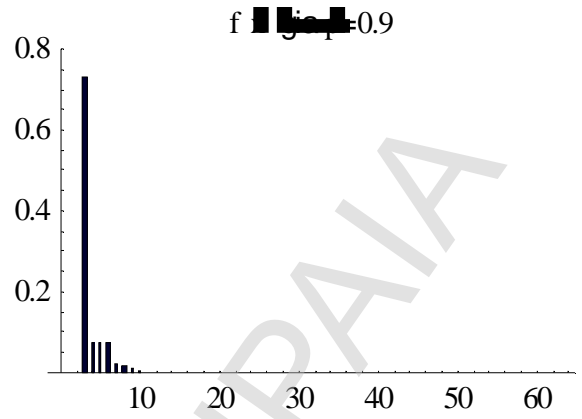
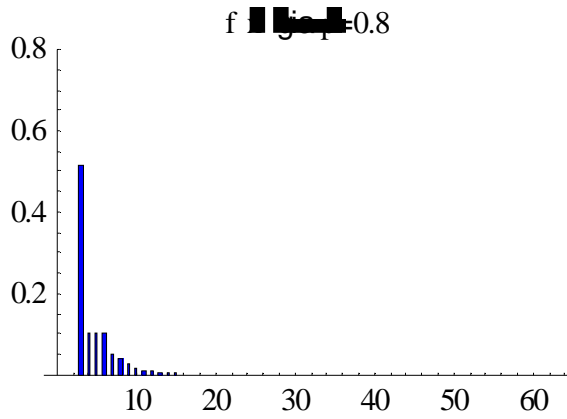
```
k=3;
p=0.6;
q=1-p;
"x>=k+1"
f[x_]:=Sum[ ((q^i)*(p^(x-i))) * Sum[ ((-1)^j) * Binomial[i,j] * Binomial[x-k*(j+1)-1,i-1], {j,0,IntegerPart[(x-i-k)/k]}], {i,1,x-k}];
x>=k+1
<<Graphics`Graphics`
c=60;
w=0.4;
list1={};
```

```
Do[If[x<=(k-
1),list1=Append[list1,{x,0,w}],If[x<=k,list1=Append[list1,{x,p^
k,w}],list1=Append[list1,{x,f[x],w}]],{x,0,k+c}]
GeneralizedBarChart[list1,BarStyle->{RGBColor[0,1,0]},PlotRange
->{0,0.8},PlotLabel->"f(x) για
p=0.6",TextStyle->{"FontFamily"->"Times New
Roman","FontSize"->12}]
```

ΔΙΑΓΡΑΜΜΑΤΑ 2.2.1

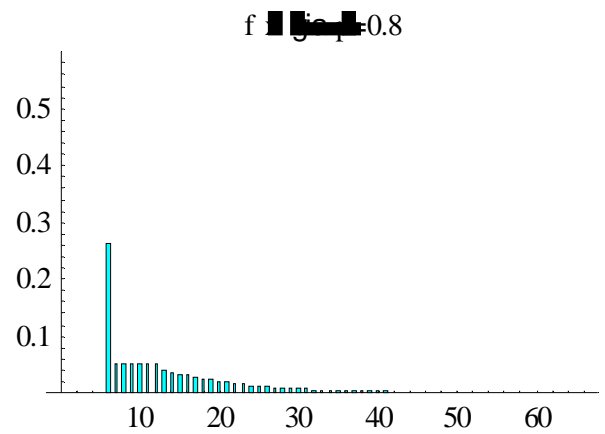
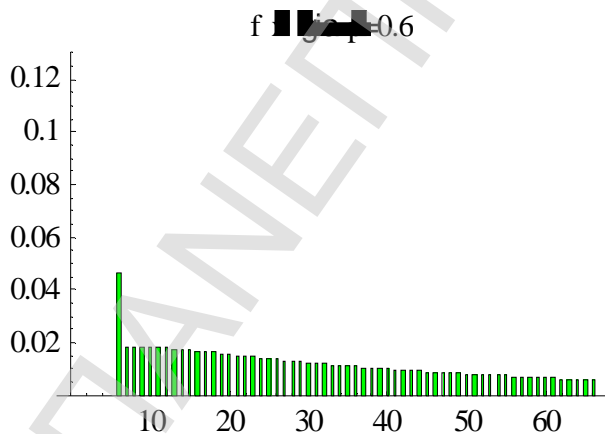
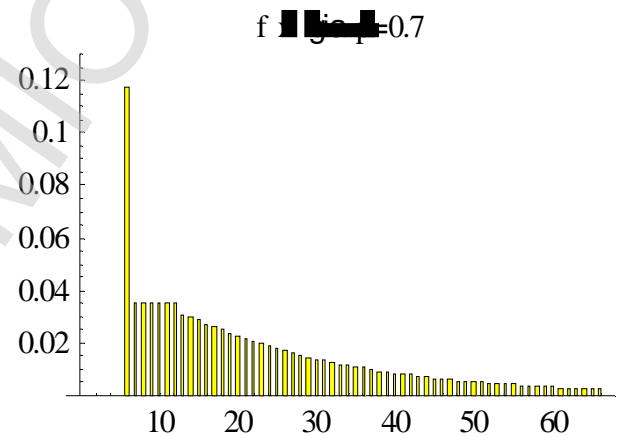
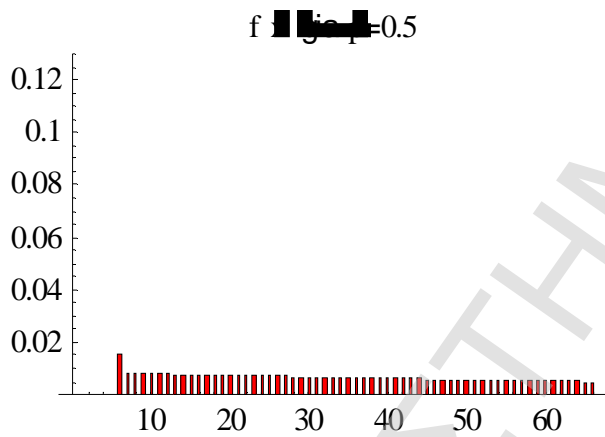
Γεωμετρική κατανομή τάξης $k=3$

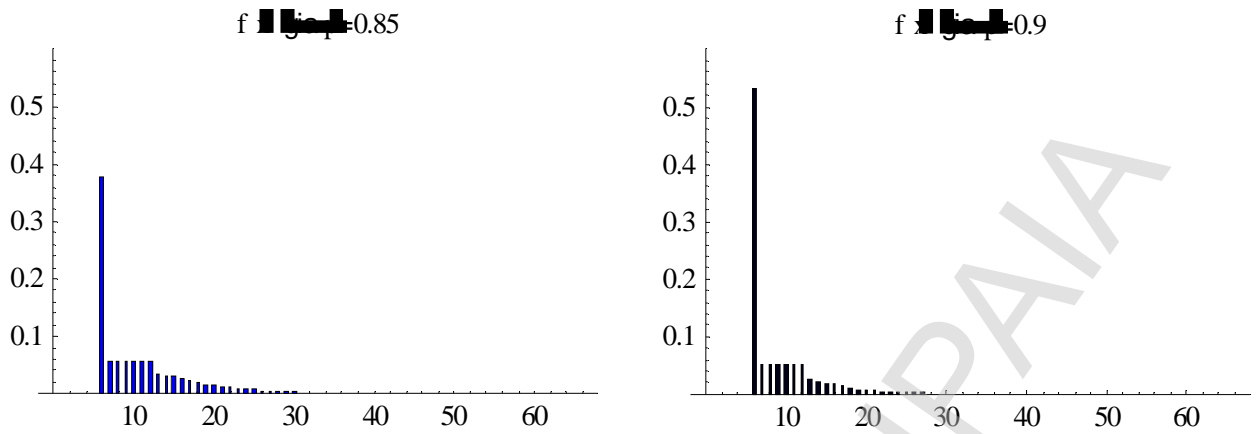




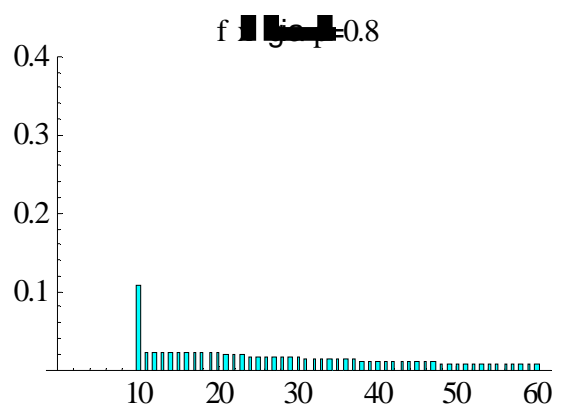
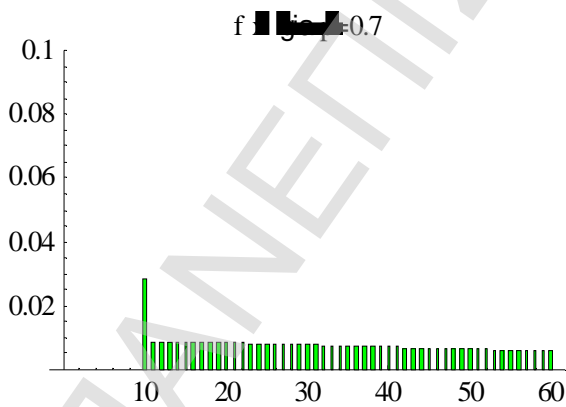
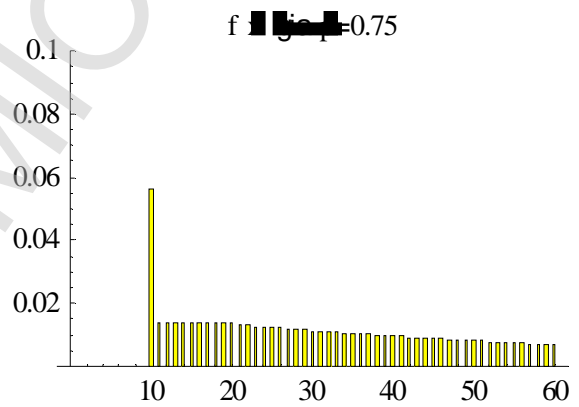
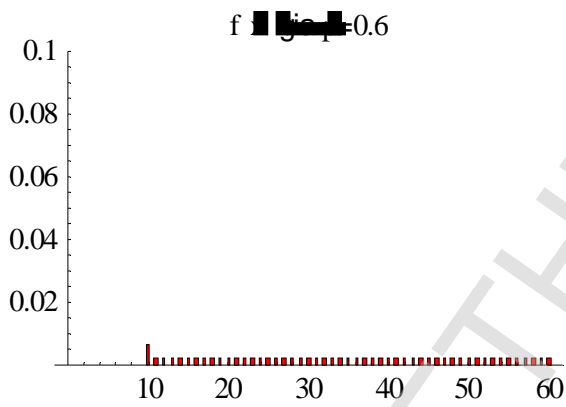
ΔΙΑΓΡΑΜΜΑΤΑ 2.2.2

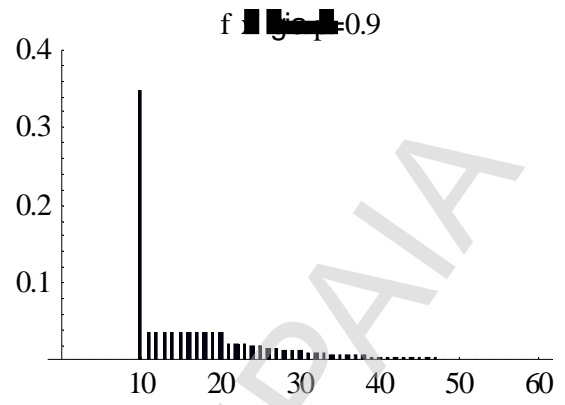
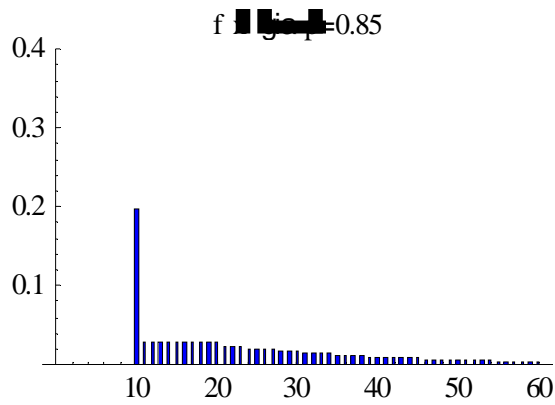
Γεωμετρική κατανομή τάξης $k=6$





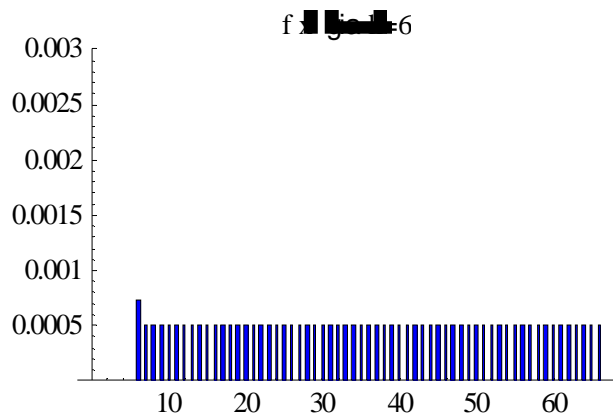
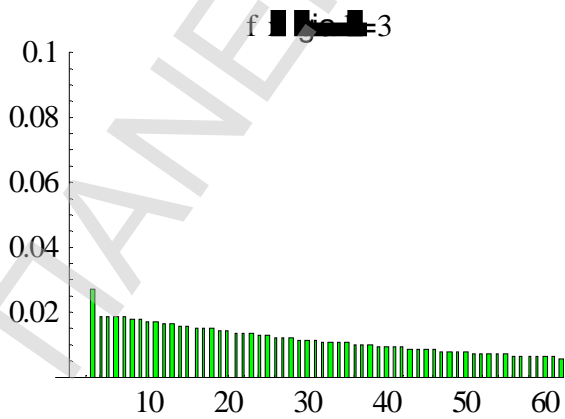
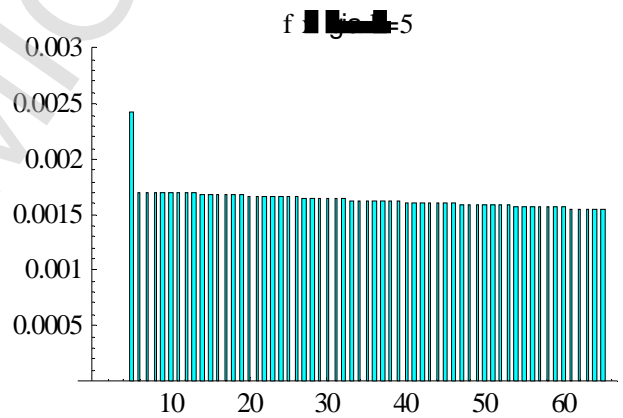
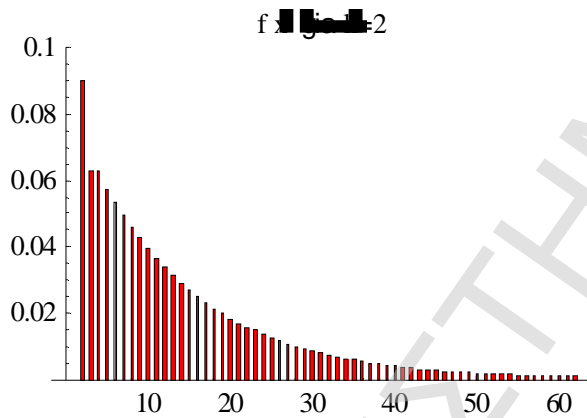
ΔΙΑΓΡΑΜΜΑΤΑ 2.2.3
Γεωμετρική κατανομή τάξης $k=10$

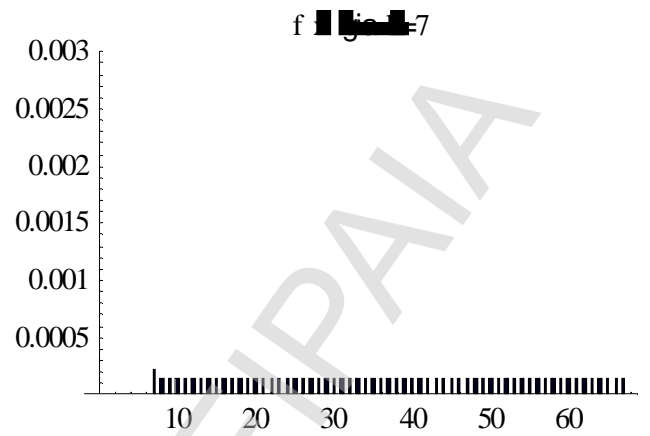
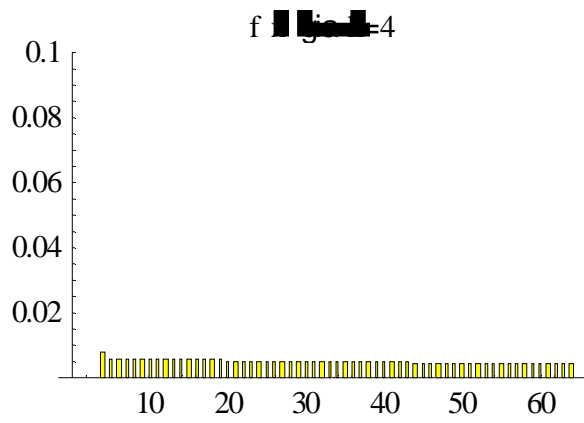




ΔΙΑΓΡΑΜΜΑΤΑ 2.2.4

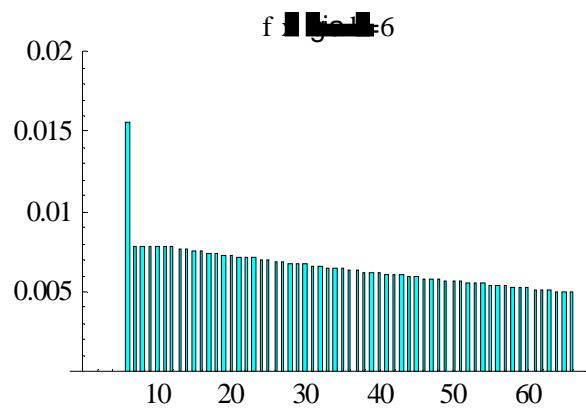
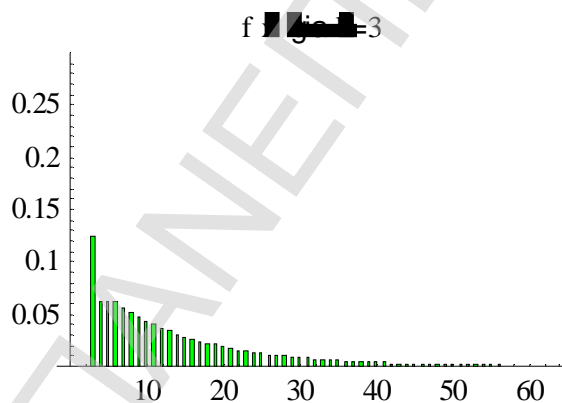
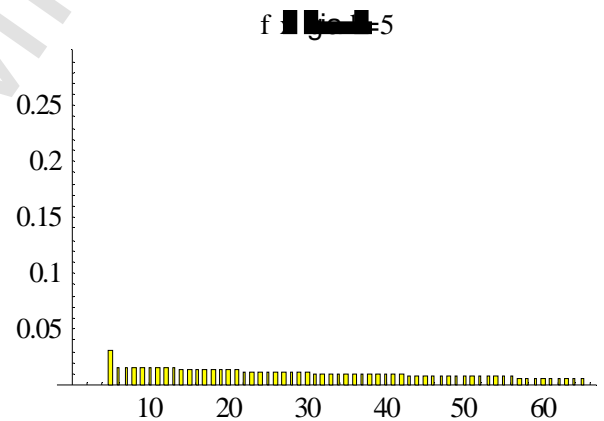
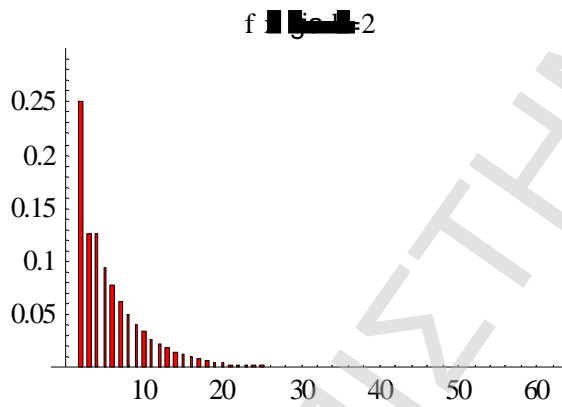
Γεωμετρική κατανομή τάξης k και πιθανότητας επιτυχίας $p=0.3$

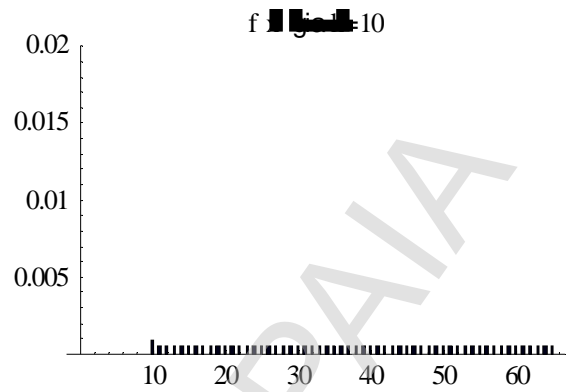
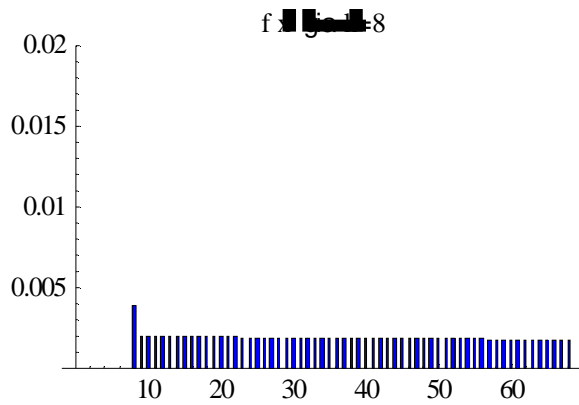




