ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ ΚΑΙ ΑΣΦΑΛΙΣΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ

Π.Μ.Σ. ΣΤΗΝ ΑΝΑΛΟΓΙΣΤΙΚΗ ΕΠΙΣΤΗΜΗ ΚΑΙ ΔΙΟΙΚΗΤΙΚΗ ΚΙΝΔΥΝΟΥ

ΕΜΠΕΙΡΙΚΗ ΜΠΕΫΖΙΑΝΗ ΕΚΤΙΜΗΣΗ ΜΕ ΕΦΑΡΜΟΓΕΣ ΣΤΗΝ ΑΣΦΑΛΙΣΤΙΚΗ ΕΠΙΣΤΗΜΗ ΚΑΙ ΤΟΝ ΑΝΑΛΟΓΙΣΜΟ

Διπλωματική Εργασία

Λίγκρης Χριστόφορος

Επιβλέπων Καθηγητής: Πιτσέλης Γεώργιος

ΠΕΙΡΑΙΑΣ 2016

UNIVERSITY OF PIRAEUS

DEPARTMENT OF STATISTICS AND INSURANCE SCIENCE

M. Sc. IN ACTUARIAL SCIENCE AND RISK MANAGEMENT

EMPIRICAL BAYES ESTIMATION WITH APPLICATIONS IN ACTUARIAL SCIENCE

M. Sc. Thesis

Ligkris Christoforos

Advisor Professor: Pitselis Georgios

PIRAEUS 2016

Στη θεία μου, Βιβή.

Ευχαριστίες

Θα ήθελα να ευχαριστήσω όλους τους Καθηγητές που με δίδαξαν κατά τη διάρκεια των ακαδημαϊκών μου σπουδών. Ιδιαίτερα, θα ήθελα να ευχαριστήσω τον Επίκουρο Καθηγητή του τμήματος Στατιστικής και Ασφαλιστικής Επιστήμης κ. Γεώργιο Πιτσέλη για τη συνεργασία, την καθοδήγηση και τη διαρκή βοήθεια που μου παρείχε κατά την περίοδο της συγγραφής αυτής της διπλωματικής εργασίας. Θα ήθελα επίσης να ευχαριστήσω ιδιαίτερα τους Καθηγητές κ. Μιχαήλ Μπούτσικα για την πολύτιμη βοήθειά του και τις συμβουλές του πάνω στην προσομοίωση του πέμπτου Κεφαλαίου, και τους κ. Δημήτριο Αντζουλάκο και κ. Γεώργιο Τζαβελά, για την τιμή που μου έκαναν να είναι μέλη της συμβουλευτικής μου επιτροπής. Τέλος θα ήθελα βαθύτατα να ευχαριστήσω την οικογένειά μου και τους φίλους μου για τη συμπαράσταση και την υποστήριξή τους κατά τη διάρκεια των σπουδών μου.

ΠΕΡΙΛΗΨΗ

Σκοπός της παρούσας εργασίας είναι η παρουσίαση και η μελέτη της εμπειρικής Μπεϋζιανής προσέγγισης, μιας μεθόδου εκτίμησης, η οποία δανείζεται στοιχεία τόσο από την κλασσική όσο και από την Μπεϋζιανή στατιστική, με στόχο να επωφεληθεί από τα δυνατά σημεία και των δυο προαναφερθέντων προσεγγίσεων. Η βασική μεθοδολογία που ακολουθείται κατά την εμπειρική Μπεϋζιανή προσέγγιση, είναι αρχικά η μοντελοποίηση ενός προβλήματος, χρησιμοποιώντας Μπεϋζιανές τεχνικές, και έπειτα ακολουθεί η εκτίμηση παραμέτρων, με μεθόδους που χρησιμοποιούνται στην κλασσική στατιστική.

Η κύρια διαφορά μεταξύ της κλασσικής και της Μπεϋζιανής στατιστικής είναι ο ορισμός των άγνωστων παραμέτρων του στατιστικού μοντέλου. Στην πρώτη περίπτωση οι παράμετροι θεωρούνται άγνωστες σταθερές ποσότητες, ενώ στην δεύτερη τυχαίες μεταβλητές. Η διαφορά αυτή έχει οδηγήσει στη διαμάχη μεταξύ των κλασσικών και των Μπεϋζιανών στατιστικών, με τους κλασικούς να μην συμμερίζονται την έλλειψη αντικειμενικότητας των αποτελεσμάτων και τους Μπεϋζιανούς να θεωρούν πως εάν υπάρχει κάποια περαιτέρω εκ των προτέρων πληροφορία για την παράμετρο, δεν πρέπει να μένει ανεκμετάλλευτη.

Αρχικά θα γίνει παρουσίαση της Μπεϋζιανής μεθόδου, της οποίας η βασική μεθοδολογία και υποθέσεις είναι κοινές με εκείνες της εμπειρικής Μπεϋζιανής προσέγγισης. Έπειτα θα γίνει παρουσίαση της μη παραμετρικής και της παραμετρικής εμπειρικής Μπεϋζιανής προσέγγισης, ενώ θα γίνει και μελέτη της απόδοσης των δυο αυτών προσεγγίσεων. Στο τελευταίο μέρος της εργασίας, γίνεται μια συνοπτική παρουσίαση της θεωρίας αξιοπιστίας χαρτοφυλακίου και διερευνάται πότε τα αποτελέσματά της είναι κοινά με εκείνα της Μπεϋζιανής και της εμπειρικής Μπεϋζιανής προσέγγισης.

ABSTRACT

In this thesis, we focus on the study and presentation of the empirical Bayes estimation methods, which borrows elements, both form classical and Bayesian statistics. These methods allow the modelling of complicated problems using Bayesian techniques, but employ classical statistics methods for obtaining parameter estimates.

The main difference between classical and Bayesian statistics, is the definition of the statistical model's unknown parameters. In the first case, the parameters are considered to be unknown fixed quantities, whereas in the second one, they are random variables. It is that difference that has led to the dispute between classical and Bayesian supporting statisticians. On one hand, the classical statistics supporters, do not agree with the lack of objectivity in the results obtain using Bayesian methods and on the other hand, the Bayesian statistics supporters believe that if there is any further information regarding the parameter beforehand, it must be utilized.

Initially, the Bayes approach is presented, which makes use of the same basic methodology and hypotheses as the empirical Bayes approach. Then the parametric and non-parametric empirical Bayes is also presented, ending with the credibility theory and its connection with the Bayes and empirical Bayes approach.

ΠΕΡΙΕΧΟΜΕΝΑ

Κεφάλαιο 1

1. ΕΙΣΑΓΩΓΗ

Στην παρούσα εργασία θα γίνει η παρουσίαση και η μελέτη της εμπειρικής Μπεϋζιανής προσέγγισης, μιας μεθόδου εκτίμησης, η οποία δανείζεται στοιχεία και από την κλασσική και από την Μπεϋζιανή στατιστική, με στόχο να εκμεταλλευθεί τα δυνατά σημεία των δυο προαναφερθέντων προσεγγίσεων. Η βασική μεθοδολογία που ακολουθείται κατά την εμπειρική Μπεϋζιανή προσέγγιση, είναι αρχικά η μοντελοποίηση ενός προβλήματος, χρησιμοποιώντας Μπεϋζιανές τεχνικές, και έπειτα ακολουθεί η εκτίμηση παραμέτρων, με μεθόδους που χρησιμοποιούνται στην κλασσική στατιστική.

Τόσο στην κλασσική, όσο και στην Μπεϋζιανή στατιστική, το γενικό μοντέλο προβλήματος, έχει να κάνει με την συμπερασματολογία πάνω σε μια άγνωστη παράμετρο θ, βασιζόμενη σε ένα τυχαίο δείγμα n παρατηρήσεων $\boldsymbol{x} = x_1, x_2, \ldots, x_n$, των οποίων η κατανομή περιγράφεται από τη γνωστή συνάρτηση $f(\boldsymbol{x}|\theta)$.

Κατά την κλασική στατιστική, η παράμετρος θ θεωρείται πως έχει μια άγνωστη αλλά συγκεκριμένη τιμή, η οποία και εκτιμάται. Κατά την Μπεϋζιανή στατιστική όμως, η παράμετρος θ αντιμετωπίζεται σαν μια τυχαία μεταβλητή, η οποία περιγράφεται από μια εκ των προτέρων κατανομή. Αυτή είναι και η κύρια διαφορά των δυο αυτών προσεγγίσεων της στατιστικής.

Η εμπειρική Μπεϋζιανή προσέγγιση χωρίζεται σε δύο τύπους. Ο πρώτος είναι εκείνος της μη παραμετρικής εμπειρικής Μπεϋζιανής, της οποίας στόχος είναι η παρουσίαση της εκ των υστέρων μέσης τιμής σε όρους της περιθώριας κατανομής. Έπειτα χρησιμοποιώντας τα δεδομένα, γίνεται ο υπολογισμός του Μπεϋζιανού κανόνα απευθείας. Ο δεύτερος τύπος είναι εκείνος της παραμετρικής εμπειρικής Μπεϋζιανής, κατά την οποία θεωρείται πως η εκ των προτέρων πληροφορία της παραμέτρου ανήκει σε μια γνωστή παραμετρική κλάση, όπου μόνο οι υπερπαράμετροί της είναι άγνωστοι. Οι υπερπαράμετροι, εκτιμώνται από τα δεδομένα, με την χρήση κλασσικών μεθόδων εκτίμησης.

Ο πρώτος που εισήγαγε την έννοια της εμπειρικής Μπεϋζιανής προσέγγισης ήταν ο Robbins (1955). Ο Morris (1983) ήταν εκείνος που μελέτησε πρώτος, την παραμετρική εκδοχή της προσέγγισης. Ο Maritz το 1970, συγκέντρωσε την μέχρι τότε μελέτη πάνω στην εμπειρική Μπευζιανή προσέγγιση. Σημαντικά βιβλία για την εισαγωγή του αναγνώστη στην εμπειρική Μπεϋζιανή προσέγγιση έχουν γραφεί από τους Berger (1985), Gelman (2004) Lwin & Maritz (1989) και Carlin & Louis (2008) ενώ μελέτη πάνω στις δυνατότητες της εμπειρικής Μπεϋζιανής προσέγγισης έχει γίνει από τους Casella (1985 &1992), Efron (2010) και Norberg (1980).

Στο πρώτο Κεφάλαιο της εργασίας, θα γίνει αναλυτική παρουσίαση της Μπεϋζιανής μεθόδου, της οποίας η βασική μεθοδολογία και υποθέσεις είναι κοινές με εκείνες της εμπειρικής Μπεϋζιανής προσέγγισης. Αρχικά θα μελετηθούν οι έννοιες της συνάρτησης πιθανοφάνειας και της εκ των προτέρων κατανομής, και θα δούμε πώς προκύπτει η εκ των υστέρων κατανομή, μέσω του Μπεϋζιανού θεωρήματος. Ύστερα θα μελετηθεί ο τρόπος με τον οποίο πραγματοποιείται η συμπερασματολογία πάνω στην παράμετρο θ, ενώ θα γίνει αναφορά και στην κατανομή πρόβλεψης. Επίσης, θα παρουσιαστεί συνοπτικά η ιεραρχική Μπεϋζιανή προσέγγιση.

Τα Κεφάλαια 3 και 4, αφορούν την μη παραμετρική και παραμετρική εμπειρική Μπεϋζιανή προσέγγιση, όπου θα γίνει αναλυτική παρουσίασή τους, και θα μελετηθεί πληθώρα παραδειγμάτων, για την καλύτερη κατανόησή τους. Έπειτα, στο Κεφάλαιο 5, θα γίνει μελέτη της απόδοσης των δυο προσεγγίσεων, συγκρίνοντας την απόκλιση των εκτιμήσεών τους από τις πραγματικές τιμές των παραμέτρων, με την απόκλιση άλλων προσεγγίσεων.

Στο τελευταίο μέρος της εργασίας, γίνεται μια συνοπτική παρουσίαση της θεωρίας αξιοπιστίας, η οποία είναι ένας θεμελιώδης τομέας της αναλογιστικής επιστήμης και διερευνάται πότε τα αποτελέσματά της ταυτίζονται με εκείνα της Μπεϋζιανής και της εμπειρικής Μπεϋζιανής προσέγγισης.

Ο κώδικας που χρησιμοποιήθηκε για τη μελέτη του πέμπτου κεφαλαίου γράφτηκε μέσω του υπολογιστικού πακέτου *R* και παρατίθεται στο Παράρτημα Γ, ενώ τα γραφήματα έγιναν μέσω της εφαρμογής *Office Excel*.

Κεφάλαιο 2

2. ΜΠΕΫΖΙΑΝΗ ΠΡΟΣΕΓΓΙΣΗ

2.1 Μπεϋζιανή Μέθοδος

Στο κεφάλαιο αυτό θα αναφερθούμε στη βασική δομή της Μπεϋζιανής ανάλυσης. Η βασική διαφορά της Μπεϋζιανής και της κλασσικής στατιστικής, εντοπίζεται στον χειρισμό της παραμέτρου θ, η οποία στην πρώτη προσέγγιση αντιμετωπίζεται σαν μια τυχαία μεταβλητή, ενώ στην κλασσική προσέγγιση ως μια άγνωστη σταθερά (βλέπε Carlin & Louis, 2008).

Τα κύρια συστατικά της Μπεϋζιανής προσέγγισης είναι η συνάρτηση πιθανοφάνειας $f(\boldsymbol{x}|\theta)$, η οποία περιγράφει το τυχαίο δείγμα n παρατηρήσεων $\boldsymbol{x} = x_1, x_2, \ldots, x_n$, δεδομένου της παραμέτρου θ, και η εκ των προτέρων κατανομή $u(\theta)$, η οποία ποσοτικοποιεί τι είναι γνωστό για την παράμετρο θ, πριν λάβουμε υπόψιν μας τις πληροφορίες που προέρχονται από την παρατήρηση των δεδομένων.

Όταν η συνάρτηση πιθανοφάνειας συνδυαστεί με την εκ των προτέρων κατανομή, μέσω του Μπεϋζιανού θεωρήματος, λαμβάνουμε την εκ των υστέρων κατανομή ως:

$$p(\theta|\boldsymbol{x}) = \frac{f(\boldsymbol{x}|\theta)u(\theta)}{g(\boldsymbol{x})},$$

όπου

$$g(\boldsymbol{x}) = \int f(\boldsymbol{x}|\theta)u(\theta)\, d\theta\,,$$

είναι η περιθώρια συνάρτηση πυκνότητας των δεδομένων $\boldsymbol{x}$. Η εκ των υστέρων κατανομή $p(\theta|\boldsymbol{x})$, αποτυπώνει τη συνολική γνώση που έχουμε για την παράμετρο θ, αφού λάβουμε υπόψιν και τις πληροφορίες που πηγάζουν από τα παρατηρούμενα δεδομένα (βλέπε Glickman & van Dyk, 2007). Εφόσον καταλήξουμε στην εκ των υστέρων κατανομή, γίνεται δυνατή η εκτίμηση των παραμέτρων και η πρόβλεψη μελλοντικών παρατηρήσεων. Επιγραμματικά, τα βήματα της Μπεϋζιανής μεθόδου είναι τα ακόλουθα:

α. Καθορισμός του μοντέλου πιθανότητας των δεδομένων

β. Κατασκευή της συνάρτησης πιθανοφάνειας μετά από την παρατήρηση των δεδομένων

γ. Επιλογή της εκ των προτέρων κατανομής

δ. Σύνθεση της εκ των υστέρων κατανομής μέσω του Μπεϋζιανού θεωρήματος

ε. Συμπερασματολογία ως προς την παράμετρο θ, μέσω της εκ των υστέρων κατανομής

στ. Πρόβλεψη μελλοντικών παρατηρήσεων

Παρακάτω παρουσιάζονται αναλυτικά τα βήματα της Μπεϋζιανής μεθόδου, όπως και παραδείγματα για την καλύτερη κατανόησή της.

2.2 Η Συνάρτηση Πιθανοφάνειας *(Likelihood function)*

Η Μπεϋζιανή προσέγγιση θεωρεί, όπως ακριβώς και η κλασσική στατιστική, ένα τυχαίο δείγμα n παρατηρήσεων, δεδομένου μιας παραμέτρου θ. Το πρώτο βήμα λοιπόν και στις δυο προσεγγίσεις, είναι ο καθορισμός της κατανομής που ακολουθούν τα δεδομένα, με το σκεπτικό πως οι παράμετροί της είναι γνωστές.

Θεωρώντας πως οι παρατηρήσεις $\boldsymbol{x} = x_1, x_2, \ldots, x_n$ είναι ανεξάρτητες, ενδιαφερόμαστε να προσδιορίσουμε την $f(x_i|\theta)$. Η γνώση της δομής των δεδομένων τις περισσότερες φορές μας βοηθά στην επιλογή της κατανομής αυτής (βλέπε Carlin & Louis, 2008)

Αφού οι τιμές $\boldsymbol{x}$ είναι γνωστές, η συνάρτηση $f(\boldsymbol{x}|\theta)$ μπορεί να θεωρηθεί ως συνάρτηση της παραμέτρου θ. Τότε, η συνάρτηση ονομάζεται συνάρτηση πιθανοφάνειας του θ και συμβολίζεται ως $L(\theta;\boldsymbol{x})$.

Επί της ουσίας η συνάρτηση πιθανοφάνειας, είναι η από κοινού συνάρτηση πιθανότητας των δεδομένων και δηλώνει το πόσο πιθανό είναι υπό την παράμετρο θ, να προκύψουν τα συγκεκριμένα δεδομένα (βλέπε Glickman & van Dyk, 2007 και Ηλιόπουλο, 2006). Ισχύει δηλαδή ότι:

$$L(\theta;\boldsymbol{x}) = L(\theta|\boldsymbol{x}) = \prod_{i=1}^{n} f(x_i|\theta).$$

Κατά την κλασσική στατιστική, είναι πιθανό κάποια τιμή του θ να μεγιστοποιεί τη συνάρτηση πιθανοφάνειας. Αυτή η τιμή ονομάζεται εκτιμητής μεγίστης πιθανοφάνειας και είναι δηλαδή εκείνη η τιμή της παραμέτρου, για την οποία η πιθανότητα των παρατηρηθέντων δεδομένων παίρνει τη μεγαλύτερή της τιμή (βλέπε Agresti, 2007).

Η μέθοδος μεγίστης πιθανοφάνειας είναι μια από τις πιο συνηθισμένες μεθόδους εκτίμησης που χρησιμοποιούνται στην κλασσική στατιστική, που όμως δεν μπορεί να χρησιμοποιηθεί στην Μπεϋζιανή ανάλυση, αφού αλλάζει η φύση της παραμέτρου θ.

Αντί της τιμής που μεγιστοποιεί την συνάρτηση πιθανοφάνειας, στην Μπεϋζιανή προσέγγιση, θα μπορούσε να ειπωθεί πως η παράμετρος θ, εκτιμάται από τον μέσο όρο της πιθανοφάνειας, σταθμισμένο ως προς την εκ των προτέρων κατανομή (βλέπε Δελλαπόρτας & Τσιαμυρτζής, 2004).

2.3 Η Εκ των Προτέρων Κατανομή *(Prior Distribution)*

Στην Μπεϋζιανή ανάλυση λοιπόν, το θ αντί να θεωρείται πως έχει μια καθορισμένη πραγματική τιμή, αντιμετωπίζεται σαν μια τυχαία μεταβλητή η οποία ακολουθεί τη δική της κατανομή. Η κατανομή που ακολουθεί το θ ονομάζεται εκ των προτέρων κατανομή *(apriori),* γιατί δεν λαμβάνει υπόψιν της, τις πληροφορίες από τα παρατηρηθέντα δεδομένα και συμβολίζεται ως $u(\theta)$.

Ο προσδιορισμός της $u(\theta)$ διαφέρει από εκείνον της $f(\boldsymbol{x}\,|\,\theta)$, αφού σπάνια υπάρχουν λεπτομερείς διαθέσιμες πληροφορίες της τυχαίας μεταβλητής θ, οπότε ο αναλυτής πρέπει να βασιστεί σε εναλλακτικές μεθόδους για να προσεγγίσει την εκ των προτέρων κατανομή (βλέπε Klugman, 1992).

Ο καθορισμός της εκ των προτέρων κατανομής από τον αναλυτή, είναι ένα πολύ σημαντικό κομμάτι της Μπεϋζιανής ανάλυσης, αφού η επιλογή της έχει άμεση επίδραση στα αποτελέσματα που θα προκύψουν.

Γενικά οι εκ των προτέρων κατανομές χωρίζονται σε δυο βασικές κατηγορίες. Η πρώτη είναι αυτή των πληροφοριακών *(informative)* εκ των προτέρων κατανομών, κατά την οποία οι κατανομές κατασκευάζονται βάσει της γνώσης που προέρχεται από άλλα δεδομένα ή παλαιότερες μελέτες. Η δεύτερη βασική κατηγορία είναι αυτή των μη πληροφοριακών *(noninformative)* εκ των προτέρων κατανομών, οι οποίες υποδηλώνουν την άγνοια για τις παραμέτρους του μοντέλου (βλέπε Glickman & van Dyk, 2007).

Παρακάτω αναλύονται περαιτέρω οι προαναφερθείσες εκ των προτέρων κατανομές, αλλά αναφέρονται και κάποιες άλλες χαρακτηριστικές κατηγορίες αυτών:

2.3.1 Πληροφοριακές - Υποκειμενικές εκ των Προτέρων Κατανομές

Μια υποκειμενική εκ των προτέρων κατανομή έχει την μορφή που θα επιλέξει ο ίδιος ο αναλυτής, βασιζόμενος στην δική του πεποίθηση για το ποιες τιμές μπορεί να πάρει η παράμετρος θ (βλέπε Berger, 1985). Ονομαστικά παρατίθενται κάποιες μέθοδοι που συγκαταλέγονται σε αυτή την κατηγορία καθορισμού εκ των προτέρων κατανομών:

- Προσέγγιση Ιστογράμματος *(Histogram approach)*
- Προσέγγιση Σχετικής Πιθανοφάνειας *(Relative likelihood approach)*
- Ταίριασμα μιας Λειτουργικής Μορφής *(Matching a given functional form)*
- Καθορισμός Αθροιστικής Συνάρτησης Κατανομής *(CDF determination)*

2.3.2 Μη Πληροφοριακές - Αντικειμενικές εκ των Προτέρων Κατανομές

Οι μη πληροφοριακές εκ των προτέρων κατανομές, είναι κατανομές οι οποίες αντιστοιχούν χονδρικά ίσες πιθανότητες σε ένα μεγάλο φάσμα ενδεχομένων τιμών. Γενικά μια μη πληροφοριακή εκ των προτέρων κατανομή δεν έχει μεγάλη επιρροή στη διαμόρφωση της εκ των υστέρων κατανομής, έχοντας σαν αποτέλεσμα η τελευταία να είναι ελάχιστα διαφοροποιημένη σε σχέση με τη συνάρτηση πιθανοφάνειας (βλέπε Πανάρετο & Ξεκαλάκη, 2000). Οι πιο συνηθισμένες μέθοδοι κατασκευής μιας τέτοιας εκ των προτέρων είναι οι κάτωθι:

- Εκ των προτέρων του Jefrey *(Jefrey's prior)*
- Ομοιόμορφη *(Uniform)*

2.3.3 Συζυγείς *(Conjugate)* εκ των Προτέρων Κατανομές

Συζυγείς ονομάζονται οι οικογένειες κατανομών, οι οποίες όταν επιλεχθούν ως εκ των προτέρων κατανομές, τότε και οι εκ των υστέρων κατανομές θα ανήκουν στην ίδια οικογένεια αλλά έχοντας διαφορετικές παραμέτρους. Οι συζυγείς εκ των προτέρων κατανομές μπορούν να είναι είτε πληροφοριακές είτε μη πληροφοριακές κατανομές, και επιλέγονται διότι διευκολύνουν το μαθηματικό υπολογισμό της εκ των υστέρων κατανομής. Η επιλογή μιας συζυγούς εκ των υστέρων κατανομής μπορεί να γίνει μόνο όταν δεν έρχεται σε αντίθεση με τις πεποιθήσεις που υπάρχουν για τη μορφή της εκ των προτέρων. Πίνακας με τους πιο συνήθεις συνδυασμούς συζυγών κατανομών παρατίθεται κάτωθι (βλέπε Robert, 2007).

$f(x\|\theta)$	$u(\theta)$	$p(\theta\|x)$
$Poisson$ $P(\theta)$	$Gamma$ $G(\alpha,\beta)$	$Gamma$ $G(\alpha+x,\beta+1)$
$Gamma$ $G(n,\theta)$	$Gamma$ $G(\alpha,\beta)$	$Gamma$ $G(\alpha+n,\beta+x)$
$Binomial$ $Bin(n,\theta)$	$Beta$ $B(\alpha,\beta)$	$Beta$ $B(\alpha+x,\beta+n-x)$
$Negative\ Binomial$ $NB(m,\theta)$	$Beta$ $B(\alpha,\beta)$	$Beta$ $B(\alpha+m,\beta+x)$
$Normal$ $N\left(\mu,\frac{1}{\theta}\right)$	$Gamma$ $G(\alpha,\beta)$	$Gamma$ $G\left(\alpha+\frac{1}{2},\beta+\frac{(\mu-x)^2}{2}\right)$
$Normal$ $N(\theta,\sigma^2)$	$Normal$ $N(\mu,r^2)$	$Normal$ $N\left(\frac{\sigma^2\mu+r^2x}{\sigma^2+r^2},\frac{\sigma^2r^2}{\sigma^2+r^2}\right)$

Πίνακας 2.3 *Φυσικές Συζυγείς εκ των Προτέρων Κατανομές. Πηγή: Robert (2001)*

2.4 Το Μπεϋζιανό Θεώρημα *(Bayes Theorem)*

Το θεώρημα του Bayes δίνει την δυνατότητα, να συσχετιστεί η πιθανότητα ενός ενδεχομένου, με το αν πραγματοποιείται ή όχι ένα άλλο ενδεχόμενο. Μέσω του θεωρήματος αυτού, γίνεται δυνατή η σύνθεση της εκ των υστέρων κατανομής. Η μαθηματική αποτύπωσή του είναι η ακόλουθη (βλέπε Bayes & Price, 1763):

Θεώρημα 2.1

Έστω $\{B_1, B_2, \ldots, B_n\}$ μια διαμέριση του δειγματικού χώρου Ω, τέτοια ώστε $P(B_i) > 0$ για όλα τα $i = 1, 2, \ldots, n$. Τότε, για κάθε ενδεχόμενο Α του ίδιου δειγματικού χώρου, με $P(A) > 0$, ισχύει

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \cdots + P(A|B_n)P(B_n)}$$

$$= \frac{P(A|B_i)P(B_i)}{\sum_{j=1}^{n} P(A|B_j)P(B_j)}, \qquad i = 1, 2, \dots , n\,.$$

Για συνεχείς τυχαίες μεταβλητές X και Y, ο παραπάνω τύπος μετατρέπεται ως εξής:

$$f_{X|Y}(x|y) = \frac{f_{Y|X}(y|x) f_X(x)}{\int f_{Y|X}(y|x) f_X(x)\, dx}$$

$$= \frac{f_{Y|X}(y|x) f_X(x)}{f_Y(y)}.$$

2.5 Η εκ των Υστέρων Κατανομή

Από τη στιγμή που έχει προσεγγιστεί το είδος και η μορφή της συνάρτησης πιθανοφάνειας $L(\theta; \boldsymbol{x})$ και έχει πραγματοποιηθεί η κατασκευή της εκ των προτέρων κατανομής $u(\theta)$, είναι πλέον δυνατή η σύνθεση της εκ των υστέρων κατανομής μέσω της εφαρμογής του Μπεϋζιανού θεωρήματος :

$$p(\theta|\boldsymbol{x}) = \frac{f(\boldsymbol{x}|\theta) u(\theta)}{\int f(\boldsymbol{x}|\theta) u(\theta) d\theta}$$

$$= \frac{L(\theta|\boldsymbol{x}) u(\theta)}{g(\boldsymbol{x})} \propto L(\theta|\boldsymbol{x}) u(\theta)\,,$$

όπου το ‘∝’ σημαίνει “ανάλογο του”. Δηλαδή καταλήγουμε στο συμπέρασμα πως η εκ των υστέρων κατανομή $p(\theta|\boldsymbol{x})$, είναι ανάλογη του γινομένου της συνάρτησης πιθανοφάνειας $L(\theta|\boldsymbol{x})$ με την εκ των προτέρων κατανομή $u(\theta)$.

Αυτό συμβαίνει επειδή, η περιθώρια συνάρτηση $g(\boldsymbol{x})$ είναι ανεξάρτητη της τυχαίας μεταβλητής θ, οπότε αντιμετωπίζεται ως σταθερά (βλέπε Glickman & van Dyk, 2007).

2.6 Μπεϋζιανή Συμπερασματολογία

Έχοντας πλέον την εκ των υστέρων κατανομή της τυχαίας μεταβλητής Θ, γίνεται εφικτή η συμπερασματολογία πάνω σε αυτήν. Η συνάρτηση πιθανότητας μιας τυχαίας μεταβλητής αποτυπώνει το ποιες τιμές μπορεί να πάρει αυτή και το πόσο πιθανό είναι να πάρει κάθε μια από αυτές. Στην πραγματικότητα λοιπόν, η ίδια η εκ των υστέρων κατανομή $p(\theta|\boldsymbol{x})$ είναι το συμπέρασμά μας για την παράμετρο θ. Όμως είναι πολύ χρήσιμο αυτή η πληροφορία να μπορεί να αποδοθεί και περιληπτικά (βλέπε Πανάρετο & Ξεκαλάκη, 2000).

2.6.1 Σημειακή Εκτίμηση

Κάνοντας μια σημειακή εκτίμηση, ουσιαστικά τίθεται μια τιμή "εκπρόσωπος" της τυχαίας μεταβλητής. Για να διαπιστωθεί ποια τιμή όμως μπορεί να εκτιμήσει καλύτερα την τυχαία μεταβλητή, πρέπει να δημιουργηθεί μια συνάρτηση απώλειας.

2.6.1.1 Συναρτήσεις απώλειας *(Loss Functions)*

Μια συνάρτηση απώλειας ουσιαστικά μετράει το πόσο εσφαλμένη είναι η εκτίμηση, δηλαδή το πόσο διαφέρει η εκτίμηση της παραμέτρου, από την πραγματική της τιμή (βλέπε Robert, 2007). Αυτό σημάνει πως στόχος είναι, η τιμή της εκτίμησης να ελαχιστοποιεί τη συνάρτηση απώλειας. Ο συμβολισμός που θα χρησιμοποιηθεί για τη συνάρτηση απώλειας, που μετρά το κόστος της εκτίμησης της παραμέτρου θ, από ένα στοιχείο απόφασης δ (βλέπε Savchuk, Tsokos, 2011), είναι ο $L(\theta, \delta)$. Οι συνηθέστερες συναρτήσεις απώλειας που χρησιμοποιούνται, είναι οι παρακάτω :

- Συνάρτηση Απώλειας Τετραγωνικού Σφάλματος *(Squared Error Loss Function)*

$$L(\theta, \delta) = (\theta - \delta)^2$$

- Συνάρτηση Απώλειας Απόλυτου Σφάλματος *(Absolute Error Loss Function)*

$$L(\theta, \delta) = |(\theta - \delta)|$$

- Γραμμική Συνάρτηση Απώλειας *(Linear Constant Loss Function)*

$$L(\theta, \delta) = \begin{cases} a(\theta - \delta) \text{ όταν } \hat{\theta} < \theta \\ b(\theta - \delta) \text{ όταν } \hat{\theta} > \theta \end{cases}$$

- Συνάρτηση Απώλειας “0-1” *(“0-1” Loss Function)*

$$L(\theta,\delta) = \begin{cases} 1 \text{ όταν } \hat{\theta} \neq \theta \\ 0 \text{ όταν } \hat{\theta} = \theta \end{cases}$$

Το ποια συνάρτηση απώλειας θα χρησιμοποιηθεί, εξαρτάται από το είδος της “ποινής” που είναι επιθυμητό να υποβληθεί. Για παράδειγμα, κατά τη συνάρτηση απώλειας απόλυτου σφάλματος, η υποεκτίμηση και η υπερεκτίμηση της παραμέτρου, τιμωρούνται με τον ίδιο τρόπο, ανάλογα με την απόκλιση της εκτίμησης. Αντίθετα με την παραπάνω, η γραμμική συνάρτηση απώλειας τιμωρεί την υποεκτίμηση, με διαφορετικό βάρος από ότι την υπερεκτίμηση (βλέπε Δελλαπόρτας & Τσιαμυρτζής, 2004).

