



Πανεπιστήμιο Αιγαίου

Ανάλυση Κατηγορικών Δεδομένων

Ενότητα 5: Μέθοδος μέγιστης πιθανοφάνειας – λόγος πιθανοφανειών

Στέλιος Ζήμερας

Τμήμα Μαθηματικών

Εισαγωγική Κατεύθυνση: Στατιστικής και Αναλογιστικών – Χρηματοοικονομικών Μαθηματικών

Σάμος, Δεκέμβριος 2014



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΣΠΑ
2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

Άδειες Χρήσης

- Το παρόν εκπαιδευτικό υλικό υπόκειται σε άδειες χρήσης Creative Commons.
- Για εκπαιδευτικό υλικό, όπως εικόνες, που υπόκειται σε άλλου τύπου άδειας χρήσης, η άδεια χρήσης αναφέρεται ρητώς.



Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Αιγαίου**» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



Μέθοδος μέγιστης πιθανοφάνειας

Αν x_i $i=1,2,\dots,n$ παρατηρήσεις των X_i $i=1,2,\dots,n$, τότε έχουμε διαθέσιμο ένα δείγμα $\mathbf{X}=\{X_i, i=1,2,\dots,n\}$ της κατανομής F μεγέθους n με από κοινού σ.κ. της X

$$f(x) = \prod_{i=1}^n f(x_i)$$

Ορισμός : Θεωρούμε ένα τυχαίο δείγμα $\mathbf{X}=(X_1, X_2, \dots, X_n)$ από πληθυσμό το οποίο έχει συνάρτηση πυκνότητας πιθανότητας $f(x_i;\theta)$, με $\theta \in \Theta$. Η συνάρτηση πιθανοφάνειας (likelihood) ορίζεται από τη σχέση,

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

Μέθοδος μέγιστης πιθανοφάνειας

και ο λογάριθμος της συνάρτησης πιθανοφάνειας (log-likelihood) ορίζεται ως

$$l(\theta) = \log L(\theta) = \log \prod_{i=1}^n f(x_i; \theta) = \sum_{i=1}^n \log f(x_i; \theta)$$

Λαμβάνοντας την πρώτη παράγωγο ως προς θ

$$\frac{\partial l(\theta)}{\partial \theta} = 0$$

και αφού βεβαιωθούμε ότι η δεύτερη παράγωγος ως προς θ είναι γνησίως αρνητική, λαμβάνεται η μέγιστη τιμή της παραμέτρου. Η εκτίμηση επομένως Μέγιστης Πιθανοφάνειας είναι η λύση της εξίσωσης πιθανοφάνειας.

Μέθοδος μέγιστης πιθανοφάνειας

Ορισμός : Ο εκτιμητής $\hat{\theta} = \hat{\theta}(X)$ που ικανοποιεί τη σχέση

$$l(\hat{\theta}) = \max l(\theta)$$

είναι η παράμετρος η οποία μεγιστοποιεί την πιθανοφάνεια του δείγματος X και καλείται **Εκτιμητής Μέγιστης Πιθανοφάνειας** (Ε.Μ.Π.) του θ .

Μέθοδος μέγιστης πιθανοφάνειας

Αν, για παράδειγμα, ληφθεί ένα τυχαίο δείγμα Y_1, Y_2, \dots, Y_n από την κατανομή Poisson με παράμετρο λ , τότε η συνάρτηση πιθανοφάνειας θα είναι

$$L(\lambda) = \prod_{i=1}^n \frac{e^{-\lambda} \lambda^{y_i}}{y_i!} = \frac{e^{-n\lambda} \lambda^{\sum_{i=1}^n y_i}}{\prod_{i=1}^n y_i!}$$

Και ο λογάριθμος της συνάρτησης πιθανοφάνειας

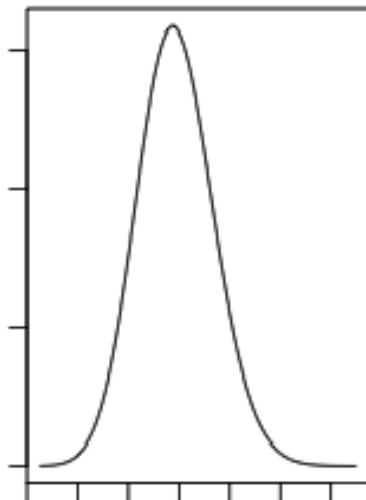
$$\ell(\lambda) = \log L(\lambda) = \log \frac{e^{-n\lambda} \lambda^{\sum_{i=1}^n y_i}}{\prod_{i=1}^n y_i!} = -n\lambda + \sum_{i=1}^n y_i \log(\lambda) - \sum_{i=1}^n \log y_i!$$

Μέθοδος μέγιστης πιθανοφάνειας

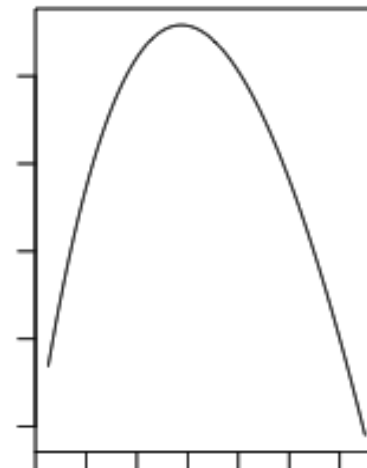
Οπότε, παίρνοντας την πρώτη παράγωγο,

$$\frac{\partial \ell(\lambda)}{\partial \lambda} = 0 \Rightarrow -n + \frac{1}{\lambda} \sum_{i=1}^n y_i = 0 \Rightarrow \hat{\lambda} = \frac{1}{n} \sum_{i=1}^n y_i = \bar{y}$$

δίνονται αντίστοιχα τα διαγράμματα για τη συνάρτηση πιθανοφάνειας και τη συνάρτηση του λογαρίθμου της πιθανοφάνειας:



likelihood



log-likelihood

Έλεγχος πολυωνυμικής με άγνωστες παραμέτρους

- Όταν οι παράμετροι π_i , $i=1,2,\dots,k$ είναι άγνωστοι, θα πρέπει να εκτιμηθούν από τα δεδομένα και μετά να χρησιμοποιηθεί το τεστ καλής προσαρμογής χ^2 όπου αντί για $e_i = n\pi_i$ γίνεται $e_i = n\hat{\pi}_i$ όπου $\hat{\pi}_i$ είναι ο εκτιμητής μέγιστης πιθανοθάνειας (Ε.Μ.Π.) των άγνωστων παραμέτρων και η χ^2 -κατανομή θα είναι χ^2_{k-1-s} όπου s ο αριθμός των εκτιμώμενων παραμέτρων.