ΔΙΑΓΡΑΜΜΑΤΑ 2.2.5

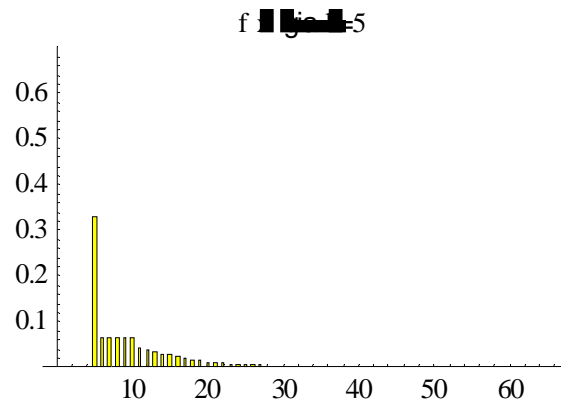
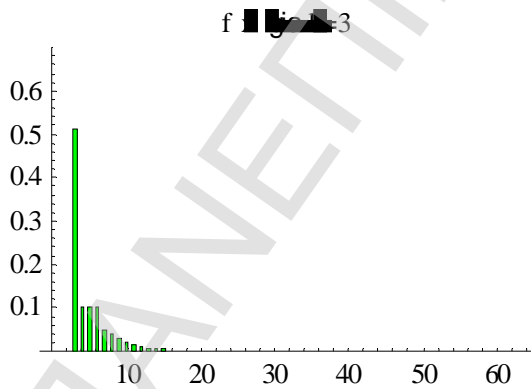
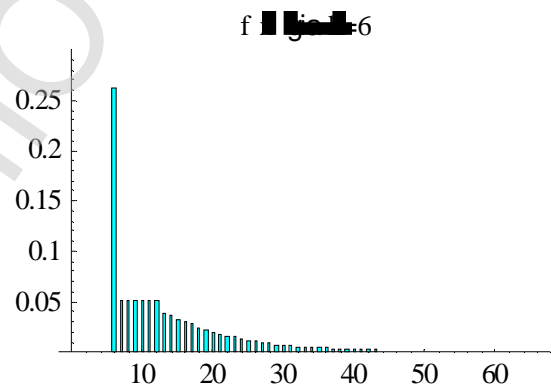
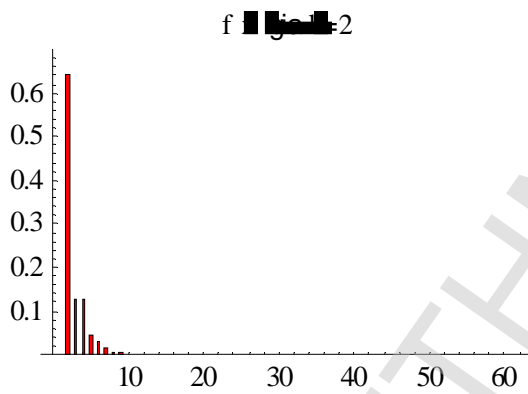
Γεωμετρική κατανομή τάξης k και πιθανότητας επιτυχίας $p=0.5$

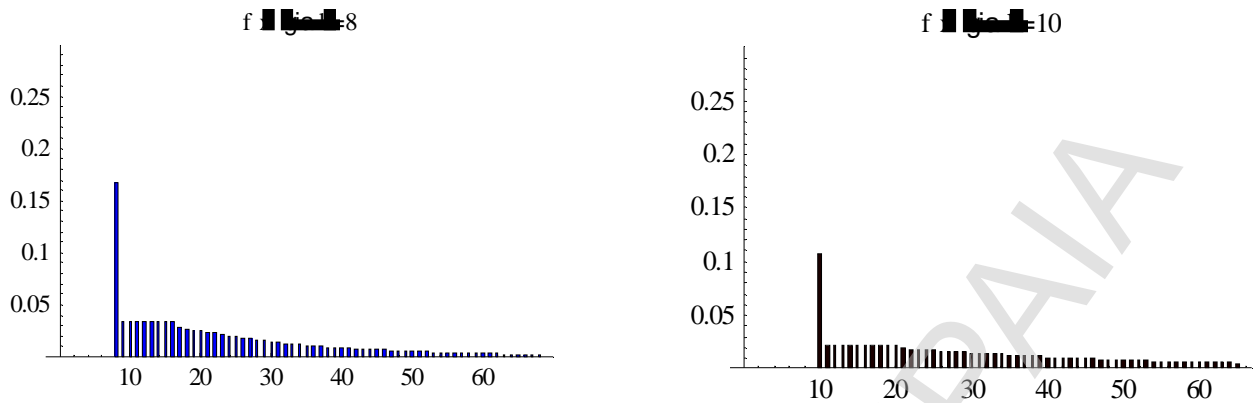




ΔΙΑΓΡΑΜΜΑΤΑ 2.2.6

Γεωμετρική κατανομή τάξης k και πιθανότητας επιτυχίας $p=0.8$





Στα διαγράμματα 2.2.1 - 2.2.3 παρατηρούμε ότι για το ίδιο k , όσο αυξάνεται η πιθανότητα επιτυχίας στην κάθε δοκιμή, τόσο μεγαλώνουν οι τιμές τις οποίες παίρνει η συνάρτηση πυκνότητας πιθανότητας $f(x)$ της T_k . Ακόμα καθώς αυξάνεται η πιθανότητα επιτυχίας p , μικραίνει το εύρος των x για τα οποία υπάρχει η $f(x)$. Δηλαδή, για ένα συγκεκριμένο k , καθώς αυξάνεται η πιθανότητα επιτυχίας p , αυξάνεται και η πιθανότητα ο χρόνος αναμονής για την πρώτη εμφάνιση ροής επιτυχιών τάξης k να είναι μικρός. Στα διαγράμματα 2.2.4 – 2.2.6 παρατηρούμε ότι για συγκεκριμένη πιθανότητα επιτυχίας p , όσο αυξάνεται η τάξη k της γεωμετρικής κατανομής, τόσο μικραίνουν οι τιμές που παίρνει γεωμετρική κατανομή τάξης k , δηλαδή η κατανομή της τ.μ. T_k . Ακόμα καθώς αυξάνεται η τάξη k της γεωμετρικής κατανομής, μεγαλώνει και το εύρος των x για τα οποία υπάρχει η $f(x)$. Δηλαδή, για ένα συγκεκριμένο p , καθώς αυξάνεται η τάξη k της ροής επιτυχιών, αυξάνεται και η πιθανότητα ο χρόνος αναμονής για την πρώτη εμφάνιση ροής επιτυχιών τάξης k να είναι μεγάλος.

2.3 Μέση τιμή και διακύμανση της τυχαίας μεταβλητής T_k

Η μέση τιμή και η διακύμανση της τυχαίας μεταβλητής T_k μπορεί να υπολογιστεί μέσω της πιθανογεννήτριας συνάρτησης της T_k . Η πιθανογεννήτρια συνάρτηση δίνεται από τον τύπο

$$G(z) = E(z^{T_k}) = \sum_{x=0}^{\infty} f(x)z^x .$$

Αντικαθιστώντας την συνάρτηση πυκνότητας πιθανότητας της μεταβλητής μας T_k και κάνοντας τις πράξεις βρίσκουμε ότι η πιθανογεννήτρια συνάρτηση της T_k έχει την μορφή

$$G(z) = \frac{(pz)^k (1-pz)}{1-z+qp^k z^{k+1}} = \frac{(pz)^k}{1-qzA(z)},$$

όπου

$$A(z) = \frac{1-(pz)^k}{1-pz}.$$

Έχουμε ότι $G(z) = E(z^{T_k}) \Rightarrow G'(z) = E(T_k z^{T_k-1}) \Rightarrow G'(1) = E(T_k)$, δηλαδή η μέση τιμή $E(T_k)$ ισούται με την πρώτη παράγωγο της πιθανογεννήτριας για $z=1$. Κάνοντας τις πράξεις καταλήγουμε στον επόμενο τύπο, ο οποίος δίνει τη μέση τιμή της T_k :

$$E(T_k) = \frac{1-p^k}{qp^k}$$

Είναι γνωστό ότι η διακύμανση υπολογίζεται από την σχέση

$$Var(T_k) = E(T_k^2) - (E(T_k))^2 \quad (2.3.1)$$

Έχουμε ότι

$$\begin{aligned} G''(z) &= E(T_k(T_k-1)z^{T_k-2}) \Rightarrow G''(1) = E(T_k(T_k-1)) = E(T_k^2) - E(T_k) \\ &\Rightarrow E(T_k^2) = G''(1) + E(T_k). \end{aligned}$$

Έτσι κάνοντας τις πράξεις έχουμε τον εξής τύπο για το $E(T_k^2)$:

$$E(T_k^2) = \frac{2 - (2k+3)p^k + (2k+1)p^{k+1} + qp^{2k}}{(qp^k)^2}.$$

Τώρα μέσω της σχέσης (2.3.1), αντικαθιστώντας τον τύπο του $E(T_k^2)$ και του $E(T_k)$ με τους παραπάνω και κάνοντας τις πράξεις, βρίσκουμε τον επόμενο τύπο για τη διακύμανση της T_k :

$$Var(T_k) = \frac{1 - (2k+1)qp^k - p^{2k+1}}{(qp^k)^2}.$$

2.4 Χρόνος αναμονής μέχρι την πρώτη εμφάνιση ροής επιτυχιών τάξης k για Μαρκοβιανά εξαρτημένες δοκιμές

Σε αυτή την περίπτωση δεν έχουμε, όπως πριν, ανεξάρτητες δοκιμές αλλά Μαρκοβιανά εξαρτημένες δοκιμές. Το αποτέλεσμα κάθε δοκιμής εξαρτάται από το αποτέλεσμα της προηγούμενης δοκιμής.

Δηλαδή έχουμε μια ομοιογενή στον χρόνο διδιάστατη Μαρκοβιανή αλυσίδα X_1, X_2, \dots με πίνακα μετάβασης

$$\mathbf{P} = \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix}.$$

Κάθε μία πιθανότητα-στοιχείο του παραπάνω πίνακα δίνεται από την σχέση

$$p_{ij} = P(X_t = j / X_{t-1} = i), \quad t \geq 2, 0 \leq i, j \leq 1$$

με αρχικές πιθανότητες $p_j = P(X_1 = j)$, $j = 0, 1$. Όπως είπαμε και στα προηγούμενα, η τυχαία μεταβλητή T_k εκφράζει το χρόνο αναμονής για την πρώτη εμφάνιση ροής επιτυχιών μήκους k . Σε αυτή την περίπτωση, όπου έχουμε Μαρκοβιανά εξαρτημένες δοκιμές, θα ονομάζουμε την κατανομή της T_k , Μαρκον-γεωμετρική κατανομή τάξης k .

Στην περίπτωση της Μαρκον-γεωμετρικής κατανομής τάξης k η συνάρτηση πυκνότητας πιθανότητας της T_k , $f(x) = P(T_k = x)$ υπολογίζεται από την παρακάτω αναδρομική σχέση:

$$f(x) = p_{00}f(x-1) + \sum_{i=2}^k p_{10}p_{01}p_{11}^{i-2}f(x-1), \quad x > k+1 \quad (2.4.1)$$

με αρχικές συνθήκες :

$$f(x) = \begin{cases} 0 & \text{αν } 0 \leq x \leq k \\ p_1 p_{11}^{k-1} & \text{αν } x = k \\ p_0 p_{01} p_{11}^{k-1} & \text{αν } x = k+1. \end{cases}$$

Σε αυτή την περίπτωση μπορούμε πάλι να υπολογίσουμε την πιθανογεννήτρια συνάρτηση της τυχαίας μεταβλητής T_k , η οποία δίνεται από τον τύπο:

$$G(z) = E(z^{T_k}) = \frac{p_{11}^{k-1} \{p_1 + (p_0 p_{01} - p_1 p_{00})z\} z^k}{1 - p_{00}z - p_{01}p_{10}z^2 - \sum_{i=0}^{k-2} (p_{11}z)^i} = \frac{\{p_1 + (p_0 p_{01} - p_1 p_{00})z\} (p_{11}z)^k}{p_{11} + (p_{01}p_{10} - p_{00}p_{11})z - p_{01}p_{10}zA(z)},$$

όπου

$$A(z) = \frac{1 - (p_{11}z)^k}{1 - p_{11}z} .$$

Υπάρχουν άλλοι δύο τρόποι για να καταλήξουμε στον παραπάνω τύπο της πιθανογεννήτριας συνάρτησης $G(z)$ της T_k , ο ένας είναι μέσω της ανανεωτικής θεωρίας (βλ. Feller (1968)) και ο άλλος είναι μέσω τη τεχνικής της "τυπικής ακολουθίας".

Η συνάρτηση πυκνότητας πιθανότητας της Markov-γεωμετρικής κατανομής τάξης k μπορεί να βρεθεί με χρήση αλυσίδας Markov, όπως στην περίπτωση των i.i.d. δοκιμών, με μια μικρή τροποποίηση. Στην περίπτωση αυτή έχουμε τον παρακάτω πίνακα μετάβασης:

$$\Lambda = \begin{bmatrix} p_{00} & p_{01} & 0 & K & 0 & 0 \\ p_{10} & 0 & p_{11} & K & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ p_{10} & 0 & 0 & K & 0 & p_{11} \\ 0 & 0 & 0 & K & 0 & 1 \end{bmatrix}_{(k+1) \times (k+1)}$$

και η συνάρτηση πυκνότητας πιθανότητας της T_k έχει τώρα την μορφή

$$f(x) = P(T_k = x) = p_{11} \pi_0 \Lambda^{x-2} \mathbf{e}'_k, \quad x \geq 2$$

όπου $\pi_0 = p_0 \mathbf{e}_1 + p_1 \mathbf{e}_2$ και $\mathbf{e}_i, i=1,2,\dots, k+1$ είναι μοναδιαία διανύσματα στο χώρο \mathfrak{R}^{k+1} .

Πάλι και εδώ μπορούμε να υπολογίσουμε την πιθανογεννήτρια συνάρτηση της τυχαίας μεταβλητής T_k και έτσι έχουμε τον εξής τύπο:

$$G(z) = p_{11} z^2 \pi_0 (\mathbf{I} - \Lambda z)^{-1} \mathbf{e}'_k .$$

Από την πιθανογεννήτρια συνάρτηση, μπορούμε να υπολογίσουμε τη μέση τιμή της τυχαίας μεταβλητής T_k μέσω της σχέσης $G'(1) = E(T_k)$. Έτσι κάνοντας τις πράξεις βρίσκουμε τη μέση τιμή της T_k για Markovιανά εξαρτημένες δοκιμές, δηλαδή τη μέση τιμή της Markov-γεωμετρικής κατανομής τάξης k , η οποία δίνεται από τον τύπο:

$$E(T_k) = \frac{(p_{01} + p_{10}) - p_1 p_{11}^{k-1} + (p_1 - p_{01}) p_{11}^k}{p_{01} p_{10} p_{11}^{k-1}}$$

Για την συνάρτηση πυκνότητας πιθανότητας της T_k , όταν έχουμε Markovιανά εξαρτημένες δοκιμές, υπάρχουν και πολυωνυμικές εκφράσεις, αλλά αν θέλουμε να υπολογίσουμε την $f(x)$ της T_k αριθμητικά, προτιμάται ο αναδρομικός τύπος (2.4.1) που δώσαμε παραπάνω.

Ακόμα μπορούμε να παρατηρήσουμε ότι αν θέσουμε $p_1 = p_{01} = p_{11} = p$ και $p_0 = p_{00} = p_{10} = q$ οδηγούμαστε στα ίδια αποτελέσματα και στους ίδιους τύπους με την περίπτωση όπου έχουμε i.i.d. δοκιμές.

Επίσης πρέπει να αναφέρουμε ότι υπάρχουν και άλλοι τρόποι για την οργάνωση και τον σχηματισμό των Μαρκοβιανά εξαρτημένων δοκιμών, Koutras (1997α, β), Aki and Hirano (1993) και Hirano and Aki (1993). Στην πρώτη αναφορά από αυτές, η ακολουθία των δοκιμών είναι X_0, X_1, K και έτσι η μεταβλητή που μας ενδιαφέρει τώρα γίνεται η $T-1$. Αυτό έχει σαν αποτέλεσμα η πιθανογεννήτρια συνάρτηση να είναι ίση τώρα με $G(z)/z$. Στις άλλες δύο αναφορές η ακολουθία των δοκιμών είναι X_0, X_1, K αλλά το μέτρημα ξεκινάει από την X_1 . Σε αυτή την περίπτωση παραμερίζουμε την X_0 και θεωρούμε την $p_0 p_{00} + p_1 p_{10}$ (για αποτυχία) και την $p_0 p_{01} + p_1 p_{11}$ (για επιτυχία) ως αρχικές πιθανότητες για την X_1 .

ΚΕΦΑΛΑΙΟ 3

Δεσμευμένες κατανομές που σχετίζονται με χρόνους αναμονής ροής επιτυχιών μήκους k

3.1 Εισαγωγή

Το πρόβλημα που θα μας απασχολήσει, σχετίζεται με το χρηματιστήριο και πιο συγκεκριμένα με την εύρεση του πιο κατάλληλου χρόνου για επένδυση σε μια μετοχή. Έστω ότι παρατηρούμε την εξέλιξη της μετοχής που μας ενδιαφέρει. Κάθε μέρα η μετοχή θα παρουσιάζει ή άνοδο ή κάθοδο. Έχουμε δηλαδή μια ακολουθία από δοκιμές με δίτιμα αποτελέσματα. Θα συμβολίζουμε την άνοδο της τιμής της μετοχής με S (success) και την κάθοδο με F (failure). Θέλουμε να βρούμε την χρονική στιγμή (ημέρα), έστω x , στην οποία θα έχουμε την μεγαλύτερη πιθανότητα στις επόμενες από αυτή y δοκιμές (ημέρες), να παρουσιαστεί ροή ανόδων (επιτυχιών) τάξης k , δεδομένου ότι μέχρι αυτή την χρονική στιγμή x δεν έχει παρουσιαστεί μια τέτοια ροή. Δηλαδή ψάχνουμε την χρονική στιγμή x στην οποία θα πρέπει να επενδύσουμε γιατί στις επόμενες από αυτή y ημέρες θα έχουμε την μεγαλύτερη πιθανότητα να εμφανιστούν για πρώτη φορά k συνεχόμενες άνοδοι και επομένως θα έχουμε την μεγαλύτερη πιθανότητα για κέρδος. Έχουμε την πιθανότητα ο χρόνος αναμονής μέχρι την εμφάνιση για πρώτη φορά μιας ροής ανόδων μήκους k , να είναι μικρότερος ή ίσος με $x+y$, δεδομένου ότι αυτός ο χρόνος αναμονής είναι μεγαλύτερος του x , και ψάχνουμε σε ποιο χρονικό σημείο x αυτή η δεσμευμένη πιθανότητα γίνεται μέγιστη.

3.2 Ένα μοντέλο πιθανοτήτων για ανεξάρτητες και ισόνομες δοκιμές

Όπως και στα προηγούμενα, η τυχαία μεταβλητή T_k εκφράζει τον χρόνο αναμονής μέχρι την εμφάνιση για πρώτη φορά μιας ροής επιτυχιών (ανόδων) μήκους k . Έστω $f_k(x) = P(T_k = x)$ η σ.π.π. και $F_k(x) = P(T_k \leq x)$ η σ.κ. του χρόνου αναμονής T_k . Θέτουμε

$$f_k(x/y) = P(T_k \leq x+y / T_k > x) = \frac{P(x < T_k \leq x+y)}{P(T_k > x)} = \frac{F_k(x+y) - F_k(x)}{1 - F_k(x)} \quad (3.2.1)$$

και για σταθερό y μας ενδιαφέρει να υπολογίσουμε το x για το οποίο έχουμε το $\max_{x=1,2,K} f_k(x/y)$. Το παραπάνω είναι ένα μοντέλο πιθανοτήτων που εκφράζει το πρόβλημα που περιγράψαμε στην προηγούμενη παράγραφο, στην περίπτωση που έχουμε ανεξάρτητες και ισόνομες (i.i.d.) δοκιμές. Δηλαδή έχουμε την ίδια πιθανότητα ανόδου p της μετοχής σε κάθε ημέρα, ανεξάρτητα από το αποτέλεσμα της προηγούμενης. Το μοντέλο αυτό μπορούμε να πούμε ότι προσεγγίζει την πραγματικότητα του χρηματιστηρίου, αν υποθέσουμε ότι το χρηματιστήριο διανύει μια περίοδο κατά την διάρκεια της οποίας δεν υπάρχουν ιδιαίτερα σημαντικοί εξωτερικοί παράγοντες που να επηρεάζουν την εξέλιξή του θετικά ή αρνητικά. Τέτοιοι παράγοντες, στους οποίους εδώ υποθέτουμε ότι επικρατεί σταθερότητα, είναι οι ρυθμοί ανάπτυξης των ισχυρών οικονομιών, η γενική κατάσταση του διεθνούς νομισματικού συστήματος, οι νομισματικές ισοτιμίες, τα επιτόκια, η ανταγωνιστικότητα της οικονομίας που εξετάζουμε στην διεθνή αγορά (μεταβολές του ακαθάριστου εθνικού προϊόντος, του ιδιωτικού διαθέσιμου εισοδήματος, του εμπορικού ισοζυγίου, των επενδύσεων) και η ανάπτυξη ή η συγκράτηση του πληθωρισμού.

Προκειμένου να πάρουμε μια έκφραση για τον αριθμητή του κλάσματος στον τύπο (3.2.1) παρατηρούμε ότι:

$$F_k(x) = f_k(0) + f_k(1) + K + f_k(x),$$

$$F_k(x+y) = f_k(0) + K + f_k(x) + f_k(x+1) + K + f_k(x+y),$$

οπότε ο αριθμητής γίνεται

$$F_k(x+y) - F_k(x) = f_k(x+1) + K + f_k(x+y)$$

και ο τύπος (3.2.1) μπορεί τώρα να γραφεί με την μορφή

$$f_k(x/y) = P(T_k \leq x+y / T_k > x) = \frac{P(x < T_k \leq x+y)}{P(T_k > x)} = \frac{f_k(x+1) + K + f_k(x+y)}{1 - (f_k(0) + f_k(1) + K + f_k(x))}.$$

Ένας άλλος τρόπος για να υπολογίσουμε το κλάσμα στον τύπο (3.2.1) είναι χρησιμοποιώντας τον αναδρομικό τύπο (2.2.4) για την $F(x)$. Από τον τύπο (2.2.4) έχουμε:

$$F(x) = F(x-1) + qp^k \{1 - F(x-k-1)\}, \quad x \geq k+1,$$

και

$$F(x+y) = F(x+y-1) + qp^k \{1 - F(x+y-k-1)\}, \quad x \geq k+1,$$

με αρχικές συνθήκες που δίνονται από τον τύπο (2.2.5).

