

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ
ΣΧΟΛΗ ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΗΣ ΚΑΙ ΣΤΑΤΙΣΤΙΚΗΣ



ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ ΚΑΙ ΑΣΦΑΛΙΣΤΙΚΗΣ
ΕΠΙΣΤΗΜΗΣ

ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ ΣΤΗΝ
ΑΝΑΛΟΓΙΣΤΙΚΗ ΕΠΙΣΤΗΜΗ ΚΑΙ ΤΗΝ ΔΙΑΧΕΙΡΙΣΗ
ΚΙΝΔΥΝΩΝ

Μεικτές Κατανομές Poisson

Κωστόπουλος Βασίλειος

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 Αναλογιστική Επιστήμη και τη Διαχείριση Κινδύνων.

Πειραιάς
Δεκέμβριος 2023

Η παρούσα Διπλωματική Εργασία εγκρίθηκε ομόφωνα από την Τριμελή Εξεταστική Επιτροπή που ορίστηκε από τη Συνέλευση του Τμήματος Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς στην υπ' αριθμ. συνεδρίασή της, σύμφωνα με τον Εσωτερικό Κανονισμό Λειτουργίας του Προγράμματος Μεταπτυχιακών Σπουδών στην Αναλογιστική Επιστήμη και Διαχείριση Κινδύνων.

Τα μέλη της Επιτροπής ήταν:

- Αντζουλάκος Δημήτριος (Αναπληρωτής Καθηγητής) (Επιβλέπων)
- Σεβρόγλου Βασίλειος (Καθηγητής)
- Ξένος Παναγιώτης (Επίκουρος Καθηγητής)

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

UNIVERSITY OF PIRAEUS
SCHOOL OF FINANCE AND STATISTICS



**DEPARTMENT OF STATISTICS AND INSURANCE
SCIENCE**

**POSTGRADUATE PROGRAM IN
ACTUARIAL SCIENCE AND RISK MANAGEMENT**

Mixed Poisson Distributions

Kostopoulos Vasileios

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 Actuarial Science and Risk Management.

Piraeus

December 2023

*Αφιερωμένη στην μνήμη της μητέρας μου, Ευσταθίας
Στον πατέρα μου, Νικόλαο
Στην οικογένεια του οικογενειακού μας φίλου, Παπαδάκη Μιχάλη*

ΕΥΧΑΡΙΣΤΙΕΣ

Με την ολοκλήρωση της παρούσας διπλωματικής εργασίας θα ήθελα ιδιαίτερα να ευχαριστήσω τον επιβλέποντα καθηγητή μου κ. Αντζουλάκο Δημήτριο, για την αμέριστη βοήθεια του κατά την εκπόνηση της διπλωματικής μου εργασίας, καθώς και τους καθηγητές της τριμελούς επιτροπής μου κ. Βασίλειο Σεβρόγλου και κ. Παναγιώτη Ξένο. Επίσης θα ήθελα να ευχαριστήσω όλους τους διδάσκοντες δηλαδή τα μέλη ΔΕΠ και τους εξωτερικούς συνεργάτες, για τις πολύτιμες γνώσεις που λάβαμε καθ' όλη την διάρκεια του μεταπτυχιακού προγράμματος.

Τέλος, οφείλω ένα μεγάλο ευχαριστώ στον πατέρα μου που με στήριξε καθ' όλη την διάρκεια των σπουδών μου στο μεταπτυχιακό πρόγραμμα.

ΠΕΡΙΛΗΨΗ

Στην παρούσα διπλωματική εργασία εξετάζουμε την οικογένεια των μεικτών κατανομών Poisson και εφαρμογές της στην αναλογιστική επιστήμη. Η εργασία αποτελείται από τρία κεφάλαια. Στο πρώτο κεφάλαιο γίνεται μια εισαγωγή στην μείξη κατανομών και σε ορισμένα βασικά χαρακτηριστικά τους. Επίσης δίνονται τα κυριότερα μοντέλα διακριτών και συνεχών κατανομών τα οποία θα μας απασχολήσουν στα επόμενα κεφάλαια. Στο δεύτερο κεφάλαιο παρουσιάζουμε τις μεικτές κατανομές Poisson, δίνουμε μια σειρά από χαρακτηριστικές ιδιότητές τους καθώς επίσης και μια αναλυτική παρουσίαση βασικών υποδειγμάτων μεικτών κατανομών Poisson. Τέλος, στο τρίτο κεφάλαιο, με τη βοήθεια της γλώσσας προγραμματισμού R, γίνεται προσαρμογή μερικών εκ των μεικτών κατανομών Poisson πάνω σε πραγματικά δεδομένα για την εξαγωγή συμπερασμάτων.

ABSTRACT

In this diploma thesis, we examine the family of mixed Poisson distributions and its applications in actuarial science. The dissertation consists of three chapters. The first chapter introduces the mixture of distributions and some of their basic characteristics. Also we give the main models of discrete and continuous distributions which will concern us in the following chapters. In the second chapter we present mixed Poisson distributions, give a series of their characteristic properties as well as a detailed presentation of basic models of mixed Poisson distributions. Finally, in the third chapter, with the help of the R programming language, some of the mixed Poisson distributions are fitted on real data and draw certain conclusions.

ΠΕΡΙΕΧΟΜΕΝΑ

ΠΕΡΙΛΗΨΗ.....	i
ABSTRACT.....	ii
ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ	v
ΚΑΤΑΛΟΓΟΣ ΣΧΗΜΑΤΩΝ	vii
ΚΕΦΑΛΑΙΟ 1	1
Μείξεις κατανομών.....	1
1.1 Εισαγωγή.....	1
1.2 Μείξη κατανομών	2
1.3 Πρότυπα διακριτών κατανομών	6
1.3.1 Διωνυμική κατανομή.....	6
1.3.2 Κατανομή Poisson	7
1.3.3 Αρνητική διωνυμική κατανομή.....	7
1.4 Πρότυπα συνεχών κατανομών	8
1.4.1 Κατανομή Γάμμα.....	8
1.4.2 Μετατοπισμένη κατανομή Γάμμα.....	9
1.4.3 Κατανομή Lindley	10
1.4.4 Κατανομή Βήτα.....	11
1.4.5 Κατανομή Pareto.....	12
1.4.6 Αντίστροφη Γάμμα κατανομή.....	12
1.4.7 Αντίστροφη κανονική κατανομή.....	13
1.4.8 Λογαριθμοκανονική κατανομή.....	14
ΚΕΦΑΛΑΙΟ 2	15
Μεικτή κατανομή Poisson	15
2.1 Εισαγωγή.....	15
2.2 Μεικτή κατανομή Poisson – Ορισμός και βασικές ιδιότητες	15
2.3 Άλλα χαρακτηριστικά της μεικτής κατανομή Poisson	20
2.4 Παραδείγματα μεικτών κατανομών Poisson	23
2.4.1 Poisson – Γάμμα κατανομή	23
2.4.2 Poisson - Μετατοπισμένη Γάμμα κατανομή.....	25
2.4.3 Poisson - Lindley κατανομή	27
2.4.4 Poisson - Beta κατανομή.....	29
2.4.5 Poisson - Pareto κατανομή.....	32
2.4.6 Poisson - Inverse Gamma κατανομή	34
2.4.7 Poisson - Inverse Gaussian κατανομή.....	37
2.4.8 Poisson – Lognormal κατανομή	38
ΚΕΦΑΛΑΙΟ 3	41
Εφαρμογή μεικτών κατανομών Poisson σε πραγματικά δεδομένα	41

3.1 Εισαγωγή	41
3.2 Ανάλυση των δεδομένων Crimes	41
3.3 Ανάλυση των δεδομένων Lemaire.....	52
3.4 Συμπεράσματα.....	63
ΠΑΡΑΡΤΗΜΑ.....	65
ΒΙΒΛΙΟΓΡΑΦΙΑ.....	72
Ελληνική	72
Ξένα	72

ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ

Πίνακας 3.1. Δεδομένα Crimes	42
Πίνακας 3.2. Δεδομένα Crimes	42
Πίνακας 3.3. Βασικά χαρακτηριστικά της προσαρμογής της κατανομής Poisson στα δεδομένα Crimes.....	43
Πίνακας 3.4. Πιθανότητες και αναμενόμενες συχνότητες της κατανομής Poisson.....	44
Πίνακας 3.5. Βασικά χαρακτηριστικά της προσαρμογής της κατανομής Poisson-Gamma στα δεδομένα Crimes.....	45
Πίνακας 3.6. Πιθανότητες και αναμενόμενες συχνότητες της κατανομής Poisson-Gamma	45
Πίνακας 3.7. Βασικά χαρακτηριστικά της προσαρμογής της κατανομής Poisson-Lindley στα δεδομένα Crimes.....	46
Πίνακας 3.8. Πιθανότητες και αναμενόμενες συχνότητες της κατανομής Poisson-Lindley	47
Πίνακας 3.9. Βασικά χαρακτηριστικά της προσαρμογής της κατανομής Poisson- Inverse Gaussian στα δεδομένα Crimes.....	48
Πίνακας 3.10. Πιθανότητες και αναμενόμενες συχνότητες της κατανομής Poisson- Inverse Gaussian.....	49
Πίνακας 3.11. Βασικά χαρακτηριστικά της προσαρμογής της κατανομής Poisson- Lognormal στα δεδομένα Crimes.....	50
Πίνακας 3.12. Πιθανότητες και αναμενόμενες συχνότητες της κατανομής Poisson- Lognormal	51
Πίνακας 3.13. Δεδομένα Lemaire	52
Πίνακας 3.14. Δεδομένα Lemaire	52
Πίνακας 3.15. Βασικά χαρακτηριστικά της προσαρμογής της κατανομής Poisson στα δεδομένα Lemaire	53
Πίνακας 3.16. Πιθανότητες και αναμενόμενες συχνότητες της κατανομής Poisson	54
Πίνακας 3.17. Βασικά χαρακτηριστικά της προσαρμογής της κατανομής Poisson- Gamma στα δεδομένα Lemaire.....	55
Πίνακας 3.18. Πιθανότητες και αναμενόμενες συχνότητες της κατανομής Poisson- Gamma.....	56
Πίνακας 3.19. Βασικά χαρακτηριστικά της προσαρμογής της κατανομής Poisson- Lindley στα δεδομένα Lemaire	57
Πίνακας 3.20. Πιθανότητες και αναμενόμενες συχνότητες της κατανομής Poisson- Lindley	57
Πίνακας 3.21. Βασικά χαρακτηριστικά της προσαρμογής της κατανομής Poisson- Inverse Gaussian στα δεδομένα Lemaire.....	58
Πίνακας 3.22. Πιθανότητες και αναμενόμενες συχνότητες της κατανομής Poisson- Inverse Gaussian.....	59
Πίνακας 3.23. Βασικά χαρακτηριστικά της προσαρμογής της κατανομής Poisson- Lognormal στα δεδομένα Lemaire	60
Πίνακας 3.24. Πιθανότητες και αναμενόμενες συχνότητες της κατανομής Poisson- Lognormal	61
Πίνακας 3.25. Βασικά χαρακτηριστικά της προσαρμογής της κατανομής Poisson- Beta στα δεδομένα Lemaire	62

Πίνακας 3.26. Πιθανότητες και αναμενόμενες συχνότητες της κατανομής Poisson-Beta	62
Πίνακας 3.27. Δεδομένα Crimes – Σύγκριση.....	64
Πίνακας 3.28. Δεδομένα Lemaire – Σύγκριση	64

ΚΑΤΑΛΟΓΟΣ ΣΧΗΜΑΤΩΝ

Σχήμα 2.1. Συναρτήσεις πιθανότητας κατανομής Poisson-Gamma.....	24
Σχήμα 2.2. Συναρτήσεις πιθανότητας κατανομής Delaporte	27
Σχήμα 2.3. Συναρτήσεις πιθανότητας κατανομής Poisson-Lindley	29
Σχήμα 2.4. Συναρτήσεις πιθανότητας κατανομής Poisson-Beta	31
Σχήμα 2.5. Συναρτήσεις πιθανότητας κατανομής Poisson-Pareto.....	34
Σχήμα 2.6. Συναρτήσεις πιθανότητας κατανομής Poisson-Inverse Gamma	37
Σχήμα 2.7. Συναρτήσεις πιθανότητας κατανομής Poisson-Inverse Gaussian.....	38
Σχήμα 2.8. Συναρτήσεις πιθανότητας κατανομής Poisson-Lognormal	39
Σχήμα 3.1. Ραβδόγραμμα και θηκόγραμμα για τα δεδομένα Crimes.....	43
Σχήμα 3.2. Ιστόγραμμα εμπειρικών/θεωρητικών πιθανοτήτων και Poisson Q-Q διάγραμμα για τα δεδομένα Crimes	44
Σχήμα 3.3. Ιστόγραμμα εμπειρικών/θεωρητικών πιθανοτήτων και Poisson-Gamma Q-Q διάγραμμα για τα δεδομένα Crimes	46
Σχήμα 3.4. Ιστόγραμμα εμπειρικών/θεωρητικών πιθανοτήτων και Poisson-Lindley Q-Q διάγραμμα για τα δεδομένα Crimes	48
Σχήμα 3.5. Ιστόγραμμα εμπειρικών/θεωρητικών πιθανοτήτων και Poisson- Inverse Gaussian Q-Q διάγραμμα για τα δεδομένα Crimes.....	50
Σχήμα 3.6. Ιστόγραμμα εμπειρικών/θεωρητικών πιθανοτήτων και Poisson-Lognormal Q-Q διάγραμμα για τα δεδομένα Crimes	52
Σχήμα 3.7. Ραβδόγραμμα και θηκόγραμμα για τα δεδομένα Lemaire	53
Σχήμα 3.8. Ιστόγραμμα εμπειρικών/θεωρητικών πιθανοτήτων και Poisson Q-Q διάγραμμα για τα δεδομένα Lemaire	55
Σχήμα 3.9. Ιστόγραμμα εμπειρικών/θεωρητικών πιθανοτήτων και Poisson-Gamma Q-Q διάγραμμα για τα δεδομένα Lemaire	56
Σχήμα 3.10. Ιστόγραμμα εμπειρικών/θεωρητικών πιθανοτήτων και Poisson-Lindley Q- Q διάγραμμα για τα δεδομένα Lemaire	58
Σχήμα 3.11. Ιστόγραμμα εμπειρικών/θεωρητικών πιθανοτήτων και Poisson- Inverse Gaussian Q-Q διάγραμμα για τα δεδομένα Lemaire	60
Σχήμα 3.12. Ιστόγραμμα εμπειρικών/θεωρητικών πιθανοτήτων και Poisson-Lognormal Q-Q διάγραμμα για τα δεδομένα Lemaire	61
Σχήμα 3.13. Ιστόγραμμα εμπειρικών/θεωρητικών πιθανοτήτων και Poisson-Beta Q-Q διάγραμμα για τα δεδομένα Lemaire	63

ΚΕΦΑΛΑΙΟ 1

Μείξεις κατανομών

1.1 Εισαγωγή

Μείξεις κατανομών έχουν χρησιμοποιηθεί ευρέως στη μοντελοποίηση παρατηρούμενων καταστάσεων των οποίων τα διάφορα χαρακτηριστικά, όπως αυτά αντανακλώνται στα δεδομένα, διαφέρουν από εκείνα που θα αναμένονταν αν αυτά περιγράφονταν ικανοποιητικά από μια συνήθη κατανομή. Για παράδειγμα, σε αναλογιστικές εφαρμογές, τα παρατηρούμενα δεδομένα που αναφέρονται στον αριθμό των αποζημιώσεων σε μια συγκεκριμένη χρονική περίοδο παρουσιάζουν συχνά αυξημένη διακύμανση που υπερβαίνει αισθητά τη μέση τιμή τους. Ως εκ τούτου, δεν μπορούμε να υποθέσουμε ότι η κατανομή Poisson (ή οποιαδήποτε άλλη μορφή κατανομής που θα συνεπαγόταν σχεδόν ισότητα μεταξύ μέσης τιμής και διακύμανσης) είναι το θεωρητικό μοντέλο που μπορεί να περιγράψει την κατανομή πιθανότητας των αποζημιώσεων.

Γενικά, η υπόθεση μιας συγκεκριμένης κατανομής για τη «μητρική» κατανομή ενός συνόλου δεδομένων επιβάλλει μια ορισμένη σχέση μεταξύ της μέσης τιμής και της διακύμανσης, η οποία στην πράξη μπορεί να παραβιάζεται. Σε τέτοιες περιπτώσεις πιο γενικές οικογένειες κατανομών, γνωστές ως μείξεις κατανομών, υιοθετούνται ως εναλλακτικά μοντέλα τα οποία προσφέρουν μεγαλύτερη ευελιξία. Αυτές αποτελούν προεκτάσεις απλών μητρικών κατανομών οι οποίες εξαρτώνται από μια παράμετρο που η τιμή της είναι η παρατηρούμενη τιμή μιας τυχαίας μεταβλητής. Μεγάλο ενδιαφέρον στη διαχείριση κινδύνων παρουσιάζουν οι μεικτές κατανομές Poisson (η μητρική κατανομή είναι η κατανομή Poisson) που έχουν χρησιμοποιηθεί ευρέως ως κατανομές συχνότητας των απαιτήσεων στην Αναλογιστική Επιστήμη.

Στο κεφάλαιο αυτό θα παρουσιαστούν τα βασικότερα θεωρητικά αποτελέσματα για τη μείξη κατανομών καθώς επίσης και τα κυριότερα μοντέλα διακριτών και συνεχών κατανομών που θα μας απασχολήσουν στην παρούσα διπλωματική εργασία. Αρκετά από

τα αποτελέσματα του παρόντος κεφαλαίου μπορούν να βρεθούν στις εργασίες των Grandell (1997), Karlis & Xelakaki (2005) και Johnson et al. (1995).

1.2 Μείξη κατανομών

Οι μείξεις κατανομών (mixture distributions), ή αλλιώς μεμειγμένες κατανομές, διακρίνονται σε διακριτές και συνεχείς.

Ορισμός 1.1 (Διακριτή Μείξη Κατανομών). Η κατανομή μιας τυχαίας μεταβλητής X θα λέμε ότι είναι μια διακριτή μείξη κατανομών εάν η συνάρτηση κατανομής της $F_X(x)$ μπορεί να γραφτεί στη μορφή

$$F_X(x) = \sum_{i=1}^k p_i F_{X_i}(x),$$

για κάποιες τυχαίες μεταβλητές X_1, X_2, \dots, X_k με συναρτήσεις κατανομής $F_{X_1}(x), F_{X_2}(x), \dots, F_{X_k}(x)$, αντίστοιχα, και p_i ($1 \leq i \leq k$) θετικές σταθερές τέτοιες ώστε $\sum_{i=1}^k p_i = 1$.

Ο παραπάνω ορισμός οδηγεί στην ακόλουθη σχέση για τη συνάρτηση πυκνότητας (αντ. συνάρτηση πιθανότητας) $f_X(x)$ της συνεχούς (αντ. διακριτής) τυχαίας μεταβλητής X ,

$$f_X(x) = \sum_{i=1}^k p_i f_{X_i}(x)$$

όπου $f_{X_i}(x)$, $1 \leq i \leq k$, είναι η συνάρτηση πυκνότητας (αντ. συνάρτηση πιθανότητας) της συνεχούς (αντ. διακριτής) τυχαίας μεταβλητής X_i .

Αν οι συναρτήσεις κατανομής $F_{X_1}(x), F_{X_2}(x), \dots, F_{X_k}(x)$ είναι όλες από την ίδια οικογένεια κατανομών και διαφέρουν μόνο ως προς την τιμή κάποιας παραμέτρου $\lambda \in \Theta$, τότε μπορούμε να γράψουμε εναλλακτικά ότι

$$F_X(x) = \sum_{i=1}^k p_i F_{X_i}(x; \lambda_i),$$

και

$$f_X(x) = \sum_{i=1}^k p_i f_{X_i}(x; \lambda_i).$$

Αξίζει να σημειώσουμε ότι ο παραπάνω τύπος για την $f_X(x)$ (δείτε Titterington et al. (1985)) μπορεί να γραφεί ως

$$f_X(x) = \int_{\Theta} f(x; \lambda) dG_p(\lambda)$$

όπου $G_p(\cdot)$ δηλώνει το μέτρο πιθανότητας πάνω στο Θ που ορίζεται από το $\mathbf{p} = (p_1, p_2, \dots, p_k)$.

Ορισμός 1.2 (Συνεχής Μείξη Κατανομών). Η κατανομή μιας τυχαίας μεταβλητής X θα λέμε ότι είναι μια συνεχής μείξη κατανομών εάν η συνάρτηση κατανομής της $F_X(x)$ μπορεί να γραφτεί στη μορφή

$$F_X(x) = \int_{\Theta} F_{X|\Lambda}(x|\lambda) g_{\Lambda}(\lambda) d\lambda$$

για κάποια συνεχή τυχαία μεταβλητή Λ με συνάρτηση πυκνότητας $g_{\Lambda}(\lambda)$ όπου $\lambda \in \Theta$.

Η τυχαία μεταβλητή Λ καλείται **μεικτική τυχαία μεταβλητή**. Ο παραπάνω ορισμός οδηγεί στην ακόλουθη σχέση

$$f_X(x) = \int_{\Theta} f_{X|\Lambda}(x|\lambda) g_{\Lambda}(\lambda) d\lambda,$$

η οποία συμβολικά γράφεται ως

$$f(x|\lambda) \underset{\lambda}{\wedge} g(\lambda).$$

Στη συνέχεια θα παρουσιάσουμε μια σειρά αξιόλογων αποτελεσμάτων που αφορούν τις μείξεις κατανομών. Το πρώτο σημαντικό αποτέλεσμα είναι η ισχύς της προσεταιριστικής ιδιότητας, δηλαδή η κατανομή

$$[\underset{\lambda}{f(x|\lambda)} \wedge \underset{\mu}{g(\lambda|\mu)}] \wedge h(\mu)$$

είναι ισοδύναμη με την κατανομή

$$\underset{\lambda}{f(x|\lambda)} \wedge [\underset{\mu}{g(\lambda|\mu)} \wedge h(\mu)]$$

όπου f, g, h συναρτήσεις πυκνότητας (ή πιθανότητας). Το παραπάνω αποτέλεσμα ισχύει με την προϋπόθεση ότι υπάρχει ανεξαρτησία μεταξύ των λ, μ .

Η παρακάτω πρόταση οφείλεται στον Sibuya (1979) και δείχνει τη σχέση που έχει μια κατανομή μείξης με το γινόμενο δύο ιδιαίτερων τυχαίων μεταβλητών.

Πρόταση 1.1: Ας υποθέσουμε ότι η δεσμευμένη συνάρτηση πυκνότητας της X είναι η $f(\cdot | \lambda)$, όπου λ είναι μια παράμετρος κλίμακας. Επίσης, ας υποθέσουμε ότι η λ είναι μια τυχαία μεταβλητή με συνάρτηση πυκνότητας $g(\cdot | \varphi)$ για κάποια παράμετρο φ (συμβολικά $\lambda \sim g(\cdot | \varphi)$). Τότε, η αδέσμευτη συνάρτηση πυκνότητας της τυχαίας μεταβλητής X είναι ίδια με τη συνάρτηση πυκνότητας της τυχαίας μεταβλητής $Z = X_1 X_2$, όπου η X_1 έχει συνάρτηση πυκνότητας $f(\cdot | 1)$ και η X_2 έχει συνάρτηση πυκνότητας $g(\cdot | \varphi)$.

Η παραπάνω πρόταση δικαιολογεί τη διπλή παραγωγή ορισμένων κατανομών ως μείξεις και ως κατανομές γινομένου δύο τυχαίων μεταβλητών. Η κατανομή Βήτα και η κατανομή t είναι τυπικά παραδείγματα.

Η επόμενη πρόταση αφορά τη σύνδεση των μίξεων κατανομών με τις σύνθετες κατανομές. Υπενθυμίζουμε ότι μια τυχαία μεταβλητή S έχει σύνθετη κατανομή αν μπορεί να γραφεί στη μορφή

$$S = I(N \geq 1) \sum_{j=1}^N X_j = \begin{cases} 0, & N = 0 \\ \sum_{j=1}^N X_j, & N \geq 1 \end{cases}$$

όπου η τυχαία μεταβλητή N είναι μια μη αρνητική διακριτή τυχαία μεταβλητή ανεξάρτητη από τις τυχαίες μεταβλητές X_1, X_2, \dots που είναι επίσης ανεξάρτητες μεταξύ τους. Αν συμβολίσουμε με f τη συνάρτηση πιθανότητας της N και με g την κοινή συνάρτηση πυκνότητας (ή πιθανότητας) των τυχαίων μεταβλητών X_1, X_2, \dots τότε η κατανομή της S συμβολίζεται με

$$f \vee g.$$

Οι δύο παρακάτω προτάσεις οφείλονται στον Gurland (1957).

Πρόταση 1.2: Έστω ότι η τυχαία μεταβλητή Λ έχει συνάρτηση πιθανότητας g και πιθανογεννήτρια συνάρτηση της μορφής $[\varphi(t)]^n$. Υποθέτουμε ότι το n είναι μια τυχαία μεταβλητή με συνάρτηση πιθανότητας f . Τότε το μοντέλο $f \vee g$ είναι ισοδύναμο με το μοντέλο $g(x|\lambda) \wedge f(n)$.

Πρόταση 1.3: Ισχύει ότι η κατανομή

$$[f(x|\lambda) \wedge g(\lambda)] \vee h$$

είναι ισοδύναμη με την

$$[f(x|\lambda) \vee h] \wedge_{\lambda} g(\lambda)$$

Η επόμενη πρόταση αφορά τη σύνδεση των μείξεων κατανομών με συνελίξεις κατανομών. Υπενθυμίζουμε ότι η συνάρτηση πιθανότητας της συνέλιξης δύο διακριτών κατανομών με συναρτήσεις πιθανότητας f και g (συμβολισμός $f * g$) δίνεται από τον τύπο

$$p(x) = \sum_{k=0}^x f(x-k)g(k)$$

(για συνεχείς κατανομές αντικαθιστούμε το άθροισμα με ολοκλήρωμα).

Πρόταση 1.4: Ισχύει ότι τα μοντέλα

$$[f(x|\lambda) * g(y|\mu)] \wedge_{\lambda} h(\lambda)$$

και

$$[f(x|\lambda) \wedge_{\lambda} h(\lambda)] * g(y|\mu)$$

είναι ισοδύναμα υπό την προϋπόθεση ότι η συνάρτηση πυκνότητας (ή πιθανότητας) $g(\cdot|\mu)$ δεν εξαρτάται από το λ .

Κλείνοντας την Παράγραφο 1.2 δίνουμε ένα αποτέλεσμα που αφορά τις ροπές και τη διακύμανση της X .