Έλεγχος πολυωνυμικής με άγνωστες παραμέτρους

- **Θεώρημα:** Αν οι παράμετροι π_i , $i=1,2,\dots,k$ εξαρτώνται από άλλες παραμέτρους άγνωστες θ δηλ: $\pi_i = \pi_i(\theta)$ τότε:

$$x_*^2 = \sum_{i=1}^I \frac{(n_i - n\pi_i(\hat{\theta}))^2}{n\pi_i(\hat{\theta})} \sim x_{k-1-s}^2$$

Άσκηση

- Σε πρόβλημα γενετικής μια ομάδα βιολόγων προτείνει μοντέλο τριωνυμικής κατανομής με $\pi_1 = \theta^2$, $\pi_2 = 2\theta(1-\theta)$, και $\pi_3 = (1-\theta)^2$, με $0 < \theta < 1$. Αν $n = 50$ με συχνότητες $n_1 = 15$, $n_2 = 10$ και $n_3 = 25$ να δειχθεί αν τα δεδομένα ακολουθούν τριωνυμική κατανομή. Επίσης να αποδειχθεί ότι

$$-\frac{\partial^2 l(\theta)}{\partial^2 \theta} = \left(\frac{2n_1 + n_2}{\theta^2} \right) + \left(\frac{n_2 + 2n_3}{(1-\theta)^2} \right)$$

καθώς και ότι η αναμενόμενη τιμή ισούται με

$$E\left(-\frac{\partial^2 l(\theta)}{\partial^2 \theta}\right) = \left(\frac{2n}{\theta(1-\theta)} \right)$$

Απάντηση

Ελέγχουμε

$$H_0 : \pi_i = \pi_i(\hat{\theta}) \quad H_0 : \pi_i \neq \pi_i(\hat{\theta})$$

Η πιθανοφάνεια δίνεται από την σχέση:

$$L(\theta) = \frac{n!}{n_1!n_2!n_3!} \pi_1^{n_1} \pi_2^{n_2} \pi_3^{n_3} = c \theta^{2n_1} [2\theta(1-\theta)]^{n_2} (1-\theta)^{2n_3}$$

Ο λογάριθμος της πιθανοφάνειας

$$\log L(\theta) = \log c + 2n_1 \log \theta + n_2 \log 2\theta + n_2 \log(1-\theta) + 2n_3 \log(1-\theta)$$

Απάντηση

με πρώτη παράγωγο

$$\frac{\partial \log L(\theta)}{\partial \theta} = \frac{2n_1}{\theta} + \frac{n_2}{\theta} - \frac{n_2}{1-\theta} - \frac{2n_3}{1-\theta} = 0$$

Οπότε

Εκτιμητής μέγιστης πιθανοφάνειας

$$\hat{\theta} = \frac{2n_1 + n_2}{2n}$$


όπου

$$n_1 + n_2 + n_3 = n$$

Άρα

$$\hat{\theta} = \frac{2n_1 + n_2}{2n} = \frac{2 * 15 + 10}{100} = 0.4$$

Απάντηση

Οπότε παίρνουμε $\hat{\pi}_1 = \hat{\theta}^2 = 0.16$
 $\hat{\pi}_2 = 2\hat{\theta}(1 - \hat{\theta}) = 0.48$
 $\hat{\pi}_3 = (1 - \hat{\theta})^2 = 0.36$

Με αντίστοιχες αναμενόμενες εκτιμήσεις

$$e_1 = n\hat{\pi}_1 = 8$$

$$e_2 = n\hat{\pi}_2 = 24$$

$$e_3 = n\hat{\pi}_3 = 18$$

Τεστ καλής προσαρμογής

$$\chi_{*}^2 = \sum_{i=1}^I \frac{(n_i - n\pi_i(\hat{\theta}))^2}{n\pi_i(\hat{\theta})} = \frac{(15 - 8)^2}{8} + \frac{(10 - 24)^2}{24} + \frac{(25 - 18)^2}{18} = 17$$

$$\chi_{k-1-s,0.05}^2 = \chi_{3-1-1,0.05}^2 = \chi_{1,0.05}^2 = 5.024$$

5.024 < 17 άρα απορρίπτω την H_0

Απάντηση

Για την δεύτερη μερική παράγωγο έχουμε

$$\frac{\partial^2 l(\theta)}{\partial^2 \theta} = -\left(\frac{2n_1 + n_2}{\theta^2}\right) - \left(\frac{n_2 + 2n_3}{(1-\theta)^2}\right) \Rightarrow$$

$$-\frac{\partial^2 l(\theta)}{\partial^2 \theta} = \left(\frac{2n_1 + n_2}{\theta^2}\right) + \left(\frac{n_2 + 2n_3}{(1-\theta)^2}\right)$$

Και για την αναμενόμενη τιμή της 2^{ης} παραγώγου έχουμε

$$\begin{aligned} E\left(-\frac{\partial^2 l(\theta)}{\partial^2 \theta}\right) &= \frac{2E(n_1) + E(n_2)}{\theta^2} + \frac{E(n_2) + 2E(n_3)}{(1-\theta)^2} = \\ &= \frac{2n\theta^2 + 2n\theta(1-\theta)}{\theta^2} + \frac{2n\theta(1-\theta) + 2n(1-\theta)^2}{(1-\theta)^2} = \frac{2n}{\theta} + \frac{2n}{(1-\theta)} = \frac{2n}{\theta(1-\theta)} \end{aligned}$$