2.6.1.2 Συνάρτηση Κινδύνου *(Risk Function)*

Η αξιολόγηση των κανόνων απόφασης γίνεται μέσω της συνάρτησης κινδύνου, η οποία ορίζεται ως η μέση τιμή της συνάρτησης απώλειας. Αν $L(\delta,\theta)$ είναι η συνάρτηση απώλειας και *δ=δ(x)* ο κανόνας απόφασης, η συνάρτηση κινδύνου του κανόνα απόφασης *δ*, αποτυπώνεται μαθηματικά ως:

$$\begin{aligned} R(\delta,\theta) &= E_\theta L(\delta,\theta) \\ &= \int L(\delta(x),\theta) f(x|\theta)\, dx\,. \end{aligned}$$

2.6.1.3 Μπεϋζιανός Κίνδυνος - Κανόνας *(Bayes Risk - Rule)*

Στην Μπεϋζιανή προσέγγιση, δίνεται η δυνατότητα να σταθμιστεί η συνάρτηση κινδύνου του κάθε κανόνα απόφασης, βάσει της εκ των προτέρων προσωπικής άποψης, μέσω της *u(θ)*. Η ποσότητα που το επιτυγχάνει ονομάζεται Μπεϋζιανός κίνδυνος. Μαθηματικά ορίζεται ως:

$$\begin{aligned} BR(\delta) &= \int_\Theta R(\delta,\theta) u(\theta)\, d\theta \\ &= \int_\Theta \int_X L(\delta(x),\theta) f(x|\theta) u(\theta)\, dx\, d\theta\,. \end{aligned}$$

Ο κανόνας απόφασης που ελαχιστοποιεί τον Μπεϋζιανό κίνδυνο ονομάζεται Μπεϋζιανός κανόνας.

Επειδή στο επόμενο κεφάλαιο θα ασχοληθούμε μόνο με τον Μπεϋζιανό κανόνα που προκύπτει υπό την συνάρτηση απώλειας τετραγωνικού σφάλματος, θα είναι ο μόνος που θα υπολογιστεί αναλυτικά στο κύριο μέρος της μελέτης.

Απόδειξη 2.1

Ο κίνδυνος Bayes του κανόνα απόφασης δ, υπό την συνάρτηση απώλειας τετραγωνικού σφάλματος ισούται με:

$$E(\delta-\theta)^2 = \int_{\Theta}\int_{X}(\delta(x)-\theta)^2\, f(x|\theta)\, dx\, u(\theta)\, d\theta$$

$$= \int_{\Theta}\int_{X}(\delta(x)-\theta)^2\, f(x|\theta)\, u(\theta) dx\, d\theta$$

$$= \int_{\Theta}\int_{X}(\delta(x)-\theta)^2\, f(x,\theta)\, dx\, d\theta$$

$$= \int_{X}\int_{\Theta}(\delta(x)-\theta)^2\, p(\theta|x) g(x)\, d\theta\, dx$$

$$= \int_{X}\int_{\Theta}(\delta(x)-\theta)^2\, p(\theta|x)\, d\theta\, g(x)\, dx\, .$$

Η αλλαγή στη σειρά ολοκλήρωσης πραγματοποιείται μέσω του θεωρήματος Fubini, θεωρώντας πως πληρούνται όλες οι απαραίτητες προϋποθέσεις για την εφαρμογή του (βλέπε Gruber, 2014). Επειδή η περιθώρια συνάρτηση $g(\boldsymbol{x})$ είναι μια μη αρνητική συνάρτηση μας ενδιαφέρει να ελαχιστοποιήσουμε την

$$I(x) = \int_{\Theta}(\delta(x)-\theta)^2\, p(\theta|x)\, d\theta\, ,$$

η οποία γίνεται

$$I(x) = \int_{\Theta}(\delta(x)^2 - 2\theta\delta(x) + \theta^2)\, p(\theta|x)\, d\theta$$

$$= \delta(x)^2 \int_{\Theta} p(\theta|x)\, d\theta - 2\delta(x) \int_{\Theta} \theta p(\theta|x)\, d\theta + \int_{\Theta} \theta^2 p(\theta|x)\, d\theta$$

$$= \int_{\Theta} p(\theta|x)\, d\theta \left(\delta(x)^2 - 2\delta(x) \frac{\int_{\Theta} \theta p(\theta|x)\, d\theta}{\int_{\Theta} p(\theta|x)\, d\theta} \right) + \int_{\Theta} \theta^2 p(\theta|x)\, d\theta$$

$$= \int_{\Theta} p(\theta|x)\, d\theta \left(\delta(x) - \frac{\int_{\Theta} \theta p(\theta|x)\, d\theta}{\int_{\Theta} p(\theta|x)\, d\theta} \right)^2 + \left[\int_{\Theta} \theta^2 p(\theta|x)\, d\theta - \frac{\left(\int_{\Theta} \theta p(\theta|x)\, d\theta\right)^2}{\int_{\Theta} p(\theta|x)\, d\theta} \right].$$

Η παραπάνω συνάρτηση ελαχιστοποιείται ως προς το $\delta(x)$ όταν

$$\delta(x) = \frac{\int_{\Theta} \theta p(\theta|x)\, d\theta}{\int_{\Theta} p(\theta|x)\, d\theta} = E(\theta|x).$$

2.6.2 Εκτίμηση Μέσω Διαστήματος

Η πληροφορία που δίνει η σημειακή εκτίμηση, πολλές φορές μπορεί να φανεί περιορισμένη ή και ανακριβής. Γι' αυτό το λόγο η συνοπτική περιγραφή των τιμών που μπορεί να πάρει η παράμετρος *θ*, μπορεί να αποδοθεί με περισσότερη λεπτομέρεια μέσω μιας εκτίμησης σε διάστημα. Στην κλασσική στατιστική η εκτίμηση σε διάστημα γίνεται μέσω των διαστημάτων εμπιστοσύνης. Η ερμηνεία ενός διαστήματος εμπιστοσύνης της μορφής

$$P\left(K_{\alpha} \leq \theta \geq K_{\beta}\right) = 1 - \alpha ,$$

είναι πως σε *n* επαναλαμβανόμενα πειράματα τύχης, η τιμή της παραμέτρου θα βρισκόταν *n(1-α)* φορές εντός του διαστήματος *K*. Η ερμηνεία της εκτίμησης σε διάστημα στη Μπεϋζιανή εκτίμηση είναι πιο βολική, αφού η παράμετρος πλέον είναι τυχαία μεταβλητή. Αυτό σημάνει πως για ένα διάστημα της μορφής

$$\int_{K} p(\theta|x)\, d\theta = 1 - \alpha,$$

η τιμή της παραμέτρου *θ* εμπεριέχεται στο διάστημα αυτό με πιθανότητα *1-α* (βλέπε Carlin & Louis, 2008).

2.6.3 Έλεγχος Υποθέσεων

Μέσω ενός ελέγχου υποθέσεων μπορούμε να διαπιστώσουμε την ορθότητα μιας υπόθεσης, που αφορά την παράμετρο θ, χωρίς να ασχοληθούμε με την εκτίμηση των πιθανών τιμών της. Στην Μπεϋζιανή ανάλυση, ο έλεγχος υποθέσεων πραγματοποιείται μέσω της αξιολόγησης του Μπεϋζιανού παράγοντα *(Bayes factor)*, ο οποίος προτάθηκε από τον Jeffreys *(1961)*.

Ας υποθέσουμε πως για τα παρατηρηθέντα δεδομένα $\boldsymbol{x}$, υπάρχουν δυο υποψήφια παραμετρικά μοντέλα, M_1 και M_2 (υπό τις αντίστοιχες υποθέσεις H_1 και H_2), που είναι πιθανό να τα περιγράφουν, και έχουν αντίστοιχα τα θ_1 και θ_2 για παραμέτρους. Κάθε παράμετρος με τη σειρά της περιγράφεται από τις εκ των προτέρων κατανομές $u_1(\theta_1)$ και $u_2(\theta_2)$ (βλέπε Carlin & Louis, 2008).

Η περιθώρια κατανομή των $\boldsymbol{x}$ είναι η κάτωθι

$$g(\boldsymbol{x}|M_i) = \int f(\boldsymbol{x}|\theta_i, M_i)\ u_i(\theta_i)d\theta_i\,, \quad \textit{όπου}\ i = 1,2\,.$$

Με την εφαρμογή του Μπεϋζιανου θεωρήματος, λαμβάνουμε τις εκ των υστέρων πιθανότητες $P(M_1|\boldsymbol{x})$ και $P(M_2|\boldsymbol{x}) = 1 - P(M_1|\boldsymbol{x})$ των δυο μοντέλων.

Το ποιο από τα δυο μοντέλα θα γίνει αποδεκτό, θα εξεταστεί μέσω του Μπεϋζιανού παράγοντα, ο οποίος ορίζεται ως ο λόγος της εκ των υστέρων απόδοσης *(posterior odds)* του M_1 προς την εκ των προτέρων απόδοση *(prior odds)* του M_1.

$$\begin{aligned} BF &= \frac{P(M_1|\boldsymbol{x})}{P(M_2|\boldsymbol{x})}\Big/\frac{P(M_1)}{P(M_2)} \\ &= \left(\frac{g(\boldsymbol{x}|M_1)P(M_1)}{g(\boldsymbol{x})}\right)P(M_2)\Big/\left(\frac{g(\boldsymbol{x}|M_2)P(M_2)}{g(\boldsymbol{x})}\right)P(M_1) \\ &= \frac{g(\boldsymbol{x}|M_1)}{g(\boldsymbol{x}|M_2)}. \end{aligned}$$

Στην περίπτωση, που τα μοντέλα M_1 και M_2 μοιράζονται την ίδια παραμετροποίηση και οι δυο υποθέσεις είναι απλές ($M_i = \theta^{(i)}$), η $u_i(\theta_i)$ έχει μάζα πιθανότητας στο σημείο $\theta^{(i)}$. Έτσι η μαθηματική απόδοση του Μπεϋζιανού παράγοντα, μετατρέπεται στο λόγο των πιθανοφανειών των δυο μοντέλων, όπως κάτωθι:

$$BF = \frac{L(\theta^{(1)}|\boldsymbol{x})}{L(\theta^{(2)}|\boldsymbol{x})}.$$

Ανάλογα με την τιμή που θα έχει ο Μπεϋζιανός παράγοντας, κρίνεται το ποιο από τα δυο μοντέλα υπερτερεί του άλλου, αλλά και κατά πόσο ισχυρός είναι ο ισχυρισμός αυτός. Όταν ισχύει $BF > 1$, σημαίνει πως το μοντέλο M_1 υποστηρίζεται από τα δεδομένα, περισσότερο από ότι το μοντέλο M_2. Πίνακας με τις τιμές που προτείνονται για την αξιολόγηση των υποθέσεων από τον Jeffreys (1961) παρατίθεται παρακάτω:

Jeffreys

BF	**Ισχύς των στοιχείων**
$< 10^0$	Καμία ισχύς (υπέρ M_2)
10^0 έως $10^{1/2}$	Οριακά άξια αναφοράς
$10^{1/2}$ έως 10^1	Ουσιώδης
10^1 έως $10^{3/2}$	Ισχυρή
$10^{3/2}$ έως 10^2	Πολύ ισχυρή
$> 10^2$	Κατηγορηματική

Πίνακας 2.6.1 *Κατηγοριοποίηση τιμών Μπεϋζιανού παράγοντα κατά Jeffreys* (1961)

Από τους Kass και Raftery (1995), προτάθηκε η κάτωθι αξιολόγηση των υποθέσεων:

Kass & Raftery

$2\ ln(BF)$	**Ισχύς των στοιχείων**
0 έως 2	Καμία ισχύς
0 έως 6	Ουσιώδης
6 έως 10	Ισχυρή
> 10	Πολύ ισχυρή

Πίνακας 2.6.2 *Κατηγοριοποίηση τιμών Μπεϋζιανού παράγοντα κατά Kass & Raftery* (1995)

2.7 Κατανομή Πρόβλεψης

Ίσως το πιο βασικό κίνητρο που μας οδηγεί στην ανάπτυξη ενός στατιστικού μοντέλου είναι η δυνατότητα πρόβλεψης των μελλοντικών τιμών της διαδικασίας. Ένα από τα πλεονεκτήματα της Μπεϋζιανής ανάλυσης είναι πως, μέσω της εκ των υστέρων κατανομής, η διαδικασία της πρόβλεψης είναι άμεση και σαν διαδικασία ακριβής.

Υποθέτοντας λοιπόν πως για τυχαίο δείγμα παρατηρήσεων $\boldsymbol{x} = x_1, x_2, \ldots, x_n$ είναι επιθυμητή η πρόβλεψη της μελλοντικής παρατήρησης y, είναι απαραίτητος ο υπολογισμός της εκ των υστέρων κατανομής πρόβλεψης $\pi(y|\boldsymbol{x})$.

Για την πλήρη κατανόηση της διαδικασίας του μαθηματικού υπολογισμού της εκ των υστέρων κατανομής πρόβλεψης, θα γίνει πρώτα η μαθηματική ανάλυση της εκ των προτέρων κατανομής πρόβλεψης $g(\boldsymbol{x})$.

$$\begin{aligned} g(\boldsymbol{x}) &= \int_\theta f(\boldsymbol{x},\theta)\, d\theta \\ &= \int_\theta f(\boldsymbol{x}|\theta)u(\theta)\, d\theta\,. \end{aligned}$$

Παρατηρώντας την εξίσωση, είναι εύκολο να διαπιστωθεί πως το ολοκλήρωμα εμπεριέχει το γινόμενο της συνάρτησης πιθανοφάνειας $f(\boldsymbol{x}|\theta)$, με την εκ των προτέρων κατανομή $u(\theta)$. Έτσι, βασιζόμενοι στην παραπάνω διαδικασία, θα γίνει ο μαθηματικός ορισμός και της εκ των υστέρων κατανομής πρόβλεψης $\pi(y|\boldsymbol{x})$.

$$\begin{aligned} \pi(y|\boldsymbol{x}) &= \int_\theta f(y,\theta|\boldsymbol{x})\, d\theta \\ &= \int_\theta f(y|\theta,\boldsymbol{x})p(\theta|\boldsymbol{x})\, d\theta \\ &= \int_\theta f(y|\theta)p(\theta|\boldsymbol{x})\, d\theta\,. \end{aligned}$$

Όπως και προηγουμένως, η συνάρτηση πιθανότητας των επερχόμενων παρατηρήσεων, είναι ίση με το ολοκλήρωμα της από κοινού συνάρτησης πιθανότητας $f(y,\theta|\boldsymbol{x})$. Πρέπει να παρατηρηθεί πως αυτή την φορά η από κοινού συνάρτηση κατανομής είναι δεσμευμένη ως προς τα παρατηρηθέντα δεδομένα.

Εφαρμόζοντας ξανά το Μπεϋζιανό θεώρημα, οδηγούμαστε στο συμπέρασμα πως η εκ των υστέρων κατανομή πρόβλεψης ισούται ξανά με το γινόμενο της πιθανοφάνειας $f(y|\theta)$, αλλά αυτή τη φορά με την εκ των υστέρων κατανομή $p(\theta|\boldsymbol{x})$, όταν αυτό ολοκληρωθεί ως προς την μεταβλητή θ. Είναι φανερή λοιπόν η επήρεια της εκ των υστέρων πληροφόρησης.

Πρέπει να σημειωθεί πως από την συνάρτηση πιθανότητας $f(y|\theta,\boldsymbol{x})$ καταλήξαμε στην συνάρτηση πιθανοφάνειας, υπό την παραδοχή πως όλη η εκ των υστέρων πληροφορία πηγάζει από την παράμετρο θ, οπότε η δέσμευση ως προς y είναι περιττή.

2.8 Ιεραρχική Μπεϋζιανή Ανάλυση *(Hierarchical Bayes Analysis)*

Κατά την Μπεϋζιανή προσέγγιση, η παράμετρος θ του μοντέλου αντιμετωπίζεται ως μια άγνωστη ποσότητα, η οποία περιγράφεται από μια εκ των προτέρων κατανομή. Στη μέχρι τώρα μελέτη, έχει γίνει η παραδοχή πως η παράμετρος (υπερπαράμετρος) της εκ των προτέρων κατανομής (έστω λ), είναι γνωστή. Στην πραγματικότητα, αυτό δεν είναι δεδομένο, με την υπερπαράμετρο να μπορεί να θεωρηθεί εξίσου μια άγνωστη ποσότητα, η οποία περιγράφεται από μια υπέρ εκ των προτέρων κατανομή *(Hyperprior)*, έστω *h(λ)*.

Σε αυτή την περίπτωση η εκ των υστέρων κατανομή $p(\theta|x)$, είναι η

$$p(\theta|x) = \frac{\int f(x|\theta)u(\theta|\lambda)\, h(\lambda)d\lambda}{\int\int f(x|\theta)u(\theta|\lambda)h(\lambda)\, d\theta d\lambda}.$$

Ορισμός 2.1

Γενικά, ιεραρχικό Μπεϋζιανό μοντέλο θεωρείται το Μπεϋζιανό μοντέλο όπου η εκ των προτέρων κατανομή $u(\theta)$ (βλέπε Robert, *2007), μπορεί να αποδομηθεί στις δεσμευμένες κατανομές*

$$u_1(\theta|\theta_1), u_2(\theta_1|\theta_2),\ \dots,\ u_n(\theta_{n-1}|\theta_n)\ ,$$

και στην περιθώρια $u_{n+1}(\theta_n)$, όπως κάτωθι

$$u(\theta) = \int_{\Theta_1\times\ldots\times\Theta_n} u_1(\theta|\theta_1)u_2(\theta_1|\theta_2)\ldots u_{n+1}(\theta_n)d\theta_1\ldots d\theta_{n+1}\,,$$

όπου με θ_i συμβολίζεται η υπερπαράμετρος του i επιπέδου, $i = 1, \dots, n$.

Από τον παραπάνω ορισμό, γίνεται φανερό πως η μελέτη μπορεί να συνεχιστεί σε μεγαλύτερο βάθος, θεωρώντας τυχαίες μεταβλητές τις υπερπαραμέτρους, περισσότερων από δυο επιπέδων. Βέβαια η ιεραρχία αυτή κάπου πρέπει να τελειώσει, θεωρώντας πάντα την υπερπαράμετρο του τελευταίου επιπέδου γνωστή.

Σημείωση

Το Κεφάλαιο αυτό ήταν αφιερωμένο στη βασική ανάλυση της Μπεϋζιανής μεθόδου. Στο υπόλοιπο μέρος του Κεφαλαίου που ακολουθεί, θα μελετηθούν παραδείγματα τα οποία θα βοηθήσουν στην κατανόηση των προαναφερθέντων θεμάτων. Πληθώρα επιπλέον παραδειγμάτων έχουν δοθεί από τους Gelman (2004), Berger (1985) και από τους Carlin & Louis (2008).

Παράδειγμα 2.1: $f(x_i|\theta) \sim Poisson(\theta), u(\theta) \sim Gamma(\alpha, \beta)$

Ας υποθέσουμε ότι παρατηρούμε τον αριθμό απαιτήσεων ενός τυχαία επιλεγόμενου ασφαλισμένου οδηγού από ένα χαρτοφυλάκιο μιας ασφαλιστικής εταιρίας. Έστω ότι ο αριθμός των απαιτήσεων ενός χρόνου ακολουθεί την κατανομή *Poisson (θ),* $\theta \in \Theta = (0, \infty)$ (βλέπε Klugman, 1992), δηλαδή

$$f(x_i|\theta) = e^{-\theta} \frac{\theta^{x_i}}{x_i!}, \qquad i = 1,2, \dots, n.$$

Για *n* χρόνια παρατήρησης, η πιθανοφάνεια του δείγματος είναι

$$\begin{aligned} L(\theta|\boldsymbol{x}) &= \prod_{i=1}^{n} f(x_i|\theta) \\ &= \prod_{i=1}^{n} e^{-\theta} \frac{\theta^{x_i}}{x_i!} \\ &= e^{-n\theta} \frac{\theta^{\sum x_i}}{\prod x_i!}. \end{aligned}$$

Έστω ότι η εκ των προτέρων κατανομή του *θ* ακολουθεί την κατανομή *Gamma (α,β)* όπου *α,β > 0*, είναι γνωστές ποσότητες, με συνάρτηση πυκνότητας πιθανότητας

$$u(\theta) = \frac{\theta^{\alpha-1} e^{-\frac{\theta}{\beta}}}{\beta^{\alpha} \Gamma(\alpha)}.$$

Επομένως η περιθώρια κατανομή είναι

$$\begin{aligned} g(\boldsymbol{x}) &= \int_0^{\infty} L(\theta|\boldsymbol{x}) \; u(\theta) \, d\theta \\ &= \int_0^{\infty} \left(e^{-n\theta} \frac{\theta^{\sum x_i}}{\prod x_i!} \right) \left(\frac{\theta^{\alpha-1} e^{-\theta/\beta}}{\beta^{\alpha} \Gamma(\alpha)} \right) d\theta \end{aligned}$$

$$= \frac{1}{\beta^{\alpha}\Gamma(\alpha)\prod x_i!}\int_0^{\infty} e^{-n\theta-\theta/\beta}\ \theta^{\sum x_i+\alpha-1}\ d\theta$$

$$= \frac{\left(\frac{\beta}{n\beta+1}\right)^{\sum x_i+\alpha}\Gamma(\sum x_i+\alpha)}{\beta^{\alpha}\Gamma(\alpha)\prod x_i!}\int_0^{\infty}\frac{e^{-\theta\left(\frac{n\beta+1}{\beta}\right)}\ \theta^{(\sum x_i+\alpha)-1}}{\left(\frac{\beta}{n\beta+1}\right)^{\sum x_i+\alpha}\Gamma(\sum x_i+\alpha)}\ d\theta$$

$$= \frac{\left(\frac{\beta}{n\beta+1}\right)^{\sum x_i+\alpha}\Gamma(\sum x_i+\alpha)}{\beta^{\alpha}\Gamma(\alpha)\prod x_i!}.$$

Συνεπώς η εκ των υστέρων κατανομή του θ είναι

$$p(\theta|\boldsymbol{x}) = \frac{L(\theta|\boldsymbol{x})u(\theta)}{g(\boldsymbol{x})}$$

$$= \frac{\left(e^{-n\theta}\frac{\theta^{\sum x_i}}{\prod x_i!}\right)\left(\frac{\theta^{\alpha-1}e^{-\theta/\beta}}{\beta^{\alpha}\Gamma(\alpha)}\right)}{g(x)}$$

$$= \left\{\frac{e^{-\theta\left(\frac{n\beta+1}{\beta}\right)}\ \theta^{\sum x_i+\alpha-1}}{\prod x_i!\,\beta^{\alpha}\Gamma(\alpha)}\right\}\Bigg/\left\{\frac{\left(\frac{\beta}{n\beta+1}\right)^{\sum x_i+\alpha}\Gamma(\sum x_i+\alpha)}{\beta^{\alpha}\Gamma(\alpha)\prod x_i!}\right\}$$

$$= \frac{e^{-\theta\left(\frac{n\beta+1}{\beta}\right)}\ \theta^{(\sum x_i+\alpha)-1}}{\left(\frac{\beta}{n\beta+1}\right)^{\sum x_i+\alpha}\Gamma(\sum x_i+\alpha)}.$$

Από την παραπάνω σχέση γίνεται φανερό πως η εκ των υστέρων κατανομή του θ είναι *Gamma* $\left(\sum x_i+\alpha,\ \frac{\beta}{n\beta+1}\right)$.

Το γεγονός ότι η εκ των υστέρων κατανομή ακολουθεί την κατανομή *Gamma* ήταν αναμενόμενο, αφού η οικογένεια κατανομών *Gamma*, όπου ανήκει η εκ των προτέρων κατανομή, είναι συζυγής με την κατανομή Poisson.

Η εκ των υστέρων μέση τιμή για την παραπάνω *Gamma* κατανομή, είναι

$$\begin{aligned} E(\theta|x) &= \left(\sum_{i=1}^{n} x_i + \alpha\right)\left(\frac{\beta}{n\beta + 1}\right) \\ &= \frac{\beta \sum x_i}{n\beta + 1} + \frac{\alpha\beta}{n\beta + 1} \\ &= \frac{n\beta}{n\beta + 1}\bar{x} + \frac{1}{n\beta + 1}\alpha\beta \,. \end{aligned}$$

Έστω $\Upsilon \sim Poisson\ (\theta)$ μια νέα ανεξάρτητη παρατήρηση. Η κατανομή πρόβλεψης δίνεται από τον τύπο

$$\pi(y|\boldsymbol{x}) \ = \ \int_{\Theta} f(y|\theta) p(\theta|\boldsymbol{x})\, d\theta \,,$$

όπου για το συγκεκριμένο στατιστικό μοντέλο είναι

$$\pi(y|\boldsymbol{x}) \ = \ \int_{\Theta} \left(e^{-\theta}\frac{\theta^{y}}{y!}\right)\left(\frac{e^{-\theta\left(\frac{n\beta+1}{\beta}\right)}\ \theta^{(\sum x_i + \alpha)-1}}{\left(\frac{\beta}{n\beta+1}\right)^{\sum x_i + \alpha}\Gamma(\sum x_i + \alpha)}\right) d\theta$$

$$= \frac{1}{y!\left(\frac{\beta}{n\beta+1}\right)^{\sum x_i + \alpha}\Gamma(\sum x_i + \alpha)} \int_{\Theta} e^{-\theta\left(\frac{n\beta+\beta+1}{\beta}\right)}\ \theta^{(\sum x_i + \alpha + y)-1}\, d\theta$$

$$= \frac{\left(\frac{\beta}{n\beta+\beta+1}\right)^{\sum x_i + \alpha + y}\Gamma(\sum x_i + \alpha + y)}{\left(\frac{\beta}{n\beta+1}\right)^{\sum x_i + \alpha}\Gamma(\sum x_i + \alpha)\Gamma(y+1)}$$

$$= \frac{\Gamma(\sum x_i + \alpha + y)}{\Gamma(\sum x_i + \alpha)\Gamma(y+1)}\left(\frac{n\beta + 1}{n\beta + \beta + 1}\right)^{\sum x_i + \alpha}\left(\frac{\beta}{n\beta + \beta + 1}\right)^{y},$$

η οποία είναι αρνητική διωνυμική με παραμέτρους $\left(\sum x_i + \alpha,\ \frac{\beta}{n\beta+\beta+1}\right)$.

Παράδειγμα 2.2: $f(x_i|\theta) \sim Normal(\theta, \sigma^2), u(\theta) \sim Normal(\mu, r^2)$

Έστω x_1, x_2, ..., x_n, ένα τυχαίο δείγμα από κατανομή *Normal* $(\theta,\ \sigma^2)$, $\theta \in \Theta = (-\infty, \infty)$ και σ > 0 δεδομένη σταθερά, δηλαδή

$$f(x_i|\theta) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i-\theta)^2}{2\sigma^2}}, \qquad i = 1,2, \dots, n\,.$$

Η πιθανοφάνεια του δείγματος είναι

$$L(\theta|\boldsymbol{x}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i-\theta)^2}{2\sigma^2}}$$

$$= \frac{1}{2\pi^{n/2}\sigma^n} e^{-\frac{1}{2\sigma^2}\sum(x_i-\theta)^2}\,.$$

Έστω πως η εκ των προτέρων κατανομή του θ είναι *Normal* $(\mu,\ r^2)$, όπου μ, r, είναι γνωστές ποσότητες. Επομένως ισχύει ότι

$$u(\theta) = \frac{1}{\sqrt{2\pi}r} e^{-\frac{(\theta-\mu)^2}{2r^2}}\,.$$

Άρα η περιθώρια κατανομή είναι

$$g(\boldsymbol{x}) = \int_{-\infty}^{\infty} L(\theta|\boldsymbol{x})\ u(\theta)\, d\theta$$

$$= \int_{-\infty}^{\infty} \left(\frac{1}{2\pi^{n/2}\sigma^n} exp\left\{-\frac{\sum(x_i-\theta)^2}{2\sigma^2}\right\}\right) \left(\frac{1}{\sqrt{2\pi}r} exp\left\{-\frac{(\theta-\mu)^2}{2r^2}\right\}\right) d\theta$$

$$= \frac{1}{2\pi^{n/2}\sigma^n} \frac{1}{\sqrt{2\pi}r} \int_{-\infty}^{\infty} exp\left\{-\frac{\sum(x_i-\theta)^2}{2\sigma^2}\right\} exp\left\{-\frac{(\theta-\mu)^2}{2r^2}\right\} d\theta$$

$$= \frac{1}{2\pi^{n/2}\sigma^n} \frac{1}{\sqrt{2\pi}r} \int_{-\infty}^{\infty} exp\left\{-\frac{\sum {x_i}^2}{2\sigma^2} + \frac{\theta\sum x_i}{\sigma^2} - \frac{n\theta^2}{2\sigma^2} - \frac{\theta^2}{2r^2} + \frac{\theta\mu}{r^2} - \frac{\mu^2}{2r^2}\right\}\ d\theta\,.$$

Συνεπώς η εκ των υστέρων κατανομή του θ είναι

$$p(\theta|\boldsymbol{x}) = \frac{L(\theta|\boldsymbol{x})u(\theta)}{g(\boldsymbol{x})}$$

$$= \frac{\left(\frac{1}{2\pi^{n/2}\sigma^n} exp\left\{-\frac{\sum(x_i-\theta)^2}{2\sigma^2}\right\}\right)\left(\frac{1}{\sqrt{2\pi}r} exp\left\{-\frac{(\theta-\mu)^2}{2r^2}\right\}\right)}{g(x)}$$

$$= \frac{exp\left(-\frac{\sum x_i{}^2}{2\sigma^2} + \frac{\theta\sum x_i}{\sigma^2} - \frac{n\theta^2}{2\sigma^2} - \frac{\theta^2}{2r^2} + \frac{\theta\mu}{r^2} - \frac{\mu^2}{2r^2}\right)}{\int_{-\infty}^{\infty} exp\left(-\frac{\sum x_i{}^2}{2\sigma^2} + \frac{\theta\sum x_i}{\sigma^2} - \frac{n\theta^2}{2\sigma^2} - \frac{\theta^2}{2r^2} + \frac{\theta\mu}{r^2} - \frac{\mu^2}{2r^2}\right)\, d\theta}$$

$$= \frac{exp\left(\frac{\theta\sum x_i}{\sigma^2} - \frac{n\theta^2}{2\sigma^2} - \frac{\theta^2}{2r^2} + \frac{\theta\mu}{r^2}\right)}{\int_{-\infty}^{\infty} exp\left(\frac{\theta\sum x_i}{\sigma^2} - \frac{n\theta^2}{2\sigma^2} - \frac{\theta^2}{2r^2} + \frac{\theta\mu}{r^2}\right)\, d\theta}$$

$$= \frac{exp\left(\frac{\theta n\bar{x}}{\sigma^2} - \frac{n\theta^2}{2\sigma^2} - \frac{\theta^2}{2r^2} + \frac{\theta\mu}{r^2}\right)}{\int_{-\infty}^{\infty} exp\left(\frac{\theta n\bar{x}}{\sigma^2} - \frac{n\theta^2}{2\sigma^2} - \frac{\theta^2}{2r^2} + \frac{\theta\mu}{r^2}\right)\, d\theta} = \frac{A}{B},$$

όπου ο αριθμητής του παραπάνω κλάσματος είναι

$$A = exp\left(\frac{\theta n\bar{x}}{\sigma^2} - \frac{n\theta^2}{2\sigma^2} - \frac{\theta^2}{2r^2} + \frac{\theta\mu}{r^2}\right)$$

$$= exp\left[\theta\left(\frac{n\bar{x}}{\sigma^2} + \frac{\mu}{r^2}\right) - \theta^2\left(\frac{n}{2\sigma^2} + \frac{1}{2r^2}\right)\right]$$

$$= exp\left[\theta\left(\frac{2n\bar{x}r^2 + 2\mu\sigma^2}{2r^2\sigma^2}\right) - \theta^2\left(\frac{nr^2+\sigma^2}{2\sigma^2r^2}\right)\right]$$

$$= exp\left\{\frac{1}{2}\left(\frac{nr^2+\sigma^2}{\sigma^2r^2}\right)\left[2\theta\left(\frac{n\bar{x}r^2+\mu\sigma^2}{nr^2+\sigma^2}\right) - \theta^2\right]\right\}$$

$$= exp\left\{-\frac{1}{2}\left(\frac{nr^2+\sigma^2}{\sigma^2r^2}\right)\left[\theta^2 - 2\theta\left(\frac{n\bar{x}r^2+\mu\sigma^2}{nr^2+\sigma^2}\right) \pm \left(\frac{n\bar{x}r^2+\mu\sigma^2}{nr^2+\sigma^2}\right)^2\right]\right\}$$

$$= exp\left\{-\frac{1}{2}\left(\frac{nr^2+\sigma^2}{\sigma^2 r^2}\right)\left[\left(\theta-\frac{n\bar{x}r^2+\mu\sigma^2}{nr^2+\sigma^2}\right)^2-\left(\frac{n\bar{x}r^2+\mu\sigma^2}{nr^2+\sigma^2}\right)^2\right]\right\}.$$

Θέτοντας

$$\frac{1}{s^2} = \frac{nr^2+\sigma^2}{\sigma^2 r^2} \;\; \kappa\alpha\iota \;\; m = \frac{n\bar{x}r^2+\mu\sigma^2}{nr^2+\sigma^2},$$

η Α γίνεται

$$A = exp\left(-\frac{(\theta-m)^2}{2s^2}-\frac{m^2}{2s^2}\right).$$

Αντίστοιχα ο παρονομαστής γίνεται

$$B = \int_{-\infty}^{\infty} exp\left(-\frac{(\theta-m)^2}{2s^2}-\frac{m^2}{2s^2}\right)$$

$$= \sqrt{2\pi}s\; exp\left(-\frac{m^2}{2s^2}\right)\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}s}exp\left(-\frac{(\theta-m)^2}{2s^2}\right)$$

$$= \sqrt{2\pi}s\; exp\left(-\frac{m^2}{2s^2}\right).$$

Τότε

$$\frac{A}{B} = \frac{exp\left(-\frac{(\theta-m)^2}{2s^2}-\frac{m^2}{2s^2}\right)}{\sqrt{2\pi}s\; exp\left(-\frac{m^2}{2s^2}\right)}$$

$$= \frac{1}{\sqrt{2\pi}s}exp\left(-\frac{(\theta-m)^2}{2s^2}\right).$$

Από την παραπάνω σχέση, γίνεται κατανοητό πως η εκ των υστέρων κατανομή ακολουθεί την κατανομή *Normal (m,* s^2*)*, όπου $m = \frac{n\bar{x}r^2+\mu\sigma^2}{nr^2+\sigma^2}$ $\kappa\alpha\iota$ $s^2 = \frac{\sigma^2 r^2}{nr^2+\sigma^2}$.

