Δικαρος Γ. Ανδρεας

Αρνητικη Διωνυμικη Κατανομη και Εκτιμηση των Παραμετρων της

Μεταπτυχιακη Διατριβη

Τμημα Μαθηματικων - Μηχανικων Η/Υ

Πανεπιστημιο Πατρων

Σεπτεμβριος 2010, Πατρα

Στη σύζυγό μου,

Αγγελική.

Περιεχόμενα

Εισαγωγή

Η παρούσα μεταπτυχιακή διατριβή εντάσσεται ερευνητικά στην περιοχή της Στατιστικής Θεωρίας Αποφάσεων και ειδικότερα στη μελέτη της αρνητικής διωνυμικής κατανομής καθώς επίσης και στην εκτίμηση των παραμέτρων της.

Στο Κεφάλαιο 1 παρουσιάζονται κάποιοι χρήσιμοι, για την πορεία της μελέτης μας, ορισμοί και θεωρήματα.

Στο Κεφάλαιο 2 μελετάται το μοντέλο της αρνητικής διωνυμικής κατανομής, δίνονται τα χαρακτηριστικά μεγέθη αυτής και παρουσιάζονται οι διαφορετικές παραμετρικοποιήσεις της.

Στο Κεφάλαιο 3, εξετάζεται το πρόβλημα εκτίμησης των παραμέτρων της αρνητικής διωνυμικής κατανομής και πιο ειδικά η εκτίμηση για τις διάφορες παραμετρικοποιήσης της . Για περισσότερη ανάλυση χρησιμοποιούνται η εκτίμηση μέγιστης πιθανοφάνειας, η εκτίμηση με τη μέθοδο των ροπών και πιο εξειδικευμένες υπολογιστικές μέθοδοι εκτίμησης.

Στο Κεφάλαιο 4, και για το ίδιο πρόβλημα εκτίμησης που πραγματεύεται το προηγούμενο κεφάλαιο, επιλέγεται ο βέλτιστος εκτιμητής των παραμέτρων της αρνητικής διωνυμικής κατανομής και παρουσιάζεται ένα παράδειγμα για την κατανόηση των μεθόδων εκτίμησης.

Α. Δίκαρος, Πάτρα 2010.

Κεφάλαιο 1

Βασικοί Ορισμοί και Θεωρήματα

Σ' αυτό το κεφαλαίο θα αναφέρουμε κάποιους βασικούς ορισμούς και Θεωρήματα,χωρίς τις αποδείξεις τους, από την πλευρά της Μαθηματικής Στατιστικής.

1.1 Εισαγωγή -Αμερόληπτοι Εκτιμητές

Έστω ότι δίνονται δεδομένα $\underset{\sim}{X} = (X_1, X_2, \ldots, X_n)$ με από κοινού πυκνότητα πιθανότητας $f_{\underset{\sim}{X}}(x;\theta)$,που εξαρτάται από μια άγνωστη παράμετρο θ,η οποία ανήκει σε κάποιο σύνολο Θ.Το θ λέγεται **άγνωστη παράμετρος** και το Θ καλείται **παραμετρικός χώρος**.Σκοπός μας είναι να εκτιμήσουμε μια συνάρτηση του θ,έστω $g(\cdot) : \Theta \to \mathbb{R}^{\kappa}$,$\kappa \geq 1$,η οποία ονομάζεται **παραμετρική συνάρτηση**.Το τυχαίο διάνυσμα $\underset{\sim}{X}$ αναφέρεται σαν **δείγμα**.Αν επιπλέον οι τυχαίες μεταβλητές $X_i, i = 1, 2, \ldots, n$ είναι ανεξάρτητες και ισόνομες,δηλαδή έχουν την ίδια κατανομή,τότε το $\underset{\sim}{X}$ αναφέρεται σαν **τυχαίο δείγμα**.

Ορισμός 1.1.1. Μια συνάρτηση μόνο του δείγματος καλείται **στατιστική συνάρτηση**.

Ορισμός 1.1.2. Μια στατιστική συνάρτηση,έστω $\delta(\underset{\sim}{X})$, που χρησιμοποιείται για την εκτίμηση της τιμής της άγνωστης παραμέτρου θ,(ή γενικότερα για την εκτίμηση της παραμετρικής συνάρτησης $g(\theta)$,όπου $g(.) : \Theta \to \mathbb{R}^{\kappa}$,$\kappa \geq 1$)αναφέρεται σαν **εκτιμητής** του θ.

Ορισμός 1.1.3. Ο εκτιμητής $T = T(\underset{\sim}{X})$,ονομάζεται **αμερόληπτος εκτιμητής** της παραμετρικής συνάρτησης $g(\theta)$,αν

$$E_\theta(T(\underset{\sim}{X})) = g(\theta), \forall\theta \in \Theta.$$

Ένα από τα πιο συνηθισμένα κριτήρια επιλογής εκτιμητών είναι το Μέσο Τετραγωνικό Σφάλμα του εκτιμητή $T(\underset{\sim}{X})$,συμβολικά $MT\Sigma(T,\theta)$, που ορίζεται παρακάτω.

Ορισμός 1.1.4. Το Μέσο Τετραγωνικό Σφάλμα του εκτιμητή $T(\underset{\sim}{X})$,ορίζεται ως εξής,

$$MT\Sigma(T,\theta) = E_\theta(T(\underset{\sim}{X}) - g(\theta))^2.$$

Πρόταση 1.1.1.

$$MT\Sigma(T,\theta) = Var_\theta(T(\underset{\sim}{X})) + (E_\theta(T(\underset{\sim}{X})) - g(\theta))^2.$$

Η ποσότητα

$$b(T,\theta) = E_\theta(T(\underset{\sim}{X})) - g(\theta)$$

καλείται μεροληψία ή συστηματικό σφάλμα του εκτιμητή T για την ποσότητα $g(\theta)$,οπότε

$$MT\Sigma(T,\theta) = Var_\theta(T(X)) + b^2(T,\theta).$$

Παρατήρηση 1.1.1. Αν T είναι αμερόληπτος εκτιμητής της παραμετρικής συνάρτησης $g(\theta)$,τότε $MT\Sigma(T,\theta) = Var_\theta(T(\underset{\sim}{X}))$.

Ορισμός 1.1.5. Ο εκτιμητής T_1 ονομάζεται καλύτερος από τον T_2(ως προς το Μέσο Τετραγωνικό Σφάλμα)για την $g(\theta)$,αν,

$$MT\Sigma(T_1,\theta) \leqslant MT\Sigma(T_2,\theta), \forall\theta \in \Theta$$

και επιπλέον

$$MT\Sigma(T_1,\theta_0) < MT\Sigma(T_2,\theta_0),\ \text{για κάποιο } \theta_0 \in \Theta.$$

Ορισμός 1.1.6. Εάν ο εκτιμητής T_1 είναι καλύτερος από τον T_2(ως προς το το Μέσο Τετραγωνικό Σφάλμα)για την $g(\theta)$,τότε ο T_2 λέγεται **μη αποδεκτός** για την εκτίμηση της παραμετρικής συνάρτησης $g(\theta)$.

Ορισμός 1.1.7. Ο T ονομάζεται **βέλτιστος εκτιμητής**(ως προς το το Μέσο Τετραγωνικό Σφάλμα)για την $g(\theta)$,αν είναι καλύτερος από κάθε άλλο εκτιμητή της παραμετρικής συνάρτησης $g(\theta)$.

Οι ακόλουθες Προτάσεις μας βοηθάνε να βρούμε αμερόληπτους εκτιμητές τόσο για την μέση τιμή,όσο και για τη διασπορά μιας κατανομής,όταν το **δείγμα** μας είναι **τυχαίο**.

Πρόταση 1.1.2. *Έστω $X_1, X_2, ..., X_n$ είναι ένα τυχαίο δείγμα από μια κατανομή με πυκνότητα πιθανότητας $f_1(x, \theta), \theta \in \Theta$ και $g(\theta) = \mu$, η μέση τιμή της κατανομής τότε ο δειγματικός μέσος $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$,είναι αμερόληπτος εκτιμητής του μ.*

Πρόταση 1.1.3. *Έστω $X_1, X_2, \ldots, X_n$ είναι ένα τυχαίο δείγμα από μια κατανομή $f_1(x, \theta), \theta \in \Theta$ και $g(\theta) = \sigma^2$ η διασπορά της κατανομής,τότε η δειγματική διασπορά $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$ είναι αμερόληπτος εκτιμητής του σ^2.*

1.2 Ανεξαρτησία Τυχαίων Μεταβλητών - Απο κοινού Συνάρτηση Πυκνότητας Πιθανότητας

Ορισμός 1.2.1. Αν έχουμε δύο τυχαίες μεταβλητές X, Y με συνάρτηση πυκνότητας πιθανότητας η κάθεμια αντίστοιχα $f_X(x) = P(X = x)$ και $f_Y(y) = P(Y = y)$. Ορίζουμε από κοινού συνάρτηση πυκνότητας πιθανότητας των τυχαίων μεταβλητών X,Y την συνάρτηση $f_{X,Y}(x, y) = P(X = x, Y = y)$

Παρατήρηση 1.2.1. Διακρίνουμε τις εξής ιδιότητες για την από κοινού συνάρτηση πυκνότητας πιθανότητας $f_{\underset{\sim}{X},\underset{\sim}{Y}}(x, y)$

1.$f_{X,Y}(x, y) \geq 0$

2.$\sum_x \sum_y f_{X,Y}(x,y) = 1$

Ορισμός 1.2.2. Θεωρούμε τις τυχαίες μεταβλητές X, Y. Αυτές λέγονται ανεξάρτητες όταν ισχύει ότι

$P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$ όπου A,B υποσύνολα του συνόλου των πραγματικών αριθμών.

Παρατήρηση 1.2.2. Διακρίνουμε τις εξής ιδιότητες για τις ανεξάρτητες τυχαίες μεταβλητές.

1.$f_{X,Y}(x,y) = P(X = x, Y = y) = P(X = x)P(Y = y) = f_{X(x)}f_{Y(y)}$

2.$E[f_X(x)f_Y(y)] = Ef_X(x)Ef_Y(y)$

3.$Var(X \pm Y) = Var(X) + Var(Y)$

1.3 Δεσμευμένη Κατανομή

Ορισμός 1.3.1. Από τη Θεωρία Πιθανοτήτων έχουμε ότι η δεσμευμένη πιθανότητα 2 ενδεχόμενα A, B ενός δειγματικού χώρου Ω είναι ίση με

$P(A|B) = \frac{P(A \cap B)}{P(B)}$

Ερμηνεύοντας την δεσμευμένη πιθανότητα, η πραγματοποίηση του ενδεχομένου A εξαρτάται από την πραγματοποίηση του ενδεχομένου B.

Ορισμός 1.3.2. Από τον ορισμό της δεσμευμένης πιθανότητας βλ. Ορισμό 1.3.1 αποδεικνύουμε εύκολα τον τύπο της δεσμευμένης κατανομής.

Έτσι για 2 τυχαίες μεταβλητές X,Y με συνάρτηση πυκνότητας πιθανότητας αντίστοιχα $f_X(x) = P(X = x)$ και $f_Y(y) = P(Y = y)$ προκύπτει ότι

$P(X = x|Y = y) = \frac{P(X=x,Y=y)}{PY=y} \Rightarrow P(X = x|Y = y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = f_{X|Y}(x|y)$

1.4 Συνάρτηση Ζημίας (Loss Function)-Συνάρτηση Κινδύνου(Risk Function)

Γενικά,η εκτίμηση της παραμετρικής παράστασης $g(\theta)$ από μια τιμή d,μετριέται από την **συνάρτηση ζημιάς (Loss function)** $L(d,\theta)$. Για την οποία ισχύουν

$$L(d,\theta) \geqslant 0 \text{ για όλα τα } \theta, d$$

και

$$L[g(\theta),\theta] = 0 \text{ για όλα τα } \theta$$

έτσι ώστε η ζημιά να είναι μηδέν όταν η παράμετρος εκτιμάται από τη σωστή τιμή.

Ορισμός 1.4.1. Η ακρίβεια ή μη-ακρίβεια,ενός εκτιμητή δ,μετριέται από την **συνάρτηση κινδύνου(risk function)**που ορίζεται ως

$$R(\delta,\theta) = E_\theta\{L[\delta(X),\theta]\}.$$

Το Τετραγωνικό Σφάλμα $L(d,\theta) = (d-\theta)^2$είναι μια συνάρτηση ζημίας. Οπότε μπορούμε να επαναδιατυπώσουμε τους παραπάνω ορισμούς αντικαθιστώντας την οποιαδήποτε συνάρτηση ζημιάς $L(d,\theta)$, με το τετραγωνικό σφάλμα.

Ορισμός 1.4.2. Μια πραγματική συνάρτηση $L(\cdot)$ ορισμένη στο σύνολο των πραγματικών αριθμών καλείται **bowl-shaped** (αντίστοιχα αυστηρά **bowl-shaped**),αν υπάρχει $t_0 \in \mathbb{R}$ τέτοιο ώστε η $L(t)$ να είναι φθίνουσα (αντίστοιχα,γνησίως φθίνουσα) για $t < t_0$ και αύξουσα (αντίστοιχα,γνησίως αύξουσα) για $t > t_0$.

1.5 ΑΟΕΔ εκτιμητές

Επειδή είναι γενικά δύσκολο να βρούμε τον βέλτιστο εκτιμητή στην κλάση όλων των εκτιμητών,περιοριζόμαστε αρχικά, σε αυτή των αμερόληπτων εκτιμητών.

Ορισμός 1.5.1. Η στατιστική συνάρτηση $T = T(\underset{\sim}{X})$ ονομάζεται **Αμερόληπτος Εκτιμητής Ελάχιστης Διασποράς**(ΑΟΕΔ) για το $g(\theta)$ εάν,

1.T αμερόληπτος , δηλαδή $E_\theta(T) = g(\theta), \forall\ \theta \in \Theta$.

2.$Var_\theta(T) \leqslant Var_\theta(T_1), \forall \in \Theta$ και για κάθε άλλο αμερόληπτο εκτιμητή T_1 του $g(\theta)$.

Από τον παραπάνω ορισμό ,φαίνεται ότι για να βρούμε ΑΟΕΔ εκτιμητή πρέπει να ελαττώσουμε όσον το δυνατόν περισσότερο τη διασπορά μίας στατιστικής συνάρτησης σε σχέση με την προς εκτίμηση ποσότητα,δηλαδή είναι επιθυμητό να βρούμε ένα κάτω φράγμα για τη διασπορά των αμερόληπτων εκτιμητών αυτής της ποσότητας.Αυτό το κάτω φράγμα μας προσφέρει το **Θεώρημα Cramer-Rao** το οποίο ισχύει όταν επαληθεύονται οι παρακάτω συνθήκες,

(I1)Ο παραμετρικός χώρος Θ είναι ανοικτό υποσύνολο του $\mathbb{R}$.

(I2)Το σύνολο $S = \{\underset{\sim}{x}; f_{\underset{\sim}{X}}(\underset{\sim}{x};\theta)\}$ δεν εξαρτάται από το θ.

(I3)$\int\limits_{\mathbb{R}^n} \frac{\partial}{\partial\theta} f_{\underset{\sim}{X}}(\underset{\sim}{x};\theta) d\underset{\sim}{x} = \frac{\partial}{\partial\theta} \int\limits_{\mathbb{R}^n} f_{\underset{\sim}{X}}(\underset{\sim}{x};\theta) d\underset{\sim}{x}, \forall\theta \in \Theta$

(I4)$\int\limits_{\mathbb{R}^n} T(\underset{\sim}{x}) \frac{\partial}{\partial\theta} f_{\underset{\sim}{X}}(\underset{\sim}{x};\theta) d\underset{\sim}{x} = \frac{\partial}{\partial\theta} \int\limits_{\mathbb{R}^n} T(\underset{\sim}{x}) f_{\underset{\sim}{X}}(\underset{\sim}{x};\theta) d\underset{\sim}{x}, \forall\theta \in \Theta$ και κάθε στατιστική συνάρτηση $T(\underset{\sim}{X})$.

(I5)Αν $I(\theta) = E_\theta(\frac{\partial}{\partial\theta} \ln f_{\underset{\sim}{X}}(\underset{\sim}{x};\theta))^2$,τότε $0 < I(\theta) < \infty, \forall\theta \in \Theta$.
Η ποσότητα $I(\theta)$ ονομάζεται **αριθμός** ή **μέτρο πληροφορίας Fisher**.

Θεώρημα 1.5.1. *(Έστω $\underset{\sim}{X} = (X_1, X_2, \ldots, X_n)$ ένα δείγμα με από κοινού πυκνότητα πιθανότητας $f_{\underset{\sim}{X}}(\underset{\sim}{x};\theta), \theta \in \Theta$.Εάν $T(\underset{\sim}{X})$ είναι στατιστική συνάρτηση με $E_\theta(T(\underset{\sim}{X})) = g(\theta), \forall\theta \in \Theta$ και ισχύουν οι συνθήκες (I1)-(I5),τότε*

$$Var_\theta(T(\underset{\sim}{X})) \geqslant \frac{(g'(\theta))^2}{I(\theta)}, \forall\theta \in \Theta.$$

Το κάτω φράγμα για την διασπορά των αμερόληπτων εκτιμητών του $g(\theta)$ ονομάζεται **Cramer-Rao Κάτω Φράγμα** (C.R.-Κ.Φ.), ενώ για τον υπολογισμό του αριθμού πληροφορίας Fisher χρησιμοποιούμε συνήθως κάποιες βοηθητικές ιδιότητες.

Ιδιότητες

1.$I(\theta) = -E_\theta(\frac{\partial^2}{\partial\theta}\ln f_{\underset{\sim}{X}}(\underset{\sim}{x};\theta))^2, \forall\theta \in \Theta.$

2.Αν το δείγμα $\underset{\sim}{X} = (X_1, X_2, \ldots, X_n)$ αποτελείται από ανεξάρτητες και τυχαίες μεταβλητές,όπου κάθε μια από τις X_i ακολουθεί μία κατανομή με πυκνότητα πιθανότητας,$f_{X_i}(x_i;\theta), i = 1, 2, \ldots, n$,τότε

$$I(\theta) = \sum_{i=1}^{n} I_i(\theta)$$

όπου $I_i(\theta) = E_\theta(\frac{\partial}{\partial\theta}\ln f_{X_i}(x_i;\theta))^2$.