Μέθοδος μέγιστης πιθανοφάνειας πολυωνυμικής δειγματοληψίας

- Όταν το μέγεθος του δείγματος n είναι γνωστό αλλά τα αθροίσματα (γραμμών ή στηλών) άγνωστα τότε πολυωνυμικό μοντέλο δειγματοληψίας εφαρμόζεται. Τα κελιά (I,J) είναι τα πιθανά αποτελέσματα με σ.π.π. να ισούται με

$$\frac{n!}{\prod_{ij} n_{ij}!} \prod_i \prod_j \pi_{ij}^{n_{ij}}$$

Επομένως αναφερόμαστε στο συνολικό αριθμό του μεγέθους του δείγματος n

Μέθοδος μέγιστης πιθανοφάνειας πολυωνυμικής δειγματοληψίας

- Αν ο πίνακας συνάφειας είναι μεγέθους IJ και δεσμεύσουμε ως προς το συνολικό αριθμό του δείγματος τότε η συνάρτηση πιθανοφάνειας δίνεται από την σχέση:

$$l = \frac{n_{..}!}{\prod_{ij} n_{ij}!} \prod_i \prod_j \pi_{ij}^{n_{ij}} \propto \prod_i \prod_j \pi_{ij}^{n_{ij}}$$

με μέγιστη τιμή για $\pi_{11} = \pi_{12} = \pi_{13} = \dots = \pi_{IJ} = 1$ δηλ: όταν όλες οι πιθανότητες εμφάνισης του συνδυασμού των ενδεχομένων ισούται με μονάδα. **Αλλά αυτό είναι αδύνατον γιατί το άθροισμα των πιθανοτήτων πρέπει να ισούται με την μονάδα.**

$$\sum_i \sum_j \pi_{ij} = 1$$

Μέθοδος μέγιστης πιθανοφάνειας πολυωνυμικής δειγματοληψίας

- Επομένως στην περίπτωση μας θέλουμε να μεγιστοποιήσουμε την συνάρτηση πιθανοφάνειας l με βάση τον περιορισμό ότι:

$$\max l \rightarrow \max \log l$$

$$\sum_{i=1}^I \sum_{j=1}^J \pi_{ij} = 1$$



Περίπτωση γραμμικού
προγραμματισμού: επίλυση
συστήματος εξισώσεων κάτω
από περιορισμούς

Για την μεγιστοποίηση χρησιμοποιούμε τον λογάριθμο της συνάρτησης πιθανοφάνειας. Στις περιπτώσεις που θέλουμε να ελέγξουμε την επίδραση των περιορισμών πάνω στην διαδικασία μεγιστοποίησης της συνάρτησής μας κάνουμε χρήση των πολλαπλασιαστών Lagrange

Μέθοδος μέγιστης πιθανοφάνειας πολυωνυμικής δειγματοληψίας

Ακρότατα με περιορισμούς ισοτήτων

- Έστω μία συνάρτηση $f: \mathbb{R}^n \rightarrow \mathbb{R}$ με συνεχείς πρώτες μερικές παραγώγους και οι περιορισμοί $g(q)=0$, $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$
- Το πρόβλημα είναι να βρεθούν τα ακρότατα της f υπό τους περιορισμούς g .
- Θα μετατρέψουμε το πρόβλημα αυτό σε κάποιο που ξέρουμε να λύνουμε: στο πρόβλημα χωρίς περιορισμούς. Αυτό επιτυγχάνεται με την χρήση των πολλαπλασιαστών Lagrange.

Μέθοδος μέγιστης πιθανοφάνειας πολυωνυμικής δειγματοληψίας

Ακρότατα με περιορισμούς ισοτήτων

- Ορίζοντας την Λαγκρανζιανή (κατ' άλλους Χαμιλτονιανή) L ως

$$L(\mathbf{q}, \boldsymbol{\lambda}) = f(\mathbf{q}) + \boldsymbol{\lambda}^T \mathbf{g}(\mathbf{q})$$

όπου το διάνυσμα $\boldsymbol{\lambda}$ καλείται πολλαπλασιαστής Lagrange, και βρίσκοντας τα ακρότατά της, έχουμε επιτύχει τον στόχο μας (Θεμελιώδες θεώρημα)

$$\frac{\partial L(\mathbf{q}, \boldsymbol{\lambda})}{\partial(\mathbf{q}, \boldsymbol{\lambda})} = \mathbf{0} \Rightarrow \left\{ \begin{array}{l} \frac{\partial f}{\partial \mathbf{q}} + \frac{\partial \mathbf{g}}{\partial \mathbf{q}} \boldsymbol{\lambda}^T = \mathbf{0} \\ \mathbf{g}(\mathbf{q}) = \mathbf{0} \end{array} \right.$$

Μέθοδος μέγιστης πιθανοφάνειας πολυωνυμικής δειγματοληψίας

Ακρότατα με περιορισμούς ισοτήτων

- Η δεύτερη προϋπόθεση εξασφαλίζει την ικανοποίηση των αρχικών περιορισμών ισοτήτων.
- Η σχέση

$$\frac{\partial L^*}{\partial g} = -\lambda$$

μας δείχνει ότι οι πολλαπλασιαστές Lagrange ερμηνεύονται σαν την αλλαγή στη βέλτιστη τιμή που επιφέρει μία αλλαγή στους περιορισμούς.