Αν τώρα στον τύπο (3.2.1) θεωρήσουμε την περίπτωση $k=1$, τότε έχουμε

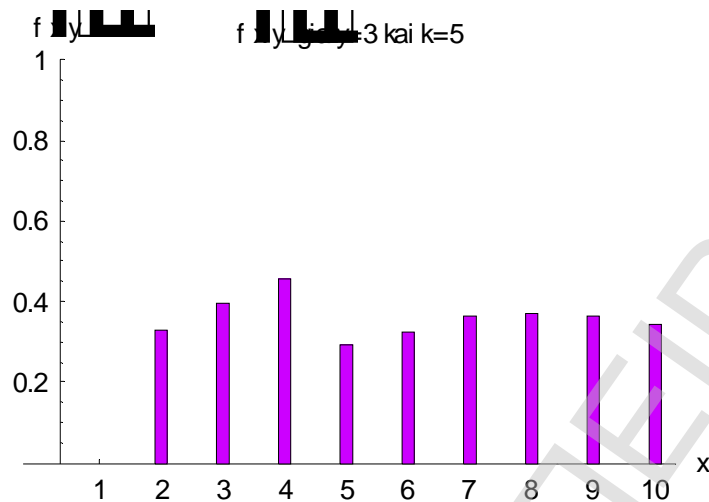
$$f_k(x/y) = \frac{P(x < T_k \leq x+y)}{P(T_k > x)} = \frac{(1-pq^{x+y}) - (1-pq^x)}{1 - (1-pq^x)} = \frac{pq^x(1-q^y)}{pq^x} = 1 - q^y.$$

Από τον παραπάνω τύπο παρατηρούμε ότι η $f_k(x/y)$ είναι ανεξάρτητη του x όταν $k=1$. Αυτό προκύπτει από την αμνήμονα ιδιότητα της γεωμετρικής κατανομής. Στην περίπτωση $k>1$ δεν αναμένεται να ισχύει αυτό. Αυτό μπορούμε να το διαπιστώσουμε με ένα παράδειγμα. Χρησιμοποιώντας τον ακόλουθο αλγόριθμο στο Mathematica θα φτιάξουμε ένα διάγραμμα που να απεικονίζει την $f_k(x/y)$ στην περίπτωση όπου έχουμε $k=5$:

```
y=3;
k=5;
"x>=k+1"
f[x_,p_]:=Sum[(((1-p)^i)*(p^(x-i)))*Sum[((-1)^j)*Binomial[i,j]*Binomial[x-k*(j+1)-1,i-1],{j,0,IntegerPart[(x-i-k)/k]}],{i,1,x-k}];
g[x_,p_]:=UnitStep[x-k]*(p^k)+UnitStep[x-(k+1)]*(Sum[f[j,p],{j,k+1,x}]);
h[x_,p_]:= (g[x+y,p]-g[x,p])/(1-g[x,p]);
c=5;
list1={};
Do[list1=Append[list1,h[x,0.80]],{x,1,k+c}];
<<Graphics`Graphics`
BarChart[list1,BarSpacing->0.5,BarGroupSpacing->0.3,BarStyle->{RGBColor[1,0,0]},PlotRange->{0,1},PlotLabel->"f(x/y) για y=3 και k=5",AxesLabel->{"x", "f(x/y)"},TextStyle->{"FontFamily"->"TimesNewRoman"}]
```

ΔΙΑΓΡΑΜΜΑ 3.2.1

Δεσμευμένη πιθανότητα $f_k(x/y)$ για $k>1$



Από ότι βλέπουμε στο παραπάνω διάγραμμα, η $f_k(x/y)$ δεν είναι ίδια για όλες τις τιμές του x , οπότε διαπιστώνουμε ότι η $f_k(x/y)$ δεν είναι ανεξάρτητη του x για $k>1$.

Παρακάτω θα δούμε κάποιες γραφικές παραστάσεις της δεσμευμένης πιθανότητας $f_k(x/y)$ για διάφορες τιμές του χρονικού ορίζοντα y , του μήκους της ροής ανόδων k και της πιθανότητας ανόδου p . Παρατηρώντας αυτές τις γραφικές παραστάσεις θα δούμε σε ποιο χρονικό σημείο x αυτή η πιθανότητα γίνεται μέγιστη, δηλαδή πότε θα έχουμε την μεγαλύτερη πιθανότητα για κέρδος και έτσι, πότε θα πρέπει να επενδύσουμε. Για τον σχεδιασμό των παρακάτω γραφικών παραστάσεων χρησιμοποιήσαμε τον ακόλουθο αλγόριθμο στο Mathematica, αλλάζοντας τις τιμές του k και του y για την κάθε περίπτωση:

```

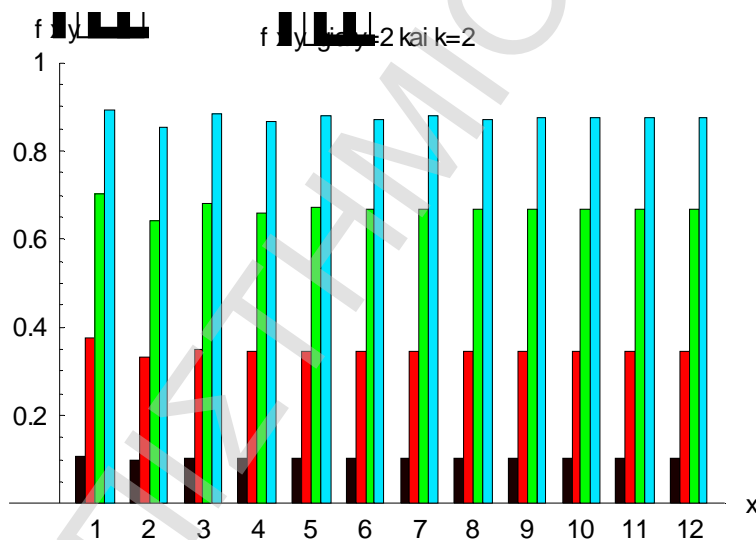
k=2;
y=2;
"x>=k+1"
f[x_,p_]:=Sum[(((1-p)^i)*(p^(x-i)))*Sum[((-1)^j)*Binomial[i,j]*Binomial[x-k*(j+1)-1,i-1],{j,0,IntegerPart[(x-i-k)/k]}],{i,1,x-k}];
g[x_,p_]:=UnitStep[x-k]*(p^k)+UnitStep[x-(k+1)]*(Sum[f[j,p],{j,k+1,x}]);
h[x_,p_]:= (g[x+y,p]-g[x,p])/(1-g[x,p]);
c=10;
list1={};
list2={};
list3={};
list4={};
Do[list1=Append[list1,h[x,0.25]],{x,1,k+c}];
Do[list2=Append[list2,h[x,0.5]],{x,1,k+c}];
Do[list3=Append[list3,h[x,0.75]],{x,1,k+c}];
Do[list4=Append[list4,h[x,0.9]],{x,1,k+c}];
<<Graphics`Graphics`
    
```

```
BarChart[list1,list2,list3,list4,BarSpacing→0,BarGroupSpacing
→0.3,BarStyle→{RGBColor[0.1,0,0],RGBColor[1,0,0],RGBColor[0,1
,0],RGBColor[0,0.9,1]},PlotRange→{0,1},PlotLabel→"f(x/y) για
y=2 και
k=2",AxesLabel→{"x","f(x/y)"},TextStyle→{"FontFamily"→"TimesN
ewRoman"}]
```

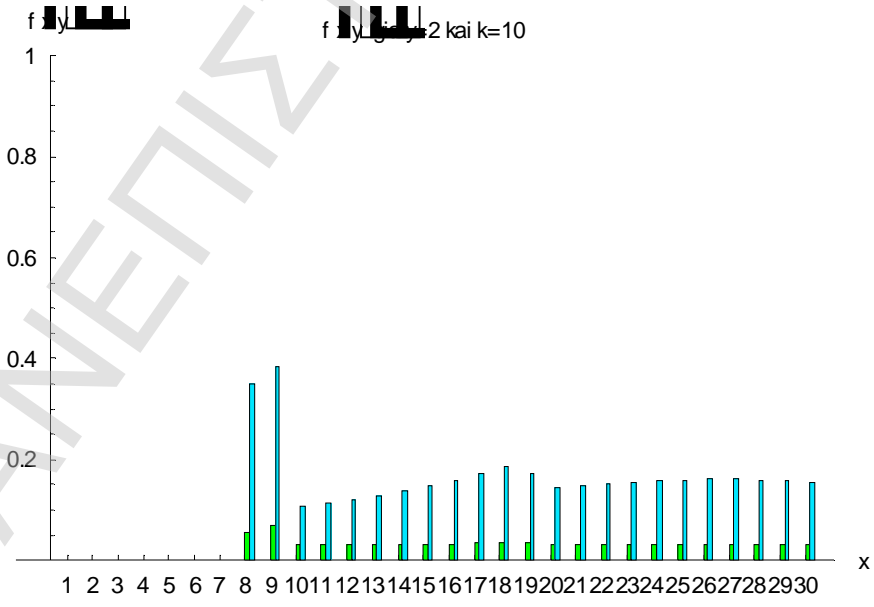
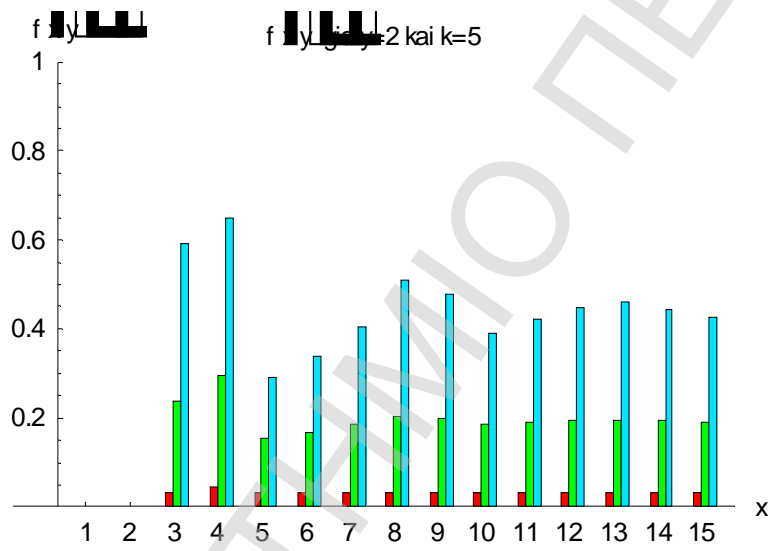
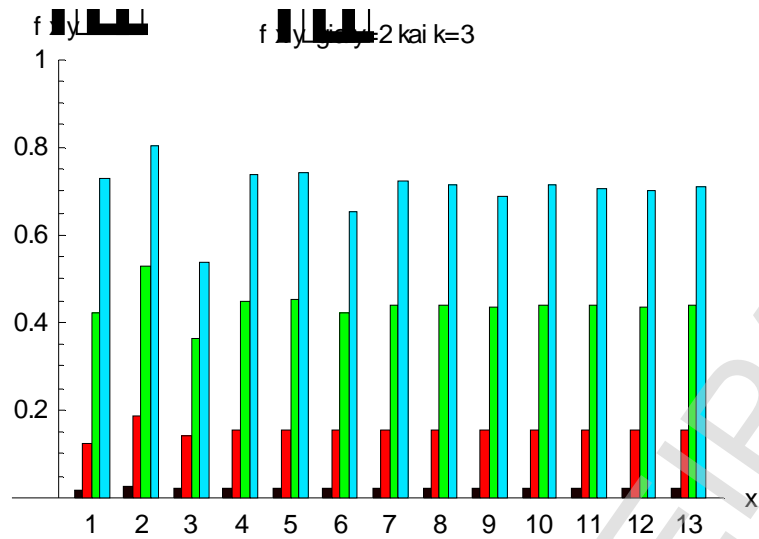
Στις παρακάτω γραφικές παραστάσεις σε κάθε χρονικό σημείο (ημέρα) x αντιστοιχούν τέσσερις μπάρες. Οι μπάρες αυτές απεικονίζουν την $f_k(x/y)$ για διαφορετικές πιθανότητες ανόδου p . Η μαύρη μπάρα αντιστοιχεί σε $p=0.25$, η κόκκινη σε $p=0.5$, η πράσινη σε $p=0.75$ και η γαλάζια σε $p=0.90$.

ΔΙΑΓΡΑΜΜΑΤΑ 3.2.2

Δεσμευμένη πιθανότητα $f_k(x/y)$ για $y=2$



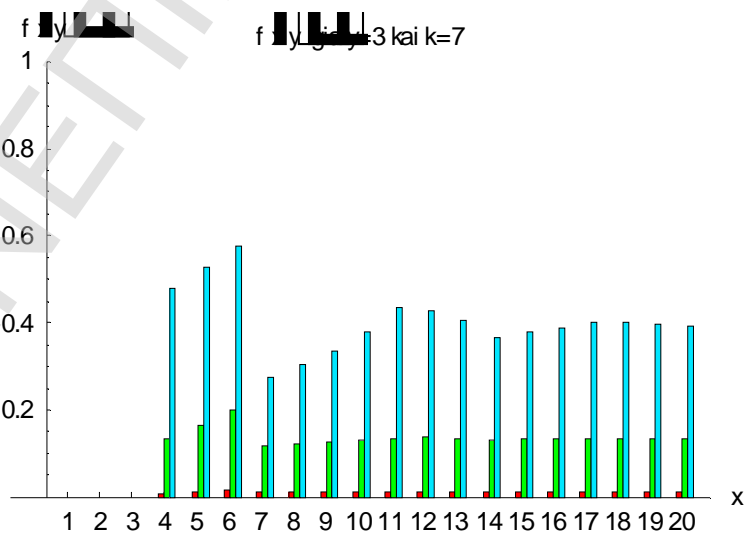
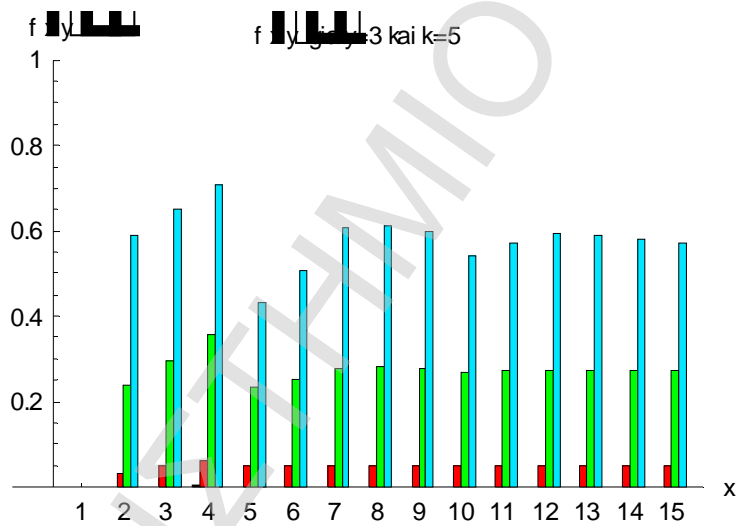
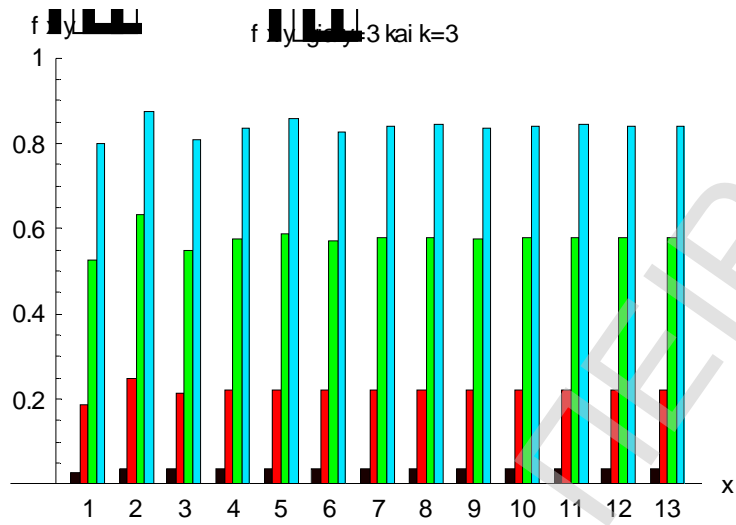
ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ:
 ΣΤΟΧΑΣΤΙΚΑ ΥΠΟΔΕΙΓΜΑΤΑ ΡΟΩΝ ΣΕ ΘΕΜΑΤΑ ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΗΣ ΑΝΑΛΥΣΗΣ

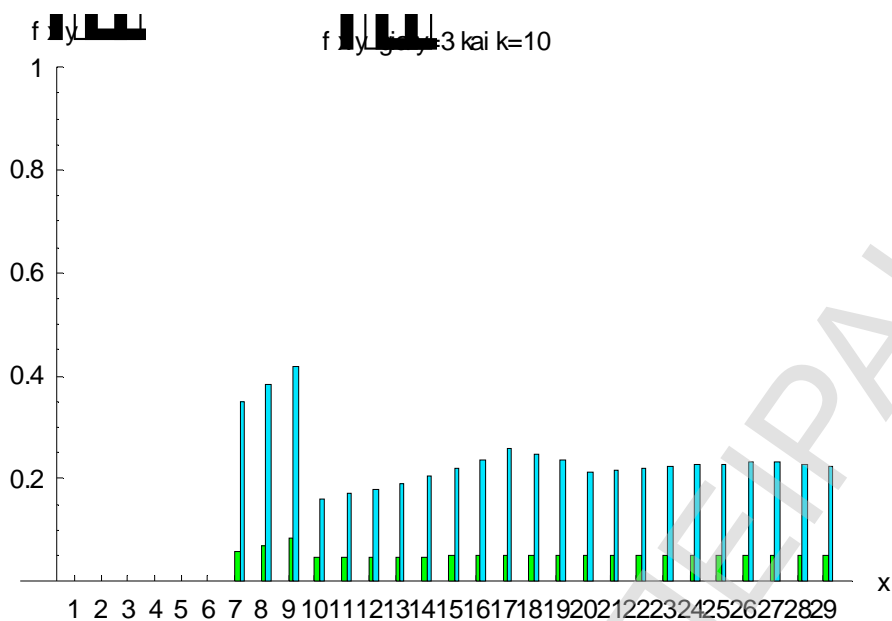


Από τα παραπάνω σχήματα παρατηρούμε ότι στην περίπτωση που έχουμε $k=2$ η $f_k(x/y)$ είναι σχεδόν ίδια για όλα τα x . Βλέπουμε ότι υπάρχουν ελάχιστες διαφορές ανάμεσα στα x , δηλαδή η $f_k(x/y)$ τείνει να γίνει ανεξάρτητη από το x . Αυτό οφείλεται στο γεγονός ότι η τιμή $k=2$ είναι πολύ κοντά στην τιμή $k=1$ και έτσι πλησιάζουμε στην αμνήμονα ιδιότητα της γεωμετρικής κατανομής που αναφέραμε προηγουμένως. Το ίδιο μπορούμε να πούμε ότι συμβαίνει, σε λίγο μικρότερο βαθμό, και στην περίπτωση όπου έχουμε $k=3$. Η τιμή $k=3$ είναι αρκετά κοντά στην τιμή $k=1$ και έτσι βλέπουμε ότι και σε αυτή την περίπτωση υπάρχουν πολύ μικρές διαφορές στην $f_k(x/y)$ ανάμεσα στα x , ειδικά για μικρές πιθανότητες ανόδου p . Στην περίπτωση όπου έχουμε $k=5$ παρατηρούμε ότι η $f_k(x/y)$ διαφοροποιείται ανάμεσα στα x . Στο διάγραμμα που αντιστοιχεί σε αυτή την περίπτωση, φαίνεται ότι η $f_k(x/y)$ παίρνει την μεγαλύτερη τιμή της στην χρονική στιγμή $x=4$. Αλλά επειδή έχουμε $k=5$, το ενδεχόμενο $T_k > x = 4$ είναι βέβαιο γι'αυτό εξετάζουμε τα x που είναι μεγαλύτερα του 5. Έτσι παρατηρούμε ότι η $f_k(x/y)$ γίνεται μέγιστη στην χρονική στιγμή $x=8$ και για τις τρεις πιθανότητες ανόδου ($p=0.5$, $p=0.75$, $p=0.90$). Ακόμα στην περίπτωση όπου έχουμε $k=5$ παρατηρούμε ότι η $f_k(x/y)$ για πιθανότητα ανόδου $p=0.25$ παίρνει πάρα πολύ μικρές τιμές γι'αυτό και δεν φαίνονται στο διάγραμμα. Στην περίπτωση όπου $k=10$, αν εξαιρέσουμε πάλι τις χρονικές στιγμές για τις οποίες το ενδεχόμενο $T_k > x$ είναι βέβαιο και έτσι η $f_k(x/y)$ παύει να είναι δεσμευμένη, παρατηρούμε από το διάγραμμα ότι η $f_k(x/y)$ παίρνει την μέγιστη τιμή της στην χρονική στιγμή $x=18$. Ακόμα για πιθανότητες ανόδου $p=0.25$ και $p=0.5$ η $f_k(x/y)$ είναι πάρα πολύ μικρή. Αν συγκρίνουμε τα παραπάνω διαγράμματα παρατηρούμε ότι όσο αυξάνεται η τάξη k της ροής ανόδων τόσο μειώνονται οι τιμές της δεσμευμένης πιθανότητας $f_k(x/y)$.