Πάλι η εκ των υστέρων κατανομή ακολουθεί την κανονική κατανομή, αφού η οικογένεια κανονικών κατανομών, είναι συζυγής με την κανονική κατανομή.

Η εκ των υστέρων μέση τιμή για την παραπάνω κανονική κατανομή είναι

$$E(\theta|x) = \frac{n\bar{x}r^2 + \mu\sigma^2}{nr^2 + \sigma^2}$$

$$= \frac{nr^2}{nr^2 + \sigma^2}\bar{x} + \frac{\sigma^2}{nr^2 + \sigma^2}\mu\,.$$

Έστω $Y \sim N(\theta, \sigma^2)$ μια νέα ανεξάρτητη παρατήρηση. Τότε η κατανομή πρόβλεψης έχει πυκνότητα

$$\pi(y|\boldsymbol{x}) =$$

$$= \int_{\Theta} f(y|\theta)p(\theta|\boldsymbol{x})\,d\theta$$

$$= \int_{\Theta} \left(\frac{1}{\sqrt{2\pi}\sigma} exp\left\{-\frac{(y-\theta)^2}{2\sigma^2}\right\}\right)\left(\frac{1}{\sqrt{2\pi}s} exp\left\{-\frac{(\theta-m)^2}{2s^2}\right\}\right) d\theta$$

$$= \frac{1}{2\pi\sigma s}\int_{\Theta} \left(exp\left\{-\frac{y^2 - 2y\theta + \theta^2}{2\sigma^2}\right\}\right)\left(exp\left\{-\frac{m^2 - 2m\theta + \theta^2}{2s^2}\right\}\right) d\theta$$

$$= \frac{1}{2\pi\sigma s}\exp -\frac{1}{2}\left\{\left(\frac{y^2}{\sigma^2} + \frac{m^2}{s^2}\right)\right\}\int_{\Theta} \left(\exp\left\{-\frac{\theta^2 - 2y\theta}{2\sigma^2}\right\}\right)\left(\exp\left\{-\frac{\theta^2 - 2m\theta}{2s^2}\right\}\right) d\theta$$

$$= \frac{1}{2\pi\sigma s}\exp -\frac{1}{2}\left\{\left(\frac{y^2}{\sigma^2} + \frac{m^2}{s^2}\right)\right\} A\,,$$

όπου

$$A = \int_{\Theta} \left(\exp\left\{-\frac{\theta^2 - 2y\theta}{2\sigma^2}\right\}\right)\left(\exp\left\{-\frac{\theta^2 - 2m\theta}{2s^2}\right\}\right) d\theta$$

Επομένως αναλύοντας το παραπάνω ολοκλήρωμα, λαμβάνουμε την κάτωθι σχέση

$$A =$$

$$= \int_{\Theta} exp -\frac{1}{2}\left\{\left(\frac{\sigma^2 + s^2}{\sigma^2 s^2}\right)\theta^2 - 2\left(\frac{ys^2 + m\sigma^2}{\sigma^2 s^2}\right)\theta\right\} d\theta$$

$$= \int_{\Theta} exp -\frac{1}{2}\left\{\left(\frac{\sigma^2 + s^2}{\sigma^2 s^2}\right)\left[\theta^2 - 2\left(\frac{ys^2 + m\sigma^2}{\sigma^2 + s^2}\right)\theta \pm \left(\frac{ys^2 + m\sigma^2}{\sigma^2 + s^2}\right)^2\right]\right\} d\theta$$

$$= exp\frac{1}{2}\left\{\left(\frac{\sigma^2 + s^2}{\sigma^2 s^2}\right)\left(\frac{ys^2 + m\sigma^2}{\sigma^2 + s^2}\right)^2\right\}\int_{\Theta} exp -\left\{\left(\frac{\sigma^2 + s^2}{2\sigma^2 s^2}\right)\left(\theta - \frac{ys^2 + m\sigma^2}{\sigma^2 + s^2}\right)^2\right\} d\theta$$

$$= \frac{1}{\sqrt{2\pi\left(\frac{\sigma^2 + s^2}{\sigma^2 s^2}\right)}} exp\left\{\frac{1}{2}\left(\frac{\sigma^2 + s^2}{\sigma^2 s^2}\right)\left(\frac{ys^2 + m\sigma^2}{\sigma^2 + s^2}\right)^2\right\}.$$

Άρα η $\pi(y|\boldsymbol{x})$ γίνεται

$$\pi(y|\boldsymbol{x}) =$$

$$= \frac{1}{2\pi\sigma s}\frac{1}{\sqrt{2\pi\left(\frac{\sigma^2 + s^2}{\sigma^2 s^2}\right)}} exp -\frac{1}{2}\left\{\left(\frac{y^2}{\sigma^2} + \frac{m^2}{s^2}\right) - \left(\frac{\sigma^2 + s^2}{\sigma^2 s^2}\right)\left(\frac{ys^2 + m\sigma^2}{\sigma^2 + s^2}\right)^2\right\}$$

$$= \frac{1}{\sqrt{2\pi(\sigma^2 + s^2)}} exp -\frac{1}{2}\left\{\frac{y^2 s^2(\sigma^2 + s^2) + m^2\sigma^2(\sigma^2 + s^2) - (ys^2 + m\sigma^2)^2}{\sigma^2 s^2(\sigma^2 + s^2)}\right\}$$

$$= \frac{1}{\sqrt{2\pi(\sigma^2 + s^2)}} exp -\frac{1}{2}\left\{\frac{y^2 s^2\sigma^2 - 2ys^2 m\sigma^2 + m^2 s^2\sigma^2}{\sigma^2 s^2(\sigma^2 + s^2)}\right\}$$

$$= \frac{1}{\sqrt{2\pi(\sigma^2 + s^2)}} exp -\frac{1}{2}\left\{\frac{(y - m)^2}{(\sigma^2 + s^2)}\right\}.$$

Από την παραπάνω σχέση, φαίνεται πως η κατανομή πρόβλεψης είναι $N(m, \sigma^2 + s^2)$.

Παράδειγμα 2.3: $f(x_i|\theta) \sim Bernoulli(\theta), u(\theta) \sim Beta(\alpha, \beta)$

Έστω $x_1, x_2, \ldots, x_n$, ένα τυχαίο δείγμα από κατανομή *Bernoulli (θ),* $\theta \in \Theta = (0,1)$, δηλαδή

$$f(x_i|\theta) = \theta^{x_i} (1-\theta)^{1-x_i}, \qquad i = 1,2,\ldots,n\,.$$

Η πιθανοφάνεια του δείγματος είναι

$$L(\theta|x) = \prod_{i=1}^{n} \theta^{x_i} (1-\theta)^{1-x_i}$$

$$= \theta^{\sum_{i=1}^{n} x_i} (1-\theta)^{n-\sum_{i=1}^{n} x_i}\,.$$

Έστω πως η εκ των προτέρων κατανομή του θ είναι *Beta* με γνωστές παραμέτρους *(α,β)*. Επομένως ισχύει ότι

$$u(\theta) = \frac{1}{B(\alpha,\beta)} \theta^{\alpha-1} (1-\theta)^{\beta-1}\,.$$

Τότε η περιθώρια κατανομή είναι

$$\begin{aligned} g(\boldsymbol{x}) &= \int_0^1 L(\theta|\boldsymbol{x})\; u(\theta)\, d\theta \\ &= \int_0^1 \left(\theta^{\sum_{i=1}^{n} x_i} (1-\theta)^{n-\sum_{i=1}^{n} x_i}\right)\left(\frac{1}{B(\alpha,\beta)} \theta^{\alpha-1} (1-\theta)^{\beta-1}\right) d\theta \\ &= \frac{1}{B(\alpha,\beta)} \int_0^1 \theta^{\left(\alpha+\sum_{i=1}^{n} x_i\right)-1} (1-\theta)^{\left(\beta+n-\sum_{i=1}^{n} x_i\right)-1}\, d\theta \\ &= \frac{B\left(\alpha+\sum_{i=1}^{n} x_i\,, \beta+n-\sum_{i=1}^{n} x_i\right)}{B(\alpha,\beta)}\,. \end{aligned}$$

Άρα η εκ των υστέρων κατανομή του θ είναι

$$p(\theta|\boldsymbol{x}) = \frac{L(\theta|\boldsymbol{x})u(\theta)}{g(\boldsymbol{x})}$$

$$= \frac{\left(\theta^{\sum_{i=1}^{n} x_i} (1-\theta)^{n-\sum_{i=1}^{n} x_i}\right)\left(\frac{1}{B(\alpha,\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}\right)}{\left(\frac{B(\alpha+\sum_{i=1}^{n} x_i,\beta+n-\sum_{i=1}^{n} x_i)}{B(\alpha,\beta)}\right)}$$

$$= \frac{\theta^{(\alpha+\sum_{i=1}^{n} x_i)-1}(1-\theta)^{(\beta+n-\sum_{i=1}^{n} x_i)-1}}{B(\alpha+\sum_{i=1}^{n} x_i,\beta+n-\sum_{i=1}^{n} x_i)}\ .$$

Επομένως η εκ των υστέρων κατανομή είναι κατανομή *Beta* με παραμέτρους $(\alpha+\sum_{i=1}^{n} x_i\,,\ \ \beta+n-\sum_{i=1}^{n} x_i)$. Η εκ των υστέρων μέση τιμή για την παραπάνω Beta κατανομή είναι

$$E(\theta|x) = \frac{\alpha+\sum_{i=1}^{n} x_i}{\alpha+\ \beta+n}$$

$$= \frac{\alpha}{(\alpha+\ \beta+n)}\frac{(\alpha+\ \beta)}{(\alpha+\ \beta)} + \frac{n}{\alpha+\ \beta+n}\frac{\sum_{i=1}^{n} x_i}{n}$$

$$= \left(\frac{\alpha+\ \beta}{\alpha+\ \beta+n}\right)\frac{a}{(\alpha+\ \beta)} + \left(\frac{n}{\alpha+\ \beta+n}\right)\bar{x}\ .$$

Εάν $Y \sim Bernoulli\,(\theta)$ μια νέα ανεξάρτητη παρατήρηση, τότε η κατανομή πρόβλεψης έχει πυκνότητα

$$\pi(y|\boldsymbol{x}) = \int_{\Theta} f(y|\theta)p(\theta|\boldsymbol{x})\,d\theta$$

$$= \int_{\Theta} (\theta^{y}(1-\theta)^{1-y})\left(\frac{\theta^{(\alpha+\sum_{i=1}^{n} x_i)-1}(1-\theta)^{(\beta+n-\sum_{i=1}^{n} x_i)-1}}{B(\alpha+\sum_{i=1}^{n} x_i,\beta+n-\sum_{i=1}^{n} x_i)}\right)d\theta$$

$$= \binom{1}{y}\frac{B(y+\alpha+\sum_{i=1}^{n} x_i\,,\ 1-y+\beta+n-\sum_{i=1}^{n} x_i)}{B(\alpha+\sum_{i=1}^{n} x_i\,,\ \beta+n-\sum_{i=1}^{n} x_i)}\,.$$

Συνεπώς η εκ των υστέρων κατανομή πρόβλεψης ακολουθεί την κατανομή Beta-Binominal με παραμέτρους $(1,\ \alpha+\sum_{i=1}^{n} x_i\,,\ \beta+n-\sum_{i=1}^{n} x_i)$.

Παράδειγμα 2.4: $f(x_i|\theta) \sim Poisson(\theta), u(\theta) \sim Gamma(\alpha, \beta)\ , h(\beta) \sim Gamma(\gamma, \delta)$

Έστω τυχαίο δείγμα $x_1, x_2, \ldots, x_n$, από κατανομή $Poisson(\theta), \theta > 0$ και η εκ των προτέρων κατανομή του θ, ακολουθεί την κατανομή $Gamma(\alpha, \beta)$. Έστω πως το α είναι μια γνωστή ποσότητα, $\alpha > 0$ και β είναι η μια άγνωστη υπερπαράμετρος η οποία έχει κατανομή $h(\beta) \sim Gamma(\gamma, \delta)$, με γ, $\delta > 0$ και γνωστά. Η από κοινού κατανομή των θ, β είναι

$$u(\theta, \beta) = u(\theta|\ \beta)h(\beta)$$

$$= \left(\frac{\beta^{\alpha}\theta^{\alpha-1}e^{-\beta\theta}}{\Gamma(\alpha)}\right)\left(\frac{\delta^{\gamma}\beta^{\gamma-1}e^{-\delta\beta}}{\Gamma(\gamma)}\right)$$

$$= \frac{\delta^{\gamma}\theta^{\alpha-1}\beta^{\alpha+\gamma-1}e^{-(\delta+\theta)\beta}}{\Gamma(\alpha)\Gamma(\gamma)}.$$

Η εκ των προτέρων κατανομή του θ, προκύπτει από την ολοκλήρωση της παραπάνω, ως προς β, δηλαδή

$$u(\theta) = \int u(\theta, \beta) d\beta$$

$$= \int \frac{\delta^{\gamma}\theta^{\alpha-1}\beta^{\alpha+\gamma-1}e^{-(\delta+\theta)\beta}}{\Gamma(\alpha)\Gamma(\gamma)} d\beta$$

$$= \frac{\delta^{\gamma}\theta^{\alpha-1}}{\Gamma(\alpha)\Gamma(\gamma)} \int \beta^{\alpha+\gamma-1}e^{-(\delta+\theta)\beta} d\beta$$

$$= \frac{\delta^{\gamma}\theta^{\alpha-1}}{\Gamma(\alpha)\Gamma(\gamma)} \frac{\Gamma(\alpha+\gamma)}{(\delta+\theta)^{\alpha+\gamma}} \int \frac{\beta^{\alpha+\gamma-1}e^{-(\delta+\theta)\beta}(\delta+\theta)^{\alpha+\gamma}}{\Gamma(\alpha+\gamma)} d\beta$$

$$= \frac{\Gamma(\alpha+\gamma)}{\Gamma(\alpha)\Gamma(\gamma)} \delta^{\gamma} \frac{\theta^{\alpha-1}}{(\delta+\theta)^{\alpha+\gamma}}.$$

Επομένως η εκ των υστέρων κατανομή είναι η κάτωθι

$$p(\theta|\boldsymbol{x}) \propto \left(e^{-n\theta}\theta^{\sum x_i}\right)\left(\frac{\theta^{\alpha-1}}{(\delta+\theta)^{\alpha+\gamma}}\right)$$

$$= e^{-n\theta}\,\frac{\theta^{\alpha-1}\theta^{\sum x_i}}{(\delta+\theta)^{\alpha+\gamma}}$$

$$= \frac{\theta^{\alpha+\sum x_i-1}e^{-n\theta}}{(\delta+\theta)^{\alpha+\gamma}}.$$

Η παραπάνω εκ των υστέρων κατανομή έχει μη αναγνωρίσιμη μορφή. Σε ένα απλοϊκό παράδειγμα σαν αυτό, δεν καταφέραμε να καταλήξουμε σε μια εύχρηστη συνάρτηση πιθανότητας.

Σημείωση:

Αν και μέσω ιεραρχικής Μπεϋζιανής μεθόδου, γίνεται δυνατή η αποφυγή της επιλογής τιμών για τις υπερπαραμέτρους του μοντέλου, και όπως θα δούμε στο επόμενο κεφάλαιο, η απόδοση των σημειακών εκτιμήσεων για την παράμετρο θ, είναι σε πολλές περιπτώσεις ανώτερη από εκείνη της Μπεϋζιανής μεθόδου, η πολυπλοκότητα του μοντέλου την κάνει σε πολλές περιπτώσεις μη προτιμητέα.

Εναλλακτική της ιεραρχικής Μπεϋζιανής μπορεί να θεωρηθεί η εμπειρική Μπεϋζιανή προσέγγιση. Στο επόμενο κεφάλαιο λοιπόν, θα μελετηθεί σε βάθος η προσέγγιση αυτή, με στόχο να διαπιστωθεί χρησιμότητα της μεθόδου.

Κεφάλαιο 3

3. ΜΗ ΠΑΡΑΜΕΤΡΙΚΗ ΕΜΠΕΙΡΙΚΗ ΜΠΕΫΖΙΑΝΗ ΠΡΟΣΕΓΓΙΣΗ

3.1 Εμπειρική Μπεϋζιανή Ανάλυση

Στο κεφάλαιο αυτό θα μελετηθεί σε βάθος η εμπειρική Μπεϋζιανή προσέγγιση. Η προσέγγιση αυτή μπορεί να θεωρηθεί ένα συνοθύλευμα της Μπεϋζιανής στατιστικής, με την κλασσική, η οποία προσπαθεί να επωφεληθεί από τα δυνατά σημεία και των δυο προσεγγίσεων (βλέπε Casella, 1985).

Η βασική διαφορά της Μπεϋζιανής και της εμπειρικής Μπεϋζιανής ανάλυσης εντοπίζεται στον καθορισμό της εκ των προτέρων πληροφορίας. Στην πρώτη περίπτωση, θεωρείται πως η εκ των προτέρων πληροφορία, προηγείται της συλλογής δεδομένων, ενώ στην εμπειρική Μπεϋζιανή κάποια χαρακτηριστικά της εκ των προτέρων πληροφορίας εκτιμώνται κατά κάποιο τρόπο από τα ίδια τα δεδομένα, μέσω μεθόδων που χρησιμοποιούνται στην κλασσική στατιστική. Στο γεγονός αυτό οφείλεται και ο όρος “εμπειρική”.

Ο Robbins (1955) ήταν εκείνος που εισήγαγε την έννοια της εμπειρικής Μπεϋζιανής ανάλυσης. Στόχος της προσέγγισης αυτής ήταν να παρουσιάσει την εκ των ύστερων μέση τιμή σε όρους της περιθώριας κατανομής, και έπειτα χρησιμοποιώντας τα δεδομένα, προχωρούσε στον εμπειρικό υπολογισμό του Μπεϋζιανού κανόνα (βλέπε Berger, 1985).

Η προσέγγιση του Robbins αργότερα μετονομάστηκε σε μη παραμετρική εμπειρική Μπεϋζιανή, όταν ο Morris (1983) την διαχώρισε από την παραμετρική της εκδοχή. Στην παραμετρική εμπειρική Μπεϋζιανή, θεωρείται πως η εκ των προτέρων πληροφορία της παραμέτρου, ανήκει σε μια γνωστή παραμετρική κλάση, όπου μόνο οι υπερπαράμετροι της είναι άγνωστοι. Έπειτα οι υπερπαράμετροι εκτιμώνται από τα δεδομένα, με τη χρήση κλασσικών μεθόδων εκτίμησης.

Παρακάτω παρουσιάζονται αναλυτικά οι δυο αυτές προσεγγίσεις, όπως επίσης και παραδείγματα που έχουν ως σκοπό την πιο λεπτομερή επεξήγησή τους.

3.2 Μη Παραμετρική Εμπειρική Μπεϋζιανή Ανάλυση

Ας υποτεθεί το παρακάτω πλαίσιο. Έστω πως για n ανεξάρτητες παρατηρήσεις $x_1,\ldots,x_n$, όπου προέρχονται από μια γνωστή κατανομή με πυκνότητα $f(x_i|\theta_i)$, στόχος είναι η συμπερασματολογία πάνω στην παράμετρο θ_n, όπου οι παράμετροι θ_i έχουν παραχθεί από την ίδια εκ των προτέρων κατανομή $u(\theta_i)$.

Το παραπάνω μοντέλο ονομάζεται σύνθετο μοντέλο δειγματοληψίας ή αλλιώς αμοιβαία ανεξάρτητο ιεραρχικό μοντέλο *(Conditionally Independent Hierarchical Model)*, (βλέπε Kass & Steffey, 1989) και χρησιμοποιείται ευρέως σε μελέτες κλινικών δοκιμών και συχνότητας ατυχημάτων (βλέπε Strickland, 2015).

Στην περίπτωση της μη παραμετρικής εμπειρικής Μπεϋζιανής προσέγγισης είναι αναγκαίο να πραγματοποιηθούν επαναλαμβανόμενες λήψεις από την εκ των προτέρων κατανομή, από τη στιγμή που εκείνη δεν θεωρείται ολοκληρωτικά γνωστή και κάποια χαρακτηριστικά της πρέπει να εκτιμηθούν (βλέπε Carlin & Louis, 2008).

Υπό την Μπεϋζιανή λογική, η εκ των προτέρων κατανομή των θ_i, θεωρείται γνωστή από τον αναλυτή, οπότε είναι δυνατή η εκτίμηση της παραμέτρου. Στην πράξη όμως, ο προσδιορισμός της εκ των προτέρων κατανομής σε κάποιες περιπτώσεις είναι αδύνατος (βλέπε Robbins, 1985).

Μια εναλλακτική προσέγγιση προτάθηκε από τον Robbins (1955) και προτείνει να αναζητηθεί μια μαθηματική αποτύπωση του Μπεϋζιανού κανόνα, η οποία θα περιέχει μόνο όρους της περιθώριας κατανομής $g(x_i)$, με αποτέλεσμα να αποφεύγεται ο υπολογισμός της εκ των προτέρων κατανομής $u(\theta_i)$. Έπειτα η περιθώρια κατανομή υπολογίζεται κατευθείαν από τα δεδομένα (βλέπε Berger, 1985).

3.2.1 Γενική Μη Παραμετρική Εμπειρική Μπεϋζιανή Προσέγγιση

Πιο αναλυτικά, στη γενική περίπτωση της μη παραμετρικής εμπειρικής Μπεϋζιανής προσέγγισης, η μόνη παραδοχή που πρέπει να γίνει είναι εκείνη της ανεξαρτησίας των θ_i, τα οποία προέρχονται από μια άγνωστη, αλλά κοινή εκ των προτέρων κατανομή.

Στόχος είναι η εκτίμηση της παραμέτρου θ_i από μια συνάρτηση $t = t(x_i)$. Υπό την συνάρτηση απώλειας τετραγωνικού σφάλματος,

$$L(t(x_i) - \theta_i) = (t(x_i) - \theta_i)^2 \,,$$

ο κίνδυνος Bayes του κανόνα απόφασης t ισούται με

$$E(t-\theta_i)^2 = \int_X \int_\Theta (t(x_i)-\theta_i)^2 f(x_i|\theta_i)u(\theta_i)\,dx\,d\theta_i\,. \qquad (3.1)$$

Ο κανόνας Bayes που ελαχιστοποιεί την παραπάνω συνάρτηση είναι η εκ των υστέρων μέση τιμή (βλέπε 2.6), η οποία υπολογίζεται όπως κάτωθι

$$\begin{aligned} t(x_i) &= E(\theta_i|x_i) \\ &= \int \theta_i p(\theta_i|x_i)d\theta_i \\ &= \int \theta_i \frac{f(x_i|\theta_i)u(\theta_i)}{g(x_i)} d\theta_i \\ &= \frac{\int \theta_i f(x_i|\theta_i)u(\theta_i)\,d\theta_i}{\int f(x_i|\theta_i)u(\theta_i)\,d\theta_i}. \end{aligned}$$

Όταν η κατανομή $f(x_i|\theta_i)$, ανήκει στην εκθετική οικογένεια κατανομών, τότε είναι δυνατό η $E(\theta_i|x_i)$ να εκφραστεί μόνο με όρους της περιθώριας κατανομής $g(x_i)$ (βλέπε Maritz, 1970).

Θα πρέπει να σημειωθεί πως θα μπορούσαν να χρησιμοποιηθούν και άλλες συναρτήσεις απώλειας οι οποίες οδηγούν σε εκτιμητές με βολική δομή, αλλά στην παρούσα μελέτη θα εξεταστεί μόνο αυτή της απώλειας τετραγωνικού σφάλματος (βλέπε Louis, 2001).

Ο Robbins πρότεινε μια ολοκληρωτικά μη παραμετρική εκτίμηση, στην οποία οι περιθώριες πιθανότητες εκτιμώνται από την εμπειρική της συχνότητα, δηλαδή

$$g_n(x_i) = \frac{\#\ των\ x\ τα\ οποια\ ισουνται\ με\ x_i}{n},$$

όπου η παραπάνω εμπειρική συχνότητα τείνει στην $g(x_i)$ με πιθανότητα 1, όταν το $n \to \infty$. Είναι φανερό, πως από τη στιγμή που η εκτίμηση της περιθώριας πυκνότητας εξαρτάται από όλα τα x_i, έτσι και η εκτίμηση του κάθε θ_i, είναι επηρεασμένη από όλα τα δεδομένα (βλέπε, Carlin & Louis, 2008).

Επομένως γίνεται δυνατός ο υπολογισμός της σημειακής εκτίμησης του θ_i, χωρίς κάποια παραμετρική υπόθεση για αυτό.

Παρακάτω θα αναφερθούν κάποιες χαρακτηριστικές περιπτώσεις, όπου είναι δυνατόν να υλοποιηθεί η μη παραμετρική εμπειρική ανάλυση (βλέπε Robbins, 1955).

Παράδειγμα 3.1: $f(x_i|\theta_i) \sim Poisson(\theta_i)$

Υποθέτοντας πως η $f(x_i|\theta_i)$ ακολουθεί την κατανομή $Poisson(\theta_i)$, ενώ η εκ των προτέρων $u(\theta_i)$ δεν είναι προσδιορισμένη, η εκ των υστέρων μέση τιμή γίνεται

$$
\begin{aligned}
E(\theta_i|x_i) &= \frac{\int \theta_i f(x_i|\theta_i)u(\theta_i)\, d\theta_i}{g(x_i)} \\
&= \frac{\displaystyle\int_{-\infty}^{\infty} \theta_i \frac{e^{-\theta_i}{\theta_i}^{x_i}}{x_i!} u(\theta_i) d\theta_i}{\displaystyle\int_{-\infty}^{\infty} \frac{e^{-\theta_i}{\theta_i}^{x_i}}{x_i!} u(\theta_i) d\theta_i} \\
&= \frac{\displaystyle\int_{-\infty}^{\infty} \theta_i \frac{e^{-\theta_i}{\theta_i}^{x_i+1}}{x_i!} u(\theta_i) d\theta_i}{\displaystyle\int_{-\infty}^{\infty} \frac{e^{-\theta_i}\theta_i^{x_i}}{x_i!} u(\theta_i) d\theta_i} \\
&= \frac{\displaystyle\int_{-\infty}^{\infty} \frac{e^{-\theta_i}{\theta_i}^{x_i+1}}{(x_i+1)!} u(\theta_i) d\theta_i}{\displaystyle\int_{-\infty}^{\infty} \frac{e^{-\theta_i}\theta_i^{x_i}}{x_i!} u(\theta_i) d\theta_i} (x_i+1) \\
&= (x_i+1)\frac{g_n(x_i+1)}{g_n(x_i)}.
\end{aligned}
$$

Όπως φαίνεται από την παραπάνω σχέση, η $E(\theta_i|x_i)$ έχει εκφραστεί μόνο με όρους της περιθώριας κατανομής. Έτσι δεν είναι πλέον αναγκαίος ο ορισμός της εκ των προτέρων κατανομής.

Παράδειγμα 3.2: $f(x_i|\theta_i) \sim Geometric(\theta_i)$

Θεωρώντας ότι η κατανομή που ακολουθούν τα δεδομένα είναι η $Geometric(\theta_i)$, καταλήγουμε πως η εκ των υστέρων μέση τιμή είναι ίση με

$$E(\theta_i|x_i) = \frac{\int \theta_i f(x_i|\theta_i)u(\theta_i)\, d\theta_i}{g(x_i)}$$

$$= \frac{\int_0^1 \theta_i(1-\theta_i){\theta_i}^{x_i}u(\theta_i)d\theta_i}{\int_0^1 (1-\theta_i){\theta_i}^{x_i}u(\theta_i)d\theta_i}$$

$$= \frac{\int_0^1 (1-\theta_i){\theta_i}^{x_i+1}u(\theta_i)d\theta_i}{\int_0^1 (1-\theta_i){\theta_i}^{x_i}u(\theta_i)d\theta_i}$$

$$= \frac{g_n(x_i+1)}{g_n(x_i)}.$$

Για να έχει νόημα η παραπάνω έκφραση, η εκ των προτέρων κατανομή $u(\theta_i)$, θα πρέπει να είναι ορισμένη στο διάστημα $(0,1)$.

Παράδειγμα 3.3: $f(x_i|\theta_i) \sim Binomial(k, \theta_i)$

Έστω πως η $f(x_i|\theta_i)$ ακολουθεί τη διωνυμική κατανομή, με παραμέτρους k και θ_i. Εδώ το k θεωρείται πως έχει μια ορισμένη τιμή, η οποία υποδηλώνει τον αριθμό των επαναλήψεων του τυχαίου πειράματος, ενώ με θ_i συμβολίζεται η άγνωστη πιθανότητα επιτυχίας του πειράματος και με X οι ο αριθμός των επιτυχιών. Ακολουθώντας την ίδια διαδικασία, η μέση τιμή υπολογίζεται ως εξής:

$$E(\theta_i|x_i) = \frac{\int \theta_i f(x_i|\theta_i)u(\theta_i)\, d\theta_i}{g(x_i)}$$

$$= \frac{\int_0^1 \theta_i \binom{k}{x_i} {\theta_i}^{x_i}(1-\theta_i)^{k-x_i}u(\theta_i)d\theta_i}{\int_0^1 \binom{k}{x_i} {\theta_i}^{x_i}(1-\theta_i)^{k-x_i}u(\theta_i)d\theta_i}$$

$$= \frac{\int_0^1 \binom{k}{x_i} {\theta_i}^{x_i+1} (1-\theta_i)^{k-x_i} u(\theta_i) d\theta_i}{\int_0^1 \binom{k}{x_i} {\theta_i}^{x_i} (1-\theta_i)^{k-x_i} u(\theta_i) d\theta_i}$$

$$= \frac{\int_0^1 \frac{k!}{x_i!\,(k-x_i)!} {\theta_i}^{x_i+1} (1-\theta_i)^{k-x_i+(1-1)}\, u(\theta_i) d\theta_i}{\int_0^1 \frac{k!}{x_i!\,(n-x_i)!} {\theta_i}^{x_i} (1-\theta_i)^{k-x_i}\, u(\theta_i) d\theta_i}$$

$$= \frac{x_i+1}{k+1} \frac{\int_0^1 \frac{k!\,(n+1)}{x_i!\,(x_i+1)(k+1-x_i-1)!} {\theta_i}^{x_i+1} (1-\theta_i)^{k+1-x_i-1}\, u(\theta_i) d\theta_i}{\int_0^1 \frac{k!}{x_i!\,(k-x_i)!} {\theta_i}^{x_i} (1-\theta_i)^{k-x_i}\, u(\theta_i) d\theta_i}$$

$$= \frac{x_i+1}{k+1} \frac{\int_0^1 \frac{(k+1)!}{(x_i+1)!\,([k+1]-[x_i+1])!} {\theta_i}^{x_i+1} (1-\theta_i)^{(k+1)-(x_i+1)} u(\theta_i) d\theta_i}{\int_0^1 \frac{k!}{x_i!\,(k-x_i)!} {\theta_i}^{x_i} (1-\theta_i)^{k-x_i}\, u(\theta_i) d\theta_i}$$

$$= \frac{x_i+1}{k+1} \frac{\int_0^1 \binom{k+1}{x_i+1} {\theta_i}^{x_i+1} (1-\theta_i)^{(k+1)-(x_i+1)}\, u(\theta_i) d\theta_i}{\int_0^1 \binom{k}{x_i} {\theta_i}^{x_i} (1-\theta_i)^{k-x_i}\, u(\theta_i) d\theta_i}$$

$$= \frac{x_i+1}{k+1} \frac{g_{n,k+1}(x_i+1)}{g_{n,k}(x_i)}.$$

Υποθέτοντας τώρα την ακολουθία των τυχαίων μεταβλητών $X_1, X_2, \ldots, X_n, \ldots$, όπου το X_r υποδηλώνει τον αριθμό των επιτυχιών στις k-1 πρώτες δοκιμές, ισχύει ότι

$$E(\theta_i|x_i) = \frac{x_i+1}{n} \frac{g_{n,k}(x_i+1)}{g_{n,k-1}(x_i)}.$$

Στην παραπάνω περίπτωση η μελέτη περιορίζεται στις k-1 επαναλήψεις, από κάθε σετ k επαναλήψεων. Για μεγάλο k αυτό δεν είναι σημαντικό ως προς την πληροφορία που θυσιάζεται.

Σημείωση

Μέχρι τώρα εξετάστηκαν παραδείγματα διακριτών κατανομών, που ανήκουν στην εκθετική οικογένεια κατανομών. Η παραπάνω μεθοδολογία μπορεί να χρησιμοποιηθεί και όταν η $f(x_i|\theta_i)$ ακολουθεί μια συνεχή κατανομή της εκθετικής οικογένειας.

Σε αυτές τις περιπτώσεις, ο Μπεϋζιανός κανόνας εκφράζεται σε όρους της περιθώριας κατανομής και όρους της παραγώγου της περιθώριας κατανομής. Παρακάτω μελετώνται τέτοια παραδείγματα (βλέπε Miyasawa, 1961).