Μέθοδος μέγιστης πιθανοφάνειας πολυωνυμικής δειγματοληψίας

Ακρότατα με περιορισμούς ισοτήτων

- Με βάση το πρόβλημα της μεγιστοποίησης που έχουμε να επιλύσουμε οι περιορισμοί μετατρέπονται σε

$$\max \log l$$

$$\sum_{i=1}^I \sum_{j=1}^J \pi_{ij} - 1 = 0$$

και η γενική μορφή της συνάρτησης Lagrange δίνεται από την σχέση

$$L(.) = f(.) + \lambda [g(.) - b]$$

όπου $g(.) = b$ είναι οι περιορισμοί μας

Στην περίπτωση την δική μας

$$\sum_{i=1}^I \sum_{j=1}^J \pi_{ij} = 1$$

Μέθοδος μέγιστης πιθανοφάνειας πολυωνυμικής δειγματοληψίας

Ακρότατα με περιορισμούς ισοτήτων

- Με βάση την γενική μορφή:

$$L(.) = f(.) + \lambda [g(.) - b]$$

η συνάρτηση που θέλουμε να μεγιστοποιήσουμε γίνεται (λαμβάνοντας τους λογαρίθμους):

$$L = \sum_{i=1}^I \sum_{j=1}^J n_{ij} \log \pi_{ij} - \lambda \left[\sum_{i=1}^I \sum_{j=1}^J \pi_{ij} - 1 \right]$$

Μέθοδος μέγιστης πιθανοφάνειας πολυωνυμικής δειγματοληψίας

- Παραγωγίζοντας ως προς π_{ij} , $i=1,2,\dots,I$ και $j=1,2,\dots,J$ και εξισώνοντας με το μηδέν έχουμε

$$\frac{\partial L}{\partial \hat{\pi}_{ij}} = \frac{n_{ij}}{\hat{\pi}_{ij}} - \lambda = 0$$

Αμα λύσουμε ως προς n_{ij} και αθροίσουμε επί του συνολικού δείγματος έχουμε:

$$\sum_{i=1}^I \sum_{j=1}^J n_{ij} = \lambda \sum_{i=1}^I \sum_{j=1}^J \hat{\pi}_{ij} = \lambda \Rightarrow \lambda = \sum_{i=1}^I \sum_{j=1}^J n_{ij} = n_{..}$$

Μέθοδος μέγιστης πιθανοφάνειας πολυωνυμικής δειγματοληψίας

- Αντικαθιστώντας της δεύτερη σχέση στην πρώτη έχουμε

$$\frac{n_{ij}}{\hat{\pi}_{ij}} = \lambda \quad \longleftrightarrow \quad \hat{\pi}_{ij} = \frac{n_{ij}}{n_{..}}$$
$$\lambda = n_{..}$$

- Οι εκτιμητές μέγιστης πιθανοφάνειας των πιθανοτήτων π_{ij} , είναι οι εμπειρικές εκτιμήσεις των πιθανοτήτων

Μέθοδος μέγιστης πιθανοφάνειας πολυωνυμικής δειγματοληψίας

- Κάτω από την υπόθεση της ανεξαρτησίας έχουμε

$$\pi_{ij} = \pi_{i.} * \pi_{.j}$$

η συνάρτηση που θέλουμε να μεγιστοποιήσουμε γίνεται (λαμβάνοντας τους λογαρίθμους):

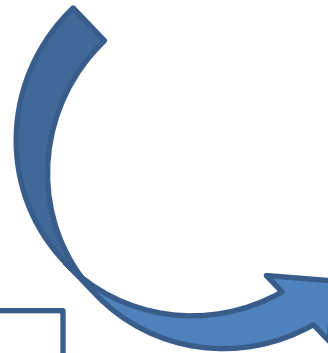
$$L = \sum_{i=1}^I \sum_{j=1}^J n_{ij} \log(\pi_{i.} * \pi_{.j}) - \lambda \left[\sum_{i=1}^I \sum_{j=1}^J (\pi_{i.} * \pi_{.j}) - 1 \right]$$

Μέθοδος μέγιστης πιθανοφάνειας πολυωνυμικής δειγματοληψίας

- Μηδενίζοντας την πρώτη παράγωγο ως προς π_i και ως π_j έχουμε:

$$\frac{\partial L}{\partial \hat{\pi}_i} = 0 \Rightarrow \sum_j \frac{n_{ij}}{\hat{\pi}_i} - \lambda = 0$$

$$\frac{\partial L}{\partial \hat{\pi}_{.j}} = 0 \Rightarrow \sum_i \frac{n_{ij}}{\hat{\pi}_{.j}} - \lambda = 0$$



$$\lambda = n_{..}$$

$$\hat{\pi}_{i.} = \frac{n_{i.}}{n_{..}}$$

$$\hat{\pi}_{.j} = \frac{n_{.j}}{n_{..}}$$

$$\hat{\pi}_{ij} = \pi_i * \pi_{.j} = \frac{n_{i.} * n_{.j}}{n_{..}^2}$$

Μέθοδος μέγιστης πιθανοφάνειας γινόμενο πολυωνυμικής

- Ας υποθέσουμε ότι σε κάθε επίπεδο της X , έστω $X = i$, διαθέτουμε n_i παρατηρήσεις. Έστω επίσης ότι οι μετρήσεις της Y σε ένα επίπεδο της X είναι ανεξάρτητες από αυτές σε ένα άλλο επίπεδο της X , έχοντας συνάρτηση πυκνότητας πιθανότητας $(\pi_{1|i}, \dots, \pi_{J|i})$. Τότε για κάθε γραμμή έχουμε ένα διαφορετικό πολυωνυμικό πείραμα. Οι μετρήσεις n_{ij} , $j = 1, \dots, J$ έχουν την πολυωνυμική κατανομή:

$$\frac{n_i!}{\prod_j n_{ij}!} \prod_j \pi_{j|i}^{n_{ij}}$$

Επομένως στην περίπτωση αυτή αναφερόμαστε σε μεταβαλλόμενα επίπεδα είτε οριζόντια είτε κάθετα στο πίνακα συνάφειας όπου το αντίθετο επίπεδο παραμένει σταθερό – “υπό συνθήκη” πιθανότητα

Μέθοδος μέγιστης πιθανοφάνειας γινόμενο πολυωνυμικής

- Για να βρούμε το μέγιστο χρησιμοποιούμε τους πολλαπλασιαστές Lagrange, όπου θέλουμε να μεγιστοποιήσουμε την συνάρτηση

$$L = \sum_{i=1}^I \sum_{j=1}^J n_{ij} \log \pi_{j|i} - \sum_{i=1}^I \lambda_i \left(\sum_{j=1}^J \pi_{j|i} - 1 \right)$$

- Παραγωγίζοντας ως προς $\pi_{j|i}$ και εξισώνοντας με το μηδέν έχουμε:

$$\frac{\partial L}{\partial \hat{\pi}_{j|i}} = \frac{n_{ij}}{\hat{\pi}_{j|i}} - \lambda_i = 0$$