ΔΙΑΓΡΑΜΜΑΤΑ 3.2.3

Δεσμευμένη πιθανότητα $f_k(x/y)$ για $y=3$

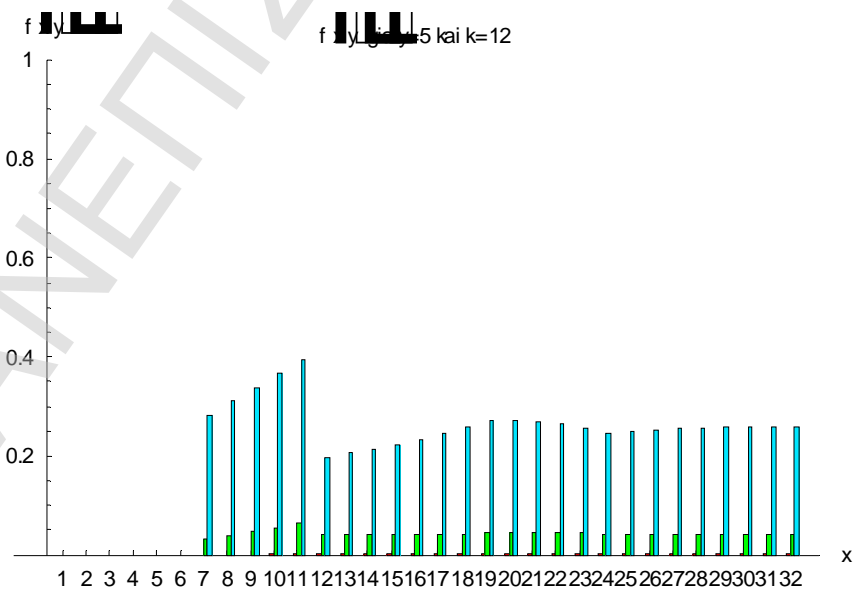
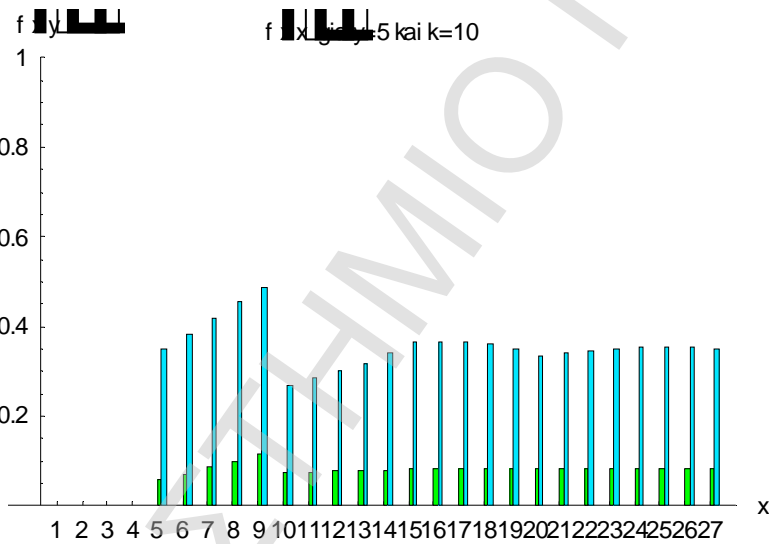
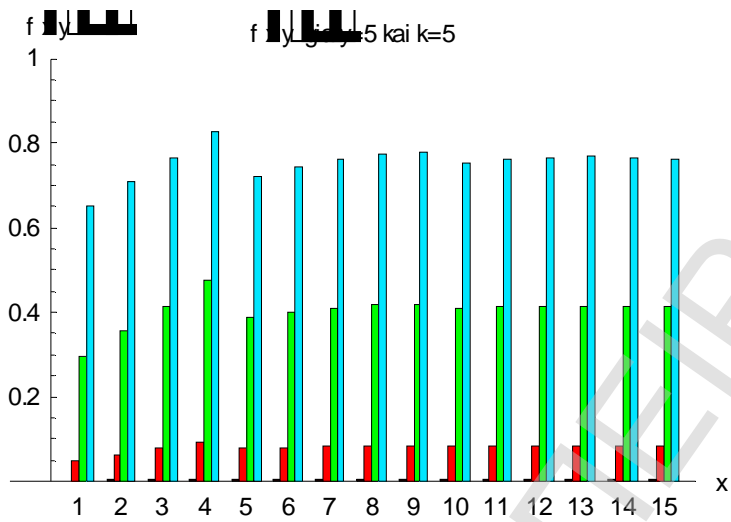


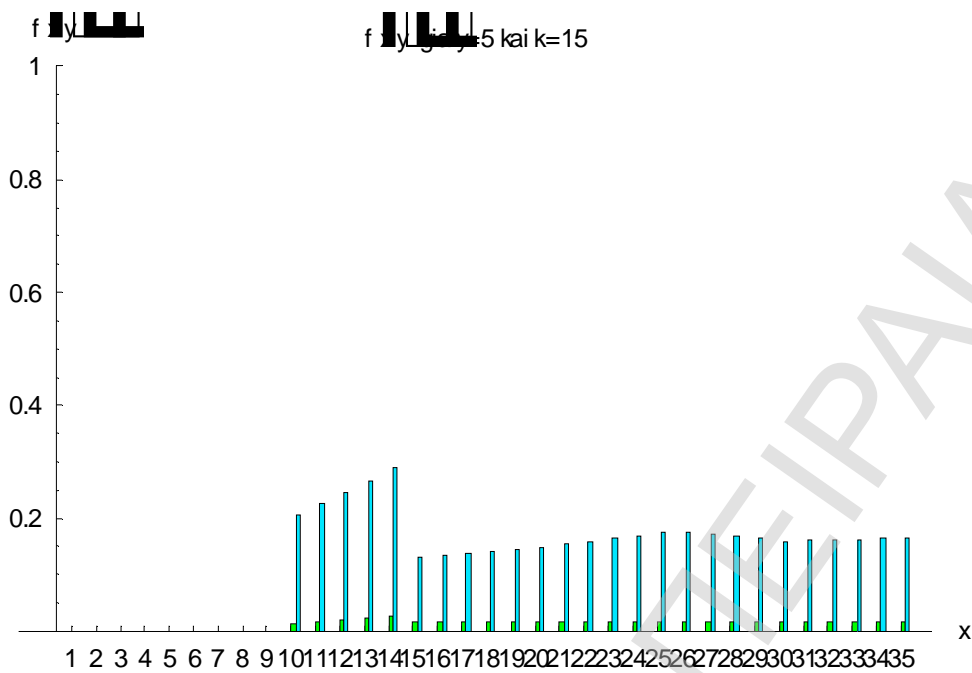


Από τα παραπάνω σχήματα παρατηρούμε ότι στην περίπτωση που έχουμε $k=3$ η $f_k(x/y)$ είναι σχεδόν ίδια για όλα τα x , δηλαδή και εδώ η $f_k(x/y)$ τείνει να γίνει ανεξάρτητη από το x . Αυτό οφείλεται, όπως αναφέραμε και στα προηγούμενα, στο ότι η τιμή $k=3$ είναι αρκετά κοντά στην τιμή $k=1$ και έτσι πλησιάζουμε στην αμνήμονα ιδιότητα της γεωμετρικής κατανομής. Στην περίπτωση όπου έχουμε $k=5$ παρατηρούνται διαφορές, όχι αρκετά σημαντικές, ανάμεσα στα x και μπορούμε να πούμε ότι η $f_k(x/y)$ παίρνει τη μέγιστη τιμή της την χρονική στιγμή $x=8$. Στην περίπτωση $k=7$ οι διαφορές ανάμεσα στα x είναι πιο έντονες και η $f_k(x/y)$ παίρνει τη μέγιστη τιμή της την χρονική στιγμή $x=11$, εξετάζοντας πάντα τις χρονικές στιγμές για τις οποίες το ενδεχόμενο $T_k > x$ δεν είναι βέβαιο. Στην περίπτωση όπου έχουμε $k=10$, η $f_k(x/y)$ παίρνει τη μέγιστη τιμή της την χρονική στιγμή $x=17$, δηλαδή αν ζητάμε ροή τάξης $k=10$ και έχουμε χρονικό ορίζοντα $y=3$ η πιο κατάλληλη χρονική στιγμή για να επενδύσουμε είναι η 17^η ημέρα.

ΔΙΑΓΡΑΜΜΑΤΑ 3.2.4

Δεσμευμένη πιθανότητα $f_k(x/y)$ για $y=5$

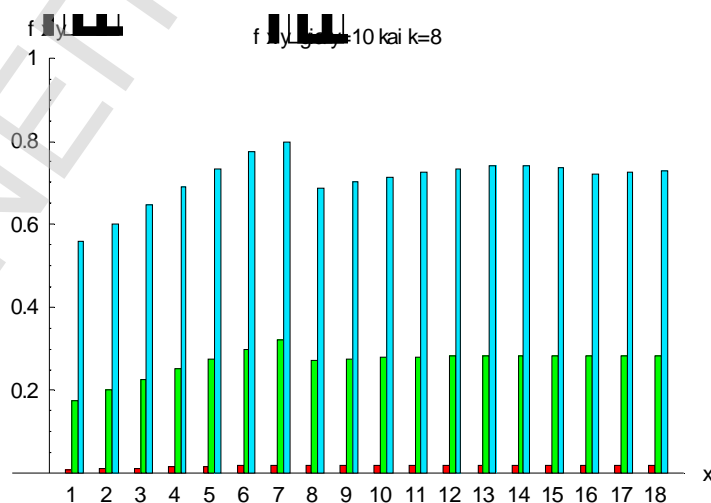




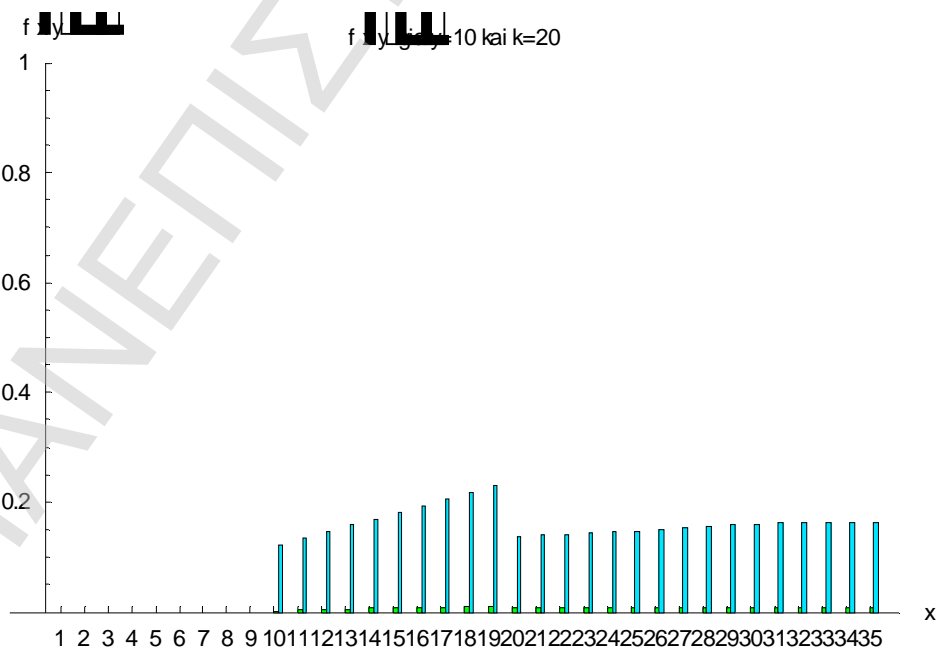
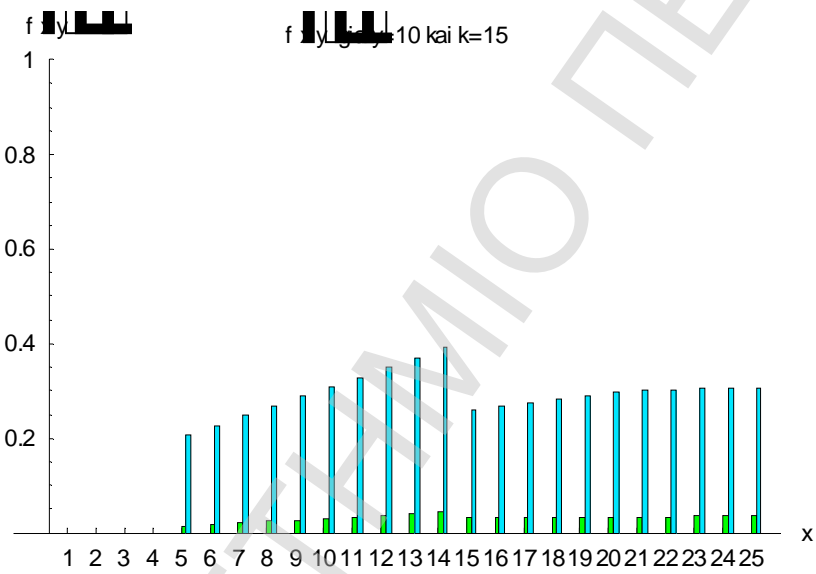
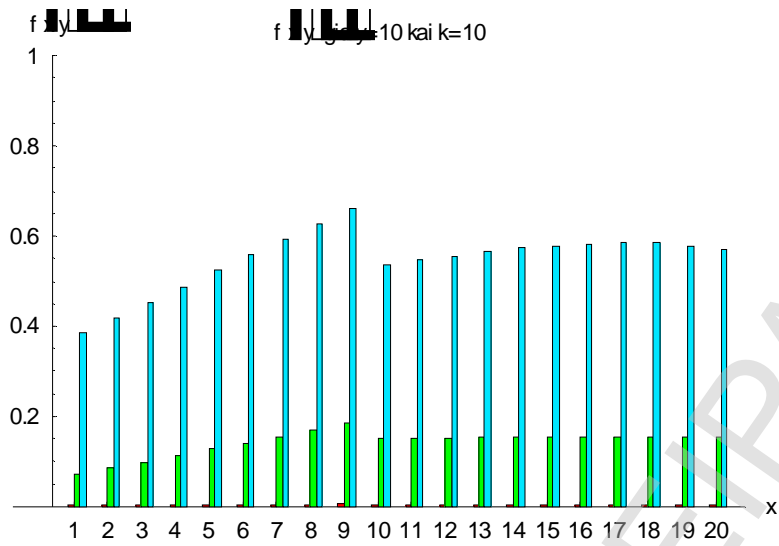
Από τα παραπάνω σχήματα παρατηρούμε ότι στην περίπτωση που έχουμε $k=5$ η $f_k(x/y)$ παρουσιάζει μικρές διαφορές ανάμεσα στα x και μέγιστη τιμή, μπορούμε να πούμε, ότι παίρνει στην χρονική στιγμή $x=9$. Στην περίπτωση όπου έχουμε $k=10$ η $f_k(x/y)$ παίρνει τη μέγιστη τιμή της, την χρονική στιγμή $x=16$, με μικρές αποκλίσεις από τις κοντινές της χρονικές στιγμές. Στην περίπτωση $k=12$, η $f_k(x/y)$ γίνεται μέγιστη την χρονική στιγμή $x=19$ και στην περίπτωση όπου έχουμε $k=15$ η $f_k(x/y)$ γίνεται μέγιστη την χρονική στιγμή $x=25$ με μικρές πάλι διαφορές από τις κοντινές χρονικές στιγμές.

ΔΙΑΓΡΑΜΜΑΤΑ 3.2.5

Δεσμευμένη πιθανότητα $f_k(x/y)$ για $y=10$



ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ:
 ΣΤΟΧΑΣΤΙΚΑ ΥΠΟΔΕΙΓΜΑΤΑ ΡΟΩΝ ΣΕ ΘΕΜΑΤΑ ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΗΣ ΑΝΑΛΥΣΗΣ



Από τα παραπάνω διαγράμματα, αν εξετάζουμε τα x για τα οποία το ενδεχόμενο $T_k > x$ δεν είναι βέβαιο, παρατηρούμε ότι σε όλες τις περιπτώσεις υπάρχουν πολύ μικρές ανάμεσα στα x . Στην περίπτωση όπου έχουμε τάξη $k=8$, φαίνεται ότι η $f_k(x/y)$ γίνεται μέγιστη την χρονική στιγμή $x=14$ με απειροελάχιστες διαφορές από τις διπλανές της χρονικές στιγμές. Πάλι με απειροελάχιστες διαφορές, στην περίπτωση όπου $k=10$, μπορούμε να πούμε ότι έχουμε την μέγιστη τιμή στην χρονική στιγμή $x=17$, στην περίπτωση $k=15$ η μέγιστη τιμή παρουσιάζεται στην χρονική στιγμή $x=24$ και στην περίπτωση $k=20$ παρουσιάζεται στην χρονική στιγμή $x=33$.

Από όλα τα παραπάνω διαγράμματα (3.2.1 - 3.2.4) παρατηρούμε ότι όταν έχουμε μικρό σχετικά χρονικό ορίζοντα y , όσο μεγαλώνει η τάξη της ροής ανόδου k , έχουμε πιο μεγάλες διαφορές στις τιμές της δεσμευμένης πιθανότητας $f_k(x/y)$ ανάμεσα στα x και η χρονική στιγμή όπου έχουμε την μέγιστη $f_k(x/y)$ ξεχωρίζει από όλες τις υπόλοιπες. Έτσι σε αυτές τις περιπτώσεις μπορούμε να υποδείξουμε την χρονική στιγμή όπου αναμένεται να έχουμε την μέγιστη πιθανότητα για κέρδος, δηλαδή την κατάλληλη χρονική στιγμή για επένδυση. Αντιθέτως, όταν έχουμε σχετικά μεγάλο χρονικό ορίζοντα y , όπως είδαμε στην περίπτωση όπου $y=10$, υπάρχουν ελάχιστες διαφορές στην $f_k(x/y)$ ανάμεσα στα x για όλες τις τιμές της τάξης ροής ανόδου k . Οπότε ερχόμαστε στο συμπέρασμα ότι όταν έχουμε μεγάλο χρονικό ορίζοντα y , η $f_k(x/y)$ τείνει να γίνει ανεξάρτητη από τον χρόνο x και έτσι δεν μας δίνεται η δυνατότητα να προσδιορίσουμε την κατάλληλη χρονική στιγμή για επένδυση. Ακόμα παρατηρώντας όλα τα παραπάνω διαγράμματα (3.2.1 - 3.2.4) συμπεραίνουμε ότι η δεσμευμένη πιθανότητα $f_k(x/y)$ έχει μια κοινή συμπεριφορά σε όλες τις περιπτώσεις συνδιασμών των y και k . Η $f_k(x/y)$ παίρνει την μέγιστη τιμή της (ολικό μέγιστο), k χρονικές στιγμές μετά από την χρονική στιγμή όπου η $f_k(x/y)$ παύει να είναι μη μηδενική. Μετά από αυτό το μέγιστο σημείο η $f_k(x/y)$ διαγράφει μια κυρτή καμπύλη, δηλαδή πρώτα φθίνει και μετά αυξάνεται, φθάνοντας μετά από k πάλι χρονικές στιγμές σε ένα τοπικό μέγιστο. Δηλαδή παρατηρούμε ότι η $f_k(x/y)$ παρουσιάζει μια περιοδικότητα μήκους k . Από την χρονική στιγμή όπου η $f_k(x/y)$ παίρνει μη μηδενικές τιμές, ανά k χρονικές στιγμές παρουσιάζει ένα μέγιστο και αυτό το μέγιστο που εμφανίζεται πρώτο έχει και την μεγαλύτερη τιμή από όλα τα άλλα. Όλη αυτή η συμπεριφορά της $f_k(x/y)$ μας υποδεικνύει τον τρόπο με τον οποίο πρέπει να επενδύσουμε, σύμφωνα με το μοντέλο αυτό. Αν έχουμε ένα συγκεκριμένο μήκος ανόδων το οποίο θέλουμε να πετύχουμε και ένα συγκεκριμένο χρονικό

ορίζοντα επένδυσης τότε στο πρώτο μέγιστο της $f_k(x/y)$, το οποίο αναφέραμε στις προηγούμενες γραμμές που εμφανίζεται, πρέπει να επενδύσουμε, γιατί σε αυτό το χρονικό σημείο έχουμε την μεγαλύτερη από όλες πιθανότητα για κέρδος. Αν δεν πάρουμε την απόφαση να επενδύσουμε αυτή την χρονική στιγμή, θα πρέπει να περιμένουμε k χρονικές στιγμές (ημέρες) για να έχουμε την αμέσως επόμενη μέγιστη πιθανότητα για κέρδος κ.ο.κ.

3.3 Ένα μοντέλο για Μαρκοβιανά εξαρτημένες δοκιμές

Στην παράγραφο αυτή θα ασχοληθούμε πάλι με το ίδιο μοντέλο που περιγράψαμε παραπάνω, μόνο που εδώ οι δοκιμές δεν θα είναι ανεξάρτητες και ισόνομες (i.i.d.), αλλά θα είναι Μαρκοβιανά εξαρτημένες. Δηλαδή έχουμε δοκιμές που το αποτέλεσμα της μίας εξαρτάται από το αποτέλεσμα της προηγούμενης δοκιμής. Αυτό το μοντέλο μπορεί να έχει εφαρμογή στην πραγματικότητα του χρηματιστηρίου υποθέτοντας ότι επικρατεί στο χρηματιστήριο μια κατάσταση συνεχούς ανόδου ή καθόδου. Αυτό μπορεί να συμβεί αν υπάρχουν εξωτερικοί παράγοντες, όπως αυτοί που αναφέραμε στην προηγούμενη παράγραφο που να επηρεάζουν το χρηματιστήριο ανοδικά ή καθοδικά. Έτσι σε αυτήν την περίπτωση η πιθανότητα ανόδου της μετοχής ή του γενικού δείκτη του χρηματιστηρίου δεν είναι ίδια για κάθε μέρα. Εξαρτάται από το αν παρατηρήθηκε άνοδος ή κάθοδος την προηγούμενη ημέρα. Στην περίπτωση που υποθέτουμε ότι επικρατεί μια ανοδική τάση στο χρηματιστήριο, αν μια μέρα παρατηρηθεί άνοδος υπάρχει μεγαλύτερη πιθανότητα να παρατηρηθεί άνοδος παρά κάθοδος και την επόμενη ημέρα.

Όπως και στα προηγούμενα, η τυχαία μεταβλητή T_k εκφράζει τον χρόνο αναμονής μέχρι την εμφάνιση για πρώτη φορά μιας ροής επιτυχιών (ανόδων) μήκους k . Εδώ όμως έχουμε Μαρκοβιανά εξαρτημένες δοκιμές, και όπως είπαμε και στα προηγούμενα θα ονομάζουμε την κατανομή της T_k , Μαρκον-γεωμετρική κατανομή τάξης k .