Παράδειγμα 3.4: $f(x_i|\theta_i) \sim Normal(\theta_i,\, \sigma^2)$

Έστω πως η $f(x_i|\theta_i)$ ακολουθεί την κανονική κατανομή με παραμέτρους $(\theta_i,\, \sigma^2)$. Μας ενδιαφέρει να υπολογίσουμε τη μερική παράγωγο του λογαρίθμου της $f(x_i|\theta_i)$, ως προς x_i (βλέπε Maritz, 1970), δηλαδή

$$log\big(f(x_i|\theta_i)\big) = log\left(\frac{1}{\sqrt{2\pi\sigma^2}} exp\left\{-\frac{(x_i-\theta_i)^2}{2\sigma^2}\right\}\right)$$

$$= -\, log\sqrt{2\pi\sigma^2} - \frac{(x_i-\theta_i)^2}{2\sigma^2},$$

και η παράγωγος της παραπάνω ως προς x, είναι

$$\frac{\partial log\big(f(x_i|\theta_i)\big)}{\partial x_i} = -\frac{x_i-\theta_i}{\sigma^2}$$

$$\Rightarrow \frac{1}{f(x_i|\theta)} \frac{\partial f(x_i|\theta_i)}{\partial x_i} = -\frac{x_i-\theta_i}{\sigma^2}.$$

Λύνοντας την παραπάνω συνάρτηση ως προς θ_i, λαμβάνουμε

$$\theta_i = x_i + \frac{\sigma^2}{f(x_i|\theta_i)} \frac{\partial f(x_i|\theta_i)}{\partial x_i}.$$

Έτσι η μέση τιμή $E(\theta_i|x_i)$ γίνεται

$$
\begin{aligned}
E(\theta_i|x_i) &= \frac{\int_{-\infty}^{\infty} \theta_i f(x_i|\theta_i)\, u(\theta_i)\, d\theta_i}{g(x_i)} \\
&= \frac{\displaystyle\int_{-\infty}^{\infty} \left(x_i + \frac{\sigma^2}{f(x_i|\theta_i)} \frac{\partial f(x_i|\theta_i)}{\partial x_i}\right) f(x_i|\theta_i)\, u(\theta_i)\, d\theta_i}{g(x_i)} \\
&= \frac{x_i \int_{-\infty}^{\infty} f(x_i|\theta_i)\pi(\theta_i) d\theta_i + \sigma^2 \displaystyle\int_{-\infty}^{\infty} \left(\frac{\partial f(x_i|\theta_i)}{\partial x_i}\right) u(\theta_i)\, d\theta_i}{g(x_i)} \\
&= x_i \frac{\int_{-\infty}^{\infty} f(x_i|\theta_i)\pi(\theta_i) d\theta_i}{g(x_i)} + \sigma^2 \frac{\displaystyle\int_{-\infty}^{\infty} \left(\frac{\partial f(x_i|\theta_i)}{\partial x_i}\right) u(\theta_i)\, d\theta_i}{g(x_i)} \\
&= x_i + \sigma^2 \, \frac{g(x_i)'}{g(x_i)}.
\end{aligned}
$$

Παράδειγμα 3.5: $f(x_i|\theta_i) \sim Exponential(\theta_i)$

Υποθέτοντας πως η $f(x_i|\theta_i) = \theta_i e^{-\theta_i x_i}$, $\theta_i > 0$ και πως η εκ των προτέρων κατανομή δεν έχει προσδιοριστεί, η εκ των υστέρων μέση τιμή δίνεται από την παρακάτω σχέση

$$
\begin{aligned}
E(\theta_i|x_i) &= \frac{\int_0^{\infty} \theta_i f(x_i|\theta_i)\, u(\theta_i)\, d\theta_i}{g(x_i)} \\
&= \frac{\int_0^{\infty} \theta_i \theta_i e^{-\theta_i x_i}\, u(\theta_i)\, d\theta_i}{g(x_i)} \\
&= \frac{\displaystyle\int_0^{\infty} \left\{-\frac{\partial \theta_i e^{-\theta_i x_i}}{\partial x_i}\right\} u(\theta_i)\, d\theta_i}{g(x_i)} \\
&= -\frac{g(x_i)'}{g(x_i)}.
\end{aligned}
$$

3.2.2 Γραμμική μη Παραμετρική Εμπειρική Μπεϋζιανή Προσέγγιση

Εξετάζοντας ξανά τη σχέση (3.1),

$$E(t-\theta_i)^2 = \int_{\Theta}\int_{X}(t(x_i)-\theta_i)^2 f(x_i|\theta_i)u(\theta_i)dx_i\, d\theta_i\ ,$$

διαπιστώνουμε, πως αν η συνάρτηση t ανήκει στην τάξη των γραμμικών συναρτήσεων της μορφής $A+Bx_i$, τότε ο κίνδυνος Bayes του κανόνα απόφασης t, παίρνει την μορφή

$$E(t-\theta_i)^2 = \int_{\Theta}\int_{X}(A+Bx_i-\theta_i)^2\, f(x_i|\theta)\, dx_i\, u(\theta_i)\, d\theta_i\ . \qquad (3.2)$$

Η μερική παράγωγος της παραπάνω εξίσωσης ως προς A (βλέπε Κούτρα, 2005), είναι

$$\begin{aligned}\frac{\partial E(t-\theta_i)^2}{\partial A} &= 2\iint (A+Bx_i-\theta_i)\, f(x|\theta_i)\, dx_i\, u(\theta_i)\, d\theta_i \\ &= 2\iint \{Af(x_i|\theta_i)+Bx_if(x_i|\theta_i)-\theta_i f(x_i|\theta_i)\}\ dx_i\ u(\theta_i)\, d\theta_i \\ &= 2\int [A+B\ E(x_i|\theta_i)-\theta_i\ ]u(\theta_i)\, d\theta_i\ ,\end{aligned}$$

ενώ μηδενίζεται όταν

$$\begin{aligned}&2\int [A+B\ E(x_i|\theta_i)-\theta_i\ ]u(\theta_i)\, d\theta_i = 0 \\ &\Rightarrow A+B\int E(x_i|\theta_i)\, u(\theta_i)\, d\theta_i = \int \theta_i\, u(\theta_i)\, d\theta_i\ ,\end{aligned}$$

επομένως ισχύει η σχέση

$$A+B\ E(x_i) = E(\theta_i)\ . \qquad (3.3)$$

Η ίδια διαδικασία ακολουθείται και ως προς τον άγνωστο B, οπότε

$$\frac{\partial E(t-\theta_i)^2}{\partial B} = 2\iint x_i(A + Bx_i - \theta_i)\, f(x_i|\theta_i)\, dx_i\, u(\theta_i)\, d\theta_i$$

$$= 2\iint Ax_i f(x_i|\theta_i) + Bx_i^2 f(x_i|\theta_i) - x_i\theta_i f(x_i|\theta_i)\ \ dx_i\ \ u(\theta_i)\, d\theta_i$$

$$= 2\int (A\, E(x_i|\theta_i) + B\, E(x_i^2|\theta_i) - \theta_i E(x_i|\theta_i)\,)u(\theta_i)\, d\theta_i\,,$$

οπότε προκύπτει ότι

$$A\, E(x_i) + B\, E(x_i^2) = \int \theta_i\, E(x_i|\theta_i)\, u(\theta_i)\, d\theta_i\,. \tag{3.4}$$

Πολλαπλασιάζοντας την σχέση (3.3) με $E(x_i)$, λαμβάνουμε

$$A\, E(x_i) + B\, \{E(x_i)\}^2 = E(x_i)E(\theta_i)\,, \tag{3.5}$$

και αφαιρώντας την (3.5) από την (3.4) έχουμε

$$B\, E(x_i^2) - B\, \{E(x_i)\}^2 = E(\theta_i x_i) - E(x_i)E(\theta_i)$$

$$\Rightarrow\ B\, Var(x_i) = Cov(\theta_i, x_i)$$

$$\Rightarrow\ B = \frac{Cov(\theta_i, x_i)}{Var(x_i)}\,. \tag{3.6}$$

Αντικαθιστώντας την παραπάνω στην (3.3) προκύπτει ότι

$$A = E(\theta_i) - \frac{Cov(\theta_i, x_i)}{Var(x_i)}\, E(x_i) \tag{3.7}$$

Επομένως ο κανόνας απόφασης t παίρνει την μορφή

$$t(x_i) = E(\theta_i) + \frac{Cov(\theta_i, x_i)}{Var(x_i)}[x_i - E(x_i)] \, . \qquad (3.8)$$

Η γραμμική μη παραμετρική εμπειρική Μπεϋζιανή προσέγγιση μελετήθηκε για πρώτη φορά από τον Robbins (1986). Κατά τον ίδιο η προσέγγιση αυτή ίσως είναι πιο βολική σε κάποιες περιπτώσεις, αφού οι ποσότητες $E(\theta_i)$, $Cov(\theta_i, x_i)$, $E(x_i)$ και $Var(x_i)$ είναι πιο εύκολο να εκτιμηθούν από τα δεδομένα $\boldsymbol{x}$, όταν η $f(x_i|\theta_i)$ ανήκει σε κάποιες συγκεκριμένες οικογένειες κατανομών.

Παράδειγμα 3.6 $f(x_i|\theta_i) \sim Exponential(1/\theta_i)$

Έστω x_i τυχαίο δείγμα που ακολουθεί την *Εκθετική* $(1/\theta_i)$ κατανομή. Επομένως ισχύει πως

$$E(x_i|\theta_i) = \theta_i \quad \text{και} \quad Var(x_i|\theta_i) = {\theta_i}^2 \, .$$

Με βάση τις παραπάνω σχέσεις προκύπτει πως η $E(x_i)$ ειναι

$$E(x_i) = E\big(E(x_i|\theta_i)\big) = E(\theta_i) \, ,$$

Ενώ η $Var(x_i)$ είναι

$$Var(x_i) = E\big(Var(x_i|\theta_i)\big) + Var\big(E(x_i|\theta_i)\big)$$

$$= E\big({\theta_i}^2\big) + Var(\theta_i) \, .$$

Με τη βοήθεια της παραπάνω η $E({x_i}^2)$ γινεται

$$E({x_i}^2) = Var(x_i) + \{E(x_i)\}^2$$

$$= E\big({\theta_i}^2\big) + Var(\theta_i) + E(\theta_i)^2$$

$$= 2E\big({\theta_i}^2\big) \, ,$$

και η $Cov(\theta_i, x_i)$ είναι

$$Cov(\theta_i, x_i) = E\big(Cov(\theta_i, x_i|\theta_i)\big) + Cov\big(E(x_i|\theta_i), E(\theta_i|\theta_i)\big)$$
$$= 0 + Cov(\theta_i, \theta_i)$$
$$= Var(\theta_i)\ .$$

Με βάση τους παραπάνω υπολογισμούς, οι (3,6) και (3,7) γίνονται

$$A = E(\theta_i) - \frac{Cov(\theta_i, x_i)}{Var(x_i)}\ E(x_i)$$

$$= E(\theta_i) - \frac{Var(\theta_i)}{E\left({\theta_i}^2\right) + Var(\theta_i)}\ E(\theta_i)$$

$$= \frac{E(\theta_i)E\left({\theta_i}^2\right)}{E\left({\theta_i}^2\right) + Var(\theta_i)}, \tag{3.9}$$

και

$$B = \frac{Cov(\theta_i, x_i)}{Var(x_i)} = \frac{Var(\theta_i)}{E\left({\theta_i}^2\right) + Var(\theta_i)}. \tag{3.10}$$

Δηλαδή ο κανόνας απόφασης t παίρνει την μορφή

$$t(x_i) = \frac{E\left({\theta_i}^2\right)}{E\left({\theta_i}^2\right) + Var(\theta_i)} E(\theta_i) + \frac{Var(\theta_i)}{E\left({\theta_i}^2\right) + Var(\theta_i)} x_i\,. \tag{3.11}$$

Αντικαθιστώντας τις δυο πρώτες ροπές του x_i, με τις εμπειρικές τους συχνότητες προκύπτει ότι

$$E(x_i) = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$\Rightarrow\ E(\theta_i) = \bar{x} = m_1\,,$$

και

$$E(x_i^2) = \frac{1}{n}\sum_{i=1}^{n} x_i^2 = m_2$$

$$\Rightarrow E(\theta_i^2) = \frac{1}{2}m_2.$$

Επομένως οι (3.9) και (3.10) παίρνουν την παρακάτω μορφή

$$A = \frac{E(\theta_i)E(\theta_i^2)}{E(\theta_i^2) + Var(\theta_i)} = \frac{m_1 m_2}{2(m_2 - m_1^2)},$$

και

$$B = \frac{Var(\theta_i)}{E(\theta_i^2) + Var(\theta_i)} = \frac{m_2 - 2m_1^2}{2(m_2 - m_1^2)}.$$

Έτσι καταλήγουμε πως

$$t(x_i) = \frac{m_1 m_2}{2(m_2 - m_1^2)} + \frac{m_2 - 2m_1^2}{2(m_2 - m_1^2)} x_i \,. \tag{3.12}$$

Συμπεράσματα

Η μη παραμετρική εμπειρική Μπεϋζιανή προσέγγιση, υπολογιστικά συνήθως είναι πολύ πιο εύκολη, από τις προσεγγίσεις που απαιτούν τον καθορισμό της εκ των προτέρων κατανομής.

Υπάρχουν όμως και πολλά μειονεκτήματα, όπως ότι η εκ των υστερών μέση τιμή, σε πολλές περιπτώσεις, δεν είναι δυνατό να εκφραστεί μόνο σε όρους της περιθώριας κατανομής και ότι η προσέγγιση αυτή έχει ικανοποιητικά αποτελέσματα μόνο για μεγάλα δείγματα (βλέπε Berger, 1985 και Robert, 2007).

Η ύπαρξη αυτών των αδυναμιών, κάνει πιο ελκυστική την παραμετρική εκδοχή της εμπειρικής Μπεϋζιανής που αναλύεται στο επόμενο Κεφάλαιο.

Κεφάλαιο 4

4. ΠΑΡΑΜΕΤΡΙΚΗ ΕΜΠΕΙΡΙΚΗ ΜΠΕΫΖΙΑΝΗ ΠΡΟΣΕΓΓΙΣΗ

4.1 Παραμετρική Εμπειρική Μπεϋζιανή *(Parametric Empirical Bayes)*

Στη Μπεϋζιανή ανάλυση, η εκ των προτέρων κατανομή του θ, θεωρείται πως ακολουθεί μια γνωστή κατανομή με γνωστές υπερπαραμέτρους. Στην περίπτωση που οι υπερπαράμετροι λ δεν μπορούν να προσδιοριστούν, οδηγούμαστε στο ιεραρχικό μοντέλο. Η ιεραρχία του μοντέλου πρέπει να τερματιστεί σε κάποιο σημείο, με τις εναπομείνασες υπερπαραμέτρους να υποτίθεται πως είναι γνωστές.

Αντί αυτής της υπόθεσης, η παραμετρική εμπειρική Μπεϋζιανή προσέγγιση εκτιμά την υπερπαράμετρο του τελευταίου επιπέδου του ιεραρχικού μοντέλου, μέσα από τα παρατηρηθέντα δεδομένα (βλέπε Carlin & Louis, 2008).

Ας υποτεθεί ξανά το αμοιβαία ανεξάρτητο ιεραρχικό μοντέλο κατά το οποίο

$$X_i|\theta_i \overset{ind}{\sim} f(x_i|\theta_i), \qquad \kappa\alpha\iota$$

$$\theta_i \overset{iid}{\sim} u(\theta_i|\lambda), \qquad \text{οπου } i = 1, \dots, n\,.$$

Σύμφωνα με τον Casella (1992), είναι αναγκαία η αλλαγή του στατιστικού μοντέλου που πραγματεύεται η παραμετρική εμπειρική Μπεϋζιανή ανάλυση, σε σχέση με αυτού της Μπεϋζιανής, επειδή οι παράμετροι που πρέπει να εκτιμηθούν στην προσέγγιση αυτή, είναι περισσότερες από τα δεδομένα.

Όπως και στη Μπεϋζιανή ανάλυση, το πρώτο βήμα είναι ο ορισμός της συνάρτησης πιθανότητας $f(x_i|\theta_i)$, δεδομένου της παραμέτρου θ_i. Η κατανομή που ακολουθεί το θ_i ονομάζεται εκ των προτέρων κατανομή και ορίζεται ως $u(\theta_i|\lambda)$, όπου με λ συμβολίζεται η υπερπαράμετρος της κατανομής αυτής.

Θα πρέπει να τονιστεί πως οι παράμετροι θ_i, συνδέονται μεταξύ τους μέσω της κοινής υπερπαραμέτρου λ. Η εκτίμηση της υπερπαραμέτρου λ γίνεται με τη βοήθεια της περιθώριας κατανομής όλων των δεδομένων, και οι μέθοδοι που ακολουθούνται είναι εκείνες που χρησιμοποιούνται στην κλασσική στατιστική, δηλαδή η μέθοδος μέγιστης πιθανοφάνειας ή η μέθοδος των ροπών. Η περιθώρια κατανομή των δεδομένων, υπό το παραπάνω μοντέλο ικανοποίει την εξίσωση

$$
\begin{aligned}
g(x|\lambda) &= \int f(\boldsymbol{x}|\boldsymbol{\theta})u(\boldsymbol{\theta}|\lambda)d\theta \\
&= \int \prod_{i=1}^{n} f_i(x_i|\theta_i) \prod_{i=1}^{n} u(\theta_i|\lambda)d\theta \\
&= \prod_{i=1}^{n} \int f_i(x_i|\theta_i)u(\theta_i|\lambda)d\theta_i \\
&= \prod_{i=1}^{n} g(x_i|\lambda) \,.
\end{aligned}
$$

Έπειτα μέσω του θεωρήματος του Bayes καταλήγουμε στην αναθεωρημένη πληροφορία του θ_i. Η κατανομή που ακολουθεί το θ_i ονομάζεται *εκτιμώμενη εκ των υστέρων κατανομή (estimated posterior*) και δίνεται από τον παρακάτω τύπο:

$$
p(\theta_i|x_i\,,\hat{\lambda}) = \frac{f(x_i|\theta_i)u(\theta_i|\hat{\lambda})}{g(x_i|\hat{\lambda})} \,.
$$

Έχοντας πλέον καταλήξει στην εκτιμώμενη εκ των υστέρων κατανομή, είναι δυνατή η συμπερασματολογία πάνω στην παράμετρο θ_i. Η μεθοδολογία που θα ακολουθηθεί για την συμπερασματολογία, είναι κοινή με εκείνη της Μπεϋζιανής ανάλυσης (βλέπε Carlin & Louis, 2008).

Παρακάτω παρατίθενται παραδείγματα τα οποία είναι κοινά με εκείνα του προηγουμένου κεφαλαίου ως προς την κατανομή που πραγματεύονται, έχοντας σαν στόχο την ανάδειξη των διαφορών της μεθοδολογίας των δυο προσεγγίσεων, αλλά και την καλύτερη κατανόηση τους.

Παράδειγμα 4.1: $f(x_i|\theta_i) \sim Poisson(\theta_i) - \ u(\theta_i|\alpha,\beta) \sim Gamma(\alpha,\beta)$

Έστω ότι η συνάρτηση πιθανότητας $f(x_i|\theta_i)$, με $i = 1, 2, \ldots, n$, ακολουθεί την κατανομή $Poisson(\theta_i)$, ισχύει δηλαδή ότι

$$
f(x_i|\theta_i) \;=\; \frac{e^{-\theta_i}{\theta_i}^{x_i}}{x_i!},
$$

ενώ και η κατανομή της εκ των προτέρων πληροφορίας του θ_i, ακολουθεί την κατανομή $Gamma(\alpha, \beta)$

$$\pi(\theta_i|\alpha,\beta) = \frac{{\theta_i}^{\alpha-1} e^{-\theta_i/\beta}}{\beta^\alpha \Gamma(\alpha)}.$$

Η περιθώρια συνάρτηση $g(x_i|\alpha,\beta)$, υπολογίζεται ως εξής

$$\begin{aligned}
g(x_i|\alpha,\beta) &= \int \left(\frac{e^{-\theta_i}{\theta_i}^{x_i}}{x_i!}\right)\left(\frac{{\theta_i}^{\alpha-1} e^{-\frac{\theta_i}{\beta}}}{\beta^\alpha \Gamma(\alpha)}\right) d\theta_i \\
&= \int \frac{e^{-\theta_i}\, e^{-\frac{\theta_i}{\beta}}\, {\theta_i}^{x_i}\, {\theta_i}^{\alpha-1}}{x_i!\, \beta^\alpha \Gamma(\alpha)} d\theta_i \\
&= \int \frac{e^{-\left(\theta_i + \frac{\theta_i}{\beta}\right)} {\theta_i}^{\alpha+x_i-1}}{x_i!\, \beta^\alpha \Gamma(\alpha)} d\theta_i \\
&= \int \frac{e^{-\theta_i\left(\frac{\beta+1}{\beta}\right)} {\theta_i}^{\alpha+x_i-1}}{x_i!\, \beta^\alpha \Gamma(\alpha)} d\theta_i \\
&= \frac{1}{x_i!\, \beta^\alpha \Gamma(\alpha)} \int e^{-\theta_i\left(\frac{\beta+1}{\beta}\right)} {\theta_i}^{\alpha+x_i-1}\, d\theta_i \\
&= \frac{\left(\frac{\beta}{\beta+1}\right)^{\alpha+x_i} \Gamma(\alpha+x_i)}{x_i!\, \beta^\alpha \Gamma(\alpha)} \int \frac{e^{-\theta_i\left(\frac{\beta+1}{\beta}\right)} {\theta_i}^{\alpha+x_i-1}}{\left(\frac{\beta}{\beta+1}\right)^{\alpha+x_i} \Gamma(\alpha+x_i)} d\theta_i \\
&= \frac{\left(\frac{\beta}{\beta+1}\right)^{\alpha+x_i} \Gamma(\alpha+x_i)}{x_i!\, \beta^\alpha \Gamma(\alpha)} \\
&= \frac{\Gamma(\alpha+x_i)}{x_i!\, \Gamma(\alpha)} \frac{\left(\frac{\beta}{\beta+1}\right)^{\alpha} \left(\frac{\beta}{\beta+1}\right)^{x_i}}{\beta^\alpha} \\
&= \frac{\Gamma(\alpha+x_i)}{\Gamma(x_i+1)\Gamma(\alpha)} \frac{\left(\frac{\beta}{\beta+1}\right)^{\alpha} \left(\frac{\beta}{\beta+1}\right)^{x_i}}{\beta^\alpha} \\
&= \frac{\Gamma(\alpha+x_i)}{\Gamma(x_i+1)\Gamma(\alpha)} \left(\frac{1}{\beta+1}\right)^{\alpha} \left(\frac{\beta}{\beta+1}\right)^{x_i}.
\end{aligned}$$

Από τον παραπάνω υπολογισμό της περιθώριας κατανομής διαπιστώνεται πως η $g(x_i|\alpha,\beta)$ ακολουθεί την αρνητική διωνυμική κατανομή $\left(\alpha, \frac{\beta}{\beta+1}\right)$.

Έχοντας υπολογίσει την περιθώρια κατανομή, γίνεται πλέον δυνατή η εκτίμηση των υπερπαραμέτρων. Η μέθοδος που θα χρησιμοποιηθεί για την εκτίμηση στην συγκεκριμένη περίπτωση είναι εκείνη της μεθόδου των ροπών, αφού η μέθοδος της μέγιστης πιθανοφάνειας στην συγκεκριμένη περίπτωση είναι πολύ πιο περίπλοκη. Έτσι λοιπών αντικαθιστώντας τις δυο πρώτες ροπές, με τις αντίστοιχες εμπειρικές συχνότητες έχουμε

$$E(x_i) = \bar{x} = m_1 \quad \kappa\alpha\iota \quad E(x_i{}^2) = \overline{x^2} = m_2 \,.$$

Άρα

$$m_1 = \alpha\left\{\frac{\beta}{\beta+1}\Big/\frac{1}{\beta+1}\right\}$$
$$= \alpha\beta$$

και

$$m_2 = \alpha\left\{\frac{\beta}{\beta+1}\Big/\left(\frac{1}{\beta+1}\right)^2\right\} + \alpha^2\beta^2$$
$$= \alpha\beta(1+\beta) + \alpha^2\beta^2$$

Άρα οι εκτιμητές των υπερπαραμέτρων *α, β* είναι οι

$$\hat{\beta} = \frac{m_2 - m_1{}^2}{m_1} - 1 \quad \kappa\alpha\iota \quad \hat{\alpha} = \frac{m_1}{\hat{\beta}} \,.$$

Μετά από εφαρμογή του Μπεϋζιανού θεωρήματος καταλήγουμε στην εκ των υστέρων κατανομή του θ_i

$$p(\theta_i|x_i,\alpha,\beta) = \frac{\left(\frac{e^{-\theta_i}\theta_i{}^{x_i}}{x_i!}\right)\left(\frac{\theta_i{}^{\alpha-1}e^{-\theta_i/\beta}}{\beta^\alpha\Gamma(\alpha)}\right)}{\left(\frac{\Gamma(\alpha+x_i)}{\Gamma(x_i+1)\Gamma(\alpha)}\left(\frac{1}{\beta+1}\right)^\alpha\left(\frac{\beta}{\beta+1}\right)^{x_i}\right)}$$

$$= \frac{\theta_i^{(\alpha+x_i)-1} e^{-\left(\frac{\beta+1}{\beta}\right)\theta_i}}{\Gamma(\alpha+x_i)\left(\frac{\beta}{\beta+1}\right)^{\alpha+x_i}}.$$

Επομένως, η εκ των υστέρων κατανομή του θ_i ακολουθεί την *Gamma* κατανομή με παραμέτρους $\left(\alpha + x_i,\ \frac{\beta}{\beta+1}\right)$.

Αντικαθιστώντας τις υπερπαραμέτρους, με τις αντίστοιχες εκτιμήτριές τους, καταλήγουμε στην εκτιμώμενη εκ των υστέρων κατανομή $p\left(\theta_i \middle| x_i, \hat{\alpha}, \hat{\beta}\right)$, η οποία ακολουθεί την $Gamma\left(\hat{\alpha} + x_i,\ \frac{\hat{\beta}}{\hat{\beta}+1}\right)$, ενώ η εκ των υστέρων εκτιμώμενη μέση τιμή είναι

$$E(\theta_i|x_i) = (\hat{\alpha} + x_i)\left(\frac{\hat{\beta}}{\hat{\beta}+1}\right)$$

$$= \left(\frac{m_1{}^2}{m_2 - m_1{}^2 - m_1} + x_i\right)\left(\frac{m_2 - m_1{}^2 - m_1}{m_2 - m_1{}^2}\right)$$

$$= \left(\frac{m_1}{m_2 - m_1{}^2}\right) m_1 + \left(1 - \frac{m_1}{m_2 - m_1{}^2}\right) x_i\,.$$

Παράδειγμα 4.2: $f(x_i|\theta_i) \sim N(\theta_i, \sigma^2) - u(\theta_i|\mu) \sim N(\mu, r^2)$

Υποθέτοντας το μοντέλο δυο επιπέδων, όπου η $f(x_i|\theta_i) \sim N(\theta_i, \sigma^2)$, με $i = 1, 2, \dots, n$, δηλαδή,

$$f(x_i|\theta_i) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-\theta_i)^2}{2\sigma^2}},$$

ενώ η εκ των προτέρων κατανομή ακολουθεί την $N(\mu, r^2)$ και έχει την παρακάτω συνάρτηση πιθανότητας

$$u(\theta_i|\mu) = \frac{1}{\sqrt{2\pi r^2}} e^{-\frac{(\theta_i-\mu)^2}{2r^2}},$$

όπου τα σ^2 και r^2 θεωρείται πως είναι γνωστά και $i = 1, \dots, n$, η περιθώρια συνάρτηση υπολογίζεται ως εξής:

$$g(x_i|\mu) =$$

$$= \int \left(\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-\theta_i)^2}{2\sigma^2}}\right)\left(\frac{1}{\sqrt{2\pi r^2}} e^{-\frac{(\theta_i-\mu)^2}{2r^2}}\right) d\theta_i$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}\,\sqrt{2\pi r^2}} \int exp\left\{-\frac{(x_i-\theta_i)^2}{2\sigma^2} - \frac{(\theta_i-\mu)^2}{2r^2}\right\} d\theta_i$$

$$= \frac{1}{2\pi\,\sqrt{\sigma^2\, r^2}} \int exp\left\{-\frac{{x_i}^2 - 2x_i\theta_i + {\theta_i}^2}{2\sigma^2} - \frac{{\theta_i}^2 - 2\mu\theta_i + \mu^2}{2r^2}\right\} d\theta_i$$

$$= \frac{1}{2\pi\,\sqrt{\sigma^2\, r^2}} \int exp\left\{-\frac{{x_i}^2}{2\sigma^2} - \frac{-2x_i\theta_i + {\theta_i}^2}{2\sigma^2} - \frac{\mu^2}{2r^2} - \frac{{\theta_i}^2 - 2\mu\theta_i}{2r^2}\right\} d\theta_i$$

$$= \frac{1}{2\pi\,\sqrt{\sigma^2\, r^2}}\ exp\left\{-\frac{{x_i}^2}{2\sigma^2} - \frac{\mu^2}{2r^2}\right\} \int exp\left\{\frac{2x_i\theta_i}{2\sigma^2} - \frac{{\theta_i}^2}{2\sigma^2} - \frac{{\theta_i}^2}{2r^2} + \frac{2\mu\theta_i}{2r^2}\right\} d\theta_i$$

$$= \frac{1}{2\pi\sqrt{\sigma^2 r^2}}\ exp - \left\{\frac{{x_i}^2}{2\sigma^2} + \frac{\mu^2}{2r^2}\right\} \int exp\left\{-{\theta_i}^2\left(\frac{1}{2\sigma^2} + \frac{1}{2r^2}\right) + 2\theta_i\left(\frac{\mu}{2r^2} + \frac{x_i}{2\sigma^2}\right)\right\} d\theta_i$$

$$= \frac{1}{2\pi\sigma r} exp - \left\{\frac{{x_i}^2 r^2 + \mu^2\sigma^2}{2\sigma^2 r^2}\right\} \int exp\left\{-{\theta_i}^2\left(\frac{r^2+\sigma^2}{2\sigma^2 r^2}\right) + 2\theta_i\left(\frac{x_i r^2 + \mu\sigma^2}{2\sigma^2 r^2}\right)\right\} d\theta_i$$

$$= \frac{1}{2\pi\sigma r} exp - \left\{\frac{{x_i}^2 r^2 + \mu^2\sigma^2}{2\sigma^2 r^2}\right\} A\,,$$

όπου

$$A = \int exp\left\{-\theta_i{}^2\left(\frac{r^2+\sigma^2}{2\sigma^2r^2}\right)+2\theta_i\left(\frac{x_ir^2+\mu\sigma^2}{2\sigma^2r^2}\right)\right\}d\theta_i$$

$$= \int exp\left\{-\frac{r^2+\sigma^2}{2\sigma^2r^2}\left[\theta_i{}^2-2\theta_i\left(\frac{x_i\,r^2+\mu\,\sigma^2}{r^2\;+\;\sigma^2}\right)\pm\left(\frac{x_i\,r^2+\mu\,\sigma^2}{r^2\;+\;\sigma^2}\right)^2\;\right]\right\}d\theta_i$$

$$= \int exp\left\{-\frac{r^2+\sigma^2}{2\sigma^2r^2}\left[\left(\theta_i-\frac{x_ir^2+\mu\sigma^2}{r^2\,\sigma^2}\right)^2-\left(\frac{x_ir^2+\mu\sigma^2}{r^2\;+\;\sigma^2}\right)^2\right]\right\}d\theta_i$$

$$= exp\left\{\frac{r^2+\sigma^2}{2\sigma^2r^2}\;\left(\frac{x_ir^2+\mu\sigma^2}{r^2\;+\;\sigma^2}\right)^2\right\}\int exp\left\{-\frac{r^2\;+\;\sigma^2}{2\sigma^2r^2}\;\left(\theta_i-\frac{x_i\,r^2+\mu\,\sigma^2}{r^2\;+\;\sigma^2}\right)^2\right\}d\theta_i$$

$$= \sqrt{2\pi}\sqrt{\frac{\sigma^2\,r^2}{r^2\;+\;\sigma^2}}\,exp\left\{\frac{r^2\;+\;\sigma^2}{2\sigma^2r^2}\;\left(\frac{x_i\,r^2+\mu\,\sigma^2}{r^2\;+\;\sigma^2}\right)^2\right\}.$$

Επομένως η $g(x_i|\mu)$ γίνεται

$g(x_i|\mu) =$

$$= \frac{1}{2\pi\sigma r}exp-\left\{\frac{x_i{}^2r^2+\mu^2\sigma^2}{2\sigma^2r^2}\right\}\sqrt{2\pi}\sqrt{\frac{\sigma^2\,r^2}{r^2\;+\;\sigma^2}}\,exp\left\{\frac{r^2\;+\;\sigma^2}{2\sigma^2r^2}\;\left(\frac{x_i\,r^2+\mu\,\sigma^2}{r^2\;+\;\sigma^2}\right)^2\right\}$$

$$= \frac{1}{\sqrt{2\pi(\sigma^2+r^2)}}\;exp-\left\{\frac{x_i{}^2r^2+\mu^2\sigma^2}{2\sigma^2r^2}\;+\;\frac{r^2\;+\;\sigma^2}{2\sigma^2r^2}\;\left(\frac{x_i\,r^2+\mu\,\sigma^2}{r^2\;+\;\sigma^2}\right)^2\right\}$$

$$= \frac{1}{\sqrt{2\pi(\sigma^2+r^2)}}\;exp-\left\{\frac{1}{2\sigma^2r^2}\left[x_i{}^2r^2+\mu^2\sigma^2\;-(r^2+\sigma^2)\left(\frac{x_i\,r^2+\mu\,\sigma^2}{r^2\;+\;\sigma^2}\right)^2\right]\right\}$$

$$= \frac{1}{\sqrt{2\pi(\sigma^2+r^2)}}\;exp-\left\{\frac{1}{2\sigma^2r^2}\left(\frac{x_i{}^2r^2\sigma^2\;-\;2x_i\,r^2\mu\,\sigma^2+\;\mu^2r^2\sigma^2}{r^2\;+\;\sigma^2}\right)\right\}$$

$$= \frac{1}{\sqrt{2\pi(\sigma^2+r^2)}}\;exp-\left\{\frac{(x_i-\mu)^2}{2(\sigma^2+r^2)}\right\}.$$

Άρα η περιθώρια κατανομή, ακολουθεί την κανονική κατανομή με παραμέτρους $(\mu, \sigma^2 + r^2)$.