Πρόταση 1.5. Ισχύει ότι

$$E(X^k) = E[E(X^k|\Lambda)],$$

$$Var(X) = E[Var(X|\Lambda)] + Var[E(X|\Lambda)].$$

Απόδειξη: Η απόδειξη θα δοθεί στην περίπτωση της συνεχής μείξης κατανομών. Έχουμε (υπό την προϋπόθεση αλλαγής της σειράς των ολοκληρωμάτων)

$$\begin{aligned} E(X^k) &= \int x^k f_X(x) dx = \int x^k \left(\int_{\Theta} f_{X|\Lambda}(x|\lambda) g_{\Lambda}(\lambda) d\lambda \right) dx \\ &= \int_{\Theta} g_{\Lambda}(\lambda) \left(\int x^k f_{X|\Lambda}(x|\lambda) dx \right) d\lambda \end{aligned}$$

$$\begin{aligned}
&= \int_{\Theta} g_{\Lambda}(\lambda) (E(X^k | \Lambda = \lambda)) d\lambda \\
&= E[E(X^k | \Lambda)]
\end{aligned}$$

και

$$\begin{aligned}
Var(X) &= E(X^2) - E^2(X) \\
&= E[E(X^2 | \Lambda)] - E^2[E(X | \Lambda)] \\
&= E[E(X^2 | \Lambda)] - E^2[E(X | \Lambda)] \pm E[E^2(X | \Lambda)] \\
&= E[E(X^2 | \Lambda)] - E[E^2(X | \Lambda)] + E[E^2(X | \Lambda)] - E^2[E(X | \Lambda)] \\
&= E[Var(X | \Lambda)] + Var[E(X | \Lambda)].
\end{aligned}$$

1.3 Πρότυπα διακριτών κατανομών

Στην παρούσα παράγραφο θα παρουσιαστούν περιληπτικά οι κυριότερες διακριτές κατανομές που θα μας απασχολήσουν σε επόμενα κεφάλαια.

1.3.1 Διωνυμική κατανομή

Λέμε ότι μια διακριτή τυχαία μεταβλητή X ακολουθεί τη διωνυμική κατανομή με παραμέτρους n και p (n θετικός ακέραιος και $0 < p < 1$), αν η συνάρτηση πιθανότητας της $f(x)$ δίνεται από τον τύπο

$$f(x) = P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, 2, \dots, n$$

Συμβολικά, γράφουμε ότι $X \sim B(n, p)$. Αν η τ.μ. $X \sim B(n, p)$, τότε

$$E(X) = np,$$

$$Var(X) = np(1-p),$$

$$a_3 = \frac{E[(X - E(X))^3]}{\sigma^3} = \frac{\mu_3}{\sigma^3} = \frac{1-2p}{\sqrt{np(1-p)}}$$

$$a_4 = \frac{E[(X - E(X))^4]}{\sigma^4} = \frac{\mu_4}{\sigma^4} = 3 - \frac{6}{n} + \frac{1}{npq},$$

$$\mu^{(k)} = E(X(X-1)\cdots(X-k+1)) = \frac{n! p^k}{(n-k)!}.$$

1.3.2 Κατανομή Poisson

Λέμε ότι μια διακριτή τυχαία μεταβλητή X ακολουθεί την κατανομή Poisson με παράμετρο $\lambda > 0$, αν η συνάρτηση πιθανότητάς της $f(x)$ δίνεται από τον τύπο

$$f(x) = P(X = x) = e^{-\lambda} \frac{\lambda^x}{x!}, \quad x = 0, 1, \dots$$

Συμβολικά, γράφουμε ότι $X \sim P(\lambda)$. Στην περίπτωση που η τυχαία μεταβλητή $X \sim B(n, p_n)$ και $np_n \rightarrow \lambda$ καθώς $n \rightarrow \infty$, τότε

$$\lim_{n \rightarrow \infty} \binom{n}{x} p_n^x (1 - p_n)^{n-x} = e^{-\lambda} \frac{\lambda^x}{x!}, \quad x = 0, 1, 2, \dots$$

Αν η τ.μ. $X \sim P(\lambda)$, τότε

$$E(X) = Var(X) = \lambda,$$

$$E(X^r) \leq \left(\frac{r}{\log\left(\frac{r}{\lambda+1}\right)} \right)^r \leq \lambda^r \exp\left(\frac{r^2}{2\lambda}\right),$$

$$\mu^{(r)} = E(X(X-1)\cdots(X-r+1)) = \lambda^r,$$

$$\mu_3 = E[(X - E(X))^3] = \lambda,$$

$$\mu_4 = E[(X - E(X))^4] = \lambda(1 + 3\lambda),$$

$$a_3 = \frac{\mu_3}{\sigma^3} = \lambda^{-1/2}, \quad a_4 = \frac{\mu_4}{\sigma^4} = 3 + \frac{1}{\lambda}.$$

1.3.3 Αρνητική διωνυμική κατανομή

Λέμε ότι μια διακριτή τυχαία μεταβλητή X ακολουθεί την αρνητική διωνυμική κατανομή με παραμέτρους r και p ($r > 0, 0 < p < 1$), αν η συνάρτηση πιθανότητάς της $f(x)$ δίνεται από τον τύπο

$$f(x) = P(X = x) = \binom{x+r-1}{x} (1-p)^x p^r, \quad x = 0, 1, 2, \dots$$

Συμβολικά, γράφουμε ότι $X \sim NB(r, p)$. Αρκετές φορές γράφουμε ότι $X \sim NB_0(r, p)$ για να δώσουμε έμφαση στο γεγονός ότι το σύνολο τιμών της X ξεκινά από το 0. Σημειώνουμε ότι σε αρκετά βιβλία η τυχαία μεταβλητή $Y = X + r$ αναφέρεται ως αρνητική διωνυμική κατανομή. Σε αυτή την περίπτωση έχουμε

$$f(y) = P(Y = y) = \binom{y-1}{r-1} (1-p)^{y-r} p^r, \quad y = r, r+1, r+2, \dots$$

Για $r = 1$, η αρνητική διωνυμική κατανομή ανάγεται στην γεωμετρική κατανομή με παράμετρο p (συμβ. $NB(r, p) \equiv G(p)$).

Αν η τ.μ. $X \sim NB(r, p)$, τότε

$$P(s) = E(s^X) = \left(\frac{p}{1-qs} \right)^r,$$

$$E(X) = \frac{r(1-p)}{p},$$

$$\text{Var}(X) = \frac{r(1-p)}{p^2},$$

$$\mu_3 = E[(X - E(X))^3] = \frac{r(1-p)(2-p)}{p^3},$$

$$\mu_4 = E[(X - E(X))^4] = \frac{r(1-p)(3rq + 6q + p^2)}{p^4},$$

$$a_3 = \frac{\mu_3}{\sigma^3} = \frac{1+q}{\sqrt{rq}}, \quad a_4 = \frac{\mu_4}{\sigma^4} = \frac{3rq + 6q + p^2}{rq},$$

$$\mu^{(k)} = E(X(X-1)\cdots(X-k+1)) = \left(\frac{q}{p} \right)^k \frac{(r+k-1)!}{(r-1)!}.$$

1.4 Πρότυπα συνεχών κατανομών

Στην παρούσα παράγραφο θα παρουσιαστούν περιληπτικά οι κυριότερες συνεχείς κατανομές οι οποίες θα μας απασχολήσουν σε επόμενα κεφάλαια.

1.4.1 Κατανομή Γάμμα

Λέμε ότι μια τυχαία μεταβλητή X ακολουθεί την κατανομή Γάμμα με παραμέτρους $a > 0$ και $\beta > 0$, αν η συνάρτηση πυκνότητάς της $f(x)$ δίνεται από τον τύπο

$$f(x) = \frac{\beta^a}{\Gamma(a)} x^{a-1} e^{-\beta x}, \quad x > 0.$$

Συμβολικά γράφουμε ότι $X \sim G(a, \beta)$. Στον παραπάνω τύπο $\Gamma(\cdot)$ είναι η συνάρτηση Γάμμα, που ορίζεται από τη σχέση

$$\Gamma(a) = \int_0^{\infty} t^{a-1} e^{-t} dt, \quad a > 0.$$

Στην περίπτωση που το a είναι ένας θετικός ακέραιος, έστω $a = n$, τότε ισχύει

$$\Gamma(n) = (n - 1)!, \quad n = 1, 2, \dots$$

Είναι φανερό ότι για $a = 1$, η κατανομή Γάμμα ανάγεται στην εκθετική κατανομή με παράμετρο β ((συμβ. $G(1, \beta) \equiv \mathcal{E}(\beta)$). Στην περίπτωση όπου η παράμετρος a είναι ένας θετικός ακέραιος η κατανομή Γάμμα αναφέρεται συνήθως ως κατανομή Erlang.

Αν η τ.μ. $X \sim G(a, \beta)$, τότε

$$E(X) = \frac{a}{\beta},$$

$$\text{Var}(X) = \frac{a}{\beta^2},$$

$$E(X^r) = \frac{\Gamma(a+r)}{\beta^r \Gamma(a)}, \quad r = 1, 2, \dots$$

$$\mu_3 = E[(X - E(X))^3] = \frac{2a}{\beta^3},$$

$$\mu_4 = E[(X - E(X))^4] = \frac{3a(a+2)}{\beta^4},$$

$$a_3 = \frac{\mu_3}{\sigma^3} = \frac{2}{\sqrt{a}}, \quad a_4 = \frac{\mu_4}{\sigma^4} = 3 + \frac{6}{a}.$$

1.4.2 Μετατοπισμένη κατανομή Γάμμα

Λέμε ότι μια τυχαία μεταβλητή X ακολουθεί τη μετατοπισμένη κατανομή Γάμμα με παραμέτρους $\mu > 0$, $a > 0$ και $\beta > 0$, αν η συνάρτηση πυκνότητάς της $f(x)$ δίνεται από τον τύπο

$$f(x) = \begin{cases} \frac{\beta^a}{\Gamma(a)} (x - \mu)^{a-1} e^{-\beta(x-\mu)}, & x \geq \mu \\ 0, & x < \mu. \end{cases}$$

Συμβολικά γράφουμε ότι $X \sim G(a, \beta, \mu)$. Αν η τ.μ. $X \sim G(a, \beta, \mu)$, τότε

$$E(X) = \mu + \frac{a}{\beta},$$

$$\text{Var}(X) = \frac{a}{\beta^2},$$

$$\mu_3 = E[(X - E(X))^3] = \frac{2a}{\beta^3},$$

$$\mu_4 = E[(X - E(X))^4] = \frac{3a(a+2)}{\beta^4},$$

$$a_3 = \frac{\mu_3}{\sigma^3} = \frac{2}{\sqrt{a}}, \quad a_4 = \frac{\mu_4}{\sigma^4} = 3 + \frac{6}{a}.$$

1.4.3 Κατανομή Lindley

Λέμε ότι μια τυχαία μεταβλητή X ακολουθεί την κατανομή Lindley με παράμετρο $\theta > 0$, αν η συνάρτηση πυκνότητάς της $f(x)$ δίνεται από τον τύπο

$$f(x) = \frac{\theta^2}{\theta + 1} (1 + x)e^{-\theta x}, \quad x > 0, \theta > 0$$

(δείτε, π.χ., Ghitany et al. (2008)). Συμβολικά γράφουμε ότι $X \sim \text{Lindley}(\theta)$. Είναι φανερό ότι η εκθετική κατανομή (συμβ. $\mathcal{E}(\theta)$) με συνάρτηση πυκνότητας που δίνεται από τον τύπο

$$f(x) = \theta e^{-\theta x}, \quad x > 0, \theta > 0$$

είναι μια πυκνότητα που μοιάζει με αυτή της κατανομής Lindley.

Αν η τ.μ. $X \sim \text{Lindley}(\theta)$, τότε

$$E(X) = \frac{\theta + 2}{\theta(\theta + 1)},$$

$$\text{Var}(X) = \frac{2 + 4\theta + \theta^2}{(\theta + 1)^2 \theta^2},$$

$$E(X^r) = \frac{r!(\theta + r + 1)}{\theta^r(\theta + 1)}, \quad r = 1, 2, 3, \dots,$$

$$\mu_3 = E[(X - E(X))^3] = \frac{2(\theta^3 + 6\theta^2 + 6\theta + 2)}{\theta^3(\theta + 1)^3},$$

$$\mu_4 = E[(X - E(X))^4] = \frac{3(3\theta^4 + 24\theta^3 + 44\theta^2 + 32\theta + 8)}{\theta^4(\theta + 1)^4},$$

$$a_3 = \frac{\mu_3}{\sigma^3} = \frac{2(\theta^3 + 6\theta^2 + 6\theta + 2)}{(\theta^2 + 4\theta + 2)^{3/2}},$$

$$a_4 = \frac{\mu_4}{\sigma^4} = \frac{3(3\theta^4 + 24\theta^3 + 44\theta^2 + 32\theta + 8)}{(\theta^2 + 4\theta + 2)^2}.$$

1.4.4 Κατανομή Βήτα

Λέμε ότι μια τυχαία μεταβλητή X ακολουθεί την κατανομή Βήτα (πρώτου είδους) με παραμέτρους $a, \beta > 0$, αν η συνάρτηση πυκνότητας της $f(x)$ δίνεται από τον τύπο

$$f(x) = \frac{x^{a-1}(1-x)^{\beta-1}}{B(a, \beta)}, \quad 0 < x < 1$$

όπου $B(a, \beta) = \frac{\Gamma(a)\Gamma(\beta)}{\Gamma(a+\beta)}$. Συμβολικά γράφουμε ότι $X \sim \text{Beta}(a, \beta)$.

Η κατανομή βήτα έχει εφαρμοστεί για να μοντελοποιήσει τη συμπεριφορά τυχαίων μεταβλητών που περιορίζονται σε διαστήματα πεπερασμένου μήκους σε μια μεγάλη ποικιλία επιστημονικών κλάδων. Η κατανομή βήτα είναι ένα κατάλληλο μοντέλο για την περιγραφή ποσοστών και αναλογιών.

Αν η τ.μ. $X \sim \text{Beta}(a, \beta)$, τότε

$$E(X) = \frac{a}{a + \beta}$$

$$\text{Var}(X) = \frac{a\beta}{(a + \beta)^2(a + \beta + 1)}$$

$$E(X^r) = \prod_{k=0}^{r-1} \frac{a + k}{a + \beta + k}$$

$$\mu_3 = E[(X - E(X))^3] = \frac{2a\beta(\beta - a)}{(a + \beta)^2(a + \beta + 1)(a + \beta + 2)}$$

$$\mu_4 = E[(X - E(X))^4] = \frac{6a\beta[(a - \beta)^2(a + \beta + 1) - a\beta(a + \beta + 2)]}{(a + \beta)^4(a + \beta + 1)^2(a + \beta + 2)(a + \beta + 3)}$$

$$a_3 = \frac{\mu_3}{\sigma^3} = \frac{2(\beta - a)\sqrt{a + \beta + 1}}{(a + \beta + 2)\sqrt{a\beta}},$$

$$a_4 = \frac{\mu_4}{\sigma^4} = 3 + \frac{6[(a - \beta)^2(a + \beta + 1) - a\beta(a + \beta + 2)]}{a\beta(a + \beta + 2)(a + \beta + 3)}.$$

1.4.5 Κατανομή Pareto

Λέμε ότι μια τυχαία μεταβλητή X ακολουθεί την κατανομή Pareto με παραμέτρους $a, \beta > 0$, αν η συνάρτηση πυκνότητας της $f(x)$ δίνεται από τον τύπο

$$f(x) = \begin{cases} \frac{a\beta^a}{x^{a+1}}, & x \geq \beta \\ 0, & x < \beta \end{cases}$$

(δείτε, π.χ., Forbes et al. (2011)). Συμβολικά γράφουμε ότι $X \sim P(a, \beta)$. Αν η τ.μ. $X \sim P(a, \beta)$, τότε

$$E(X) = \frac{a\beta}{a-1}, \quad a > 1$$

$$Var(X) = \frac{a\beta^2}{(a-1)^2(a-2)}, \quad a > 2$$

$$E(X^r) = \frac{a\beta^r}{a-r} \quad r < a$$

$$\mu_3 = E[(X - E(X))^3] = \frac{2a(a+1)\beta^3}{(a-1)^3(a-2)(a-3)}, \quad a > 3$$

$$\mu_4 = E[(X - E(X))^4] = \frac{3a(3a^2 + a + 2)\beta^4}{(a-1)^4(a-2)(a-3)(a-4)}, \quad a > 4$$

$$a_3 = \frac{\mu_3}{\sigma^3} = \frac{2(1+a)}{a-3} \cdot \sqrt{\frac{a-2}{a}}, \quad a > 3,$$

$$a_4 = \frac{\mu_4}{\sigma^4} = 3 + \frac{6(a^3 + a^2 - 6a - 2)}{a(a-3)(a-4)}, \quad a > 4.$$

1.4.6 Αντίστροφη Γάμμα κατανομή

Λέμε ότι μια τυχαία μεταβλητή X ακολουθεί την αντίστροφη Γάμμα κατανομή με παραμέτρους $a, \beta > 0$, αν η συνάρτηση πυκνότητάς της $f(x)$ δίνεται από τον τύπο

$$f(x) = \frac{\beta^a}{\Gamma(a)} x^{-a-1} e^{(-\beta/x)}, \quad x > 0$$

(δείτε, π.χ., Witkovsky (2001)). Συμβολικά γράφουμε ότι $X \sim IGamma(a, \beta)$. Η αντίστροφη Γάμμα κατανομή είναι η κατανομή της τ.μ. $Y = g(W) = 1/W$, όπου $W \sim G(a, \beta)$. Πράγματι

$$f_Y(y) = f_W(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right| = \frac{\beta^a}{\Gamma(a)} \left(\frac{1}{y} \right)^{a-1} e^{(-\beta/y)} \frac{1}{y^2} = \frac{\beta^a}{\Gamma(a)} y^{-a-1} e^{(-\beta/y)}.$$

Αν η τ.μ. $X \sim IG(a, \beta)$, τότε

$$E(X) = \frac{\beta}{a-1}, \quad a > 1,$$

$$Var(X) = \frac{\beta^2}{(a-1)^2(a-2)}, \quad a > 2,$$

$$E(X^r) = \beta^r \frac{\Gamma(a-r)}{\Gamma(a)} = \frac{\beta^r}{(a-1)(a-2)\dots(a-r)}, \quad a > r$$

$$\mu_3 = E[(X - E(X))^3] = \frac{4\beta^{5/2}}{(a-1)^3(a-2)(a-3)}, \quad a > 3,$$

$$\mu_4 = E[(X - E(X))^4] = \frac{6\beta^4(5a-11)}{(a-1)^4(a-2)^2(a-3)(a-4)}, \quad a > 4,$$

$$a_3 = \frac{\mu_3}{\sigma^3} = \frac{4\sqrt{a-2}}{a-3}, \quad a > 3,$$

$$a_4 = \frac{\mu_4}{\sigma^4} = \frac{(3a+15)(a-2)}{(a-3)(a-4)}, \quad a > 4.$$

1.4.7 Αντίστροφη κανονική κατανομή

Λέμε ότι μια τυχαία μεταβλητή X ακολουθεί την αντίστροφη κανονική κατανομή με παραμέτρους $\mu, \theta > 0$, αν η συνάρτηση πυκνότητάς της $f(x)$ δίνεται από τον τύπο

$$f_X(x) = \left(\frac{\theta}{2\pi x^3} \right)^{1/2} \exp\left(-\frac{\theta(x-\mu)^2}{2x\mu^2} \right), \quad x > 0,$$

(δείτε, π.χ., Chhikara & Folk (1989), Panjer & Willmot (1992), Johnson et al. (1994)). Συμβολικά γράφουμε ότι $X \sim IG(\mu, \theta)$.

Αν η τ.μ. $X \sim IG(\mu, \theta)$, τότε

$$E(X) = \mu,$$

$$Var(X) = \frac{\mu^3}{\theta},$$

$$a_3 = 3\sqrt{\frac{\mu}{\theta}},$$

$$a_4 = \frac{15\mu + 3\theta}{\theta},$$

$$k_r = \frac{\mu^{2r-1}}{\theta^{r-1}} (1)(3) \cdots (2r-3).$$

Σημειώνουμε ότι

$$k_r = \left. \frac{d^r \log M(t)}{dt^r} \right|_{t=0}$$

όπου

$$M(t) = E(e^{tX}) = \exp \left[\frac{\theta}{\mu} \left(1 - \sqrt{1 - \frac{2\mu^2 t}{\theta}} \right) \right].$$

1.4.8 Λογαριθμοκανονική κατανομή

Λέμε ότι μια τυχαία μεταβλητή X ακολουθεί την λογαριθμοκανονική κατανομή (LogNormal) με παραμέτρους $\mu \in \mathbb{R}$ και $\sigma > 0$ αν η συνάρτηση πυκνότητάς της $f(x)$ δίνεται από τον τύπο

$$f_X(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right), \quad x \geq 0,$$

(δείτε Johnson et al. (1995)). Συμβολικά γράφουμε ότι $X \sim LN(\mu, \sigma)$.

Αν η τ.μ. $X \sim LN(\mu, \sigma)$, τότε

$$E(X) = \exp\left(\mu + \frac{\sigma^2}{2}\right),$$

$$Var(X) = (\exp(\sigma^2) - 1) \exp(2\mu + \sigma^2),$$

$$a_3 = (\exp(\sigma^2) + 2)\sqrt{(\exp(\sigma^2) - 1)}$$

$$a_4 = \exp(4\sigma^2) + 2 \exp(3\sigma^2) + 3 \exp(2\sigma^2) - 3$$

$$E(X^r) = \exp\left(r\mu + \frac{r^2\sigma^2}{2}\right).$$

ΚΕΦΑΛΑΙΟ 2

Μεικτή κατανομή Poisson

2.1 Εισαγωγή

Μία από τις πιο σημαντικές κλάσεις κατανομών στην αναλογιστική επιστήμη είναι η κλάση των μεικτών κατανομών Poisson. Χρησιμοποιείται κυρίως στη μοντελοποίηση του αριθμού των απαιτήσεων που προκύπτουν από μη ομογενείς πληθυσμούς, όταν σε κάθε ένα από αυτούς ο αριθμός των απαιτήσεων ακολουθεί κατανομή Poisson.

Στο παρόν κεφάλαιο θα παρουσιαστούν οι βασικές ιδιότητες και ορισμένα άλλα χαρακτηριστικά της μεικτής κατανομής Poisson. Στη συνέχεια θα δοθούν μερικά από τα πιο σημαντικά παραδείγματα μεικτών κατανομών Poisson και θα γίνει μια αναλυτική παρουσίασή τους στην οποία συμπεριλαμβάνουμε αποδείξεις βασικών αποτελεσμάτων. Αρκετά αποτελέσματα του παρόντος κεφαλαίου μπορούν να βρεθούν στις εργασίες των Grandell (1997), Karlis & Xelakaki (2005), Johnson et al. (1995), Dean et al. (1989), Bulmer (1974) και Ong et al. (2021).

2.2 Μεικτή κατανομή Poisson – Ορισμός και βασικές ιδιότητες

Στην παρούσα παράγραφο θα δοθεί ο ορισμός της μεικτής κατανομής Poisson και θα παρουσιαστούν ορισμένα βασικά χαρακτηριστικά της που θα μας βοηθήσουν στην περαιτέρω μελέτη της. Ένας τυπικός ορισμός για την μεικτή κατανομή Poisson είναι ο ακόλουθος.

Ορισμός 2.1 (Μεικτή κατανομή Poisson). Θα λέμε ότι η διακριτή τυχαία μεταβλητή X έχει μεμειγμένη (ή πιο απλά μεικτή) κατανομή Poisson με μεικτική (συνεχή) μια μη αρνητική τυχαία μεταβλητή Λ που έχει συνάρτηση πυκνότητας $g(\lambda)$, αν η συνάρτηση πιθανότητάς της δίνεται από τον τύπο

$$f(x) = P(X = x) = \int_0^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} g(\lambda) d\lambda, \quad x = 0, 1, 2, \dots$$

Τη συνάρτηση κατανομής της τυχαίας μεταβλητής Λ θα τη συμβολίζουμε ως $G(\lambda)$, δηλαδή

$$G(\lambda) = P(\Lambda \leq \lambda) = \int_0^{\lambda} g(t) dt, \quad \lambda > 0.$$

Τη μεικτή κατανομή Poisson με μεικτική την μη αρνητική τυχαία μεταβλητή Λ θα τη συμβολίζουμε με $MP(G)$.

Σύμφωνα με τον Grandell (1997) έχουμε την ακόλουθη πρόταση.

Πρόταση 2.1. Έστω ότι η τυχαία μεταβλητή X ακολουθεί την κατανομή $MP(G)$ όπου G είναι η συνάρτηση κατανομής μιας μη αρνητικής τυχαίας μεταβλητής Λ . Τότε

(i) $E[X] = E[\Lambda].$

(ii) $Var[X] = E[\Lambda] + Var[\Lambda].$

(iii) $P\{X > x\} = \int_0^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} (1 - G(\lambda)) d\lambda$

(iv) $P\{X \leq x\} = \int_0^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} G(\lambda) d\lambda$

(v) $P\{\Lambda \leq y | X = x\} = \frac{\int_0^y \lambda^x e^{-\lambda} dG(\lambda)}{\int_0^{\infty} \lambda^x e^{-\lambda} dG(\lambda)}.$

(vi) $E\{\Lambda | X = x\} = \frac{\int_0^{\infty} \lambda^{x+1} e^{-\lambda} U(d\lambda)}{\int_0^{\infty} \lambda^x e^{-\lambda} U(d\lambda)}.$

(vii) $P_X(s) = \int_0^{\infty} e^{\lambda(s-1)} dG(\lambda) = L_{\Lambda}(1 - s) = M_{\Lambda}(s - 1)..$

Απόδειξη. (i)-(ii). Προκύπτουν άμεσα από τις σχέσεις

$$E[X] = E[E(X|\Lambda)] = E[\Lambda]$$

και

$$Var[X] = E[Var(X|\Lambda)] + Var[E(X|\Lambda)] = E[\Lambda] + Var[\Lambda]$$

αφού $E(X|\Lambda = \lambda) = Var(X|\Lambda = \lambda) = \lambda.$

(iii) Τα βασικά βήματα της απόδειξης είναι τα εξής (δείτε Willmot (1990))

$$P\{X > x\} = \sum_{k=x+1}^{\infty} P(N = k) = \sum_{k=x+1}^{\infty} \int_0^{\infty} \frac{e^{-\lambda} \lambda^k}{k!} g(\lambda) d\lambda = \int_0^{\infty} \sum_{k=x+1}^{\infty} \frac{e^{-\lambda} \lambda^k}{k!} dG(\lambda)$$

$$= \int_0^{\infty} \sum_{k=x+1}^{\infty} \left(\frac{e^{-\lambda} \lambda^{k-1}}{(k-1)!} - \frac{e^{-\lambda} \lambda^k}{k!} \right) (1 - G(\lambda)) d\lambda = \int_0^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} (1 - G(\lambda)) d\lambda.$$

(iv) Έχουμε ότι

$$\begin{aligned} P\{X \leq x\} &= 1 - P\{X > x\} = \int_0^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} d\lambda - \int_0^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} (1 - G(\lambda)) d\lambda \\ &= \int_0^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} G(\lambda) d\lambda. \end{aligned}$$

(v) Έχουμε ότι

$$P\{\Lambda \leq y | X = x\} = \frac{P\{\Lambda \leq y, X = x\}}{P\{X = x\}} = \frac{\int_0^y \frac{\lambda^x}{x!} e^{-\lambda} dG(\lambda)}{\int_0^{\infty} \frac{\lambda^x}{x!} e^{-\lambda} dG(\lambda)} = \frac{\int_0^y \lambda^x e^{-\lambda} dG(\lambda)}{\int_0^{\infty} \lambda^x e^{-\lambda} dG(\lambda)}.$$

(vi) Προκύπτει από τη βασική σχέση

$$E(\Lambda | X = x) = \int_0^{\infty} y dP(\Lambda \leq y | X = x).$$

(vii) Για την πιθανογεννήτρια της X έχουμε

$$\begin{aligned} P_X(s) &= E(s^X) = \sum_{k=0}^{\infty} s^k P(X = k) = \int_0^{\infty} \sum_{k=0}^{\infty} s^k \frac{\lambda^k}{k!} e^{-\lambda} dG(\lambda) \\ &= \int_0^{\infty} e^{\lambda(s-1)} dG(\lambda) = L_{\Lambda}(1-s) = M_{\Lambda}(s-1) \end{aligned}$$

όπου $L_{\Lambda}(t) = \int_0^{\infty} e^{-\lambda t} dG(\lambda)$ είναι ο μετασχηματισμός Laplace της Λ , και $M_{\Lambda}(t) = \int_0^{\infty} e^{\lambda t} dG(\lambda)$ είναι η ροπογεννήτρια συνάρτηση της Λ .