Έστω $f_k(x) = P(T_k = x)$ η σ.π.π. η οποία δίνεται από τον τύπο (2.4.1) και $F_k(x) = P(T_k \leq x)$ η σ.κ. του χρόνου αναμονής T_k . Θέτουμε και πάλι

$$f_k(x/y) = P(T_k \leq x+y / T_k > x) = \frac{P(x < T_k \leq x+y)}{P(T_k > x)} = \frac{F_k(x+y) - F_k(x)}{1 - F_k(x)}$$

και για σταθερό y μας ενδιαφέρει να υπολογίσουμε το x για το οποίο έχουμε το

$$\max_{x=1,2,K} f_k(x/y).$$

Παρακάτω θα δούμε κάποιες γραφικές παραστάσεις της δεσμευμένης πιθανότητας $f_k(x/y)$ για διάφορες τιμές του χρονικού ορίζοντα y , του μήκους της ροής ανόδων k και των πιθανοτήτων p_{11}, p_{10} και p_{01}, p_{00} σε κάθε ημέρα. Η πιθανότητα p_{11} είναι η πιθανότητα να έχουμε άνοδο ενώ την προηγούμενη ημέρα είχαμε και πάλι άνοδο και η πιθανότητα p_{10} είναι η πιθανότητα να έχουμε κάθοδο ενώ την προηγούμενη ημέρα είχαμε άνοδο. Είναι λογικό ότι ισχύει $p_{11} + p_{10} = 1$. Η πιθανότητα p_{01} είναι η πιθανότητα να έχουμε άνοδο ενώ την προηγούμενη ημέρα είχαμε κάθοδο και η πιθανότητα p_{00} είναι η πιθανότητα να έχουμε κάθοδο ενώ την προηγούμενη ημέρα είχαμε πάλι κάθοδο. Ισχύει και εδώ $p_{01} + p_{00} = 1$. Όταν η πιθανότητα p_{11} είναι μεγάλη, αυτό σημαίνει ότι η μετοχή που παρακολουθούμε έχει μπει σε ανοδική πορεία. Αντίστοιχα όταν η πιθανότητα p_{00} είναι μεγάλη, συμπεραίνουμε ότι η μετοχή ακολουθεί καθοδική πορεία. Στην περίπτωση τώρα, όπου οι πιθανότητες p_{10} και p_{01} , είναι και οι δύο μεγάλες, καταλαβαίνουμε ότι η μετοχή διανύει μια περίοδο σταθεροποίησης. Για τον σχεδιασμό των παρακάτω γραφικών παραστάσεων χρησιμοποιήσαμε τον ακόλουθο αλγόριθμο στο Mathematica, αλλάζοντας τις τιμές του k και του y για την κάθε περίπτωση:

```
p1=0.6;
p0=0.4;
p00=0.3;
p01=0.7;
p10=0.75;
p11=0.25;
k=3;
c=10;
l1={};
y=2;
Do[If[x<k,F=0,If[x<=k,F=p1*p11^(k-1),If[x<=k+1,F=p1*p11^(k-1)+p0*p01*(p11^(k-1)),If[x>=k+2,L=Table[0,{x+1}];L[[k+1]]=p1*p11^(k-1);L[[k+2]]=p0*p01*(p11^(k-1));Do[L[[j+1]]=p00*L[[j]]+Sum[p10*p01*(p11^(i-2))*L[[j-i+1]],{i,2,k}],{j,k+2,x}];
F=Sum[L[[i+1]],{i,k,x}];Fx=F;x=x+y;If[x<k,F=0,If[x<=k,F=p1*p11^(k-1),If[x<=k+1,F=p1*p11^(k-1)+p0*p01*(p11^(k-1)),If[x>=k+2,L=Table[0,{x+1}];L[[k+1]]=p1*p11^(k-1);L[[k+2]]=p0*p01*(p11^(k-1));Do[L[[j+1]]=p00*L[[j]]+Sum[p10*p01*(p11^(i-2))*L[[j-i+1]],{i,2,k}],{j,k+2,x}];
Fxy=Sum[L[[i+1]],{i,k,x}];Fxy=(Fxy-Fx)/(1-Fx);l1=Append[l1,fxy],{x,0,k+c}];
l1;
```

```

d1=0.6;
d0=0.4;
d00=0.3;
d01=0.7;
d10=0.45;
d11=0.55;
k=3;
c=10;
l2={};
y=2;
Do[If[x<k,F=0,If[x□k,F=d1*d11^(k-1),If[x□k+1,F=d1*d11^(k-1)+d0*d01*(d11^(k-1))],If[x≥k+2,L=Table[0,{x+1}];L[[k+1]]=d1*d11^(k-1);L[[k+2]]=d0*d01*(d11^(k-1));Do[L[[j+1]]=d00*L[[j]]+Sum[d10*d01*(d11^(i-2))*L[[j-i+1]],{i,2,k}],{j,k+2,x}];

F=Sum[L[[i+1]],{i,k,x}];Fx=F;x=x+y;If[x<k,F=0,If[x□k,F=d1*d11^(k-1),If[x□k+1,F=d1*d11^(k-1)+d0*d01*(d11^(k-1))],If[x≥k+2,L=Table[0,{x+1}];L[[k+1]]=d1*d11^(k-1);L[[k+2]]=d0*d01*(d11^(k-1));Do[L[[j+1]]=d00*L[[j]]+Sum[d10*d01*(d11^(i-2))*L[[j-i+1]],{i,2,k}],{j,k+2,x}];
Fxy=Sum[L[[i+1]],{i,k,x}];Fxy=F;fxy=(Fxy-Fx)/(1-Fx);l2=Append[l2,fxy],{x,0,k+c}];
l2;
r1=0.6;
r0=0.4;
r00=0.3;
r01=0.7;
r10=0.25;
r11=0.75;
k=3;
c=10;
l3={};
y=2;
Do[If[x<k,F=0,If[x□k,F=r1*r11^(k-1),If[x□k+1,F=r1*r11^(k-1)+r0*r01*(r11^(k-1))],If[x≥k+2,L=Table[0,{x+1}];L[[k+1]]=r1*r11^(k-1);L[[k+2]]=r0*r01*(r11^(k-1));Do[L[[j+1]]=r00*L[[j]]+Sum[r10*r01*(r11^(i-2))*L[[j-i+1]],{i,2,k}],{j,k+2,x}];

F=Sum[L[[i+1]],{i,k,x}];Fx=F;x=x+y;If[x<k,F=0,If[x□k,F=r1*r11^(k-1),If[x□k+1,F=r1*r11^(k-1)+r0*r01*(r11^(k-1))],If[x≥k+2,L=Table[0,{x+1}];L[[k+1]]=r1*r11^(k-1);L[[k+2]]=r0*r01*(r11^(k-1));Do[L[[j+1]]=r00*L[[j]]+Sum[r10*r01*(r11^(i-2))*L[[j-i+1]],{i,2,k}],{j,k+2,x}];

```



```

F=Sum[L[[i+1]],{i,k,x}];Fxy=F;fxy=(Fxy-Fx)/(1-
Fx);l3=Append[l3,fxy},{x,0,k+c}
l3;
n1=0.6;
n0=0.4;
n00=0.3;
n01=0.7;
n10=0.1;
n11=0.9;
k=3;
c=10;
l4={};
y=2;
Do[If[x<k,F=0,If[x<=k,F=n1*n11^(k-1),If[x<=k+1,F=n1*n11^(k-
1)+n0*n01*(n11^(k-
1)),If[x>=k+2,L=Table[0,{x+1}];L[[k+1]]=n1*n11^(k-
1);L[[k+2]]=n0*n01*(n11^(k-
1));Do[L[[j+1]]=n00*L[[j]]+Sum[n10*n01*(n11^(i-2))*L[[j-
i+1]],{i,2,k}],{j,k+2,x}];

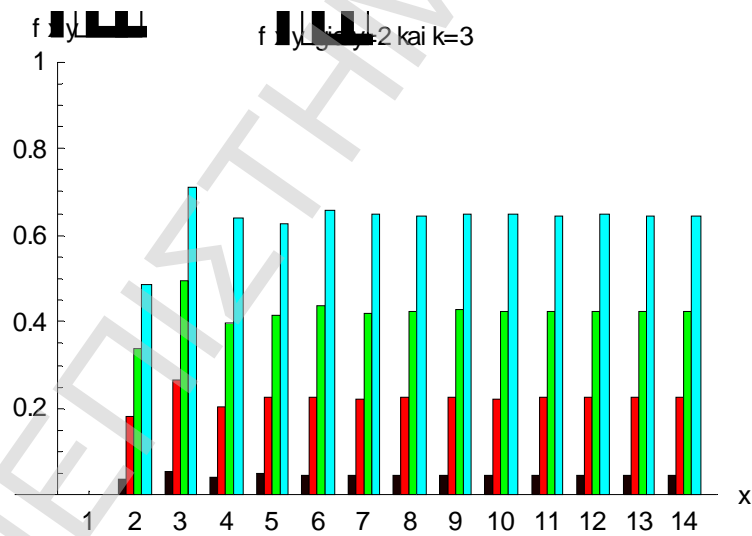
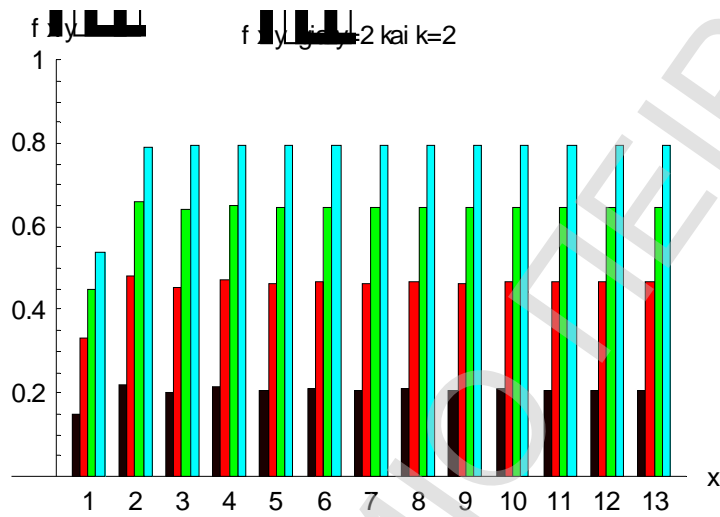
F=Sum[L[[i+1]],{i,k,x}];Fxy=F;x=x+y;If[x<k,F=0,If[x<=k,F=n1
*n11^(k-1),If[x<=k+1,F=n1*n11^(k-1)+n0*n01*(n11^(k-
1)),If[x>=k+2,L=Table[0,{x+1}];L[[k+1]]=n1*n11^(k-
1);L[[k+2]]=n0*n01*(n11^(k-
1));Do[L[[j+1]]=n00*L[[j]]+Sum[n10*n01*(n11^(i-2))*L[[j-
i+1]],{i,2,k}],{j,k+2,x}];
F=Sum[L[[i+1]],{i,k,x}];Fxy=F;fxy=(Fxy-Fx)/(1-
Fx);l4=Append[l4,fxy},{x,0,k+c}
l4;
<<Graphics`Graphics`
BarChart[l1,l2,l3,l4,BarSpacing->0,BarGroupSpacing->0.3,BarStyl
e->{RGBColor[0.1,0,0],RGBColor[1,0,0],RGBColor[0,1,0],RGBColor[
0,1,1]},PlotRange->{0,1},PlotLabel->"f(x/y) για y=2 και
k=3",AxesLabel->{"x","f(x/y)",TextStyle->{"FontFamily"->"TimesN
ewRoman"}]

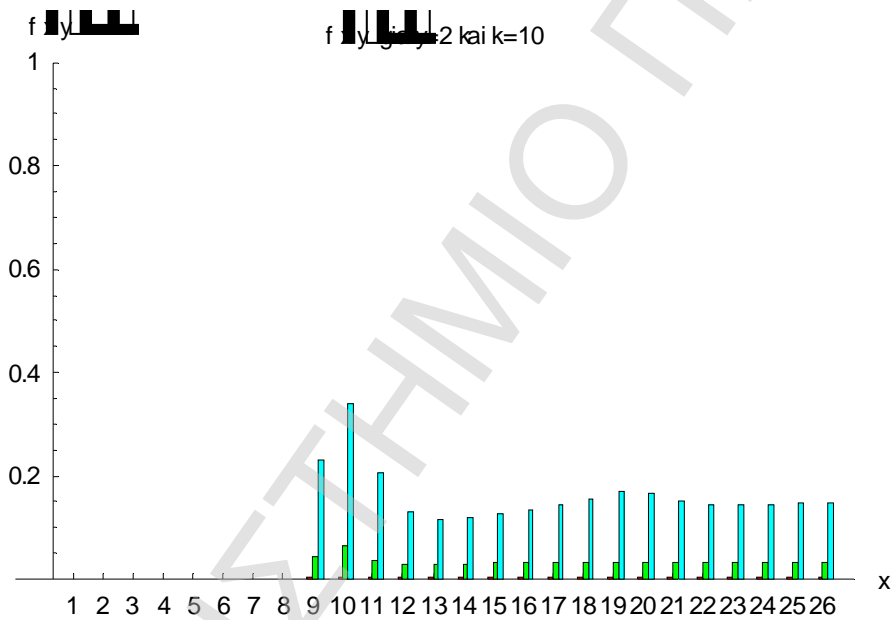
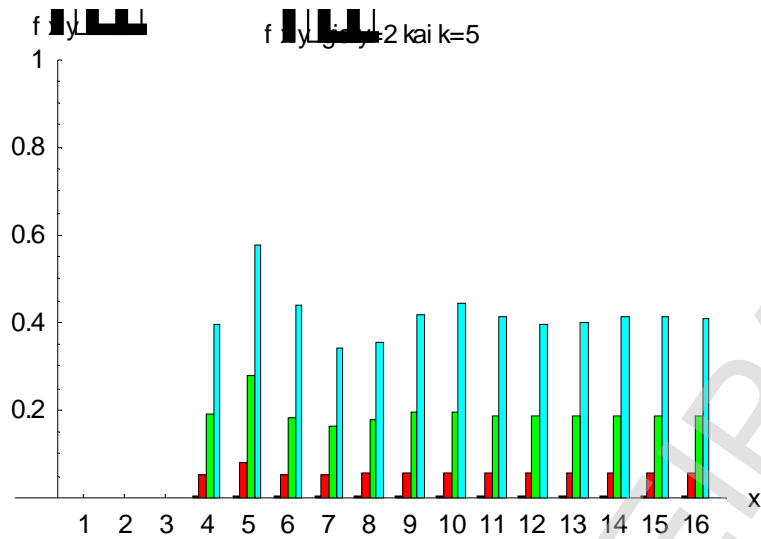
```

Παρατηρώντας αυτές τις γραφικές παραστάσεις θα δούμε σε ποιο χρονικό σημείο x αυτή η πιθανότητα γίνεται μέγιστη, δηλαδή τότε θα έχουμε την μεγαλύτερη πιθανότητα για κέρδος και έτσι τότε θα πρέπει να επενδύσουμε. Στις παρακάτω γραφικές παραστάσεις σε κάθε χρονικό σημείο (ημέρα) x αντιστοιχούν τέσσερις μπάρες. Οι μπάρες αυτές απεικονίζουν την $f_k(x/y)$ για διαφορετικές πιθανότητες p_{11}, p_{10} . Οι τιμές των p_{01}, p_{00} και των αρχικών πιθανοτήτων p_1, p_0 είναι κοινές για όλες τις μπάρες και είναι οι εξής: $p_{01} = 0.7, p_{00} = 0.3$ και $p_1 = 0.6, p_0 = 0.4$. Η μαύρη μπάρα αντιστοιχεί σε $p_{11} = 0.25, p_{10} = 0.75$, η κόκκινη σε $p_{11} = 0.55, p_{10} = 0.45$, η πράσινη σε $p_{11} = 0.75, p_{10} = 0.25$ και η γαλάζια σε $p_{11} = 0.90, p_{10} = 0.10$.

ΔΙΑΓΡΑΜΜΑΤΑ 3.3.1

Δεσμευμένη πιθανότητα $f_k(x/y)$ για $y=2$ σε Μαρκοβιανά εξαρτημένες δοκιμές

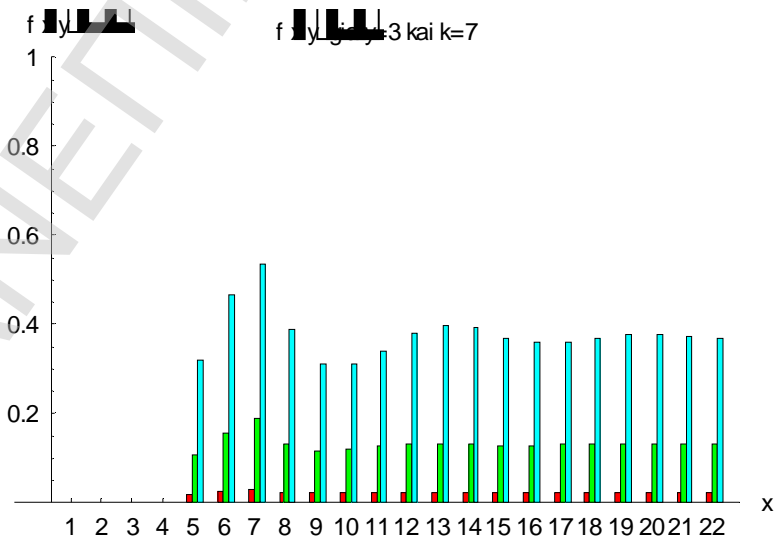
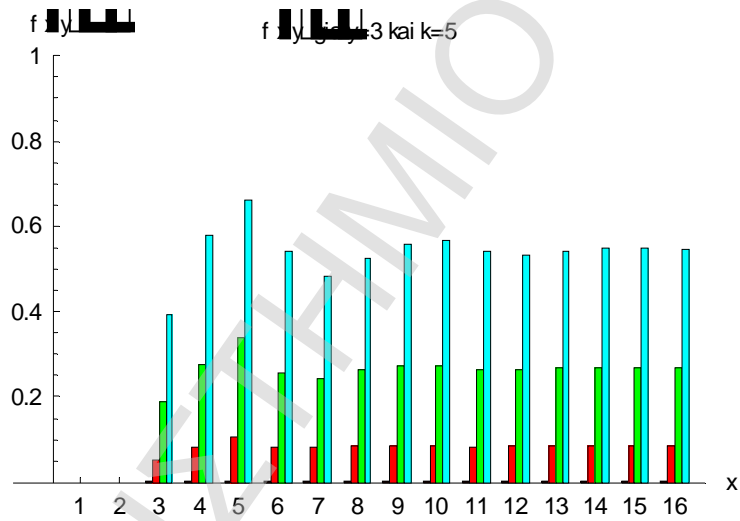
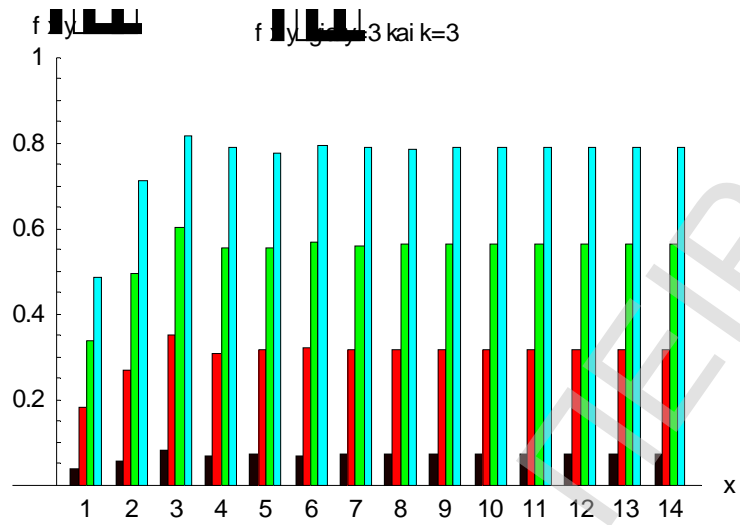


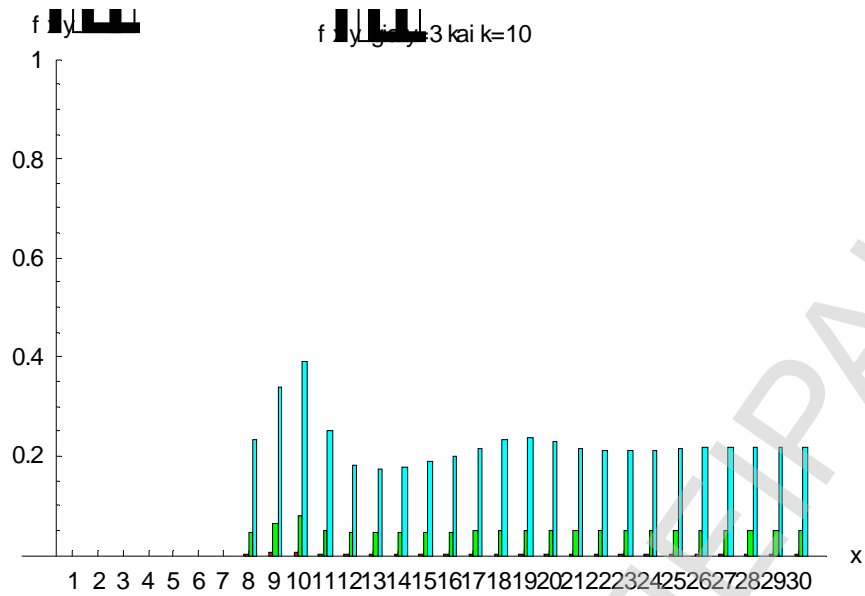


Από τα παραπάνω σχήματα παρατηρούμε ότι στην περίπτωση όπου $k=2$ δεν έχουμε διαφορές στην $f_k(x/y)$ και το ίδιο σχεδόν συμβαίνει και στην περίπτωση όπου $k=3$ με την τιμή $x=6$ να ξεχωρίζει ελάχιστα από τις άλλες. Στην περίπτωση όπου $k=5$ η $f_k(x/y)$ γίνεται μέγιστη την χρονική στιγμή $x=10$ και στην περίπτωση όπου $k=10$ η μέγιστη τιμή της $f_k(x/y)$ βρίσκεται στην χρονική στιγμή $x=19$.