3.Αν το δείγμα $\underset{\sim}{X} = (X_1, \ldots, X_n)$ είναι τυχαίο ,τότε

$$I(\theta) = nI_1(\theta)$$

όπου $I_1(\theta)$ είναι ο αριθμός πληροφορίας Fisher για κάθε μία από τις $X_1, X_2, \ldots, X_n$.

Η δυσκολία του Θεωρήματος Cramer-Rao βρίσκεται στη επαλήθευση των συνθηκών (I1)-(I5),η οποία άρεται όταν η οικογένεια του τυχαίου διανύσματος $\underset{\sim}{X}$ ανήκει στην **Μονοπαραμετρική Εκθετική Οικογένεια Κατανομών** (ΜΕΟΚ).

Ορισμός 1.5.2. Η οικογένεια κατανομών $\{f_{\underset{\sim}{X}}(\underset{\sim}{x};\theta), \theta \in \Theta\}$ ανήκει στην **Μονοπαραμετρική Εκθετική Οικογένεια Κατανομών** (ΜΕΟΚ) αν,

1.Το σύνολο $S = \{x; f_{\underset{\sim}{X}}(\underset{\sim}{x};\theta) > 0\}$ δεν εξαρτάται από το Θ.

2.$f_{\underset{\sim}{X}}(\underset{\sim}{x};\theta) = e^{A(\theta)+B(\underset{\sim}{x})+c(\theta)D(\underset{\sim}{x})}, \forall \underset{\sim}{x} \in S, \theta \in \Theta.$

Θεώρημα 1.5.2. *Αν το δείγμα* $\underset{\sim}{X} = (X_1, X_2, \ldots, X_n)$ *έχει κατανομή με πυκνότητα πιθανότητας* $f_{\underset{\sim}{X}}(\underset{\sim}{x};\theta)$ *η οποία ανήκει στην ΜΕΟΚ και η* $c(\theta)$ *(που εμφανίζεται στον τύπο της* $f_{\underset{\sim}{X}}(\underset{\sim}{x};\theta)$*) έχει συνεχή και μη μηδενική παράγωγο* $\forall\theta \in \Theta$*,τότε οι συνθήκες (Ι2),(Ι3) και (Ι4) του Θεωρήματος Cramer-Rao ισχύουν και η (Ι4) ισχύει για κάθε στατιστική συνάρτηση* $T = T(\underset{\sim}{X})$*.*

Η παρακάτω Πρόταση δίνει,ουσιαστικά,έναν τρόπο εύρεσης του ΑΟΕΔ εκτιμητή για μια παραμετρική συνάρτηση $g(\theta)$ και γραμμικούς συνδυασμούς αυτής.

Πρόταση 1.5.1. *Αν το δείγμα* $\underset{\sim}{X} = (X_1, X_2, \ldots, X_n)$ *έχει κατανομή με πυκνότητα πιθανότητας* $f_{\underset{\sim}{X}}(\underset{\sim}{x};\theta)$ *η οποία ανήκει στην ΜΕΟΚ (*$f_{\underset{\sim}{X}}(\underset{\sim}{x};\theta) = e^{A(\theta)+B(\underset{\sim}{x})+c(\theta)D(\underset{\sim}{x})}$*) και ισχύουν,*

a)Το σύνολο Θ *είναι ανοικτό υποσύνολο του* $\mathbb{R}$*.*

β)Το $c(\theta)$ *έχει συνεχή και μη μηδενική παράγωγο* $\forall\theta \in \Theta$*.*

γ)$0 < I(\theta) < \infty$*.*

Τότε,

1.Η στατιστική συνάρτηση $D(\underset{\sim}{X})$ *είναι ΑΟΕΔ εκτιμητής της* $g(\theta) = E_\theta(D(\underset{\sim}{X}))$*.*

2.Η στατιστική συνάρτηση $c_1 D(\underset{\sim}{X}) + c_2$ *με* c_1, c_2 *σταθερές* $c_1 \neq 0$ *είναι ΑΟΕΔ εκτιμητής της* $c_1 g(\theta) + c_2$*.*

Επίσης, ισχύει και η εξής Πρόταση.

Πρόταση 1.5.2. *Έστω ότι ισχύουν οι συνθήκες (Ι1),(Ι2),(Ι3) και (Ι5) του Θεωρήματος Cramer-Rao και η (Ι4) ισχύει για κάποια στατιστική συνάρτηση* $T(\underset{\sim}{X})$*,αμερόληπτο εκτιμητή του* $g(\theta)$*. Έστω,ακόμα,η παραμετρική συνάρτηση* $g(\theta)$ *είναι μη σταθερά (σαν συνάρτηση του* θ*) και η* $T(\underset{\sim}{X})$

επιτυγχάνει το C.R.-K.Φ.,δηλαδή

$$Var_\theta(T(\underset{\sim}{X})) = \frac{g'(\theta))^2}{I(\theta)}, \forall\theta \in \Theta,$$

τότε, $f_{\underset{\sim}{X}}(\underset{\sim}{x};\theta) = e^{A(\theta)+B(\underset{\sim}{x})+c(\theta)T(\underset{\sim}{x})}, \forall \underset{\sim}{x} \in S, \theta \in \Theta$,*δηλαδή η κατανομή του δείγματος* $\underset{\sim}{X}$ *ανήκει στην ΜΕΟΚ.*

Παρατήρηση 1.5.1. Οι Προτάσεις 1.5.1 και 1.5.2 συνεπάγονται το γεγονός ότι η εύρεση του εκτιμητή για κάποια παραμετρική συνάρτηση $g(\theta)$ είναι δυνατή με τη χρήση του Θεωρήματος Cramer-Rao αν και μόνο αν η κατανομή του δείγματος $\underset{\sim}{X} = (X_1, X_2, \ldots, X_n)$ ανήκει στην ΜΕΟΚ και η $g(\theta)$ έχει μια συγκεκριμένη μορφή $g(\theta) = E_\theta(D(\underset{\sim}{X}))$ ή κάποιος γραμμικός μετασχηματισμός της $E_\theta(D(\underset{\sim}{X}))$.

Όπως γίνεται εύκολα αντιληπτό από την παραπάνω Παρατήρηση η μέθοδος εύρεσης ΑΟΕΔ εκτιμητή με χρήση του Θεωρήματος Cramer-Rao (Θεώρημα 1.5.1) μας περιορίζει τόσο ως προς την οικογένεια του δείγματος,όσο και ως προς την μορφή των παραμετρικών συναρτήσεων για τις οποίες βρίσκουμε ΑΟΕΔ εκτιμητές, οπότε χρειάζεται μια διαφορετική μέθοδος από την προηγούμενη η οποία να μην έχει αυτού του είδους τα προβλήματα.Αρχικά,εισάγουμε δύο έννοιες (Επάρκεια και Πληρότητα) οι οποίες μας βοηθούν προς αυτήν την κατεύθυνση.

1.6 Επάρκεια

Ορισμός 1.6.1. Έστω το δείγμα $\underset{\sim}{X} = (X_1, X_2, \ldots, X_n)$ έχει κατανομή με πυκνότητα πιθανότητας $f_{\underset{\sim}{X}}(\underset{\sim}{x};\theta), \theta \in \Theta$ τότε η στατιστική συνάρτηση $T(\underset{\sim}{X})$ ονομάζεται **επαρκής** αν η δεσμευμένη κατανομή του $\underset{\sim}{X}|T(\underset{\sim}{X}) = t$ δεν εξαρτάται από το θ για κάθε τιμή t για την οποία μπορεί να οριστεί η δεσμευμένη κατανομή.

Ένας τρόπος εύρεσης μιας επαρκούς στατιστικής συνάρτησης,εκτός του ορισμού,δίνεται απο την παρακάτω πρόταση,η οποία αναφέρεται και ως **παραγοντικό κριτήριο των Neyman-Fisher**.

Θεώρημα 1.6.1 (παραγοντικό κριτήριο των Neyman-Fisher). *Η στατιστική συνάρτηση $T(\underset{\sim}{X})$ είναι επαρκής αν και μόνο αν $f_{\underset{\sim}{X}}(\underset{\sim}{x};\theta) = q(T(\underset{\sim}{X});\theta)h(\underset{\sim}{x}), \forall \underset{\sim}{x}$ και $\theta \in \Theta$,όπου q και h είναι συναρτήσεις.*

Παρατήρηση 1.6.1. Ισχύουν οι παρακάτω ιδιότητες για τις επαρκείς στατιστικές συναρτήσεις.

1)Το δείγμα $\underset{\sim}{X} = (X_1, \dots, X_n)$ είναι τετριμμένα επαρκής στατιστική συνάρτηση.

2)Η στατιστική συνάρτηση $T(\underset{\sim}{X}) = (X_{(1)}, \dots, X_{(n)})$ είναι επαρκής ,όπου οι $X_{(i)}, i = 1, \dots, n$ είναι οι διατεταγμένες παρατηρήσεις.

3)Έστω $T_1 = T_1(\underset{\sim}{X})$ επαρκής στατιστική συνάρτηση και $T_2 = K(T_1)(\underset{\sim}{X})$,όπου $K(\cdot)$ είναι "1-1" συνάρτηση,τότε η στατιστική συνάρτηση $T_2(\underset{\sim}{X})$ είναι επαρκής.

Συνήθως,όταν μιλάμε για επαρκή στατιστική συνάρτηση αναφερόμαστε στην ελάχιστη επαρκή.

Ορισμός 1.6.2. Ελάχιστη επαρκής στατιστική συνάρτηση είναι μια επαρκής στατιστική συνάρτηση η οποία προέρχεται από την μεγαλύτερη δυνατή "σύμπτηξη"(δηλ. έχει την μικρότερη δυνατή διάσταση).

Παρατήρηση 1.6.2. Σχεδόν πάντα,η διάσταση της παραμετρικής συνάρτησης $g(\theta)$ συμπίπτει με την διάσταση της ελάχιστης επαρκούς στατιστικής συνάρτησης.

Στο παρακάτω Θεώρημα χρησιμοποιείται η έννοια της επάρκειας στη βελτίωση εκτιμητών.

Θεώρημα 1.6.2. *Έστω $T = T(\underset{\sim}{X})$ είναι μια επαρκής στατιστική συνάρτηση και $S = S(\underset{\sim}{X})$ είναι εκτιμητής της παραμετρικής συνάρτησης $g(\theta)$.Θέτουμε $S^* = E_\theta(S|T)$.Τότε,*

1.Η S^* είναι στατιστική συνάρτηση.

2.$E_\theta(S^*) = E_\theta(S), \forall\theta \in \Theta$,έτσι αν S είναι αμερόληπτος εκτιμητής για την $g(\theta)$,τότε S^* είναι αμερόληπτος εκτιμητής για την $g(\theta)$.

3.$Var_\theta(S^*) \leqslant Var_\theta(S), \forall\theta \in \Theta$ και ισχύει αυστηρή ανισότητα,εκτός εάν S είναι συνάρτηση της στατιστικής συνάρτησης T,οπότε $S^* = S$.

4.ΜΤΣ$(S^*, \theta) \leqslant$ΜΤΣ$(S, \theta), \forall\theta \in \Theta$ και ισχύει αυστηρή ανισότητα εκτός εάν S είναι συνάρτηση της στατιστικής συνάρτησης T,οπότε $S^* = S$.

Επομένως,αν S είναι ένας εκτιμητής της $g(\theta)$ ο οποίος δεν είναι συνάρτηση της επαρκούς στατιστικής συνάρτησης T,τότε ο S είναι μη αποδεκτός και βελτιώνεται από τον $S^* = E_\theta(S|T)$ που ονομάζεται βελτίωση του S κατά Rao-Blackwell ή **Rao-Blackwell βελτίωση του S**.

Παρατήρηση 1.6.3. Έστω T_1 και T_2 είναι επαρκείς στατιστικές συναρτήσεις και S είναι αμερόληπτος εκτιμητής της $g(\theta)$.Τότε $S_1^* = E_\theta(S|T_1)$ είναι η Rao-Blackwell βελτίωση του S μέσω της T_1 και $S_2^* = E_\theta(S|T_2)$ είναι η η Rao-Blackwell βελτίωση του S,μέσω της T_2.Όμως,μέσω του Θεωρήματος 1.6.2 δεν μπορούμε να συγκρίνουμε αυτές τις δύο βελτιώσεις.Η έννοια της **πληρότητας** θα βοηθήσει σε αυτή την σύγκριση.

1.7 Πληρότητα

Ορισμός 1.7.1. Η στατιστική συνάρτηση $T = T(\underset{\sim}{X})$ ονομάζεται **πλήρης**,αν ισχύει η ακόλουθη σχέση,

$$E_\theta(\phi(T)) = 0, \forall \in \Theta \Rightarrow \phi(t) = 0$$

για κάθε δυνατή τιμή t της T,δηλαδή $\phi(T) = 0$.

Θεώρημα 1.7.1. *Έστω $T = T(\underset{\sim}{X})$ είναι επαρκής και πλήρης στατιστική συνάρτηση και S είναι ένας αμερόληπτος εκτιμητής της $g(\theta)$.Τότε $S^* = E_\theta(S|T)$ είναι μοναδικός ΑΟΕΔ εκτιμητής της $g(\theta)$.*

Άρα με τη βοήθεια του Θεωρήματος των Lehmann-Scheffé μπορούμε να βρούμε ΑΟΕΔ εκτιμητή με την χρήση επαρκούς και πλήρους στατιστικής συνάρτησης και μάλιστα,αν υπάρχει αυτός ο ΑΟΕΔ εκτιμητής,είναι και μοναδικός.

Πόρισμα 1.7.1. *Έστω $T = T(\underset{\sim}{X})$ είναι επαρκής και πλήρης στατιστική συνάρτηση και S είναι ένας αμερόληπτος εκτιμητής της $g(\theta)$,ο οποίος είναι συνάρτηση της επαρκούς και πλήρους T.Τότε S είναι μοναδικός ΑΟΕΔ εκτιμητής της $g(\theta)$.*

Όπως καταλαβαίνουμε,σε αυτή την μεθοδολογία είναι σημαντική η εύρεση μιας επαρκούς και πλήρους στατιστικής συνάρτησης και μέσω του ορισμού δεν είναι πάντα εύκολο,αλλά αν η κατανομή του δείγματος $\underset{\sim}{X}$ ανήκει στην **Πολυπαραμετρική Εκθετική Οικογένεια Κατανομών**(ΠΕΟΚ) τα πράγματα απλοποιούνται.

Ορισμός 1.7.2. Η οικογένεια κατανομών $\{f_{\underset{\sim}{X}}(\underset{\sim}{x};\theta), \theta \in \Theta\}$ ανήκει στην Πολυπαραμετρική Εκθετική Οικογένεια Κατανομών(ΠΕΟΚ),διάστασης k,αν

1.Το σύνολο $S = \{\underset{\sim}{x}; f_{\underset{\sim}{X}}(\underset{\sim}{x};\theta) > 0\}$ δεν εξαρτάται από το θ.

2.$f_{\underset{\sim}{X}}(\underset{\sim}{x};\theta) = e^{A(\theta)+B(\underset{\sim}{x})+\sum_{j=1}^{k} c_j D_j(\underset{\sim}{x})}, \forall \underset{\sim}{x} \in S, \theta \in \Theta.$

Παρατήρηση 1.7.1. Η ΠΕΟΚ διάστασης 1 συμπίπτει με την ΜΕΟΚ.

Πρόταση 1.7.1. *Έστω ότι το δείγμα $\underset{\sim}{X} = (X_1, X_2, \ldots, X_n)$ έχει κατανομή η οποία ανήκει στην ΠΕΟΚ διάστασης k,τότε ισχύουν τα εξής:*

1.Η στατιστική συνάρτηση $T(\underset{\sim}{X}) = (D_1(\underset{\sim}{X}), D_2(\underset{\sim}{X}), \ldots, D_k(\underset{\sim}{X}))$ είναι επαρκής.

2.Αν το πεδίο τιμών του διανύσματος $(c_1(\theta), c_2(\theta), \ldots, c_k(\theta))$ περιέχει ένα ανοικτό υποσύνολο του $\mathbb{R}^k$,τότε η $T(\underset{\sim}{X})$ είναι πλήρης.

Το παρακάτω Θεώρημα,γνωστό και ως **Θεώρημα Basu** ,πιστοποιεί και μια άλλη χρήση της επάρκειας και της πληρότητας,αυτής της απόδειξης ανεξαρτησίας μεταξύ στατιστικών συναρτήσεων(δηλαδή τυχαίων μεταβλητών).

Θεώρημα 1.7.2. *Έστω $T(\underset{\sim}{X})$ επαρκής και πλήρης στατιστική συνάρτηση και $S(\underset{\sim}{X})$ είναι μία στατιστική συνάρτηση,η κατανομή της οποίας δεν εξαρτάται από το θ,τότε οι στατιστικές συναρτήσεις $T(\underset{\sim}{X})$ και $S(\underset{\sim}{X})$ είναι ανεξάρτητες.*

1.8 Συνέπεια

Ορισμός 1.8.1. Έστω $T_n = T(X_1, X_2, \ldots, X_n), n = 1, 2, \ldots$ ένας εκτιμητής της παραμετρικής συνάρτησης $g(\theta)$.Τότε ο εκτιμητής T_n ονομάζεται **συνεπής** αν

$$\lim_{n\to\infty} P(|T_n - g(\theta)| > \varepsilon) = 0, \forall \varepsilon > 0.$$

Η παρακάτω πρόταση δίνει ικανές συνθήκες έτσι ώστε ένας εκτιμητής για την $g(\theta)$ να είναι συνεπής.