Μέθοδος μέγιστης πιθανοφάνειας γινόμενο πολυωνυμικής

- Αθροίζοντας ως προς όλα τα j έχουμε:

$$\frac{n_{ij}}{\hat{\pi}_{j|i}} = \lambda_i \Rightarrow n_{ij} = \lambda_i * \hat{\pi}_{j|i} \Rightarrow \sum_j n_{ij} = \lambda_i * \sum_j \hat{\pi}_{j|i} \Rightarrow \lambda_i = n_i.$$

- Αντικαθιστώντας την δεύτερη στην πρώτη σχέση έχουμε:

$$\hat{\pi}_{j|i} = \frac{n_{ij}}{n_i}.$$

Μέθοδος μέγιστης πιθανοφάνειας γινόμενο πολυωνυμικής

- Αν ισχύει η υπόθεση της ανεξαρτησίας τότε η πιθανοφάνεια γράφεται:

$$l \propto \prod \pi_j^{n_{\cdot j}}$$

όπου π_j είναι η κοινή τιμή των j πιθανοτήτων $\pi_{j|1}, \dots, \pi_{j|I}$ οι οποίες κάτω από την μηδενική υπόθεση είναι ίσες

- Με βάση τους πολλαπλασιαστές Lagrange η συνάρτηση που πρέπει να μεγιστοποιηθεί είναι:

$$L = \sum_{i=1}^I n_{\cdot j} \log \pi_j - \lambda \left(\sum_{j=1}^J \pi_j - 1 \right)$$

Μέθοδος μέγιστης πιθανοφάνειας γινόμενο πολυωνυμικής

- Παραγωγίζοντας ως προς την π_j και μηδενίζοντας έχουμε:

$$\frac{\partial L}{\partial \hat{\pi}_j} = \frac{n_{\cdot j}}{\hat{\pi}_j} - \lambda = 0 \Rightarrow \lambda = \frac{n_{\cdot j}}{\hat{\pi}_j}$$

με λύσεις των εξισώσεων μετά από αντικατάσταση:

$$\lambda = n_{..}$$

$$\hat{\pi}_j = \frac{n_{\cdot j}}{n_{..}}$$

Άσκηση

- Έστω ότι μας ενδιαφέρει να εκτιμήσουμε το ποσοστό των ανθρώπων ενός μεγάλου ομογενούς πληθυσμού που πάσχουν από AIDS. Για τον λόγο αυτό επιλέγεται ένα τυχαίο δείγμα n ατόμων του πληθυσμού αυτού. Τα άτομα αυτά υποβάλλονται στο σχετικό τεστ. Έστω ότι από τον έλεγχο αυτό προκύπτει ότι x από τα άτομα αυτά πάσχουν από την ασθένεια. Να βρεθεί η κατανομή τους και οι αντίστοιχες εκτιμήσεις.

Απάντηση

- Λόγω της υπόθεσης ότι ο πληθυσμός είναι μεγάλος και ομοιογενής, μπορούμε να υποθέσουμε ότι τα άτομα του δείγματος που υποβλήθηκαν στον έλεγχο είναι ανεξάρτητα και κάθε ένα από αυτά έχει πιθανότητα p να πάσχει από την ασθένεια.
- Στην περίπτωση αυτή η πιθανότητα του παρατηρηθέντος ενδεχομένου (του παρατηρηθέντος δείγματος) είναι

$$P(E; \theta) = P(x \text{ από } n \text{ άτομα έχουν AIDS})$$

$$= \binom{n}{x} p^x (1-p)^{n-x}$$

Απάντηση

- Ο παραμετρικός χώρος στην περίπτωση αυτή είναι

$$\Theta = [0, 1] \text{ (δηλ. } 0 \leq p \leq 1)$$

- Ως συνάρτηση πιθανοφάνειας μπορεί να ορισθεί οποιοδήποτε θετικό πολλαπλάσιο της $P(E; p)$

$$L(p) = p^x (1 - p)^{n-x}, \quad 0 \leq p \leq 1$$

$$\ell(p) = \ln L(p) = x \ln p + (n - x) \ln(1 - p)$$

$$\ell'(p) = \frac{x}{p} - \frac{n-x}{1-p}, \quad \ell''(p) = -\frac{x}{p^2} - \frac{n-x}{(1-p)^2}, \quad p \neq 0, p \neq 1$$

Απάντηση

Για $1 \leq x \leq n-1$, η εξίσωση $l'(p)=0$ έχει ως μοναδική λύση την $p = \frac{x}{n}$.

Δεδομένου ότι $l''(p) < 0$ για $p = \frac{x}{n}$ το σημείο αυτό είναι ένα (τοπικό) μέγιστο.

Εξάλλου, δοθέντος ότι $L(p)=0$ για $p=0$ και για $p=1$, το $p = \frac{x}{n}$ είναι το (ολικό) απόλυτο μέγιστο και επομένως

$$\hat{p} = \frac{x}{n}$$

Απάντηση

- Αν $x=0$, η εξίσωση $l'(p) = 0$ δεν έχει λύση και το μέγιστο εμφανίζεται στο σύνορο του παραμετρικού χώρου $\Theta=[0,1]$.
- Στην περίπτωση αυτή έχουμε

$$P(E; \theta) = (1 - p)^n \text{ για } 0 \leq p \leq 1$$

- Προφανώς, η συνάρτηση αυτή παίρνει την μέγιστη τιμή της όταν $p=0$.
- Ομοίως, για $x=n$ θα είναι $p = 1$ και έτσι, τελικά,

$$\hat{p} = \frac{x}{n} \text{ για } x = 0, 1, 2, \dots, n$$

Άσκηση

Στον στατιστικό έλεγχο ποιότητας ενός εργοστασίου, προκειμένου να ελεγχθεί η ποιότητα των προϊόντων, επιλέγονται καθημερινά και επί 200 συνεχείς ημέρες 10 αντικείμενα με τυχαίο τρόπο και ελέγχονται για πιθανές ατέλειες. Τα αποτελέσματα δίνονται στον πίνακα που ακολουθεί:

Αριθμός ελαττωματικών	0	1	2	3	≥ 4	Σύνολο
Παρατηρηθείσα συχνότητα	133	52	12	3	0	200

Έστω p η πιθανότητα ένα αντικείμενο να είναι ελαττωματικό. Να βρεθεί η εκτιμήτρια μεγίστης πιθανοφάνειας του p .