Πρόταση 2.2. Η καθοδική παραγοντική ροπή r -τάξης της μεικτής κατανομής Poisson $MP(G)$ δίνεται από τη σχέση

$$E[X(X-1)(X-2)\cdots(X-r+1)] = E(\Lambda^r).$$

Απόδειξη. Έστω $P_X(s) = P(s)$ η πιθανογεννήτρια συνάρτηση της X . Γνωρίζουμε ότι

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

Από την Πρόταση 2.1 (vii) παίρνουμε

$$P'(s) = \int_0^{\infty} \lambda e^{-\lambda} e^{\lambda s} g(\lambda) d\lambda.$$

Έτσι

$$P''(s) = \int_0^{\infty} \lambda^2 e^{-\lambda} e^{\lambda s} g(\lambda) d\lambda$$

και γενικότερα μπορεί να διαπιστωθεί η ισχύς της σχέσης

$$P^{(r)}(s) = \int_0^{\infty} \lambda^r e^{-\lambda} e^{\lambda s} g(\lambda) d\lambda$$

Επομένως

$$E[X(X-1)\cdots(X-r+1)]P^{(r)}(1) = \int_0^{\infty} \lambda^r g(\lambda) d\lambda = E(\Lambda^r).$$

Για τις πρώτες τέσσερις ροπές ως προς την αρχή της μεικτής κατανομής Poisson ισχύει η ακόλουθη πρόταση.

Πρόταση 2.3. Οι 4 πρώτες ροπές (γύρω από το 0) της μεικτής κατανομής Poisson $MP(G)$ δίνονται από τις σχέσεις

- (i) $E(X) = E(\Lambda)$.
- (ii) $E(X^2) = E(\Lambda^2) + E(\Lambda)$.
- (iii) $E(X^3) = E(\Lambda^3) + 3E(\Lambda^2) + E(\Lambda)$.
- (iv) $E(X^4) = E(\Lambda^4) + 6E(\Lambda^3) + 7E(\Lambda^2) + E(\Lambda)$.

Απόδειξη. (i) Αφού

$$E(X|\Lambda = \lambda) = \sum_{x=0}^{\infty} x \frac{e^{-\lambda} \lambda^x}{x!}$$

έχουμε ότι

$$E(X) = E(E(X|\Lambda)) = E\left(\sum_{x=0}^{\infty} x \frac{e^{-\Lambda} \Lambda^x}{x!}\right) = E\left(e^{-\Lambda} \Lambda \sum_{x=1}^{\infty} \frac{\Lambda^{x-1}}{(x-1)!}\right) = E(e^{-\Lambda} \Lambda e^{\Lambda}) = E(\Lambda).$$

(ii) Για τη δεύτερη ροπή, με τον ίδιο τρόπο, παίρνουμε

$$\begin{aligned} E(X^2) &= E\left(\sum_{x=0}^{\infty} x^2 \frac{e^{-\Lambda} \Lambda^x}{x!}\right) = E\left(e^{-\Lambda} \Lambda \sum_{x=1}^{\infty} \frac{(x-1+1)\Lambda^{x-1}}{(x-1)!}\right) \\ &= E\left(e^{-\Lambda} \Lambda^2 \sum_{x=2}^{\infty} \frac{\Lambda^{x-2}}{(x-2)!} + e^{-\Lambda} \Lambda \sum_{x=1}^{\infty} \frac{\Lambda^{x-1}}{(x-1)!}\right) = E(\Lambda^2) + E(\Lambda). \end{aligned}$$

(iii) Για την τρίτη ροπή έχουμε

$$\begin{aligned}
E(X^3) &= E\left(e^{-\Lambda} \sum_{x=0}^{\infty} \frac{x^3 \Lambda^x}{x!}\right) = E\left(e^{-\Lambda} \sum_{x=1}^{\infty} \frac{(x-1+1)^2 \Lambda^x}{(x-1)!}\right) \\
&= E\left(e^{-\Lambda} \left[\sum_{x=2}^{\infty} \frac{(x-2)+3}{(x-2)!} + \sum_{x=1}^{\infty} \frac{1}{(x-1)!} \right] \Lambda^x\right) \\
&= E\left(e^{-\Lambda} \left[\sum_{x=3}^{\infty} \frac{1}{(x-3)!} + \sum_{x=2}^{\infty} \frac{3}{(x-2)!} + \sum_{x=1}^{\infty} \frac{1}{(x-1)!} \right] \Lambda^x\right) \\
&= E\left(e^{-\Lambda} \left[\Lambda^3 \sum_{x=3}^{\infty} \frac{\Lambda^{x-3}}{(x-3)!} + 3\Lambda^2 \sum_{x=2}^{\infty} \frac{\Lambda^{x-2}}{(x-2)!} + \Lambda \sum_{x=1}^{\infty} \frac{\Lambda^{x-1}}{(x-1)!} \right]\right) \\
&= E(\Lambda^3) + 3E(\Lambda^2) + E(\Lambda).
\end{aligned}$$

(iii) Για την τέταρτη ροπή έχουμε

$$\begin{aligned}
E(X^4) &= E\left(e^{-\Lambda} \sum_{x=0}^{\infty} \frac{x^4 \Lambda^x}{x!}\right) = E\left(e^{-\Lambda} \sum_{x=1}^{\infty} \frac{(x-1+1)^3 \Lambda^x}{(x-1)!}\right) \\
&= E\left(e^{-\Lambda} \sum_{x=1}^{\infty} [(x-1)^3 + 3(x-1)^2 + 3(x-1) + 1] \frac{\Lambda^x}{(x-1)!}\right) \\
&= E\left(e^{-\Lambda} \left[\sum_{x=4}^{\infty} \frac{1}{(x-4)!} + \sum_{x=3}^{\infty} \frac{6}{(x-3)!} + \sum_{x=2}^{\infty} \frac{7}{(x-2)!} + \sum_{x=1}^{\infty} \frac{1}{(x-1)!} \right] \Lambda^x\right) \\
&= E(\Lambda^4) + 6E(\Lambda^3) + 7E(\Lambda^2) + E(\Lambda).
\end{aligned}$$

Για τις κεντρικές ροπές της μεικτής κατανομής Poisson ισχύει η ακόλουθη πρόταση.

Πρόταση 2.4. Η 2^η, 3^η και η 4^η κεντρική ροπή της μεικτής κατανομής Poisson $MP(G)$

δίνονται από τις σχέσεις

(i) $\mu_2 = Var(\Lambda) + E(\Lambda).$

(ii) $\mu_3 = E[\Lambda - E(\Lambda)]^3 + 3Var(\Lambda) + E(\Lambda).$

(iii) $\mu_4 = E[\Lambda - E(\Lambda)]^4 + 6[E[\Lambda - E(\Lambda)]^3 + Var(\Lambda)E(\Lambda)] + [7Var(\Lambda) + 3[E(\Lambda)]^2] + E(\Lambda).$

Απόδειξη. (i) Για τη δεύτερη κεντρική ροπή έχουμε

$$\mu_2 = E[(X - E(X))^2] = Var(X) = Var(\Lambda) + E(\Lambda).$$

(ii) Για την τρίτη κεντρική ροπή, χρησιμοποιώντας την Πρόταση 2.3, έχουμε

$$\mu_3 = E[X - E(X)]^3 = E(X^3) - 3E(X^2)E(X) + 2[E(X)]^3$$

$$= E[\Lambda - E(\Lambda)]^3 + 3\text{Var}(\Lambda) + E(\Lambda).$$

(iii) Για την τέταρτη κεντρική ροπή, έχουμε

$$\mu_4 = E[X - E(X)]^4 = E(X^4) - 4E(X^3)E(X) + 6E(X^2)[E(X)]^2 - 3[E(X)]^4.$$

Χρησιμοποιώντας την Πρόταση 2.3, παίρνουμε

$$\begin{aligned} \mu_4 &= \{E(\Lambda^4) - 4E(\Lambda^3)E(\Lambda) + 6E(\Lambda^2)[E(\Lambda)]^2 - 3[E(\Lambda)]^4\} \\ &\quad + 6\{E(\Lambda^3) - 2E(\Lambda^2)E(\Lambda) + [E(\Lambda)]^3\} \\ &\quad + \{7E(\Lambda^2) - 4[E(\Lambda)]^2\} + E(\Lambda) \\ &= E[\Lambda - E(\Lambda)]^4 + 6[E[\Lambda - E(\Lambda)]^3 + \text{Var}(\Lambda)E(\Lambda)] + [7\text{Var}(\Lambda) + \\ &\quad + 3[E(\Lambda)]^2] + E(\Lambda). \end{aligned}$$

2.3 Άλλα χαρακτηριστικά της μεικτής κατανομή Poisson

Σε αυτή την παράγραφο θα δώσουμε κάποιες περαιτέρω ιδιότητες των μεικτών κατανομών Poisson. Θα δούμε κάποια θεωρητικά αποτελέσματα όπως είναι η σύγκριση των μεικτών κατανομών Poisson με τις απλές κατανομές Poisson με την ίδια μέση τιμή, πως μπορούμε να προσεγγίσουμε τη συνάρτηση πιθανότητας της μεικτής κατανομής Poisson, κλπ.

Έστω $p(x)$, $x = 0, 1, 2, \dots$, η συνάρτηση πιθανότητας της μεικτής κατανομής Poisson $MP(G)$ με μέση τιμή m , και έστω $p(x; m)$ η συνάρτηση πιθανότητας της κατανομής Poisson με παράμετρο m ($P(m)$). Για τη σύγκριση των δύο συναρτήσεων πιθανότητας (δείτε Feller (1943)) έχουμε:

- (i) $p(0) \geq p(0; m)$, δηλαδή η πιθανότητα παρατήρησης της μηδενικής τιμής είναι πάντα υψηλότερη σε μια μεικτή κατανομή Poisson από ότι σε μια απλή κατανομή Poisson με την ίδια μέση τιμή,
- (ii) $\frac{p(1)}{p(0)} \leq \frac{p(1; m)}{p(0; m)} = m$, δηλαδή ο λόγος της πιθανότητας του 1 ως προς αυτή του 0, είναι μικρότερος από την μέση τιμή της μεικτής κατανομής Poisson.

Επίσης ο Shaked (1980) έδειξε ότι η συνάρτηση $p(x) - p(x; m)$ έχει ακριβώς δύο εναλλαγές πρόσημων της μορφής $+ - +$, υπονοώντας ότι η μεικτή κατανομή Poisson δίνει υψηλότερη πιθανότητα στο ενδεχόμενο $\{X = 0\}$, και έχει μακρύτερη δεξιά ουρά.

Αυτό το αποτέλεσμα μπορεί να χρησιμοποιηθεί για να ελέγξουμε αν μια μεικτή κατανομή Poisson είναι επαρκής για την περιγραφή ενός συνόλου δεδομένων.

Ο αναλυτικός υπολογισμός των πιθανοτήτων των μεικτών κατανομών Poisson σύμφωνα με τον Ορισμό 2.1 είναι αρκετές φορές περίπλοκος και για αυτό το λόγο χρησιμοποιούμε άλλες μεθόδους.

Πρόταση 2.5: Η συνάρτηση πιθανότητας της μεικτής κατανομής Poisson $MP(G)$ ικανοποιεί τη σχέση

$$P(X = x) = \frac{1}{x!} \sum_{r=0}^{\infty} \frac{(-1)^r}{r!} E(\Lambda^{x+r}).$$

Απόδειξη: Έχουμε

$$\begin{aligned} P(X = x) &= \frac{1}{x!} \int_0^{\infty} e^{-\lambda} \lambda^x g(\lambda) d\lambda = \int_0^{\infty} \left(\sum_{r=0}^{\infty} \frac{(-\lambda)^r}{r!} \right) \frac{\lambda^x}{x!} g(\lambda) d\lambda \\ &= \sum_{r=0}^{\infty} \int_0^{\infty} (-1)^r \frac{\lambda^{x+r}}{x! r!} g(\lambda) d\lambda = \sum_{r=0}^{\infty} \frac{(-1)^r}{x! r!} \int_0^{\infty} \lambda^{x+r} g(\lambda) d\lambda = \sum_{r=0}^{\infty} \frac{(-1)^r}{x! r!} E(\Lambda^{x+r}). \end{aligned}$$

Η ακόλουθη πρόταση οφείλεται στον Ong (1995) και βασίζεται στο ανάπτυγμα του Taylor.

Πρόταση 2.6. Έστω $g(\lambda)$ η συνάρτηση πυκνότητας πιθανότητας της μεικτικής τυχαίας μεταβλητής Λ της κατανομής $MP(G)$. Αν η $g(\lambda)$ έχει πεπερασμένες παραγώγους στο σημείο k έως και τουλάχιστον τάξεως n , τότε η συνάρτηση πιθανότητας της μεικτής κατανομής Poisson $MP(G)$ δίνεται προσεγγιστικά από το ανάπτυγμα

$$P(X = k) \simeq g(k) + \frac{1}{k} \sum_{i=2}^n \frac{\mu_i h^{(i)}(k)}{i!}$$

όπου $h(k) = kg(k)$, $h^{(i)}(k)$ δηλώνει την i -οστή παράγωγο της $h(k)$ υπολογισμένη στο σημείο k , και μ_i είναι η i -οστή κεντρική ροπή της Γάμμα κατανομής $G(k, 1)$.

Η παραπάνω προσέγγιση έχει το μειονέκτημα ότι δεν δίνει την πιθανότητα $P(X = 0)$, και ενδεχομένως ο υπολογισμός των παραγώγων της συνάρτησης $h(k)$ να μην είναι εύκολος σε ορισμένες περιπτώσεις.

Τέλος για τον υπολογισμό των πιθανοτήτων μια μεικτής κατανομής Poisson μπορούν να χρησιμοποιηθούν αναδρομικές σχέσεις. Η ακόλουθη πρόταση οφείλεται στον Willmot (1993).

Πρόταση 2.7. Αν η συνάρτηση πυκνότητας πιθανότητας $g(\lambda)$ της μεικτικής τυχαίας μεταβλητής Λ ικανοποιεί τη σχέση

$$\frac{d}{d\lambda} \log g(\lambda) = \frac{\sum_{i=0}^k s_i \lambda^i}{\sum_{i=0}^k w_i \lambda^i}, \quad \lambda \in (0, +\infty)$$

για κάποιες σταθερές $s_i, w_i, i = 0, 1, \dots, k, k > 0$, τότε συνάρτηση πιθανότητας της κατανομής $MP(G)$ ικανοποιεί τον αναδρομική τύπο

$$\sum_{n=-1}^k \{\varphi_n + m w_{n+1}\} (m+n)^{(n)} P(X = m+n) = 0$$

όπου $\varphi_{-1} = 0, \varphi_n = s_n + (n+1)w_{n+1} - w_n$ για $n = 0, 1, \dots, k$, και $a^{(b)} = \prod_{i=1}^b (a+1-i)$.

Η ακόλουθη πρόταση οφείλεται στον Feller (1943) και αναφέρεται στην ιδιότητα της αναπαρωσιμότητας (reproducibility) της μεικτής κατανομής Poisson.

Πρόταση 2.8. Το άθροισμα δύο ανεξάρτητων τυχαίων μεταβλητών X και Y , με κατανομές $MP(G_1)$ και $MP(G_2)$, αντίστοιχα, είναι μια μεικτή Poisson κατανομή $MP(G)$ όπου $G = G_1 * G_2$.

Σημειώνουμε ότι η πράξη “*” δηλώνει συνέλιξη, δηλαδή

$$G_1(\lambda) * G_2(\lambda) = \int_{-\infty}^{\infty} G_1(\lambda - \mu) dG_2(\mu).$$

Σχετική είναι και η ακόλουθη πρόταση που οφείλεται στους Willmot & Sundt (1989).

Πρόταση 2.9. Το άθροισμα δύο ανεξάρτητων τυχαίων μεταβλητών X και Y , με κατανομές $MP(G)$ και $P(\lambda)$, αντίστοιχα, είναι μια μεικτή Poisson κατανομή $MP(U)$ όπου $U(y) = G(y - \lambda), y \geq \lambda$.

Η ακόλουθη πρόταση οφείλεται στον Grandell (1997).

Πρόταση 2.10. Έστω δυο τυχαίες μεταβλητές N_1 και N_2 με κατανομές $MP(G_1)$ και $MP(G_2)$ και μεικτικές τυχαίες μεταβλητές Λ_1 και Λ_2 , αντιστοίχως. Τότε η μεικτή κατανομή Poisson $MP(G_1)$ συγκλίνει υπό την έννοια της κατανομής στη μεικτή κατανομή Poisson $MP(G_2)$ (δηλαδή $N_1 \xrightarrow{d} N_2$), αν και μόνο αν $\Lambda_1 \xrightarrow{d} \Lambda_2$.

Η ακόλουθη πρόταση οφείλεται στον Maceda (1948).

Πρόταση 2.11. Αν η μεικτική τυχαία μεταβλητή είναι απείρως διαιρετή τότε και η αντίστοιχη μεικτή κατανομή Poisson είναι απείρως διαιρετή.

Σημειώνουμε ότι μια κατανομή πιθανότητας θεωρείται απείρως διαιρετή (infinitely divisible) αν για κάθε θετικό ακέραιο n , η κατανομή αυτή μπορεί να προκύψει ως η κατανομή του αθροίσματος n ανεξάρτητων και ισόνομων τυχαίων μεταβλητών.

Κλείνοντας την παρούσα παράγραφο τονίζουμε ότι περαιτέρω σημαντικές ιδιότητες των μεικτών κατανομών Poisson μπορούν να βρεθούν στον Grandell (1997) και Karlis & Xekalaki (2005).

2.4 Παραδείγματα μεικτών κατανομών Poisson

2.4.1 Poisson – Γάμμα κατανομή

Η πρώτη και πιο συνηθισμένη επιλογή για τη μεικτική τυχαία μεταβλητή Λ είναι η κατανομή $G(a, \beta)$, η οποία μελετήθηκε από τους Greenwood & Yule (1920). Σε αυτή την περίπτωση ισχύει η ακόλουθη πρόταση.

Πρόταση 2.12. (i) Η συνάρτηση πιθανότητας της μεικτής κατανομής Poisson με μεικτική την τυχαία μεταβλητή $\Lambda \sim G(a, \beta)$ δίνεται από τον τύπο

$$f(x) = \frac{\Gamma(a+x)}{\Gamma(a)x!} \left(\frac{1}{1+\beta}\right)^x \left(\frac{\beta}{1+\beta}\right)^a, \quad x = 0, 1, 2, \dots$$

(ii) Ισχύει ότι

$$P(s) = E(s^X) = \left(\frac{\beta}{1+\beta-s}\right)^a,$$

$$E(X) = \frac{a}{\beta},$$

$$Var(X) = \frac{a(1 + \beta)}{\beta^2},$$

$$\mu_3 = E[(X - E(X))^3] = \frac{a(\beta + 1)(\beta + 2)}{\beta^3},$$

$$\mu_4 = E[(X - E(X))^4] = \frac{3a^2(\beta + 1)^2 + a(6 + 12\beta + 7\beta^2 + \beta^3)}{\beta^4}.$$

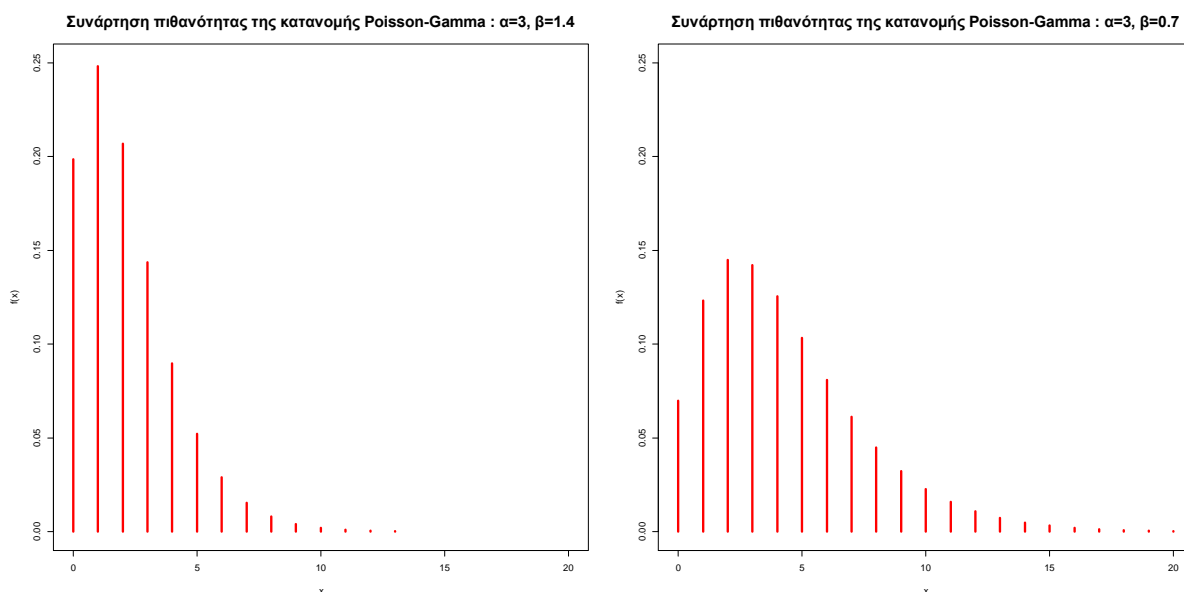
Απόδειξη. (i) Για $x = 0, 1, 2, \dots$ έχουμε

$$\begin{aligned} f(x) &= \int_0^\infty \frac{e^{-\lambda} \lambda^x}{x!} g_\lambda(\lambda) d\lambda = \int_0^\infty \frac{e^{-\lambda} \lambda^x}{x!} \frac{\beta^a}{\Gamma(a)} \lambda^{a-1} e^{-\beta\lambda} d\lambda \\ &= \frac{\beta^a}{\Gamma(a)x!} \int_0^\infty \lambda^{x+a-1} e^{-\lambda(1+\beta)} d\lambda. \end{aligned}$$

Με την αλλαγή μεταβλητής $\lambda = t/(1 + \beta)$, και χρησιμοποιώντας τον ορισμό της συνάρτησης $\Gamma(\cdot)$ παίρνουμε

$$\begin{aligned} f(x) &= \frac{\beta^a}{\Gamma(a)x!} \cdot \frac{1}{(1 + \beta)^{x+a}} \int_0^\infty t^{x+a-1} e^{-t} dt \\ &= \frac{\Gamma(a + x)}{\Gamma(a)x!} \left(\frac{1}{1 + \beta}\right)^x \left(\frac{\beta}{1 + \beta}\right)^a. \end{aligned}$$

(ii) Από το (i) προκύπτει ότι η τυχαία μεταβλητή X ακολουθεί την αρνητική διωνυμική κατανομή με παραμέτρους $r = a$ και $p = \frac{\beta}{1+\beta}$, δηλαδή $X \sim Nb(a, \frac{\beta}{1+\beta})$. Επομένως η πιθανογεννήτρια $P(s)$, η μέση τιμή $E(X)$, η διακύμανση $Var(X)$, η τρίτη και η τέταρτη κεντρική ροπή μ_3 και μ_4 της X είναι οι αντίστοιχες ποσότητες της κατανομής $Nb(a, \frac{\beta}{1+\beta})$.



Σχήμα 2.1. Συναρτήσεις πιθανότητας κατανομής Poisson-Gamma

2.4.2 Poisson - Μετατοπισμένη Γάμμα κατανομή

Η περίπτωση της μετατοπισμένης Γάμμα κατανομής ως κατανομής της μεικτικής τυχαίας μεταβλητής Λ μελετήθηκε από τον Ruohonen (1988) (δείτε επίσης Willmot & Sundt (1989)). Σε αυτή την περίπτωση ισχύει η ακόλουθη πρόταση.