Πρόταση 1.8.1. *Έστω ότι ο εκτιμητής T_n ικανοποιεί τις παρακάτω συνθήκες,*

1.$Var_\theta T_n \to 0, n \to +\infty$

2.$b(T_n, \theta) = E_\theta T_n - g(\theta) \to 0, n \to +\infty$

Τότε ο T_n είναι συνεπής εκτιμητής της παραμετρικής συνάρτησης $g(\theta)$.

1.9 Εκτίμηση με την Μέθοδο Μεγίστης Πιθανοφάνειας

Ορισμός 1.9.1. Θεωρούμε ότι το δείγμα $\underset{\sim}{X} = (X_1, X_2, \ldots, X_n)$ έχει συνάρτηση πυκνότητας πιθανότητας $f_{\underset{\sim}{X}}(\underset{\sim}{x}; \theta), \theta \in \Theta$,τότε **η συνάρτηση πιθανοφάνειας (ή απλά πιθανοφάνεια)** του θ ορίζεται από την σχέση,

$$L(\theta) = L(\theta|\underset{\sim}{x}) = f_{\underset{\sim}{X}}(\underset{\sim}{x}; \theta)$$

Αναφέρουμε παρακάτω τον ορισμό του Εκτιμητή Μεγίστης Πιθανοφάνειας (Ε.Μ.Π.).

Ορισμός 1.9.2. Ο εκτιμητής $\hat{\theta} = \hat{\theta}(\underset{\sim}{X})$,που ικανοποιεί τη σχέση

$$L(\hat{\theta}) = \sup_{\theta \in \Theta} L(\theta)$$

ονομάζεται **Εκτιμητής Μεγίστης Πιθανοφάνειας**(Ε.Μ.Π.) του θ.

Παρατήρηση 1.9.1. Από τον προηγουμένο ορισμό φαίνεται ότι ο Ε.Μ.Π. του θ είναι εκείνη η τιμή του θ,η οποία μεγιστοποιεί τη συνάρτηση πιθανοφάνειας.Επειδή η συνάρτηση $\ln x$ είναι γνησίως αύξουσα συνάρτηση του x,η τιμή του θ που μεγιστοποιεί την $L(\theta)$ είναι η ίδια με αυτήν που μεγιστοποιεί την $\ln L(\theta)$.Συνήθως ακολουθούμε αυτήν την διαδικασία όταν το μέγιστο μπορεί να βρεθεί με παραγώγιση.

Παρατήρηση 1.9.2. 1.Η μέθοδος μεγίστης πιθανοφάνειας ισχύει και για το διάνυσμα $\underset{\sim}{\theta} = (\theta_1, \theta_2, \ldots, \theta_k)$
2.Είναι δυνατόν ο εκτιμητής $\hat{\theta}$ να μην μπορεί να βρεθεί σε αναλυτική μορφή,τότε η τιμή του θ για την οποία επιτυγχάνεται η μεγιστοποιήση της $L(\theta)$ βρίσκεται με μεθόδους αριθμητικής ανάλυσης.
3.Ορισμένες φορές υπάρχουν "παθολογικές καταστάσεισ" με την έννοια ότι είτε δεν υπάρχει τιμή του θ η οποία να μεγιστοποιεί τη συνάρτηση πιθανοφάνειας,είτε υπάρχουν περισσότερα μέγιστα για την $L(\theta)$ και συνεπώς περισσότεροι του ενός Ε.Μ.Π.

Παρατήρηση 1.9.3. Σε αυτό το σημείο αναφέρουμε κάποιες γενικές ιδιότητες των Ε.Μ.Π.
1.Από τον Ορισμό 1.9.2 προκύπτει ότι ο Ε.Μ.Π(αν υπάρχει) παίρνει τιμές μέσα στον παραμετρικό χώρο Θ.
2.Αν ο Ε.Μ.Π. του θ είναι μοναδικός,τότε είναι συνάρτηση της επαρκούς στατιστικής συνάρτησης.
3.Αν $\hat{\theta} = \hat{\theta}(\underset{\sim}{X})$ είναι Ε.Μ.Π. του θ,τότε ο Ε.Μ.Π. της παραμετρικής συνάρτησης $g(\theta)$ είναι ο $g(\hat{\theta})$.
4.Οι Ε.Μ.Π. είναι(υπό ορισμένες συνθήκες) **συνεπείς** εκτιμητές (βλ. Ορισμό 1.8.1).

Παρατήρηση 1.9.4. Οι Ε.Μ.Π έχουν(υπό ορισμένες συνθήκες) κάποιες ασυμπτωτικές ιδιότητες.Αν $X_1, X_2, \ldots, X_n$ είναι ένα τυχαίο δείγμα από κατανομή με πυκνότητα πιθανοτητας $f_1(x;\theta)$ και συμβολίσουμε με $\hat{\theta}$ τον Ε.Μ.Π. του θ,τότε
1.Η κατανομή του $\hat{\theta}$ είναι κατά προσεγγίση $(n \to +\infty)$ η κανονική κατανομή δηλαδή

$$\hat{\theta} \sim \mathrm{N}(\theta, \frac{1}{I(\theta)})$$

όπου $I(\theta)$ είναι ο αριθμός πληροφορίας του Fisher.

2.Ο $\hat{\theta}$ είναι ασυμπτωτικά αποτελεσματικός εκτιμητής, δηλαδή αν κάποιος άλλος εκτιμητής του θ,έστω S_n,έχει κατά προσέγγιση κανονική κατανομή $N(\theta^2, \sigma^2(\theta))$,τότε υπό ορισμένες συνθήκες $\sigma^2(\theta) \geq \frac{1}{I(\theta)}$.

Οι παραπάνω ιδιότητες των Ε.Μ.Π. συνεπάγονται ότι ο $\hat{\theta}$ είναι ασυμπτωτικά ΑΟΕΔ για το θ,δηλαδή αν υπάρχουν ΑΟΕΔ και Ε.Μ.Π. για κάποια $g(\theta)$,τότε αυτοί δεν διαφέρουν ασυμπτωτικά.

1.10 Εκτίμηση με την Μέθοδο των Ροπών

Ορισμός 1.10.1. Θεωρούμε την τυχαία μεταβλητή Χ, τότε ορίζουμε ροπή r-τάξης (Ε.Μ.Ρ.) τη συνάρτηση $\mu_r = (E(X^r))$. Για το τυχαίο δείγμα $\underset{\sim}{X} = (X_1, X_2, \ldots, X_n)$ ορίζουμε δειγματική ροπή r-τάξης τη συνάρτηση $m_r = \frac{1}{n}\sum_{i=1}^{n} X_i^r$

Παρατήρηση 1.10.1. Σε αυτό το σημείο αναφέρουμε κάποιες γενικές ιδιότητες της δειγματικής ροπής r-τάξης

Αν διαθέτουμε ένα τυχαίο δείγμα $\underset{\sim}{X} = (X_1, X_2, \ldots, X_n)$ όπου X_i με $\imath = 1, 2, \ldots, n$ τότε προκύπτει ότι:

1.Από τον Ορισμό της μέσης τιμής έχουμε ότι $E(m_r) = \frac{1}{n}\sum_{i=1}^{n} E(X_i^r) = \mu_r$ όπου συμπεραίνουμε ότι η m_r είναι **αμερόληπτος εκτιμητής** για τη μ_r (βλ. Ορισμό 1.1.3)

2.Η δειγματική ροπή είναι **συνεπής** εκτιμητής (βλ. Ορισμό 1.8.1).

Παρατήρηση 1.10.2. 1. Οι Ε.Μ.Ρ. είναι εύκολοι στον υπολογισμό.

2. Οι Ε.Μ.Ρ. είναι(υπό ορισμένες συνθήκες) **συνεπείς** εκτιμητές (βλ. Ορισμό 1.8.1).

3. Οι Ε.Μ.Ρ. δεν είναι κατ΄ ανάγκη συνάρτηση της επαρκούς στατιστικής συνάρτησης (βλ. Όρισμό 1.6.1)

4. Η λυση του συστήματος δεν είναι κατ΄ ανάγκη μοναδική και οι τιμές που μπορεί να είναι εκτός του συνόλου των τιμών της εκτιμώμενης παραμέτρου.

Παρατήρηση 1.10.3. Για τον υπολογισμό των Ε.Μ.Ρ. ακουλουθούμε τα εξής βήματα: για τυχαίο δείγμα $\underset{\sim}{X} = (X_1, X_2, \ldots, X_n)$ με συνάρτηση πυκνότητας πιθανότητας $f_{\underset{\sim}{X}}(\underset{\sim}{x};\theta), \theta \in \Theta$ και έστω $\underset{\sim}{\theta} = (\theta_1, \theta_2, \ldots, \theta_k)$

Αρχικά υπολογίζουμε k ροπές της κατανομής, συνήθως τις πρώτες k, δηλαδή $\mu_1, \mu_2, \ldots, \mu_k$

$\mu_1 = \mu_1(\theta_1, \theta_2, \ldots, \theta_k)$

$\mu_2 = \mu_2(\theta_1, \theta_2, \ldots, \theta_k)$

$\mu_3 = \mu_3(\theta_1, \theta_2, \ldots, \theta_k)$

$\vdots$

$\mu_k = \mu_k(\theta_1, \theta_2, \ldots, \theta_k)$

Κατόπιν θέτουμε $\mu_1(\theta_1, \theta_2, \ldots, \theta_k) = \frac{1}{n}\sum_{i=1} X_i = m_1$

$\mu_2(\theta_1, \theta_2, \ldots, \theta_k) = \frac{1}{n}\sum_{i=1} X_i^2 = m_2$

$\mu_3(\theta_1, \theta_2, \ldots, \theta_k) = \frac{1}{n}\sum_{i=1} X_i^3 = m_3$

$\vdots$

$\mu_k(\theta_1, \theta_2, \ldots, \theta_k) = \frac{1}{n}\sum_{i=1} X_i^k = m_k$

Προκύπτει λοιπόν ένα σύστημα k εξισώσεων με k αγνώστους η λύση του οποίου μας δίνει τους εκτιμητές μεθόδου ροπών των παραμέτρων $\theta_1, \theta_2, \ldots, \theta_k$, οι οποίοι συμβολιζόμενοι είναι $\widehat{\theta_1}, \widehat{\theta_2}, \widehat{\theta_3}, \ldots, \widehat{\theta_k}$

1.11 Εκτιμητές Bayes και minimax

Η εκτίμηση κατά Bayes γίνεται από μια διαφορετική σκοπιά σε σχέση με ότι έχουμε αντιμετωπίσει μέχρι τώρα,που αντιλαμβανόμασταν το θ απλά σαν ένα πραγματικό αριθμό χωρίς καμιά ιδιότητα.Αν π.χ. θεωρήσουμε μια βιομηχανία η οποία παράγει ηλεκτρικούς λαμπτήρες,τότε ο χρόνος αυτών των

λαμπτήρων ακολουθεί την εκθετική κατανομή με άγνωστη παράμετρο θ και αυτή η παράμετρος εκφράζει τον μέσο χρόνο ζωής των λαμπτήρων. Επομένως δεν πρέπει να αναμένουμε ούτε ¨μεγάλεσ' τιμές για το θ,αλλά ούτε και ¨μικρέσ¨.Δηλαδή σε σχέση με το πρόβλημα και την εμπειρία που διαθέτουμε πρέπει να δώσουμε μια διαφορετική βαρύτητα στις διάφορες τιμές του θ για να εκμεταλλευτούμε αυτήν την εμπειρία ώστε να δώσουεμε καλύτερη εκτίμηση για το θ.
Οπότε,θεωρούμε το θ σαν μια τυχαία μεταβλητή με πυκνότητα πιθανότητας $\pi(\theta), \theta \in \Theta$,και τις εξής ιδιότητες,

$$(i)\pi(\theta) \geq 0, \forall\theta \in \Theta \text{ και } (ii) \int\limits_{\Theta} \pi(\theta)d\theta = 1 \text{ (ή} \sum_{\theta} \pi(\theta) = 1)$$

Η συνάρτηση $\pi(\theta)$ ονομάζεται **εκ των προτέρων κατανομή** του θ και εκφράζει την είτε την προσωπική μας αντίληψη για την πιθανή τιμή του θ,είτε συνοψίζει κάποιες εκ των προτέρων(δηλ. πριν την συλλογή δεδομένων)πληροφορίες για το θ.Θεωρούμε μια **συνάρτηση κινδύνου** $L(t, \theta)$ και προσπαθούμε να ελαχιστοποιήσουμε τη **συνάρτηση κινδύνου** $R(T, \theta) = \mathbb{E}_\theta(L(T(\underset{\sim}{X}), \theta)$ Επειδή έχουμε θεωρήσει ότι το θ είναι μια τυχαία μεταβλητή,προφανώς,η συνάρτηση κινδύνου είναι και αυτή μία τυχαία μεταβλητή,επομένως είναι λογικό σε αυτή την περίπτωση,να προσπαθούμε να ελαχιστοποιήσουμε την μέση τιμή της,δηλαδή την

$$BR(T) = \mathbb{E}(R(T, \theta)) = \int\limits_{\Theta} R(T, \theta)\pi(\theta)d\theta$$

η οποία ονομάζεται κίνδυνος Bayes του εκτιμητή T.Συνεπώς βέλτιστος εκτιμητής είναι εκείνος που ελαχιστοποιεί τον κίνδυνο Bayes,οπότε καταλήγουμε στον εξής ορισμό για τον εκτιμητή Bayes

Ορισμός 1.11.1. Ο εκτιμητής $T^* = T^*(\underset{\sim}{X})$ ονομάζεται εκτιμητής Bayes του $g(\theta)$,ως προς τη συνάρτηση ζημίας $L(t, \theta)$ και την εκ των προτέρων κατανομή $\pi(\theta)$ αν,

$$\int\limits_{\Theta} R(T^*, \theta)\pi(\theta)d\theta \leq \int\limits_{\Theta} R(T, \theta)\pi(\theta)d\theta$$

για κάθε εκτιμητή $T = T(\underset{\sim}{X})$.

Συνήθως,για να υπολογίσουμε αυτόν τον εκτιμητή Bayes πρέπει να βρούμε πρώτα την **εκ των υστέρων κατανομή** του θ

$$\pi(\theta|\underset{\sim}{x}) = \frac{f(\underset{\sim}{x};\theta)\pi(\theta)}{f(\underset{\sim}{x})}$$

όπου $f_(\underset{\sim}{x}) = \int\limits_{\Theta} f(\underset{\sim}{x};\theta)\pi(\theta)d\theta$.Η εκ των υστέρων κατανομή συνοψίζει την πληροφορία για το θ μετά την συλλογή των δεδομένων και έχει τις ιδιότητες της συνάρτησης πυκνότητα πιθανότητας.

Παρατήρηση 1.11.1. Είναι σημαντικό να τονίσουμε,σε αυτό το σημείο,ότι δεν μας ενδιαφέρει ιδιαίτερα η ακριβής συνάρτηση $\pi(\theta|\underset{\sim}{x})$,αλλά η μορφή της εκ των υστέρων κατανομής για την οποία διαπιστώνουμε,συνήθως,ότι ακολουθεί κάποια από τις γνωστές κατανομές.

Στο επόμενο θεώρημα δίνουμε έναν διαφορετικό τρόπο υπολογισμού του εκτιμητή Bayes.

Θεώρημα 1.11.1. *Για $\underset{\sim}{X} = \underset{\sim}{x}$ ο εκτιμητής Bayes $T^* = T^*(\underset{\sim}{X})$ της παραμετρικής συνάρτησης $g(\theta)$ ως προς τη συνάρτηση $L(t,\theta)$ και την εκ των προτέρων κατανομή $\pi(\theta)$ έχει τιμή $T^*(\underset{\sim}{x}) = t^*$,όπου t^* είναι η τιμή του t που ελαχιστοποιεί τη συνάρτηση*

$$h^*(t) = \int\limits_{\Theta} L(t,\theta)\pi(\theta|\underset{\sim}{x})d\theta$$

Αν επιπλέον,η συνάρτηση ζημίας είναι το τετραγωνικό σφάλμα,δηλαδή $L(t,\theta) = (t - g(\theta))^2$,τότε η εύρεση του εκτιμητή Bayes γίνεται πιο απλά,όπως φαίνεται και στο παρακάτω Θεώρημα.

Θεώρημα 1.11.2. *Έστω ότι η συνάρτηση ζημίας για την εκτίμηση του $g(\theta)$ είναι το τετραγωνικό σφάλμα $L(t,\theta) = (t - g(\theta))^2$. Τότε για $\underset{\sim}{X} = \underset{\sim}{x}$ ο εκτιμητής Bayes $T^* = T^*(\underset{\sim}{X})$ της παραμετρικής συνάρτησης $g(\theta)$ έχει τιμή $T^*(\underset{\sim}{x}) = E_\theta(g(Y))$,όπου Y είναι μια τυχαία μεταβλητή με κατανομή την εκ των υστέρων κατανομή $\pi(\theta|\underset{\sim}{x})$.*

1.12 Θεώρημα Μετασχηματισμού

Θεώρημα 1.12.1. *Έστω X μια συνεχής τυχαία μεταβλητή με συνάρτηση πυκνότητας πιθανότητας $f_X(x)$. Θέτουμε $S = \{x : f_X(x) > 0\}$.*

Υποθέτουμε ότι,

(i) $y = h(x)$ *είναι ένας αμφιμονοσήμαντος (ένα -προς -ένα) μετασχηματισμός (μετρήσιμη συνάρτηση) που απεικονίζει το σύνολο* S *σε ένα σύνολο* T *των* y.

(ii) η αντίστροφη συνάρτηση $x = h^{-1}(y)$ *είναι παραγωγίσιμη και η παράγωγος της συνεχής και μη μηδενική για κάθε* $y \in T$.

Τότε η τυχαία μεταβλητή $Y = h(X)$ *είναι συνεχής με συνάρτηση πυκνότητας πιθανότητας*

$$f_Y(y) = \begin{cases} f_X[h^{-1}(y)] \left| \frac{dh^{-1}(y)}{dy} \right| & , y \in T \\ 0 & , \text{αλλού} \end{cases}$$

όπου $|\cdot|$ *σημαίνει την απόλυτο τιμή της συνάρτησης.*

1.13 Κεντρικό Οριακό Θεώρημα

Θεώρημα 1.13.1. *Έστω* $X_1, X_2, \ldots, X_n$ *ανεξάρτητα ταυτοτικά κατανεμημένες τυχαίες μεταβλητές με μέσο* μ *και διασπορά* σ^2.