Απάντηση

Μπορούμε να υποθέσουμε ότι ο αριθμός των ελαττωματικών αντικειμένων σε κάθε παρτίδα 10 αντικειμένων ακολουθεί την διωνυμική κατανομή με παράμετρο p . Στην περίπτωση αυτή, η πιθανότητα P_j στα 10 αντικείμενα να παρατηρηθούν j ελαττωματικά δίνεται από την διωνυμική κατανομή.

$$P_j = P(X = j) = \binom{10}{j} p^j (1-p)^{10-j} \quad j=0, 1, \dots, 10$$

Η πιθανότητα να παρατηρηθούν 4 ή περισσότερα ελαττωματικά αντικείμενα στο δείγμα των 10 αντικειμένων δίνεται από τον τύπο

$$P(X \geq 4) = 1 - [P(X=0) + P(X=1) + P(X=2) + P(X=3)]$$

Απάντηση

Για τα συγκεκριμένα δεδομένα έχουμε ότι η συνάρτηση πιθανοφάνειας της παραμέτρου p παίρνει την γενική μορφή:

$$l = \frac{n_{..}!}{\prod_{ij} n_{ij}!} \prod_i \prod_j \pi_{ij}^{n_{ij}} \propto k \prod_i \prod_j \pi_{ij}^{n_{ij}}$$

με $L(p) = k [P(X=0)]^{133} [P(X=1)]^{52} [P(X=2)]^{12} [P(X=3)]^3 [P(X \geq 4)]^0$
 $= k P_0^{133} P_1^{52} P_2^{12} P_3^3 P_4^0, \quad 0 \leq p \leq 1$



Αριθμός ελαττωματικών	0	1	2	3	≥ 4	Σύνολο
Παρατηρηθείσα συχνότητα	133	52	12	3	0	200

Απάντηση

Αντικαθιστώντας τις τιμές των P_0, P_1, P_2, P_3, P_4 , έχουμε

$$L(p) = \left[(1-p)^{10} \right]^{133} \left[p(1-p)^9 \right]^{52} \left[p^2(1-p)^8 \right]^{12} \left[p^3(1-p)^7 \right]^3$$
$$= p^{85} (1-p)^{1915}$$

Γενική μορφή $P_j = P(X = j) = \binom{10}{j} p^j (1-p)^{10-j} \quad j=0, 1, \dots, 10$

$$j=0 \Rightarrow P_0 = P(X = 0) = \binom{10}{0} p^0 (1-p)^{10-0} \propto (1-p)^{10}$$

Απάντηση

Η συνάρτηση πιθανοφάνειας είναι ο λόγος των παρατηρηθεισών ευνοϊκών περιπτώσεων διά του συνολικού αριθμού των δοκιμών

$$\hat{p} = \frac{x}{n}$$

Άρα $x=85$, $n=2000=1915+85$

$$\begin{aligned} L(p) &= \left[(1-p)^{10} \right]^{133} \left[p(1-p)^9 \right]^{52} \left[p^2(1-p)^8 \right]^{12} \left[p^3(1-p)^7 \right]^3 \\ &= p^{85} (1-p)^{1915} \end{aligned}$$

Επομένως

$$\hat{p} = \frac{x}{n} = \frac{85}{2000} = 0.0425$$

Απάντηση

Χρησιμοποιώντας την τιμή αυτή μπορούμε να βρούμε τις αναμενόμενες συχνότητες

$$nP_j = 200 * P_j, j = 0,1,2,3$$

Οι αναμενόμενες συχνότητες δίνονται στον πίνακα που ακολουθεί:

Αριθμός Ελαττωματικών	0	1	2	3	≥4	Σύνολο
Παρατηρούμενη Συχνότητα	133	52	12	3	0	200
Αναμενόμενη Συχνότητα	129.54	57.50	11.48	1.36	0.12	200

Άσκηση

- Έστω η διακριτή τ.μ. $X \sim f(x; \theta) = \theta^2(x+1)(1-\theta)^x$, $x = 0, 1, \dots$
όπου $\theta \in \Omega = \{0, 1\}$. Να βρεθεί ο Ε.Μ.Π. της θ

$$L(\theta) = \theta^{2n} \left[\prod_{i=1}^n (x_i + 1) \right] (1-\theta)^{\sum_{i=1}^n x_i}$$

$$l(\theta) = 2n \log \theta + \left(\sum_{i=1}^n x_i \right) \log(1-\theta) + \log \prod_{i=1}^n (x_i + 1)$$

$$\frac{\partial l(\theta)}{\partial \theta} = \frac{2n}{\theta} - \frac{\sum_{i=1}^n x_i}{(1-\theta)} = 0 \Rightarrow \theta = \frac{2}{2 + \bar{x}} \Rightarrow \bar{x} = \frac{2(1 - \hat{\theta})}{\hat{\theta}}$$

$$\frac{\partial^2 l(\theta)}{\partial^2 \theta} = -\frac{2n}{\left(\frac{2}{2 + \bar{x}}\right)^2} - \frac{\sum_{i=1}^n x_i}{\left(\frac{\bar{x}}{2 + \bar{x}}\right)} < 0$$

Άσκηση

- Έστω $X_i, i=1,2,\dots,n$ τ.δ. από την κατανομή $B(N,\theta)$ με $\theta \in (0,1)$. Να βρεθεί ο Ε.Μ.Π. Έστω τ.δ. από την κατανομή $B(3,\theta)$ με τιμές $x_1=3, x_2=0, x_3=1, x_4=2, x_5=1$, να βρεθεί ο Ε.Μ.Π. της $P(X=0)$

$$L(\theta) = \theta^{\sum_{i=1}^n x_i} (1-\theta)^{nN - \sum_{i=1}^n x_i} \prod_{i=1}^n \binom{N}{x_i},$$

$$l(\theta) = \left(\sum_{i=1}^n x_i \right) \ln \theta + \left(nN - \sum_{i=1}^n x_i \right) \ln(1-\theta) + \ln \prod_{i=1}^n \binom{N}{x_i},$$

$$\frac{d}{d\theta} l(\theta) = \frac{\sum_{i=1}^n x_i}{\theta} - \frac{nN - \sum_{i=1}^n x_i}{1-\theta} = 0 \Rightarrow \theta = \frac{\bar{x}}{N}.$$

$$\frac{d^2}{d^2\theta} l(\theta) = -\frac{nN^2}{\bar{x}} - \frac{n(N-\bar{x})}{\left(1 - \frac{\bar{x}}{N}\right)^2} < 0$$

Άσκηση

Επομένως η $\hat{\theta} = \frac{\bar{X}}{N}$ είναι η Ε.Μ.Π. του θ .