Πρόταση 2.13. (i) Η συνάρτηση πιθανότητας της μεικτικής κατανομής Poisson με μεικτική την τυχαία μεταβλητή $\Lambda \sim G(a, \beta, \mu)$ δίνεται από τον τύπο

$$f(x) = \sum_{k=0}^x \left(\frac{e^{-\mu} \mu^{x-k}}{(x-k)!} \right) \frac{\Gamma(a+k)}{\Gamma(a)k!} \left(\frac{1}{1+\beta} \right)^k \left(\frac{\beta}{1+\beta} \right)^a, \quad x = 0, 1, 2, \dots$$

(ii) Ισχύει ότι

$$P(s) = E(s^X) = e^{-\mu(1-s)} \left(\frac{\beta}{1+\beta-s} \right)^a,$$

$$E(X) = \mu + \frac{a}{\beta},$$

$$\text{Var}(X) = \frac{a}{\beta^2} + \mu + \frac{a}{\beta},$$

$$\mu_3 = E[(X - E(X))^3] = \frac{2a}{\beta^3} + \frac{3a}{\beta^2} + \mu + \frac{a}{\beta},$$

$$\mu_4 = E[(X - E(X))^4]$$

$$= \frac{3a^2(\beta+1)^2 + a(6+12\beta+7\beta+\beta^3) + \mu\beta^2(6a+3\mu\beta^2+6a\beta+\beta^2)}{\beta^4}.$$

Απόδειξη. (i) Για $x = 0, 1, 2, \dots$ έχουμε

$$\begin{aligned} f(x) &= \int_{\mu}^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} g_{\Lambda}(\lambda) d\lambda = \int_{\mu}^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} \frac{\beta^a}{\Gamma(a)} (\lambda - \mu)^{a-1} e^{-\beta(\lambda-\mu)} d\lambda \\ &= \frac{\beta^a}{\Gamma(a)x!} \int_{\mu}^{\infty} e^{-\lambda} \lambda^x (\lambda - \mu)^{a-1} e^{-\beta(\lambda-\mu)} d\lambda. \end{aligned}$$

Με την αλλαγή μεταβλητής $z = \lambda - \mu$, παίρνουμε

$$\begin{aligned} f(x) &= \frac{\beta^a}{\Gamma(a)x!} \int_0^{\infty} e^{-(z+\mu)} (z+\mu)^x z^{a-1} e^{-\beta z} dz \\ &= \frac{\beta^a e^{-\mu}}{\Gamma(a)x!} \int_0^{\infty} (z+\mu)^x z^{a-1} e^{-(1+\beta)z} dz \end{aligned}$$

$$\begin{aligned}
&= \frac{\beta^a e^{-\mu}}{\Gamma(a)x!} \int_0^\infty \sum_{k=0}^x \binom{x}{k} \mu^{x-k} z^k z^{a-1} e^{-(1+\beta)z} dz \\
&= \frac{\beta^a e^{-\mu}}{\Gamma(a)x!} \sum_{k=0}^x \binom{x}{k} \mu^{x-k} \int_0^\infty z^{k+a-1} e^{-(1+\beta)z} dz.
\end{aligned}$$

Με την αλλαγή μεταβλητής $z = t/(1 + \beta)$, και χρησιμοποιώντας τον ορισμό της συνάρτησης $\Gamma(\cdot)$ παίρνουμε

$$\begin{aligned}
f(x) &= \frac{\beta^a e^{-\mu}}{\Gamma(a)x!} \cdot \frac{1}{(1 + \beta)^{k+a}} \sum_{k=0}^x \binom{x}{k} \mu^{x-k} \int_0^\infty t^{k+a-1} e^{-t} dt \\
&= \sum_{k=0}^x \left(\frac{e^{-\mu} \mu^{x-k}}{(x-k)!} \right) \frac{\Gamma(k+a)}{\Gamma(a)k!} \left(\frac{1}{1+\beta} \right)^k \left(\frac{\beta}{1+\beta} \right)^a.
\end{aligned}$$

(ii) Για την πιθανογεννήτρια της τυχαίας μεταβλητής X έχουμε

$$\begin{aligned}
P(s) = E(s^X) &= \sum_{x=0}^\infty s^x f(x) = \sum_{x=0}^\infty s^x \int_\mu^\infty \frac{e^{-\lambda} \lambda^x}{x!} \frac{\beta^a}{\Gamma(a)} (\lambda - \mu)^{a-1} e^{-\beta(\lambda-\mu)} d\lambda \\
&= \frac{\beta^a}{\Gamma(a)} \int_\mu^\infty e^{-\lambda} \left(\sum_{x=0}^\infty \frac{(\lambda s)^x}{x!} \right) e^{-\beta(\lambda-\mu)} (\lambda - \mu)^{a-1} d\lambda \\
&= \frac{\beta^a}{\Gamma(a)} \int_\mu^\infty e^{-\lambda(1-s)} e^{-\beta(\lambda-\mu)} (\lambda - \mu)^{a-1} d\lambda.
\end{aligned}$$

Με την αλλαγή μεταβλητής $z = \lambda - \mu$, παίρνουμε

$$P(s) = \frac{\beta^a}{\Gamma(a)} e^{-\mu(1-s)} \int_0^\infty z^{a-1} e^{-(1+\beta-s)z} dz.$$

Με την αλλαγή μεταβλητής $z = t/(1 + \beta - s)$, παίρνουμε

$$\begin{aligned}
P(s) &= \frac{\beta^a}{\Gamma(a)} e^{-\mu(1-s)} \frac{1}{(1 + \beta - s)^a} \int_0^\infty z^{a-1} e^{-t} dt \\
&= e^{-\mu(1-s)} \left(\frac{\beta}{1 + \beta - s} \right)^a.
\end{aligned}$$

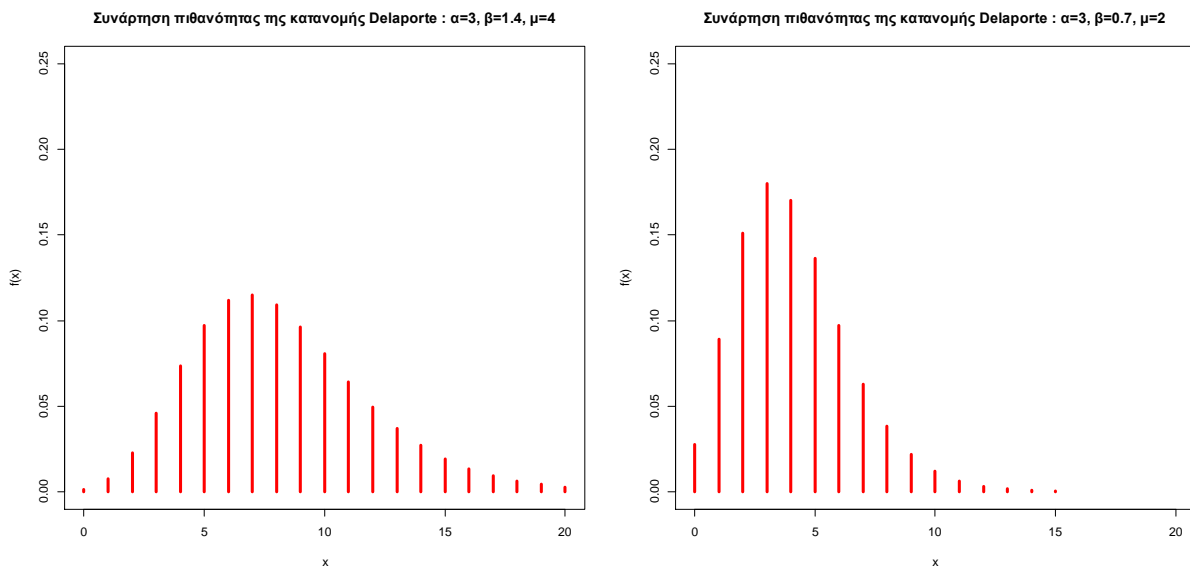
Επειδή η μεικτική τυχαία μεταβλητή $\Lambda \sim G(a, \beta, \mu)$, τότε από την Πρόταση 2.3 την Πρόταση 2.4 και την Παράγραφο 1.4.2 έχουμε τα εξής:

$$\begin{aligned}
E(X) &= E(\Lambda) = \mu + \frac{a}{\beta}, \\
\text{Var}(X) &= \text{Var}(\Lambda) + E(\Lambda) = \frac{a}{\beta^2} + \mu + \frac{a}{\beta}, \\
\mu_3 &= E[(X - E(X))^3] = \frac{2a}{\beta^3} + \frac{3a}{\beta^2} + \mu + \frac{a}{\beta},
\end{aligned}$$

$$\mu_4 = E[(X - E(X))^4] = \frac{3a^2(\beta + 1)^2 + a(6 + 12\beta + 7\beta + \beta^3) + \mu\beta^2(6a + 3\mu\beta^2 + 6a\beta + \beta^2)}{\beta^4}.$$

Η κατανομή της μεικτής κατανομής Poisson με μεικτική την μετατοπισμένη κατανομή Γάμμα είναι γνωστή και ως κατανομή Delaporte. Από τη μορφή της πιθανογεννήτριας της προκύπτει ότι η κατανομή Delaporte προκύπτει ως η κατανομή του αθροίσματος (συνέλιξη) δύο ανεξάρτητων τυχαίων μεταβλητών που η μία έχει κατανομή $P(\mu)$ και η άλλη κατανομή $Nb(a, \frac{\beta}{1+\beta})$.

Παρατήρηση. Αν στα αποτελέσματα της παρούσας παραγράφου θέσουμε $\mu = 0$, τότε προκύπτουν τα αποτελέσματα της προηγούμενης παραγράφου.



Σχήμα 2.2. Συναρτήσεις πιθανότητας κατανομής Delaporte

2.4.3 Poisson - Lindley κατανομή

Η περίπτωση της Lindley κατανομής ως κατανομής της μεικτικής τυχαίας μεταβλητής L μελετήθηκε από τον Sankaran (1970). Σε αυτή την περίπτωση έχουμε τα ακόλουθα αποτελέσματα.

Πρόταση 2.14. (i) Η συνάρτηση πιθανότητας της μεικτής κατανομής Poisson με μεικτική την τυχαία μεταβλητή $L \sim Lindley(\theta)$ δίνεται από τον τύπο

$$f(x) = \frac{\theta^2}{\theta + 1} \cdot \frac{x + 2 + \theta}{(1 + \theta)^{x+2}}, \quad x = 0, 1, 2, \dots, \quad \theta > 0.$$

(ii) Ισχύει ότι

$$\begin{aligned}
P(s) &= \frac{\theta^2}{(\theta+1)} \frac{(\theta+2-s)}{(\theta+1-s)^2}, \\
E(X) &= \frac{(2+\theta)}{(\theta+1)\theta}, \\
\text{Var}(X) &= \frac{(2+4\theta+\theta^2)}{(\theta+1)^2\theta^2} + \frac{(2+\theta)}{(\theta+1)\theta}, \\
\mu_3 &= E[(X-E(X))^3] = \frac{2(8+6\theta+6\theta^2+\theta^3)}{(\theta+1)^3\theta^3} + \frac{3(2+4\theta+\theta^2)}{(\theta+1)^2\theta^2} + \frac{(2+\theta)}{(\theta+1)\theta}, \\
\mu_4 &= E[(X-E(X))^4] \\
&= \frac{12(6+16\theta+17\theta^2+8\theta^3+\theta^4)}{(\theta+1)^4\theta^4} + \frac{6(20+22\theta+18\theta^2+3\theta^3)}{(\theta+1)^3\theta^3} \\
&\quad + \frac{2(13+20\theta+5\theta^2)}{(\theta+1)^2\theta^2} + \frac{(2+\theta)}{(\theta+1)\theta}.
\end{aligned}$$

Απόδειξη. (i) Για $x = 0, 1, 2, \dots$ έχουμε

$$\begin{aligned}
f(x) &= \int_0^\infty \frac{e^{-\lambda}\lambda^x}{x!} \frac{\theta^2}{\theta+1} (\lambda+1)e^{-\lambda\theta} d\lambda \\
&= \frac{\theta^2}{x!(\theta+1)} \int_0^\infty (\lambda^{x+1} + \lambda^x) e^{-(1+\theta)\lambda} d\lambda \\
&= \frac{\theta^2}{x!(\theta+1)} \left[\frac{\Gamma(x+2)}{(1+\theta)^{x+2}} + \frac{\Gamma(x+1)}{(1+\theta)^{x+1}} \right] \\
&= \frac{\theta^2}{x!(\theta+1)} \frac{\Gamma(x+1)}{(1+\theta)^{x+1}} \left(\frac{x+1}{1+\theta} + 1 \right) \\
&= \frac{\theta^2}{(\theta+1)} \cdot \frac{x+2+\theta}{(1+\theta)^{x+2}}.
\end{aligned}$$

(ii) Για την πιθανογεννήτρια της τυχαίας μεταβλητής X έχουμε

$$\begin{aligned}
P(s) &= \int_0^\infty e^{-\lambda(1-s)} \frac{\theta^2}{\theta+1} (\lambda+1)e^{-\lambda\theta} d\lambda \\
&= \frac{\theta^2}{\theta+1} \left\{ \int_0^\infty \lambda e^{-\lambda(\theta+1-s)} d\lambda + \int_0^\infty e^{-\lambda(\theta+1-s)} d\lambda \right\} \\
&= \frac{\theta^2}{\theta+1} \left\{ \frac{1}{(\theta+1-s)^2} + \frac{1}{\theta+1-s} \right\} \\
&= \frac{\theta^2}{\theta+1} \cdot \frac{(\theta+2-s)}{(\theta+1-s)^2}
\end{aligned}$$

Επειδή η μεικτική τυχαία μεταβλητή $\Lambda \sim \text{Lindley}(\theta)$, τότε από την Πρόταση 2.3 την Πρόταση 2.4 και την Παράγραφο 1.4.3 έχουμε τα εξής:

$$E(X) = E(\Lambda) = \frac{(2 + \theta)}{(\theta + 1)\theta},$$

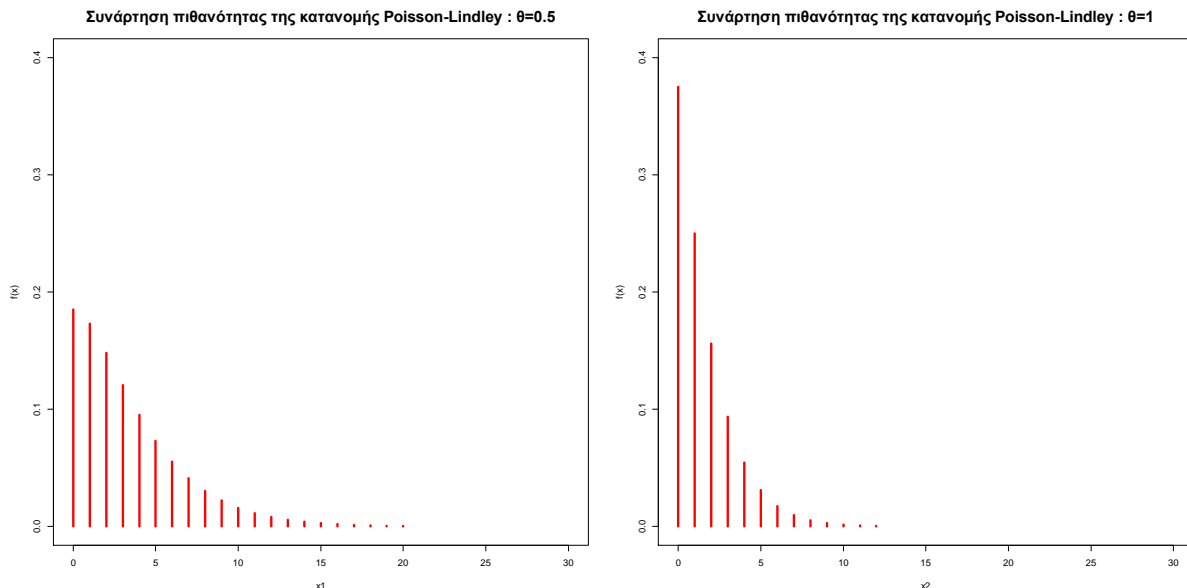
$$Var(X) = Var(\Lambda) + E(\Lambda) = \frac{(2 + 4\theta + \theta^2)}{(\theta + 1)^2\theta^2} + \frac{(2 + \theta)}{(\theta + 1)\theta},$$

$$\mu_3 = E[(X - E(X))^3] = \frac{2(8 + 6\theta + 6\theta^2 + \theta^3)}{(\theta + 1)^3\theta^3} + \frac{3(2 + 4\theta + \theta^2)}{(\theta + 1)^2\theta^2} + \frac{(2 + \theta)}{(\theta + 1)\theta},$$

$$\mu_4 = E[(X - E(X))^4]$$

$$= \frac{12(6 + 16\theta + 17\theta^2 + 8\theta^3 + \theta^4)}{(\theta + 1)^4\theta^4} + \frac{6(20 + 22\theta + 18\theta^2 + 3\theta^3)}{(\theta + 1)^3\theta^3}$$

$$+ \frac{2(13 + 20\theta + 5\theta^2)}{(\theta + 1)^2\theta^2} + \frac{(2 + \theta)}{(\theta + 1)\theta}.$$



Σχήμα 2.3. Συναρτήσεις πιθανότητας κατανομής Poisson-Lindley

2.4.4 Poisson - Beta κατανομή

Η περίπτωση της Beta κατανομής ως κατανομής της μεικτικής τυχαίας μεταβλητής Λ μελετήθηκε από τους Holla & Bhattacharya (1965). Σε αυτή την περίπτωση έχουμε τα ακόλουθα αποτελέσματα.

Πρόταση 2.15. (i) Η συνάρτηση πιθανότητας της μεικτικής κατανομής Poisson με μεικτική την τυχαία μεταβλητή $\Lambda \sim \text{Beta}(a, \beta)$ δίνεται από τον τύπο

$$f(x) = \frac{1}{x!} \frac{\Gamma(a + \beta)}{\Gamma(a)} \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \frac{\Gamma(x + a + n)}{\Gamma(x + a + n + \beta)}, \quad x = 0, 1, 2, \dots$$

(ii) Ισχύει ότι

$$P(s) = E(s^X) = \frac{1}{B(a, \beta)} {}_1F_1[a, a + \beta; -(1 - s)]$$

$$E(X) = \frac{a}{a + \beta}$$

$$Var(X) = \frac{a[(a + \beta)^2 + a + 2\beta]}{(a + \beta)^2(a + \beta + 1)}$$

$$\mu_3 = E[(X - E(X))^3]$$

$$= \frac{2a\beta(\beta - a)}{(a + \beta)^2(a + \beta + 1)(a + \beta + 2)} + 3 \frac{a\beta}{(a + \beta)^2(a + \beta + 1)} + \frac{a}{a + \beta}$$

$$\mu_4 = E[(X - E(X))^4]$$

$$= \frac{6a\beta[(a - \beta)^2(a + \beta + 1) - a\beta(a + \beta + 2)]}{(a + \beta)^4(a + \beta + 1)^2(a + \beta + 2)(a + \beta + 3)}$$

$$+ 6 \left[\frac{2a\beta(\beta^2 - a^2) + a^2\beta(a + \beta + 2)}{(a + \beta)^3(a + \beta + 1)(a + \beta + 2)} \right] + \frac{7a\beta + 3a^2(a + \beta + 1)}{(a + \beta)^2(a + \beta + 1)} + \frac{a}{a + \beta}$$

Απόδειξη. (i) Για $x = 0, 1, 2, \dots$ έχουμε

$$\begin{aligned} f(x) &= \int_0^1 \frac{e^{-\lambda} \lambda^x}{x!} g_\Lambda(\lambda) d\lambda = \int_0^1 \frac{e^{-\lambda} \lambda^x}{x!} \cdot \frac{\lambda^{a-1}(1-\lambda)^{\beta-1}}{B(a, \beta)} d\lambda \\ &= \frac{1}{B(a, \beta)x!} \int_0^1 \lambda^{x+a-1}(1-\lambda)^{\beta-1} \sum_{n=0}^{\infty} \frac{\lambda^n (-1)^n}{n!} d\lambda \\ &= \frac{1}{B(a, \beta)x!} \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \int_0^1 \lambda^{x+a+n-1}(1-\lambda)^{\beta-1} d\lambda \\ &= \frac{1}{B(a, \beta)x!} \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} B(x+a+n, \beta) \\ &= \frac{1}{x!} \frac{\Gamma(a+\beta)}{\Gamma(a)} \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \frac{\Gamma(x+a+n)}{\Gamma(x+a+n+\beta)}. \end{aligned}$$

(iii) Έχουμε ότι

$$G(s) = \sum_{x=0}^{\infty} s^x f(x) = \sum_{x=0}^{\infty} s^x \frac{1}{x!} \frac{\Gamma(a+\beta)}{\Gamma(a)} \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \frac{\Gamma(x+a+n)}{\Gamma(x+a+n+\beta)}$$

Μετά από αρκετές πράξεις προκύπτει ο τελικός τύπος

$$P(s) = \frac{1}{B(a, \beta)} {}_1F_1[a, a + \beta; -(1 - s)]$$

(για λεπτομέρειες δείτε Holla & Bhattacharya (1965)). Σημειώνουμε ότι

$${}_1F_1(a, b; s) = \sum_{k=0}^{\infty} \frac{a^{(k)} s^k}{b^{(k)} k!}$$

είναι η συρρέουσα (confluent) υπεργεωμετρική σειρά όπου $a^{(0)} = 1$ και $a^{(k)} = a(a+1)\cdots(a+k-1)$.

Επειδή η μεικτική τυχαία μεταβλητή $\Lambda \sim \text{Beta}(a, \beta)$, τότε από την Πρόταση 2.3 την Πρόταση 2.4 και την Παράγραφο 1.4.4 έχουμε τα εξής:

$$E(X) = E(\Lambda) = \frac{a}{a + \beta},$$

$$\text{Var}(X) = \text{Var}(\Lambda) + E(\Lambda) = \frac{a[(a + \beta)^2 + a + 2\beta]}{(a + \beta)^2(a + \beta + 1)},$$

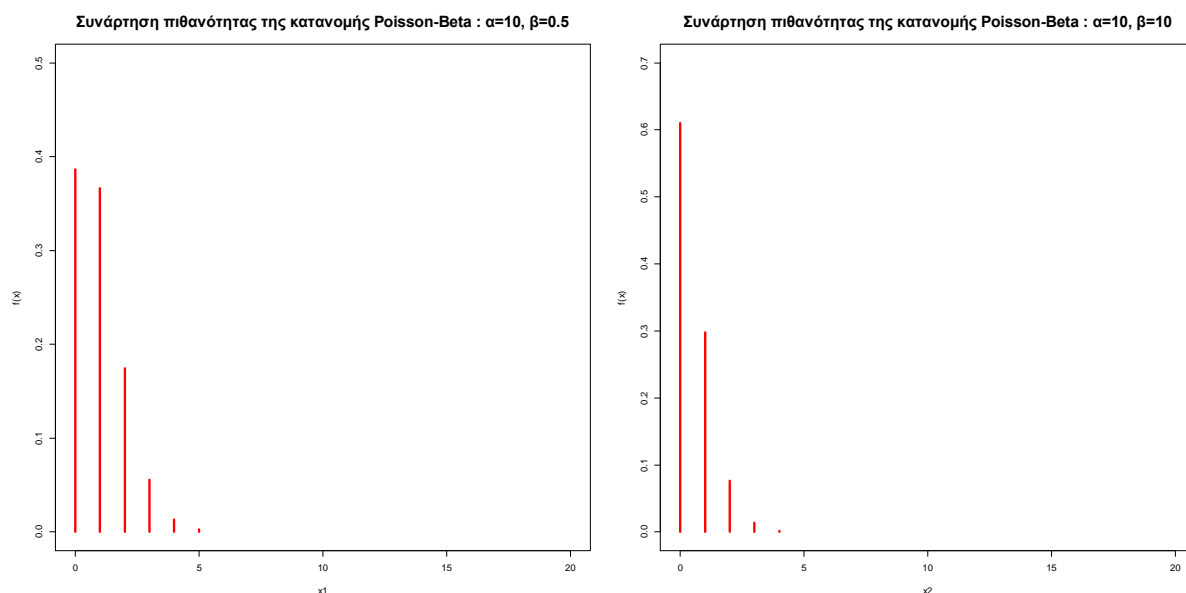
$$\mu_3 = E[(X - E(X))^3] = E[\Lambda - E(\Lambda)]^3 + 3\text{Var}(\Lambda) + E(\Lambda)$$

$$= \frac{2a\beta(\beta - a)}{(a + \beta)^2(a + \beta + 1)(a + \beta + 2)} + 3 \frac{a\beta}{(a + \beta)^2(a + \beta + 1)} + \frac{a}{a + \beta},$$

$$\mu_4 = E[(X - E(X))^4]$$

$$= \frac{6a\beta[(a - \beta)^2(a + \beta + 1) - a\beta(a + \beta + 2)]}{(a + \beta)^4(a + \beta + 1)^2(a + \beta + 2)(a + \beta + 3)}$$

$$+ 6 \left[\frac{2a\beta(\beta^2 - a^2) + a^2\beta(a + \beta + 2)}{(a + \beta)^3(a + \beta + 1)(a + \beta + 2)} \right] + \frac{7a\beta + 3a^2(a + \beta + 1)}{(a + \beta)^2(a + \beta + 1)} + \frac{a}{a + \beta}.$$



Σχήμα 2.4. Συναρτήσεις πιθανότητας κατανομής Poisson-Beta

2.4.5 Poisson - Pareto κατανομή

Η περίπτωση της Pareto κατανομής ως κατανομής της μεικτικής τυχαίας μεταβλητής Λ μελετήθηκε από τον Willmot (1993). Σε αυτή την περίπτωση έχουμε τα ακόλουθα αποτελέσματα.

Πρόταση 2.16. (i) Η συνάρτηση πιθανότητας της μεικτικής κατανομής Poisson με μεικτική την τυχαία μεταβλητή $\Lambda \sim P(a, \beta)$ δίνεται από τον τύπο

$$f(x) = \frac{a\beta^x e^{-\beta}}{x!} \psi(1, x - a + 1; \beta)$$

όπου

$$\psi(a, b; c) = \frac{1}{\Gamma(a)} \int_0^\infty y^{a-1} (1+y)^{b-a-1} e^{-cy} dy.$$

(ii) Ισχύει ότι

$$P(s) = E(s^X) = a e^{-\beta(1-s)} \psi(1, 1 - a; \beta(1 - s)),$$

$$E(X) = \frac{a\beta}{a-1},$$

$$\text{Var}(X) = \frac{a\beta(a^2 - 3a + 2 + \beta)}{(a-1)^2(a-2)},$$

$$\mu_3 = E[(X - E(X))^3] = \frac{2a(a+1)\beta^3}{(a-1)^3(a-2)(a-3)} + 3 \frac{\beta^2 a}{(a-1)^2(a-2)} + \frac{a\beta}{a-1},$$

$$\mu_4 = E[(X - E(X))^4]$$

$$= \frac{3a(3a^2 + a + 2)\beta^4}{(a-1)^4(a-2)(a-3)(a-4)}$$

$$+ 6 \left[\frac{2a(a+1)\beta^3}{(a-1)^3(a-2)(a-3)} + \frac{\beta^2 a}{(a-1)^2(a-2)} \frac{a\beta}{a-1} \right] + 7 \frac{\beta^2 a}{(a-1)^2(a-2)}$$

$$+ 3 \frac{a^2 \beta^2}{(a-1)^2} + \frac{a\beta}{a-1}.$$

Απόδειξη. (i) Για $x = 0, 1, 2, \dots$ έχουμε

$$f(x) = \frac{a\beta^a}{x!} \int_\beta^\infty e^{-\lambda} \lambda^{x-a-1} d\lambda.$$

Θέτοντας $z = \lambda - \beta$ παίρνουμε

$$f(x) = \frac{a\beta^a e^{-\beta}}{x!} \int_0^\infty (z + \beta)^{x-a-1} e^{-z} dz.$$

Κάνοντας επιπλέον την αντικατάσταση $z = \beta y$ παίρνουμε

$$\begin{aligned}
f(x) &= \frac{a\beta^a e^{-\beta}}{x!} \int_0^\infty \beta^{x-a-1} (1+y)^{x-a-1} e^{-\beta y} dy \\
&= \frac{a\beta^x e^{-\beta}}{x!} \int_0^\infty (1+y)^{x+a-1} e^{-\beta y} dy \\
&= \frac{a\beta^x e^{-\beta}}{x!} \psi(1, x-a+1; \beta).
\end{aligned}$$

(ii) Για την πιθανογεννήτρια της τυχαίας μεταβλητής X έχουμε

$$P(s) = \sum_{x=0}^{\infty} \int_{\beta}^{\infty} \frac{e^{-\lambda} (\lambda s)^x}{x!} \frac{a\beta^a}{\lambda^{a+1}} d\lambda = a\beta^a \int_{\beta}^{\infty} \lambda^{-a-1} e^{-\lambda(1-s)} d\lambda.$$

Θέτοντας $z = \lambda - \beta$ παίρνουμε

$$P(s) = a\beta^a e^{-\beta(1-s)} \int_0^\infty (z + \beta)^{-a-1} e^{-z(1-s)} dz$$

Κάνοντας επιπλέον την αντικατάσταση $z = \beta y$ παίρνουμε

$$\begin{aligned}
P(s) &= a\beta^a e^{-\beta(1-s)} \int_0^\infty \beta^{-a-1} (1+y)^{-a-1} e^{-(1-s)\beta y} \beta dy \\
&= a e^{-\beta(1-s)} \psi(1, 1-a; \beta(1-s)).
\end{aligned}$$

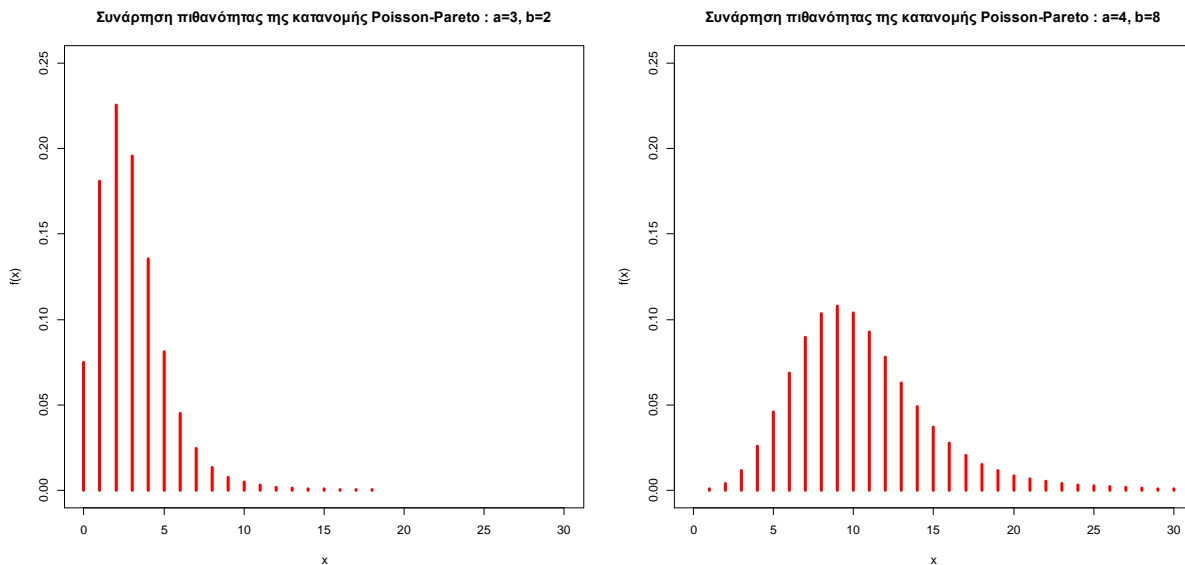
Επειδή η μεικτική τυχαία μεταβλητή $\Lambda \sim P(a, \beta)$, τότε από την Πρόταση 2.3 την Πρόταση 2.4 και την Παράγραφο 1.4.5 έχουμε τα εξής:

$$E(X) = E(\Lambda) = \frac{a\beta}{a-1},$$

$$Var(X) = Var(\Lambda) + E(\Lambda) = \frac{a\beta(a^2 - 3a + 2 + \beta)}{(a-1)^2(a-2)},$$

$$\mu_3 = E[(X - E(X))^3] = \frac{2a(a+1)\beta^3}{(a-1)^3(a-2)(a-3)} + 3 \frac{\beta^2 a}{(a-1)^2(a-2)} + \frac{a\beta}{a-1},$$

$$\begin{aligned}
\mu_4 = E[(X - E(X))^4] &= \frac{3a(3a^2 + a + 2)\beta^4}{(a-1)^4(a-2)(a-3)(a-4)} \\
&+ 6 \left[\frac{2a(a+1)\beta^3}{(a-1)^3(a-2)(a-3)} + \frac{\beta^2 a}{(a-1)^2(a-2)} \frac{a\beta}{a-1} \right] \\
&+ 7 \frac{\beta^2 a}{(a-1)^2(a-2)} + 3 \frac{a^2 \beta^2}{(a-1)^2} + \frac{a\beta}{a-1}.
\end{aligned}$$



Σχήμα 2.5. Συναρτήσεις πιθανότητας κατανομής Poisson-Pareto

2.4.6 Poisson - Inverse Gamma κατανομή

Η περίπτωση της αντίστροφης Γάμμα κατανομής ως κατανομής της μεικτικής τυχαίας μεταβλητής Λ μελετήθηκε από τον Willmot (1993). Σε αυτή την περίπτωση έχουμε τα ακόλουθα αποτελέσματα.