ΔΙΑΓΡΑΜΜΑΤΑ 3.3.2

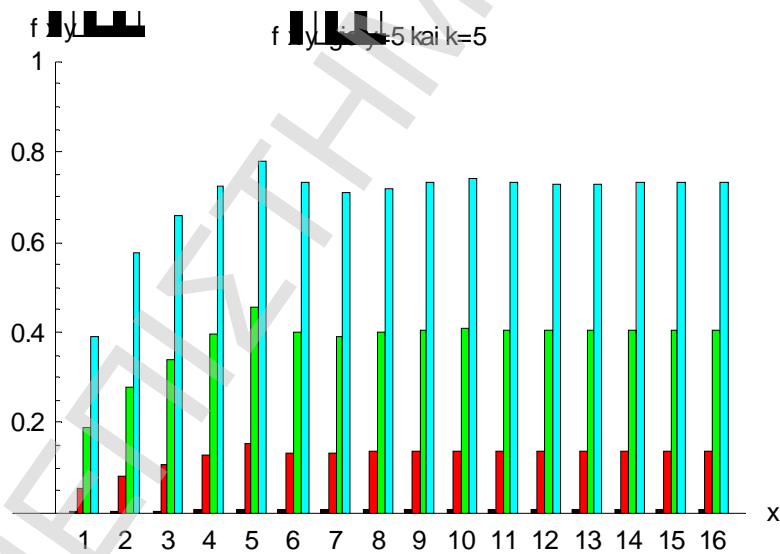
Δεσμευμένη πιθανότητα $f_k(x/y)$ για $y=3$ σε Μαρκοβιανά εξαρτημένες δοκιμές



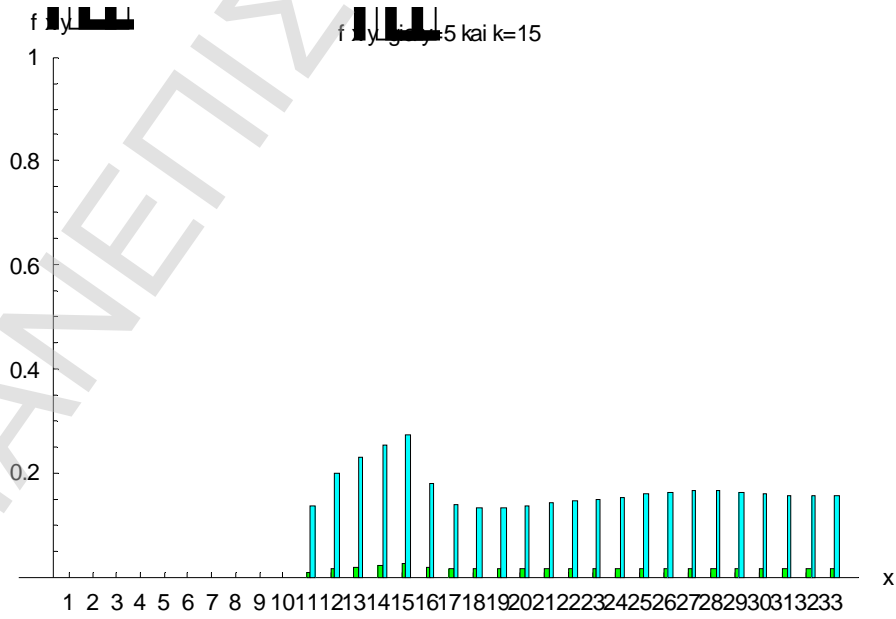
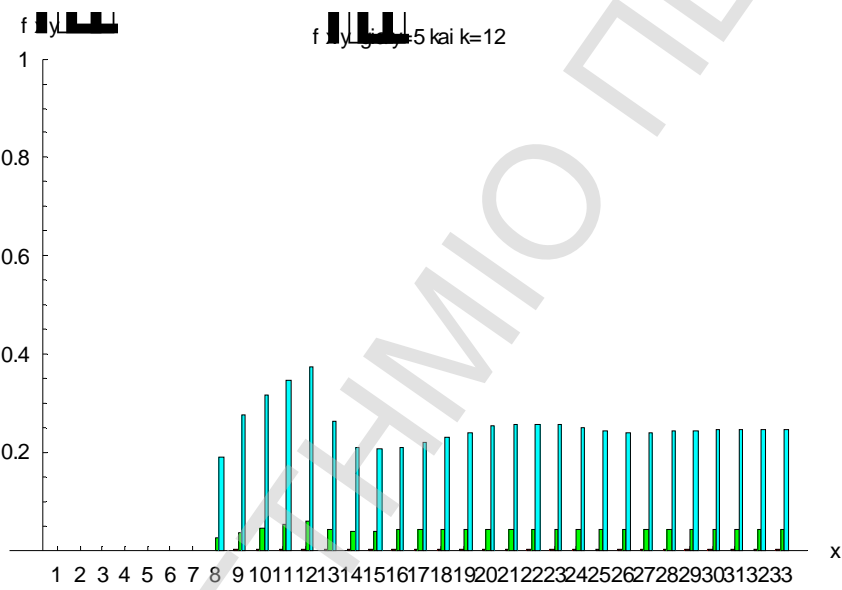
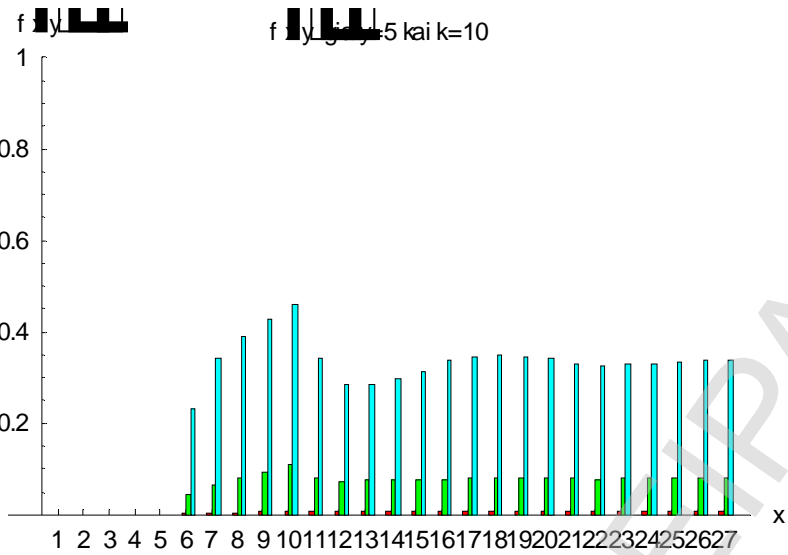


ΔΙΑΓΡΑΜΜΑΤΑ 3.3.3

Δεσμευμένη πιθανότητα $f_k(x/y)$ για $y=5$ σε Μαρκοβιανά εξαρτημένες δοκιμές

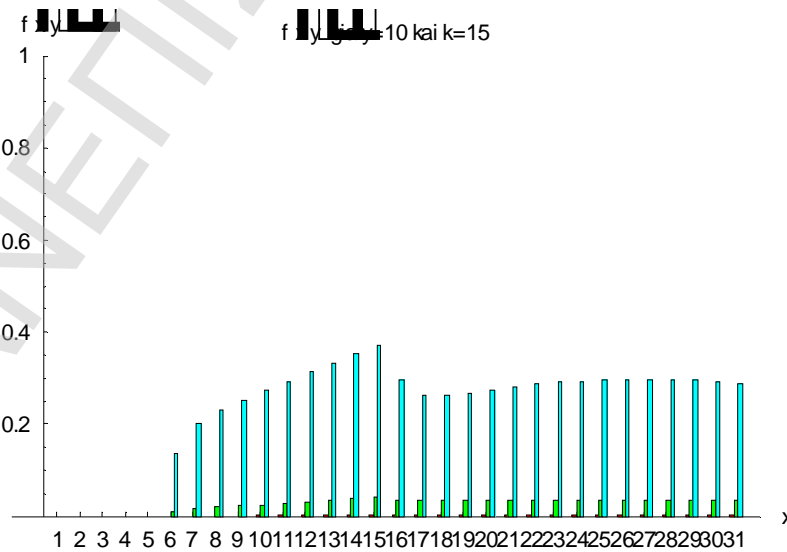
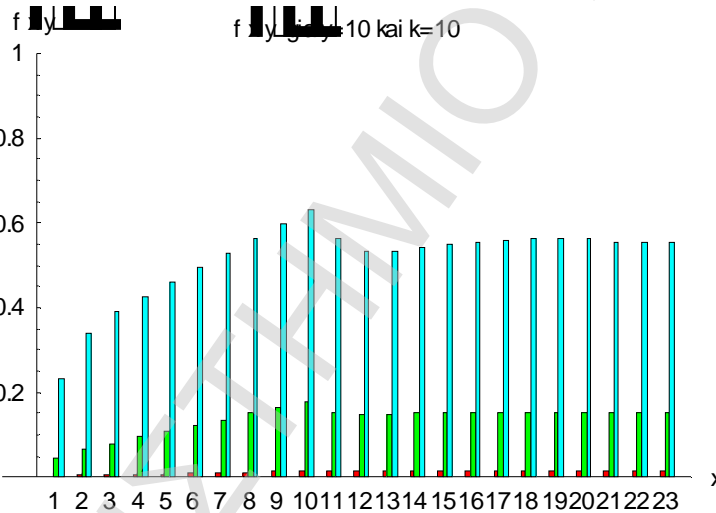
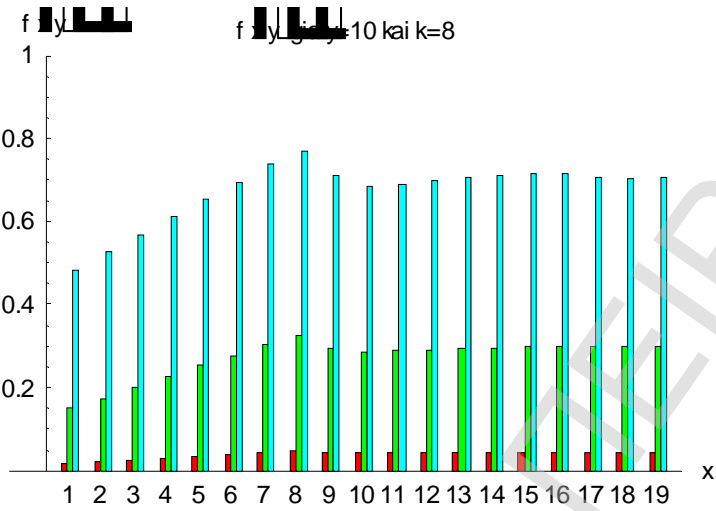


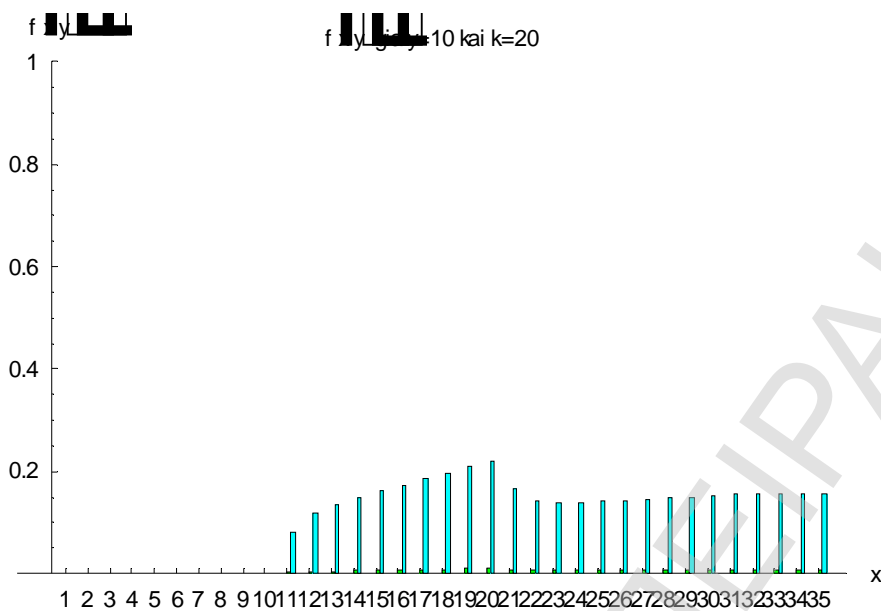
ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ:
 ΣΤΟΧΑΣΤΙΚΑ ΥΠΟΔΕΙΓΜΑΤΑ ΡΟΩΝ ΣΕ ΘΕΜΑΤΑ ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΗΣ ΑΝΑΛΥΣΗΣ



ΔΙΑΓΡΑΜΜΑΤΑ 3.3.4

Δεσμευμένη πιθανότητα $f_k(x/y)$ για $y=10$ σε Μαρκοβιανά εξαρτημένες δοκιμές





Από όλα τα παραπάνω διαγράμματα (3.3.1 - 3.3.4) παρατηρούμε ότι η $f_k(x/y)$ σε Μαρκοβιανά εξαρτημένες δοκιμές παρουσιάζει λιγότερο έντονες διαφορές ανάμεσα στα x σε σχέση με την περίπτωση των i.i.d. δοκιμών. Ακόμα βλέπουμε και εδώ, ότι όταν έχουμε μεγάλο χρονικό ορίζοντα y , η $f_k(x/y)$ τείνει να γίνει ανεξάρτητη από τον χρόνο x και έτσι δεν μας δίνεται η δυνατότητα να προσδιορίσουμε την κατάλληλη χρονική στιγμή για επένδυση. Η $f_k(x/y)$ σε όλες τις παραπάνω περιπτώσεις παρουσιάζει μια κοινή συμπεριφορά, για τιμές του k μεγαλύτερες από τις τιμές του χρονικού ορίζοντα y . Μετά από $k+1$ περίπου χρονικές στιγμές από την στιγμή που η $f_k(x/y)$ αρχίζει και παίρνει μη μηδενικές τιμές, εμφανίζεται το ολικό μέγιστο της $f_k(x/y)$. Μετά από αυτήν την χρονική στιγμή ανά k ή $k+1$ χρονικές στιγμές (ημέρες) παρουσιάζεται ένα τοπικό μέγιστο της $f_k(x/y)$. Οπότε η πρώτη μεγαλύτερη ευκαιρία για κέρδος, η χρονική στιγμή δηλαδή που πρέπει να επενδύσουμε εμφανίζεται μετά από περίπου $k+1$ μη μηδενικές τιμές της $f_k(x/y)$ και οι επόμενες, όλο και μικρότερες, ευκαιρίες για κέρδος, εμφανίζονται ανά περίπου k χρονικές στιγμές. Δηλαδή και στην περίπτωση των Μαρκοβιανά εξαρτημένων δοκιμών η $f_k(x/y)$ παρουσιάζει μια περιοδικότητα με μήκος που κυμαίνεται από $k-1$ έως $k+2$. Έχουμε πάλι περιοδικότητα στην συμπεριφορά της δεσμευμένης πιθανότητας $f_k(x/y)$ αλλά μικρότερης ακρίβειας από την περίπτωση των i.i.d. δοκιμών.

ΚΕΦΑΛΑΙΟ 4

Εφαρμογή του μοντέλου σε πραγματικά δεδομένα

4.1 Εισαγωγή

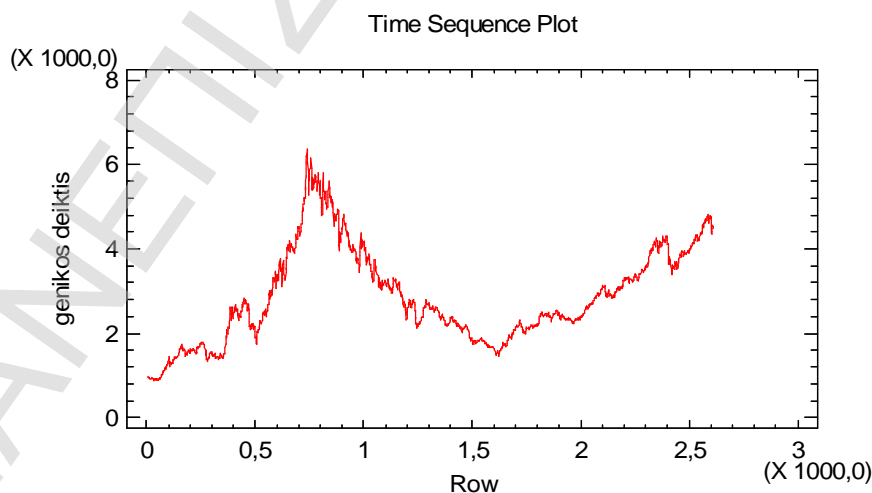
Στο κεφάλαιο αυτό θα εφαρμόσουμε το μοντέλο πιθανοτήτων, που περιγράψαμε στο προηγούμενο κεφάλαιο, σε πραγματικά δεδομένα του χρηματιστηρίου. Θα πάρουμε τρεις χρονοσειρές, την εξέλιξη του γενικού δείκτη του χρηματιστηρίου Αθηνών, την εξέλιξη της μετοχής του ΟΤΕ και της μετοχής της Εθνικής. Θα τις αναλύσουμε και θα δούμε τελικά κατά πόσο η εφαρμογή του μοντέλου είναι επιτυχής. Θα πάρουμε ορισμένες υποπεριόδους των χρονοσειρών αυτών με συγκεκριμένη πιθανότητα ανόδου p , συγκεκριμένο χρονικό ορίζοντα y και μήκος ροής k και για συγκεκριμένα x θα δούμε κατά πόσο τα αποτελέσματα από τα πραγματικά δεδομένα προσεγγίζουν τα θεωρητικά αποτελέσματα του Κεφαλαίου 3. Παίρνοντας δηλαδή κάθε φορά μια υποπερίοδο των χρονοσειρών αυτών, θα δούμε κατά πόσο θα μπορούσαμε να προβλέψουμε με επιτυχία την κατάλληλη χρονική στιγμή για επένδυση, όπου έχουμε την μεγαλύτερη πιθανότητα για κέρδος, με την βοήθεια των ροών και της δεσμευμένης πιθανότητας $f_k(x/y)$, όπως περιγράψαμε στα προηγούμενα. Θα ασχοληθούμε με την περίπτωση i.i.d. δοκιμών, θα υποθέσουμε δηλαδή ότι δεν υπάρχουν εξωτερικοί παράγοντες που να επηρεάζουν την εξέλιξη του χρηματιστηρίου ανοδικά ή καθοδικά. Δηλαδή θα εξετάσουμε την περίπτωση όπου βρισκόμαστε σε μια περίοδο σταθερότητας. Για τον λόγο αυτό, στην επόμενη παράγραφο θα εξετάσουμε την φύση των δεδομένων των τριών αυτών χρονοσειρών, για να δούμε αν σε κάθε χρονοσειρά από αυτές επικρατεί σχετική σταθερότητα, για να μπορούμε επομένως, να προσεγγίσουμε την περίπτωση των i.i.d. δοκιμών.

4.2 Φύση των δεδομένων

Τα δεδομένα που θα χρησιμοποιήσουμε στην εφαρμογή μας είναι οι τρεις χρονοσειρές που αναφέραμε παραπάνω. Πρώτα θα ασχοληθούμε με τον γενικό δείκτη του χρηματιστηρίου. Έχουμε μια χρονοσειρά που αποτελείται από 2609 παρατηρήσεις οι οποίες αντιστοιχούν στην εξέλιξη του γενικού δείκτη του χρηματιστηρίου για περίπου 9,5 χρόνια, από τον Οκτώβριο του 1996 έως τον Μάρτιο του 2007. Κάθε στοιχείο της χρονοσειράς αυτής είναι 1 ή 0 και δείχνει αν είχαμε άνοδο ή κάθοδο του δείκτη αντίστοιχα, σε κάθε ημέρα λειτουργίας του χρηματιστηρίου. Επειδή υποθέτουμε ότι έχουμε την περίπτωση i.i.d. δοκιμών θα κάνουμε ένα p διάγραμμα ελέγχου για τις παρατηρήσεις μας, παίρνοντας μέγεθος δείγματος $n=100$. Εκτιμούμε την πιθανότητα επιτυχίας από όλες τις παρατηρήσεις μας και την θέτουμε ως κεντρική τιμή (δηλαδή φτιάχνουμε p διάγραμμα ελέγχου φάσης II με 3σ όρια). Η εκτιμημένη από όλες τις παρατηρήσεις πιθανότητα ανόδου είναι $\bar{P}=0,5161$. Ακόμα φτιάχνουμε το διάγραμμα της χρονοσειράς αυτής (Time Sequence Plot), παίρνοντας τις τιμές κλεισίματος του γενικού δείκτη για κάθε ημέρα, για να έχουμε μια γενική εικόνα της εξέλιξής της. Τα δύο αυτά διαγράμματα που αφορούν το γενικό δείκτη του χρηματιστηρίου είναι τα παρακάτω. Για τον σχεδιασμό των παρακάτω διαγραμμάτων χρησιμοποιήσαμε το στατιστικό πρόγραμμα Statgraphics.

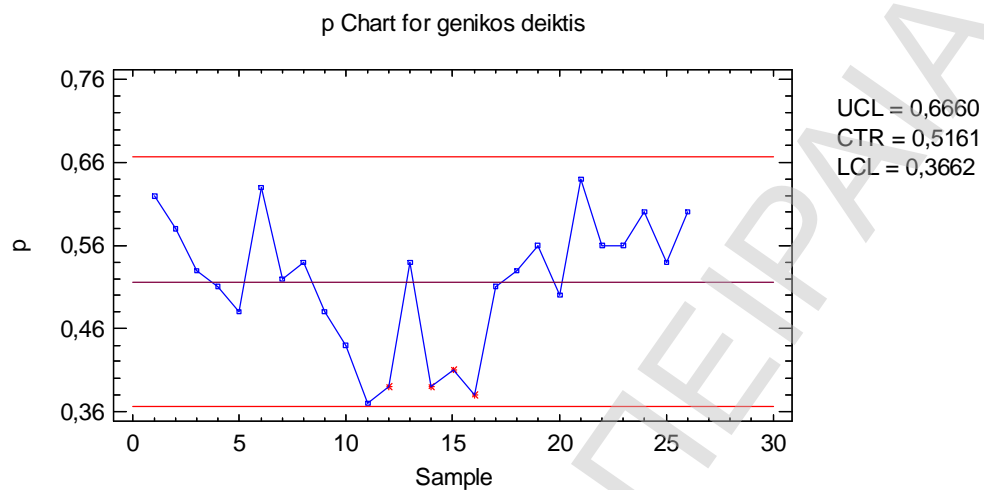
ΔΙΑΓΡΑΜΜΑ 4.2.1

Διάγραμμα χρονοσειράς του γενικού δείκτη του χρηματιστηρίου



ΔΙΑΓΡΑΜΜΑ 4.2.2

p διάγραμμα ελέγχου για τον γενικό δείκτη με μέγεθος δείγματος $n=100$

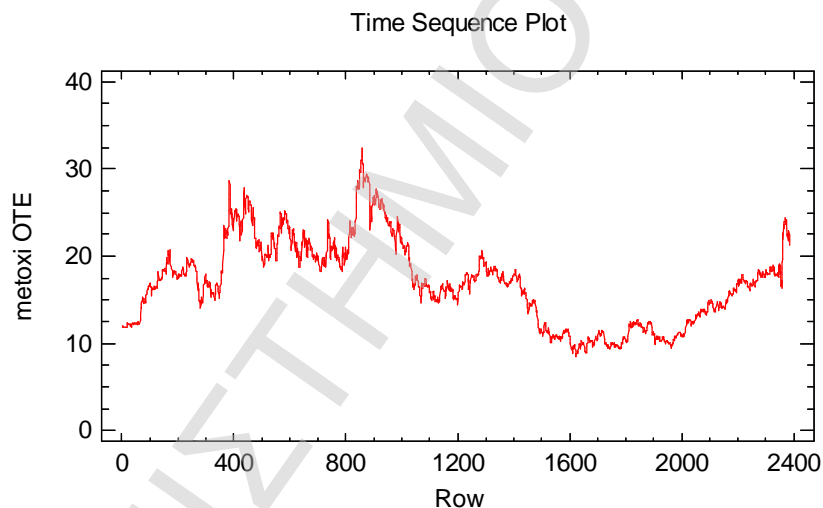


Από το διάγραμμα της χρονοσειράς παρατηρούμε ότι ο γενικός δείκτης παρουσιάζει μια πολύ μεγάλη, απότομη άνοδο η οποία κορυφώνεται τον Σεπτέμβριο του 1999, (η περίοδος αυτή στο διάγραμμα αντιστοιχεί στις τιμές από 0,6 ως 0,75 περίπου). Μετά επέρχεται μια κάθοδος που διαρκεί μέχρι, το τέλος Μαρτίου του 2003, (η περίοδος αυτή στο διάγραμμα αντιστοιχεί στις τιμές από 0,8 ως 1,62 περίπου), για να ακολουθήσει μια περίοδος ομαλής ανόδου, η οποία περιέχει και μικρές περιόδους καθόδου του δείκτη. Φτάνοντας στο τέλος της χρονοσειράς, τους πρώτους μήνες του έτους 2007, δημιουργείται μία δεύτερη κορυφή που βέβαια δεν βρίσκεται σε τόσο υψηλές τιμές όπως το Σεπτέμβρη του 1999. Από το p διάγραμμα ελέγχου παρατηρούμε ότι οι πιθανότητες ανόδου, υπολογιζόμενες με μέγεθος δείγματος $n=100$, βρίσκονται όλες εντός ορίων. Βέβαια την περίοδο του 1999 όπου είχαμε την μεγάλη άνοδο του γενικού δείκτη και την περίοδο μετά τον Μάρτιο του 2003 όπου πάλι εμφανίζεται σχετική άνοδος, παρατηρούμε ότι οι πιθανότητες ανόδου είναι αρκετά πιο υψηλές σε σχέση με την περίοδο της καθόδου του δείκτη. Ακόμα παρατηρούμε ότι, μπορεί να μην έχουμε σημεία εκτός των ορίων, αλλά έχουμε σημεία πολύ κοντά στα όρια ή και πάνω στα όρια. Μια άλλη παρατήρηση που μπορούμε να κάνουμε είναι ότι τα σημεία που αντιπροσωπεύουν τις πιθανότητες μας, δεν εναλλάσσονται συχνά πάνω και κάτω από την κεντρική τιμή. Δηλαδή έχουμε αρκετά συνεχόμενα σημεία πάνω ή κάτω από την κεντρική γραμμή ανάλογα με το αν ο δείκτης διανύει περίοδο ανόδου ή καθόδου. Για αυτούς τους λόγους, συμπεραίνουμε ότι δεν επικρατεί μια σχετική σταθερότητα των πιθανοτήτων ανόδου και έτσι δεν μπορούμε να θεωρήσουμε ότι βρισκόμαστε στην περίπτωση των i.i.d. δοκιμών.