Από τις μετρήσεις $x_1=3, x_2=0, x_3=1, x_4=2, x_5=1$ παίρνουμε ότι $\hat{\theta} = \frac{\bar{x}}{N} = 0.466$.

$$P(X = 0) = \binom{3}{0} 0.466^0 (1 - 0.466)^{3-0} = 0.15227$$

Λογος πιθανοφανειων πολυωνυμικη δειγματοληψια

- Οι έλεγχοι αυτοί στηρίζονται στη συνάρτηση μέγιστης πιθανοφάνειας. Έστω θ το σύνολο των παραμέτρων του υποδείγματος και $L(\theta)$ η συνάρτηση πιθανοφάνειας. Με τη μέθοδο της μέγιστης πιθανοφάνειας εκτιμώνται (i) το υπόδειγμα χωρίς περιορισμούς $L(\theta_1)$ και (ii) το υπόδειγμα με τους περιορισμούς, $L(\theta_2)$.

Ο έλεγχος του λόγου συναρτήσεων πιθανοφάνειας προκύπτει από τη διαίρεση:

$$\Lambda = \frac{\max L(\theta_2)}{\max L(\theta_1)}, \lambda \in [0,1]$$

Λογος πιθανοφανειων πολυωνυμικη δειγματοληψία

- Η τιμή του λόγου Λ είναι απαραίτητα μικρότερη της μονάδας, αφού η τιμή του αριθμητή δεν μπορεί να είναι μεγαλύτερη από αυτήν του παρονομαστή. Εάν η τιμή του Λ είναι πλησίον του μηδενός, τότε η μηδενική υπόθεση δεν μπορεί να γίνει αποδεκτή, που σημαίνει ότι οι περιορισμοί δεν ασκούν επίδραση στη μέγιστη τιμή της συνάρτησης πιθανοφάνειας.
- Αντίθετα, εάν πλησιάζει τη μονάδα, τότε οι περιορισμοί ασκούν επίδραση και η μηδενική υπόθεση μπορεί να γίνει αποδεκτή.

Λογος πιθανοφανειων πολυωνυμικη δειγματοληψια

- Η μεγιστοποιημένη πιθανοφάνεια κάτω από την μηδενική υπόθεση H_0 γίνεται:

$$\frac{n_{..}!}{n_{11}! \cdots n_{IJ}!} \prod_{i=1}^I \prod_{j=1}^J \left(\frac{n_{i.} n_{.j}}{n_{..} n_{..}} \right)^{n_{ij}}$$

$\hat{\pi}_{ij}$

$$\hat{\pi}_{ij} = \pi_{i.} * \pi_{.j} = \frac{n_{i.} * n_{.j}}{n_{..}^2}$$

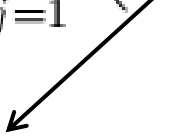
Υπόθεση ανεξαρτησίας

Λογος πιθανοφανειων πολυωνυμικη δειγματοληψια

- Η μεγιστοποιημένη πιθανοφάνεια κάτω από την υπόθεση H_1 των περιορισμών γίνεται:

$$\frac{n_{..}!}{n_{11}! \cdots n_{IJ}!} \prod_{i=1}^I \prod_{j=1}^J \left(\frac{n_{ij}}{n_{..}} \right)^{n_{ij}}$$

$\hat{\pi}_{ij}$



αναφερόμαστε στο συνολικό αριθμό του
μεγέθους του δείγματος n

Λόγος πιθανοφανειών πολυωνυμική δειγματοληψία

- Ο λόγος πιθανοφανειών δίνεται από την σχέση

$$\Lambda = \frac{\prod_i \prod_j (n_{i \cdot} n_{\cdot j})^{n_{ij}}}{n^{n_{\cdot\cdot}} \prod_i \prod_j n_{ij}^{n_{ij}}}$$

- Η στατιστική συνάρτηση

$$G^2 = -2 \log \Lambda \sim \chi_{d_1 - d_0}^2$$

d_k είναι ο αριθμός των εκτιμώμενων παραμέτρων κάτω από από την υπόθεση H_k για $k = 0, 1$.

Λογος πιθανοφανειων πολυωνυμικη δειγματοληψια

- Οταν έχουμε ΙJ πίνακες οι υποθέσεις είναι
 H_0 : X, Y ανεξάρτητες μεταβλητές έναντι της
 H_1 : X, Y εξαρτημένες μεταβλητές

Οταν η H_1 ισχύει, πρέπει να εκτιμήσουμε όλες τις από κοινού πιθανότητες π_{ij}

Όταν η H_0 ισχύει, πρέπει να εκτιμήσουμε τις περιθώριες κατανομές $\pi_{i \cdot}$ και $\pi_{\cdot j}$