Η μέθοδος που θα χρησιμοποιηθεί για την εκτίμηση στη συγκεκριμένη περίπτωση είναι εκείνη της μεγίστης πιθανοφάνειας. Γενικά η μέθοδος της μεγίστης πιθανοφάνειας έχει σαν στόχο τη μεγιστοποίηση της συνάρτησης πιθανοφάνειας ή ισοδύναμα της λογοπιθανοφάνειας, αφού μεγιστοποιούνται στο ίδιο σημείο.

Η μεγιστοποίηση γίνεται ως προς την παράμετρο που είναι επιθυμητό να εκτιμηθεί. Όσον αφορά την $g(x_i|\mu)$, η μοναδική παράμετρος που χρειάζεται να εκτιμηθεί είναι η μ, αφού στο συγκεκριμένο παράδειγμα η r θεωρείται γνωστή. Η πιθανοφάνεια της περιθώριας κατανομής είναι

$$g(\boldsymbol{x}|\mu) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi(\sigma^2 + r^2)}} \, exp - \left\{\frac{(x_i - \mu)^2}{2(\sigma^2 + r^2)}\right\}$$

$$= \frac{1}{[2\pi(\sigma^2 + r^2)]^{n/2}} exp - \left\{\frac{1}{2(\sigma^2 + r^2)} \sum_{i=1}^{n} (x_i - \mu)^2\right\},$$

ενώ ο λογάριθμος της $g(\boldsymbol{x}|\mu)$ είναι

$$log\big(g(\boldsymbol{x}|\mu)\big) = log\left(\frac{1}{[2\pi(\sigma^2 + r^2)]^{n/2}} exp - \left\{\frac{1}{2(\sigma^2 + r^2)} \sum_{i=1}^{n} (x_i - \mu)^2\right\}\right)$$

$$= - nlog\sqrt{(\sigma^2 + r^2)} - nlog\sqrt{2\pi} - \frac{1}{2(\sigma^2 + r^2)} \sum_{i=1}^{n} (x_i - \mu)^2 .$$

Η παράγωγος της παραπάνω ως προς μ, είναι

$$\frac{\partial log\big(g(\boldsymbol{x}|\mu)\big)}{\partial \mu} = - \frac{1}{2(\sigma^2 + r^2)} \frac{\partial(\sum_{i=1}^{n} (x_i - \mu)^2)}{\partial \mu}$$

$$= \frac{1}{2(\sigma^2 + r^2)} \sum_{i=1}^{n} 2(x_i - \mu)$$

$$= \frac{1}{(\sigma^2 + r^2)} \left\{\left(\sum_{i=1}^{n} x_i\right) - n\mu\right\}.$$

Συνεπώς η παράγωγος της περιθώριας πιθανοφάνειας μηδενίζεται όταν $\mu = \bar{x}$.

Μετά από εφαρμογή του Μπεϋζιανού θεωρήματος καταλήγουμε στην εκ των υστέρων κατανομή του θ_i

$$p(\theta_i|x_i,\alpha,\beta) = \frac{\left(\frac{1}{\sqrt{2\pi\sigma^2}} exp\left\{-\frac{(x_i-\theta_i)^2}{2\sigma^2}\right\}\right)\left(\frac{1}{\sqrt{2\pi r^2}} exp\left\{-\frac{(\theta_i-\mu)^2}{2r^2}\right\}\right)}{\left(\frac{1}{\sqrt{2\pi(\sigma^2+r^2)}} exp\left\{-\frac{(x_i-\mu)^2}{2(\sigma^2+r^2)}\right\}\right)}$$

$$= \frac{\sqrt{2\pi(\sigma^2+r^2)}\ exp\left\{-\frac{(x_i-\theta_i)^2}{2\sigma^2} - \frac{(\theta_i-\mu)^2}{2r^2}\right\}}{2\pi\sqrt{\sigma^2 r^2}\ exp\left\{-\frac{(x_i-\mu)^2}{2(\sigma^2+r^2)}\right\}}$$

$$= \frac{1}{\sqrt{2\pi\frac{\sigma^2 r^2}{\sigma^2+r^2}}}\ exp\left\{-\frac{(x_i-\theta_i)^2}{2\sigma^2} + \frac{(x_i-\mu)^2}{2(\sigma^2+r^2)} - \frac{(\theta_i-\mu)^2}{2r^2}\right\}$$

$$= \frac{1}{\sqrt{2\pi\frac{\sigma^2 r^2}{\sigma^2+r^2}}}\ exp\{B\},$$

όπου ο εκθέτης γίνεται

$B =$

$$= -\frac{(x_i-\theta_i)^2}{2\sigma^2} + \frac{(x_i-\mu)^2}{2(\sigma^2+r^2)} - \frac{(\theta_i-\mu)^2}{2r^2}$$

$$= \frac{-r^2(\sigma^2+r^2)(x_i-\theta_i)^2 + \sigma^2 r^2(x_i-\mu)^2 - \sigma^2(\sigma^2+r^2)(\theta_i-\mu)^2}{2\sigma^2 r^2(\sigma^2+r^2)}$$

$$= \left\{\frac{-(r^4+\sigma^2 r^2)(x_i-\theta_i)^2 + \sigma^2 r^2(x_i-\mu)^2 - (\sigma^2 r^2+\sigma^4)(\theta_i-\mu)^2}{2\sigma^2 r^2(\sigma^2+r^2)}\right\}$$

$$= \left\{\frac{-(\sigma^4+2\sigma^2 r^2+r^4){\theta_i}^2 + 2(\sigma^2 r^2 x_i + \sigma^4\mu + r^4 x_i + \sigma^2 r^2\mu)\theta_i - r^4 x_i^2 - 2\sigma^2 r^2 x_i\mu - \sigma^4\mu^2}{2\sigma^2 r^2(\sigma^2+r^2)}\right\}$$

$$= \frac{1}{2\frac{\sigma^2 r^2}{\sigma^2+r^2}}\left\{\frac{-(\sigma^2+r^2)^2{\theta_i}^2+2(\sigma^2+r^2)(r^2x_i+\sigma^2\mu)\theta_i-(r^2x_i+\sigma^2\mu)^2}{(\sigma^2+r^2)^2}\right\}$$

$$= \frac{1}{2\frac{\sigma^2 r^2}{\sigma^2+r^2}}\left\{-{\theta_i}^2+2\frac{r^2x_i+\sigma^2\mu}{\sigma^2+r^2}\theta_i-\left(\frac{r^2x_i+\sigma^2\mu}{\sigma^2+r^2}\right)^2\right\}$$

$$= -\frac{1}{2\frac{\sigma^2 r^2}{\sigma^2+r^2}}\left(\theta_i-\frac{r^2x_i+\sigma^2\mu}{\sigma^2+r^2}\right)^2 .$$

Άρα η $p(\theta_i|x_i,\alpha,\beta)$ γίνεται

$$p(\theta_i|x_i,\alpha,\beta) = \frac{1}{\sqrt{2\pi\frac{\sigma^2r^2}{\sigma^2+r^2}}}\, exp - \left\{\frac{1}{2\frac{\sigma^2 r^2}{\sigma^2+r^2}}\left(\theta_i-\frac{r^2x_i+\sigma^2\mu}{\sigma^2+r^2}\right)^2\right\}$$

Έτσι η εκτιμώμενη εκ των υστέρων κατανομή $p(\theta_i|x_i,\hat{\mu})$ είναι $N\left(\frac{x_ir^2+\bar{x}\sigma^2}{r^2+\sigma^2},\frac{\sigma^2r^2}{r^2+\sigma^2}\right)$, με σημειακή εκτίμηση του θ_i την

$$\hat{\theta}_i^{\mu} = \frac{x_ir^2+\bar{x}\sigma^2}{r^2+\sigma^2}$$

$$= \frac{\sigma^2}{r^2+\sigma^2}\bar{x}+\frac{r^2}{r^2+\sigma^2}x_i .$$

Θα πρέπει να σημειωθεί πως η παραπάνω σχέση, διαφέρει από την αντίστοιχη μέση τιμή της Μπεϋζιανής προσέγγισης (βλέπε Παράδειγμα 2.2), στο ότι η γνωστή εκ των προτέρων μέση τιμή μ, αντικαταστάθηκε με τον δειγματικό μέσο $\bar{x}$. Έτσι επιβεβαιώνεται η παρατήρηση που αναφέρεται παραπάνω, δηλαδή ότι η παραμετρική εμπειρική Μπεϋζιανή συμπερασματολογία της παραμέτρου θ_i, εμπεριέχει πληροφορίες από όλα τα δεδομένα, μέσω της υπερπαραμέτρου (βλέπε Carlin & Louis, 2008).

Σημείωση

Επίσης, αν η υπερπαράμετρος του μοντέλου r^2 ήταν επίσης άγνωστη, υπό την ίδια διαδικασία θα διαπιστωνόταν πως $r^2=(s^2-\sigma^2)^+$, όπου $s^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$.

Παράδειγμα 4.3: $f(x_i|\theta_i) \sim Binomial(k, \theta_i) - \ u(\theta_i|\alpha, \beta) \sim \ Beta(\alpha, \beta)$

Έστω το παρακάτω σύνθετο μοντέλο δειγματοληψίας δυο επιπέδων, όπου η συνάρτηση πιθανότητας του μοντέλου είναι $Binomial(k, \theta_i)$, με $i = 1, 2, \dots, n$, δηλαδή,

$$f(x_i|\theta_i) \ = \ \binom{k}{x_i} \theta_i^{\ x_i} (1 - \theta_i)^{k - x_i},$$

ενώ και η κατανομή της εκ των προτέρων πληροφορίας του θ_i, ακολουθεί την κατανομή $Beta(\alpha, \beta)$

$$u(\theta_i|\alpha, \beta) = \frac{1}{B(\alpha, \beta)} \theta_i^{\ \alpha - 1} (1 - \theta_i)^{\beta - 1} \, .$$

Τότε η περιθώρια συνάρτηση $g(x_i|\alpha, \beta)$, υπολογίζεται ως εξής:

$$g(x_i|\mu) \ = \ \int \left(\binom{k}{x_i} \theta_i^{\ x_i} (1 - \theta_i)^{k - x_i} \right) \left(\frac{1}{B(\alpha, \beta)} \theta_i^{\ \alpha - 1} (1 - \theta_i)^{\beta - 1} \right) d\theta_i$$

$$= \binom{k}{x_i} \frac{1}{B(\alpha, \beta)} \int \theta_i^{\ (\alpha + x_i) - 1} (1 - \theta_i)^{(\beta + k - x_i) - 1} \, d\theta_i$$

$$= \binom{k}{x_i} \frac{B(\alpha + x_i, \ \beta + k - x_i)}{B(\alpha, \beta)} \int \frac{\theta_i^{\ (\alpha + x_i) - 1} (1 - \theta_i)^{(\beta + k - x_i) - 1}}{B(\alpha + x_i, \ \beta + k - x_i)} d\theta_i$$

$$= \binom{k}{x_i} \frac{B(\alpha + x_i, \ \beta + k - x_i)}{B(\alpha, \beta)} \, .$$

Συνεπώς η περιθώρια ακολουθεί την κατανομή *Beta-Binominal* με παραμέτρους $(k, \ \alpha, \ \beta)$. Και σε αυτή την περίπτωση, για την εκτίμηση των υπερπαραμέτρων, είναι περισσότερο εύχρηστη η μέθοδος των ροπών. Έτσι λοιπόν, αντικαθιστώντας τις δυο πρώτες ροπές με τις αντίστοιχες εμπειρικές συχνότητες έχουμε

$$E(x_i) = \bar{x} \ = m_1 \quad \text{και} \quad E(x_i^{\ 2}) = \overline{x^2} = m_2 \, .$$

Άρα

$$m_1 = \frac{k\alpha}{\alpha+\beta}$$
$$\Rightarrow \hat{\alpha} = \frac{m_1}{k-m_1}\beta$$

και

$$m_2 = \frac{k\alpha(k\alpha+k+\beta)}{(\alpha+\beta)(\alpha+\beta+1)}$$
$$= m_1\frac{(k\alpha+k+\beta)}{(\alpha+\beta+1)}$$
$$= m_1\frac{\left(k\frac{m_1}{k-m_1}\beta+k+\beta\right)}{\left(\frac{m_1}{k-m_1}\beta+\beta+1\right)}.$$

Επομένως οι εκτιμητές των υπερπαραμέτρων *α, β* είναι οι

$$\hat{\alpha} = \frac{m_1(m_2-km_1)}{(k-1){m_1}^2+k(m_1-m_2)}$$

και

$$\hat{\beta} = \frac{(k-m_1)(m_2-k\,m_1)}{(k-1){m_1}^2+k(m_1-m_2)}.$$

Μετά από εφαρμογή του Μπεϋζιανού θεωρήματος και την αντικατάσταση των υπερπαραμέτρων με τους εκτιμητές τους, καταλήγουμε στην εκτιμώμενη εκ των υστέρων κατανομή του θ_i

$$p\left(\theta_i\middle|x_i,\hat{\alpha},\hat{\beta}\right) = \frac{{\theta_i}^{(\hat{\alpha}+x_i)-1}(1-\theta_i)^{(\hat{\beta}+k-x_i)-1}}{B(\hat{\alpha}+x_i,\ \hat{\beta}+k-x_i)},$$

η οποία είναι $Beta\left(\hat{\alpha}+x_i,\ \hat{\beta}+k-x_i\right)$, ενώ η εκ των υστέρων εκτιμώμενη μέση τιμή υπολογίζεται όπως κάτωθι

$$E(\theta_i|x_i) = \frac{(\hat{\alpha}+x_i)}{\left(\hat{\alpha}+\hat{\beta}+k\right)}.$$

Παράδειγμα 4.4: $f(x_i|\theta_i) \sim Geometric(\theta_i) - u(\theta_i|\alpha,\beta) \sim Beta(\alpha,\beta)$

Ας υποθέσουμε το μοντέλο όπου $f(x_i|\theta_i)$ είναι Γεωμετρική με παράμετρο (θ_i), με $i = 1, 2, \dots, n$, δηλαδή η συνάρτηση πιθανότητας του μοντέλου

$$f(x_i|\theta_i) = (1-\theta_i){\theta_i}^{x_i},$$

ενώ και η κατανομή της εκ των προτέρων πληροφορίας του θ_i, ακολουθεί την κατανομή $Beta(\alpha,\beta)$

$$u(\theta_i|\alpha,\beta) = \frac{1}{B(\alpha,\beta)}{\theta_i}^{\alpha-1}(1-\theta_i)^{\beta-1}\,.$$

Μια ενδιαφέρουσα πρόταση για τον υπολογισμό των εκτιμητών των υπερπαραμέτρων, δίνεται από τον Maritz (1970) και διατυπώνεται όπως κάτωθι.

Όπως και προηγουμένως θα υπολογιστεί η περιθώρια κατανομή

$$\begin{aligned} g(x_i|\alpha,\beta) &= \int \left((1-\theta_i){\theta_i}^{x_i}\right)\left(\frac{1}{B(\alpha,\beta)}{\theta_i}^{\alpha-1}(1-\theta_i)^{\beta-1}\right)d\theta_i \\ &= \frac{1}{B(\alpha,\beta)}\int {\theta_i}^{(\alpha+x_i)-1}(1-\theta_i)^{(\beta+1)-1}\,d\theta_i \\ &= \frac{B(\alpha+x_i,\ \beta+1)}{B(\alpha,\ \beta)}\,. \end{aligned}$$

Για την περιθώρια ισχύει

$$\begin{aligned} g(0|\alpha,\beta) &= \frac{B(\alpha,\ \beta+1)}{B(\alpha,\ \beta)} \\ &= \frac{(\alpha-1)!\,(\beta)!}{(\alpha+\beta)!}\Big/\frac{(\alpha-1)!\,(\beta-1)!}{(\alpha+\beta-1)!} \\ &= \frac{\beta}{\alpha+\beta} = w_0 \end{aligned}$$

και

$$g(1|\alpha,\beta) = \frac{B(\alpha+1,\ \beta+1)}{B(\alpha,\ \beta)}$$

$$= \frac{(\alpha)!\,(\beta)!}{(\alpha+\beta+1)!}\Big/\frac{(\alpha-1)!\,(\beta-1)!}{(\alpha+\beta-1)!}$$

$$= \frac{\alpha\,\beta}{(\alpha+\beta+1)(\alpha+\beta)} = w_1\,.$$

Με τη βοήθεια των παραπάνω σχέσεων, είναι δυνατό να εκφράσουμε τις υπερπαραμέτρους, σε όρους του w_0 και w_1, δηλαδή

$$w_0 = \frac{\beta}{\alpha+\beta}$$

$$\Rightarrow \beta = \frac{w_0}{1-w_0}\alpha \tag{4.1}$$

και

$$w_1 = \frac{\alpha\,\beta}{(\alpha+\beta+1)(\alpha+\beta)} = \frac{\alpha}{\left(\alpha+\frac{w_0}{1-w_0}\alpha+1\right)}w_0\,.$$

Συνεπώς, ισχύει ότι

$$\alpha = \frac{w_1(w_0-1)}{{w_0}^2-w_0+w_1},$$

ενώ η (4.1), με τη βοήθεια της παραπάνω γίνεται

$$\beta = -\frac{w_0 w_1}{{w_0}^2-w_0+w_1}.$$

Κεφάλαιο 5

5. Η ΑΠΟΔΟΣΗ ΤΗΣ ΕΜΠΕΙΡΙΚΗΣ ΜΠΕΫΖΙΑΝΗΣ

5.1 Απόδοση *(performance)* της Παραμετρικής Εμπειρικής Μπεϋζιανής

Οι Carlin και Louis (2008), προχώρησαν στον έλεγχο της απόδοσης της παραμετρικής εμπειρικής Μπεϋζιανής μεθόδου. Για να το καταφέρουν αυτό, αξιολόγησαν και σύγκριναν τα αποτελέσματα, τεσσάρων διαφορετικών μεθόδων εκτίμησης.

5.1.1 Μοντέλο *Normal - Normal*

Από τη στιγμή που η μέθοδος προς αξιολόγηση, είναι εκείνη της παραμετρικής εμπειρικής Μπεϋζιανής, η σύγκριση έγινε υπό το πλαίσιο του σύνθετου μοντέλου δειγματοληψίας. Το μοντέλο που χρησιμοποίησαν ήταν εκείνο της Κανονικής/Κανονικής κατανομής, δηλαδή

$$X_i|\theta_i \sim N(\theta_i, 1), \qquad i = 1, \dots, k$$
$$\text{και}$$
$$\theta_i|\mu \sim N(\mu, r^2), \qquad i = 1, \dots, k,$$

όπου για το συγκεκριμένο πείραμα, το $k = 5$.

Με βάση το παραπάνω μοντέλο, στην κλασσική προσέγγιση, φυσικός σημειακός εκτιμητής του θ_i, μέσω της μεθόδου μέγιστης πιθανοφάνειας (βλέπε Παράρτημα Β.1), είναι ο

$$\hat{\theta}_i^F = x_i \ .$$

Η Μπεϋζιανή σημειακή εκτίμηση για το συγκεκριμένο μοντέλο, είναι η εκ των υστέρων μέση τιμή (βλέπε Παράδειγμα 2.2), όπου είναι

$$\hat{\theta}_i^B = \frac{x_i r^2 + \mu}{r^2 + 1}.$$

Ο υπολογισμός του $\hat{\theta}_i^B$, προϋποθέτει την επιλογή των υπερπαραμέτρων του μοντέλου. Η προφανέστερη επιλογή υπερπαραμέτρων για το παραπάνω μοντέλο, είναι $\mu = 0$ και $r^2 = 1$, επομένως ο εκτιμητής γίνεται

$$\hat{\theta}_i^B = \frac{x_i}{2}.$$

Οι παραπάνω σχέσεις φανερώνουν, πως υπό το σύνθετο μοντέλο δειγματοληψίας, στην κλασσική και Μπεϋζιανή προσέγγιση, τα δεδομένα δεν συνδυάζονται.

Για την αποφυγή της επιλογής των υπερπαραμέτρων, οδηγούμαστε στην ιεραρχική Μπεϋζιανή προσέγγιση. Στην προσέγγιση αυτή, οι υπερπαράμετροι ακολουθούν τη δική τους κατανομή. Έτσι εκτιμητής που προκύπτει, αν η υπέρ εκ των προτέρων κατανομή που περιγράφει τις υπερπαραμέτρους είναι η $h(\mu,\, r^2) = 1$ (βλέπε Παράρτημα Β.2), είναι

$$\hat{\theta}_i^{HB} = Z\bar{x} + (1 - Z)x_i\,,$$

με

$$Z = \frac{\kappa - 3}{(\kappa - 1)s^2}\left[1 - g(s^2)\right],$$

και

$$g(s^2) = \frac{(\kappa - 1)s^2}{2}\left[exp\left\{\frac{(\kappa - 1)s^2}{2}\right\} - 1\right]^{-1},$$

όπου $\bar{x} = \frac{1}{k}\sum_{i=1}^{k} x_i$ και $s^2 = \frac{1}{k-1}\sum_{i=1}^{k}(x_i - \bar{x})^2$.

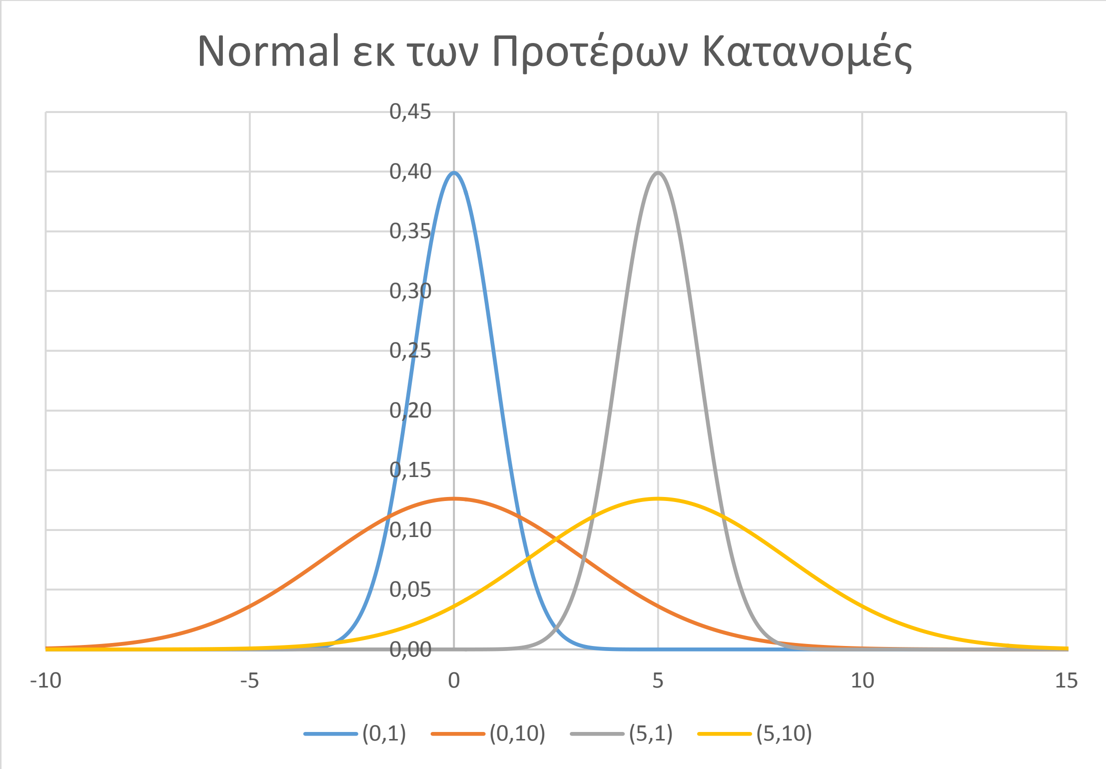
Ο εκτιμητής της παραμέτρου θ_i, αν οι υπερπαράμετροι εκτιμηθούν από τα ίδια τα δεδομένα, δηλαδή ο εκτιμητής που προκύπτει μέσω της παραμετρικής εμπειρικής Μπεϋζιανής προσέγγισης (βλέπε Παράδειγμα 4.2), είναι

$$\hat{\theta}_i^{EB} = \frac{x_i\hat{r}^2 + \hat{\mu}}{\hat{r}^2 + 1},$$

όπου $\hat{\mu} = \bar{x}$ και $\hat{r}^2 = (s^2 - 1)^+$, δηλαδη $\hat{r}^2 = max\{0,\; s^2 - 1\}$.

Για να αξιολογηθούν οι εκτιμητές των παραπάνω προσεγγίσεων, επιλέχθηκαν τιμές για τις υπερπαραμέτρους μ και r^2. Έπειτα παράχθηκε τυχαίο δείγμα $\boldsymbol{\theta}_j = (\theta_{1j}, \dots, \theta_{kj})$ από την εκ των προτέρων κατανομή $u(\theta_i) \sim N(\mu, r^2)$ και ύστερα τυχαίο δείγμα $\boldsymbol{X}_j = (X_{1j}, \dots, X_{kj})$ από την πιθανοφάνεια $f(x_i|\theta_i) \sim N(\theta_i, 1)$, για $j = 1, \dots, N$, μέσω προσομοίωσης.

Οι τιμές των υπερπαραμέτρων (μ, r^2) που επιλέχθηκαν ήταν οι (0,1), (0,10), (5,1) και (5,10), και $N = 10000$ ο αριθμός των επαναλήψεων. Κάτωθι παρατίθεται γράφημα των συναρτήσεων πυκνότητας πιθανότητας των παραπάνω κατανομών.



Γράφημα 5.1: *Normal Εκ των Προτέρων Κατανομές*

Αποτέλεσμα της παραπάνω προσομοίωσης, ήταν η παραγωγή N προσομοιωμένων τιμών για κάθε έναν από τους εκτιμητές, τους οποίους και χρησιμοποίησαν, για να εκτιμήσουν τον κίνδυνο υπό τη συνάρτηση τετραγωνικού σφάλματος. Για παράδειγμα, ο μέσος κίνδυνος ενός εκτιμητή $\hat{\theta}_i^{Est}$, ο οποίος είναι

$$r = E_{\boldsymbol{\theta}} E_{\mathbf{X}|\boldsymbol{\theta}} \left[\frac{1}{k} \sum_{i=1}^{k} \left(\hat{\theta}_i^{Est} - \theta_\iota \right)^2 \right],$$

εκτιμήθηκε από την προσομοιωμένη τιμή

$$\hat{r} = \frac{1}{Nk} \sum_{j=1}^{N} \sum_{i=1}^{k} \left(\hat{\theta}_{ij}^{Est} - \theta_{\iota j} \right)^2 .$$

Η εκτιμώμενη τυπική απόκλιση Monte Carlo είναι

$$\widehat{se}(\hat{r}) = \sqrt{\frac{1}{(Nk)(Nk-1)}\sum_{j=1}^{N}\sum_{i=1}^{k}\left[\left(\hat{\theta}_{ij}^{Est} - \theta_{ij}\right)^2 - \hat{r}\right]^2} .$$

Τα αποτελέσματα της μελέτης, αποτυπώνονται στον Πίνακα 5.1.

Συμπεράσματα

Ο προσομοιωμένος κίνδυνος των εκτιμητών που προκύπτουν από την κλασσική προσέγγιση, είναι σε όλες τις περιπτώσεις κοντά στο 1. Το γεγονός επιβεβαιώνει την ορθότητα της μελέτης, αφού

$$\begin{aligned} \boldsymbol{r}\left(\hat{\theta}^F\right) &= E_\Theta E_{X|\Theta}\left[\frac{1}{k}\sum_{i=1}^{k}(X_i - \theta_i)^2\right] \\ &= E_\Theta\left[\frac{1}{k}\sum_{i=1}^{k}Var_{X_i|\Theta_i}(X_i)\right] \\ &= E_\Theta\left[\frac{1}{k}k\right] = 1 . \end{aligned}$$

Ο εκτιμητής $\hat{\theta}^B$ που προκύπτει από την Μπεϋζιανή προσέγγιση, αποδίδει εντυπωσιακά όταν οι παράμετροι μ και r^2 που επιλέχθηκαν για την Μπεϋζιανή μελέτη είναι ίδιοι με εκείνους του προσομοιωμένου τυχαίου δείγματος. Όταν όμως οι τιμές των παραμέτρων απομακρύνονται από την πραγματική διακύμανση, ή τον πραγματικό μέσο του πληθυσμού, υστερεί σημαντικά σε απόδοση, σε σχέση με τους εκτιμητές των άλλων προσεγγίσεων.

Ο αντίστοιχος εκτιμητής $\hat{\theta}^{EB}$ που προκύπτει από την εμπειρική Μπεϋζιανή προσέγγιση, έχει άκρως ικανοποιητικά αποτελέσματα σε όλες τις περιπτώσεις, και συγχρόνως φαίνεται πως είναι ανώτερος από τον $\hat{\theta}^F$, όταν η r^2 περιορίζεται σε μικρές τιμές, χωρίς να σημαίνει πως στις υπόλοιπες περιπτώσεις δεν αποδίδει ικανοποιητικά.

Για τον εκτιμητή $\hat{\theta}^{HB}$ της ιεραρχικής Μπεϋζιανής προσέγγισης, υπάρχουν ενδείξεις πως αποδίδει καλύτερα από τον $\hat{\theta}^{EB}$, όταν το r^2 είναι μεγάλο, πιθανώς λόγω της επιλογής μιας επίπεδης υπέρ εκ των προτέρων.

Για τους σκοπούς της παρούσας εργασίας, έγινε αναπαραγωγή του πρωτότυπου κώδικα της παραπάνω προσομοίωσης σε προγραμματιστικό περιβάλλον R και παρατίθεται στο Παράρτημα Γ.1.

Προσέγγιση	$\hat{r}$	$\widehat{se}(\hat{r})$
$(\mu, r^2) = (0, 1)$		
Κλασσική	1.005	0.006
Μπεϋζιανή	0.500	0.003
Ιεραρχική Μπεϋζιανή	0.762	0.005
Εμπειρική Μπεϋζιανή	0.705	0.005
$(\mu, r^2) = (0, 10)$		
Κλασσική	0.994	0.006
Μπεϋζιανή	2.736	0.017
Ιεραρχική Μπεϋζιανή	0.951	0.006
Εμπειρική Μπεϋζιανή	0.976	0.006
$(\mu, r^2) = (5, 1)$		
Κλασσική	1.000	0.006
Μπεϋζιανή	6.774	0.016
Ιεραρχική Μπεϋζιανή	0.758	0.005
Εμπειρική Μπεϋζιανή	0.704	0.005
$(\mu, r^2) = (5, 10)$		
Κλασσική	0.987	0.006
Μπεϋζιανή	9.053	0.041
Ιεραρχική Μπεϋζιανή	0.945	0.006
Εμπειρική Μπεϋζιανή	0.975	0.006

Πίνακας 5.1 *Προσομοιωμένος κίνδυνος Normal - Normal.*

Βασιζόμενοι στην παραπάνω μελέτη, για τους σκοπούς της παρούσας εργασίας, θα επαναληφθεί η ίδια διαδικασία, συγκρίνοντας την απόδοση της εμπειρικής παραμετρικής Μπεϋζιανής προσέγγισης, με εκείνη της κλασσικής και της Μπεϋζιανής, στα μοντέλα *Poisson/Gamma* και *Binomial/Beta*.

5.1.2 Μοντέλο *Poisson - Gamma*

Αρχικά θα μελετηθεί το μοντέλο Poisson/Gamma, δηλαδή

$$X_i|\theta_i \sim Poisson(\theta_i), \qquad i = 1, \dots, k$$

$$και$$

$$\theta_i|a, \beta \sim Gamma(a, \beta), \qquad i = 1, \dots, k,$$

όπου για το συγκεκριμένο πείραμα, το $k = 10$.

Στόχος είναι και πάλι, η σύγκριση του προσομοιωμένου κινδύνου $\boldsymbol{r}(\hat{\theta}^F)$ που προκύπτει από την εμπειρική Μπεϋζιανή εκτίμηση, με εκείνον των άλλων εκτιμήσεων.