Έστω οι συναρτήσεις $G_n(x) = P[\frac{\sqrt{n}(X-\mu)}{\sigma} \leqslant x]$ *και* $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-t^2}{2}} dt$

Τότε $G_n(x) \longrightarrow \Phi(x)$ *για* $n \to \infty$

Κεφάλαιο 2

Αρνητική Διωνυμική Κατανομή

2.1 Συνάρτηση πυκνότητας πιθανότητας

Ορισμός 2.1.1. Η τυχαία μεταβλητή X ακολουθεί την αρνητική διωνυμική κατανομή (Negative Binomial) με παραμέτρους $0 \leq p \leq 1$ και $k \geq 1$, αν η πυκνότητα πιθανότητας της δίνεται από τον τύπο,

$$f(x;k,p) = \binom{x+k-1}{x} p^k q^x \tag{2.1}$$

όπου

$$\binom{x+k-1}{x} = \frac{\Gamma(k+x)}{\Gamma(k)x!}, \Gamma(k) = (k-1)! \tag{2.2}$$

Παρατήρηση 2.1.1. Το διωνυμικό πείραμα τύχης που περιγράφει την αρνητική διωνυμική κατανομή επαναλαμβάνεται μέχρι να εμφανισθούν οι πρώτες k επιτυχίες, όπου $k \geq 1$. Για να συμβεί το παραπάνω πρέπει το πείραμα τύχης να επαναληφθεί k φορές συν το πλήθος των αποτυχιών πριν την k-οστή επιτυχία. Σε κάθε επανάληψη του πειράματος έχουμε στη διάθεσή μας 2 πιθανά αποτελέσματα, επιτυχία $\rightarrow S$ και αποτυχία $\rightarrow F$.

Ο δειγματικός χώρος Ω του πειράματος τύχης περιλαμβάνει όλες τις $(k+x)$-σειρές των S,F έτσι ώστε η k-οστή επιτυχία να είναι στο τέλος της σειράς. Η τυχαία μεταβλητή X παριστά το πλήθος των αποτυχιών πριν την k-οστή επιτυχία ενώ παίρνει διακριτές τιμές, X=0,1,2,.... Μια τέτοια τυχαία μεταβλητή X ονομάζεται αρνητική διωνυμική, το p δηλώνει την πιθανότητα επιτυχίας ενώ το $q = 1-p$ την πιθανότητα αποτυχίας.

Παρατήρηση 2.1.2. Αποδεικνύεται ότι η $f(x;k,p) = \binom{x+k-1}{x}p^k q^x$ είναι συνάρτηση πυκνότητας πιθανότητας επειδή ισχύουν οι ιδιότητες

1. Η $f(x;k,p)$ είναι μη αρνητική, $f(x;k,p) \geq 0$, $\forall x \in \{0,1,2,\dots\}$.
2. $$\sum_{x=0}^{\infty} \binom{x+k-1}{x} q^x = (1-q)^{-k} = p^{-k} \tag{2.3}$$

   ενώ πολλαπλασιάζοντας και τα 2 μέλη με p^k προκύπτει ότι

   $$\sum_{x=0}^{\infty} f(x) = 1 \tag{2.4}$$

Παρατήρηση 2.1.3. Το γεγονός $\{X = x\}$ σημαίνει ότι εμφανίστηκαν k επιτυχίες, από τις από τις οποίες μια στην τελευταία θέση και ότι εμφανίστηκαν x αποτυχίες πριν από την τελευταία επιτυχία.

Λόγω ανεξαρτησίας, βλ. Ορισμό 1.2.2, η πιθανότητα να συμβεί ένα τέτοιο γεγονός είναι

$$p^{k-1}q^x p = p^k q^x \tag{2.5}$$

Το πλήθος των δυνατών περιπτώσεων εμφάνισης του γεγονότος $\{X = x\}$ είναι ίσο με τον αριθμό επιλογής x θέσεων από $x+k-1$ θέσεις για την τοποθέτηση των πρώτων $k-1$ επιτυχιών και δίνεται από τον τύπο $\binom{x+k-1}{x}$.

Παρατήρηση 2.1.4. Για $k = 1$ προκύπτει η **Γεωμετρική Κατανομή** (ή *Pascal*) που δίνεται από τον τύπο

$$f(x) = pq^x \tag{2.6}$$

Πρόταση 2.1.1. *Έστω $X \sim NB(k,p)$ τότε η $EX = k\left(\frac{1}{p}-1\right)$ και $VarX = k\frac{1}{p}\left(\frac{1}{p}-1\right)$*

Απόδειξη. Για τη μέση τιμή έχουμε

$$EX = \sum_{x=0}^{\infty} x\binom{x+k-1}{x}q^x p^k = \sum_{x=1}^{\infty} x\frac{(x+k-1)!}{x!(k-1)!}q^x p^k = \sum_{x=1}^{\infty} \frac{(x+k-1)!}{(x+1)!(k-1)!}q^x p^k =$$
$$\sum_{x=0}^{\infty} \frac{(x+k-1)!}{x!(k-1)!}q^{x+1}p^k = k\frac{q}{p}\sum_{x=0}^{\infty}\binom{x+k}{x}q^x p^{k+1} = \frac{kq}{p} = k\frac{1-p}{p} = k\left(\frac{1}{p}-1\right) \tag{2.7}$$

Επίσης για τον υπολογισμό της διασποράς έχουμε

$$\sigma^2 = VarX = EX^2 - (EX)^2 = E(X(X-1)) + EX - (EX)^2 \tag{2.8}$$

Όπου υπολογίζοντας την τιμή $E(X(X-1))$ παρακάτω

$$E(X(X-1)) = \sum_{x=0}^{\infty} x(x-1)\binom{x+k-1}{x}q^x p^k = \sum_{x=2}^{\infty} x(x-1)\frac{(x+k-1)!}{x!(k-1)!}q^x p^k =$$
$$\sum_{x=2}^{\infty}\frac{(x+k-1)!}{(x-2)!(k-1)!}q^x p^k = \sum_{x=0}^{\infty}\frac{(x+k+1)!}{x!(k-1)!}q^{x+2}p^k = \frac{k(k+1)q^2}{p^2}\sum_{x=0}^{\infty}\binom{x+k+1}{x}q^x p^{k+2} =$$
$$\frac{k(k+1)(1-p)^2}{p^2} \tag{2.9}$$

Άρα από τη Σχέση (2.7) και τη Σχέση (2.9), η Σχέση (2.8) γίνεται

$$VarX = k(k+1)\frac{(1-p)^2}{p^2} + k\frac{1-p}{p} - k^2\frac{(1-p)^2}{p^2} = k\frac{(1-p)^2}{p^2} + k\frac{1-p}{p} =$$
$$k\left(\frac{1}{p}-1\right)\left(\frac{1}{p}-1+1\right) = k\frac{1}{p}\left(\frac{1}{p}-1\right) \tag{2.10}$$

□

Πρόταση 2.1.2. *Η ροπογεννήτρια της τυχαίας μεταβλητής $X \sim NB(k,p)$, με πυκνότητα πιθανότητας $f(x;k,p) = \binom{x+k-1}{x}p^k q^x$ είναι ίση με*

$$M_X(t) = E(e^{tX}) = \frac{p^k}{(1-qe^t)^k} \tag{2.11}$$

Απόδειξη.

$$E(e^{tX}) = \sum_{x=0}^{\infty} e^{tx}\binom{x+k-1}{x}p^k q^x = \binom{k-1}{0}p^k + e^t\binom{k}{1}p^k q + e^{2t}\binom{k+1}{2}p^k q^2 + \dots =$$
$$p^k\left[1+\binom{k}{1}qe^t+\binom{k+1}{2}(qe^t)^2+\dots\right] = p^k(1-qe^t)^{-k} = \frac{p^k}{(1-qe^t)^k} \quad (2.12)$$

□

Παρατήρηση 2.1.5. Έχουμε ότι

$$\sum_{x=0}^{\infty}\binom{x+k-1}{x}p^x q^k = 1 \Leftrightarrow 1+\binom{k}{1}q+\binom{k+1}{2}q^2+\dots = \frac{1}{p^k} \Leftrightarrow 1+\binom{k}{1}q+\binom{k+1}{2}q^2+\dots = (1-q)^{-k}$$

Παρατήρηση 2.1.6. Από την παραμετρικοποίηση που έχουμε ως προς μ και k προέκυψε ότι $p = \frac{k}{\mu+k}$ και $q = \frac{\mu}{\mu+k}$. Άρα έχουμε ότι,

$$E(c^X) = E(e^{Xlnc}) = \frac{\left(\frac{k}{\mu+k}\right)^k}{\left(1-\frac{\mu}{\mu+k}c\right)^k} = \left(\frac{k}{k+\mu(1-c)}\right)^k \quad (2.13)$$

Πρόταση 2.1.3. *Έστω $X_1, X_2, \dots, X_n$ τυχαίο δείγμα της Αρνητικής Διωνυμικής Κατανομής $NB(1,p)$, τότε η κατανομή της τυχαίας μεταβλητής Y είναι, $Y = \sum_{j=1}^{k} X_j \sim NB(k,p)$.*

Απόδειξη.

$$M_Y(t) = \prod_{j=1}^{k} M_{X_j}(t) = \prod_{i=1}^{k}\frac{p}{1-qe^t} = \frac{p^k}{(1-qe^t)^k} \Rightarrow Y \sim NB(k,p) \quad (2.14)$$

□

2.2 Παραμετρικοποίηση της Αρνητικής Διωνυμικής Κατανομής

1η Παραμετρικοποίηση

Μια από τις παραμετρικοποιήσεις της Αρνητικής Διωνυμικής Κατανομής είναι αυτή με χρήση της μέσης τιμής της κατανομής λ, όπου $\lambda = k(p^{-1} - 1)$ ενώ σε αυτή την περίπτωση η παράμετρος p είναι ίση με $p = \frac{k}{k+\lambda}$.

Οπότε η συνάρτηση πυκνότητας πιθανότητας από τη Σχέση (2.1) παίρνει τη μορφή

$$f(x; \lambda, k) = \frac{\lambda^x}{x!} \frac{\Gamma(k+x)}{\Gamma(k)(k+\lambda)^x} \frac{1}{(1+\frac{\lambda}{k})^k}, \tag{2.15}$$

όπου λ και k είναι θετικοί πραγματικοί αριθμοί.

Υπό αυτή την παραμετρικοποίηση και θέλοντας να βρούμε τη σύγκλιση της πυκνότητας πιθανότητας $f(x; \lambda, k)$ όταν $k \to \infty$, προκύπτει ότι,

$$\lim_{k \to \infty} f(x; \lambda, k) = \frac{\lambda^x}{x!} 1 \frac{1}{e^\lambda} = e^{-\lambda} \frac{\lambda^x}{x!}. \tag{2.16}$$

Συμπερασματικά η Αρνητική Διωνυμική Κατανομή συγκλίνει με την εναλλακτική παραμετρικοποίηση στην Κατανομή *Poisson* με παράμετρο λ. Το γεγονός αυτό καθιστά την Αρνητική Διωνυμική Κατανομή (NBD) κατάλληλη εναλλακτική κατανομή της *Poisson* καθώς την προσεγγίζει για μεγάλο k. Μειονέκτημα στην προσέγγιση αυτή αποτελεί το ότι η NBD έχει μεγαλύτερες, σε τιμή, παραμέτρους από την *Poisson* για μικρά k.

2η Παραμετρικοποίηση

Μια δεύτερη παραμετρικοποίηση της Αρνητικής Διωνυμικής Κατανομής είναι αυτή κατά *Anscombe* (1949). Σύμφωνα με αυτή λοιπόν η παράμετρος της κατανομής θα είναι η $p = \frac{k}{\mu+k}$, όπου μ και k θετικές ποσότητες, ενώ η συνάρτηση πυκνότητας πιθανότητας θα δίνεται από τη σχέση

$$P(X=x) = \binom{k+x-1}{k-1} \left(\frac{k}{\mu+k}\right)^k \left(\frac{\mu}{\mu+k}\right)^x \tag{2.17}$$

Παρατήρηση 2.2.1. Μια απλή προέλευση της Αρνητικής Διωνυμικής Κατανομής μπορεί να βρεθεί αν θεωρήσουμε ανεξάρτητες και ισόνομες τ.μ. *Bernoulli*, όπου κάθε μια έχει πιθανότητα επιτυχίας p. Η πιθανότητα απαίτησης x ($x \in \{0,1,\dots\}$) αποτυχιών πριν εμφανισθούν k ($k \in \{0,1,\dots\}$) επιτυχίες δίνεται από τη Σχέση (2.17) με πιθανότητα επιτυχίας $p = \frac{k}{\mu+k}$. Αυτή η παραμετρικοποίηση της Αρνητικής Διωνυμικής Κατανομής, όταν το k περιορίζεται στους θετικούς ακέραιους, καλείται κατανομή του *Polya*.

Παρατήρηση 2.2.2. Να σημειωθεί ότι βάσει της παραπάνω παραμετρικοποίησης κατά *Anscombe* (1949), με την προϋπόθεση ότι το k είναι γνωστό, η Αρνητική Διωνυμική Κατανομή θα ανήκει στην οικογένεια των Εκθετικών Κατανομών.

Παρατήρηση 2.2.3. Όταν το k είναι άγνωστο και πάντα υπό την παραμετρικοποίηση κατά *Anscombe* (1949), η Αρνητική Διωνυμική Κατανομή δεν ανήκει στην οικογένεια των εκθετικών κατανομών ενώ η εκτίμηση του k είναι αδύνατη.

3η Παραμετρικοποίηση

Η τρίτη παραμετρικοποίηση της Αρνητικής Διωνυμικής Κατανομής έγινε προκειμένου να εκτιμηθεί η παράμετρος σκέδασης α από τους *Clark* και *Perry* (1989). Μέσω της Σχέσης (2.17) ο *Anscombe* (1950), βρήκε ότι η συνάρτηση πυκνότητας πιθανότητας είναι

$$P[X=x] = \left(\frac{\mu}{\mu+k}\right)^x \left(1+\frac{\mu}{k}\right)^{-k} \frac{\Gamma(x+k)}{x!\Gamma(k)} \tag{2.18}$$

όπου $k > 0$ και $\mu > 0$.

Να σημειωθεί ότι μ είναι η μέση τιμή της κατανομής και ότι η διασπορά της κατανομής υπολογίζεται από τη σχέση,

$$VarX = \mu + \frac{\mu^2}{k} \tag{2.19}$$

4η Παραμετρικοποίηση

Για το λόγο ότι δεν μπορεί να υπολογιστεί ο εκτιμητής της παραμέτρου k της κατανομής, όπως διαπίστωσε ο *Anscombe* (1950), δημιουργήθηκε μια νέα παραμετρικοποίηση από τους *Bliss* και *Owen* (1958) η οποία κάνει χρήση της παραμέτρου $\alpha = \frac{1}{k}$.

Χρησιμοποιώντας τη Σχέση (2.17), η συνάρτηση πυκνότητας πιθανότητας βάσει της νέας παραμετρικοποίησης είναι,

$$P[X = x] = \frac{\Gamma(x + \alpha^{-1})}{x!\Gamma(\alpha^{-1})} \left(\frac{\alpha\mu}{1 + \alpha\mu}\right)^x (1 + \alpha\mu)^{-\frac{1}{a}} \tag{2.20}$$

Συμβολικά έχουμε ότι $X \sim NB(\mu, \alpha)$

Παρατήρηση 2.2.4. Η παραπάνω παραμετρικοποίηση επιτρέπει την ευθεία ταυτοποίηση των αποτελεσμάτων της παραμέτρου α στη διαδικασία της εκτίμησης.

Παρατήρηση 2.2.5. Όταν η τιμή της παραμέτρου α τείνει στο μηδέν ($\alpha \to 0$), η Σχέση (2.19) μας οδηγεί στην κατανομή *Poisson* με παράμετρο μ.

Παρατήρηση 2.2.6. Βάσει των παραμέτρων μ και k της Αρνητικής Διωνυμικής Κατανομής ο *Oldham* (1968), έκανε μια νέα παραμετρικοποίηση με παραμέτρους $a = \frac{\mu}{k}$ και $b = k$, βάσει της οποίας η συνάρτηση πυκνότητας πιθανότητας γίνεται

$$P(X = x) = \binom{b + x - 1}{x} \left(\frac{a}{a+1}\right)^b \left(\frac{1}{a+1}\right)^x \tag{2.21}$$

Ενώ ο *Fisher* (1941) χρησιμοποιώντας $p = \frac{\mu}{k}$ και $q = 1 - p$ δημιούργησε μια νέα παραμετρικοποίηση με συνάρτηση πυκνότητας πιθανότητας

$$P(X = x) = p_0 \binom{x+k-1}{x} (\frac{p}{q})^x \tag{2.22}$$

όπου $x = 0, 1, 2, \ldots$

και

$$p_0 = q^{-k} \tag{2.23}$$

Μια επιπλέον παραμετρικοποίηση είναι αυτή των *Hastings* και *Peacock* (1974) οι οποίοι έκαναν χρήση των παραμέτρων

$$p = \frac{k}{\mu + k} \tag{2.24}$$

$$q = 1 - p \tag{2.25}$$

με αντίστοιχη συνάρτηση πιθανότητας

$$P(X = x) = \binom{k+x-1}{x} p^k q^x \tag{2.26}$$

με $x = 0, 1, 2, \ldots$

Παρατήρηση 2.2.7. Η βασική παραμετρικοποίηση που θα χρησιμοποιήσουμε είναι αυτή με παραμέτρους μ, k, κατά *Anscombe*(1949) και με αυτή θα συνεχίσουμε την αναλυσή μας βάσει της Σχέσης (2.17).