Πρόταση 2.17. (i) Η συνάρτηση πιθανότητας της μεικτικής κατανομής Poisson με μεικτική την τυχαία μεταβλητή $\Lambda \sim IGamma(a, \beta)$ δίνεται από τον τύπο

$$f(x) = \frac{2}{x!} \frac{\beta^{\frac{x+a}{2}}}{\Gamma(a)} K_{x-a}(2\sqrt{\beta})$$

όπου

$$K_\nu(t) = \frac{1}{2} \int_0^\infty x^{\nu-1} e^{-(t/2)(x+x^{-1})} dx$$

είναι η τροποποιημένη συνάρτηση Bessel τρίτου είδους.

(ii) Ισχύει ότι

$$P(s) = E(s^X) = \frac{2\beta^a}{\Gamma(a)} \left(\sqrt{\frac{\beta}{(1-s)}} \right)^{-a} K_{-a}(2\sqrt{\beta(1-s)}),$$

$$E(X) = \frac{\beta}{a-1}$$

$$\text{Var}(X) = \frac{\beta(\beta + (a-1)(a-2))}{(a-1)^2(a-2)}$$

$$\mu_3 = E[(X - E(X))^3]$$

$$= \frac{4\sqrt{a-2}(a-1)^2(a-2) + 3\beta^2(a-3) + \beta(a-1)(a-2)(a-3)}{(a-1)^2(a-2)(a-3)}$$

$$\mu_4 = E[(X - E(X))^4]$$

$$\begin{aligned} &= \frac{6(a-1)^3(a-2)(5a-11)}{(a-1)^3(a-2)(a-3)(a-4)} \\ &+ \frac{6(a-4)[4(a-1)^3(a-2)\sqrt{a-2} + \beta^3(a-3)]}{(a-1)^3(a-2)(a-3)(a-4)} \\ &+ \frac{(a-1)(a-3)(a-4)[7\beta^2 + 3\beta^2(a-2)]}{(a-1)^3(a-2)(a-3)(a-4)} \\ &+ \frac{\beta(a-1)^2(a-2)(a-3)(a-4)}{(a-1)^3(a-2)(a-3)(a-4)}. \end{aligned}$$

Απόδειξη. (i) Για $x = 0, 1, 2, \dots$ έχουμε

$$\begin{aligned} f(x) &= \int_0^\infty \frac{e^{-\lambda} \lambda^x}{x!} g_\lambda(\lambda) d\lambda = \int_0^\infty \frac{e^{-\lambda} \lambda^x}{x!} \frac{\beta^a}{\Gamma(a)} \lambda^{-a-1} e^{-\frac{\beta}{\lambda}} d\lambda \\ &= \frac{\beta^a}{x! \Gamma(a)} \int_0^\infty \lambda^{x-a-1} e^{-(\lambda + \frac{\beta}{\lambda})} d\lambda. \end{aligned}$$

Θέτοντας $\lambda = \sqrt{\beta}z$, οπότε $d\lambda = dz\sqrt{\beta}$, παίρνουμε

$$\begin{aligned} f(x) &= \frac{1}{x!} \frac{\beta^a}{\Gamma(a)} \int_0^\infty (\sqrt{\beta}z)^{x-a-1} e^{-\sqrt{\beta}(z + \frac{1}{z})} dz \sqrt{\beta} \\ &= \frac{1}{x!} \frac{\beta^a}{\Gamma(a)} (\sqrt{\beta})^{x-a} \int_0^\infty z^{x-a-1} e^{-\frac{2\sqrt{\beta}}{2}(z + \frac{1}{z})} dz \\ &= \frac{2}{x!} \frac{\beta^{\frac{x+a}{2}}}{\Gamma(a)} K_{x-a}(2\sqrt{\beta}). \end{aligned}$$

(ii) Για την πιθανογεννήτρια τη τυχαίας μεταβλητής X έχουμε

$$P(s) = \frac{\beta^a}{\Gamma(a)} \int_0^\infty \lambda^{-a-1} e^{-\lambda(1-s) - \beta/\lambda} d\lambda$$

$$= \frac{\beta^a}{\Gamma(a)} \int_0^{\infty} \lambda^{-a-1} e^{-(1-s)\left[\lambda + \frac{\beta}{(1-s)\lambda}\right]} d\lambda$$

Για $\lambda = z \sqrt{\frac{\beta}{(1-s)}}$, οπότε $d\lambda = dz \sqrt{\frac{\beta}{(1-s)}}$, παίρνουμε

$$\begin{aligned} P(s) &= \frac{\beta^a}{\Gamma(a)} \left(\sqrt{\frac{\beta}{(1-s)}} \right)^{-a} \int_0^{\infty} z^{-a-1} e^{-\frac{2\sqrt{\beta(1-s)}}{2}\left(z + \frac{1}{z}\right)} dz \\ &= \frac{2\beta^a}{\Gamma(a)} \left(\sqrt{\frac{\beta}{(1-s)}} \right)^{-a} K_{-a}(2\sqrt{\beta(1-s)}). \end{aligned}$$

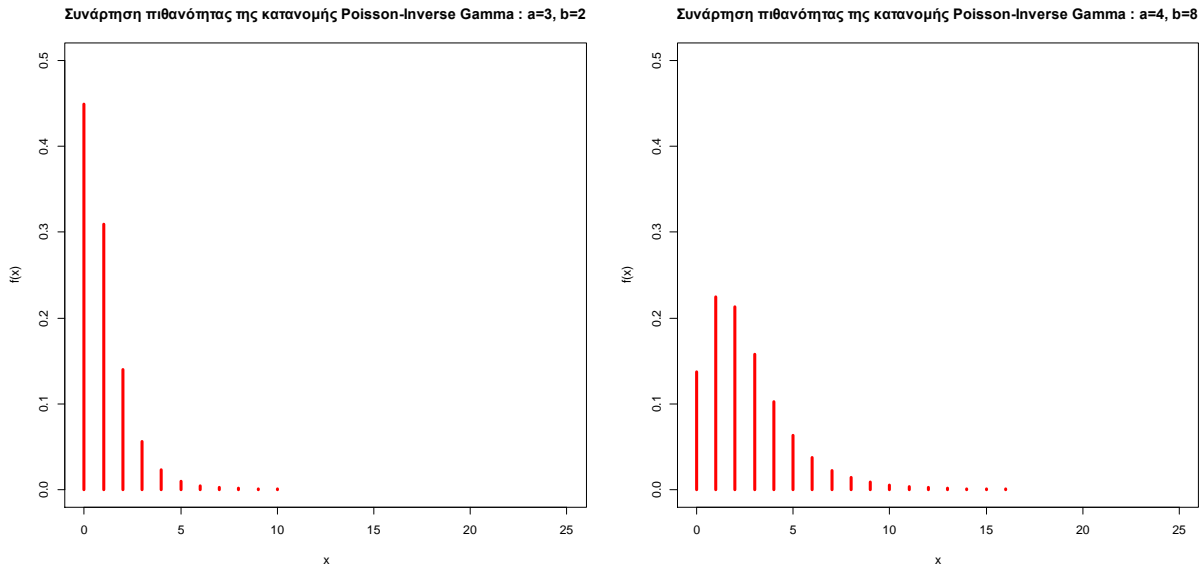
Επειδή η μεικτική τυχαία μεταβλητή $\Lambda \sim IGamma(a, \beta)$ τότε από την Πρόταση 2.3 την Πρόταση 2.4 και την Παράγραφο 1.4.6 έχουμε τα εξής:

$$E(X) = E(\Lambda) = \frac{\beta}{a-1},$$

$$Var(X) = Var(\Lambda) + E(\Lambda) = \frac{\beta(\beta + (a-1)(a-2))}{(a-1)^2(a-2)},$$

$$\begin{aligned} \mu_3 &= E[(X - E(X))^3] = E[\Lambda - E(\Lambda)]^3 + 3Var(\Lambda) + E(\Lambda) \\ &= \frac{4\sqrt{a-2}(a-1)^2(a-2) + 3\beta^2(a-3) + \beta(a-1)(a-2)(a-3)}{(a-1)^2(a-2)(a-3)}, \end{aligned}$$

$$\begin{aligned} \mu_4 &= E[(X - E(X))^4] = E[\Lambda - E(\Lambda)]^4 + 6\{E[\Lambda - E(\Lambda)]^3 + Var(\Lambda)E(\Lambda)\} \\ &\quad + \{7Var(\Lambda) + 3[E(\Lambda)]^2\} + E(\Lambda) \\ &= \frac{6(a-1)^3(a-2)(5a-11)}{(a-1)^3(a-2)(a-3)(a-4)} + \frac{6(a-4)[4(a-1)^3(a-2)\sqrt{a-2} + \beta^3(a-3)]}{(a-1)^3(a-2)(a-3)(a-4)} \\ &\quad + \frac{(a-1)(a-3)(a-4)[7\beta^2 + 3\beta^2(a-2)]}{(a-1)^3(a-2)(a-3)(a-4)} \\ &\quad + \frac{\beta(a-1)^2(a-2)(a-3)(a-4)}{(a-1)^3(a-2)(a-3)(a-4)}. \end{aligned}$$



Σχήμα 2.6. Συναρτήσεις πιθανότητας κατανομής Poisson-Inverse Gamma

2.4.7 Poisson - Inverse Gaussian κατανομή

Η περίπτωση της Poisson - Inverse Gaussian κατανομής μελετήθηκε από τον Holla (1966) (δείτε, επίσης, Shaban (1981), Zha et al. (2014)). Στη συνέχεια δίνουμε τη συνάρτηση πιθανότητας, την πιθανογεννήτρια, τη μέση τιμή και τη διακύμανση της Poisson - Inverse Gaussian κατανομής, καθώς επίσης και ένα ενδεικτικό σχήμα για τη συνάρτηση πιθανότητάς της.

Πρόταση 2.18. (i) Η συνάρτηση πιθανότητας της μεικτής κατανομής Poisson με μεικτική την τυχαία μεταβλητή $A \sim IG(\mu, \theta)$ δίνεται από τον τύπο

$$f(x) = \sqrt{\frac{2\theta}{\pi}} \frac{e^{(\theta/\mu)}}{x!} \left(\frac{2}{\theta} \sqrt{1 + \frac{\theta}{2\mu^2}} \right)^{-(x-\frac{1}{2})} K_{x-\frac{1}{2}} \left(\sqrt{2\theta \left(1 + \frac{\theta}{2\mu^2} \right)} \right)$$

όπου

$$K_\nu(t) = \frac{1}{2} \int_0^\infty x^{\nu-1} e^{-(t/2)(x+x^{-1})} dx$$

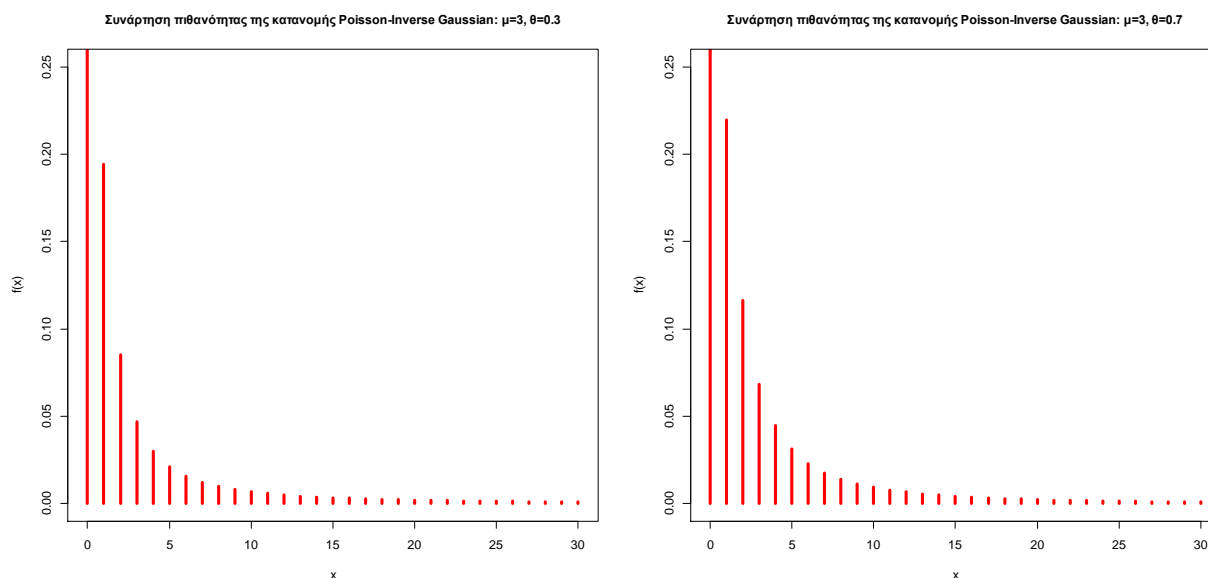
είναι η τροποποιημένη συνάρτηση Bessel τρίτου είδους.

(ii) Ισχύει ότι

$$P(s) = \exp\left(\frac{\theta}{\mu} - \frac{\theta}{\mu} \sqrt{1 + \frac{2\mu^2}{\theta}(1-s)}\right),$$

$$E(X) = \mu,$$

$$Var(X) = \frac{\mu^3}{\theta} + \mu.$$



Σχήμα 2.7. Συναρτήσεις πιθανότητας κατανομής Poisson-Inverse Gaussian

2.4.8 Poisson - Lognormal κατανομή

Η περίπτωση της Poisson-Lognormal κατανομής μελετήθηκε από τον Bulmer (1974) (δείτε, επίσης, Alonso et al. (2008), Engen et al. (2002)). Στη συνέχεια δίνουμε τη συνάρτηση πιθανότητας, μέση τιμή και διακύμανση της Poisson-Lognormal κατανομής, καθώς επίσης και ένα ενδεικτικό σχήμα για τη συνάρτηση πιθανότητάς της.

Πρόταση 2.19. (i) Η συνάρτηση πιθανότητας της μεικτής κατανομής Poisson με μεικτική την τυχαία μεταβλητή $\Lambda \sim LN(\mu, \sigma)$ δίνεται από τον τύπο

$$f(x) = \frac{e^{x\mu + x^2\sigma/2} (2\pi\sigma)^{-1/2}}{x!} g(y)$$

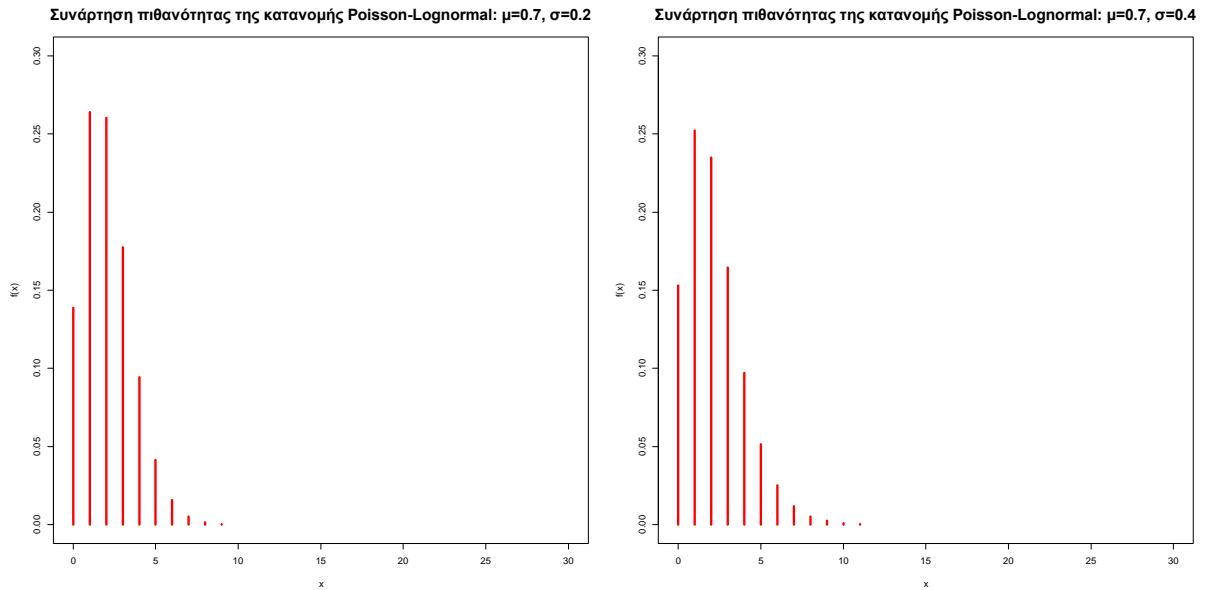
όπου

$$g(y) = \int_{-\infty}^{\infty} e^{-ey} \frac{e^{(-y-\mu-x\sigma)^2}}{2\sigma} dy.$$

(ii) Ισχύει ότι

$$E(X) = \exp\left(\mu + \frac{\sigma^2}{2}\right),$$

$$\text{Var}(X) = \exp\left(\mu + \frac{\sigma^2}{2}\right) + (\exp(\sigma^2) - 1) \exp(\sigma^2 + 2\mu).$$



Σχήμα 2.8. Συναρτήσεις πιθανότητας κατανομής Poisson-Lognormal

ΚΕΦΑΛΑΙΟ 3

Εφαρμογή μεικτών κατανομών Poisson σε πραγματικά δεδομένα

3.1 Εισαγωγή

Στο παρόν κεφάλαιο θα προσαρμόσουμε μεικτές κατανομές Poisson σε δυο σύνολα δεδομένων. Για την ανάλυσή μας θα χρησιμοποιήσουμε την γλώσσα προγραμματισμού R (δείτε Αντζουλάκος (2018)). Το πιο σημαντικό εργαλείο που θα χρησιμοποιήσουμε από την R είναι το πακέτο `fitdistrplus`. Στην αρχή θα δούμε κάποια βασικά περιγραφικά μέτρα των δεδομένων ώστε να λάβουμε μια σαφή εικόνα του εκάστοτε δείγματος και στην συνέχεια θα προσαρμόσουμε στο εκάστοτε δείγμα τις μεικτές κατανομές Poisson, με απώτερο σκοπό να παρατηρήσουμε πόσο καλή είναι η εφαρμογή τους και ποια κατανομή από αυτές εφαρμόζει καλύτερα στο κάθε ένα από τα σύνολα δεδομένων μας.

Οι εκτιμήσεις των παραμέτρων των κατανομών που προσαρμόζονται στα δεδομένα γίνεται με τη μέθοδο των εκτιμητών μέγιστης πιθανοφάνειας.

Το πρώτο σύνολο δεδομένων που θα μας απασχολήσει είναι το σύνολο δεδομένων Crimes που αναφέρεται στο μηνιαίο αριθμό εγκλημάτων από τον Ιανουάριο του 1982 έως και τον Ιανουάριο του 1993 στην Ελλάδα (Karlis (2005)). Το δεύτερο σύνολο δεδομένων που θα μας απασχολήσει είναι το ένα σύνολο δεδομένων που παρουσίασε ο Lemaire (δείτε Lemaire & Gossiaux (1981)) από τις έξι ομάδες δεδομένων που ανέλυσαν αναφορικά με τον αριθμό των απαιτήσεων που εγείρουν οι ασφαλισμένοι του κλάδου αυτοκινήτου ανά ασφαλιστικό συμβόλαιο.

3.2 Ανάλυση των δεδομένων Crimes

Τα δεδομένα Crimes δίνονται στον ακόλουθο πίνακα.

Πίνακας 3.1. Δεδομένα Crimes

Τιμές	0	1	2	3	4	5	6	7	8	9
Συχνότητα	21	41	32	16	19	8	4	1	2	1

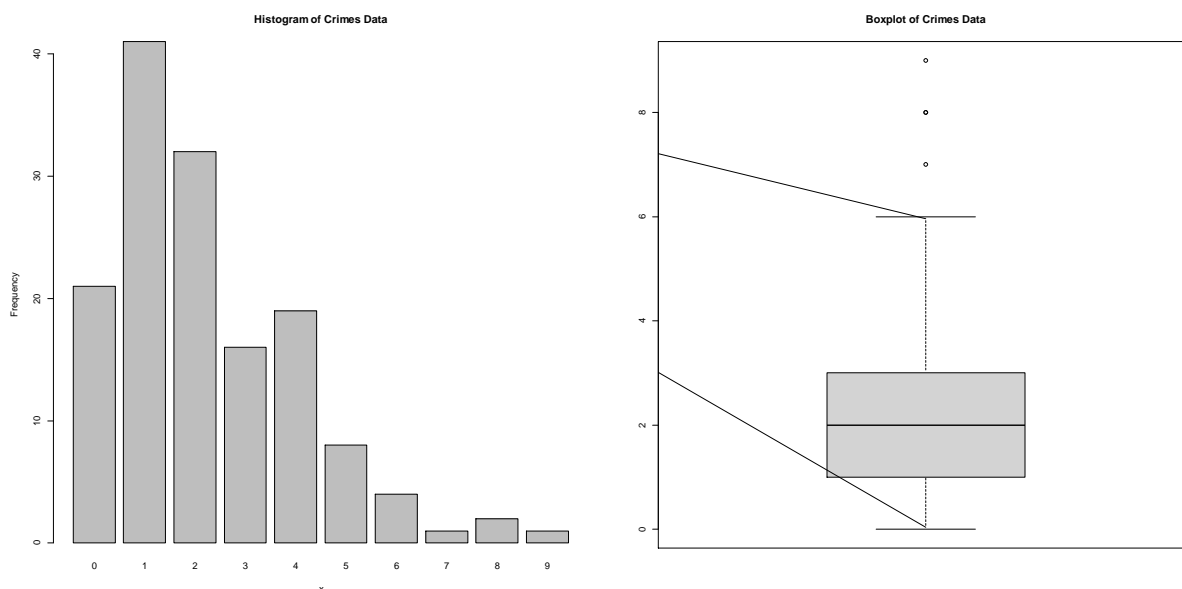
Ο Πίνακας 3.2 δίνει κάποια βασικά περιγραφικά μέτρα των δεδομένων.

Πίνακας 3.2. Περιγραφικά μέτρα για τα δεδομένα Crimes

Περιγραφικά μέτρα	Τιμή
Ελάχιστο	0
Μέγιστο	9
Μέση τιμή	2.241
Διάμεσος	2
1 ^ο Τεταρτημόριο	1
3 ^ο Τεταρτημόριο	3
Διακύμανση	3.407
Τυπική Απόκλιση	1.8457
Συντελεστής ασυμμετρίας	1.079
Συντελεστής κύρτωσης	4.142

Από τον Πίνακα 3.2 προκύπτει ότι η διακύμανση είναι μεγαλύτερη της μέσης τιμής γεγονός που υποδεικνύει την πιθανή ακαταλληλότητα της κατανομής Poisson να περιγράψει ικανοποιητικά τα δεδομένα, κάνοντας έτσι πιο πιθανή την καλή προσαρμογή στα δεδομένα μιας μεικτής κατανομής Poisson. Στο Σχήμα 3.1 δίνουμε ραβδόγραμμα και θηκόγραμμα για τα δεδομένα Crimes όπου φαίνεται ξεκάθαρα ότι πρόκειται για μια κατανομή λοξή προς τα δεξιά (θετική ασυμμετρία).

Στη συνέχεια θα προσαρμόσουμε στα δεδομένα Crimes μια σειρά από μεικτές κατανομές Poisson και θα διαπιστώσουμε ότι κάποιες από αυτές έχουν καλή προσαρμογή στα δεδομένα έναντι της αρχικής επιλογής για μια απλή κατανομή Poisson. Οι κώδικες που χρησιμοποιήθηκαν δίνονται στο Παράρτημα.



Σχήμα 3.1. Ραβδόγραμμα και θηκόγραμμα για τα δεδομένα Crimes

3.2.1 Προσαρμογή κατανομής Poisson

Ο κάτωθι πίνακας δίνει τα βασικά αποτελέσματα της προσαρμογής της κατανομής $P(\lambda)$ στα δεδομένα Crimes.