Υπό το παραπάνω μοντέλο, στην κλασσική προσέγγιση, φυσικός σημειακός εκτιμητής του θ_i, μέσω της μεθόδου μέγιστης πιθανοφάνειας (βλέπε Παράρτημα Β.3), είναι ο

$$\hat{\theta}_i^F = x_i \ .$$

Η Μπεϋζιανή σημειακή εκτίμηση για το συγκεκριμένο μοντέλο είναι (βλέπε Παράδειγμα 2.1),

$$\hat{\theta}_i^B = \frac{\alpha\beta}{\beta + 1} + \frac{\beta}{\beta + 1} x_i \ .$$

Οι υπερπαράμετροι που θα επιλεχθούν είναι οι $\alpha = 1$ και $\beta = 1$, επομένως ο εκτιμητής γίνεται

$$\hat{\theta}_i^B = \frac{x_i + 1}{2} .$$

Ο εκτιμητής που προκύπτει μέσω της παραμετρικής εμπειρικής Μπεϋζιανής προσέγγισης (βλέπε Παράδειγμα 4.1), είναι ο

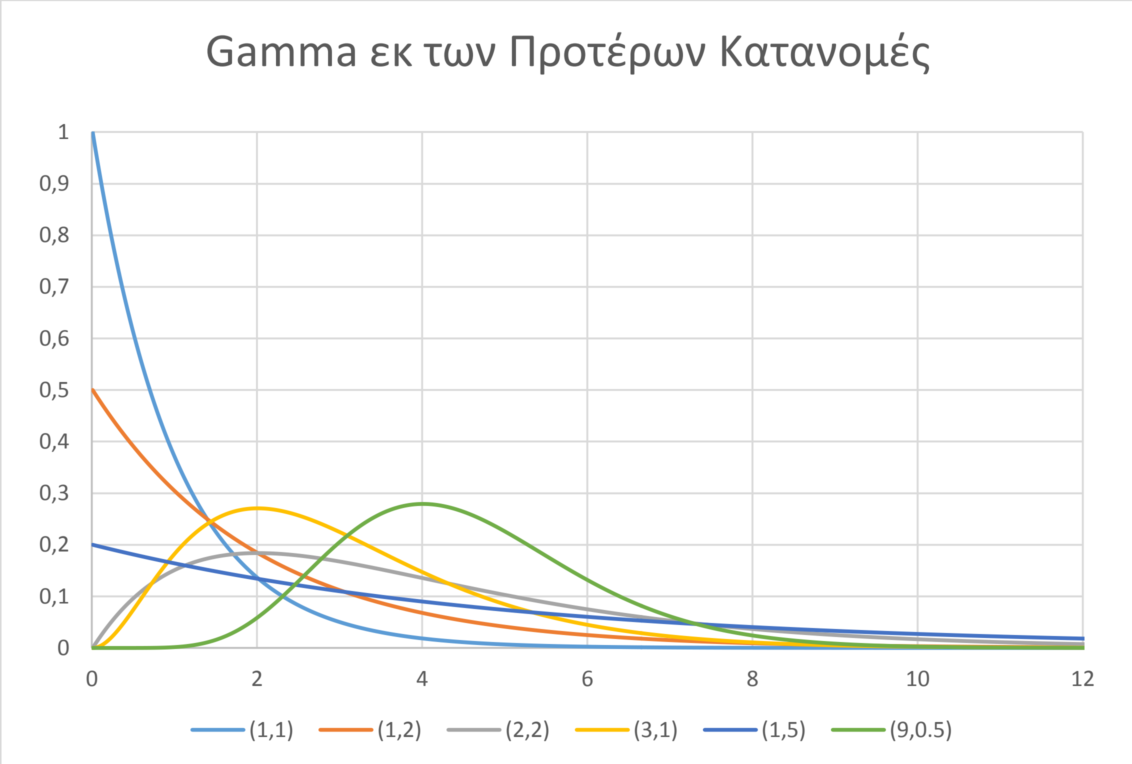
$$\hat{\theta}_i^{EB} = \frac{\hat{\alpha}\hat{\beta}}{\hat{\beta} + 1} + \frac{\hat{\beta}}{\hat{\beta} + 1} x_i \ ,$$

όπου

$$\hat{\beta} = \frac{m_2 - {m_1}^2}{m_1} - 1 \quad και \quad \hat{\alpha} = \frac{m_1}{\hat{\beta}} .$$

Η διαδικασία της προσομοίωσης που θα ακολουθηθεί, είναι ίδια με αυτή του προηγουμένου πειράματος. Αρχικά θα παραχθεί τυχαίο δείγμα $\boldsymbol{\theta}_j = (\theta_{1j}, \dots, \theta_{kj})$ από την εκ των προτέρων κατανομή $u(\theta_i) \sim Gamma(a, \beta)$ και έπειτα τυχαίο δείγμα $\boldsymbol{X}_j = (X_{1j}, \dots, X_{kj})$ από την πιθανοφάνεια $f(x_i|\theta_i) \sim Poisson(\theta_i)$, για $j = 1, \dots, N$, μέσω προσομοίωσης.

Οι τιμές των υπερπαραμέτρων (α, β) που επιλέχθηκαν ήταν οι (1,1), (1,2), (2,2), (3,1), (5,1) και (9,0.5), και $N = 10000$ ο αριθμός των επαναλήψεων. Κάτωθι παρατίθεται γράφημα των συναρτήσεων πυκνότητας πιθανότητας των παραπάνω κατανομών.



Γράφημα 5.2: *Gamma Εκ των Προτέρων Κατανομές*

Με τη χρήση των N προσομοιωμένων τιμών που παρήχθησαν, υπολογίστηκε προσομοιωμένος κινδύνους $\boldsymbol{r}(\hat{\theta}^F)$ και η εκτιμώμενη τυπική απόκλιση Monte Carlo, υπό τη συνάρτηση τετραγωνικού σφάλματος, και για τις τρεις μεθόδους.

Τα αποτελέσματα της μελέτης, αποτυπώνονται στον Πίνακα 5.2

Προσέγγιση	$\hat{r}$	$\widehat{se}(\hat{r})$
	$(\alpha, \beta) = (1, 1)$	
Κλασσική	0.998	0.011
Μπεϋζιανή	0.499	0.006
Εμπειρική Μπεϋζιανή	0.610	0.007
	$(\alpha, \beta) = (1, 2)$	
Κλασσική	1.995	0.021
Μπεϋζιανή	1.753	0.022
Εμπειρική Μπεϋζιανή	1.492	0.017
	$(\alpha, \beta) = (2, 2)$	
Κλασσική	3.987	0.035
Μπεϋζιανή	5.274	0.046
Εμπειρική Μπεϋζιανή	2.943	0.026
	$(\alpha, \beta) = (3, 1)$	
Κλασσική	3.005	0.017
Μπεϋζιανή	2.516	0.014
Εμπειρική Μπεϋζιανή	2.055	0.012
	$(\alpha, \beta) = (1, 5)$	
Κλασσική	4.977	0.036
Μπεϋζιανή	11.466	0.094
Εμπειρική Μπεϋζιανή	4.545	0.034
	$(\alpha, \beta) = (9, 0.5)$	
Κλασσική	4.519	0.032
Μπεϋζιανή	4.743	0.025
Εμπειρική Μπεϋζιανή	1.989	0.014

Πίνακας 5.2 *Προσομοιωμένος κίνδυνος Poisson - Gamma.*

Συμπεράσματα

Ο προσομοιωμένος κίνδυνος των εκτιμητών που προκύπτουν από την κλασσική προσέγγιση, είναι σε όλες τις περιπτώσεις κοντά στο γινόμενο των υπερπαραμέτρων. Το γεγονός αυτό είναι λογικό αφού

$$
\begin{aligned}
\boldsymbol{r}(\hat{\theta}^F) &= E_{\Theta} E_{X|\Theta}\left[\frac{1}{k}\sum_{i=1}^{k}(X_i-\theta_i)^2\right] \\
&= E_{\Theta}\left[\frac{1}{k}\sum_{i=1}^{k} Var_{X_i|\Theta_i}(X_i)\right] \\
&= E_{\Theta}[\theta_i] = \alpha\beta\ .
\end{aligned}
$$

Ο εκτιμητής $\hat{\theta}^B$ που προκύπτει από την Μπεϋζιανή προσέγγιση, αποδίδει πολύ καλά, όταν οι παράμετροι α και β που επιλέχθηκαν για την Μπεϋζιανή μελέτη, είναι ίδιοι με εκείνους του προσομοιωμένου τυχαίου δείγματος. Όταν όμως οι τιμές των παραμέτρων *α, β,* απομακρύνονται από τις πραγματικές τους τιμές, υστερεί σημαντικά σε απόδοση, σε σχέση με τους εκτιμητές, των υπόλοιπων προσεγγίσεων. Φαίνεται επίσης πως σημαντικότερη επίδραση στα αποτελέσματα έχει η μεταβολή της παραμέτρου *β*.

Ο αντίστοιχος εκτιμητής $\hat{\theta}^{EB}$ που προκύπτει από την εμπειρική Μπεϋζιανή προσέγγιση, έχει άκρως ικανοποιητικά αποτελέσματα σε όλες τις περιπτώσεις και συγχρόνως φαίνεται πως είναι ανώτερος από τον $\hat{\theta}^F$, κυρίως όταν η παράμετρος α απομακρύνεται από την μονάδα.

Ο κώδικας σε R της παραπάνω προσομοίωσης, παρατίθεται στο Παράρτημα Γ.2.

Σημείωση

Όταν ισχύει $x_{1j} = x_{2j} = \cdots = x_{kj}$, ο εκτιμητής $\hat{\theta}^{EB}$ δεν μπορεί να υπολογιστεί. Στις περιπτώσεις αυτές, τα δεδομένα που προέκυψαν από την προσομοίωση, δεν λήφθηκαν υπόψιν για καμία από τις μεθόδους εκτίμησης. Αυτό το φαινόμενο, ήταν και ο λόγος που επιλέχθηκε το $k = 10$, αφού η απώλεια της πληροφορίας σε αυτήν τη περίπτωση, κρίνεται αμελητέα.

5.1.3 Μοντέλο *Beta - Binomial*

Η τελευταία μελέτη που θα πραγματοποιηθεί, αφορά το μοντέλο

$$X_i|\theta_i \sim Binomial(n,\theta_i), \qquad i=1,\dots,k$$
$$και$$
$$\theta_i|a,\beta \sim Beta(a,\beta), \qquad i=1,\dots,k,$$

όπου για το συγκεκριμένο πείραμα, το $k = 20 \; και \; n = 5$.

Υπό την κλασσική προσέγγιση, για το παραπάνω μοντέλο, κάνοντας ξανά χρήση της μεθόδου μέγιστης πιθανοφάνειας, καταλήγουμε ότι φυσικός σημειακός εκτιμητής του θ_i (βλέπε Παράρτημα Β.4), είναι ο

$$\hat{\theta}_i^F = \frac{x_i}{n},$$

ενώ η Μπεϋζιανή σημειακή εκτίμηση για το συγκεκριμένο μοντέλο είναι (βλέπε Παράδειγμα 2.3),

$$\hat{\theta}_i^B = \frac{a + x_i}{a+\beta+n}.$$

Επιλέγοντας για υπερπαραμέτρους τα $\alpha = 1$ και $\beta = 1$, καταλήγουμε στον εκτιμητή

$$\hat{\theta}_i^B = \frac{x_i + 1}{2+n}.$$

Ο εκτιμητής που προκύπτει μέσω της παραμετρικής εμπειρικής Μπεϋζιανής προσέγγισης (βλέπε Παράδειγμα 4.3), είναι ο

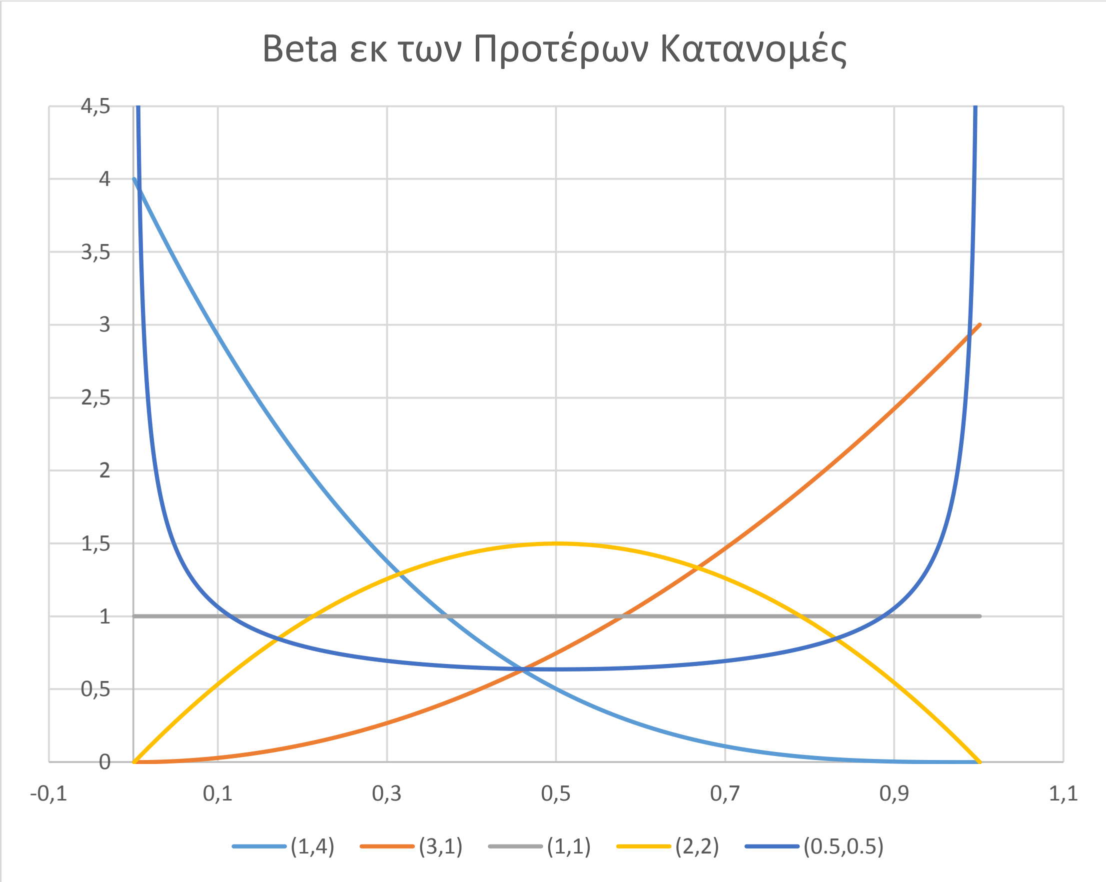
$$\hat{\theta}_i^{EB} = \frac{\hat{\alpha} + x_i}{\hat{\alpha}+\hat{\beta}+n},$$

όπου

$$\hat{\alpha} = \frac{m_1(m_2 - n\, m_1)}{(n-1){m_1}^2 + n(m_2 - m_1)} \quad και \quad \hat{\beta} = \frac{(n-m_1)(m_2 - n\, m_1)}{(n-1){m_1}^2 + n(m_2 - m_1)}.$$

Και πάλι θα γίνει παραγωγή τυχαίου δείγματος $\boldsymbol{\theta}_j = (\theta_{1j}, \dots, \theta_{kj})$ από την εκ των προτέρων κατανομή $u(\theta_i) \sim Beta(a, \beta)$ και έπειτα τυχαίου δείγματος $\boldsymbol{X}_j = (X_{1j}, \dots, X_{kj})$ από την πιθανοφάνεια $f(x_i|\theta_i) \sim Binomial(n, \theta_i)$, για $j = 1, \dots, N$, μέσω προσομοίωσης.

Οι τιμές των υπερπαραμέτρων (α, β) που επιλέχθηκαν ήταν οι (0.5,0.5), (1,1), (2,2), (3,1) και (1,4), και $N = 10000$ ο αριθμός των επαναλήψεων. Κάτωθι παρατίθεται γράφημα των συναρτήσεων πυκνότητας πιθανότητας των παραπάνω κατανομών.



Γράφημα 5.3: *Beta Εκ των Προτέρων Κατανομές*

Με τη χρήση των N προσομοιωμένων τιμών που παρήχθησαν, υπολογίστηκε προσομοιωμένος κινδύνους $\boldsymbol{r}(\hat{\theta}^F)$ και η εκτιμώμενη τυπική απόκλιση Monte Carlo, υπό την συνάρτηση τετραγωνικού σφάλματος, και για τις τρεις μεθόδους.

Τα αποτελέσματα της μελέτης για το παραπάνω μοντέλο, αποτυπώνονται στον Πίνακα 5.3

Προσέγγιση	$\hat{r}$	$\widehat{se}(\hat{r})$
$(\alpha, \beta) = (0.5, 0.5)$		
Κλασσική	0.025	0.00010
Μπεϋζιανή	0.023	0.00007
Εμπειρική Μπεϋζιανή	0.021	0.00008
$(\alpha, \beta) = (1, 1)$		
Κλασσική	0.033	0.00011
Μπεϋζιανή	0.024	0.00007
Εμπειρική Μπεϋζιανή	0.026	0.00008
$(\alpha, \beta) = (2, 2)$		
Κλασσική	0.040	0.00012
Μπεϋζιανή	0.024	0.00008
Εμπειρική Μπεϋζιανή	0.025	0.00007
$(\alpha, \beta) = (3,1)$		
Κλασσική	0.030	0.00011
Μπεϋζιανή	0.024	0.00008
Εμπειρική Μπεϋζιανή	0.019	0.00007
$(\alpha, \beta) = (1, 4)$		
Κλασσική	0.027	0.00010
Μπεϋζιανή	0.023	0.00008
Εμπειρική Μπεϋζιανή	0.016	0.00006

Πίνακας 5.3 *Προσομοιωμένος κίνδυνος Beta - Binomial.*

Συμπεράσματα

Ο προσομοιωμένος κίνδυνος των εκτιμητών που προκύπτουν από την κλασσική προσέγγιση, πρέπει σε όλες τις περιπτώσεις να συμβαδίζει με τις τιμές που προκύπτουν από την παρακάτω σχέση

$$\boldsymbol{r}(\hat{\theta}^F) = E_\Theta E_{X|\Theta}\left[\frac{1}{k}\sum_{i=1}^{k}\left(\frac{X_i}{n} - \theta_\iota\right)^2\right] = E_\Theta\left[\frac{1}{k}\sum_{i=1}^{k} Var_{X_i|\Theta_\iota}\left(\frac{X_i}{n}\right)\right]$$

$$= E_\Theta\left[\frac{1}{k\,n^2}\sum_{i=1}^{k} Var_{X_i|\Theta_\iota}(X_i)\right] = E_\Theta\left[\frac{1}{k\,n^2} k\,n\,\theta_i(1-\theta_i)\right]$$

$$= \frac{1}{n}E_\Theta\left(\theta_i - {\theta_i}^2\right) = \frac{1}{n}E_\Theta(\theta_i) - \frac{1}{n}E_\Theta\left({\theta_i}^2\right)$$

$$= \frac{\alpha}{n(\alpha+\beta)} - \frac{\alpha\beta + \alpha^2(\alpha+\beta+1)}{n(\alpha+\beta)^2(\alpha+\beta+1)}$$

$$= \frac{\alpha\beta}{n(\alpha+\beta)(\alpha+\beta+1)}.$$

Ο εκτιμητής $\hat{\theta}^B$ που προκύπτει από την Μπεϋζιανή προσέγγιση, αλλά και ο $\hat{\theta}^{EB}$ που προκύπτει από την εμπειρική Μπεϋζιανή προσέγγιση, φαίνεται πως για τις παραμέτρους που επιλέχθηκαν, αποδίδουν και οι δυο καλύτερα από τον $\hat{\theta}^F$.

Προχωρώντας στην μεταξύ τους σύγκριση, είναι και πάλι εμφανές πως ο $\hat{\theta}^B$, υπερτερεί έναντι του $\hat{\theta}^{EB}$, μόνο στην περίπτωση όπου οι παράμετροι α και β που επιλέχθηκαν για την Μπεϋζιανή μελέτη είναι ίδιες με εκείνες του προσομοιωμένου τυχαίου δείγματος. Στις υπόλοιπες περιπτώσεις, ο $\hat{\theta}^{EB}$ υπερτερεί σημαντικά.

Ο κώδικας σε R της παραπάνω προσομοίωσης, παρατίθεται στο Παράρτημα Γ.3.

Σημείωση

Σε κάποιες περιπτώσεις, κατά την εμπειρική Μπεϋζιανή προσέγγιση, υπό το παραπάνω μοντέλο, καταλήγουμε σε $\hat{\alpha}, \hat{\beta} < 0$. Στις περιπτώσεις αυτές, υποτέθηκε πως ισχύει $\hat{\alpha}, \hat{\beta} \cong 0$.

Επίσης σε πολύ λίγες περιπτώσεις, ο παρονομαστής των $\hat{\alpha}, \hat{\beta}$, τυχαίνει να μηδενίζεται. Στις περιπτώσεις αυτές τα δεδομένα που προέκυψαν από την προσομοίωση, δεν λήφθηκαν υπόψιν για καμία από τις μεθόδους εκτίμησης.

5.2 Σύγκριση απόδοσης Μη Παραμετρικής και Παραμετρικής Εμπειρικής Μπεϋζιανής

Αξιοποιώντας την παραπάνω μελέτη, θα προχωρήσουμε στη σύγκριση των αποτελεσμάτων που προκύπτουν από την παραμετρική και την μη παραμετρική εμπειρική Μπεϋζιανή προσέγγιση. Η μελέτη που θα πραγματοποιηθεί αφορά το μοντέλο Poisson/Gamma, που χρησιμοποιήθηκε προηγουμένως.

Ο εκτιμητής που προκύπτει μέσω της παραμετρικής εμπειρικής Μπεϋζιανής προσέγγισης, είναι εκείνος που παρουσιάστηκε στην ενότητα 5.1.2, ενώ ο εκτιμητής που προκύπτει μέσω της μη παραμετρικής εμπειρικής Μπεϋζιανής προσέγγισης (βλέπε Παράδειγμα 3.1), είναι ο

$$\hat{\theta}_i^{NPEB} = (x_i + 1)\frac{g_n(x_i + 1)}{g_n(x_i)},$$

όπου η περιθώρια κατανομή $g_n(x_i)$, θα εκτιμηθεί από την εμπειρική της συχνότητα

$$g_n(x_i) = \frac{\#\, των\ x\ τα\ οποια\ ισουνται\ με\ x_i}{n}.$$

Η διαδικασία της προσομοίωσης που θα ακολουθηθεί, είναι ίδια με αυτή της ενότητας 5.1.2. Στόχος της συγκεκριμένης μελέτης, είναι ο έλεγχος της επίδρασης που έχει η μεταβολή του μεγέθους, του παραγόμενου δείγματος k, πάνω στην απόδοση των δυο προσεγγίσεων. Πιο συγκεκριμένα το k θα πάρει τις τιμές 500, 1000 και 5000, ενώ ο αριθμός επαναλήψεων του πειράματος είναι $N = 100$.

Οι τιμές των υπερπαραμέτρων (α, β) που επιλέχθηκαν, είναι και πάλι οι (1,1), (1,2), (2,2), (3,1), (5,1) και (9,0.5).

Συμπεράσματα

Τα αποτελέσματα της μελέτης, αποτυπώνονται στον Πίνακα 5.4. Από τη μελέτη μπορούν να εξαχθούν τα εξής συμπεράσματα. Ο εκτιμητής $\hat{\theta}^{EB}$ φαίνεται πως έχει βελτιωμένη απόδοση σε σχέση με αυτή της ενότητας 5.2. Αυτό οφείλεται στην αύξηση του αριθμού του παραγόμενου δείγματος k, σε σχέση με το προηγούμενο πείραμα. Βέβαια τα αποτελέσματα της παρούσας μελέτης, δείχνουν μια σταθερότητα της απόδοσης. Αυτό προδίδει, πως η αύξηση του δείγματος πέρα από κάποιο μέγεθος, δεν επιφέρει βελτίωση στην απόδοση του εκτιμητή.

Ο εκτιμητής $\hat{\theta}^{NPEB}$ φαίνεται να έχει θεαματική βελτίωση όσο το δείγμα αυξάνεται, όμως ακόμα και με δείγμα $k = 5000$, δεν προσεγγίζει την απόδοση του $\hat{\theta}^{EB}$.

	$\hat{r}$		
Προσέγγιση	$k = 500$	$k = 1000$	$k = 5000$
	$(\alpha, \beta) = (1, 1)$		
Παραμετρική Εμπειρική	0.505	0.501	0.501
Μη Παραμετρική Εμπειρική	0.999	0.791	0.602
	$(\alpha, \beta) = (1, 2)$		
Παραμετρική Εμπειρική	1.347	1.322	1.330
Μη Παραμετρική Εμπειρική	3.565	2.886	1.939
	$(\alpha, \beta) = (2, 2)$		
Παραμετρική Εμπειρική	2.711	2.691	2.661
Μη Παραμετρική Εμπειρική	8.779	6.446	3.946
	$(\alpha, \beta) = (3, 1)$		
Παραμετρική Εμπειρική	1.527	1.504	1.504
Μη Παραμετρική Εμπειρική	3.405	2.752	1.876
	$(\alpha, \beta) = (1, 5)$		
Παραμετρική Εμπειρική	4.191	4.178	4.211
Μη Παραμετρική Εμπειρική	24.572	19.676	11.427
	$(\alpha, \beta) = (9, 0.5)$		
Παραμετρική Εμπειρική	1.517	1.523	1.503
Μη Παραμετρική Εμπειρική	4.046	3.087	1.934

Πίνακας 5.4 *Προσομοιωμένος κίνδυνος εμπειρικής Μπεϋζιανής.*

Κεφάλαιο 6

6. ΘΕΩΡΙΑ ΑΞΙΟΠΙΣΤΙΑΣ ΚΑΙ ΜΠΕΫΖΙΑΝΗ ΣΤΑΤΙΣΤΙΚΗ

6.1 Θεωρία Αξιοπιστίας

Τα μοντέλα αξιοπιστίας, είναι βασικά αναλογιστικά εργαλεία, τα οποία εξυπηρετούν στη δίκαιη κατανομή των ασφαλίστρων, ανάμεσα σε ετερογενείς ομάδες ασφαλισμένων.

Σύμφωνα με τους Miller και Hickman (1975), η θεωρία αξιοπιστίας μπορεί να θεωρηθεί ως η ανάπτυξη ενός συστήματος βαρών, τα οποία είναι συνάρτηση του μεγέθους της διαθέσιμης εμπειρίας της ασφαλιστικής εταιρείας. Στόχος της διαδικασίας αυτής, είναι η εκτίμηση του μελλοντικού κόστους των απαιτήσεων, ως σταθμισμένος μέσος των πραγματικών απαιτήσεων μιας ομάδας υπό εξέταση, και των αναμενόμενων απαιτήσεων με βάση την εκ των προτέρων γνώση.

Ουσιαστικά είναι μία τεχνική τιμολόγησης ασφαλίστρων, σε ομάδες ασφαλιστήριων συμβολαίων ετερογενών χαρτοφυλακίων, στην περίπτωση που υπάρχει περιορισμένη εμπειρία εξέλιξης ζημιών για κάθε ομάδα, αλλά απεριόριστη εμπειρία εξέλιξης ζημιών σε μεγαλύτερη ομάδα ασφαλιστήριων συμβολαίων (χαρτοφυλάκιο), που έχουν όμοια χαρακτηριστικά (βλέπε Πιτσέλη, 2016).

Μαθηματικά η παραπάνω πρόταση μπορεί να αποτυπωθεί ως

$$C = ZR + (1 - Z)H, \tag{6.1}$$

όπου με R, συμβολίζεται ο μέσος των παρατηρήσεων του τρέχοντος δείγματος, με H, η εκ των προτέρων μέση τιμή, με Z, ο συντελεστής αξιοπιστίας, όπου $0 \leq Z \leq 1$, και με C, η σταθμισμένη εκτίμηση.

Σε αυτό το κεφάλαιο, θα μελετηθεί η θεωρία αξιοπιστίας χαρτοφυλακίου και θα διερευνηθεί σε ποιες περιπτώσεις τα αποτελέσματά της συμπίπτουν, με εκείνα της Μπεϋζιανής και εμπειρικής Μπεϋζιανής προσέγγισης.

6.2 Το Μοντέλο του Bühlmann

Ας υποτεθεί το παρακάτω μοντέλο. Έστω το j-οστο ασφαλιστήριο συμβόλαιο που περιγράφεται από το διάνυσμα $(\theta_j, \boldsymbol{X}_j) = (\theta_j, X_{1j}, X_{2j}, \ldots)$, όπου τα διανύσματα $(\theta_j, \boldsymbol{X}_j)$ είναι ανεξάρτητα και ισόνομα. Οι δεσμευμένες μεταβλητές $X_{ij}|\theta_j$ και οι παράμετροι κινδύνου θ_j, θεωρούνται πως είναι ανεξάρτητα και ισόνομα κατανεμημένες. Επίσης οι παράμετροι κινδύνου θ_j, υποθέτεται πως είναι ομογενείς στον χρόνο. Υπό τις παραπάνω υποθέσεις μπορεί να εφαρμοστεί το μοντέλο του Bühlmann (1967).

Γενικά οι παράμετροι κινδύνου θ_j, περιγράφουν τα χαρακτηριστικά του κινδύνου τα οποία είναι δύσκολο να παρατηρηθούν (π.χ. ικανότητα οδηγών) ή δεν επιτρέπεται να παρατηρηθούν (π.χ. απόρρητο προσωπικών δεδομένων), τα οποία δημιουργούν ετερογένεια στο χαρτοφυλάκιο. Η εφαρμογή των τεχνικών αξιοπιστίας γίνεται με το σκεπτικό πως οι παράμετροι κινδύνου των ομάδων διαφέρουν μεταξύ τους.

Οι παράμετροι κινδύνου θεωρούνται τυχαίες μεταβλητές, οι οποίες προέρχονται από μια κοινή εκ των προτέρων κατανομή, η οποία ονομάζεται συνάρτηση δόμησης. Υπό αυτή την έννοια τα ασφαλιστήρια συμβόλαια θεωρούνται πως είναι όμοια μεταξύ τους, αλλά διαφέρουν στο γεγονός ότι τα θ_j έχουν διαφορετικές τιμές μεταξύ τους.

Για την ευκολότερη κατανόηση του παραπάνω μοντέλου, θα παρουσιαστεί η περίπτωση της μελέτης ενός ασφαλιστηρίου συμβολαίου, το οποίο περιγράφεται από το διάνυσμα $(\theta, \boldsymbol{X}) = (\theta, X_1, X_2, \ldots)$. Το μοντέλο του Bühlmann με ένα ασφαλιστήριο συμβόλαιο δεν έχει πρακτική εφαρμογή, γιατί εμπεριέχει παραμέτρους που δεν μπορούν να εκτιμηθούν από τις παρατηρούμενες τυχαίες μεταβλητές, του ίδιου του μοντέλου. Η μεθοδολογία για περισσότερα ασφαλιστήρια συμβόλαια όμως, βασίζεται σε αυτήν του ενός ασφαλιστηρίου συμβολαίου. Για αυτόν τον λόγο κάτωθι παρατίθεται η μεθοδολογία του μοντέλου του Bühlmann για ένα ασφαλιστήριο συμβόλαιο.

Με βάση το παραπάνω πλαίσιο, ο Bühlmann πρότεινε, ο συντελεστής αξιοπιστίας της σχέσης (6.1) να είναι της μορφής

$$Z = \frac{n}{n+k},$$

με $0 \le Z \le 1$, όπου με n συμβολίζεται ο αριθμός των δοκιμών ή των μονάδων έκθεσης στον κίνδυνο, και

$$k = \frac{E_\theta\big(Var_X(X|\theta)\big)}{Var_\theta\big(E_X(X|\theta)\big)}.$$

Ο αριθμητής υποδηλώνει την αναμενόμενη τιμή της διακύμανσης της διαδικασίας *(Expected value of the process variance),* ενώ ο παρονομαστής είναι η διασπορά της υποθετικής μέσης τιμής *(Variance of hypothetical means).*

Πιο αναλυτικά, με τον όρο υποθετική μέση τιμή, αναφερόμαστε στην μέση συχνότητα, μέση σφοδρότητα ή μέσο αθροιστικό ποσό των απαιτήσεων (καθαρό ασφάλιστρο), από έναν ατομικό αριθμό συνδυασμών των χαρακτηριστικών του κινδύνου.

Επίσης με τον όρο διακύμανση της διαδικασίας, μπορεί να αναφερόμαστε είτε στη διακύμανση της συχνότητας, είτε της σφοδρότητας, είτε του ποσού των αθροιστικών απαιτήσεων από έναν ατομικό αριθμό συνδυασμών των χαρακτηριστικών του κινδύνου. Ο όρος διαδικασία, αναφέρεται στη διαδικασία παραγωγής αριθμού απαιτήσεων ή ποσού απαιτήσεων (βλέπε Herzog, 1989).

Όταν ο όρος k είναι πολύ μεγάλος, ο συντελεστής αξιοπιστίας του Bühlmann τείνει στον μηδενισμό, δηλαδή $Z \cong 0$, συμπεραίνοντας πως το ατομικό ασφάλιστρο τείνει να ταυτιστεί με το ολικό ασφάλιστρο. Σε αντίθετη περίπτωση όταν το k μηδενίζεται, ο συντελεστής αξιοπιστίας του Bühlmann είναι $Z = 1$, πράγμα που σημαίνει πως το ατομικό ασφάλιστρο τιμολογείται μόνο με βάση τη δική του εμπειρία εξέλιξης ζημιών.