2.3 Επάρκεια -πληρότητα

Θεώρημα 2.3.1. *Η στατιστική συνάρτηση* $T(\underset{\sim}{x}) = \left(X_{(1)}, X_{(2)}, ...X_{(n)}\right)$ *είναι ελάχιστα επαρκής αλλά όχι πλήρης για μέγεθος δείγματος* $n \geq 3$*, όπου* $X_{(1)}, X_{(2)}, ...X_{(n)}$ *είναι οι διατεταγμένες παρατη-ρήσεις.*

Απόδειξη. Για τον προσδιορισμό της ελάχιστα επαρκούς στατιστικής συνάρτησης (βλ. Ορισμό 1.6.2) θα κάνουμε χρήση της μεθόδου των $Lehmann - Scheffe$ (1950).

Για οποιοδήποτε $\underset{\sim}{x}^0$ ορίζουμε το $D(\underset{\sim}{x}^0)$ ως το σύνολο των σημείων $\underset{\sim}{x}$ για τα οποία υπάρχει μια συνάρτηση $k(\underset{\sim}{x}, \underset{\sim}{x}^0)$ η τιμή της οποίας είναι 0,

$$k(\underset{\sim}{x}, \underset{\sim}{x}^0) = 0 \tag{2.27}$$

Επιπλέον η συνάρτηση $k(\underset{\sim}{x}, \underset{\sim}{x}^0)$ είναι ανεξάρτητη των παραμέτρων της Αρνητικής Διωνυμικής Κατανομής μ, k. Τέλος, η συνάρτηση $k(\underset{\sim}{x}, \underset{\sim}{x}^0)$ είναι τέτοια ώστε

$$P_{\mu,k}\left(\underset{\sim}{x}\right) = k(\underset{\sim}{x}, \underset{\sim}{x}^0) P_{\mu,k}\left(\underset{\sim}{x}^0\right) \tag{2.28}$$

δηλαδή το σύνολο $D(\underset{\sim}{x}^0)$ είναι,

$$D(\underset{\sim}{x}^0) = \{\underset{\sim}{x} | P_{\mu,k}\left(\underset{\sim}{x}\right) = k(\underset{\sim}{x}, \underset{\sim}{x}^0) P_{\mu,k}\left(\underset{\sim}{x}^0\right)\} \tag{2.29}$$

για όλα τα μ,k.

Συμπερασματικά το σύνολο $D(\underset{\sim}{x}^0)$ αποτελείται από όλα εκείνα τα $\underset{\sim}{x}$ για τα οποία ο λόγος $\frac{P_{\mu,k}(\underset{\sim}{x})}{P_{\mu,k}(\underset{\sim}{x}^0)}$ είναι ανεξάρτητος των μ,k.

Για την ελάχιστα επαρκή στατιστική συνάρτηση $T(\underset{\sim}{x})$ θέλουμε να αποδείξουμε ότι

$$T(\underset{\sim}{x}) = T(\underset{\sim}{x}^0) \tag{2.30}$$

αν και μόνο αν το κάθε διάνυσμα $\underset{\sim}{x} \in D(\underset{\sim}{x}^0)$ ενώ,

$$T(\underset{\sim}{x}) \neq T(\underset{\sim}{x}^0) \tag{2.31}$$

αν και μόνο αν το κάθε διάνυσμα $\underset{\sim}{x} \notin D(\underset{\sim}{x}^0)$.

Ξεκινώντας λοιπόν από το λόγο $\frac{P_{\mu,k}(\underset{\sim}{x})}{P_{\mu,k}(\underset{\sim}{x}^0)}$, έχουμε ότι

$$\frac{P_{\mu,k}\,(\underset{\sim}{x})}{P_{\mu,k}\,(\underset{\sim}{x}^0)} = \frac{(\frac{k}{k+\mu})^{nk}(\frac{\mu}{\mu+k})^{\sum_{i=1}^n x_i}(\frac{1}{(k-1)!})^n \prod_{i=1}^n \frac{(x_i+k-1)!}{x_i!}}{(\frac{k}{k+\mu})^{nk}(\frac{\mu}{\mu+k})^{\sum_{i=1}^n x_i{}^0}(\frac{1}{(k-1)!})^n \prod_{i=1}^n \frac{(x_i{}^0+k-1)!}{x_i{}^0!}} =$$

$$(\frac{\mu}{\mu+k})^{\sum_{i=1}^n x_i - \sum_{i=1}^n x_i{}^0} \prod_{i=1}^n \frac{x_i{}^0!(x_i+k-1)!}{x_1!(x_i{}^0+k-1)!} = k(\underset{\sim}{x}, \underset{\sim}{x}^0) \tag{2.32}$$

Προκειμένου να αποδείξουμε ότι η στατιστική συνάρτηση που επιλέξαμε είναι ελάχιστα επαρκής αρκεί να αποδείξουμε ότι

$$\sum_{i=1}^n x_i = \sum_{i=1}^n x_i{}^0 \tag{2.33}$$

και

$$\prod_{i=1}^n \frac{x_i{}^0!(x_i+k-1)!}{x_1!(x_i{}^0+k-1)!} = c \tag{2.34}$$

όπου c μια σταθερά.

Για να ισχύει η παραπάνω σχέση θα πρέπει για κάθε $x_i{}^0$ με $i = 1, 2, \ldots, n$ να αντιστοιχίζεται σε αυτό ένα x_j με $j = 1, 2, \ldots, n$ τέτοιο ώστε να ισχύει

$$x_j = x_i{}^0. \tag{2.35}$$

Έτσι οι διατεταγμένες παρατηρήσεις του διανύσματος $\underset{\sim}{x}$ πρέπει να ταυτίζονται με τις μεταβλητές του διανύσματος $\underset{\sim}{x}^0$ για κάθε $\underset{\sim}{x} \in D(\underset{\sim}{x}^0)$.

Έτσι πετυχαίνουμε να αποδείξουμε ότι η $T(\underset{\sim}{x}) = \big(X_{(1)}, X_{(2)}, ...X_{(n)}\big)$ είναι ελάχιστα επαρκής στατιστική συνάρτηση.

Στη συνέχεια αποδεικνύουμε ότι η στατιστική συνάρτηση $T(\underset{\sim}{x})$ δεν είναι πλήρης. Για να το αποδείξουμε αυτο θεωρούμε μια συνάρτηση $g(x_{(1)}, x_{(2)}, \ldots, x_{(n)})$ τέτοια ώστε να ισχύει

$$E(g(X_{(1)}, \ldots, X_{(n)})) = 0 \tag{2.36}$$

Έχουμε λοιπόν ότι

$$0 = \sum g(x_1, x_2, \ldots, x_n) n! \frac{\prod_i (k + x_i - 1)!}{((k-1)!)^n \prod_i x_i!} (\frac{k}{\mu + k})^{nk} (\frac{\mu}{\mu + k})^{\sum_{i=1}^n x_i} \tag{2.37}$$

Η ανάλυσή μας παρακάτω θα γίνει βάσει του παραπάνω αθροίσματος και θα θέσουμε για ευκολία στην παρουσίαση $p = \frac{k}{\mu+k}$ και $q = \frac{\mu}{\mu+k}$

$$\begin{aligned} 0 = \frac{n! p^{nk}}{((k-1)!)^n} &[((k-1)!)^n g(0, \ldots, 0) q^0 + ((k-1)!)^{n-1} k! g(0, \ldots, 0, 1) q^1 \\ &+ (\frac{((k-1)!)^{n-1} (k+1)!}{2!} g(0, \ldots, 0, 2) + ((k-1)!)^{n-2} (k!)^2 g(0, \ldots, 0, 1, 1)) q^2 \\ &+ (\frac{((k-1)!)^{n-1} (k+2)!}{3!} g(0, \ldots, 0, 3) + \frac{((k-1)!)^{n-2} k! (k+1)!}{2!} g(0, \ldots, 0, 1, 2) \\ &+ ((k-1)!)^{n-3} (k!)^3 g(0, \ldots, 0, 1, 1, 1)) q^3 + \ldots] \end{aligned} \tag{2.38}$$

Θεωρούμε την παραπάνω σχέση ως πολυώνυμο ως προς q και συμπεραίνουμε ότι για να ισχύει η ισότητα πρέπει οι συντελεστές των δυνάμεων του q να είναι ίσοι με το 0.

Να σημειωθεί ότι βάσει του θεωρήματος η στατιστική συνάρτηση δεν είναι πλήρης για μέγεθος δείγματος $n \geq 3$.

Συνεχίζουμε την απόδειξη για τους συντελεστές του q^c για $4 \leq c \leq n$ και $n \geq 4$ και παράλληλα χρησιμοποιώντας τον συμβολισμό $[X]$ για τον μεγαλύτερο ακέραιο μικρότερο ή ίσο του X.

$$\begin{aligned}
0 = \frac{((k-1)!)^{n-1}(k+c-1)!}{c!}g(0,\dots,0,c) \\
+\frac{((k-1)!)^{n-2}k!(k+c-2)!}{(c-1)!}g(0,\dots,0,1,c-1) \\
+\frac{((k-1)!)^{n-2}(k+1)!(k+c-3)!}{2!(c-2)!}g(0,\dots,0,2,c-2) \\
+\frac{((k-1)!)^{n-2}(k+2)!(k+c-4)!}{3!(c-3)!}g(0,\dots,0,3,c-3)+\dots \\
+\frac{((k-1)!)^{n-2}(k+[\frac{c}{2}]-1)!(k+c-[\frac{c}{2}]-1)!}{[\frac{c}{2}]!(c-[\frac{c}{2}])!}g(0,\dots,0,[\frac{c}{2}],c-[\frac{c}{2}]) \\
+\dots+\frac{((k-1)!)^{n-m}(k!)^{m-1}(k+c-m)!}{(c-m+1)!}g(0,\dots,0,1,\dots,1,c-m+1) \\
+\frac{((k-1)!)^{n-m}(k!)^{m-2}(k+1)!(k+c-m-1)!}{2!(c-m)!}g(0,\dots,0,1,\dots,1,2,c-m) \\
+\dots+\frac{((k-1)!)^{n-m}[(k+[\frac{c}{m}]-1)!]^{m-1}(k+c-(m-1)[\frac{c}{m}]-1)!}{([\frac{c}{m}]!)^{m-1}(c-(m-1)[\frac{c}{m}])!} \\
g(0,\dots,0,[\frac{c}{m}],\dots,[\frac{c}{m}],c-(m-1)[\frac{c}{m}])^3 \\
+\dots+\frac{((k-1)!)^{n-c+1}(k!)^{c-2}(k+1)!}{2!}g(0,\dots,0,1,\dots,1,2) \\
+((k-1)!)^{n-c}(k!)^c g(0,\dots,0,1,\dots,1) \qquad (2.39)
\end{aligned}$$

Απαλοίφοντας από την παραπάνω σχέση τον παράγοντα $((k-1)!)^{n-1}k!$ και θεωρώντας πλέον την παραπάνω σχέση πολυώνυμο ως προς k έχουμε

$$g(0,\dots,0,c)=0. \qquad (2.40)$$

Επιπλέον βγάζοντας κοινό παράγοντα το k, το άθροισμα της παραπάνω σχέσης είναι ίσο με

$$\begin{aligned}
0 =& \frac{(k+c-2)\dots(k+1)}{(c-1)!}g(0,\dots,0,1,c-1)\\
&+\frac{(k+1)(k+c-3)\dots(k+1)}{2!(c-2)!}g(0,\dots,0,2,c-2)\\
&+\frac{(k+2)(k+1)(k+c-4)\dots(k+1)}{3!(c-3)!}g(0,\dots,0,3,c-3)\\
&+\dots+\frac{(k+[\frac{c}{2}]-1)\dots(k+1)(k+c-[\frac{c}{2}]-1)\dots(k+1)}{[\frac{c}{2}]!(c-[\frac{c}{2}])!}\\
&g(0,\dots,0,[\frac{c}{2}],c-[\frac{c}{2}])\\
&+\dots+\frac{k^{m-2}(k+c-m)\dots(k+1)}{(c-m+1)!}g(0,\dots,0,1,\dots,1,c-m+1)\\
&+\frac{k^{m-3}(k+1)k(k+c-m-1)\dots(k+1)}{2!(c-m)!}g(0,\dots,0,1,\dots,1,2,c-m)\\
&+\dots+\frac{k^{m-2}[(k+[\frac{c}{m}]-1)\dots(k+1)]^{m-1}(k+c-(m-1)[\frac{c}{m}]-1)\dots(k+1)}{[([\frac{c}{m}])!]^{m-1}(c-(m-1)[\frac{c}{m}])!}\\
&g(0,\dots,0,[\frac{c}{m}],\dots,[\frac{c}{m}],c-(m-1)[\frac{c}{m}])\\
&+\dots+\frac{k^{c-3}(k+1)}{2!}g(0,\dots,0,1,\dots,1,2)+k^{c-2}g(0,\dots,0,1,\dots,1) \qquad (2.41)
\end{aligned}$$

Παρατηρούμε ότι από το παραπάνω πολυώνυμο, ως προς k, έχουμε $c-1$ περιορισμούς των g. Αν υπήρχαν παραπάνω από $c-1$ συντελεστές τότε θα είχαμε άπειρο πλήθος λύσεων. Παρατίθεται παράδειγμα για $c=4$.

$$\begin{aligned}
&\frac{(k+2)(k+1)}{3!}g(0,\dots,0,1,3)+\frac{(k+1)(k+1)}{(2!)^2}g(0,\dots,0,2,2)\\
&+\frac{k(k+1)}{2!}g(0,\dots,0,1,1,2)+k^2g(0,\dots,0,1,1,1,1)=0 \qquad (2.42)
\end{aligned}$$

Καταλήγουμε λοιπόν ότι οι συντελεστές του πολυωνύμου παράγουν 3 περιορισμούς για συνολικά 4 g. Έχουμε δηλαδή ένα ομογενές σύστημα 3 εξισώσεων με 4 αγνώστους άρα υπάρχουν άπειροι τρόποι έκφρασης της συνάρτησης g που να ικανοποιούν το παραπάνω σύστημα.

Συμπεραίνουμε λοιπόν ότι ισχύει η υπόθεση της ύπαρξης μιας συνάρτησης $g(x_{(1)}, x_{(2)}, \dots, x_{(n)}) \neq 0$

τέτοια ώστε η αναμενόμενη τιμή της να είναι μηδέν, $E(g(X_{(1)}, \ldots, X_{(n)})) = 0$. Συνεπώς η στατιστική συνάρτηση $T(\underset{\sim}{x})$ δεν είναι πλήρης.

$\square$

Κεφάλαιο 3

Εκτίμηση Παραμέτρων Αρνητικής Διωνυμικής Κατανομής

Σε αυτό το κεφάλαιο θα ασχοληθούμε με την εκτίμηση των παραμέτρων k, μ της Αρνητικής Διωνυμικής Κατανομής $NB(\mu, k)$. Αρχικά θα αποδείξουμε ότι δεν υπάρχει Αμερόληπτος Εκτιμητής για την παράμετρο k. Στη συνέχεια θα εκτιμήσουμε με τη χρήση των μεθόδων, Μέγιστης Πιθανοφάνειας, Μεγίστων Ροπών τις παραμέτρους μ, k.

3.1 Αμερόληπτος Εκτιμητής του k

Οι Willson et al. (1986), ανέφεραν ότι δεν μπορεί να βρεθεί ΑΟΕΔ εκτιμητής για το k με τη συνήθη διαδικασία. Ουσιαστικά όπως απέδειξε ο Wang (1996), δεν υπάρχει καν αμερόληπτος εκτιμητής για το k.

Πρόταση 3.1.1. *Έστω τ.μ. $X \sim NB(\mu, k)$, τότε δεν υπάρχει αμερόληπτος εκτίμητής του k.*

Απόδειξη. Έστω $T(\underset{\sim}{X})$ ένας εκτιμητής της παραμέτρου k, όπου $\underset{\sim}{X} = (X_1, X_2, \ldots, X_n)$. Αν υποθέσουμε ότι ο $T(\underset{\sim}{X})$ είναι αμερόληπτος εκτιμητής του k έχουμε,

$ET(\underset{\sim}{X}) = k, \forall k = 1, 2, \ldots$ ή διαφορετικά

$$\sum_{\underset{\sim}{x}} \left[\prod_{i=1}^{n} \binom{x_i + k - 1}{k - 1} \right] (1-q)^{nk} q^{\sum x_i} T(\underset{\sim}{x}) = k \tag{3.1}$$

για όλα τα $k \in (0, \infty)$ και $q = 1 - p$ όπου $q \in (0, 1)$.

Μπορεί εύκολα να αποδειχθεί ότι το αριστερό μέλος της Σχέσης (3.1) συγκλίνει για $|q| < 1$. Έστω $k = 1$ και $q \to 0$, τότε από τη Σχέση (3.1) προκύπτει ότι $T(0, 0, \ldots, 0) = 1$. Στη συνέχεια για $k = 2$ και ομοίως για $q \to 0$ η Σχέση (3.1) παίρνει πλέον την τιμή $T(0, 0, \ldots, 0) = 2$. Συνεχίζοντας με τον ίδιο τρόπο, θέτοντας διάφορες τιμές στο k η Σχέση (3.1) δεν επαληθεύεται ποτέ και άρα δεν υπάρχει ο Αμερόληπτος εκτιμητής του k. □

3.2 Εκτιμητές Μέγιστης Πιθανοφάνειας

Σε αυτή την ενότητα υπολογίζονται οι εκτιμητές μέγιστης πιθανοφάνειας των παραμέτρων μ και k, χρησιμοποιώντας την εργασία του $Haldane$(1941).