Λογος πιθανοφανειων πολυωνυμικη δειγματοληψια

- Υποθέτουμε πολυωνυμική κατανομή με

$$L_0 = \prod_i \prod_j (\pi_{i.} * \pi_{.j})^{n_{ij}}$$

$$\log L_0 = \sum_i \sum_j n_{ij} \log \pi_{i.} + \sum_i \sum_j n_{ij} \log \pi_{.j}$$

$$L_1 = \prod_i \prod_j (\pi_{ij})^{n_{ij}}$$

$$\log L_1 = \sum_i \sum_j n_{ij} \log \pi_{ij}$$

Λογος πιθανοφανειων πολυωνυμικη δειγματοληψια

- Επομένως

$$\begin{aligned}
 G^2 &= -2(\log L_0 - \log L_1) = -2 \sum_i \sum_j n_{ij} \log \frac{\pi_{i.} * \pi_{.j}}{\pi_{ij}} = \\
 &= -2 \sum_i \sum_j n_{ij} \log \frac{e_{ij}}{\pi_{ij}} = -2 \sum_i \sum_j n_{ij} \log \frac{e_{ij}}{n \pi_{ij}} \Rightarrow \\
 \Rightarrow G^2 &= -2 \sum_i \sum_j n_{ij} \log \frac{e_{ij}}{n_{ij}} \sim \chi_{(I-1)(J-1)}^2
 \end{aligned}$$


$\pi_{i.} * \pi_{.j} = \frac{n_{i.} * n_{.j}}{n^2} = \frac{n_{i.} * n_{.j}}{n} \frac{1}{n} = \frac{e_{ij}}{n}$

$\pi_{ij} = \frac{n_{ij}}{n} \Rightarrow n_{ij} = n \pi_{ij}$

Λογος πιθανοφανειων γινομενο πολυωνυμικων

- Η μεγιστοποιημένη πιθανοφάνεια κάτω από την μηδενική υπόθεση H_0 γίνεται:

$$\frac{n_{..}!}{n_{11}! \cdots n_{IJ}!} \prod_{j=1}^J \left(\frac{n_{.j}}{n_{..}} \right)^{n_{.j}}$$

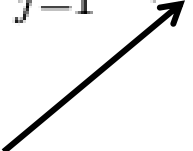
$\hat{\pi}_j$ 

Υπόθεση ανεξαρτησίας

Λογος πιθανοφανειων γινομενο πολυωνυμικων

- Η μεγιστοποιημένη πιθανοφάνεια κάτω από την υπόθεση H_1 των περιορισμών γίνεται:

$$\frac{n_{..}!}{n_{11}! \cdots n_{IJ}!} \prod_{i=1}^I \prod_{j=1}^J \left(\frac{n_{ij}}{n_{i.}} \right)^{n_{ij}}$$

$\hat{\pi}_{ji}$ 

Επομένως στην περίπτωση αυτή αναφερόμαστε σε μεταβαλλόμενα επίπεδα είτε οριζόντια είτε κάθετα στο πίνακα συνάφειας όπου το αντίθετο επίπεδο παραμένει σταθερό – “υπό συνθήκη” πιθανότητα

Λόγος πιθανοφανειών γινομενο πολυωνυμικών

- Ο λόγος πιθανοφανειών δίνεται από την σχέση

$$\Lambda = \frac{\prod_{j=1}^J n_{.j}^{n_{.j}} \prod_{i=1}^I n_{i.}^{n_{i.}}}{n_{..}^{n_{..}} \prod_{i=1}^I \prod_{j=1}^J n_{ij}^{n_{ij}}}$$

- Η στατιστική συνάρτηση

$$G^2 = -2 \sum_i \sum_j n_{ij} \log \frac{e_{ij}}{n_{ij}} \sim \chi_{(I-1)(J-1)}^2$$

Ιδιότητες

- Τα στατιστικά τεστ χ^2 και G^2 είναι ίσα ασυμπτωτικά.
- Το στατιστικό τεστ χ^2 συγκλίνει γρηγορότερα στην χ^2 – κατανομή από ότι το G^2 .
- Η χ^2 σύγκλιση της κατανομής του G^2 δεν είναι καλή όταν

$$\frac{n_{..}}{IJ} \leq 5$$

- Αντίθετα το στατιστικό τεστ χ^2 προσεγγίζεται καλύτερα από την χ^2 – κατανομή ακόμα και για

$$\frac{n_{..}}{IJ} = 1$$

υπό την προϋπόθεση ο πίνακας να μην παίρνει πολύ μικρές τιμές ή σχετικά μεγάλες αναμενόμενες συχνότητες.

Άσκηση

- Σε μια προοπτική μελέτη κατά την οποία εξετάσθηκαν 368 άνδρες καπνιστές ηλικίας κάτω των 60 ετών οι οποίοι έπαθαν μια καρδιακή ανακοπή και επιβίωσαν. Μετά από 2 έτη εξετάσθηκαν πόσοι από αυτούς είχαν επιβιώσει και τους χωρίσαμε ανάλογα εάν είχαν κόψει το τσιγάρο ή όχι. Έτσι εδώ μας ενδιαφέρει να εξετάσουμε αν το σταμάτημα του καπνίσματος (X) είχε ευνοϊκή επίδραση στην επιβίωση μετά από δύο έτη (Y). Τα δεδομένα δίνονται στον 2×2 Πίνακα που ακολουθεί:

X: Συνέχισαν το κάπνισμα	Y: Επιβίωση σε 2 χρόνια		Σύνολο
	1: Πεθαμένος	2: Ζωντανός	
1: Ναι	19 (12.3%)	135 (87.7%)	154 (41.8%)
2: Όχι	15 (7.0%)	199 (93.0%)	214 (58.2%)
Σύνολο	34 (9.2%)	334 (90.8%)	368

Άσκηση

Έχουμε:

$$\begin{aligned}x_{obs}^2 &= \frac{(19 - 14.23)^2}{14.23} + \frac{(135 - 139.77)^2}{139.77} + \frac{(15 - 19.77)^2}{19.77} + \frac{(199 - 194.23)^2}{194.23} \\ &= 3.03 \Leftrightarrow p\text{-value} = 0.082\end{aligned}$$

άρα δεν απορρίπτουμε την υπόθεση ανεξαρτησίας.

$$\begin{aligned}G_{obs}^2 &= -2 \left(\begin{aligned} &19 \log 14.231 + 135 \log 139.771 + \\ &15 \log 19.771 + 199 \log 194.231 \end{aligned} \right) = \\ &= 2.98 \Leftrightarrow p\text{-value} = 0.0842\end{aligned}$$

άρα δεν απορρίπτουμε την υπόθεση ανεξαρτησίας.