Η αξιοπιστία του Bühlmann μπορεί να εφαρμοστεί υπό ένα μοντέλο εντελώς απαλλαγμένο από κατανομές *(distribution free)*, όμως στην παρούσα εργασία θα παρουσιαστεί η παραμετρική του εκδοχή, αφού στόχος είναι η ανάδειξη των ομοιοτήτων της θεωρίας αξιοπιστίας χαρτοφυλακίου, με την Μπεϋζιανή και εμπειρική Μπεϋζιανή προσέγγιση (βλέπε Πιτσέλη, 2016).

Θεώρημα 6.1

Υπό τις παραπάνω υποθέσεις, ο εκτιμητής C, είναι η βέλτιστη γραμμική προσέγγιση της Μπεϋζιανής εκτίμησης. Αυτό συμβαίνει επειδή το σταθμισμένο άθροισμα του τετραγωνικού σφάλματος, μεταξύ της γραμμικής προσέγγισης και της Μπεϋζιανής εκτίμησης, ελαχιστοποιείται. Απόδειξη αυτής της πρότασης ακολουθεί κάτωθι (βλέπε Herzog, 1989).

Απόδειξη 6.1

Ας υποτεθεί η περίπτωση, στην οποία εξετάζεται ένα χαρτοφυλάκιο ασφαλιστήριων συμβολαίων, το οποίο περιγράφεται από μια παράμετρο κινδύνου θ, και $X_1, X_2, \dots, X_n$ τυχαίες μεταβλητές οι οποίες εκφράζουν τις αθροιστικές απώλειες των ετών $1, 2, \dots, n$ αντίστοιχα.

Στόχος είναι να βρεθεί ο βέλτιστος γραμμικός εκτιμητής της $E_{\theta|X}[\mu(\theta|\boldsymbol{X})]$, δηλαδή να προσδιοριστούν οι σταθερές *α, β* οι οποίες ελαχιστοποιούν την παρακάτω σχέση,

$$Q = E_X\left\{\left(E_{\theta|X}[\mu(\theta|\boldsymbol{X})] - a - b\bar{X}\right)^2\right\}. \qquad (6.2)$$

Έστω ότι

$$E_{X|\theta}[X] = \mu(\theta)\,,$$

άρα ισχύει

$$E_{X|\theta}[\bar{X}] = \mu(\theta), \qquad E_X\{E_{\theta|X}[\mu(\theta|\boldsymbol{X})]\} = E_\theta[\mu(\theta)]$$

και

$$E_X[\bar{X}] = E_\theta\{E_{X|\theta}[\bar{X}]\} = E_\theta[\mu(\theta)]\,.$$

Παραγωγίζοντας τη σχέση (6.2) ως προς α λαμβάνουμε

$$\begin{aligned}\frac{\partial E_X\{Q\}}{\partial a} &= -2\, E_X\{(E_{\theta|X}[\mu(\theta|\boldsymbol{X})] - a - b\bar{X})\} \\ &= -2\, E_X\{E_{\theta|X}[\mu(\theta|\boldsymbol{X})]\} -2\, E_X\{-a - b\bar{X}\} \\ &= -2\, E_X\{E_{\theta|X}[\mu(\theta|\boldsymbol{X})]\} + 2a + 2b(E_X\{\bar{X}\})\,,\end{aligned}$$

όπου η παραπάνω μηδενίζεται όταν

$$\begin{aligned}\hat{\alpha} &= E_X\{E_{\theta|X}[\mu(\theta|\boldsymbol{X})]\} - b\, E_X\{\bar{X}\} \\ &= E_\theta[\mu(\theta)] - bE_\theta[\mu(\theta)] \\ &= (1-b)E_\theta[\mu(\theta)]\,.\end{aligned}$$

Με τη βοήθεια της παραπάνω, η (6.2) γίνεται

$$\begin{aligned}Q &= E_X\left\{(E_{\theta|X}[\mu(\theta|\boldsymbol{X})] - a - b\bar{X})^2\right\} \\ &= E_X\left\{(E_{\theta|X}[\mu(\theta|\boldsymbol{X})] - (1-b)E_\theta[\mu(\theta)] - b\bar{X})^2\right\} \\ &= E_X\left\{\left(E_{\theta|X}[\mu(\theta|\boldsymbol{X})] - E_\theta[\mu(\theta)] - b(\bar{X} - E_\theta[\mu(\theta)])\right)^2\right\}.\end{aligned}$$

Παραγωγίζοντας τη παραπάνω σχέση ως προς b λαμβάνουμε

$$\begin{aligned}\frac{\partial E_X\{Q\}}{\partial b} &= \\ &= 2E_X\left\{\left(E_{\theta|X}[\mu(\theta|\boldsymbol{X})] - E_\theta[\mu(\theta)] - b(\bar{X} - E_\theta[\mu(\theta)])\right)(E_\theta[\mu(\theta)] - \bar{X})\right\} \\ &= 2E_X\{(E_{\theta|X}[\mu(\theta|\boldsymbol{X})] - E_\theta[\mu(\theta)])(E_\theta[\mu(\theta)] - \bar{X})\} - 2b\, E_X\{(\bar{X} - E_\theta[\mu(\theta)])^2\}\,,\end{aligned}$$

όπου η παραπάνω μηδενίζεται όταν

$$\hat{b} = \frac{E_X\{(E_{\theta|X}[\mu(\theta|\boldsymbol{X})] - E_\theta[\mu(\theta)])(\bar{X} - E_\theta[\mu(\theta)])\}}{E_X\{\bar{X} - E_\theta[\mu(\theta)]\}^2} = \frac{A}{B}.$$

Όμως ο αριθμητής της παραπάνω γίνεται

$$\begin{aligned}
A &= E_X\{(E_{\theta|X}[\mu(\theta|\boldsymbol{X})] - E_\theta[\mu(\theta)])(\bar{X} - E_\theta[\mu(\theta)])\} \\
&= E_X\{(E_{\theta|X}[\mu(\theta|\boldsymbol{X})] - E_\theta[\mu(\theta)])(\bar{X} - E_\theta[\mu(\theta)])\} \\
&= E_X\{(E_{\theta|X}[\mu(\theta|\boldsymbol{X})]\bar{X} - E_\theta[\mu(\theta)]\bar{X} - E_{\theta|X}[\mu(\theta|\boldsymbol{X})]E_\theta[\mu(\theta)] + (E_\theta[\mu(\theta)])^2)\} \\
&= E_\theta\{(E_{X|\theta}[\bar{X}\mu(\theta)])\} - (E_\theta[\mu(\theta)])^2 - (E_\theta[\mu(\theta)])^2 + (E_\theta[\mu(\theta)])^2 \\
&= E_\theta\left[(\mu(\theta))^2\right] - (E_\theta[\mu(\theta)])^2 = Var_\theta[\mu(\theta)]\,,
\end{aligned}$$

και ο παρονομαστής

$$\begin{aligned}
B &= E_X\{\bar{X} - E_\theta[\mu(\theta)]\}^2 \\
&= E_X\{\bar{X} - E_X[\bar{X}]\}^2 \\
&= Var_X[\bar{X}]\,.
\end{aligned}$$

Επομένως καταλήγουμε πως

$$\begin{aligned}
\hat{b} &= \frac{Var[\mu(\theta)]}{Var[\bar{X}]} = \frac{Var[\mu(\theta)]}{\frac{1}{n}E(Var[X|\theta]) + Var[\mu(\theta)]} \\
&= \frac{nVar[\mu(\theta)]}{E(Var[X|\theta]) + nVar[\mu(\theta)]} \\
&= \frac{n}{n + \frac{E(Var[X|\theta])}{Var[\mu(\theta)]}} = Z\,.
\end{aligned}$$

Άρα ο βέλτιστος γραμμικός εκτιμητής της $E_{\theta|X}[\mu(\theta|\boldsymbol{X})]$ είναι ο

$$\hat{\alpha} + \hat{b}\bar{X} = (1 - Z)E_\theta[\mu(\theta)] + Z\bar{X}.$$

Παράδειγμα 6.1 $f(x_i|\theta) \sim Poisson(\theta), u(\theta) \sim Gamma(\alpha, \beta)$

Έστω *X* τυχαία μεταβλητή η οποία ακολουθεί την κατανομή Poisson(θ) και η θ είναι κατανομή *Gamma* με παραμέτρους *(α, β)*. Έτσι ισχύει ότι

$$E_X(X|\theta) = Var_X(X|\theta) = \theta \, ,$$

και

$$E_\theta(\theta) = \alpha\beta \ \ \kappa\alpha\iota \ \ Var_\theta(\theta) = \alpha\beta^2 \, .$$

Τότε σύμφωνα με τα παραπάνω έχουμε

$$\begin{aligned} k &= \frac{E_\theta\big(Var_X(X|\theta)\big)}{Var_\theta\big(E_X(X|\theta)\big)} \\ &= \frac{E_\theta(\theta)}{Var_\theta(\theta)} \\ &= \frac{\alpha\beta}{\alpha\beta^2} = \frac{1}{\beta} \, . \end{aligned}$$

Έτσι ο συντελεστής αξιοπιστίας γίνεται

$$Z = \frac{n}{n+k} = \frac{n}{n+\frac{1}{\beta}} = \frac{n\beta}{n\beta+1} \, .$$

Οπότε η εκτίμηση αξιοπιστίας του Bühlmann γίνεται

$$\begin{aligned} C &= Z\bar{x} + (1-Z)E_\theta(\theta) \\ &= \frac{n\beta}{n\beta+1}\bar{x} + \frac{1}{n\beta+1}\alpha\beta \, . \end{aligned}$$

Παρατηρούμε ότι η παραπάνω εκτίμηση, είναι κοινή με την εκ των υστέρων μέση τιμή της κατανομής Gamma$\left(\sum x_i + a, \frac{\beta}{n\beta+1}\right)$, που μελετήθηκε στο Παράδειγμα 2.1.

Παράδειγμα 6.2 $f(x_i|\theta) \sim Normal(\theta, \sigma^2), u(\theta) \sim Normal(\mu, r^2)$

Έστω X τυχαία μεταβλητή η οποία ακολουθεί την Κανονική κατανομή με παραμέτρους (θ, σ^2), οπου $\sigma>0$ δεδομένη σταθερά και η θ είναι κατανομή Normal(μ, r^2), με τα μ και r^2 να θεωρούνται γνωστές σταθερές. Ισχύει δηλαδή ότι

$$E_X(X|\theta) = \theta \; και \; Var_X(X|\theta) = \sigma^2 \,,$$

και

$$E_\theta(\theta) = \mu \;\; και \;\; Var_\theta(\theta) = r^2 \,.$$

Τότε σύμφωνα με τα παραπάνω έχουμε

$$\begin{aligned} k &= \frac{E_\theta\big(Var_X(X|\theta)\big)}{Var_\theta\big(E_X(X|\theta)\big)} \\ &= \frac{E_\theta(\sigma^2)}{Var_\theta(\theta)} \\ &= \frac{\sigma^2}{r^2} \,. \end{aligned}$$

Έτσι ο συντελεστής αξιοπιστίας γίνεται

$$Z = \frac{n}{n+k} = \frac{n}{n+\frac{\sigma^2}{r^2}} = \frac{nr^2}{nr^2+\sigma^2} \,.$$

Οπότε η εκτίμηση αξιοπιστίας του Bühlmann γίνεται

$$\begin{aligned} C &= Z\bar{x} + (1-Z)E_\theta(\theta) \\ &= \frac{nr^2}{nr^2+\sigma^2}\bar{x} + \frac{\sigma^2}{nr^2+\sigma^2}\mu \,. \end{aligned}$$

Διαπιστώνεται ξανά, ότι η παραπάνω εκτίμηση είναι κοινή με την εκ των υστέρων μέση τιμή της Κανονικής κατανομής με παραμέτρους (m, s^2), όπου $m = \frac{n\bar{x}r^2+\mu\sigma^2}{nr^2+\sigma^2} \;\; και \;\; s^2 = \frac{\sigma^2 r^2}{nr^2+\sigma^2}$, όπως υπολογίστηκε στο Παράδειγμα 2.2.

Παράδειγμα 6.3 $f(x_i|\theta) \sim Bernoulli(\theta), u(\theta) \sim Beta(\alpha, \beta)$

Έστω *X* τυχαία μεταβλητή η οποία ακολουθεί την κατανομή *Bernouli(θ)* και η θ είναι κατανομή Beta(α, β). Ισχύει δηλαδή πως

$$E_X(X|\theta) = \theta \;\; \kappa\alpha\iota \;\; Var_X(X|\theta) = \theta(1-\theta)\,,$$

όπως επίσης

$$E_\theta(\theta) = \frac{\alpha}{\alpha+\beta}\,, \quad E_\theta(\theta^2) = \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)}$$

και

$$Var_\theta(\theta) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}\,.$$

Τότε σύμφωνα με τα παραπάνω έχουμε

$$\begin{aligned} k &= \frac{E_\theta\big(Var_X(X|\theta)\big)}{Var_\theta\big(E_X(X|\theta)\big)} \\ &= \frac{E_\theta(\theta(1-\theta))}{Var_\theta(\theta)} = \frac{E_\theta(\theta-\theta^2)}{Var_\theta(\theta)} \\ &= \frac{E_\theta(\theta)-E_\theta(\theta^2)}{Var_\theta(\theta)} = \frac{E_\theta(\theta)-E_\theta(\theta^2)}{Var_\theta(\theta)} \\ &= \frac{\dfrac{\alpha}{\alpha+\beta} - \dfrac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)}}{\dfrac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}} \\ &= \alpha+\beta\,. \end{aligned}$$

Έτσι ο συντελεστής αξιοπιστίας γίνεται

$$Z = \frac{n}{n+k} = \frac{n}{n+\alpha+\beta}\,.$$

Οπότε η εκτίμηση αξιοπιστίας του Bühlmann είναι

$$C = Z\bar{x} + (1-Z)E_\theta(\theta)$$

$$= \frac{n}{n+\alpha+\beta}\bar{x} + \frac{\alpha+\beta}{n+\alpha+\beta}\left(\frac{\alpha}{\alpha+\beta}\right).$$

Γίνεται φανερό πως η παραπάνω εκτίμηση είναι πάλι κοινή με την εκ των υστέρων μέση τιμή της Beta$(\alpha + \sum_{i=1}^{n} x_i\,,\ \beta + n - \sum_{i=1}^{n} x_i)$, όπως υπολογίστηκε στο Παράδειγμα 2.3.

Σημείωση

Μέσα από τα παραπάνω παραδείγματα, έγινε κατανοητό, πως η εκτίμηση αξιοπιστίας του Bühlmann, όχι απλά είναι η βέλτιστη γραμμική προσέγγιση της Μπεϋζιανής εκτίμησης, αλλά για τους παραπάνω τρεις συνδυασμούς κατανομών, συμπίπτει με τη σημειακή προσέγγιση της Μπεϋζιανής προσέγγισης. Ο πρώτος που διαπίστωσε αυτό το γεγονός, για τους παραπάνω συγκεκριμένους συνδυασμούς κατανομών, ήταν ο Bailey (1945 και 1950).

Ο Jewell (1975) διεύρυνε το παραπάνω συμπέρασμα, σε ένα ευρύτερο φάσμα κατανομών. Πιο συγκεκριμένα απέδειξε πως εάν η συνάρτηση πιθανοφάνειας είναι μέλος της απλής εκθετικής οικογένειας κατανομών και είναι συζυγής με την εκ των προτέρων κατανομή, τότε ο εκτιμητής αξιοπιστίας του Bühlmann συμπίπτει με τον Μπεϋζιανό σημειακό εκτιμητή. Η απόδειξη της παραπάνω πρότασης παρατίθεται κάτωθι για την απλή οικογένεια εκθετικών κατανομών.

Θεώρημα 6.2

Στην απλή εκθετική οικογένεια κατανομών, ανήκουν οι κατανομές των οποίων η συνάρτηση πιθανότητας, έχει την μορφή

$$f(x|\theta) = \frac{\alpha(x)e^{-\theta x}}{c(\theta)}\,, \quad (x \,\varepsilon\, X)$$

Όπου το $c(\theta)$ είναι μια σταθερά κανονικοποίησης, τέτοια ώστε $\int f(x|\theta)\,dx = 1$. Η φυσική συζυγής εκ των προτέρων κατανομή της παραπάνω, είναι της μορφής

$$u(\theta) = \frac{c(\theta)^{-t_0} e^{-\theta x_0}}{d(t_0, x_0)}$$

Στην περίπτωση που η συνάρτηση πιθανοφάνειας είναι μέλος της παραπάνω οικογένειας κατανομών και είναι συζυγής με την εκ των προτέρων κατανομή, τότε ο εκτιμητής αξιοπιστίας του Bühlmann συμπίπτει με τον Μπεϋζιανό σημειακό εκτιμητή (βλέπε Jewell, 1975).

Απόδειξη 6.2

Η οικογένεια κατανομών που θα μελετηθεί, είναι εκείνη της απλής οικογένειας εκθετικών κατανομών, με $f_1(x) = x$, και φυσική παραμετροποίηση $\varphi_1(\theta) = -\theta$, δηλαδή

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

όπου $\alpha(x)$ είναι μια τυχαία μη αρνητική συνάρτηση.

Η φυσική συζυγής εκ των προτέρων κατανομή της παραπάνω, είναι της μορφής

$$u(\theta) = \frac{c(\theta)^{-t_0}e^{-\theta x_0}}{d(t_0, x_0)}\ ,\ \ \theta > 0\ ,$$

όπου t_0 και x_0 είναι θετικές σταθερές, και $d(t_0, x_0)$ είναι μια σταθερά παραμετροποίησης.

Επομένως με βάση τα παραπάνω, η εκ των υστέρων κατανομή είναι

$$\begin{aligned} p(\theta|x) &\propto \prod_{i=1}^{t}\left(\frac{e^{-\theta x_i}}{c(\theta)}\right)c(\theta)^{-t_0}e^{-\theta x_0} \\ &\propto \frac{e^{-\theta\sum x_i}}{c(\theta)^t}c(\theta)^{-t_0}e^{-\theta x_0} \\ &\propto c(\theta)^{-(t_0+t)}e^{-\theta(x_0+\sum x_i)}\ . \end{aligned} \tag{6.3}$$

Γνωρίζουμε ότι

$$\int f(x|\theta)\,dx = \int \frac{\alpha(x)e^{-\theta x}}{c(\theta)}dx = 1\ ,$$

επομένως

$$c(\theta) = \int \alpha(x)e^{-\theta x}\,dx\ .$$

Η πρώτη παράγωγος της παραπάνω ως προς θ, είναι

$$c'(\theta) = \int \frac{d}{d\theta} \alpha(x) e^{-\theta x}\, dx = -\int x \cdot \alpha(x) e^{-\theta x}\, dx\,,$$

ενώ η δεύτερη παράγωγος ως προς θ, είναι

$$c''(\theta) = \int \frac{d}{d\theta} x \cdot \alpha(x) e^{-\theta x}\, dx = \int x^2 \cdot \alpha(x) e^{-\theta x}\, dx.$$

Επίσης ισχύει

$$E(X|\Theta = \theta) = \int x f(x|\theta)\, dx = \frac{1}{c(\theta)} \int x \cdot \alpha(x) e^{-\theta x}\, dx$$

$$= -\frac{c'(\theta)}{c(\theta)} = \mu(\theta)\,,$$

και

$$E(X^2|\Theta = \theta) = \int x^2 f(x|\theta)\, dx = \frac{1}{c(\theta)} \int x^2 \cdot \alpha(x) e^{-\theta x}\, dx$$

$$= \frac{c''(\theta)}{c(\theta)}\,.$$

Χρησιμοποιώντας τις παραπάνω δυο ροπές, είναι γνωστό ότι

$$Var(X|\Theta = \theta) = E(X^2|\Theta = \theta) - [E(X|\Theta = \theta)]^2 = v(\theta)$$

$$= \frac{c''(\theta)}{c(\theta)} - \left(-\frac{c'(\theta)}{c(\theta)}\right)^2$$

$$= \frac{c''(\theta)c(\theta) - c'(\theta)c'(\theta)}{[c(\theta)]^2}$$

$$= \frac{d}{d\theta}\left(\frac{c'(\theta)}{c(\theta)}\right)$$

$$= -\frac{d}{d\theta}\big(\mu(\theta)\big)$$

$$= -\mu'(\theta)\,.$$

Παραγωγίζοντας την εκ των προτέρων κατανομή $u(\theta)$, λαμβάνουμε

$$
\begin{aligned}
u'(\theta) &= \frac{d}{d\theta}\frac{c(\theta)^{-t_0}e^{-\theta x_0}}{d(t_0,x_0)} \\
&= \frac{-t_0 c(\theta)^{-t_0-1}c'(\theta)e^{-\theta x_0} - x_0 c(\theta)^{-t_0}e^{-\theta x_0}}{d(t_0,x_0)} \\
&= -t_0\frac{c'(\theta)}{c(\theta)}\frac{c(\theta)^{-t_0}e^{-\theta x_0}}{d(t_0,x_0)} - x_0\frac{c(\theta)^{-t_0}e^{-\theta x_0}}{d(t_0,x_0)} \\
&= t_0\mu(\theta)u(\theta) - x_0 u(\theta)\,.
\end{aligned}
$$

Ολοκληρώνοντας την $u'(\theta)$ ως προς θ, και υποθέτοντας ότι το ολοκλήρωμα ισούται με 0 έχουμε

$$
\int_0^{\infty} u'(\theta)\; d\theta = \int_0^{\infty}\big(t_0\mu(\theta)u(\theta) - x_0 u(\theta)\big)d\theta = 0\,.
$$

Επομένως η παραπάνω γίνεται

$$
\begin{aligned}
\int_0^{\infty} t_0\mu(\theta)u(\theta)d\theta &= \int_0^{\infty} x_0 u(\theta)d\theta \\
t_0\int_0^{\infty} \mu(\theta)u(\theta)d\theta &= x_0\int_0^{\infty} u(\theta)d\theta \\
t_0 E\big(\mu(\theta)\big) &= x_0 \\
E\big(\mu(\theta)\big) &= \frac{x_0}{t_0} = \mu\,.
\end{aligned}
$$

Ανανεώνοντας το παραπάνω αποτέλεσμα με τις πληροφορείς της (6.3), λαμβάνουμε την εκ των υστέρων μέση τιμή

$$
\begin{aligned}
E\big(\mu(\theta)\big) &= \frac{x_0 + \sum x_i}{t_0 + t} \\
&= \frac{t_0}{t_0 + t}\frac{x_0}{t_0} + \frac{t}{t_0 + t}\frac{\sum x_i}{t} \\
&= (1 - Z)m + Z\bar{x}\,.
\end{aligned}
$$

Το παραπάνω αποτέλεσμα, δείχνει ότι η αξιοπιστία είναι ακριβής για την απλή οικογένεια εκθετικών κατανομών. Η δεύτερη παράγωγος της $u(\theta)$ είναι

$$u''(\theta) = \frac{d}{d\theta}(t_0\mu(\theta)u(\theta) - x_0u(\theta)\,)$$

$$= \big(t_0\mu'(\theta)u(\theta) + t_0\mu(\theta)u'(\theta)\big) - x_0u'(\theta)$$

$$= -t_0v(\theta)u(\theta) + (t_0\mu(\theta) - x_0)(t_0\mu(\theta) - x_0)u(\theta)$$

$$= {t_0}^2(\mu(\theta) - \mu)^2u(\theta) - t_0v(\theta)u(\theta)\,.$$

Ολοκληρώνοντας την $u''(\theta)$ ως προς θ, και υποθέτοντας ότι το ολοκλήρωμα ισούται με 0 έχουμε

$$\int_0^\infty u''(\theta)\ d\theta = \int_0^\infty ({t_0}^2(\mu(\theta) - \mu)^2u(\theta) - t_0v(\theta)u(\theta)\,)d\theta = 0\,,$$

και καταλήγουμε πως

$${t_0}^2\int_0^\infty (\mu(\theta) - \mu)^2u(\theta)d\theta = t_0\int_0^\infty v(\theta)u(\theta)d\theta$$

$${t_0}^2Var\big(E(X|\Theta)\big) = t_0E\big(Var(X|\Theta)\big)$$

$$t_0\big\{t_0Var\big(E(X|\Theta)\big) - E\big(Var(X|\Theta)\big)\big\} = 0\,,$$

το οποίο σημαίνει πως η παραπάνω σχέση ισχύει είτε για $t_0 = 0,$ είτε για

$$t_0Var\big(E(X|\Theta)\big) - E\big(Var(X|\Theta)\big) = 0$$

$$t_0 = \frac{E\big(Var(X|\Theta)\big)}{Var\big(E(X|\Theta)\big)}.$$

Γενίκευση του παραπάνω θεωρήματος, δίνεται από τους Jewell (1975) και Ericson (1970).

Σημείωση

Το παραπάνω αποτέλεσμα, είναι εύκολο να διαπιστωθεί πως επεκτείνεται και στην εμπειρική Μπεϋζιανή προσέγγιση, όταν κάποια χαρακτηριστικά της εκ των προτέρων κατανομής, εκτιμηθούν εμπειρικά (βλέπε Norberg, 1980), όπως στο παράδειγμα που ακολουθεί.

Παράδειγμα 6.4 $f(x_i|\theta_i) \sim Exponential(1/\theta_i)$

Ας υποθέσουμε το σύνθετο μοντέλο δειγματοληψίας, όπου x_i τυχαίο δείγμα, οποίο ακολουθεί την Εκθετική$(1/\theta_i)$, με $i = 1, 2, \dots, n$. Επομένως ισχύει ότι

$$E_X(x_i|\theta_i) = \theta_i \quad \kappa\alpha\iota \quad Var_X(x_i|\theta_i) = {\theta_i}^2 .$$

όπως επίσης (βλέπε Παράδειγμα 3.6),

$$E_\theta(\theta_i) = E_X(x_i) \quad \kappa\alpha\iota \quad E_\theta\left({\theta_i}^2\right) = \frac{1}{2} E_X({x_i}^2).$$

Τότε σύμφωνα με τα παραπάνω έχουμε

$$k = \frac{E_\theta\left(Var_X(x_i|\theta_i)\right)}{Var_\theta\left(E_X(x_i|\theta_i)\right)} = \frac{E_\theta\left({\theta_i}^2\right)}{Var_\theta(\theta_i)} .$$

Έτσι ο συντελεστής αξιοπιστίας, για $n = 1$, γίνεται

$$Z = \frac{1}{1+k} = \frac{1}{1 + \frac{E_\theta\left({\theta_i}^2\right)}{Var_\theta(\theta_i)}} = \frac{Var_\theta(\theta_i)}{Var_\theta(\theta_i) + E_\theta\left({\theta_i}^2\right)} .$$

Οπότε η εκτίμηση αξιοπιστίας του Bühlmann είναι

$$\begin{aligned} C &= Zx_i + (1 - Z)E_\theta(\theta_i) \\ &= \frac{E_\theta\left({\theta_i}^2\right)}{Var_\theta(\theta_i) + E_\theta\left({\theta_i}^2\right)} E_\theta(\theta_i) + \frac{Var_\theta(\theta_i)}{Var_\theta(\theta_i) + E_\theta\left({\theta_i}^2\right)} x_i . \end{aligned}$$

Παρατηρούμε πως η παραπάνω σχέση είναι κοινή με την (3.11) του παραδείγματος (3.6), ενώ εύκολα καταλήγουμε στην σχέση (3.12) ακολουθώντας την ίδια διαδικασία.

ΕΠΙΛΟΓΟΣ

Το πρώτο μέρος αυτής της εργασίας, ήταν αφιερωμένο στη μελέτη της Μπεϋζιανής προσέγγισης. Αρχικά έγινε η εισαγωγή στις έννοιες της συνάρτησης πιθανοφάνειας και της εκ των προτέρων κατανομής, ενώ είδαμε πώς γίνεται η σύνθεση της εκ των υστέρων κατανομής, μέσω του Μπεϋζιανού θεωρήματος. Ύστερα μελετήθηκε το πώς πραγματοποιείται η συμπερασματολογία πάνω στην παράμετρο θ και πώς προσδιορίζεται η κατανομή πρόβλεψης. Τέλος παρουσιάστηκε συνοπτικά η ιεραρχική Μπεϋζιανή προσέγγιση, με στόχο τη σύγκρισή της με την εμπειρική Μπεϋζιανή.

Στο τρίτο κεφάλαιο έγινε η εισαγωγή στις βασικές έννοιες και παραδοχές της εμπειρικής Μπεϋζιανής προσέγγισης, και μελετήθηκε σε βάθος μέσω πληθώρας παραδειγμάτων η γενική και η γραμμική εκδοχή, της μη παραμετρικής εμπειρικής Μπεϋζιανής προσέγγισης.

Στο Κεφάλαιο 4 μελετήθηκε η παραμετρική εμπειρική Μπεϋζιανή προσέγγιση και παρουσιάστηκαν παραδείγματα κοινά ως προς την κατανομή που πραγματεύονται, με εκείνα που παρουσιάστηκαν στο Κεφάλαιο 2, με στόχο την ανάδειξη των διαφορών και των ομοιοτήτων των δυο προσεγγίσεων.

Στο επόμενο μέρος, βασιζόμενοι στη μελέτη των Carlin και Louis (2008), έγινε έλεγχος της απόδοσης της εμπειρικής Μπεϋζιανής προσέγγισης, όπου τα αποτελέσματα της παραμετρικής εκδοχής ήταν πολύ ενθαρρυντικά, για όλα τα μοντέλα που εξετάστηκαν. Όσο για την μη παραμετρική εμπειρική Μπεϋζιανή προσέγγιση, έγινε εμφανές πόσο επηρεάζεται η ποιότητα των αποτελεσμάτων από το μέγεθος του εκάστοτε δείγματος.

Στο τελευταίο κεφάλαιο, αφού έγινε μια συνοπτική παρουσίαση της θεωρίας αξιοπιστίας χαρτοφυλακίου, διαπιστώθηκε πως όταν η συνάρτηση πιθανοφάνειας είναι μέλος της απλής εκθετικής οικογένειας κατανομών και είναι συζυγής με την εκ των προτέρων κατανομή, τότε ο εκτιμητής αξιοπιστίας του Bühlmann συμπίπτει με τον Μπεϋζιανο σημειακό εκτιμητή.

Εύχομαι η παρούσα εργασία να φανεί ένα χρήσιμο εγχειρίδιο και πηγή έμπνευσης για όποιον φοιτητή επιθυμεί να ασχοληθεί με την περαιτέρω έρευνα πάνω στην εμπειρική Μπεϋζιανή προσέγγιση και τις δυνατότητές της.

ΒΙΒΛΙΟΓΡΑΦΙΑ

Ελληνική Βιβλιογραφία

Ηλιόπουλος, Γ. (2006): *Βασικές Μέθοδοι Εκτίμησης Παραμέτρων (με σημείο και με διάστημα).* Εκδόσεις Αθ. Σταμούλης, Αθήνα.

Κούτρας, Μ. (2005): *Εισαγωγή στις Πιθανότητες Θεωρία και Εφαρμογές*, τόμος ΙΙ, Β΄ έκδοση. Εκδόσεις Αθ. Σταμούλης, Αθήνα.

Πανάρετος, Ι. & Ξεκαλάκη, Ε. (2000): *Εισαγωγή στη Στατιστική Σκέψη*, τόμος ΙΙ (Εκτιμητική και Έλεγχοι Υποθέσεων), Αθήνα.

Ξένη Βιβλιογραφία

Agresti, A. (2007): *Categorical Data Analysis*, 2nd ed., New York: John Wiley.

Bailey, A.L. (1945): *A generalized theory of credibility*. Proceedings of the Casualty Actuarial Society 32, 13-20.

Bailey, A.L. (1950): *Credibility procedures, La Place's generalization of Bayes' rule, and the combination of collateral knowledge with observed data*. Proceedings of the Casualty Actuarial Society 37, 7-23. Discussion in 37, 94-115.

Bayes, M. & Price, M. (1763): *An Essay towards solving a Problem in the Doctrine of Chances*. By the late Rev. Mr. Bayes, communicated by Mr. Price, in a letter to John Canton, M. A. and F. R. S.

Berger, J.O. (1985): *Statistical Decision Theory and Bayesian Analysis*, 2nd ed., New York: Springer-Verlag.

Buhlmann, H. (1967): *Experience rating and credibility*. ASTIN Bulletin 4, 199-207.

Carlin, B.P. & Louis, T.A. (2008): *Bayes and Empirical Bayes Methods for Data Analysis*, 3rd ed. Boca Raton, Chapman & Hall.

Casella, G. (1985): *An introduction to empirical Bayes data analysis.* The American Statistician, 39, 83-87.

Casella, G. (1992): *Illustrating Empirical Bayes Methods*. Chemometrics and Intelligent Laboratory Systems, 107-125.

Ericson, W.A. (1970): *On the Posterior Mean and Variance of a Population Mean.* Journal of American Statistical Association, 649-652.

Gelman, A., Carlin, J., Stern, H. & Rubin, D.B. (2004): *Bayesian Data Analysis*, 2nd ed. Boca Raton, FL: Chapman and Hall/CRC Press.

Glickman, M.E. & Dyk, D. A. (2007): *Basic Bayesian Methods*. Topics in Biostatistics Methods in Molecular Biology, 319-338.

Gruber, M.H.J. (2014): *Regression Estimators: A Comparative Study*. Academic Press

Herzog, T.N. (1989): Credibility: *The Bayesian model versus Buhlmann's model*, Vol. 41. Transactions of Society of Actuaries.

Jeffreys, H. (1961). *Theory of Probability*, 3rd ed. Oxford: University Press.