Θεώρημα 3.2.1. *Θεωρούμε τυχαίο δείγμα $X_1, X_2, \ldots, X_n$ της Αρνητικής Διωνυμικής Κατανομής, $NB(\mu, k)$. Έστω f_j η σχετική συχνότητα του πλήθους x των αποτυχιών του πειράματος σε μέγεθος δείγματος n, ίσο με j και $F_j = \sum_{i=j}^{\infty} f_i$ η αθροιστική σχετική συχνότητα των x μεγαλύτερων ή ίσων του j, τότε ο εκτιμητής μέγιστης πιθανοφάνειας του μ είναι ο δειγματικός μέσος,*

$$\hat{\mu} = \bar{X} = \sum_{j=0}^{\infty} f_j X_j = \sum_{j=1}^{\infty} F_j \tag{3.2}$$

ενώ ο εκτιμητής μέγιστης πιθανοφάνειας του k είναι η ρίζα της παρακάτω εξίσωσης ως προς $\hat{k}$.

$$log(1 + \bar{X}\hat{k}^{-1}) = \sum_{j=1}^{\infty} (\hat{k} + j - 1)^{-1} F_j \tag{3.3}$$

Απόδειξη. Για να υπολογίσουμε τον εκτιμητή μέγιστης πιθανοφάνειας της παραμέτρου μ, βρίσκουμε τη συνάρτηση πιθανοφάνειας,

$$L(\mu,k) = f(\underset{\sim}{x};\mu,k) = \prod_{i=1}^{n} f_i(x_i;\mu,k) = \prod_{i=1}^{n} \binom{x_i+k-1}{k-1} \left(\frac{k}{\mu+k}\right)^k \left(\frac{\mu}{\mu+k}\right)^{x_i} = \frac{\prod_{i=1}^{n}(x_i+k-1)!}{((k-1)!)^n \prod_{i=1}^{n} x_i!} \left(\frac{k}{\mu+k}\right)^{nk} \left(\frac{\mu}{\mu+k}\right)^{\sum_{i=1}^{n} x_i} \tag{3.4}$$

Λογαριθμίζουμε την συνάρτηση πιθανοφάνειας και προκύπτει ότι

$$\begin{aligned} lnL(\mu,k) &= ln\frac{\prod_{i=1}^{n}(x_i+k-1)!}{((k-1)!)^n \prod_{i=1}^{n} x_i!} + ln\left(\frac{k}{\mu+k}\right)^{nk} + ln\left(\frac{\mu}{\mu+k}\right)^{\sum_{i=1}^{n} x_i} = \\ & ln\frac{\prod_{i=1}^{n}(x_i+k-1)!}{((k-1)!)^n \prod_{i=1}^{n} x_i!} + nkln\left(\frac{k}{\mu+k}\right) + \sum_{i=1}^{n} x_i ln\left(\frac{\mu}{\mu+k}\right) = \\ & ln\frac{\prod_{i=1}^{n}(x_i+k-1)!}{((k-1)!)^n \prod_{i=1}^{n} x_i!} + nk(lnk - ln(\mu+k)) + \sum_{i=1}^{n} x_i(ln\mu - ln(\mu+k)) = \\ & ln\frac{\prod_{i=1}^{n}(x_i+k-1)!}{((k-1)!)^n \prod_{i=1}^{n} x_i!} + nklnk - nkln(\mu+k) + \sum_{i=1}^{n} x_i ln\mu - \sum_{i=1}^{n} x_i ln(\mu+k) \end{aligned} \tag{3.5}$$

Κατόπιν βρίσκουμε τη μερική παράγωγο ως προς μ και μηδενίζουμε προκειμένου να βρούμε την τιμή του εκτιμητή μέγιστης πιθανοφάνειας $\hat{\mu}$,

$$\begin{aligned} &\frac{\partial lnL(\mu,k)}{\partial \mu} = 0 \Longleftrightarrow -nk\frac{1}{\mu+k} + \frac{1}{\mu}\sum_{i=1}^{n} x_i - \frac{1}{\mu+k}\sum_{i=1}^{n} x_i = 0 \Longleftrightarrow \\ &-\mu(\mu+k)nk\frac{1}{\mu+k} + \mu(\mu+k)\frac{1}{\mu}\sum_{i=1}^{n} x_i - \mu(\mu+k)\frac{1}{\mu+k}\sum_{i=1}^{n} x_i = 0 \Longleftrightarrow \\ &-n\mu k + (\mu+k)\sum_{i=1}^{n} x_i - \mu\sum_{i=1}^{n} x_i = 0 \Longleftrightarrow -n\mu k + \mu\sum_{i=1}^{n} x_i + k\sum_{i=1}^{n} x_i - \mu\sum_{i=1}^{n} x_i = 0 \Longleftrightarrow \\ &k\left(-n\mu + \sum_{i=1}^{n} x_i\right) = 0 \Longleftrightarrow -n\mu + \sum_{i=1}^{n} x_i = 0 \Longleftrightarrow \hat{\mu} = \frac{\sum_{i=1}^{n} x_i}{n} \Longleftrightarrow \hat{\mu} = \bar{X} \end{aligned} \tag{3.6}$$

Αντίστοιχα, για να υπολογίσουμε τον εκτιμητή μέγιστης πιθανοφάνειας $\hat{k}$ του k, θεωρούμε $m = max\{X_i : i = 1,2,\ldots,n\}$. Για λόγους απλότητας και κατανόησης της απόδειξης θεωρούμε $m = 2$, κάτι το οποίο και σημαίνει ότι το δείγμα $\underset{\sim}{X} = (X_1, X_2, \ldots, X_n)$ παίρνει τιμές μόνο από το σύνολο $0, 1, 2$. Συμβολίζουμε με,

Y_0: το πλήθος εμφάνισης του 0 στο δείγμα

Y_1: το πλήθος εμφάνισης του 1 στο δείγμα

Y_2: το πλήθος εμφάνισης του 2 στο δείγμα.

Οπότε η πιθανοφάνεια γίνεται,

$$f_{\underset{\sim}{X}}(\underset{\sim}{x}) = f(x_1, x_2, \ldots, x_n) = [f(0)]^{Y_0} [f(1)]^{Y_1} [f(2)]^{Y_2} \tag{3.7}$$

όπου

$$\begin{aligned} f(0) = P(X=0) = \left(\frac{\bar{X}}{k}+1\right)^{-k} &= f_0(k) \\ f(1) = P(X=1) = k\left(\frac{\bar{X}}{k}+1\right)^{-k}\left(\frac{\bar{X}}{\bar{X}+k}\right) &= f_1(k) \\ f(1) = P(X=2) = \frac{k(k+1)}{2}\left(\frac{\bar{X}}{k}+1\right)^{-k}\left(\frac{\bar{X}}{\bar{X}+k}\right)^2 &= f_2(k) \end{aligned} \tag{3.8}$$

με $X \sim NB(\mu, k)$ και έχοντας αντικαταστήσει τον εκτιμητή μέγιστης πιθανοφάνειας του μ, δηλαδή $\hat{\mu} = \bar{X}$. Λογαριθμίζοντας την συνάρτηση της πιθανοφάνειας στη Σχέση (3.7) έχουμε ότι

$$L(k) = lnf(\underset{\sim}{x}) = y_0 lnf_0(k) + y_1 lnf_1(k) + y_2 lnf_2 \tag{3.9}$$

Παραγωγίζουμε ως προς k και μηδενίζοντας παράλληλα βρίσκουμε ότι,

$$L'(k) = 0 \Leftrightarrow y_0\frac{f_0'(k)}{f_0 k} + y_1\frac{f_1'(k)}{f_1(k)} + y_2\frac{f_2'(k)}{f_2(k)} = 0 \tag{3.10}$$

Στη συνέχεια υπολογίζουμε κάθε μια παράγωγο χωριστά, για την συνάρτηση $f_0(k)$,

$$f_0(k) = \left(\frac{\bar{X}}{k}+1\right)^{-k} = e^{-kln\left(\frac{\bar{X}}{k}+1\right)} \tag{3.11}$$

έχουμε ότι

$$f_0'(k) = \left(\frac{\bar{X}}{k}+1\right)^{-k}\left[-ln\left(\frac{\bar{X}}{k}+1\right) + \frac{\bar{X}}{\bar{X}+k}\right] \tag{3.12}$$

από τις σχέσεις (3.11),(3.12) προκύπτει ότι

$$\frac{f_0'(k)}{f_0(k)} = -ln\left(\frac{\bar{X}}{k}+1\right) + \frac{\bar{X}}{\bar{X}+k} \tag{3.13}$$

για τη συνάρτηση $f_1(k)$ και δουλεύοντας παρόμοια έχουμε,

$$f_1(k) = k\left(\frac{\bar{X}}{k}+1\right)^{-k}\left(\frac{\bar{X}}{\bar{X}+k}\right) \tag{3.14}$$

παραγωγίζουμε εκ νέου οπότε,

$$f_1'(k) = \left(\frac{\bar{X}}{k}+1\right)^{-k}\left(\frac{\bar{X}}{\bar{X}+k}\right) + k\left(\frac{\bar{X}}{k}+1\right)^{-k}\left[-ln\left(\frac{\bar{X}}{k}+1\right)+\frac{\bar{X}}{\bar{X}+k}\right]\left(\frac{\bar{X}}{\bar{X}+k}\right)$$
$$+k\left(\frac{\bar{X}}{k}+1\right)^{-k}\left(\frac{\bar{X}}{\bar{X}+k}\right)\left(-\frac{1}{\bar{X}+k}\right) \tag{3.15}$$

από τις Σχέσεις (3.14),(3.15) και διαιρώντας κατά μέλη βρίσκουμε,

$$\frac{f_1'(k)}{f_1(k)} = \frac{1}{k} - ln\left(\frac{\bar{X}}{k}+1\right) + \frac{\bar{X}}{\bar{X}+k} - \frac{1}{\bar{X}+k} \tag{3.16}$$

τέλος για τη συνάρτηση $f_2(k)$ και ακολουθώντας την ίδια διαδικασία, δηλαδή παραγωγίζοντας τη σχέση

$$f_2(k) = \frac{k(k+1)}{2}\left(\frac{\bar{X}}{k}+1\right)^{-k}\left(\frac{\bar{X}}{\bar{X}+k}\right)^2 \tag{3.17}$$

βρίσκουμε ότι

$$f_2'(k) = \frac{k+1}{2}\left(\frac{\bar{X}}{k}+1\right)^{-k}\left(\frac{\bar{X}}{\bar{X}+k}\right)^2 + \frac{k}{2}\left(\frac{\bar{X}}{k}+1\right)^{-k}\left(\frac{\bar{X}}{\bar{X}+k}\right)^2 +$$
$$+k\frac{(k+1)}{2}\left(\frac{\bar{X}}{k}+1\right)^{-k}\left[-ln\left(\frac{\bar{X}}{k}+1\right)+\frac{\bar{X}}{\bar{X}+k}\right]\left(\frac{\bar{X}}{\bar{X}+k}\right)^2 +$$
$$+k\frac{(k+1)}{2}\left(\frac{\bar{X}}{k}+1\right)^{-k}\left(\frac{\bar{X}}{\bar{X}+k}\right)^2\left(\frac{-2}{\bar{X}+k}\right) \tag{3.18}$$

διαιρώντας κατά μέλη τις Σχέσεις (3.17),(3.18),

$$\frac{f_2'(k)}{f_2(k)} = \frac{1}{k} + \frac{1}{k+1} - ln\left(\frac{\bar{X}}{k}+1\right) + \frac{\bar{X}}{\bar{X}+k} - \frac{2}{\bar{X}+k} \tag{3.19}$$

Αντικαθιστώντας τις Σχέσεις (3.13),(3.16),(3.19) στη Σχέση (3.10)

$$-y_0 ln\left(\frac{\bar{X}}{k}+1\right) + y_0\frac{\bar{X}}{\bar{X}+k} + y_1\frac{1}{k} - y_1 ln\left(\frac{\bar{X}}{k}+1\right) + y_1\frac{\bar{X}}{\bar{X}+k} - y_1\frac{1}{\bar{X}+k} + y_2\frac{1}{k}$$
$$y_2\frac{1}{k+1} - y2ln\left(\frac{\bar{X}}{k}+1\right) + y_2\frac{\bar{X}}{\bar{X}+k} - \frac{2y_2}{\bar{X}+k} = 0 \Longleftrightarrow$$
$$-nln\left(\frac{\bar{X}}{k}+1\right) + n\frac{\bar{X}}{\bar{X}+k} - (y_1+2y_2)\frac{1}{\bar{X}+k} + (y_1+y_2)\frac{1}{k} + y_2\frac{1}{k+1} = 0 \tag{3.20}$$

Παρατήρηση 3.2.1. ισχύει ότι $y_0 + y_1 + y_2 = n$

Παρατήρηση 3.2.2. $f_0 = \frac{y_0}{n}$,$f_1 = \frac{y_1}{n}$,$f_2 = \frac{y_2}{n}$

$$\bar{X} = \sum_{j=0}^{\infty} f_j X_j = f_0 0 + f_1 1 + f_2 2 = \frac{y_1 + 2y_2}{n} \Longleftrightarrow n\bar{X} = \frac{y_1 + 2y_2}{n} \tag{3.21}$$

Επίσης ισχύει ότι από τον ορισμό της αθροιστικής συχνότητας βρίσκουμε, $F_2 = f_2$,$F_1 = f_1 + f_2$ ή

$$F_1 = y_1 + y_2, nF_2 = y_2 \tag{3.22}$$

Από τις Σχέσεις (3.21) και (3.22) η Σχέση (3.20) γίνεται,

$$-nln\left(\frac{\bar{X}}{k} + 1\right) + \frac{y_1 + 2y_2}{\bar{X} + k} - \frac{y_1 + 2y_2}{\bar{X} + k} + n\frac{F_1}{k} + n\frac{F_2}{k+1} = 0 \tag{3.23}$$

Δηλαδή,

$$\sum_{j=1}^{2} \frac{F_j}{k + j - 1} - ln\left(\frac{\bar{X}}{k} + 1\right) = 0 \tag{3.24}$$

Η Εξίσωση (3.24) γενικεύεται για οποιοδήποτε m, δηλαδή καταλήγουμε στη Σχέση (3.3). □

Ο Anscombe(1950) *υπέθεσε ότι υπάρχει μόνο μια θετική ακέραια ρίζα της Σχέσης (3.3), αν και μόνο αν* $s^2 = \sum_{j=1}^{\infty} j^2 f_j - \left(\sum_{j=1}^{\infty} j f_j\right)^2 > \bar{X}$. *Οι Johnson και Kotz* (1969) *απέδειξαν ότι υπάρχει μια τουλάχιστον ρίζα της Σχέσης (3.3) όταν* $s^2 > \bar{X}$. *Παρόλα αυτά όμως η απόδειξή τους περιείχε πολλά λάθη οπότε παρατίθεται το παρακάτω Θεώρημα προκειμένου να τεκμηριωθεί ο ισχυρισμός των Johnson και Kotz* (1969).

Θεώρημα 3.2.2. *Υπάρχει μια τουλάχιστον θετική ακέραια ρίζα της Σχέσης (3.3) αν* $s^2 > \bar{X}$.

Απόδειξη. Για μικρά $\hat{k}$ ισχύει

$$log(1 + \bar{X}\hat{k}^{-1}) < 1 + log(\bar{X}\hat{k}^{-1}) = 1 + log(\hat{X}e^c) \tag{3.25}$$

όπου $c = log(\hat{k}^{-1})$. Χρησιμοποιώντας την ιδιότητα των γραμμικών συναρτήσεων, ότι αυτές φράσονται από τις εκθετικές, η Σχέση (3.7) μπορεί να γραφεί πάλι με τη μορφή

$$log(1 + \bar{X}\hat{k}^{-1}) < 1 + log\bar{X} + c < F_1 e^c < \sum_{j=1}^{\infty} (\hat{k} + j - 1)^{-1} F_j \tag{3.26}$$

Στη συνέχεια μελετάμε τη Σχέση (3.3) για μεγάλες τιμές του $\hat{k}$. Εργαζόμαστε στο αριστερό μέλος της Σχέσης (3.3) και έχουμε

$$log(1+\bar{X}\hat{k}^{-1}) > \bar{X}\hat{k}^{-1} - \frac{1}{2}\left(\bar{X}\hat{k}^{-1}\right)^2 \tag{3.27}$$

Να σημειωθεί ότι το δεξιό μέλος της Σχέσης (3.3) μπορεί να γραφτεί ως,

$$\begin{aligned}
\sum_{j=1}^{\infty}(\hat{k}+j-1)^{-1}F_j = \hat{k}^{-1}\sum_{j=1}^{\infty}(1+(j-1)\hat{k}^{-1})^{-1}F_j = \\
= \hat{k}^{-1}\sum_{j=1}^{\infty}F_j[1-(j-1)\hat{k}^{-1}+(j-1)^2\hat{k}^{-2}-(j-1)^3\hat{k}^{-3}+\dots] \\
< \hat{k}^{-1}\sum_{j=1}^{\infty}F_j[1-(j-1)\hat{k}^{-1}+(j-1)^2\hat{k}^{-2}] = \\
\hat{k}^{-1}\sum_{j=1}^{\infty}F_j[1-(j-1)\hat{k}^{-1}] + C\hat{k}^3 = \\
\hat{k}^{-1}\left(\sum_{j=1}^{\infty}jF_j - \left(\sum_{j=1}^{\infty}jF_j - \sum_{j=1}^{\infty}F_j\right)\hat{k}^{-1}\right) + C\hat{k}{-3} = \\
\hat{k}^{-1}\left(\bar{X} - \left(\sum_{j=1}^{\infty}j\frac{(j+1)}{2}f_j - \sum_{j=1}^{\infty}jf_j\right)\hat{k}^{-1}\right) + C\hat{k}^{-3} = \\
\hat{k}^{-1}\left(\bar{X} - \frac{1}{2}\left(\sum_{j=1}^{\infty}j^2f_j - \sum_{j=1}^{\infty}jf_j\right)\hat{k}^{-1}\right) + C\hat{k}{-3} = \\
\hat{k}^{-1}\left(\bar{X} - \frac{1}{2}(s^2+\bar{X}^2-\bar{X})\hat{k}^{-1}\right) + C\hat{k}{-3} = \\
\bar{X}\hat{k}^{-1} - \frac{1}{2}(\bar{X}\hat{k}^{-1})^2 - \frac{1}{2}\hat{k}^{-2}(s^2-\bar{X}+2C\hat{k}^{-1})
\end{aligned} \tag{3.28}$$

όπου C είναι μια σταθερά ανεξάρτητη του k και εξαρτημένη από τις σχετικές συχνότητες f_j. Από τις Σχέσεις (3.27) και (3.28) αποδεικνύεται ότι αν $s^2 > \bar{X}$, το δεξιό μέλος της Σχέσης (3.3) θα είναι μικρότερο του αριστερού μέλους, για k αρκετά μεγάλο. Το γεγονός αυτό σε συνδυασμό με τη Σχέση (3.25) και το γεγονός ότι και τα 2 μέλη της Σχέσης (3.3) είναι συνεχή ως προς $\hat{k}$, μας αποδεικνύει ότι υπάρχει τουλάχιστον μια θετική ρίζα της Σχέσης (3.3). □

Παρατήρηση 3.2.3. Αν και εξασφαλίζουμε την ύπαρξη του Εκτιμητή Μέγιστης Πιθανοφάνειας του k, δεν μπορούμε να εξασφαλίσουμε τη μοναδικότητά του.

