Μαθηματικά Αγοράς
Πληροφορικής

Μαθηματικά Υποδείγματα Εμπορολογίας

Σεπτέμβριος 2009

Σημειώσεις Παραδόσεων

Ε. Φ. Μαγείρου

- Υποδείγματα συμπεριφοράς
- Μεικτή Πιθανοφάνεια
- Κύριες Συνιστώσες
- Yield Management

Υποδείγματα Συμπεριφοράς Καταναλωτή

Το υπόδειγμα NBD (Negative Binomial Distribution)

· Βασική παραδοχή: θα γίνει πώληση σε ένα "σύντομο" διάστημα Δt με πιθανότητα $\lambda \Delta t$.

· Σε ένα πεπερασμένο διάστημα διάρκειας T η πιθανότητα καμμίας πώλησης είναι $(1-\lambda T/N)^N$ (γιατί; θεωρείστε $\Delta t = 1/N$) και στο όριο

$$\lim_{N\to\infty}\left(1-\frac{\lambda T}{N}\right)^N = e^{-\lambda T}$$

· Έτσι ο χρόνος μέχρι την επόμενη πώληση έχει πυκνότητα κατανομής $f(T|\lambda) = \lambda e^{-\lambda T}$ (γιατί;)

· Η πιθανότητα x πωλήσεων σε ένα διάστημα T είναι (χωρίζοντας το διάστημα σε N υποδιαστήματα)

$$P_{Poisson}(x \text{ πωλήσεις}) = \binom{N}{x}\left(\frac{\lambda T}{N}\right)^x\left(1-\frac{\lambda T}{N}\right)^{N-x}$$

που για $N \to \infty$ συγκλίνει στο

$$P_{Poisson}(\tilde{X}=x) = \lambda^x e^{-\lambda T}/x!$$

· Έτσι αν κάνουμε παρατηρήσεις των πωλήσεων σε καταναλωτές $i=1,2,\ldots,N$ που όλοι υπακούουν στο υπόδειγμα αυτό με το ίδιο λ περιμένουμε ότι

$$N_x/N \simeq \lambda^x e^{-\lambda T}/x!$$

N_x: ο αριθμός των καταναλωτών που πραγματοποίησαν x αγορές.

Συχνά όμως τα εμπειρικά δεδομένα δεν επιβεβαιώνουν την θεωρία αυτή, ακόμα και αν το λ εκτιμηθεί με στατιστικά σωστές μεθόδους (π.χ. το λ μπορεί να εκτιμηθεί με την μέθοδο της μέγιστης πιθανοφάνειας αλλά στην συνέχεια ο έλεγχος καταλληλότητας χ^2 να απορρίψει την υπόθεση ότι τα στοιχεία προέκυψαν από αυτές τις κατανομές).

Μία εύλογη παραδοχή είναι αυτή της ετερογένειας του πληθυσμού, δηλαδή ότι το λ είναι κατανεμημένο στον πληθυσμό με πυκνότητα $g(\lambda \mid r_1, r_2, \ldots, a_1, a_2, \ldots)$ όπου $r_1, r_2, \ldots, a_1, a_2, \ldots$ είναι άγνωστοι παράμετροι. Τότε η πιθανότητα πωλήσεων x σε ένα καταναλωτή επιλεγμένο τυχαία είναι

$$P(\tilde{X}=x \mid r_1, r_2, \ldots, a_1, a_2, \ldots) = \int_0^\infty p_P(X=x \mid \lambda)\, g(\lambda \mid r \ldots a)\, d\lambda$$

Έτσι έχουμε στην διάθεσή μας περισσότερες παραμέτρους για προσαρμογή στα στοιχεία. Προφανώς πρέπει να γίνει οικονομία στο πλήθος των παραμέτρων, καθώς υπερβολικός αριθμός αυξάνει το σφάλμα εκτίμησης.