Jewell, W.S. (1975): *Credible Means Are Exact Bayesian for Exponential Families*. Vol. 8. The Astin Bulletin, 77-90.

Kass, R.E. & Raftery, A.E. (1995): *Bayes factors*. J. Amer. Statist. Assoc., 773-795.

Kass, R.E. & Steffey, D. (1989): *Approximate Bayesian inference in conditionally independent hierarchical models (parametric empirical Bayes models)*. J. Amer. Statist. Assoc., 717-726.

Klugman, S.A. (1992): *Bayesian statistics in actuarial science: With emphasis on credibility*. Boston: Kluwer Academic.

Louis T.A. (2001): *Empirical Bayes and Likelihood Inference*. Springer Science+Business Media, LLC.

Maritz, J.S. (1970): *Empirical Bayes Methods*. London: Methuen.

Miller, R.B. & Hickman, J.C. (1975): *Insurance credibility theory and Bayesian estimation.* Credibility Theory and Applications, 249-270, New York: Academic Press

Miyasawa K. (1961*): An empirical Bayes estimator of the mean of Normal population*. Bull. Internat. Statist.,38 , 181-188

Morris, C.N. (1983a): *Parametric empirical Bayes inference: Theory and applications*. J. Amer. Statist. Assoc., 47–65.

Norberg, R. (1980): *Empirical Bayes credibility*. Scandinavian Actuarial Journal, 177-194.

Robbins, H. (1955): *An empirical Bayes approach to statistics*. In Proc. 3rd Berkeley Symp. on Math. Statist. and Prob., Berkeley, CA: Univ. of California Press, 157-164.

Robbins, H. (1983): *Some thoughts on empirical Bayes estimation*. Ann. Statist., 713-723.

Robbins, H. (1985): *An Empirical Bayes Approach to Statistics*. Herbert Robbins Selected Papers, 41-47.

Robert, C.P. (2001): *The Bayesian Choice*, 2nd ed. New York: Springer-Verlag.

Savchuk, V. & Tsokos, C.P. (2011): *Bayesian Theory and Methods with Applications.* Atlantis Studies in Probability and Statistics.

Strickland, J.S. (2014): *Predictive Analytics using R*. Lulu.

Πανεπιστημιακές Σημειώσεις

Δελλαπόρτας Π. & Τσιαμυρτζής Π. (2004): *Στατιστική κατά Bayes*.

Πιτσέλης Γ. (2016): *Θεωρία Αξιοπιστίας Χαρτοφυλακίου*.

Παράρτημα A

A.1 Βασικές Διακριτές Κατανομές

Όνομα κατανομής	**Παράμετροι**	**Συνάρτηση πιθανότητας**	**Σύνολο τιμών**	$\boldsymbol{E(X)}$	$\boldsymbol{V(X)}$
Διωνυμική *(Binomial)* $B(n,p)$	n θετικός ακέραιος, $0<p<1$	$\binom{n}{x}p^x(1-p)^{n-x}$	$x=0,1,\dots,n$	np	$np(1-p)$
Poisson $P(\lambda)$	$\lambda>0$	$e^{-\lambda}\frac{\lambda^x}{x!}$	$x=0,1,\dots$	λ	λ
Γεωμετρική (Geometric) $Geo(p)$	$0<p<1$	$p(1-p)^{x-1}$	$x=1,2,\dots$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Αρνητική Διωνυμική *(Negative Binomial)* $NB(r,p)$	r θετικός ακέραιος, $0<p<1$	$\binom{x-1}{r-1}p^r(1-p)^{x-r}$	$x=r,r+1,\dots$	$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$
Βήτα - Διωνυμική *(Beta - Binomial)* $BB(n,\alpha,\beta)$	n θετικός ακέραιος, $\alpha>0,\beta>0$	$\binom{n}{x}\frac{B(x+\alpha,n-x+\beta)}{B(\alpha,\beta)}$	$x=0,1,\dots,n$	$\frac{n\alpha}{\alpha+\beta}$	$\frac{n\alpha\beta(\alpha+\beta+n)}{(\alpha+\beta)^2(\alpha+\beta+1)}$

A.2 Βασικές Συνεχείς Κατανομές

Όνομα κατανομής	**Παράμετροι**	**Συνάρτηση πυκνότητας πιθανότητας**	**Σύνολο τιμών**	$\boldsymbol{E(X)}$	$\boldsymbol{V(X)}$
Κανονική *(Normal)* $N(\mu,\sigma^2)$	$-\infty<\mu<\infty$, $\sigma>0$	$\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	$-\infty<x<\infty$,	μ	σ^2
Εκθετική *(Exponential)* $Exp(\lambda)$	$\lambda>0$	$\lambda e^{-\lambda x}$	$0\le x<\infty$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Γάμμα (Gamma) $Gamma(\alpha,\beta)$	$\alpha>0,\beta>0$	$\frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}$	$0\le x<\infty$	$\frac{\alpha}{\beta}$	$\frac{\alpha}{\beta^2}$
Βήτα *(Beta)* $Beta(\alpha,\beta)$	$\alpha>0,\beta>0$	$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}$	$0\le x\le 1$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta(\alpha+\beta)^{-2}}{(1+\alpha+\beta)}$

Παράρτημα Β

Β.1 Μέθοδος Μεγίστης Πιθανοφάνειας για $\boldsymbol{f(x_i|\theta_i) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i-\theta_i)^2}{2\sigma^2}}}$

Ο λογάριθμος της πιθανοφάνειας είναι

$$log\big(f(x_i|\theta_i)\big) = -\log\big(\sqrt{2\pi}\big) - \log(\sigma) + \frac{(x_i-\theta_i)^2}{2\sigma^2}.$$

Η παράγωγος του λογαρίθμου της πιθανοφάνειας είναι

$$\frac{\partial log\big(f(x_i|\theta_i)\big)}{\partial\theta_i} = \frac{x_i-\theta}{\sigma^2}.$$

Η παραπάνω σχέση μηδενίζεται όταν

$$\frac{\partial log\big(f(x_i|\theta_i)\big)}{\partial\theta_i} = 0$$
$$\Rightarrow \frac{x_i-\theta_i}{\sigma^2} = 0$$
$$\Rightarrow \theta_i = x_{i\,.}$$

Β.2 Βασισμένοι στο θεώρημα 2 του Berger (1985, σελ. 183), έχουμε

$$f_{2.2}(r^2|x) = K(\sigma^2+r^2)^{-(k-1)/2} epx\left\{-\frac{(\kappa-1)s^2}{2(\sigma^2+r^2)}\right\},$$

όπου

$$K = \int_0^\infty (\sigma^2+r^2)^{-(k-1)/2} epx\left\{-\frac{(\kappa-1)s^2}{2(\sigma^2+r^2)}\right\} dr^2$$

και

$$\hat{\theta}_i^{HB} = \big(1-E(B)\big)x_i + E(B)\bar{x},$$

όπου

$$B = \frac{\sigma^2}{\sigma^2+r^2}.$$

Άρα ισχύει

$$E(B) =$$

$$= \int_0^{\infty} \frac{\sigma^2}{\sigma^2 + r^2}\, f_{2.2}(r^2|x)dr^2$$

$$= K\sigma^2 \int_0^{\infty} \frac{1}{\sigma^2 + r^2}\, (\sigma^2 + r^2)^{-(k-1)/2} epx\left\{-\frac{(\kappa-1)s^2}{2(\sigma^2 + r^2)}\right\} dr^2$$

$$= K\sigma^2 \int_0^{\infty}\, (\sigma^2 + r^2)^{-(k+1)/2} epx\left\{-\frac{(\kappa-1)s^2}{2(\sigma^2 + r^2)}\right\} dr^2$$

$$= \frac{\sigma^2 \int_0^{\infty}\, (\sigma^2 + r^2)^{-(k+1)/2} epx\left\{-\frac{(\kappa-1)s^2}{2(\sigma^2 + r^2)}\right\} dr^2}{\int_0^{\infty} (\sigma^2 + r^2)^{-(k-1)/2} epx\left\{-\frac{(\kappa-1)s^2}{2(\sigma^2 + r^2)}\right\} dr^2}$$

$$= \frac{\sigma^2 \int_0^{\infty} \left\{\frac{1}{(\sigma^2 + r^2)}\right\}^{(k+1)/2} epx\left\{-\frac{(\kappa-1)s^2}{2(\sigma^2 + r^2)}\right\} dr^2}{\int_0^{\infty} \left\{\frac{1}{(\sigma^2 + r^2)}\right\}^{(k-1)/2} epx\left\{-\frac{(\kappa-1)s^2}{2(\sigma^2 + r^2)}\right\} dr^2}$$

$$= \frac{\sigma^2 \int_0^{\infty} \left\{\frac{(\kappa-1)s^2}{2(\sigma^2 + r^2)}\right\}^{(\kappa+1)/2} \left(\frac{2}{(\kappa-1)s^2}\right)^{(\kappa+1)/2} epx\left\{-\frac{(\kappa-1)s^2}{2(\sigma^2 + r^2)}\right\} dr^2}{\int_0^{\infty} \left\{\frac{(\kappa-1)s^2}{2(\sigma^2 + r^2)}\right\}^{(\kappa-1)/2} \left(\frac{2}{(\kappa-1)s^2}\right)^{(\kappa-1)/2} epx\left\{-\frac{(\kappa-1)s^2}{2(\sigma^2 + r^2)}\right\} dr^2}$$

$$= \frac{2\sigma^2 \int_0^{\infty} \left\{\frac{(\kappa-1)s^2}{2(\sigma^2 + r^2)}\right\}^{(\kappa-3)/2} \left(\frac{(\kappa-1)s^2}{2(\sigma^2 + r^2)}\right)' epx\left\{-\frac{(\kappa-1)s^2}{2(\sigma^2 + r^2)}\right\} \left(\frac{(\kappa-1)s^2}{2(\sigma^2 + r^2)}\right)' dr^2}{s^2 \int_0^{\infty} \left\{\frac{(\kappa-1)s^2}{2(\sigma^2 + r^2)}\right\}^{(\kappa-5)/2} \left(\frac{(\kappa-1)s^2}{2(\sigma^2 + r^2)}\right)' epx\left\{-\frac{(\kappa-1)s^2}{2(\sigma^2 + r^2)}\right\} \left(\frac{(\kappa-1)s^2}{2(\sigma^2 + r^2)}\right)' dr^2}.$$

Θέτοντας $y = \frac{(\kappa-1)s^2}{2(\sigma^2+r^2)}$ και $u = \frac{(\kappa-1)s^2}{2\sigma^2}$, λαμβάνουμε

$$E(B) = \frac{\int_0^u y^{(\kappa-3)/2} e^{-y}\, dy}{u\int_0^u y^{(\kappa-5)/2} e^{-y}\, dy}$$

$$= \frac{y^{(\kappa-3)/2}(-e^{-y})\big|_0^u + \int_0^u \frac{(\kappa-3)}{2} y^{(\kappa-5)/2} e^{-y}\ dy}{u\int_0^u y^{(\kappa-5)/2} e^{-y}\, dy}$$

$$= \frac{u^{(\kappa-3)/2}(-e^{-u})}{u\int_0^u y^{(\kappa-5)/2} e^{-y}\, dy} + \frac{(\kappa-3)}{2u}$$

$$= \frac{(\kappa-3)}{2u}\left\{1 - \frac{2u^{(\kappa-3)/2} e^{-u}}{(p-3)\int_0^u y^{(\kappa-5)/2} e^{-y}\, dy}\right\}$$

$$= \frac{(\kappa-3)}{2u}\{1 - g(s^2)\}$$

Όπου

$$g(s^2) = \frac{2u^{(\kappa-3)/2} e^{-u}}{(\kappa-3)\int_0^u y^{(\kappa-5)/2} e^{-y}\, dy}$$

Για $\kappa = 5$ η $g(s^2)$ γίνεται

$$g(s^2) = \frac{u\, e^{-u}}{\int_0^u e^{-y}\, dy} = \frac{u\, e^{-u}}{(1-e^{-u})}$$

$$= \frac{\left(\frac{u}{e^u}\right)}{\left(\frac{e^u - 1}{e^u}\right)} = \frac{u}{e^u - 1}.$$

Αντικαθιστώντας όπου $u = \frac{(\kappa-1)s^2}{2\sigma^2}$ και $\sigma^2 = 1$ στην $g(s^2)$, λαμβάνουμε

$$g(s^2) = \frac{u}{e^u - 1}$$

$$= \frac{(\kappa-1)s^2}{2}\left[exp\left\{\frac{(\kappa-1)s^2}{2}\right\} - 1\right]^{-1}.$$

Αντικαθιστώντας όπου $u = \frac{(\kappa-1)s^2}{2\sigma^2}$ και $\sigma^2 = 1$ στην $E(B)$, λαμβάνουμε

$$E(B) = \frac{(\kappa-3)}{(\kappa-1)s^2}\{1 - g(s^2)\}.$$

B.2 Μέθοδος Μεγίστης Πιθανοφάνειας για $\boldsymbol{f(x_i|\theta_i)} = \frac{e^{-\theta_i}\theta_i^{x_i}}{x_i!}$

Ο λογάριθμος της πιθανοφάνειας είναι

$$log\big(f(x_i|\theta_i)\big) = -\,log(x_i!) - \theta_i + x_i\,log(\theta_i).$$

Η παράγωγος του λογαρίθμου της πιθανοφάνειας είναι

$$\frac{\partial log\big(f(x_i|\theta_i)\big)}{\partial\theta_i} = -1 + \frac{x_i}{\theta_i}.$$

Η παραπάνω σχέση μηδενίζεται όταν

$$\frac{\partial log\big(f(x_i|\theta_i)\big)}{\partial\theta_i} = 0$$
$$\Rightarrow -1 + \frac{x_i}{\theta_i} = 0$$
$$\Rightarrow \theta_i = x_i\,.$$

B.3 Μέθοδος Μεγίστης Πιθανοφάνειας για $\boldsymbol{f(x_i|\theta_i)} = \binom{k}{x_i}\theta_i^{x_i}(1-\theta_i)^{k-x_i}$

Ο λογάριθμος της πιθανοφάνειας είναι

$$log\big(f(x_i|\theta_i)\big) = log\binom{k}{x_i} + x_i\,log(\theta_i) + (k - x_i)\,log(1-\theta_i).$$

Η παράγωγος του λογαρίθμου της πιθανοφάνειας είναι

$$\frac{\partial log\big(f(x_i|\theta_i)\big)}{\partial\theta_i} = \frac{x_i}{\theta_i} + \frac{k - x_i}{1-\theta_i}.$$

Η παραπάνω σχέση μηδενίζεται όταν

$$\frac{\partial log\big(f(x_i|\theta_i)\big)}{\partial\theta_i} = 0$$
$$\Rightarrow \frac{x_i}{\theta_i} + \frac{k - x_i}{1-\theta_i} = 0$$
$$\Rightarrow x_i(1-\theta_i) = \theta_i(k - x_i)$$
$$\Rightarrow \theta_i = \frac{x_i}{k}.$$

Παράρτημα Γ

Γ.1 Κώδικας R Απόδοσης Normal-Normal

```
theta<-0;psi<-0

FtotError<-0;FMontePart<-data.frame(1:5)
BtotError<-0;BMontePart<-data.frame(1:5)
HBtotError<-0;HBMontePart<-data.frame(1:5)
EBtotError<-0;EBMontePart<-data.frame(1:5)

for(k in 1:10000){
     for(i in 1:5){theta[i]<-rnorm(1,0,1);psi[i]<-rnorm(1,theta[i],1)}

######################### -Κλασσική Προσέγγιση- #########################

FestTheta<-0;FgroupError<-0

##Κλασσική σημειακή εκτίμηση##
FestTheta<-psi

##Κίνδυνος ομάδας i##
FgroupError<-(sum((FestTheta-theta)^2)/5)

##Άθροισμα κινδύνου ομάδων##
FtotError<-FtotError+FgroupError

##Εκτιμώμενο τετραγωνικό σφάλμα##
FMontePart[k]<-((FestTheta-theta)^2)

######################## - Μπεϋζιανή Προσέγγιση - #######################

BestTheta<-0;BgroupError<-0

##Μπεϋζιανή σημειακή εκτίμηση##
BestTheta<-(psi/2)

##Κίνδυνος ομάδας i##
BgroupError<-(sum((BestTheta-theta)^2)/5)

##Άθροισμα κινδύνου ομάδων##
BtotError<-BtotError+BgroupError

##Εκτιμώμενο τετραγωνικό σφάλμα##
BMontePart[k]<-((BestTheta-theta)^2)
```

```
################### -Ιεραρχική  Μπεϋζιανή Προσέγγιση - ###################

HBestMean<-0;HBestVar<-0;Gs<-0;HBestTheta<-0;HBgroupError<-0

##Εκτίμηση μέσης τιμής##
HBestMean<-mean(psi)

##Εκτίμηση διακύμανσης##
HBestVar<-var(psi)

##g(s^2) εκτίμηση##
Gs<-(2*HBestVar/((exp(2*HBestVar))-1))

##Ιεραρχική Μπεϋζιανή σημειακή εκτίμηση##
HBestTheta<-((((1-Gs)/(2*HBestVar))*HBestMean)+(1-((1-
Gs)/(2*HBestVar)))*psi)

##Κίνδυνος ομάδας i##
HBgroupError<-sum((HBestTheta-theta)^2)/5

##Άθροισμα κινδύνου ομάδων##
HBtotError<-HBtotError+HBgroupError

##Εκτιμώμενο τετραγωνικό σφάλμα##
HBMontePart[k]<-((HBestTheta-theta)^2)


################### -Εμπειρική  Μπεϋζιανή Προσέγγιση - ###################

EBestMean<-0;EBvar<-0;EBestVar<-0;B<-0;EBestTheta<-0;EBgroupError<-0

##Εκτίμηση μέσης τιμής##
EBestMean<-mean(psi)

##Εκτίμηση διακύμανσης##
EBvar<-var(psi)-1;EBestVar<-if(EBvar>0){var(psi)-1}else{0}

##B εκτίμηση##
B<-(1/(1+EBestVar))

##Εμπειρική Μπεϋζιανή σημειακή εκτίμηση ##
EBestTheta<-((B*EBestMean)+(1-B)*psi)

##Κίνδυνος ομάδας i ##
EBgroupError<-sum((EBestTheta-theta)^2)/5

##Άθροισμα κινδύνου ομάδων ##
EBtotError<-EBtotError+EBgroupError

##Εκτιμώμενο τετραγωνικό σφάλμα ##
EBMontePart[k]<-((EBestTheta-theta)^2)

############################# -Τέλος λούπας- #############################
}
```

```
##Εκτιμώμενος μέσος κίνδυνος Κλασσικής προσέγγισης##
FestRisk<-FtotError/10000

##Εκτιμώμενος μέσος κίνδυνος Μπεϋζιανής προσέγγισης##
BestRisk<-BtotError/10000

##Εκτιμώμενος μέσος κίνδυνος Ιεραρχικής Μπεϋζιανής προσέγγισης##
HBestRisk<-HBtotError/10000

##Εκτιμώμενος μέσος κίνδυνος Εμπειρικής Μπεϋζιανής προσέγγισης##
EBestRisk<-EBtotError/10000



##Εκτιμώμενο Monte Carlo St. Error Κλασσικής προσέγγισης##
FMontePartB<-((FMontePart-FestRisk)^2);FMontePartG<-sum(FMontePartB)
FMCerror<-sqrt(FMontePartG/(50000*49999))

##Εκτιμώμενο Monte Carlo St. Μπεϋζιανής προσέγγισης##
BMontePartB<-((BMontePart-BestRisk)^2);BMontePartG<-sum(BMontePartB)
BMCerror<-sqrt(BMontePartG/(50000*49999))

##Εκτιμώμενο Monte Carlo St. Error Ιεραρχικής Μπεϋζιανής προσέγγισης##
HBMontePartB<-((HBMontePart-HBestRisk)^2);HBMontePartG<-sum(HBMontePartB)
HBMCerror<-sqrt(HBMontePartG/(50000*49999))

##Εκτιμώμενο Monte Carlo St. Error Εμπειρικής Μπεϋζιανής προσέγγισης##
EBMontePartB<-((EBMontePart-EBestRisk)^2);EBMontePartG<-sum(EBMontePartB)
EBMCerror<-sqrt(EBMontePartG/(50000*49999))

####################### -Εμφάνιση Αποτελεσμάτων- #######################

Tags<-c("r","se(r)")
SumrF<-c(FestRisk,FMCerror)
SumrB<-c(BestRisk,BMCerror)
SumrHB<-c(HBestRisk,HBMCerror)
SumrEB<-c(EBestRisk,EBMCerror)
SumrsA<-
data.frame(Tags,Frequentist=SumrF,Bayes=SumrB,HBayes=SumrHB,EBayes=SumrEB)
```

Γ.2 Κώδικας R Απόδοσης Poisson-Gamma

```
theta<-0;psi<-0

FtotError<-0;FMontePart<-data.frame(1:10)
BtotError<-0;BMontePart<-data.frame(1:10)
EBtotError<-0;EBMontePart<-data.frame(1:10)
ZeroVar<-0

for(k in 1:10000){
   for(i in 1:10){theta[i]<-rgamma(1,shape=1,scale=1);psi[i]<-rpois(1,theta[i])}
```

```
######################### -Έλεγχος σφάλματος- ########################

  if(var(psi)==0){

########### -Σε περίπτωση σφάλματος, παραγωγή νέου δείγματος- ##########

   while(var(psi)==0){
    for(i in 1:10){theta[i]<-rgamma(1,shape=1,scale=1);psi[i]<-rpois(1,theta[i])}
    ZeroVar<-ZeroVar+1
                      }
                   }

######################## -Κλασσική Προσέγγιση- ########################

FestTheta<-0;FgroupError<-0

##Κλασσική σημειακή εκτίμηση##
FestTheta<-psi

##Κίνδυνος ομάδας i##
FgroupError<-(sum((FestTheta-theta)^2)/10)

##Άθροισμα κινδύνου ομάδων##
FtotError<-FtotError+FgroupError

##Εκτιμώμενο τετραγωνικό σφάλμα ##
FMontePart[k]<-((FestTheta-theta)^2)

####################### - Μπεϋζιανή Προσέγγιση - ######################

BestTheta<-0;BgroupError<-0

##Μπεϋζιανή σημειακή εκτίμηση##
BestTheta<-(psi*(1/2)+(1/2))

##Κίνδυνος ομάδας i##
BgroupError<-(sum((BestTheta-theta)^2)/10)

##Άθροισμα κινδύνου ομάδων##
BtotError<-BtotError+BgroupError

##Εκτιμώμενο τετραγωνικό σφάλμα ##
BMontePart[k]<-((BestTheta-theta)^2)


################### -Εμπειρική  Μπεϋζιανή Προσέγγιση - #################

EBestMean<-0;EBestVar<-0;B<-0;EBestTheta<-0;EBgroupError<-0

##Εκτίμηση μέσης τιμής##
EBestMean<-mean(psi)

##Εκτίμηση διακύμανσης##
EBestVar<-var(psi)
```

```
##B εκτίμηση##
B<-(EBestMean/EBestVar)

##Εμπειρική Μπεϋζιανή σημειακή εκτίμηση ##
EBestTheta<-((B*EBestMean)+(1-B)*psi)

##Κίνδυνος ομάδας i##
EBgroupError<-sum((EBestTheta-theta)^2)/10

##Άθροισμα κινδύνου ομάδων##
EBtotError<-EBtotError+EBgroupError

##Εκτιμώμενο τετραγωνικό σφάλμα ##
EBMontePart[k]<-((EBestTheta-theta)^2)

############################# -Τέλος λούπας- #############################
}

##Εκτιμώμενος μέσος κίνδυνος Κλασσικής προσέγγισης##
FestRisk<-FtotError/10000

##Εκτιμώμενος μέσος κίνδυνος Μπεϋζιανής προσέγγισης##
BestRisk<-BtotError/10000

##Εκτιμώμενος μέσος κίνδυνος Εμπειρικής Μπεϋζιανής προσέγγισης##
EBestRisk<-EBtotError/10000

##Εκτιμώμενο Monte Carlo St. Error Κλασσικής προσέγγισης##
FMontePartB<-((FMontePart-FestRisk)^2);FMontePartG<-sum(FMontePartB)
FMCerror<-sqrt(FMontePartG/(100000*99999))

##Εκτιμώμενο Monte Carlo St. Error Μπεϋζιανής προσέγγισης##
BMontePartB<-((BMontePart-BestRisk)^2);BMontePartG<-sum(BMontePartB)
BMCerror<-sqrt(BMontePartG/(100000*99999))

##Εκτιμώμενο Monte Carlo St. Error Εμπειρικής Μπεϋζιανής προσέγγισης##
EBMontePartB<-((EBMontePart-EBestRisk)^2);EBMontePartG<-sum(EBMontePartB)
EBMCerror<-sqrt(EBMontePartG/(100000*99999))

####################### -Εμφάνιση Αποτελεσμάτων- ########################

Tags<-c("r","se(r)")
SumrF<-c(FestRisk,FMCerror)
SumrB<-c(BestRisk,BMCerror)
SumrEB<-c(EBestRisk,EBMCerror)
Sumrs<-data.frame(Tags,Frequentist=SumrF,Bayes=SumrB,EBayes=SumrEB)
Sumrs
```

Γ.3 Κώδικας R Απόδοσης Beta-Binomial

```
m<-5
theta<-0;psi<-0
FtotError<-0;FMontePart<-data.frame(1:20)
BtotError<-0;BMontePart<-data.frame(1:20)
EBtotError<-0;EBMontePart<-data.frame(1:20)

EBestMean<-0;EBestVar<-0;B<-0;EBestTheta<-0;EBgroupError<-0

for(k in 1:10000){
     for(i in 1:20){theta[i]<-rbeta(1,2,2);psi[i]<-rbinom(1,m,theta[i])}

########################## -Έλεγχος σφάλματος- ##########################

     EBestMean<-mean(psi);EBestVar<-var(psi);EBestSqMean<-
var(psi)+(EBestMean)^2
     check<-((m-1)*((EBestMean)^2)+(m*(EBestMean-EBestSqMean)))

########### -Σε περίπτωση σφάλματος, παραγωγή νέου δείγματος- ###########

          if(check==0){
               while(check1==0){
                for(i in 1:5){theta[i]<-rbeta(1,1,1);psi[i]<-
rbinom(1,m,theta[i])}
                EBestMean<-mean(psi);EBestVar<-var(psi);EBestSqMean<-
var(psi)+(EBestMean)^2
                check<-((m-1)*((EBestMean)^2)+(m*(EBestMean-
EBestSqMean)))
                                                         }
                              }

######################### -Κλασσική Προσέγγιση- #########################

FestTheta<-0;FgroupError<-0

##Κλασσική σημειακή εκτίμηση##
FestTheta<-psi/m

##Κίνδυνος ομάδας i##
FgroupError<-(sum((FestTheta-theta)^2)/20)


##Άθροισμα κινδύνου ομάδων##
FtotError<-FtotError+FgroupError

##Εκτιμώμενο τετραγωνικό σφάλμα ##
FMontePart[k]<-((FestTheta-theta)^2)

######################## - Μπεϋζιανή Προσέγγιση - #######################

BestTheta<-0;BgroupError<-0

##Μπεϋζιανή σημειακή εκτίμηση##
BestTheta<-(1+psi)/(2+m)

##Κίνδυνος ομάδας i##
BgroupError<-(sum((BestTheta-theta)^2)/20)

##Άθροισμα κινδύνου ομάδων##
BtotError<-BtotError+BgroupError
```

```
##Εκτιμώμενο τετραγωνικό σφάλμα ##
BMontePart[k]<-((BestTheta-theta)^2)

################### -Εμπειρική  Μπεϋζιανή Προσέγγιση - ##################

EBestMean<-0;EBestVar<-0;B<-0;EBestTheta<-0;EBgroupError<-0

##Εκτίμηση μέσης τιμής##
EBestMean<-mean(psi)

##Εκτίμηση διακύμανσης##
EBestVar<-var(psi)

##Εκτίμηση δεύτερης ροπής##
EBestSqMean<-var(psi)+(EBestMean)^2

##Εκτίμηση  α##
a<-(EBestMean*(EBestSqMean-m*EBestMean))/((m-
1)*((EBestMean)^2)+(m*(EBestMean-EBestSqMean)))
      if(a<0){a<-0}

##Εκτίμηση  β##
b<-((m-EBestMean)*(EBestSqMean-(m*EBestMean)))/((m-
1)*((EBestMean)^2)+(m*(EBestMean-EBestSqMean)))
      if(b<0){b<-0}

##Εμπειρική Μπεϋζιανή σημειακή εκτίμηση##
EBestTheta<-(a+psi)/(a+b+m)

##Κίνδυνος ομάδας i##
EBgroupError<-sum((EBestTheta-theta)^2)/20

##Άθροισμα κινδύνου ομάδων##
EBtotError<-EBtotError+EBgroupError

##Εκτιμώμενο τετραγωνικό σφάλμα##
EBMontePart[k]<-((EBestTheta-theta)^2)

############################## -Τέλος λούπας- ############################
}

##Εκτιμώμενος μέσος κίνδυνος Κλασσικής προσέγγισης##
FestRisk<-FtotError/10000

##Εκτιμώμενος μέσος κίνδυνος Μπεϋζιανής προσέγγισης##
BestRisk<-BtotError/10000

##Εκτιμώμενος μέσος κίνδυνος Εμπειρικής Μπεϋζιανής προσέγγισης##
EBestRisk<-EBtotError/10000

##Εκτιμώμενο Monte Carlo St. Error Κλασσικής προσέγγισης##
FMontePartB<-((FMontePart-FestRisk)^2);FMontePartG<-sum(FMontePartB)
FMCerror<-sqrt(FMontePartG/(200000*199999))

##Εκτιμώμενο Monte Carlo St. Error Μπεϋζιανής προσέγγισης##
BMontePartB<-((BMontePart-BestRisk)^2);BMontePartG<-sum(BMontePartB)
BMCerror<-sqrt(BMontePartG/(200000*199999))

##Εκτιμώμενο Monte Carlo St. Error Εμπειρικής Μπεϋζιανής προσέγγισης##
EBMontePartB<-((EBMontePart-EBestRisk)^2);EBMontePartG<-sum(EBMontePartB)
EBMCerror<-sqrt(EBMontePartG/(200000*199999))
```

```
###################### -Εμφάνιση Αποτελεσμάτων- #######################

Tags<-c("r","se(r)")
SumrF<-c(round(FestRisk,6),round(FMCerror,6))
SumrB<-c(round(BestRisk,6),round(BMCerror,6))
SumrEB<-c(round(EBestRisk,6),round(EBMCerror,6))
Sumrs<-data.frame(Tags,Frequentist=SumrF,Bayes=SumrB,EBayes=SumrEB)
Sumrs
```

Γ.4 Κώδικας R Σύγκριση Παραμετρικής και Μη Εμπειρικής Μπεϋζιανής

```
## -Sample Size- ##
L<-500

theta<-0;psi<-0;FtotError<-0;BtotError<-0;EBtotError<-0;NEBtotError<-0

#################### -Παραγωγή Τυχαίου Δείγματος- #####################

for(k in 1:100){
    for(i in 1:L){theta[i]<-rgamma(1,shape=1,scale=1);psi[i]<-rpois(1,theta[i])}

######################## -Έλεγχος σφάλματος- #######################

    if(var(psi)==0){


########## -Σε περίπτωση σφάλματος, παραγωγή νέου δείγματος- ##########
         while(var(psi)==0){
             for(i   in   1:L){theta[i]<-rgamma(1,shape=1,scale=1);psi[i]<-
rpois(1,theta[i])}
                         }
               }

############ -Παραμετρική Εμπειρική  Μπεϋζιανή Προσέγγιση - ############

EBestMean<-0;EBestVar<-0;B<-0;EBestTheta<-0;EBgroupError<-0

##Εκτίμηση μέσης τιμής##
EBestMean<-mean(psi)

##Εκτίμηση διακύμανσης##
EBestVar<-var(psi)

##B εκτίμηση##
B<-(EBestMean/EBestVar)

##Παραμετρική Εμπειρική Μπεϋζιανή σημειακή εκτίμηση##
EBestTheta<-((B*EBestMean)+(1-B)*psi)

##Κίνδυνος ομάδας i##
EBgroupError<-sum((EBestTheta-theta)^2)/L
```

```
##Άθροισμα κινδύνου ομάδων##
EBtotError<-EBtotError+EBgroupError

########### -Μη Παραμετρική Εμπειρική  Μπεϋζιανή Προσέγγιση - ###########

NEBestTheta<-0;NEBgroupError<-0

for(i in 1:L){

##Μη Παραμετρική Εμπειρική Μπεϋζιανή σημειακή εκτίμηση##
NEBestTheta[i]<-
(psi[i]+1)*(length(which(psi==psi[i]+1)))/(length(which(psi==psi[i])))
}

##Κίνδυνος ομάδας i##
NEBgroupError<-sum((NEBestTheta-theta)^2)/L

##Άθροισμα κινδύνου ομάδων##
NEBtotError<-NEBtotError+NEBgroupError

############################ -Τέλος λούπας- ############################
}

##Εκτιμώμενος μέσος κίνδυνος Παρ. Εμπειρικής Μπεϋζιανής προσέγγισης##
EBestRisk<-EBtotError/100

##Εκτιμώμενος μέσος κίνδυνος Μη Παρ. Εμπειρικής Μπεϋζιανής προσέγγισης##
NEBestRisk<-NEBtotError/100
```