Παρατήρηση 3.2.4. Δεν αποδεικνύουμε ότι δεν υπάρχει λύση της εξίσωσης στη Σχέση (3.3) αν $s^2 < \bar{X}$. Αν και γνωρίζουμε ότι $\sigma^2 > \mu$, η δειγματική διασπορά s^2 μπορεί να μην είναι μεγαλύτερη από το δειγματικό μέσο $\bar{X}$, γεγονός που μας οδηγεί στην ερώτηση για την καταλληλότητα της Αρνητικής Διωνυμικής Κατανομής, σχετικά με τη χρήση της από ερευνητές προκειμένου να κάνουν απαραίτητες εκτιμήσεις.

3.3 Εκτιμητές Μεθόδου Ροπών

Στη συγκεκριμένη ενότητα θα ασχοληθούμε με την εκτίμηση των παραμέτρων της Αρνητικής Διωνυμικής Κατανομής μέσω μιας κλάσης εκτιμητών ροπών. Αποδεικνύουμε ότι σε αυτή την κλάση είναι πιθανό να αποκομίσουμε εκτιμητές το ίδιο ικανοποιητικούς με τους εκτιμητές μέγιστης πιθανοφάνειας της προηγουμενης παραγράφου. Επιπροσθέτως και καθώς αποδείξαμε ότι η εύρεση του Ε.Μ.Π. του k δεν είναι απλή υπόθεση, θα καταλήξουμε ότι η εκτίμηση των k, μ μέσω των μεθόδων των ροπών, είναι απλούστερη αν και πλέον οι εκτιμητές δεν είναι ασυμπτωτικά ασυσχέτιστοι.

Θεώρημα 3.3.1. *Θεωρούμε τυχαίο δείγμα $X_1, X_2, \ldots, X_n$ από την Αρνητική Διωνυμική Κατανομή, $NB(\mu, k)$. Ο εκτιμητής για το μ με τη μέθοδο ροπών είναι ίσος με το δειγματικό μέσο*

$$\hat{\mu}_{MOM} = \bar{X} \tag{3.29}$$

ενώ ο εκτιμητής με τη μέθοδο ροπών για το k είναι ίσος με

$$\hat{k}_{MOM} = \frac{\bar{X}^2}{s^2 - \bar{X}} \tag{3.30}$$

Απόδειξη. Θεωρούμε τη δειγματική ροπή $\overline{f} = \frac{1}{n}\sum_{i=1}^{n} f(x_i)$, όπου n το μέγεθος δείγματος. Προκειμένου να βρούμε τον εκτιμητή του k, εξισώνουμε τη δειγματική ροπή $\bar{f}$ με την αναμενόμενη τιμή

$Ef(X)$ και λύνουμε την εξίσωση $\bar{f} = Ef(X)$, αντικαθιστώντας όπου $\hat{\mu} = \bar{X}$.

Ο $Anscombe$(1950) απέδειξε ότι αν η συνάρτηση $f(x)$ είναι μια ολοκληρώσιμη, κυρτή ή κοίλη συνάρτηση όταν x είναι μη αρνητικός ακέραιος, τότε η εξίσωση $\bar{f} = Ef(X)$ έχει το πολύ μια λύση. Προς το παρόν θεωρούμε 3 συναρτήσεις για την εκτίμηση του k,

$$\overline{f_j} = \frac{1}{n}\sum_{i=1}^{n} f_j(x_i) \tag{3.31}$$

με $f_1(x) = x^2$,$f_2(x) = I_{[x=0]}$, όπου $f_2(x) = 1$ αν $x = 0$ και $f_2(x) = 0$ αν $x \neq 0$, και $f_3(x) = c^x$ με $c > 0$ και $c \neq 1$. Οι δειγματικές ροπές είναι οι εξής,

$$\overline{f_1} = \overline{x^2} = \frac{1}{n}\sum_{i=1}^{n} {x_i}^2, \overline{f_2} = \widehat{p_0} = \frac{n_0}{n}, \overline{f_3} = \widehat{c^X} = \frac{1}{n}\sum_{i=1}^{n} c^{x_i}, \tag{3.32}$$

ενώ οι αναμενόμενες τιμές $Ef_j(X)$ δίνονται κάτωθι,

$$Ef_1(X) = \alpha k(\alpha + 1 + k\alpha), Ef_2(X) = (\alpha + 1)^{-k}, Ef_3(X) = (1 + \alpha - \alpha c)^{-k}. \tag{3.33}$$

Λύνουμε διαδοχικά την εξίσωση $\overline{f_j} = Ef_j(X)$ για $j = 1$, για $j = 2$ και για $j = 3$ και βρίσκουμε τον εκτιμητή του k, με την μέθοδο των ροπών (MOM), με τη $ZeroTerm$ μέθοδο (ZTM) και με τη μέθοδο της δυνάμεως (PM) αντίστοιχα. Οι συμβολισμοί που χρησιμοποιούμε είναι $\hat{k}_{MOM}$, $\hat{k}_{ZTM}$ και $\hat{k}_{PM(c)}$.

Για το λόγο ότι δεν υπάρχει αναλυτική αριθμητική έκφραση για τους εκτιμητές $\hat{k}_{ZTM}$ και $\hat{k}_{PM(c)}$, υπολογίζουμε τον εκτιμητή με τη μέθοδο των ροπών $\hat{k}_{MOM}$. Θέτουμε $s^2 = \overline{x^2} - \bar{x}^2$ και έχουμε ότι

$$\hat{k}_{MOM} = \frac{\bar{x}^2}{s^2 - \bar{x}} \tag{3.34}$$

Πιο αναλυτικά για τον υπολογισμό του $\hat{k}_{MOM}$, έχουμε

$$EX = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X} \Rightarrow \mu = \bar{X} \Rightarrow \hat{\mu} = \bar{X} \tag{3.35}$$

$$\begin{aligned} EX^2 = \frac{1}{n}\sum_{i=1}^{n} {X_i}^2 \Rightarrow \mu + \frac{\mu^2}{k} + \mu^2 = \frac{1}{n}\sum_{i=1}^{n} {X_i}^2 \Rightarrow \hat{\mu} + \frac{\hat{\mu}^2}{k} + \hat{\mu}^2 = \frac{1}{n}\sum_{i=1}^{n} {X_i}^2 \\ \Rightarrow \bar{X} + \frac{\bar{X}^2}{k} + \bar{X}^2 = \frac{1}{n}\sum_{i=1}^{n} {X_i}^2 \Rightarrow \hat{k}_{MOM} = \frac{\bar{X}^2}{s^2 - \bar{X}} \end{aligned} \tag{3.36}$$

όπου $s^2 = \frac{1}{n}\sum_{i=1}^{n} {X_i}^2 - \bar{X}^2$

Για τη $ZeroTerm$ μέθοδο (ZTM) με $\hat{\mu} = \bar{X}$ έχουμε

$Ef_2(X) = \bar{f}_2 \Leftrightarrow P(X=0) = \begin{cases} 1 & \text{για} \quad X = 0 \\ 0 & \text{για} \quad X \neq 0 \end{cases} \Leftrightarrow \left(\frac{k}{\hat{\mu}+k}\right)^k = \begin{cases} 1 & \text{για} \quad X = 0 \\ 0 & \text{για} \quad X \neq 0 \end{cases}$

ενώ για τη $PowerMethod$ (PM) με $\hat{\mu} = \bar{X}$ έχουμε

$Ef_3(X) = \bar{f}_3 \Leftrightarrow Ec^x = \frac{1}{n}\sum_{i=1}^{n} c^{x_i}$ □

Παρατήρηση 3.3.1. Ο εκτιμητής PM, $\hat{k}_{PM(c)}$, για $c = 0$ είναι ίσος με τον εκτιμητή ZTM, $\hat{k}_{ZTM}$, ενώ για $c \to 1$, ο εκτιμητής PM γίνεται εκτιμητής MOM.

Παρατήρηση 3.3.2. Όταν θέλουμε να συγκρίνουμε την ασυμπτωτική κατανομή και την ασυμπτωτική αποτελεσματικότητα των διαφορετικών εκτιμητών του k, είναι συνετό να ασχοληθούμε μόνο με την μέθοδο της δυνάμεως PM, με $c \in [0,1]$.

3.4 Ασυμπτωτικές Ιδιότητες των Εκτιμητών Μεθόδου Ροπών

Η ασυμπτωτική κατανομή των εκτιμητών μπορεί να αποδειχθεί αν χρησιμοποιήσουμε μια εκδοχή της πολυμεταβλητής δ-μεθόδου. Γνωρίζουμε ότι η από κοινού κατανομή του $\underset{\sim}{x}$ και της δειγματικής ροπής $\overline{f} = \frac{1}{n}\sum_{i=1}^{n} f(x_i)$ για $j = 1,2,3$ ακολουθεί μια διμεταβλητή κανονική κατανομή καθώς το $n \to \infty$. Για αυτό το λόγο οι εκτιμητές των (μ,k),(α,k),(p,k) και (b,w) επίσης ακολουθούν μια διμεταβλητή κατανομή για $n \to \infty$.

Να σημειωθεί ότι όταν το k είναι μεγάλο σε σχέση με το μ, οι πεπερασμένες δειγματικές κατανομές των εκτιμητών είναι λοξές, ενώ παράλληλα συγκλίνουν αργά στην ασυμπτωτική κανονική κατανομή.

Οι ασυμπτωτικές διασπορές των δειγματικών ροπών δίνονται από τις σχέσεις

$$Var(\bar{X}) = E\bar{X}^2 - (E\bar{X})^2 \tag{3.37}$$

και

$$Var(\overline{f_j}) = E\overline{f_j}^2 - (E\overline{f_j})^2 \tag{3.38}$$

για $j = 1, 2, 3$. Κάθε ένας από τους εκτιμητές $(\hat{\mu}, \hat{k}_{MOM})$, $(\hat{\mu}, \hat{k}_{ZTM})$, $(\hat{\mu}, \hat{k}_{PM})$, με $\hat{\mu} = \bar{X}$ ακολουθούν ασυμπτωτικά μια διμεταβλητή κανονική κατανομή με μέσο το διάνυσμα (μ, k). Παράλληλα ισχύει ότι $\lim_{n\to\infty} nVar(\hat{\mu}) = k\alpha(\alpha+1)$ και $\lim_{n\to\infty} nCov(\hat{\mu}, \hat{k}) = 0$. Για τον εκτιμητή PM ισχύει,

$$\lim_{n\to\infty} nVar(\hat{k}_{PM(c)}) = u_{PM}(c) = \frac{(1+\alpha-\alpha c^2)^{-k} r^{2k+2} - r^2 - k\alpha(\alpha+1)(1-c)^2}{[r\log(r) - r + 1]^2} \tag{3.39}$$

όπου $r = 1 + \alpha - \alpha c$.

Την κανονικοποιημένη διασπορά των εκτιμητών $\hat{k}_{MOM}$ και $\hat{k}_{ZTM}$ μπορούμε να τη βρούμε μέσω των ορίων, $\lim_{c\to 1} u_{PM}(c)$ και $\lim_{c\to 0} u_{PM}(c)$, όπου έχουμε ότι,

$$u_{MOM} = \lim_{n\to\infty} nVar(\hat{k}_{MOM}) = \frac{2k(k+1)(\alpha+1)^2}{\alpha^2} \tag{3.40}$$

και

$$u_{ZTM} = \lim_{n\to\infty} nVar(\hat{k}_{ZTM}) = \frac{(\alpha+1)^{k+2} - (\alpha+1)^2 - k\alpha(\alpha+1)}{[(\alpha+1)\log(\alpha+1) - \alpha]^2}. \tag{3.41}$$

Οι εκτιμητές MOM και ZTM είναι ασυμπτωτικά αποδοτικοί όταν ισχύουν τα ακόλουθα, $\frac{u_{MOM}}{u_{MLE}} \to 1$ καθώς το $k \to \infty$ και $\frac{u_{ZTM}}{u_{MLE}} \to 1$ καθώς το $k \to 0$.

Στην εργασία των Savani και Zhigljavsky (2006), αποδεικνύεται ότι για σταθερό $\mu > 0$ και k, $(0 < k < \infty)$, ο εκτιμητής PM μπορεί να είναι πιο αποδοτικός από τους εκτιμητές MOM και ZTM. Για την ακρίβεια, αποδεικνύεται ότι υπάρχει πάντα ένα $\tilde{c}$, με $0 < \tilde{c} < 1$, τέτοιο ώστε $u_{PM}(\tilde{c}) < \min\{u_{ZTM}, u_{MOM}\}$. Θεωρούμε ως $\tilde{c}$ την τιμή η οποία και ελαχιστοποιεί την $u_{PM}(c)$.

3.5 Εκτιμητής Bayes

Για να εκτιμήσουμε την παράμετρο k δεν θα κάνουμε χρήση της ευθείας μεθόδου κατά Bayes, αλλά θα χρησιμοποιήσουμε μια προσεγγιστική μέθοδο εκτίμησης αυτής. Για το λόγο ότι για μεγάλες τιμές του ακεραίου k, η τυχαία μεταβλητή X_i τείνει να ακολουθεί την κανονική κατανομή, μπορούμε να βρούμε κατάλληλη εκ των προτέρων συνάρτηση από το σύνολο $\{(\mu, \sigma^2) : 0 < \mu < \sigma^2\}$ και με χρήση της μεθόδου κατά Bayes, να αποκομίσουμε τον εκτιμητή για τη μεταβλητή k. Μάλιστα αυτός ο εκτιμητής αποδεικνύεται ότι έχει αποδεκτές τιμές για όλα τα δείγματα. Επιπρόσθετα ο εκτιμητής που θα βρούμε με την παραπάνω προσεγγιστική μέθοδο είναι καλύτερος του εκτιμητή με τη μέθοδο των ροπών και του εκτιμητή μέγιστης πιθανοφάνειας.

Θεώρημα 3.5.1. *Θεωρούμε τυχαίο δείγμα $X_1, X_2, \ldots, X_n$ της Αρνητικής Διωνυμικής Κατανομής, $NB(p,k)$. Τότε ο εκτιμητής κατά Bayes της παραμέτρου k είναι ίσος με $\hat{k}_b = \frac{\hat{\mu}^2}{\hat{\sigma}^2 - \hat{\mu}}$.*

Απόδειξη. Θεωρούμε τυχαίο δείγμα $X_1, X_2, \ldots, X_n$ της Αρνητικής Διωνυμικής Κατανομής, $NB(p,k)$. Αν ο k είναι φυσικός αριθμός, τότε κάθε X_i μπορεί να εκφραστεί μέσω τυχαίας μεταβλητής Y_j, $Y_j \sim NB(p,1)$, από τη σχέση

$$X_i = \sum_{j=1}^{k} Y_j. \tag{3.42}$$

Για μεγάλες τιμές του k και αν λάβουμε υπόψη το Κεντρικό Οριακό Θεώρημα, έχουμε ότι η τυχαία μεταβλητή X_i ακολουθεί την κανονική κατανομή $N(\mu, \sigma^2)$ με $0 < \mu < \sigma^2$. Υποθέτουμε ότι η εκ των προτέρων κατανομή του μ δοθέντος του σ^2 είναι η $\pi(\mu|\sigma^2) = \frac{1}{\sigma^2}$, για $0 \le \mu < \sigma^2$ και 0 αλλού. Επιπλέον θεωρούμε εκ των προτέρων κατανομή του σ^2, $\pi(\sigma^2) = \frac{1}{\sigma^2}$ για $\sigma^2 > 0$. Τότε προκύπτει ότι η από κοινού εκ των προτέρων κατανομή των μ και σ^2 είναι η $\pi(\mu, \sigma^2) = \frac{1}{\sigma^4}$ για $0 \le \mu < \sigma^2 < \infty$. Την εκ των υστέρων κατανομή που προκύπτει για (μ, σ^2) δοθέντος $\underset{\sim}{x}$ την υπολογίζουμε ως εξής,

$$\pi(\mu, \sigma^2 | \underset{\sim}{x}) \simeq \frac{(\frac{1}{\sigma^2})^2 (\frac{1}{\sigma^2})^{\frac{n}{2}} e^{-\frac{1}{2\sigma^2}\sum(x_i-\mu)^2}}{\int_0^\infty \int_0^{\sigma^2} (\frac{1}{\sigma})^{\frac{n}{2}+2} e^{-\frac{1}{2\sigma^2}} d\mu d\sigma^2} \tag{3.43}$$

η οποία μπορεί να γραφεί ως,

$$\begin{aligned}\pi(\mu, \sigma^2 | \underset{\sim}{x}) &\simeq \frac{h_1(\sigma^2)\frac{\sqrt{n}}{\sigma}\phi\left(\frac{\sqrt{n}(\mu-\bar{x})}{\sigma}\right)}{\int_0^\infty \int_0^{\sigma^2} h_1(\sigma^2)\frac{\sqrt{n}}{\sigma}\phi\left(\frac{\sqrt{n}(\mu-\bar{x})}{\sigma}\right) d\mu d\sigma^2} \\ &\simeq \frac{h_1(\sigma^2)\frac{\sqrt{n}}{\sigma}\phi\left(\frac{\sqrt{n}(\mu-\bar{x})}{\sigma}\right)}{E\left(\Phi\left(\frac{(\sigma^2-\bar{x})\sqrt{n}}{\sigma}\right) - \Phi\left(-\frac{\bar{x}\sqrt{n}}{\sigma}\right)\right)}.\end{aligned} \tag{3.44}$$

Στην παραπάνω σχέση, $h_1(\sigma^2)$ είναι η συνάρτηση πυκνότητας της αντίστροφης Γάμμα κατανομής $IG(\alpha, \beta)$ με $\alpha = \frac{(n+1)}{2}$, $\beta = \frac{2}{\sum(x_i - \bar{x})^2}$, $\phi(z)$ η συνάρτηση πυκνότητας πιθανότητας της τυπικής κανονικής κατανομής και Φ η συνάρτηση κατανομής της τυπικής κανονικής κατανομής.