Πίνακας 3.3. Βασικά χαρακτηριστικά της προσαρμογής της κατανομής Poisson στα δεδομένα Crimes

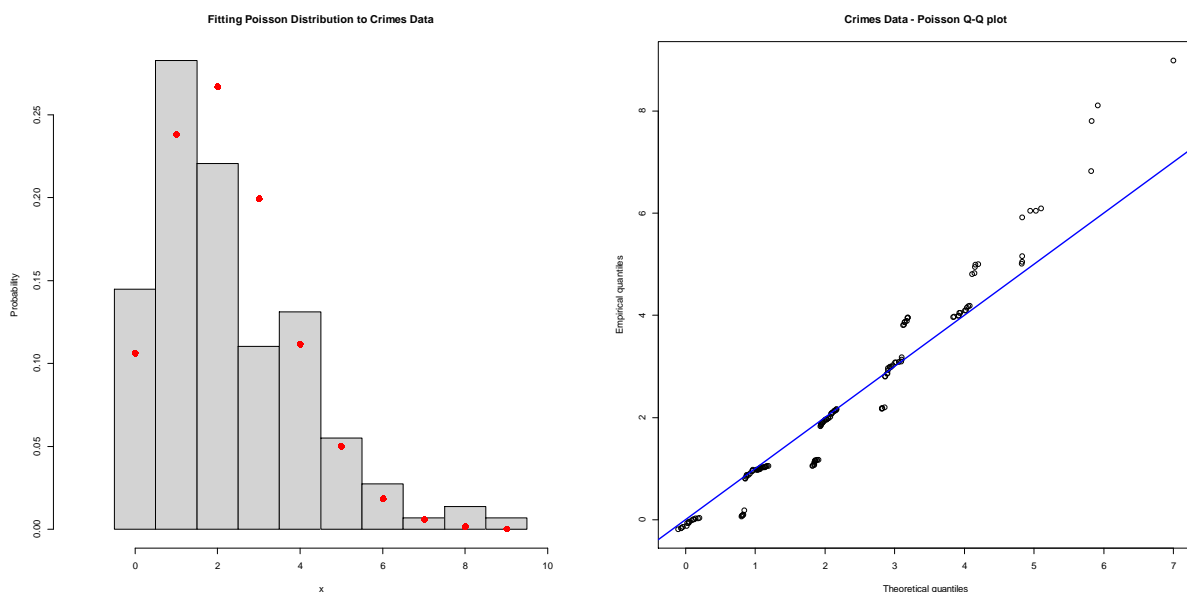
Παράμετρος λ	2.241379
Loglikelihood	-281.0803
AIC	564.1606
p – value χ^2 ελέγχου	0.01262442
Στατιστική χ^2 ελέγχου	12.73894
Βαθμοί ελευθερίας	4

Από τον παραπάνω πίνακα προκύπτει ότι η κατανομή Poisson δεν φαίνεται να έχει καλή προσαρμογή στα δεδομένα αφού το p -value του χ^2 ελέγχου είναι μικρότερο του 0.05 (τιμή του συνήθους επιπέδου σημαντικότητας). Σημειώνουμε ότι για την εφαρμογή του χ^2 ελέγχου οι τιμές ομαδοποιούνται έτσι ώστε να προκύπτει αναμενόμενη συχνότητα μεγαλύτερη ή ίση του 5. Οι αναμενόμενες συχνότητες δίνονται στον ακόλουθο πίνακα.

Πίνακας 3.4. Πιθανότητες και αναμενόμενες συχνότητες της κατανομής Poisson

X	Πιθανότητα	Αναμενόμενες Συχνότητες	Παρατηρηθείσες Συχνότητες
X=0	0.1063	15.4152	21
X=1	0.2383	34.5513	41
X=2	0.2670	38.7213	32
X=3	0.1995	28.9297	16
X=4	0.1118	16.2106	19
X>=5	0.0770	11.1718	16

Στο ακόλουθο σχήμα δίνεται το ιστόγραμμα των εμπειρικών πιθανοτήτων μαζί με τις θεωρητικές πιθανότητες (κόκκινες κουκίδες) της προσαρμοσμένης κατανομής Poisson, καθώς επίσης και το Poisson Q-Q διάγραμμα. Από το Poisson Q-Q διάγραμμα διαπιστώνουμε τη φτωχή εφαρμογή της κατανομής Poisson στα δεδομένα.



Σχήμα 3.2. Ιστόγραμμα εμπειρικών/θεωρητικών πιθανοτήτων και Poisson Q-Q διάγραμμα για τα δεδομένα Crimes

3.2.2 Προσαρμογή κατανομής Poisson-Gamma

Ο κάτωθι πίνακας δίνει τα βασικά αποτελέσματα της προσαρμογής της μεικτής

κατανομής Poisson με μεικτική την τυχαία μεταβλητή $L \sim G(a, \beta)$ (δείτε Παράγραφο 2.4.1) στα δεδομένα Crimes.

Πίνακας 3.5. Βασικά χαρακτηριστικά της προσαρμογής της κατανομής Poisson-Gamma στα δεδομένα Crimes

Παράμετρος a	4.494424
Παράμετρος β	2.005007
Loglikelihood	-274.5055
AIC	553.011
p – value χ^2 ελέγχου	0.2974202
Στατιστική χ^2 ελέγχου	4.90277
Βαθμοί ελευθερίας	4

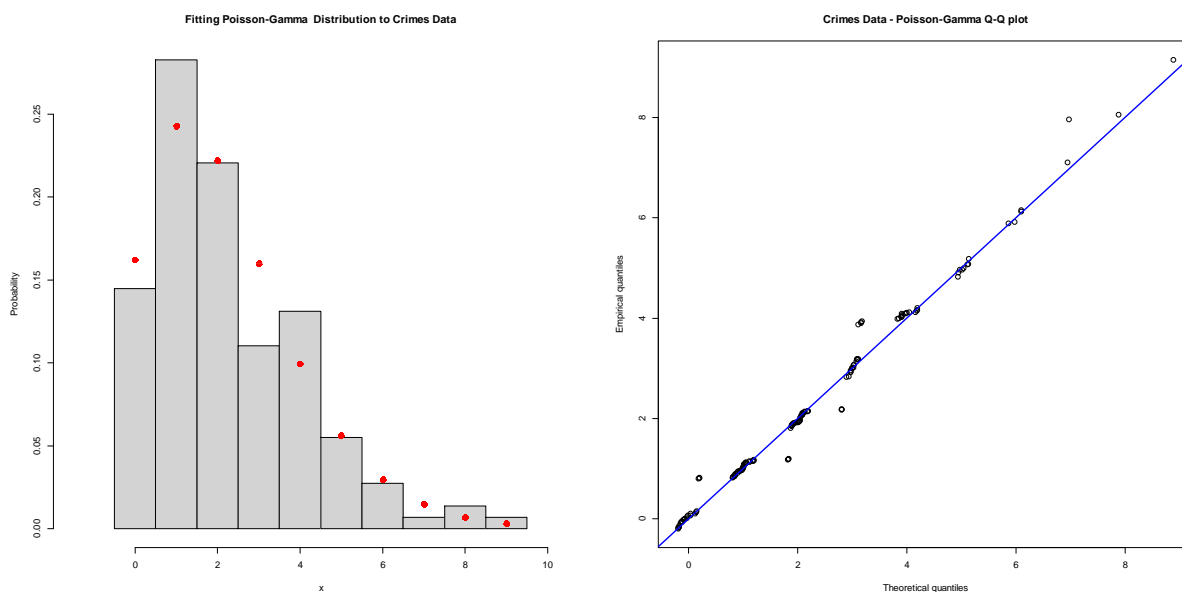
Σημειώνουμε ότι η προσαρμοσμένη κατανομή Poisson-Gamma είναι ως γνωστόν η αρνητική διωνυμική κατανομή $NB_0(r, p)$ με παραμέτρους $r = a = 4.494424$ και $p = \frac{1}{1+\beta} = 0.667222$. Από τον παραπάνω πίνακα προκύπτει ότι η κατανομή Poisson-Gamma φαίνεται να έχει καλή προσαρμογή στα δεδομένα αφού το p -value του χ^2 ελέγχου δεν είναι μικρότερο του 0.05 (τιμή του συνήθους επιπέδου σημαντικότητας). Οι αναμενόμενες συχνότητες δίνονται στον ακόλουθο πίνακα.

Πίνακας 3.6. Πιθανότητες και αναμενόμενες συχνότητες της κατανομής Poisson-Gamma

X	Πιθανότητα	Αναμενόμενες Συχνότητες	Παρατηρηθείσες Συχνότητες
X=0	0.1623	23.5269	21
X=1	0.2427	35.1879	41
X=2	0.2219	32.1692	32
X=3	0.1598	23.1747	16
X=4	0.0997	14.4493	19

$X=5$	0.0563	8.1689	8
$X \geq 6$	0.0574	8.3231	8

Στο ακόλουθο σχήμα δίνεται το ιστόγραμμα των εμπειρικών πιθανοτήτων μαζί με τις θεωρητικές πιθανότητες (κόκκινες κουκίδες) του προσαρμοσμένου μοντέλου, καθώς επίσης και το Poisson-Gamma Q-Q διάγραμμα. Από το Q-Q διάγραμμα διαπιστώνουμε την καλή εφαρμογή της κατανομής Poisson-Gamma στα δεδομένα.



Σχήμα 3.3. Ιστόγραμμα εμπειρικών/θεωρητικών πιθανοτήτων και Poisson-Gamma Q-Q διάγραμμα για τα δεδομένα Crimes

3.2.3 Προσαρμογή κατανομής Poisson-Lindley

Ο κάτωθι πίνακας δίνει τα βασικά αποτελέσματα της προσαρμογής της μεικτής κατανομής Poisson με μεικτική την τυχαία μεταβλητή $\Lambda \sim Lindley(\theta)$ (δείτε Παράγραφο 2.4.3) στα δεδομένα Crimes.

Πίνακας 3.7. Βασικά χαρακτηριστικά της προσαρμογής της κατανομής Poisson-Lindley στα δεδομένα Crimes

Παράμετρος θ	0.7012227
Loglikelihood	-284.2483
AIC	570.4965

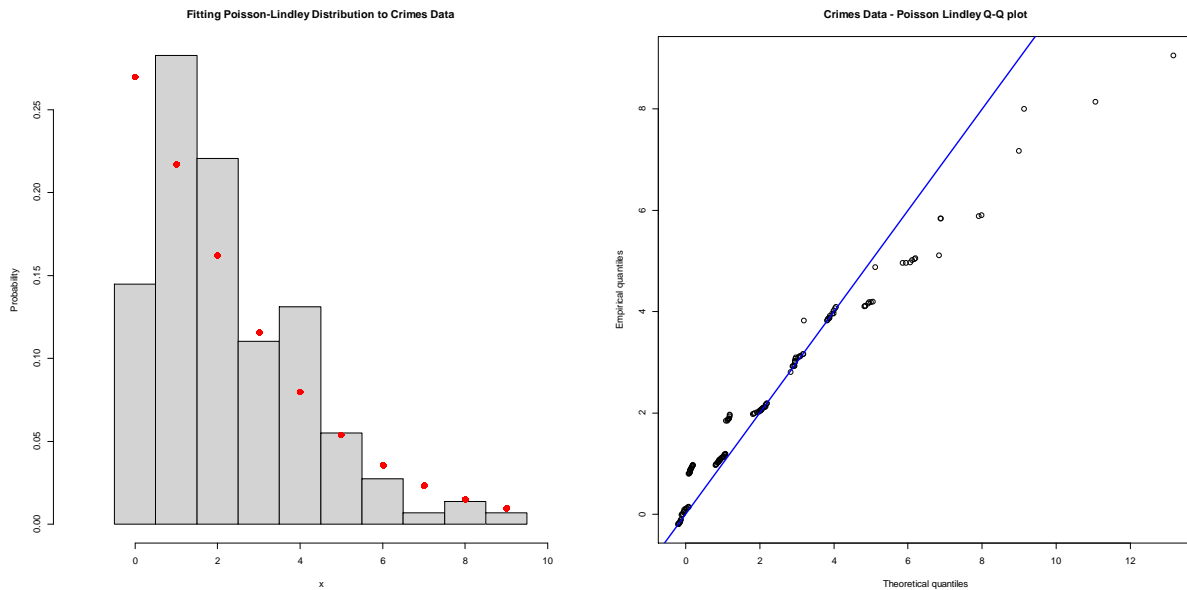
p – value χ^2 ελέγχου	0.0009683159
Στατιστική χ^2 ελέγχου	22.53461
Βαθμοί ελευθερίας	6

Από τον παραπάνω πίνακα προκύπτει ότι η κατανομή Poisson-Lindley δεν φαίνεται να έχει καλή προσαρμογή στα δεδομένα αφού το p -value του χ^2 ελέγχου είναι πολύ μικρότερο του 0.05. Οι αναμενόμενες συχνότητες δίνονται στον ακόλουθο πίνακα.

Πίνακας 3.8. Πιθανότητες και αναμενόμενες συχνότητες της κατανομής Poisson-Lindley

X	Πιθανότητα	Αναμενόμενες Συχνότητες	Παρατηρηθείσες Συχνότητες
X=0	0.2698	39.1162	21
X=1	0.2173	31.5051	41
X=2	0.1622	23.5226	32
X=3	0.1156	16.7680	16
X=4	0.0799	11.5853	19
X=5	0.0540	7.8262	8
X=6	0.0358	5.1977	4
X>=7	0.0654	9.4790	4

Στο ακόλουθο σχήμα δίνεται το ιστόγραμμα των εμπειρικών πιθανοτήτων μαζί με τις θεωρητικές πιθανότητες (κόκκινες κουκίδες) του προσαρμοσμένου μοντέλου, καθώς επίσης και το Poisson-Lindley Q-Q διάγραμμα με το οποίο διαπιστώνεται η φτωχή εφαρμογή της κατανομής Poisson-Lindley στα δεδομένα.



Σχήμα 3.4. Ιστόγραμμα εμπειρικών/θεωρητικών πιθανοτήτων και Poisson-Lindley Q-Q διάγραμμα για τα δεδομένα Crimes

3.2.4 Προσαρμογή κατανομής Poisson-Inverse Gaussian

Ο κάτωθι πίνακας δίνει τα βασικά αποτελέσματα της προσαρμογής της μεικτής κατανομής Poisson με μεικτική την τυχαία μεταβλητή $\Lambda \sim IG(\mu, \theta)$ (δείτε Παράγραφο 2.4.7) στα δεδομένα Crimes.

Πίνακας 3.9. Βασικά χαρακτηριστικά της προσαρμογής της κατανομής Poisson-Inverse Gaussian στα δεδομένα Crimes

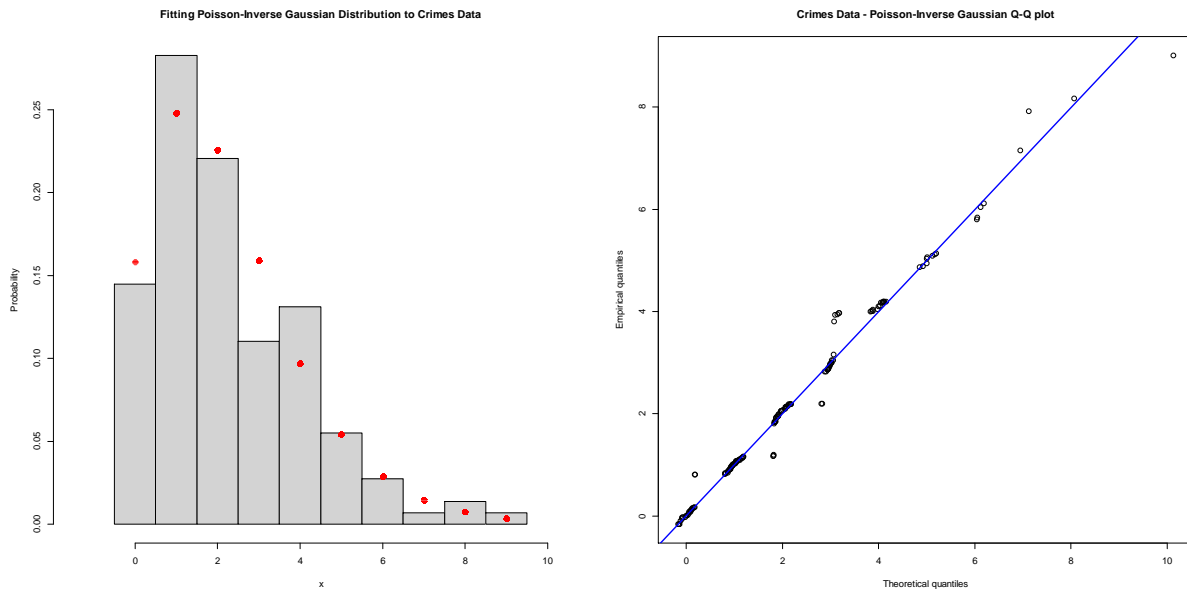
Παράμετρος μ	2.241187
Παράμετρος θ	9.574963
Loglikelihood	-274.4575
AIC	552.9151
p – value χ^2 ελέγχου	0.3070627
Στατιστική χ^2 ελέγχου	4.812685
Βαθμοί ελευθερίας	4

Από τον παραπάνω πίνακα προκύπτει ότι η κατανομή Poisson- Inverse Gaussian φαίνεται να έχει καλή προσαρμογή στα δεδομένα αφού το p-value του χ^2 ελέγχου είναι πολύ μεγαλύτερο του 0.05. Οι αναμενόμενες συχνότητες δίνονται στον ακόλουθο πίνακα.

Πίνακας 3.10. Πιθανότητες και αναμενόμενες συχνότητες της κατανομής Poisson-Inverse Gaussian

X	Πιθανότητα	Αναμενόμενες Συχνότητες	Παρατηρηθείσες Συχνότητες
X=0	0.1583	22.9489	21
X=1	0.2478	35.9294	41
X=2	0.2257	32.7250	32
X=3	0.1590	23.0559	16
X=4	0.0970	14.0625	19
X=5	0.0542	7.8657	8
X>=6	0.0580	8.4127	8

Στο ακόλουθο σχήμα δίνεται το ιστόγραμμα των εμπειρικών πιθανοτήτων μαζί με τις θεωρητικές πιθανότητες (κόκκινες κουκίδες) του προσαρμοσμένου μοντέλου, καθώς επίσης και το Poisson-Inverse Gaussian Q-Q διάγραμμα με το οποίο διαπιστώνεται η αρκετά καλή εφαρμογή της κατανομής Poisson-Inverse Gaussian στα δεδομένα.



Σχήμα 3.5. Ιστόγραμμα εμπειρικών/θεωρητικών πιθανοτήτων και Poisson-Inverse Gaussian Q-Q διάγραμμα για τα δεδομένα Crimes

3.2.5 Προσαρμογή κατανομής Poisson-Lognormal

Ο κάτωθι πίνακας δίνει τα βασικά αποτελέσματα της προσαρμογής της μεικτής κατανομής Poisson με μεικτική την τυχαία μεταβλητή $\Lambda \sim LN(\mu, \sigma)$ (δείτε Παράγραφο 2.4.8) στα δεδομένα Crimes.

Πίνακας 3.11. Βασικά χαρακτηριστικά της προσαρμογής της κατανομής Poisson-Lognormal στα δεδομένα Crimes

Παράμετρος μ	0.7021677
Παράμετρος σ	0.4584693
Loglikelihood	-274.4967
AIC	552.9935
p – value χ^2 ελέγχου	0.3046507
Στατιστική χ^2 ελέγχου	4.834996
Βαθμοί ελευθερίας	4

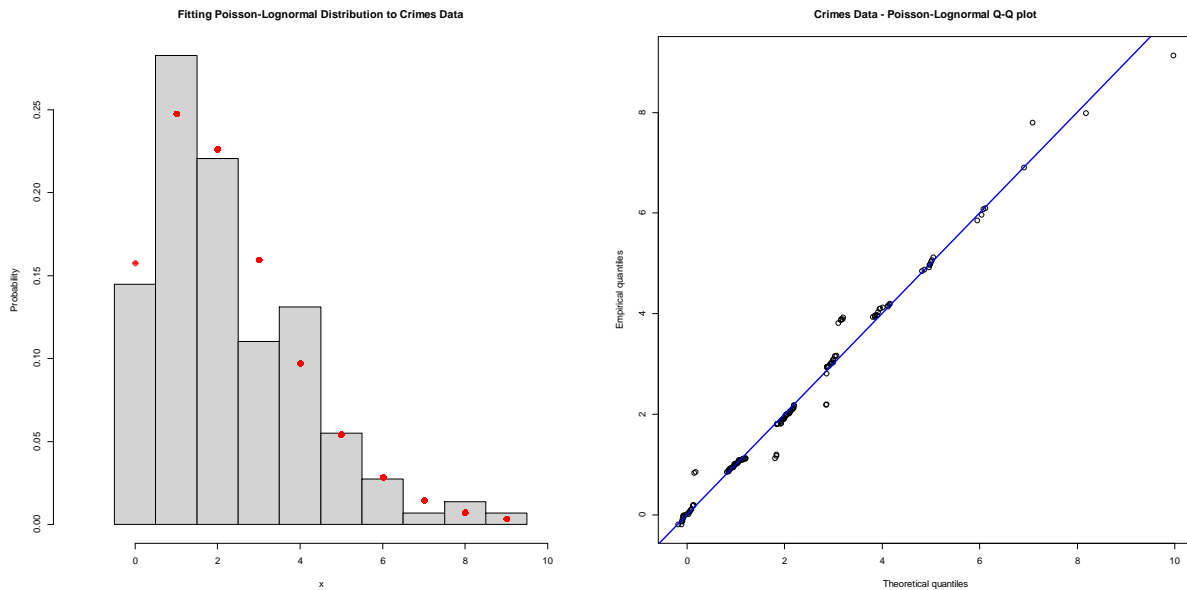
Από τον παραπάνω πίνακα προκύπτει ότι η κατανομή Poisson-Lognormal φαίνεται να έχει καλή προσαρμογή στα δεδομένα αφού το p -value του χ^2 ελέγχου είναι πολύ

μεγαλύτερο του 0.05. Οι αναμενόμενες συχνότητες σύμφωνα με το μοντέλο Poisson-Lognormal δίνονται στον ακόλουθο πίνακα.

Πίνακας 3.12. Πιθανότητες και αναμενόμενες συχνότητες της κατανομής Poisson-Lognormal

X	Πιθανότητα	Αναμενόμενες Συχνότητες	Παρατηρηθείσες Συχνότητες
X=0	0.1577	22.8691	21
X=1	0.2475	35.8905	41
X=2	0.2263	32.8092	32
X=3	0.1596	23.1482	16
X=4	0.0972	14.0913	19
X=5	0.0541	7.8446	8
X>=6	0.0576	8.3472	8

Στο ακόλουθο σχήμα δίνεται το ιστόγραμμα των εμπειρικών πιθανοτήτων μαζί με τις θεωρητικές πιθανότητες (κόκκινες κουκίδες) του προσαρμοσμένου μοντέλου, καθώς επίσης και το Poisson-Lognormal Q-Q διάγραμμα από το οποίο διαπιστώνεται η αρκετά καλή εφαρμογή της κατανομής Poisson-Lognormal στα δεδομένα.



Σχήμα 3.6. Ιστόγραμμα εμπειρικών/θεωρητικών πιθανοτήτων και Poisson-Lognormal Q-Q διάγραμμα για τα δεδομένα Crimes

3.3 Ανάλυση των δεδομένων Lemaire

Τα δεδομένα Lemaire δίνονται στον ακόλουθο πίνακα.

Πίνακας 3.13. Δεδομένα Lemaire

Τιμές	0	1	2	3	4	5	6	7
Συχνότητα	7840	1317	239	42	14	4	4	1

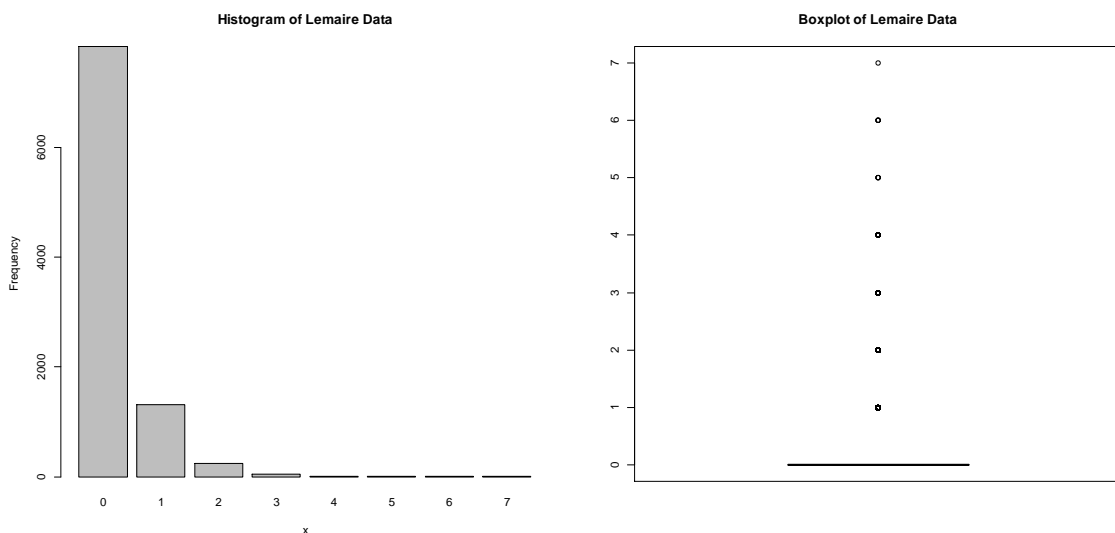
Ο Πίνακας 3.14 δίνει κάποια βασικά περιγραφικά μέτρα των δεδομένων.

Πίνακας 3.14. Περιγραφικά μέτρα για τα δεδομένα Lemaire

Περιγραφικά μέτρα	Τιμή
Ελάχιστο	0
Μέγιστο	7
Μέση τιμή	0.214
Διάμεσος	0
1 ^ο Τεταρτημόριο	0

3 ^ο Τεταρτημόριο	0
Διακύμανση	0.289
Τυπική Απόκλιση	0.538
Συντελεστής ασυμμετρίας	3.482
Συντελεστής κύρτωσης	21.595

Από τον Πίνακα 3.14 προκύπτει ότι η διακύμανση είναι μεγαλύτερη της μέσης τιμής γεγονός που υποδεικνύει την πιθανή ακαταλληλότητα της κατανομής Poisson να περιγράψει ικανοποιητικά τα δεδομένα, κάνοντας έτσι πιο πιθανή την καλή προσαρμογή στα δεδομένα μιας μεικτής κατανομής Poisson. Στο Σχήμα 3.7 δίνουμε το ραβδόγραμμα και το θηκόγραμμα για τα δεδομένα Lemaire όπου φαίνεται ξεκάθαρα ότι πρόκειται για μια κατανομή με έντονη λοξότητα προς τα δεξιά (θετική ασυμμετρία).



Σχήμα 3.7. Ραβδόγραμμα και θηκόγραμμα για τα δεδομένα Lemaire

3.3.1 Προσαρμογή κατανομής Poisson

Ο κάτωθι πίνακας δίνει τα βασικά αποτελέσματα της προσαρμογής της κατανομής $P(\lambda)$ στα δεδομένα Lemaire.