Η δεύτερη χρονοσειρά με την οποία θα ασχοληθούμε είναι η εξέλιξη της μετοχής του ΟΤΕ. Αποτελείται από 2357 παρατηρήσεις που αντιστοιχούν στην εξέλιξη της μετοχής του ΟΤΕ για περίπου 9,5 χρόνια, από τον Οκτώβριο του 1996 ως τον Μάρτιο του 2007. Κάθε στοιχείο της χρονοσειράς αυτής δείχνει και εδώ, αν είχαμε άνοδο ή κάθοδο της τιμής της μετοχής σε κάθε ημέρα λειτουργίας του χρηματιστηρίου. Θα φτιάξουμε και εδώ το διάγραμμα της χρονοσειράς (Time Sequence Plot) και ένα p διάγραμμα ελέγχου (φάσης II με 3σ όρια) παίρνοντας πάλι μέγεθος δείγματος $n=100$. Η εκτιμημένη από όλες της παρατηρήσεις πιθανότητα ανόδου την οποία θέτουμε ως κεντρική τιμή, είναι $\bar{P}=0,4787$. Τα διαγράμματα αυτά είναι τα ακόλουθα.

ΔΙΑΓΡΑΜΜΑ 4.2.3

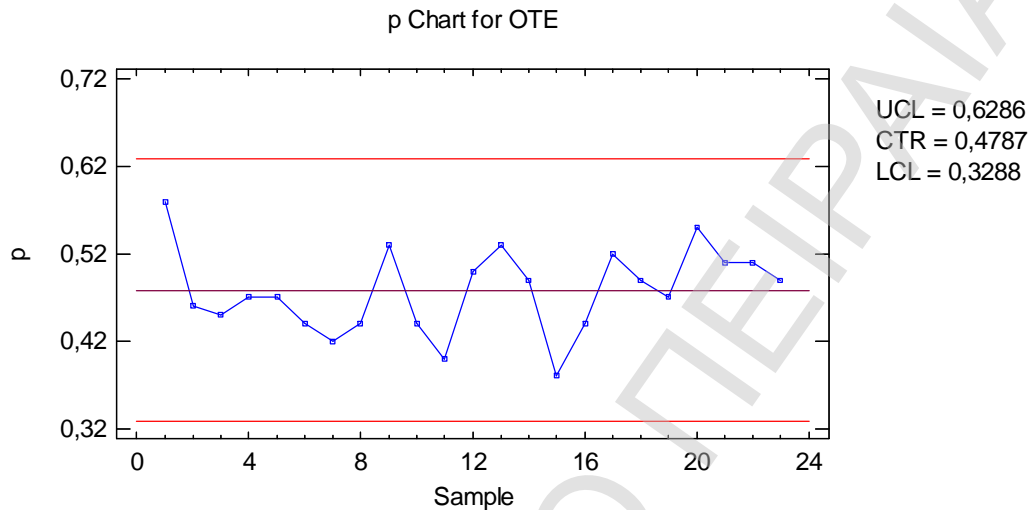
Διάγραμμα χρονοσειράς της μετοχής του ΟΤΕ



ΔΙΑΓΡΑΜΜΑ 4.2.4

p διάγραμμα ελέγχου για την μετοχή του ΟΤΕ με μέγεθος δείγματος

$$n=100$$

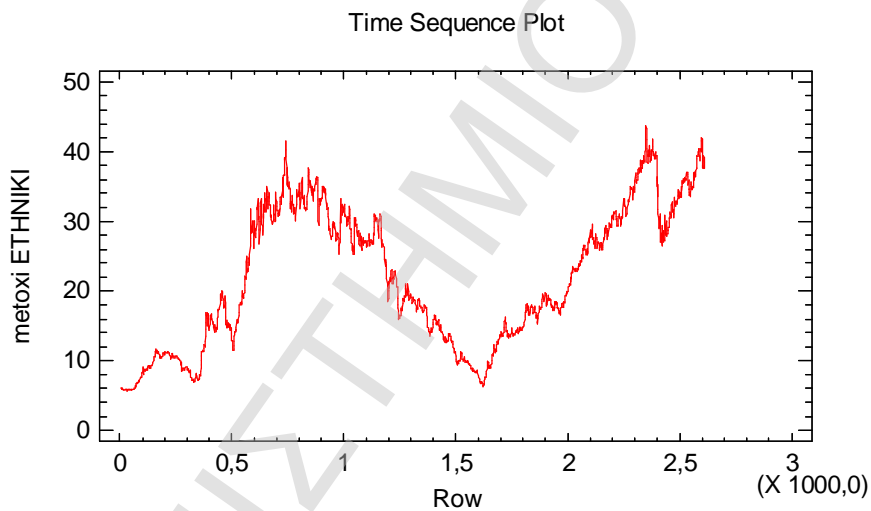


Από το παραπάνω διάγραμμα της χρονοσειράς παρατηρούμε ότι η μετοχή του ΟΤΕ παρουσιάζει μια πιο ομαλή εξέλιξη και δεν συμβαδίζει σε μεγάλο βαθμό με την εξέλιξη του γενικού δείκτη του χρηματιστηρίου. Αυτό συμβαίνει γιατί είναι μια αρκετά σταθερή μετοχή που δεν επηρεάζεται πολύ από την γενική πορεία του χρηματιστηρίου. Παρατηρούμε ότι η εξέλιξη της μετοχής αυτής παρουσιάζει πέντε με έξι κορυφές, με μεγαλύτερη αυτή που εμφανίζεται στους πρώτους μήνες του έτους 2000, (η περίοδος αυτή στο διάγραμμα αντιστοιχεί στις τιμές από 820 ως 840 περίπου). Ακόμα μια περίοδος που η μετοχή είχε πάρει υψηλές τιμές είναι κατά τους μήνες Απρίλιο και Μάιο του 1998, (η περίοδος αυτή στο διάγραμμα αντιστοιχεί στις τιμές από 390 ως 400 περίπου). Μετά την μεγάλη άνοδο στις αρχές του 2000 η μετοχή παρουσιάζει κάθοδο και έπειτα κυμαίνεται, χωρίς μεγάλες αυξομειώσεις, σε πιο χαμηλά επίπεδα για αρκετά μεγάλο διάστημα. Τέλος από τον Οκτώβριο του 2004, (η περίοδος αυτή στο διάγραμμα αντιστοιχεί στις τιμές γύρω από το 1600), η μετοχή αρχίζει να παρουσιάζει μια σταδιακή άνοδο που φτάνει μέχρι το τέλος της χρονοσειράς, δηλαδή τον Μάρτιο του 2007. Από το p διάγραμμα ελέγχου παρατηρούμε ότι οι πιθανότητες ανόδου, υπολογιζόμενες με μέγεθος δείγματος $n=100$, εξελίσσονται αρκετά πιο ομαλά σε σχέση με τον γενικό δείκτη. Βρίσκονται όλες εντός ορίων και κινούνται, εναλλάσσονται γύρω από την κεντρική τιμή. Άρα μπορούμε να θεωρήσουμε ότι στην μετοχή του ΟΤΕ επικρατεί σχετική σταθερότητα στις πιθανότητες ανόδου. Επομένως μπορούμε να την αναλύσουμε σαν περίπτωση i.i.d. δοκιμών.

Η τρίτη χρονοσειρά με την οποία θα ασχοληθούμε είναι η εξέλιξη της μετοχής της Εθνικής. Αποτελείται από 2577 παρατηρήσεις που αντιστοιχούν στην εξέλιξη της μετοχής της Εθνικής για περίπου 9,5 χρόνια, από τον Οκτώβριο του 1996 ως τον Μάρτιο του 2007. Κάθε στοιχείο της χρονοσειράς αυτής, δείχνει και εδώ αν είχαμε άνοδο ή κάθοδο της τιμής της μετοχής σε κάθε ημέρα λειτουργίας του χρηματιστηρίου. Θα φτιάξουμε και εδώ το διάγραμμα της χρονοσειράς (Time Sequence Plot) και ένα p διάγραμμα ελέγχου (φάσης II με 3σ όρια) παίρνοντας πάλι μέγεθος δείγματος $n=100$. Η εκτιμημένη από όλες της παρατηρήσεις πιθανότητα ανόδου την οποία θέτουμε ως κεντρική τιμή, είναι $\bar{P}=0,4994$. Τα διαγράμματα αυτά είναι τα ακόλουθα.

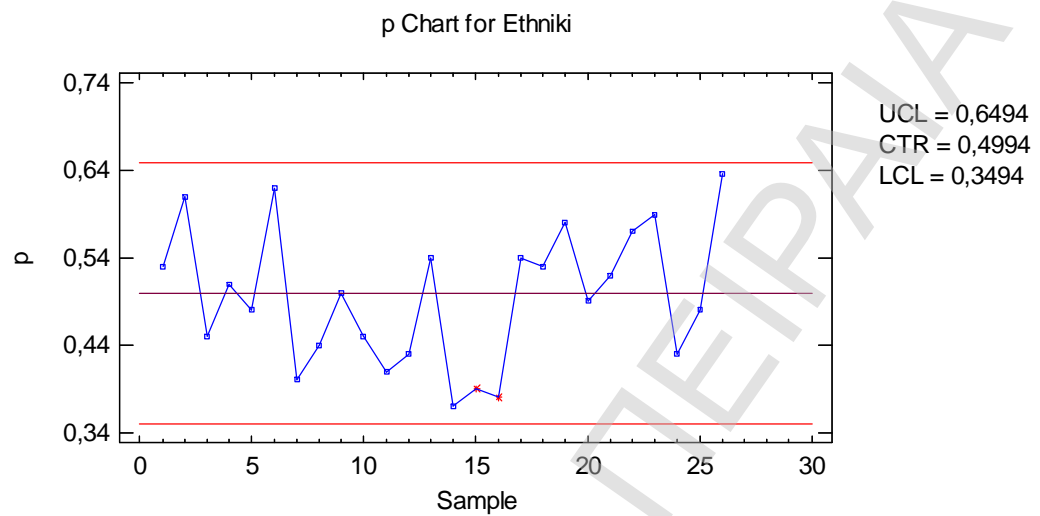
ΔΙΑΓΡΑΜΜΑ 4.2.5

Διάγραμμα χρονοσειράς της μετοχής της Εθνικής



ΔΙΑΓΡΑΜΜΑ 4.2.6

p διάγραμμα ελέγχου για την μετοχή της Εθνικής με μέγεθος δείγματος $n=100$



Από το παραπάνω διάγραμμα της χρονοσειράς (Time Sequence Plot) βλέπουμε ότι υπάρχουν πολύ μεγάλες και έντονες αυξομειώσεις στην τιμή της μετοχής της Εθνικής. Η πορεία της μετοχής δεν είναι ομαλή. Ακόμα παρατηρούμε ότι συμβαδίζει αρκετά με την πορεία του γενικού δείκτη. Στο διάγραμμα μας βλέπουμε ότι υπάρχουν τρεις κορυφές σε αρκετά υψηλές τιμές για την μετοχή. Η πρώτη μεγάλη άνοδος που παρατηρείται κορυφώνεται τον Σεπτέμβριο του 1999, (η περίοδος αυτή στο διάγραμμα αντιστοιχεί στις τιμές από 0,71 ως 0,75 περίπου), όπως συμβαίνει και με τον γενικό δείκτη του χρηματιστηρίου. Έπειτα ακολουθεί μια μεγάλη κάθοδος της τιμής της μετοχής, που παίρνει την ελάχιστη τιμή της στο τέλος Μαρτίου με αρχές Απριλίου του 2003, (η περίοδος αυτή στο διάγραμμα αντιστοιχεί στις τιμές γύρω από το 1,61). Παρόμοια συμβαίνει και με τον γενικό δείκτη. Μετά παρατηρείται μια εξίσου μεγάλη άνοδος η οποία κορυφώνεται τον Φεβρουάριο του 2006, (η περίοδος αυτή στο διάγραμμα αντιστοιχεί στις τιμές από 2,32 ως 2,35 περίπου) και εκεί η τιμή της μετοχής παίρνει την μέγιστη τιμή της σε όλα αυτά τα χρόνια που εξετάζουμε. Ακολουθεί μια μικρή αλλά απότομη κάθοδος που φτάνει μέχρι της αρχές Ιουνίου του 2006, (τιμές στο διάγραμμα γύρω από το 1,61), και έπειτα έχουμε συνεχή άνοδο, με την τιμή της μετοχής να παίρνει πάλι αρκετά μεγάλες τιμές, μέχρι το τέλος της χρονοσειράς μας, στα μέσα Μαρτίου του 2007. Από το p διάγραμμα ελέγχου παρατηρούμε ότι οι πιθανότητες ανόδου, υπολογιζόμενες και εδώ με μέγεθος δείγματος $n=100$, είναι μεν εντός ορίων αλλά έχουμε πιο έντονες αυξομειώσεις σε σχέση με την μετοχή του ΟΤΕ. Όπως και στο p διάγραμμα ελέγχου του γενικού δείκτη, έτσι και εδώ έχουμε πολλά σημεία τα οποία είναι κοντά στα όρια. Ακόμα

τα σημεία αυτά βλέπουμε ότι δεν εναλλάσσονται συχνά γύρω από την κεντρική τιμή, αλλά έχουμε πολλά συνεχόμενα σημεία πάνω ή κάτω από την κεντρική γραμμή. Έτσι συμπεραίνουμε ότι και στην μετοχή της Εθνικής, δεν επικρατεί μια σχετική σταθερότητα των πιθανοτήτων ανόδου και δεν μπορούμε να θεωρήσουμε ότι βρισκόμαστε στην περίπτωση των i.i.d. δοκιμών.

Επομένως η μετοχή στην οποία θα εφαρμόσουμε το μοντέλο της δεσμευμένης πιθανότητας στην περίπτωση i.i.d. δοκιμών, θα είναι η μετοχή του ΟΤΕ.

4.3 Επιλογή δεδομένων και σύγκριση θεωρητικών και πειραματικών πιθανοτήτων

Στην παράγραφο αυτή θα επιλέξουμε κάποιες υποπεριόδους, από την χρονοσειρά της μετοχής του ΟΤΕ, οι οποίες να έχουν κάποιες συγκεκριμένες πιθανότητες ανόδου. Έπειτα θα εφαρμόσουμε το θεωρητικό μοντέλο πιθανοτήτων με αυτές τις πιθανότητες ανόδου για να κάνουμε αργότερα τη σύγκριση ανάμεσα στα θεωρητικά και στα πειραματικά αποτελέσματα. Επιλέγουμε περιοχές από την χρονοσειρά του ΟΤΕ που να δίνουν πιθανότητες ανόδου $p=0,40$, $p=0,60$ και $p=0,80$. Οι περιοχές αυτές, για τις τρεις διαφορετικές πιθανότητες, δίνονται στον παρακάτω πίνακα.

ΠΙΝΑΚΑΣ 4.3.1

Επιλεγμένες από τις χρονοσειρά του ΟΤΕ περιοχές με πιθανότητες ανόδου $p=0,40$, $p=0,60$ και $p=0,80$

Περιοχές επιλεγμένες από τις χρονοσειρές για πιθανότητες ανόδου p	Μετοχή ΟΤΕ
$p=0,40$	11/12/2001 - 1/03/2001
$p=0,60$	9/12/1996 - 12/03/1997
$p=0,80$	31/12/1996 - 12/03/1997

Τώρα με αυτές τις πιθανότητες ανόδου θα εφαρμόσουμε το μοντέλο πιθανοτήτων για τρεις συνδιασμούς του χρονικού οριζόντα y και του μήκους ροής k για να πάρουμε τα θεωρητικά αποτελέσματα για την δεσμευμένη πιθανότητα $f_k(x/y)$ για συγκεκριμένες χρονικές στιγμές x . Πρώτα έχουμε τον συνδιασμό $y=2$, $k=3$ και θα υπολογίσουμε την $f_k(x/y)$ για τις χρονικές στιγμές $x=4$, $x=5$, $x=6$, $x=7$ και $x=8$. Παρακάτω δίνεται διάγραμμα της $f_k(x/y)$ και πίνακας

με τις ακριβείς τιμές της. Για τον σχεδιασμό των παρακάτω διαγραμμάτων χρησιμοποιήσαμε τον ακόλουθο αλγόριθμο στο Mathematica, αλλάζοντας τις τιμές του y και του k για την κάθε περίπτωση:

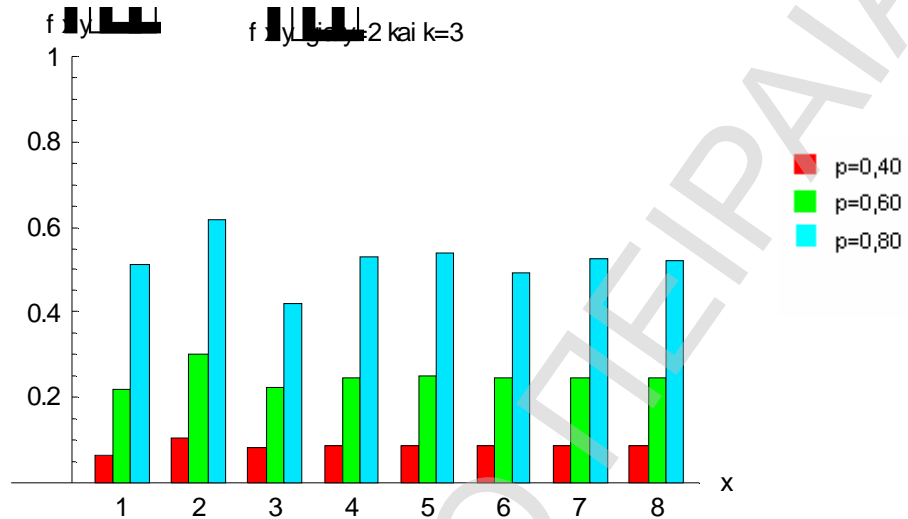
```
y=2;
k=3;
"x>=k+1"
f[x_,p_]:=Sum[(((1-p)^i)*(p^(x-i)))*Sum[((-1)^j)*Binomial[i,j]*Binomial[x-k*(j+1)-1,i-1],{j,0,IntegerPart[(x-i-k)/k]}],{i,1,x-k}];
g[x_,p_]:=UnitStep[x-k]*(p^k)+UnitStep[x-(k+1)]*(Sum[f[j,p],{j,k+1,x}]);
h[x_,p_]:= (g[x+y,p]-g[x,p])/(1-g[x,p]);
c=5;
list1={};
list2={};
list3={};
Do[list1=Append[list1,h[x,0.40]],{x,1,k+c}];
Do[list2=Append[list2,h[x,0.60]],{x,1,k+c}];
Do[list3=Append[list3,h[x,0.80]],{x,1,k+c}];
<<Graphics`Graphics`
BarChart[list1,list2,list3,BarSpacing->0,BarGroupSpacing->0.3,BarStyle->{RGBColor[1,0,0],RGBColor[0,1,0],RGBColor[0,0.9,1]},PlotRange->{0,1},PlotLabel->"f(x/y) για y=2 και k=3",AxesLabel->{"x","f(x/y)"},TextStyle->{"FontFamily"->"TimesNewRoman"}]
```

Ακόμα, για να υπολογίσουμε τις ακριβείς τιμές της $f_k(x/y)$ χρησιμοποιήσαμε τον ακόλουθο αλγόριθμο στο Mathematica, αλλάζοντας τις τιμές του y , του k , του p και του x για την κάθε περίπτωση:

```
p=0.40;
x=4;
y=2;
k=3;
"x>=k+1"
f[x_,p_]:=Sum[(((1-p)^i)*(p^(x-i)))*Sum[((-1)^j)*Binomial[i,j]*Binomial[x-k*(j+1)-1,i-1],{j,0,IntegerPart[(x-i-k)/k]}],{i,1,x-k}];
g[x_,p_]:=UnitStep[x-k]*(p^k)+UnitStep[x-(k+1)]*(Sum[f[j,p],{j,k+1,x}]);
h[x_,p_]:= (g[x+y,p]-g[x,p])/(1-g[x,p]);
h[x,p]
```

ΔΙΑΓΡΑΜΜΑ 4.3.1

Δεσμευμένη πιθανότητα $f_k(x/y)$ για $y=2$ και $k=3$



ΠΙΝΑΚΑΣ 4.3.2

Ακριβείς τιμές της $f_k(x/y)$ για $y=2$ και $k=3$

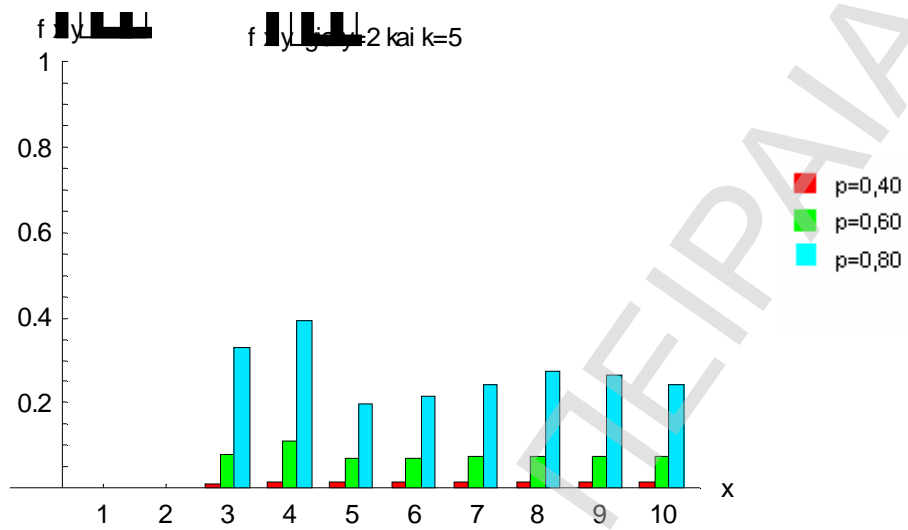
$f_k(x/y)$ για τον συνδιασμό $y=2, k=3$	$x=4$	$x=5$	$x=6$	$x=7$	$x=8$
$p=0,40$	0.0855615	0.0865251	0.0857825	0.0859533	0.0859713
$p=0,60$	0.247706	0.252188	0.243922	0.247407	0.247361
$p=0,80$	0.53112	0.538034	0.494782	0.523471	0.520165

Δεύτερος συνδιασμός που θα χρησιμοποιήσουμε είναι αυτός με χρονικό ορίζοντα $y=2$ και $k=5$ και θα υπολογίσουμε την $f_k(x/y)$ για τις χρονικές στιγμές $x=6, x=7, x=8, x=9$ και $x=10$.