Στη συνέχεια βρίσκουμε ότι, όταν η συνάρτηση ζημίας είναι το τετραγωνικό σφάλμα τότε ο εκτιμητής *Bayes* του μ, βλέπε Θεώρημα (1.11.1) υπολογίζεται από τον τύπο,

$$E(\mu|\underset{\sim}{X}) = \int_0^\infty \int_0^{\sigma^2} \mu\pi(\mu,\sigma^2|x)d\mu d\sigma^2 \Leftrightarrow$$

$$E(\mu|\underset{\sim}{X}) = \hat{\mu} \approxeq \bar{X} - \frac{E\left[\frac{\sigma}{\sqrt{n}}\left(\phi\left(\frac{(\sigma^2-\bar{X})\sqrt{n}}{\sigma}\right) - \phi\left(-\frac{\bar{X}\sqrt{n}}{\sigma}\right)\right)\right]}{E\left(\Phi\left(\frac{(\sigma^2-\bar{X})\sqrt{n}}{\sigma}\right) - \phi\left(-\frac{\bar{X}\sqrt{n}}{\sigma}\right)\right)} \tag{3.45}$$

ομοίως βρίσκουμε ότι ο εκτιμητής *Bayes* της διασποράς υπολογίζεται από τον τύπο,

$$E(\sigma^2|\underset{\sim}{X}) = \int_0^\infty \int_0^{\sigma^2} \sigma^2\pi(\mu,\sigma^2|\bar{X})d\mu d\sigma^2 \Leftrightarrow$$

$$E(\sigma^2|\underset{\sim}{X}) = \hat{\sigma^2} \approxeq \frac{E\left[\sigma^2\left(\phi\left(\frac{(\sigma^2-\bar{X})\sqrt{n}}{\sigma}\right) - \phi\left(-\frac{\bar{X}\sqrt{n}}{\sigma}\right)\right)\right]}{E\left(\Phi\left(\frac{(\sigma^2-\bar{X})\sqrt{n}}{\sigma}\right) - \phi\left(-\frac{\bar{X}\sqrt{n}}{\sigma}\right)\right)} \tag{3.46}$$

Οι αριθμοί που υπολογίσαμε παραπάνω, δηλαδή οι $\hat{\mu}$ και $\hat{\sigma}^2$ ονομάζονται γενικευμένοι ασυμπτωτικοί εκτιμητές *Bayes* των μ και σ^2 αντίστοιχα. Ενώ μέσω αυτών υπολογίζουμε τον εκτιμητή *Bayes* του k από τον τύπο,

$$\hat{k}_b = \frac{\hat{\mu}^2}{\hat{\sigma}^2 - \hat{\mu}}. \tag{3.47}$$

χρησιμοποιώντας τις ασυμπτωτικές ιδιότητες των εκτιμητών *Bayes*.

Παρατήρηση 3.5.1. Σημειώνουμε ότι $\hat{\mu} \leqslant \hat{\sigma}^2$ και για αυτό το λόγο καταλήγουμε ότι $\hat{k}_b \geqslant 0$.

$\square$

Κεφάλαιο 4

Αριθμητικά Αποτελέσματα

4.1 Αριθμητικές Συγκρίσεις Εκτιμητών Μεθόδου Ροπών

Στον Πίνακα (4.1) θεωρούμε τη μελέτη μιας προσομοίωσης της Αρνητικής Διωνυμικής Κατανομής $NB(\mu, k)$ για $R = 1000$ δείγματα, με μέγεθος δείγματος $n = 10000$. Επιπλέον η προσομοίωση γίνεται για διάφορες τιμές των παραμέτρων μ, k με $\mu \in \{0.1, 0.5, 1, 5, 10\}$ και $k \in \{0.01, 0.25, 0.5, 1, 3, 5\}$. Η αποτελεσματικότητα των εκτιμητών του k, διερευνάται συγκρίνοντας τους συντελεστές της διασποράς των εκτιμητών, MLE, MOM, PM στο c_* σε σχέση με τον ZTM εκτιμητή.

Παρατήρηση 4.1.1. Ο Πίνακας (4.1) μας δίνει τις τιμές των συντελεστών της διασποράς $\hat{\kappa} = \sqrt{N}\sqrt{\frac{1}{R}\sum_{i=1}^{R}\frac{(\hat{k_i}-k)^2}{k}} = \frac{\sqrt{NMSE}}{l}$ για τους εκτιμητές MLE, MOM, PM, ZTM, έναντι του συντελεστή της θεωρητικής διασποράς $\kappa = \frac{\sqrt{u_{ML}}}{k}$, όπου $u_{ML} = \lim_{n\to\infty} nVar\hat{k}_{ML} = \frac{2k(k+1)(\alpha+1)^2}{\alpha^2\left(1+2\sum_{j=2}^{\infty}\left(\frac{\alpha}{\alpha+1}\right)^{j-1}\frac{j!\Gamma(k+2)}{(j+1)\Gamma(k+j-1)}\right)}$ με $\alpha = \frac{\mu}{k}$.

Παρατήρηση 4.1.2. Μια τιμή του $\hat{\kappa} = \infty$ μας υποδηλώνει ότι $\hat{k}_i \leq 0$ ή $\hat{k}_i = \infty$ για τουλάχιστον ένα δείγμα. Για όλα τα δείγματα με $\hat{\kappa} < \infty$, ο εκτιμητής PM στο c_*, έχει σημαντικά χαμηλότερο $\hat{\kappa}$ από τους εκτιμητές MOM και ZTM. Η μεγαλύτερη ποσοτική διαφορά μεταξύ του εκτιμητή PM στο c_* και το συνδυασμό των εκτιμητών MOM/ZTM, παρουσιάζεται όταν $k = 1$ και $\mu = 10$.

Πίνακας 4.1: Σύγκριση εκτιμητών

	$\sqrt{u_{ML}}/k$	MLE	ZTM	PM	MOM	$\sqrt{u_{ML}}/k$	MLE	ZTM	PM	MOM	$\sqrt{u_{ML}}/k$	MLE	ZTM	PM	MOM
μ			$k = 0.01$					$k = 0.25$					$k = 0.5$		
0.1	8.27	8.78	8.78	8.78	15.21	10.05	10.38	10.38	10.38	11.50	14.01	15.34	15.34	15.34	16.25
0.5	5.90	5.91	5.91	5.91	13.70	3.56	3.62	3.63	3.62	4.76	4.14	4.06	4.11	4.07	4.86
1	5.28	5.30	5.30	5.30	13.86	2.66	2.65	2.67	2.65	4.00	2.85	2.85	2.92	2.86	3.60
5	4.41	4.48	4.48	4.48	13.75	1.77	1.80	1.83	1.81	3.34	1.70	1.77	1.89	1.78	2.66
10	4.15	4.19	4.20	4.19	13.74	1.59	1.61	1.66	1.61	3.30	1.51	1.51	1.70	1.53	2.56
			$k = 1$					$k = 3$					$k = 5$		
0.1	21.55	∞	∞	∞	∞	50.40	∞	∞	∞	∞	78.86	∞	∞	∞	∞
0.5	5.51	5.52	5.68	5.53	6.00	11.21	11.88	12.55	11.88	12.07	16.89	19.92	22.26	20.00	19.90
1	3.49	3.51	3.65	3.52	4.07	6.30	6.63	7.45	6.63	6.82	9.14	9.28	10.73	9.29	9.38
5	1.79	1.81	2.07	1.83	2.38	2.36	2.41	3.78	2.41	2.67	2.94	2.96	5.63	2.96	3.09
10	1.55	1.58	2.02	1.60	2.22	1.86	1.80	4.25	1.81	2.07	2.16	2.14	7.41	2.15	2.34

4.2 Αριθμητικές Συγκρίσεις Εκτιμητών Bayes

Στην παράγραφο αυτή μελετάμε τις αριθμητικές συγκρίσεις των εκτιμητών $Bayes$, MLE και MME της παραμέτρου k. Αυτό γίνεται μέσω προσομοίωσης η οποία περιγράφεται συνοπτικά παρακάτω.
Αρχικά για δοθέντες τιμές των παραμέτρων k, p και n, παίρνουμε τυχαίο δείγμα μεγέθους n της Αρνητικής Διωνυμικής Κατανομής $NB(k,p)$. Στη συνέχεια προσομοιώνουμε L τιμές της διασποράς σ^2 και υπολογίζουμε τις αντίστοιχες L τιμές της συνάρτησης πυκνότητας της Αντίστροφης Γάμμα κατανομής. Κατόπιν υπολογίζουμε τη μέση τιμή αυτών των τιμών δοθέντος του τυχαίου δείγματος και επαναλαμβάνοντας m φορές την παραπάνω διαδικασία παίρνουμε τις τιμές $\hat{k_1}, \hat{k_2}, \ldots, \hat{k_m}$. Υπολογίζουμε το μέσο τετραγωνικό σφάλμα $MSE(\hat{k_m}) \simeq \frac{\sum_{i=1}^{m}(\hat{k}-k)^2}{m}$. Η αποτελεσματικότητα του $\hat{k_b}$ σε σχέση με το $\hat{k_m}$, προσεγγίζεται από τη σχέση

$$\text{eff}(\hat{k_b};\hat{k_m}) = \frac{MSE(\hat{k_m})}{MSE(\hat{k_b})} \tag{4.1}$$

με ανάλογο τρόπο υπολογίζουμε την αποτελεσματικότητα του $\hat{k_b}$ σε σχέση με το $\hat{k_1}$.
Στον Πίνακα (4.2) και στον Πίνακα (4.3) καταγράφονται οι τιμές του μέσου τετραγωνικού σφάλματος MSE του εκτιμητή $Bayes$, σε σχέση με τον εκτιμητή μεθόδου ροπών $MME(\hat{k_m})$ και σε σχέση με

τον εκτιμητή μέγιστης πιθανοφάνειας $MLE(\hat{k_1})$ για τις περιπτώσεις όπου το μέγεθος δείγματος είναι ίσο με 50 και ίσο με 100. Το L το θεωρούμε ίσο με 2000 και το m ίσο με 10000. Το μέσο τετραγωνικό σφάλμα MSE, ο εκτιμητής μεθόδου ροπών MME και ο εκτιμητής μέγιστης πιθανοφάνειας MLE, έχουν χρησιμοποιηθεί από την εργασία της $Willson$ (1984). Το εύρος των μ και k, επιλέχθηκε από τους συγγραφείς έτσι ώστε να αντιπροσωπεύει το εύρος τιμών που συνήθως χρησιμοποιείται στην εντομολογία.

Πίνακας 4.2: Σύγκριση εκτιμητών για μέγεθος δείγματος ίσο με 50

μ	k	Count	MSE($\hat{k_l}$)	MSE($\hat{k_m}$)	Count	MSE($\hat{k_b}$)	eff($\hat{k_b}, \hat{k_l}$)	eff($\hat{k_b}, \hat{k_l}$)
1	1	66	18	3.6	76	0.58	31.0	6.6
	3	1397	150	23	1529	1.5	100	15.3
	5	3278	240	31	2747	4.6	52.2	6.7
3	1	0	0.15	0.22	0	0.26	0.58	0.85
	3	19	56	16	37	5.2	10.8	3.08
	5	254	740	97	267	13.2	56.8	7.5
5	1	0	0.10	0.15	0	0.16	0.56	0.83
	3	0	3.3	3.2	0	0.67	1.0	0.97
	5	15	330	52	23	14.1	23.1	3.6

Πίνακας 4.3: Σύγκριση εκτιμητών για μέγεθος δείγματος ίσο με 100

μ	k	Count	MSE($\hat{k_l}$)	MSE($\hat{k_m}$)	Count	MSE($\hat{k_b}$)	eff($\hat{k_b}, \hat{k_l}$)	eff($\hat{k_b}, \hat{k_l}$)
1	1	3	0.57	0.55	2	0.334	1.71	1.65
	3	489	150	38	515	2.26	66.4	16.8
	5	1561	240	68	1547	3.5	68.3	19.36
3	1	0	0.059	0.084	0	0.09	0.644	0.92
	3	0	2.2	2.3	2	2.16	1.02	1.07
	5	16	170	40	32	10.6	16.91	3.98
5	1	0	0.04	0.064	0	0.07	0.563	0.9
	3	0	0.85	0.94	0	0.96	0.885	0.98
	5	0	6	5.7	0	5.3	1.13	1.08

Παρατήρηση 4.2.1. Το μέσο τετραγωνικό σφάλμα των $\hat{k_b}$ και $\hat{k_m}$ αυξάνεται όταν αυξάνεται το k

και το μ παραμένει σταθερό και μειώνεται όταν το μ μειώνεται και το k παραμένει σταθερό. Συμπεραίνουμε ότι είναι δύσκολο να εκτιμήσουμε το k, όταν το μ είναι μικρό και το k μεγάλο.

Παρατήρηση 4.2.2. Η αποτελεσματικότητα του $\hat{k_b}$ είναι κατά πολύ μεγαλύτερη από 1 για μικρές τιμές του μ και μεγάλες τιμές του k. Δηλαδή προτιμούμε σε αυτή την περίπτωση τη χρήση του $\hat{k_b}$ παρά τη χρήση του $\hat{k_m}$. Για τις υπόλοιπες περιπτώσεις η αποτελεσματικότητα του $\hat{k_b}$ είναι κοντά στο 1, που σημαίνει ότι οι 2 εκτιμητές έχουν την ίδια συμπεριφορά στην εκτίμηση του k.

Παρατήρηση 4.2.3. Στην περίπτωση που έχουμε ότι το μ είναι μικρό και το k μεγάλο, το πλήθος των ακατάλληλων δειγμάτων είναι μεγάλο. Έτσι, υπάρχουν περιπτώσεις που δεν μπορεί να χρησιμοποιηθεί ο $\hat{k_m}$ σε αντίθεση με τον $\hat{k_b}$.

Παρατήρηση 4.2.4. Το μέσο τετραγωνικό σφάλμα MSE των $\hat{k_b}$ και $\hat{k_m}$ μειώνεται όταν από δείγμα μεγέθους 50 δουλεύουμε με δείγμα μεγάθους 100 εκτός από τις περιπτώσεις μεγάλου πλήθους ακατάλληλων δειγμάτων. Σε αυτή την περίπτωση και αφού απαλαχθούμε από αυτά, δεν περιμένουμε το μέσο τετραγωνικό σφάλμα που υπολογίζεται για τον $\hat{k_m}$ να συμπεριφέρεται όπως αναμέναμε, καθώς οι τιμές που υπολογίζονται δεν είναι οι πραγματικές

Παρατήρηση 4.2.5. Οι παραπάνω παρατηρήσεις ισχύουν και στην περίπτωση του εκτιμητή μέγιστης πιθανοφάνειας MSE.

Βιβλιογραφία

[1] Al-Saleh, Mohammad F. and Al-Batainah, Fatima K. (2003). Estimation of the shape parameter k of the negative binomial distribution. , *Applied Mathematics and Computation*, **143**, 431-441.

[2] Anscombe, F.J. (1950). Sampling Theory of the Negative Binomial and Logarithmic Series Distributions. , *Biometrika*, **37**, 358-382.

[3] Johnson, N.L. and Kotz, S. (1969). Descrete Distribution. , *Houghton Mifflin Company, Boston.*

[4] Ross, G.J.S. and Preece, D.A. (1985). The negative binomial distribution. , *The Statistician*, **34**, 323-336.

[5] Savani, V. and Zhigljavsky A.(2006). Efficient Estimation of Parameters of the Negative Binomial Distribution. , *Communications in Statistics-Theory and Methods*,**35**, 767-783.

[6] Wang, Yining. (1996). Estimation problems for the two-parameter negative binomial distribution. , *Statistics and Probability Letters*, **26**, 113-114.

[7] Walter W. Piegorsch (1990). Maximum Likelihood Estimation for the Negative Binomial Dispersion Parameters. , *BIOMETRICS*, **46**, 863-867.

[8] Willson,L.J. and Folks, J.L. and Young, J.H. (1986). Complete Sufficiency and Maximum Likelihood Estimation for the Two-Parameter Negative Binomial Distribution. , *Metrika*, **33**, 349-362.

[9] Anscombe, F.J. (1949). The statistical analysis of insect counts based on the negative binomial distribution. , *Biometrics*, **5**, 165-173.

[10] Bliss, C.I. and Owen, A.R.G (1958). Negative Binomial distributions with a common k. , *Biometrika*, **45**, 37-58.

[11] Clark, S.J. and Perry, J.N. (1989). Estimation of the negative binomial parameter k by maximum quasi-likelihood. , *Biometrics*, **45**, 309-316.

[12] Fisher, R.A. (1941). The negative binomial distribution. , *Annals of Eugenics*, **11**, 182-187.

[13] Haldane, J.B.S. (1941). The fitting of binomial distributions. , *Annals of Eugenics*, **11**, 179-181.

[14] Lehmann, E.L. and Scheffe, H. (1950). Completeness, similar regions and unbiased estimation, part 1. , *Sankhya*, **10**, 304-340.

[15] Oldham, P.D. (1968). Measurement in Medicine: The Interpretation of Numerical data. , *English University Press, London.*