Πίνακας 3.15. Βασικά χαρακτηριστικά της προσαρμογής της κατανομής Poisson στα δεδομένα Lemaire

Παράμετρος λ	0.2143537
----------------------	-----------

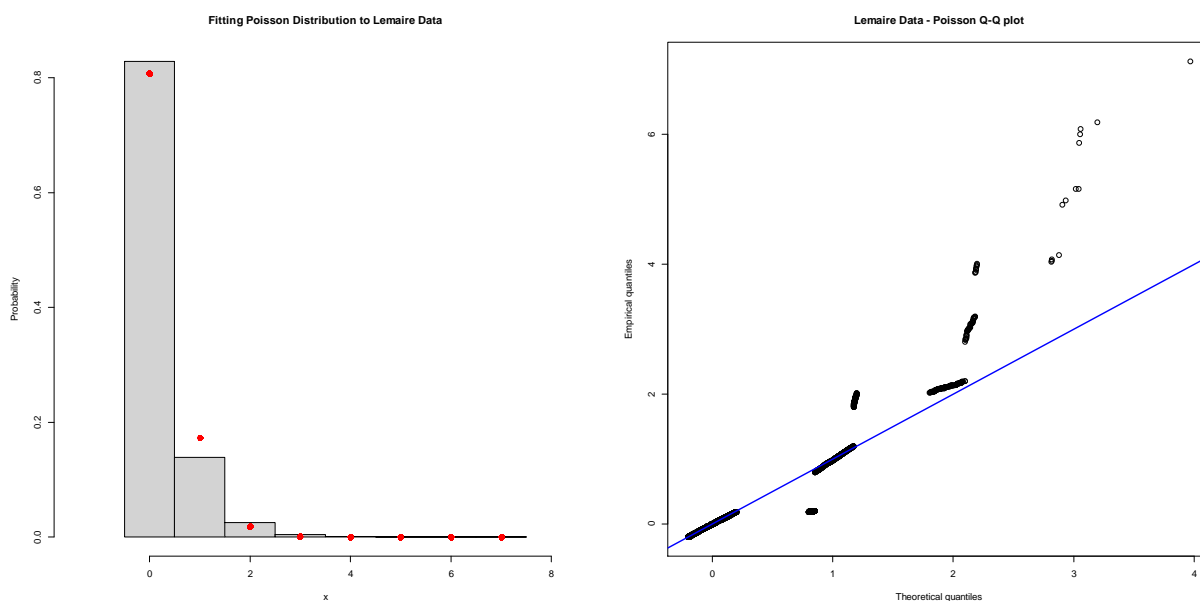
Loglikelihood	-5490.781
AIC	10983.56
p – value χ^2 ελέγχου	0
Στατιστική χ^2 ελέγχου	293.4263
Βαθμοί ελευθερίας	2

Από τον παραπάνω πίνακα προκύπτει ότι η κατανομή Poisson δεν έχει καθόλου καλή προσαρμογή στα δεδομένα αφού το p -value του χ^2 ελέγχου είναι 0 το οποίο είναι μικρότερο του 0.05 (τιμή του συνήθους επιπέδου σημαντικότητας). Σημειώνουμε ότι για την εφαρμογή του χ^2 ελέγχου οι τιμές ομαδοποιούνται έτσι ώστε να προκύπτει αναμενόμενη συχνότητα μεγαλύτερη ή ίση του 5 σε κάθε τάξη. Οι αναμενόμενες συχνότητες δίνονται στον ακόλουθο πίνακα.

Πίνακας 3.16. Πιθανότητες και αναμενόμενες συχνότητες της κατανομής Poisson

X	Πιθανότητα	Αναμενόμενες Συχνότητες	Παρατηρηθείσες Συχνότητες
X=0	0.8071	7635.6221	7840
X=1	0.1730	1636.7236	1317
X=2	0.0185	175.4188	239
X>=3	0.0014	13.2354	65

Στο ακόλουθο σχήμα δίνεται το ιστόγραμμα των εμπειρικών πιθανοτήτων μαζί με τις θεωρητικές πιθανότητες (κόκκινες κουκίδες) της προσαρμοσμένης κατανομής Poisson, καθώς επίσης και το Poisson Q-Q διάγραμμα. Από το Poisson Q-Q διάγραμμα διαπιστώνουμε την ακαταλληλότητα της προσαρμογής της κατανομής Poisson στα δεδομένα.



Σχήμα 3.8. Ιστόγραμμα εμπειρικών/θεωρητικών πιθανοτήτων και Poisson Q-Q διάγραμμα για τα δεδομένα Lemaire

3.3.2 Προσαρμογή κατανομής Poisson-Gamma

Ο κάτωθι πίνακας δίνει τα βασικά αποτελέσματα της προσαρμογής της μεικτής κατανομής Poisson με μεικτική την τυχαία μεταβλητή $L \sim G(a, \beta)$ (δείτε Παράγραφο 2.4.1) στα δεδομένα Lemaire.

Πίνακας 3.17. Βασικά χαρακτηριστικά της προσαρμογής της κατανομής Poisson-Gamma στα δεδομένα Lemaire

Παράμετρος a	0.70218
Παράμετρος β	3.275098
Loglikelihood	-5348.04
AIC	10700.08
p – value χ^2 ελέγχου	0.01251451
Στατιστική χ^2 ελέγχου	8.761734
Βαθμοί ελευθερίας	2

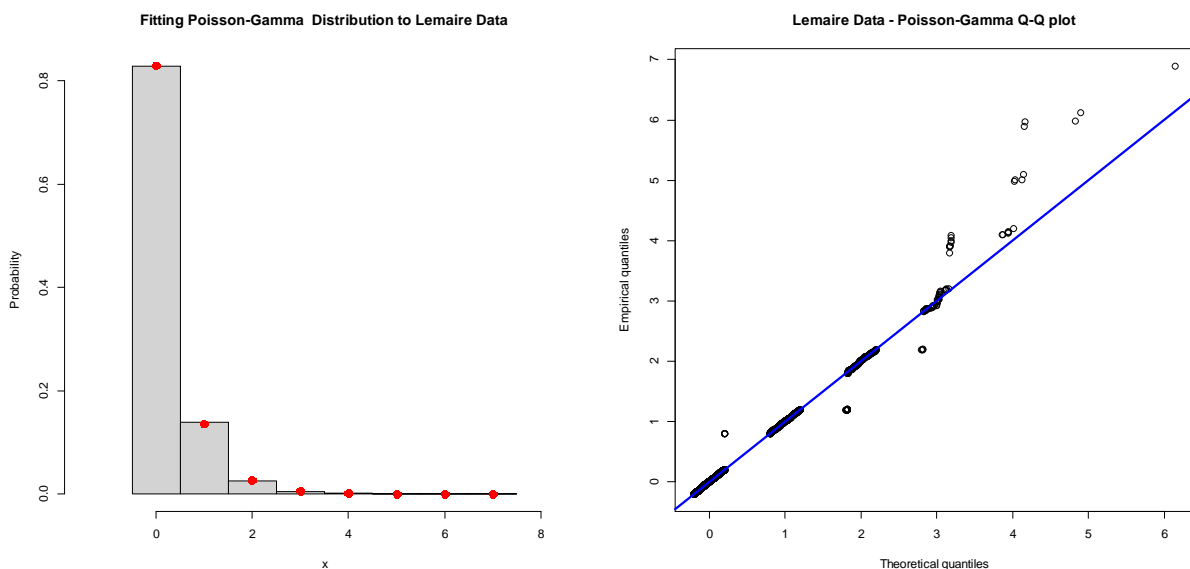
Σημειώνουμε ότι η προσαρμοσμένη κατανομή Poisson-Gamma είναι ως γνωστόν η αρνητική διωνυμική κατανομή $NB_0(r, p)$ με παραμέτρους $r = a = 0.70218$ και $p =$

$\frac{1}{1+\beta} = 0.2339127$. Από τον παραπάνω πίνακα προκύπτει ότι η κατανομή Poisson-Gamma φαίνεται να μην έχει καλή προσαρμογή στα δεδομένα αφού το p-value του χ^2 ελέγχου είναι μικρότερο του 0.05 (τιμή του συνήθους επιπέδου σημαντικότητας). Οι αναμενόμενες συχνότητες δίνονται στον ακόλουθο πίνακα.

Πίνακας 3.18. Πιθανότητες και αναμενόμενες συχνότητες της κατανομής Poisson-Gamma

X	Πιθανότητα	Αναμενόμενες Συχνότητες	Παρατηρηθείσες Συχνότητες
X=0	0.8294	7846.5643	7840
X=1	0.1362	1288.7891	1317
X=2	0.0271	256.5732	239
X=3	0.0057	54.0578	42
X>=4	0.0016	15.0156	23

Στο ακόλουθο σχήμα δίνεται το ιστόγραμμα των εμπειρικών πιθανοτήτων μαζί με τις θεωρητικές πιθανότητες (κόκκινες κουκίδες) του προσαρμοσμένου μοντέλου, καθώς επίσης και το Poisson-Gamma Q-Q διάγραμμα. Από το Q-Q διάγραμμα διαπιστώνουμε τη σχετικά μη καλή εφαρμογή της κατανομής Poisson-Gamma στα δεδομένα.



Σχήμα 3.9. Ιστόγραμμα εμπειρικών/θεωρητικών πιθανοτήτων και Poisson-Gamma Q-Q διάγραμμα για τα δεδομένα Lemaire

3.3.3 Προσαρμογή κατανομής Poisson-Lindley

Ο κάτωθι πίνακας δίνει τα βασικά αποτελέσματα της προσαρμογής της μεικτής κατανομής Poisson με μεικτική την τυχαία μεταβλητή $A \sim Lindley(\theta)$ (δείτε Παράγραφο 2.4.3) στα δεδομένα Lemaire.

Πίνακας 3.19. Βασικά χαρακτηριστικά της προσαρμογής της κατανομής Poisson-Lindley στα δεδομένα Lemaire

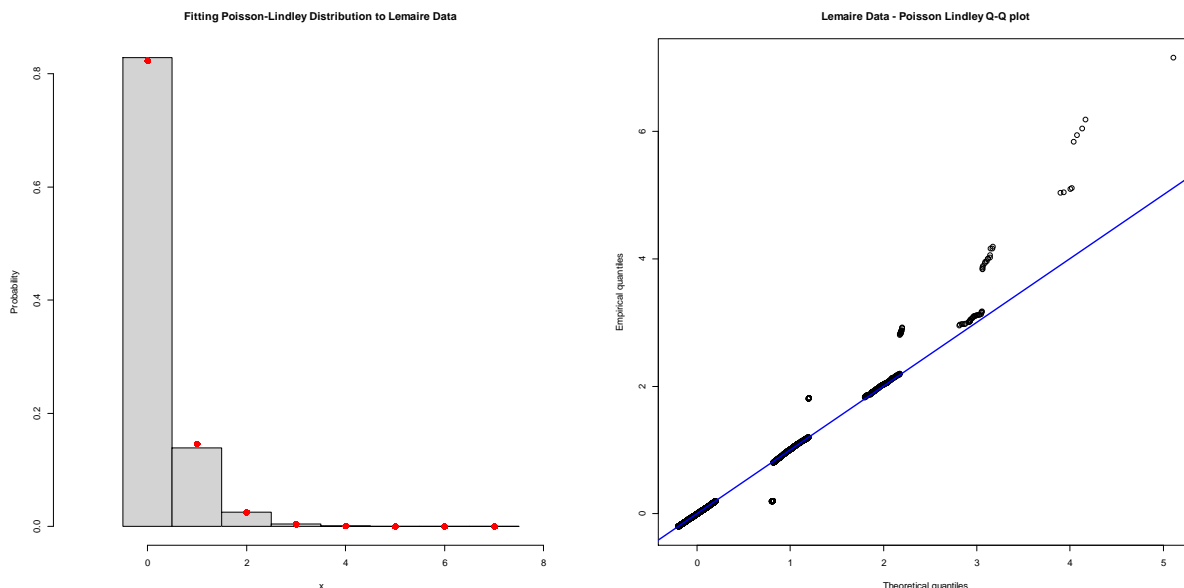
Παράμετρος θ	5.399384
Loglikelihood	-5356.278
AIC	10714.56
p – value χ^2 ελέγχου	4.869671e-06
Στατιστική χ^2 ελέγχου	27.39305
Βαθμοί ελευθερίας	3

Από τον παραπάνω πίνακα προκύπτει ότι η κατανομή Poisson-Lindley δεν φαίνεται να έχει καλή προσαρμογή στα δεδομένα αφού το p -value του χ^2 ελέγχου είναι πολύ μικρότερο του 0.05. Οι αναμενόμενες συχνότητες δίνονται στον ακόλουθο πίνακα.

Πίνακας 3.20. Πιθανότητες και αναμενόμενες συχνότητες της κατανομής Poisson-Lindley

X	Πιθανότητα	Αναμενόμενες Συχνότητες	Παρατηρηθείσες Συχνότητες
X=0	0.8231	7787.6518	7840
X=1	0.1460	1381.4025	1317
X=2	0.0255	241.5650	239
X=3	0.0044	41.7642	42
X>=4	0.0009	8.6165	23

Στο ακόλουθο σχήμα δίνεται το ιστόγραμμα των εμπειρικών πιθανοτήτων μαζί με τις θεωρητικές πιθανότητες (κόκκινες κουκίδες) του προσαρμοσμένου μοντέλου, καθώς επίσης και το Poisson-Lindley Q-Q διάγραμμα με το οποίο διαπιστώνεται η μη καλή προσαρμογή της κατανομής Poisson-Lindley στα δεδομένα.



Σχήμα 3.10. Ιστόγραμμα εμπειρικών/θεωρητικών πιθανοτήτων και Poisson-Lindley Q-Q διάγραμμα για τα δεδομένα Lemaire

3.3.4 Προσαρμογή κατανομής Poisson-Inverse Gaussian

Ο κάτωθι πίνακας δίνει τα βασικά αποτελέσματα της προσαρμογής της μεικτής κατανομής Poisson με μεικτική την τυχαία μεταβλητή $\Lambda \sim IG(\mu, \theta)$ (δείτε Παράγραφο 2.4.7) στα δεδομένα Lemaire.

Πίνακας 3.21. Βασικά χαρακτηριστικά της προσαρμογής της κατανομής Poisson- Inverse Gaussian στα δεδομένα Lemaire

Παράμετρος μ	0.2143947
Παράμετρος θ	0.1397237
Loglikelihood	-5343.511
AIC	10691.02
p – value χ^2 ελέγχου	0.2051549

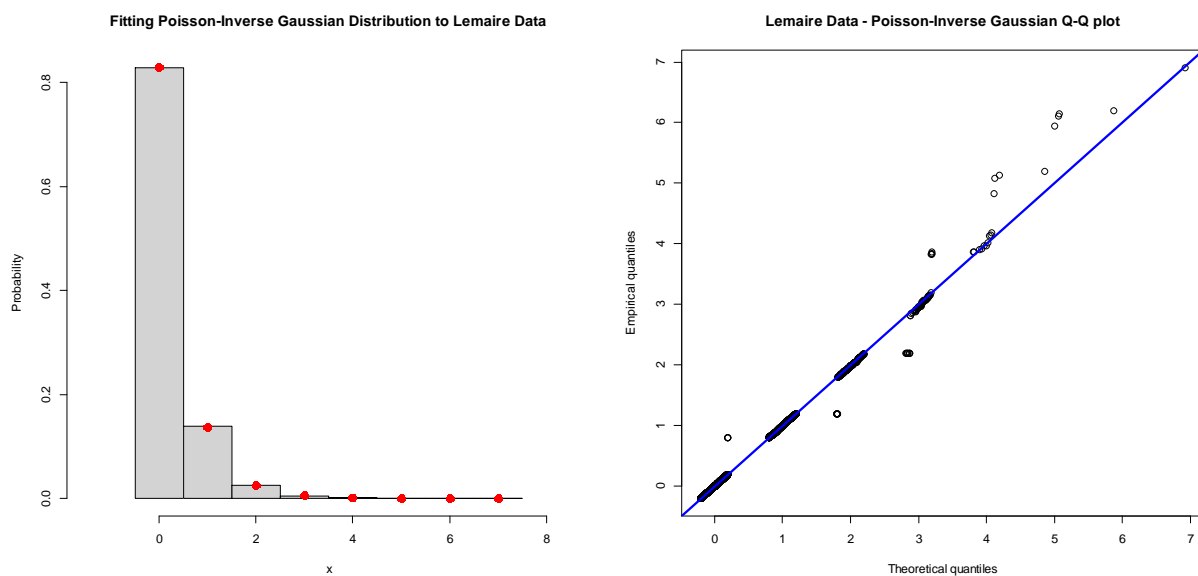
Στατιστική χ^2 ελέγχου	3.16798
Βαθμοί ελευθερίας	2

Από τον παραπάνω πίνακα προκύπτει ότι η κατανομή Poisson- Inverse Gaussian φαίνεται να έχει καλή προσαρμογή στα δεδομένα αφού το p-value του χ^2 ελέγχου είναι πολύ μεγαλύτερο του 0.05. Οι αναμενόμενες συχνότητες δίνονται στον ακόλουθο πίνακα.

Πίνακας 3.22. Πιθανότητες και αναμενόμενες συχνότητες της κατανομής Poisson-Inverse Gaussian

X	Πιθανότητα	Αναμενόμενες Συχνότητες	Παρατηρηθείσες Συχνότητες
X=0	0.8291	7843.9166	7840
X=1	0.1380	1306.0575	1317
X=2	0.0252	238.3079	239
X=3	0.0056	53.3203	42
X>=4	0.0021	19.3977	23

Στο ακόλουθο σχήμα δίνεται το ιστόγραμμα των εμπειρικών πιθανοτήτων μαζί με τις θεωρητικές πιθανότητες (κόκκινες κουκίδες) του προσαρμοσμένου μοντέλου, καθώς επίσης και το Poisson-Inverse Gaussian Q-Q διάγραμμα με το οποίο διαπιστώνεται η καλή εφαρμογή της κατανομής Poisson-Inverse Gaussian στα δεδομένα.



Σχήμα 3.11. Ιστόγραμμα εμπειρικών/θεωρητικών πιθανοτήτων και Poisson- Inverse Gaussian Q-Q διάγραμμα για τα δεδομένα Lemaire

3.3.5 Προσαρμογή κατανομής Poisson-Lognormal

Ο κάτωθι πίνακας δίνει τα βασικά αποτελέσματα της προσαρμογής της μεικτής κατανομής Poisson με μεικτική την τυχαία μεταβλητή $\Lambda \sim LN(\mu, \sigma)$ (δείτε Παράγραφο 2.4.8) στα δεδομένα Lemaire.

Πίνακας 3.23. Βασικά χαρακτηριστικά της προσαρμογής της κατανομής Poisson-Lognormal στα δεδομένα Lemaire

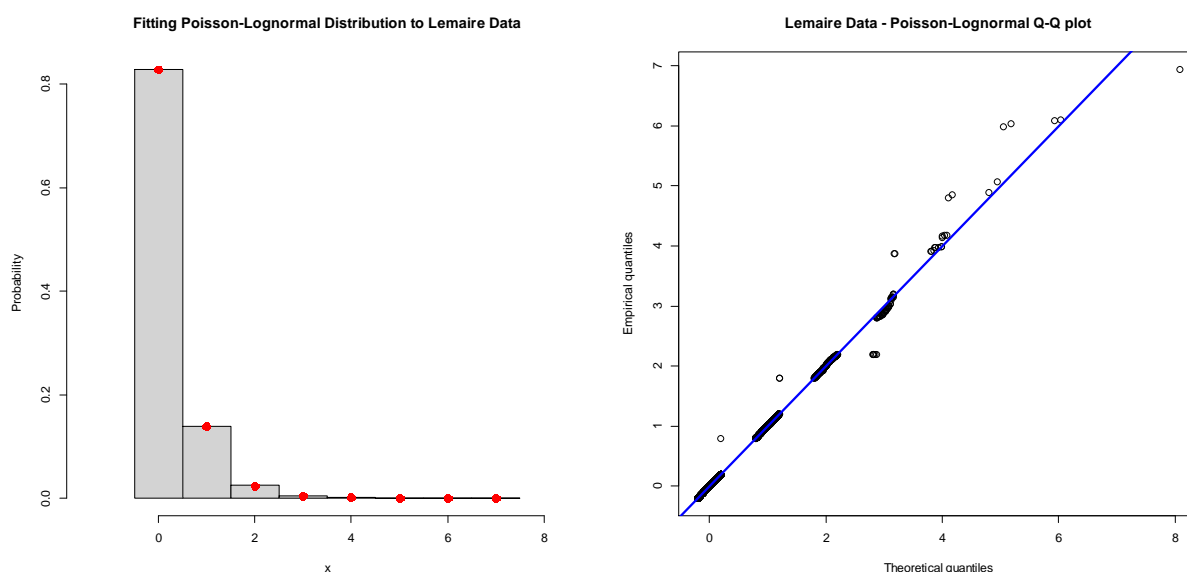
Παράμετρος μ	-2.024666
Παράμετρος σ	0.9842833
Loglikelihood	-5342.363
AIC	10688.73
p – value χ^2 ελέγχου	0.5724823
Στατιστική χ^2 ελέγχου	1.999636
Βαθμοί ελευθερίας	3

Από τον παραπάνω πίνακα προκύπτει ότι η κατανομή Poisson-Lognormal φαίνεται να έχει αρκετά καλή προσαρμογή στα δεδομένα αφού το p -value του χ^2 ελέγχου είναι πολύ μεγαλύτερο του 0.05. Οι αναμενόμενες συχνότητες σύμφωνα με το μοντέλο Poisson-Lognormal δίνονται στον ακόλουθο πίνακα.

Πίνακας 3.24. Πιθανότητες και αναμενόμενες συχνότητες της κατανομής Poisson-Lognormal

X	Πιθανότητα	Αναμενόμενες Συχνότητες	Παρατηρηθείσες Συχνότητες
X=0	0.8287	7840.6712	7840
X=1	0.1393	1317.9553	1317
X=2	0.0245	231.3518	239
X=3	0.0053	50.2049	42
X=4	0.0014	13.4987	14
X>=5	0.0008	7.3181	9

Στο ακόλουθο σχήμα δίνεται το ιστόγραμμα των εμπειρικών πιθανοτήτων μαζί με τις θεωρητικές πιθανότητες (κόκκινες κουκίδες) του προσαρμοσμένου μοντέλου, καθώς επίσης και το Poisson-Lognormal Q-Q διάγραμμα από το οποίο διαπιστώνεται η αρκετά καλή προσαρμογή της κατανομής Poisson-Lognormal στα δεδομένα.



Σχήμα 3.12. Ιστόγραμμα εμπειρικών/θεωρητικών πιθανοτήτων και Poisson-Lognormal Q-Q διάγραμμα για τα δεδομένα Lemaire

3.3.6 Προσαρμογή κατανομής Poisson-Beta

Ο κάτωθι πίνακας δίνει τα βασικά αποτελέσματα της προσαρμογής της μεικτής κατανομής Poisson με μεικτική την τυχαία μεταβλητή $\Lambda \sim \text{Beta}(a, \beta)$ (δείτε Παράγραφο 2.4.4) στα δεδομένα Lemaire.

Πίνακας 3.25. Βασικά χαρακτηριστικά της προσαρμογής της κατανομής Poisson- Beta στα δεδομένα Lemaire

Παράμετρος a	0.4188788
Παράμετρος β	1.546422
Loglikelihood	-5358.571
AIC	10721.14
p – value χ^2 ελέγχου	4.96893e-06
Στατιστική χ^2 ελέγχου	24.42461
Βαθμοί ελευθερίας	2

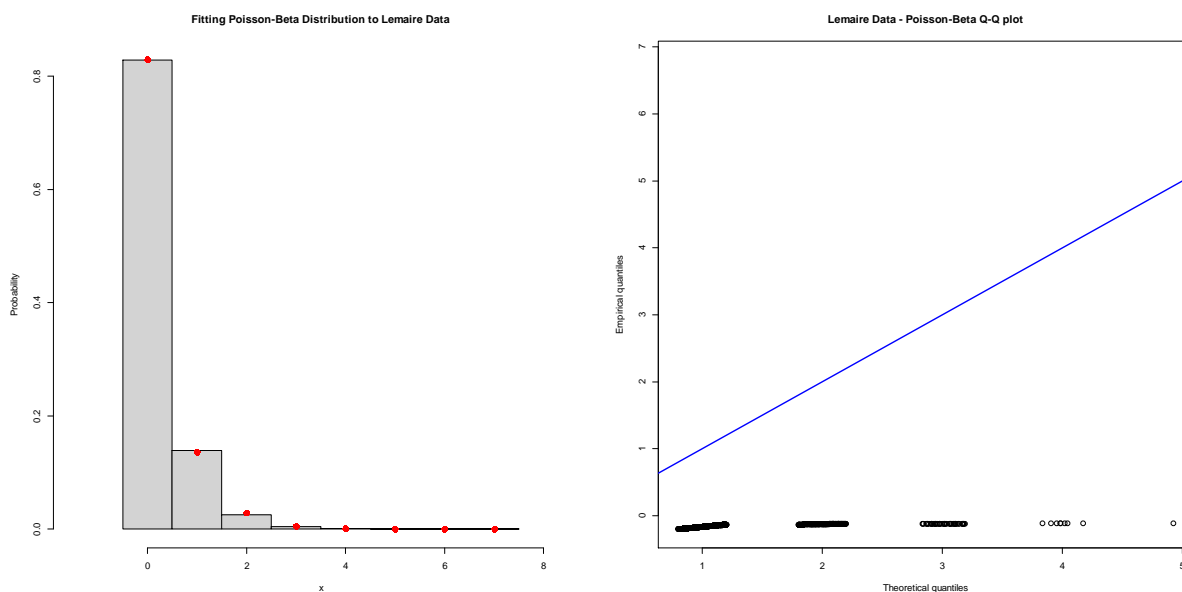
Από τον παραπάνω πίνακα προκύπτει ότι η κατανομή Poisson-Beta δεν έχει καλή προσαρμογή στα δεδομένα αφού το p -value του χ^2 ελέγχου είναι πολύ μικρότερο του 0.05. Οι αναμενόμενες συχνότητες σύμφωνα με το μοντέλο Poisson-Beta δίνονται στον ακόλουθο πίνακα.

Πίνακας 3.26. Πιθανότητες και αναμενόμενες συχνότητες της κατανομής Poisson-Beta

X	Πιθανότητα	Αναμενόμενες Συχνότητες	Παρατηρηθείσες Συχνότητες
X=0	0.8290	7843.5257	7840
X=1	0.1363	1289.1561	1317
X=2	0.0284	268.6247	239
X=3	0.0053	50.1999	42

$X \geq 4$	0.0010	9.4936	23
------------	--------	--------	----

Στο ακόλουθο σχήμα δίνεται το ιστόγραμμα των εμπειρικών πιθανοτήτων μαζί με τις θεωρητικές πιθανότητες (κόκκινες κουκίδες) του προσαρμοσμένου μοντέλου, καθώς επίσης και το Poisson-Lognormal Q-Q διάγραμμα από το οποίο διαπιστώνεται η καθόλου καλή προσαρμογή της κατανομής Poisson-Beta στα δεδομένα.



Σχήμα 3.13. Ιστόγραμμα εμπειρικών/θεωρητικών πιθανοτήτων και Poisson-Beta Q-Q διάγραμμα για τα δεδομένα Lemaire

3.4 Συμπεράσματα

Ο στόχος μας σε αυτό το κεφάλαιο ήταν η εφαρμογή μεικτών κατανομών Poisson σε δύο επιλεγμένα σύνολα δεδομένων. Παρατηρήθηκε λοιπόν ότι κάποιες από αυτές έχουν καλή προσαρμογή όπως παρατηρούμε και από τους δυο συγκεντρωτικούς πίνακες που δίνονται παρακάτω, με κριτήριο την p – value του σχετικού χ^2 ελέγχου, αλλά και με το κριτήριο AIC. Με έντονα γράμματα σημειώνονται οι κατανομές με την καλύτερη προσαρμογή. Κατά τον έλεγχο συμβατότητας των δεδομένων, παρατηρήσαμε ότι η κατανομή Poisson δεν έχει καλή προσαρμογή.

Πιο συγκεκριμένα, στα δεδομένα Crimes η κατανομή Poisson-Inverse Gaussian έχει την καλύτερη προσαρμογή με βάση το κριτήριο AIC, ενώ καλή προσαρμογή έχουν και οι κατανομές Poisson-Gamma και Poisson-Lognormal. Στα δεδομένα Lemaire η κατανομή

Poisson-Lognormal έχει την καλύτερη προσαρμογή με βάση το κριτήριο AIC, ενώ καλή προσαρμογή έχει και η κατανομή Poisson-Inverse Gaussian.