Παρακάτω δίνεται διάγραμμα της $f_k(x/y)$ και πίνακας με τις ακριβείς τιμές της.

ΔΙΑΓΡΑΜΜΑ 4.3.2

Δεσμευμένη πιθανότητα $f_k(x/y)$ για $y=2$ και $k=5$



ΠΙΝΑΚΑΣ 4.3.3

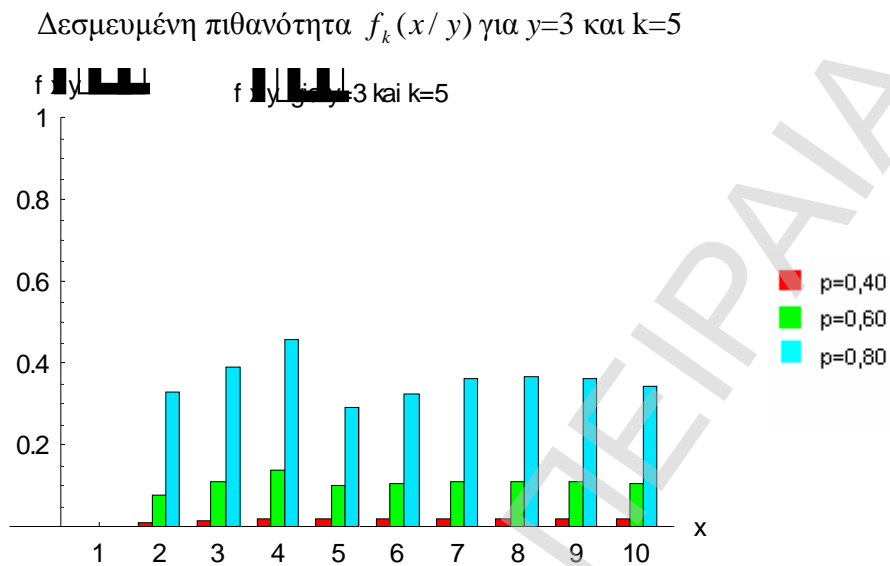
Ακριβείς τιμές της $f_k(x/y)$ για $y=2$ και $k=5$

$f_k(x/y)$ για τον συνδιασμό $y=2, k=5$	$x=6$	$x=7$	$x=8$	$x=9$	$x=10$
$p=0,40$	0.0124927	0.0125712	0.0126507	0.0126661	0.0126422
$p=0,60$	0.0698075	0.0723322	0.0750463	0.0749405	0.0735643
$p=0,80$	0.216011	0.242166	0.275528	0.267195	0.243232

Τελευταίος συνδιασμός που θα χρησιμοποιήσουμε είναι αυτός με χρονικό ορίζοντα $y=3$ και $k=5$ και θα υπολογίσουμε την $f_k(x/y)$ για τις χρονικές στιγμές $x=6, x=7, x=8, x=9$ και $x=10$.

Παρακάτω δίνεται διάγραμμα της $f_k(x/y)$ και πίνακας με τις ακριβείς τιμές της.

ΔΙΑΓΡΑΜΜΑ 4.3.3



ΠΙΝΑΚΑΣ 4.3.4

Ακριβείς τιμές της $f_k(x/y)$ για $y=3$ και $k=5$

$f_k(x/y)$ για τον συνδιασμό $y=3, k=5$	$x=6$	$x=7$	$x=8$	$x=9$	$x=10$
$p=0,40$	0.018739	0.0188568	0.0189113	0.0189274	0.0189043
$p=0,60$	0.104711	0.108498	0.109652	0.109682	0.108454
$p=0,80$	0.324016	0.363249	0.36815	0.364145	0.346154

Από τα παραπάνω διαγράμματα και τους παραπάνω πίνακες μπορούμε να δούμε σε ποια χρονική στιγμή μεγιστοποιείται η πιθανότητα να εμφανιστεί ροή μήκους k ανόδων, δηλαδή ποια χρονική στιγμή έχουμε τη μέγιστη πιθανότητα για κέρδος σε κάθε περίπτωση. Τα αποτελέσματα που παρατηρούνται είναι παρόμοια με αυτά της παραγράφου 3.2 με μόνη διαφορά ότι εδώ έχουμε άλλες τιμές για τις πιθανότητες ανόδου.

Τώρα θα επιλέξουμε μια από τις πιθανότητες p και θα υπολογίσουμε τις πειραματικές τιμές της $f_k(x/y)$ από την χρονοσειρά του ΟΤΕ. Παίρνουμε πιθανότητα $p=0,80$ και η επιλεγμένη περιοχή από την χρονοσειρά μας η οποία δίνει τέτοια πιθανότητα ανόδου είναι από 31/12/1996 ως 2/03/1997, όπως είδαμε στον πίνακα 4.3.1. Για τον υπολογισμό των τιμών της $f_k(x/y)$ από αυτή την περιοχή, χρησιμοποιήσαμε τον ακόλουθο αλγόριθμο στο Mathematica, αλλάζοντας τις τιμές του y , του k και του x για την κάθε περίπτωση:

```

l={1,1,1,1,1,1,0,1,1,1,0,1,1,0,0,0,1,1,1,1,1,0,0,1,1,1,1,1,1,1
};
le=Length[l]
x=4;
y=2;
k=3;
c1=0;
c2=0;
Do[If[Sum[Product[l[[j]],{j,i,i+k-1}],{i,n-x+1,n-
k+1}]==0,(c1=c1+1)&&If[Sum[Product[l[[j]],{j,i,i+k-1}],{i,n-
x+1,n+y-k+1}]==0,c2=c2,c2=c2+1],c1=c1],{n,x,le-y}]
"ΟΤΕ periodos me p=0,80"
c1
c2
N[c2/c1]
    
```

Οι ακριβείς τιμές της $f_k(x/y)$ υπολογιζόμενες από την περίοδο της μετοχής του ΟΤΕ που δίνει πιθανότητα ανόδου $p=0,80$, όπως αναφέραμε παραπάνω, δίνονται στον ακόλουθο πίνακα:

ΠΙΝΑΚΑΣ 4.3.5

Πειραματικές τιμές της $f_k(x/y)$ για την μετοχή του ΟΤΕ σε περίοδο με $p=0,80$

Μετοχή ΟΤΕ, Περίοδος: 31/12/1996 - 12/02/1997 με $p=0,80$	$x=4$	$x=5$	$x=6$	$x=7$	$x=8$
$y=2, k=3$	0.5	0.555556	0.5	0.5	0.666667
Μετοχή ΟΤΕ, Περίοδος: 31/12/1996 - 12/02/1997 με $p=0,80$	$x=6$	$x=7$	$x=8$	$x=9$	$x=10$
$y=2, k=5$	0.222222	0.25	0.285714	0.333333	0.3
$y=3, k=5$	0.333333	0.375	0.428571	0.416667	0.4

Για να μπορούμε να κάνουμε καλύτερα την σύγκριση ανάμεσα στα θεωρητικά και στα πειραματικά αποτελέσματα, θα φτιάξουμε για τον κάθε συνδιασμό συγκριτικούς πίνακες. Οι τιμές πάνω στους πίνακες που είναι με έντονη γραφή (Bold) δείχνουν που η $f_k(x/y)$ παίρνει την μέγιστη τιμή της.

ΠΙΝΑΚΑΣ 4.3.6

Θεωρητικές και πειραματικές τιμές της $f_k(x/y)$ για $y=2$, $k=3$ και $p=0,80$

$f_k(x/y)$ για τον συνδυασμό $y=2$, $k=3$	$x=4$	$x=5$	$x=6$	$x=7$	$x=8$
Θεωρητικές τιμές με $p=0,80$	0.53112	0.538034	0.494782	0.523471	0.520165
Πειραματικές τιμές, Μετοχή ΟΤΕ, Περίοδος: 31/12/1996 - 12/02/1997 με $p=0,80$	0.5	0.555556	0.5	0.5	0.666667

Από τον παραπάνω πίνακα, που αντιστοιχεί στον συνδυασμό με χρονικό ορίζοντα $y=2$ και μήκος ροής $k=3$, παρατηρούμε ότι οι θεωρητικές τιμές είναι αρκετά κοντά με τις πειραματικές, με αποκλίσεις από 0,005 έως 0,03 περίπου όσον αφορά τα $x=4, 5, 6, 7$. Στην περίπτωση όπου έχουμε $x=8$, η απόκλιση μεγαλώνει και φτάνει το ύψος του 0,14 περίπου. Ακόμα παρατηρούμε ότι η $f_k(x/y)$ και στα θεωρητικά και στα πειραματικά αποτελέσματα παρουσιάζει μέγιστο στην ίδια χρονική στιγμή $x=5$. Δηλαδή συμπίπτουν τα αποτελέσματα και στις δύο περιπτώσεις και μας δείχνουν ότι για τον συγκεκριμένο συνδυασμό των y και k , η χρονική στιγμή όπου έχουμε την μέγιστη πιθανότητα για κέρδος και στην οποία πρέπει να επενδύσουμε, είναι η χρονική στιγμή $x=5$.

ΠΙΝΑΚΑΣ 4.3.7

Θεωρητικές και πειραματικές τιμές της $f_k(x/y)$ για $y=2$, $k=5$ και $p=0,80$

$f_k(x/y)$ για τον συνδυασμό $y=2$, $k=5$	$x=6$	$x=7$	$x=8$	$x=9$	$x=10$
Θεωρητικές τιμές με $p=0,80$	0.216011	0.242166	0.275528	0.267195	0.243232
Πειραματικές τιμές, Μετοχή ΟΤΕ, Περίοδος: 31/12/1996 - 12/02/1997 με $p=0,80$	0.222222	0.25	0.285714	0.333333	0.3

Από τον παραπάνω πίνακα, που αντιστοιχεί στον συνδυασμό με χρονικό ορίζοντα $y=2$ και μήκος ροής $k=5$, παρατηρούμε ότι οι θεωρητικές τιμές είναι αρκετά κοντά με τις πειραματικές, με αποκλίσεις από 0,006 έως 0,01 περίπου όσον αφορά τα $x=6, 7, 8$. Στην περίπτωση όπου έχουμε $x=9$ και $x=10$, η απόκλιση μεγαλώνει λίγο και φτάνει το ύψος του 0,06 περίπου. Όπως βλέπουμε στον συνδυασμό αυτό, οι αποκλίσεις είναι μικρότερες από τον προηγούμενο συνδυασμό ($y=2$ και $k=3$), και εδώ η μέγιστη απόκλιση φτάνει το ύψος του 0,06. Στην περίπτωση αυτή βέβαια βλέπουμε ότι η μέγιστη πιθανότητα τιμή της $f_k(x/y)$ δεν παρουσιάζεται την ίδια χρονική στιγμή στα θεωρητικά και στα πειραματικά αποτελέσματα. Παρουσιάζεται με μια χρονική στιγμή, μια ημέρα παρακολούθησης, διαφορά. Τα θεωρητικά αποτελέσματα δείχνουν για την μέγιστη πιθανότητα για κέρδος την χρονική στιγμή $x=8$, ενώ τα πειραματικά αποτελέσματα δείχνουν την χρονική στιγμή $x=9$. Μπορούμε να θεωρήσουμε ότι η διαφορά αυτή δεν είναι αρκετά σημαντική.

ΠΙΝΑΚΑΣ 4.3.8

Θεωρητικές και πειραματικές τιμές της $f_k(x/y)$ για $y=3$, $k=5$ και $p=0,80$

$f_k(x/y)$ για τον συνδυασμό $y=3$, $k=5$	$x=6$	$x=7$	$x=8$	$x=9$	$x=10$
Θεωρητικές τιμές με $p=0,80$	0.324016	0.363249	0.36815	0.364145	0.346154
Πειραματικές τιμές, Μετοχή ΟΤΕ, Περίοδος: 31/12/1996 - 12/02/1997 με $p=0,80$	0.333333	0.375	0.428571	0.416667	0.4

Από τον παραπάνω πίνακα, που αντιστοιχεί στον συνδυασμό με χρονικό ορίζοντα $y=3$ και μήκος ροής $k=5$, παρατηρούμε ότι οι θεωρητικές τιμές είναι και εδώ αρκετά κοντά με τις πειραματικές, με αποκλίσεις από 0,009 έως 0,06 περίπου. Στην περίπτωση αυτή, όπως και στην περίπτωση του πρώτου συνδυασμού ($y=2$ και $k=3$) που εξετάσαμε, η $f_k(x/y)$ και στα θεωρητικά και στα πειραματικά αποτελέσματα παρουσιάζει μέγιστο στην ίδια χρονική στιγμή $x=8$. Δηλαδή συμπίπτουν τα αποτελέσματα και στις δύο περιπτώσεις και μας δείχνουν ότι για τον συγκεκριμένο συνδυασμό των y και k , η χρονική στιγμή όπου έχουμε την μέγιστη πιθανότητα για κέρδος και στην οποία πρέπει να επενδύσουμε, είναι η χρονική στιγμή $x=8$.

Για να μπορούμε να δούμε συγκεντρωτικά τα παραπάνω, φτιάχνουμε έναν πίνακα με τις αποκλίσεις ανάμεσα στις θεωρητικές και τις πειραματικές τιμές της $f_k(x/y)$.

ΠΙΝΑΚΑΣ 4.3.9

Αποκλίσεις θεωρητικών και πειραματικών τιμών της $f_k(x/y)$ με $p=0,80$

Αποκλίσεις θεωρητικών και πειραματικών τιμών της $f_k(x/y)$	$x=4$	$x=5$	$x=6$	$x=7$	$x=8$
$y=2, k=3$	0,03112	0,017522	0,005218	0,023471	0,146502
Αποκλίσεις θεωρητικών και πειραματικών τιμών της $f_k(x/y)$	$x=6$	$x=7$	$x=8$	$x=9$	$x=10$
$y=2, k=5$	0,006211	0,00783	0,01019	0,06614	0,05677
$y=3, k=5$	0,009317	0,01175	0,06042	0,05252	0,05385

Από τον παραπάνω συγκεντρωτικό πίνακα παρατηρούμε ότι η ελάχιστη απόκλιση είναι 0,005218 και η μέγιστη 0,146502. Η απόκλιση 0,146502 όμως, είναι μια μεμονωμένη περίπτωση γιατί όλες οι άλλες αποκλίσεις κυμαίνονται από 0,005218 έως 0,06614. Το γεγονός ότι υπάρχουν αποκλίσεις, όχι μεγάλες βέβαια, οφείλεται στο ότι στην μετοχή του ΟΤΕ, που εξετάσαμε, δεν επικρατεί κατάσταση απόλυτης σταθερότητας αλλά σχετικής σταθερότητας. Δηλαδή προσεγγίζει την περίπτωση των i.i.d. δοκιμών, γ' αυτό δικαιολογείται μια μικρή απόκλιση από τα θεωρητικά αποτελέσματα.

Από όλα τα παραπάνω, ερχόμαστε στο τελικό συμπέρασμα, ότι μετά την εφαρμογή του μοντέλου της δεσμευμένης πιθανότητας σε πραγματικά δεδομένα του χρηματιστηρίου και την σύγκρισή της με τα θεωρητικά αποτελέσματα, το μοντέλο αυτό είναι επιτυχές. Δηλαδή χρησιμοποιώντας τις ροές και το μοντέλο της δεσμευμένης πιθανότητας $f_k(x/y)$, μπορούμε να προβλέψουμε με σχετική επιτυχία την χρονική στιγμή που αναμένεται η μεγαλύτερη πιθανότητα για κέρδος, οπότε να προβλέψουμε την κατάλληλη χρονική στιγμή για επένδυση πάνω στην μετοχή που παρακολουθούμε.

Βιβλιογραφία

- Aki, S. and Hirano, K. (1993). Discrete distributions related to succession events in a two-state Markov chain, In *Statistical Science and Data Analysis* (Eds., K. Matusita, M.L. Puri and T. Hayakawa), pp.467-474, VSP International Science Publishers, Zeist.
- Aki, S., Kuboki H. and Hirano, K. (1984). On discrete distributions of order k , *Annals of Institute of Statistical Mathematics*, **36**, 431-440.
- Balakrishnan, N. and Koutras, M.V. (2002). *Runs and Scans with Applications* Wiley Series in probability and Statistics, John Wiley & Sons, New York
- Barry, M. J. J. and Lo Bello, A. J. (1993). The moment generating function of the geometric distribution of order k , *The Fibonacci Quarterly*, **31**, 178-180.
- Chao, M. T, and Fu, J. C. (1989). A limit theorem of certain repairable systems, *Annals of the Institute of Statistical Mathematics*, **41**, 809-818.
- Chao, M. T. and Fu, J. C. (1991). The reliability of large series system under o Markovian structure, *Advances in Applied Probability*, **23**, 894-908.
- Chao, M. T. and Fu, J. C. and Koutras, M. V. (1995). A survey of the reliability studies of consecutive- k -out-of- n : F systems and its related systems, *IEEE Transactions on Reliability*, **44**, 120-127.
- Chiang, D. T. and Niu, S. (1981). Reliability of consecutive- k -out-of- n : F system, *IEEE Transactions on Reliability*, **30**, 87-89.
- Cohen, J. P. and Menjoge, S. S. (1988). One-sample run tests of symmetry, *Journal of Statistical Planning and Inference*, **18**, 93-100.
- de Forest, E. L. (1976). *Interpolation and Adjustment of Series*, Tuttle, Morehouse and Taylor, New Haven.
- de Moivre, A. (1738). *The Doctrine of Chance*, Chelsea Publishing Co., Third edition, New York.
- Feder, P. I. (1974). Problem solving: Markov Chain method, *Industrial Engineering*, **6**, 23-25.
- Feller, W. (1968). *An Introduction to Probability Theory and its Applications*, Vol. I, Third edition, John Wiley & Sons, New York.
- Fu, J. C. and Lou, W. Y. W. (1991). On reliabilities of certain large linearly connected engineering systems, *Statistics & Probability Letters*, **12**, 291-296.
- Hahn, G. J. and Gage, J. B. (1983). Evaluation of a start-up demonstration test, *Journal of Quality Technology*, **15**, 103-105.

- Henze, N. (1993). On the consistency of a test for symmetry based on a runs statistic. *Nonparametric Statistics*, **3**, 195-199.
- Hirano, K. and Aki, S. (1993). On number of occurrences of success runs of specified length in a two-state Markov chain, *Statistic Sinica*, **3**, 313-320.
- Johnson, N. L., Kotz, S. and Kemp, A. W. (1992). *Univariate Discrete Distributions*, Second edition, John Wiley & Sons, New York.
- Kemp, A. W. (1997). Generalizations of classical discrete distributions, In *Encyclopedia of Statistical Sciences*, **1** (updated), (Eds., S. Kotz, C.B. Read and D.L. Banks), pp. 93-110, John Wiley & Sons, New York.
- Koutras, M. V. (1996a). On a Markov chain approach for the study of reliability structures, *Journal of Applied Probability* **33**, 357-367.
- Koutras, M. V. (1997a). Waiting time distributions associated with runs of fixed length in two-state Markov chains, *Annals of the Institute of Statistical Mathematics*, **49**, 123-139.
- Koutras, M. V. (1997b). Waiting times and numbers of appearances of events in a sequence of discrete random variables, In *Advances in Combinatorial Methods and Applications to Probability and Statistics* (Ed., N. Balakrishnan), pp. 363-384, Birkhäuser, Boston.
- Laplace, P. S. de (1812). *Théorie Analytique des Probabilités*, Paris. Second edition, 1814; Third edition, 1820; Reprinted in *Oeuvres*, **7**, 1886.
- Laurencelle, L. (1987). La loi des succès consécutifs dans un processus de Bernouli, *Lettre Statistique*, **8**, 23-47.
- McWilliams, T. P. (1990). A distribution-free test for symmetry based on a run statistic, *Journal of the American Statistical Association*, **85**, 1130-1133.
- Modarres, R. and Gastwirth, J. L. (1996). A modified runs test for symmetry, *Statistics & Probability Letters*, **31**, 107-112.
- Modarres, R. and Gastwirth, J. L. (1998). Hybrid test for the hypothesis of symmetry, *Journal of Applied Statistics*, **25**, 777-783.
- Muselli, M. (1996). Simple expressions for success run distributions in Bernouli trials, *Statistics & Probability Letters*, **31**, 121-128.
- Philippou, A. N., Georgiou, C. and Philippou, G. N. (1995). Fibonacci-type polynomials of order k with probability applications, *The Fibonacci Quarterly*, **23**, 100-105.
- Philippou, A. N. and Makri, F. S. (1985). Longest success runs and Fibonacci-type polynomials, *The Fibonacci Quarterly*, **23**, 338-346.
- Philippou, A. N. and Makri, F. S. (1986). Success runs and longest runs, *Statistics & Probability Letters*, **4**, 211-215.

- Philippou, A. N. and Muwafi, A. A. (1982). Waiting for the k -th consecutive success and fibonacci sequence of order k , *The Fibonacci Quarterly*, **20**, 28-32.
- Rajarshi, M. B. (1974). Success runs in two-state Markov chain, *Journal of Applied Probability*, **11**, 190-194.
- Simpson, T. (1740). The natural and Laws of Chance. *The Whole after a new, general, and conspicuous Manner, and illustrated with a great Variety of Examples*, Cave Publishers, London. Reprinted, 1792.
- Tobhunter, I. (1865). *A History of the Mathematical Theory of Probability from the Time of Pascal to that of Laplace*, Macmillan, London. Reprinted by Chelsea Publishing Company, New York, 1949.
- Uppuluri, V. R. R. and Patil, S. A. (1983). Waiting times and generalized Fibonacci sequences, *The Fibonacci Quarterly*, **21**, 242-249.
- Wald, A. and Wolfowitz, J. (1940). On a test whether two samples are from the same population, *Annals of Mathematical Statistics*, **11**, 147-162.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