Πίνακας 3.27. Δεδομένα Crimes – Σύγκριση

	p – value χ^2 ελέγχου	AIC
Poisson	0.01262442	564.1606
Poisson-Gamma	0.2974202	553.011
Poisson-Lindley	0.0009683159	570.4965
Poisson-Inverse Gaussian	0.3070627	552.9151
Poisson-Lognormal	0.3046507	552.9935

Πίνακας 3.28. Δεδομένα Lemaire – Σύγκριση

	p – value χ^2 ελέγχου	AIC
Poisson	0	10983.56
Poisson-Gamma	0.01251451	10700.08
Poisson-Lindley	4.869671e-06	10714.56
Poisson-Inverse Gaussian	0.2051549	10691.02
Poisson-Lognormal	0.5724823	10688.73
Poisson-Beta	4.96893e-06	10721.14

Άρα, εύκολα μπορούμε να συμπεράνουμε ότι οι μεικτές κατανομές Poisson που είναι κατανομές με διάφορες μορφές και χαρακτηριστικά, αν τις επιλέξουμε και τις προσαρμόσουμε στα εκάστοτε δεδομένα, μπορούμε να εξάγουμε αποτελέσματα και συμπεράσματα συμβατά με τα εμπειρικά δεδομένα ιδιαίτερα στην περίπτωση που η κλασική κατανομή Poisson έχει φτωχή προσαρμογή. Συνεπώς, οποιαδήποτε από τις μεικτές κατανομές Poisson θα πρέπει να συμπεριλαμβάνεται στις κατανομές που χρησιμοποιεί ένας αναλογιστής ή διαχειριστής κινδύνων στις έρευνές του.

ΠΑΡΑΡΤΗΜΑ

Κεφάλαιο 2

Κώδικας για τη διεξαγωγή των Σχημάτων 2.1 έως 2.8

ΣΧΗΜΑ 2.1

```
par(mfrow=c(1,2))
alpha1 <- 3; beta1 <- 1.4
alpha2 <- 3; beta2 <- 0.7
x <- 0:20
z1 <- dnbinom(x,size=alpha1, prob=(beta1/(1+beta1)))
z2 <- dnbinom(x,size=alpha2, prob=(beta2/(1+beta2)))
plot(x,z1,type="h", ylim = c(0,0.25), col="red", main="Συνάρτηση
πιθανότητας της κατανομής Poisson-Gamma :  $\alpha=3$ ,  $\beta=1.4$ ", ylab="f(x)", lwd=4,
cex.main=1.7)
plot(x,z2,type="h", ylim = c(0,0.25), col="red", main="Συνάρτηση
πιθανότητας της κατανομής Poisson-Gamma :  $\alpha=3$ ,  $\beta=0.7$ ", ylab="f(x)", lwd=4,
cex.main=1.7)

# Alternative way
library(extraDistr)
par(mfrow=c(1,2))
w1 <- dgpois(x,alpha1,beta1)
w2 <- dgpois(x,alpha2,beta2)
plot(x,w1,type="h", ylim = c(0,0.25), col="red",
main="Συνάρτηση πιθανότητας της κατανομής Poisson-Gamma :  $\alpha=3$ ,  $\beta=1.4$ ",
ylab="f(x)", lwd=4, cex.main=1.7)
plot(x,w2,type="h", ylim = c(0,0.25), col="red",
main="Συνάρτηση πιθανότητας της κατανομής Poisson-Gamma :  $\alpha=3$ ,  $\beta=0.7$ ",
ylab="f(x)", lwd=4, cex.main=1.7)
```

ΣΧΗΜΑ 2.2

```
par(mfrow=c(1,2))
alpha1 <- 3; beta1 <- 1.4; lambda1 <- 4
alpha2 <- 3; beta2 <- 0.7; lambda2 <- 2
x <- 0:20
library(Delaporte)
n1 <- ddelap(x,alpha=alpha1, beta=beta1, lambda=lambda1)
n2 <- ddelap(x,alpha=alpha2, beta=beta2, lambda=lambda2)
plot(x,n1,type="h", ylim = c(0,0.25), col="red", main="Συνάρτηση
πιθανότητας της κατανομής Delaporte :  $\alpha=3$ ,  $\beta=1.4$ ,  $\mu=4$ ", ylab="f(x)", lwd=4,
cex.main=1.2)
plot(x,n2,type="h", ylim = c(0,0.25), col="red", main="Συνάρτηση
πιθανότητας της κατανομής Delaporte :  $\alpha=3$ ,  $\beta=0.7$ ,  $\mu=2$ ", ylab="f(x)", lwd=4,
cex.main=1.2)
```

ΣΧΗΜΑ 2.3

```
par(mfrow=c(1,2))
theta1 = 0.5; theta2 = 1
x1 <- 0:30
w1 <- theta1^2*(x1+theta1+2)/(theta1+1)^(x1+3)
plot(x1,w1,type="h", ylim = c(0,0.4), col="red",
main="Συνάρτηση πιθανότητας της κατανομής Poisson-Lindley :  $\theta=0.5$ ",
ylab="f(x)", lwd=4, cex.main=1.7)
x2 <- 0:30
w2 <- theta2^2*(x2+theta2+2)/(theta2+1)^(x2+3)
plot(x2,w2,type="h", ylim = c(0,0.4), col="red",
```

```

main="Συνάρτηση πιθανότητας της κατανομής Poisson-Lindley :  $\theta=1$ ",
ylab="f(x)", lwd=4, cex.main=1.7)

# Alternative way
library(tolerance)
par(mfrow=c(1,2))
theta1 = 0.5; theta2 = 1
x1 <- 0:30
y1 <- dpoislind(x1, theta = theta1)
plot(x1,y1,type="h", ylim = c(0,0.4), col="red",
main="Συνάρτηση πιθανότητας της κατανομής Poisson-Lindley :  $\theta=0.5$ ",
ylab="f(x)", lwd=4, cex.main=1.7)
x2 <- 0:30
y2 <- dpoislind(x2, theta = theta2)
plot(x2,y2,type="h", ylim = c(0,0.4), col="red",
main="Συνάρτηση πιθανότητας της κατανομής Poisson-Lindley :  $\theta=1$ ",
ylab="f(x)", lwd=4, cex.main=1.7)

```

ΣΧΗΜΑ 2.4

```

library(scModels)
par(mfrow=c(1,2))
a1 = 10; b1 = 0.5
x1 <- 0:20
y1 <- dpb(x1,a1,b1)
plot(x1,y1,type="h", ylim = c(0,0.5), col="red", main="Συνάρτηση
πιθανότητας της κατανομής Poisson-Beta :  $\alpha=10$ ,  $\beta=0.5$ ", ylab="f(x)", lwd=4,
cex.main=1.7)

a2 = 10; b2 = 10
x2 <- 0:20
y2 <- dpb(x2, a2,b2)
plot(x2,y2,type="h", ylim = c(0,0.7), col="red", main="Συνάρτηση
πιθανότητας της κατανομής Poisson-Beta :  $\alpha=10$ ,  $\beta=10$ ", ylab="f(x)", lwd=4,
cex.main=1.7)

```

ΣΧΗΜΑ 2.5

```

a1 <- 3; b1 <- 2; f1 <- c()
for (x in 0:30) {
z <- function(t) { (t^(x-a1-1)) * exp(-t) }
val <- integrate(z, lower=b1, upper=Inf)
f1[x+1] <- a1*(b1^a1)*(1/factorial(x)) * val$value
}

a2 <- 4; b2 <- 8; f2 <- c()
for (x in 0:30) {
z <- function(t) { (t^(x-a2-1)) * exp(-t) }
val <- integrate(z, lower=b2, upper=Inf)
f2[x+1] <- a2*(b2^a2)*(1/factorial(x)) * val$value
}

par(mfrow=c(1,2))
x1 <- 0:30
plot(x1,f1,type="h", ylim = c(0,0.25), col="red", main="Συνάρτηση
πιθανότητας της κατανομής Poisson-Pareto : a=3, b=2", ylab="f(x)",
xlab="x", lwd=4, cex.main=1.2)
x2 <- 0:30
plot(x2,f2,type="h", ylim = c(0,0.25), col="red", main="Συνάρτηση
πιθανότητας της κατανομής Poisson-Pareto : a=4, b=8", ylab="f(x)",
xlab="x", lwd=4, cex.main=1.2)

```

ΣΧΗΜΑ 2.6

```
a1 <- 3; b1 <- 2; v <- 2*sqrt(b1); f1 <- c()
for (x in 0:25) {
  u <- x-a1
  z <- function(t) { (1/2)*(t^(u-1)) * exp(-(v/2)*(t+(1/t))) }
  val <- integrate(z, lower=0, upper=Inf)
  f1[x+1] <- (2/factorial(x)) * (b1^((x+a1)/2)) * (1/gamma(a1)) * val$value
}

a2 <- 4; b2 <- 8; v <- 2*sqrt(b2); f2 <- c()
for (x in 0:25) {
  u <- x-a2
  z <- function(t) { (1/2)*(t^(u-1)) * exp(-(v/2)*(t+(1/t))) }
  val <- integrate(z, lower=0, upper=Inf)
  f2[x+1] <- (2/factorial(x)) * (b2^((x+a2)/2)) * (1/gamma(a2)) * val$value
}

par(mfrow=c(1,2))
x1 <- 0:25
plot(x1,f1,type="h", ylim = c(0,0.5), col="red", main="Συνάρτηση
πιθανότητας της κατανομής Poisson-Inverse Gamma : a=3, b=2", ylab="f(x)",
xlab="x", lwd=4, cex.main=1.2)
x2 <- 0:25
plot(x2,f2,type="h", ylim = c(0,0.5), col="red", main="Συνάρτηση
πιθανότητας της κατανομής Poisson-Inverse Gamma : a=4, b=8", ylab="f(x)",
xlab="x", lwd=4, cex.main=1.2)
```

ΣΧΗΜΑ 2.7

```
library(actuar)
par(mfrow=c(1,2))
mean1 <- 3; shapel <- 0.3
mean2 <- 3; shape2 <- 0.7
x <- 0:30
w1 <- dpoisinvgauss(x,mean1,shapel)
w2 <- dpoisinvgauss(x,mean2,shape2)
plot(x,w1,type="h", ylim = c(0,0.25), col="red",
main="Συνάρτηση πιθανότητας της κατανομής Poisson-Inverse Gaussian: μ=3,
θ=0.3", ylab="f(x)", lwd=4, cex.main=1.0)
plot(x,w2,type="h", ylim = c(0,0.25), col="red",
main="Συνάρτηση πιθανότητας της κατανομής Poisson-Inverse Gaussian: μ=3,
θ=0.7", ylab="f(x)", lwd=4, cex.main=1.0)
```

ΣΧΗΜΑ 2.8

```
library(sads)
par(mfrow=c(1,2))
mu1 <- 0.7; sig1 <- 0.2
mu2 <- 0.7; sig2 <- 0.4
x <- 0:30
w1 <- dpoilog(x,mu1,sig1)
w2 <- dpoilog(x,mu2,sig2)
plot(x,w1,type="h", ylim = c(0,0.25), col="red",
main="Συνάρτηση πιθανότητας της κατανομής Poisson-Lognormal: μ=0.7, σ=0.2",
ylab="f(x)", lwd=4, cex.main=1.0)
plot(x,w2,type="h", ylim = c(0,0.25), col="red",
main="Συνάρτηση πιθανότητας της κατανομής Poisson-Lognormal: μ=0.7, σ=0.4",
ylab="f(x)", lwd=4, cex.main=1.0)
```

Κεφάλαιο 3

Κώδικας για τα δεδομένα Crimes

Κατανομή Poisson

```
x <- 0:9
Observed <- c(21,41,32,16,19,8,4,1,2,1)
d <- rep(x,times=Observed)
summary(d)
var(d)
sd(d)
library(moments)
skewness(d)
kurtosis(d)
par(mfrow=c(1,2))
barplot(table(d), xlab="x", main="Histogram of Crimes Data",
ylab="Frequency")
boxplot(d, main="Boxplot of Crimes Data")

library(fitdistrplus)
poisMLE <- fitdist(d, "pois", method="mle")
fitdist(d, "pois", method="mle", start=list(lambda=1))
summary(poisMLE)
(LAMBDA <- poisMLE$estimate)

par(mfrow=c(1,2))
class <- seq(-0.5, 9.5, 1)
hist(d, breaks=class, xlim=c(-1.5, 10.5), prob=T, xlab="x",
ylab="Probability", main="Fitting Poisson Distribution to Crimes Data" )
y <- dpois(x, lambda=LAMBDA)
points(x,y, pch=16, col="red", cex=1.5)
qqplot(jitter(qpois(ppoints(length(d)),lambda=LAMBDA)), jitter(d),
xlab="Theoretical quantiles", ylab="Empirical quantiles", main="Crimes Data
- Poisson Q-Q plot", cex=1.2)
abline(0,1, col="blue", lwd=3)

CUT <- 4
LIM <- length(x)
(PROBABILITY <- c(dpois(0:CUT, LAMBDA), 1-ppois(CUT,LAMBDA)))
(EXPECTED <- length(d)*PROBABILITY)
CUT1 <- CUT+1; CUT2 <- CUT+2
(OBSERVED <- c(Observed[1:CUT1], sum(Observed[CUT2:LIM])))
MAT <- cbind(PROBABILITY,EXPECTED,OBSERVED)
row.names(MAT) <- c("X=0","X=1","X=2","X=3","X=4","X>=5")
round(MAT, 4)
(chi.obs <- sum((OBSERVED-EXPECTED)^2/EXPECTED))
(DF <- 4)
(p.value <- 1-pchisq(chi.obs,DF))
```

Κατανομή Poisson-Gamma

```
x <- 0:9
Observed <- c(21,41,32,16,19,8,4,1,2,1)
d <- rep(x,times=Observed)

library(fitdistrplus)
library(extraDistr)
mystartPG <- list(shape=2,rate=0.5)
PGMLE <- fitdist(d, "gpois", method="mle",start=mystartPG)
```

```

summary(PGMLE)
(sh <- PGMLE$estimate[1])
(ra <- PGMLE$estimate[2])
( r <- sh )
( p <- ra/(1+ra))

par(mfrow=c(1,2))
class <- seq(-0.5, 9.5, 1)
hist(d, breaks=class, xlim=c(-1.5, 10.5), prob=T, xlab="x",
ylab="Probability", main="Fitting Poisson-Gamma Distribution to Crimes
Data" )
y <- dnbinom(x, r, p)
points(x,y, pch=16, col="red", cex=1.5)
qqplot(jitter(qnbinom(ppoints(length(d)),r, p)), jitter(d),
xlab="Theoretical quantiles", ylab="Empirical quantiles", main="Crimes Data
- Poisson-Gamma Q-Q plot", cex=1.2)
abline(0,1, col="blue", lwd=3)
# qqline(d, distribution = function(p) qpoislind(p, THETA), prob = c(0.25,
0.75), col = "red")
par(mfrow=c(1,1))

CUT <- 5
LIM <- length(x)
(PROBABILITY <- c(dnbinom(0:CUT, r, p), 1-pnbinom(CUT,r,p)))
(EXPECTED <- length(d)*PROBABILITY)
CUT1 <- CUT+1
CUT2 <- CUT+2
(OBSERVED <- c(Observed[1:CUT1], sum(Observed[CUT2:LIM])))
MAT <- cbind(PROBABILITY,EXPECTED,OBSERVED)
row.names(MAT) <- c("X=0", "X=1", "X=2", "X=3", "X=4", "X=5", "X>=6")
round(MAT, 4)

(chi.obs <- sum((OBSERVED-EXPECTED)^2/EXPECTED))
(DF <- 4)
(p.value <- 1-pchisq(chi.obs,DF))

```

Κατανομή Poisson-Lindley

```

x <- 0:9
Observed <- c(21,41,32,16,19,8,4,1,2,1)
d <- rep(x,times=Observed)

library(tolerance)
LL <- function(th) {
  dens <- dpoislind(d, theta=th)
  return(-sum(log(dens)))
}
library(stats4)
mystartPL <- list(th=0.5)
mle(LL, start=mystartPL)
THETA <- 0.7012227
LogLik <- -LL(THETA); LogLik
AIC <- 2*LL(THETA)+2*1; AIC

par(mfrow=c(1,2))
class <- seq(-0.5, 9.5, 1)
hist(d, breaks=class, xlim=c(-1.5, 10.5), prob=T, xlab="x",
ylab="Probability", main="Fitting Poisson-Lindley Distribution to Crimes
Data" )
y <- dpoislind(x, theta=THETA)
points(x,y, pch=16, col="red", cex=1.5)
qqplot(jitter(qpoislind(ppoints(length(d)),theta=THETA)), jitter(d),

```

```

xlab="Theoretical quantiles", ylab="Empirical quantiles", main="Crimes Data
- Poisson Lindley Q-Q plot", cex=1.2)
abline(0,1, col="blue", lwd=3)

CUT <- 6
LIM <- length(x)
(PROBABILITY <- c(dpoislind(0:CUT, THETA), 1-ppoislind(CUT,THETA)))
(EXPECTED <- length(d)*PROBABILITY)
CUT1 <- CUT+1
CUT2 <- CUT+2
(OBSERVED <- c(Observed[1:CUT1], sum(Observed[CUT2:LIM])))
MAT <- cbind(PROBABILITY,EXPECTED,OBSERVED)
row.names(MAT) <- c("X=0","X=1","X=2","X=3","X=4","X=5", "X=6", "X>=7")
round(MAT, 4)

(chi.obs <- sum((OBSERVED-EXPECTED)^2/EXPECTED))
(DF <- 6)
(p.value <- 1-pchisq(chi.obs,DF))

```

Κατανομή Poisson-Inverse Gaussian

```

x <- 0:9
Observed <- c(21,41,32,16,19,8,4,1,2,1)
d <- rep(x,times=Observed)

library(fitdistrplus)
library(actuar)
mystartPIG <- list(mean=3, shape=0.7)
PIG <- fitdist(d, "poisinvgauss", method="mle", start=mystartPIG)
summary(PIG)
(m <- PIG$estimate[1])
(s <- PIG$estimate[2])

LL <- function(m,s) {
  dens <- dpoisinvgauss(d, m, s)
  return(-sum(log(dens)))
}
LogLik <- -LL(m,s); LogLik
AIC <- 2*LL(m,s)+2*2; AIC

par(mfrow=c(1,2))
class <- seq(-0.5, 9.5, 1)
hist(d, breaks=class, xlim=c(-1.5, 10.5), prob=T, xlab="x",
ylab="Probability", main="Fitting Poisson-Inverse Gaussian Distribution to
Crimes Data")
y <- dpoisinvgauss(x, m, s)
points(x,y, pch=16, col="red", cex=1.5)
qqplot(jitter(qpoisinvgauss(ppoints(length(d)),m, s)), jitter(d),
xlab="Theoretical quantiles", ylab="Empirical quantiles", main="Crimes Data
- Poisson-Inverse Gaussian Q-Q plot",cex=1.2)
abline(0,1, col="blue", lwd=3)

CUT <- 5
LIM <- length(x)
(PROBABILITY <- c(dpoisinvgauss(0:CUT, m, s), 1-ppoisinvgauss(CUT,m,s)))
(EXPECTED <- length(d)*PROBABILITY)
CUT1 <- CUT+1
CUT2 <- CUT+2
(OBSERVED <- c(Observed[1:CUT1], sum(Observed[CUT2:LIM])))
MAT <- cbind(PROBABILITY,EXPECTED,OBSERVED)
row.names(MAT) <- c("X=0","X=1","X=2","X=3","X=4","X=5", "X>=6")
round(MAT, 4)

(chi.obs <- sum((OBSERVED-EXPECTED)^2/EXPECTED))

```



```
(DF <- 4)
(p.value <- 1-pchisq(chi.obs,DF))
```

Κατανομή Poisson-Lognormal

```
x <- 0:9
Observed <- c(21,41,32,16,19,8,4,1,2,1)
d <- rep(x,times=Observed)

library(fitdistrplus)
library(sads)
mystartPL <- list(mu=0.7, sig=0.2)
PL <- fitdist(d, "poilog", method="mle", start=mystartPL)
summary(PL)
(m <- PL$estimate[1])
(s <- PL$estimate[2])

LL <- function(mu,sig) {
  dens <- dpoilog(d, mu, sig)
  return(-sum(log(dens)))
}
LogLik <- -LL(m,s); LogLik
AIC <- 2*LL(m,s)+2*2; AIC

par(mfrow=c(1,2))
class <- seq(-0.5, 9.5, 1)
hist(d, breaks=class, xlim=c(-1.5, 10.5), prob=T, xlab="x",
ylab="Probability", main="Fitting Poisson-Lognormal Distribution to Crimes
Data")
y <- dpoilog(x, m, s)
points(x,y, pch=16, col="red", cex=1.5)
qqplot(jitter(qpoilog(ppoints(length(d)),m, s)), jitter(d),
xlab="Theoretical quantiles", ylab="Empirical quantiles", main="Crimes Data
- Poisson-Lognormal Q-Q plot",cex=1.2)
abline(0,1, col="blue", lwd=3)

CUT <- 5
LIM <- length(x)
(PROBABILITY <- c(dpoilog(0:CUT, m, s), 1-ppoilog(CUT,m,s)))
(EXPECTED <- length(d)*PROBABILITY)
CUT1 <- CUT+1
CUT2 <- CUT+2
(OBSERVED <- c(Observed[1:CUT1], sum(Observed[CUT2:LIM])))
MAT <- cbind(PROBABILITY,EXPECTED,OBSERVED)
row.names(MAT) <- c("X=0","X=1","X=2","X=3","X=4","X=5", "X>=6")
round(MAT, 4)

(chi.obs <- sum((OBSERVED-EXPECTED)^2/EXPECTED))
(DF <- 4)
(p.value <- 1-pchisq(chi.obs,DF))
```

ΒΙΒΛΙΟΓΡΑΦΙΑ

Ελληνική

Αντζουλάκος, Δ., (2018), *Σημειώσεις Μαθήματος Γλώσσα Προγραμματισμού R, Γλώσσα Προγραμματισμού R Με Εφαρμογές Στον Αναλογισμό, ΠΜΣ Αναλογιστική Επιστήμη & Διαχείριση Κινδύνων.*

Ξένη

Alonso, D., Ostling, A., & Etienne, R. S., (2008), The implicit assumption of symmetry and the species abundance distribution. *Ecology Letters*, 11: 93-105.

Bulmer, M. G., (1974), On Fitting the Poisson Lognormal Distribution to Species-Abundance Data. *Biometrics*, 30, 101-110.

Chhikara, R. S., & Folk, T. L., (1989), *The Inverse Gaussian Distribution: Theory, Methodology and Applications*, Decker.

Dean, C., Lawless, J. F., & Willmot, G. E., (1989), A mixed poisson-inverse-gaussian regression model. *Canadian Journal of Statistics*, 17(2), 171–181.

Engen, S., Lande, R., Walla, T., & DeVries, P. J., (2002), Analyzing spatial structure of communities using the two-dimensional Poisson lognormal species abundance model. *American Naturalist*, 160: 60-73.

Feller, W., (1943), On a Generalized Class of Contagious Distributions. *Annals of Mathematical Statistics*, 14, 389–400.

Feller, W., (1971), *An Introduction to Probability Theory and its Applications*. Vol II. 2nd ed. Wiley, New York.

Forbes, C., M. Evans, N. Hastings, & B. Peacock, (2011), *Statistical Distributions*. Fourth Edition. John Wiley and Sons, Hoboken, NJ.

Greenwood, M., & Yule, G., (1920), An inquiry into the nature of frequency distributions representative of multiple happenings with particular reference to the occurrence of multiple attacks of disease or of repeated accidents. *Journal of the Royal Statistical Society A*, 83, 255–279.

Ghitany, M. E., Atieh, B., & Nadarajah, S., (2008), Lindley distribution and its application. *Mathematics and Computers in Simulation*, 78, 4, 49-506.

Grandell, J., (1997), *Mixed Poisson Processes*. Chapman and Hall.

- Gurland, J., (1957), Some interrelations among compound and generalized distributions. *Biometrika*, 44, 263–268.
- Holla, M. S., (1966), On a Poisson-Inverse Gaussian distribution. *Metrika*, Vol. 15, 377-384.
- Johnson, N. L., Kotz, S. & Balakrishnan, N., (1995), *Continuous Univariate Distributions*, Volume 1, chapter 14. Wiley, New York.
- Holla, M.S., & Bhattacharya, S.K., (1965). On a discrete compound distribution. *Annals of the Institute of Statistical Mathematics*, 15, 377–384.
- Kallenberg, O., (1983), Random Measures. 3rd ed. Akademie-Verlag, Berlin and Academic Press, New York.
- Karlis, D., (2005), EM algorithm for mixed Poisson and other discrete distributions. *ASTIN Bulletin: The Journal of the IAA*, 35(1), 3-24.
- Karlis, D., & Xelakaki, E., (2005), Mixed Poisson Distributions. *International Statistical Review*, Vol. 73, No. 1, pp. 35-58.
- Maceda, E.C., (1948), On the compound and generalized Poisson distributions. *Annals of Mathematical Statistics*, 19, 414–416.
- Ong, S.H., (1995), Computation of Probabilities of a Generalized Log-Series and Related Distributions. *Communication in Statistics-Theory and Methods*, 24, 253–271.
- Ong, S-H., Low, Y-C. & Toh, K-K., (2021), Recent developments in mixed Poisson distributions. *ASM Science Journal*, 14:1-10
- Panjer, H. H., & G. E. Willmot, G. E., (1992), *Insurance risk models*. Society of Actuaries, Chicago.
- Puri, P. S., & Goldie, C. M., (1979), Poisson mixtures and quasi-infinite divisibility of distributions. *J. Appl. Prob.* 16, 138-153.
- Ruohonen, M., (1988), A model for the claim number process. *ASTIN Bulletin*, 18, 57–68.
- Sankaran, M., (1970), The discrete Poisson–Lindley distribution. *Biometrics*, 26, 145–149.
- Sarguta, R. J., & Ottieno, J. A. M., (2015), Mixed Poisson distributions in terms of special functions. *Mathematical theory and modelling*, Vol.5, 245-274.
- Shaban, S. A., (1981), Computation of the Poisson-Inverse Gaussian distribution, *Communications in Statistics - Theory and Methods*. Vol. 10, no. 14, 1389-1399.
- Shaked, M., (1980), On mixtures from exponential families. *Journal of the Royal Statistical Society B*, 42, 192–198.
- Sibuya, M., (1979), Generalized hypergeometric, digamma and trigamma distributions. *Annals of the Institute of Statistical Mathematics*, 31, 373–390.

- Titterton, D.M., Smith, A.F.M. & Makov, U.E. (1985). *Statistical Analysis of Finite Mixture Distributions*. New York: John Wiley and Sons.
- Willmot, G.E., (1990), Asymptotic tail behaviour of Poisson mixtures with applications. *Advances in Applied Probability*, 22, 147–159.
- Willmot, G.E., (1993), On recursive evaluation of mixed Poisson probabilities and related quantities. *Scandinavian Actuarial Journal*, 18, 114–133.
- Willmot, G.E. & Sundt, B., (1989), On evaluation of the Delaporte distribution and related distributions. *Scandinavian Actuarial Journal*, 2, 101-113.
- Willmot, G.E. & Sundt, B., (1989), On posterior probabilities and moments in mixed Poisson processes. *Scandinavian Actuarial Journal*, 3, 139–146.
- Witkovsky, V., (2001), Computing the Distribution of a Linear Combination of Inverted Gamma Variables. *Kybernetika*, 37 (1): 79–90.
- Zha, L., Lord, D., & Zou, Y., (2014), The Poisson inverse Gaussian (PIG) generalized linear regression model for analyzing motor vehicle crash data. *Journal of Transportation Safety & Security*, 8(1), 18–35.