Εμπειρικά, μια καλή κατανομή του λ είναι η Γάμμα (δύο παραμέτρων)

$$g(\lambda \mid r, a) = \frac{a^r \lambda^{r-1} e^{-a\lambda}}{\Gamma(r)} \qquad \lambda \geq 0$$

(Βλέπε LKM Appendix A)

Στην περίπτωση αυτή

$$P(\tilde{X}=x \mid r, a) = \int_0^\infty \frac{\lambda^x e^{-\lambda}}{x!} \cdot \frac{a^r \lambda^{r-1} e^{-a\lambda}}{\Gamma(r)}\, d\lambda = \frac{a^r}{x!\,\Gamma(r)} \int_0^\infty \lambda^{r+x-1} e^{-(a+1)\lambda}\, d\lambda$$

(Η συνάρτηση Γ ορίζεται ως $\Gamma(z) = \int_0^\infty e^{-t}\, t^{z-1}\, dt \quad z>0$)

Η κατανομή Γάμμα έχει $E(\tilde{X}) = r/a$ $\quad Var(\tilde{X}) = r/a^2$

Παρατηρούμε ότι το ολοκλήρωμα προηγουμένως είναι τύπου αναμενόμενης τιμής μιάς κατανομής Γ, (με παραμέτρους $\hat{r} = r+x$ $\quad \hat{a} = a+1$) οπότε προκύπτει μετά από πράξεις (κάντε τις...)

$$P_{NBD}(X=x \mid r,a) = \binom{x+r-1}{x}\left(\frac{a}{a+1}\right)^r\left(\frac{1}{a+1}\right)^x \qquad x=0,1,\ldots$$

$a, r > 0$, $\quad r$: ακέραιος

Η κατανομή αυτή ονομάζεται Negative Binomial Distribution (βλέπε LKM App A) (είναι η πιθανότητα x αποτυχιών πριν r επιτυχίες σε Bernoulli δοκιμές με $p = a/(a+1)$). Γιά αυτήν ισχύει

$$E_{NBD}(\tilde{X}) = E(\lambda) = r/a$$

$$Var_{NBD}(\tilde{X}) = r/a + r/a^2$$

$$\left(= Var(\lambda) + Var(\tilde{X} \mid \lambda)\right)$$

Έτσι αν έχουμε μετρήσεις N_x από ημιδιορό, μπορούμε να εκτιμήσουμε τις παραμέτρους r, a με τις $\hat{r}, \hat{a}$ έτσι ώστε:

$$\hat{r}/\hat{a} = \sum_{x=0,1,\ldots} x \frac{N_x}{N} = \hat{x}$$

και

$$\hat{r}/\hat{a} + \hat{r}/\hat{a}^2 \cong \sum_{x=0,1,\ldots} (x-\hat{x})^2 \frac{N_x}{N}$$

Αυτό που έχει ιδιαίτερη σημασία στις εφαρμογές είναι ο υπολογισμός της ποσότητας

$$E(\tilde{X}_2 \mid \tilde{X}_1 = x)$$

όπου $\tilde{X}_2$ οι αναμενόμενες αγορές την $2^η$ περίοδο αν ο καταναλωτής την $1^η$ περίοδο έχει πραγματοποιήσει x αγορές ($\tilde{X}_1 = x$)

Η απάντηση του θέματος έχει ως εξής: Αν κάποιος πελάτης έχει κάνει x αγορές, το λ που τον αντιστοιχεί <u>δεν</u> περιγράφεται από την $g(\lambda \mid r, a)$ αλλά από την πυκνότητα

$$g(\lambda \mid r, a, \tilde{X}_1 = x) = \frac{P(\lambda, r, a, \tilde{X}_1 = x)}{P(r, a, \tilde{X}_1 = x)}$$

$$= \frac{P(\tilde{X}_1 = x \mid \lambda, r, a)\, g(\lambda \mid r, a)}{P(\tilde{X}_1 = x \mid r, a)}$$

$$= \frac{\lambda^x e^{-\lambda} \cdot a^r \lambda^{r-1} e^{-a\lambda}}{x! \binom{x+r-1}{x} \left(\frac{a}{a+1}\right)^r \left(\frac{1}{a+1}\right)^x \Gamma(r)}$$

$$= \frac{\lambda^{x+r-1} e^{-(a+1)\lambda} (a+1)^{r+x}}{\cancel{x!} \frac{(x+r-1)!}{\cancel{x!}\,\cancel{(r-1)!}} \cancel{(r-1)!}}$$

$$= \frac{(a+1)^{r+x} \lambda^{r+x-1} e^{-(a+1)\lambda}}{\Gamma(r+x)}$$

που είναι κατανομή Γ με παραμέτρους $\hat{a} = a+1$ $\hat{r} = r+x$ (Αν γνωρίζουμε n ανεξάρτητες μετρήσεις του X, η $g(\lambda \mid r, a, x_1, \ldots, x_n)$ είναι Γ με $\hat{a} \leftarrow a+n$ $r \leftarrow r + \sum_{i=1}^{n} x_i$)

Έτσι η $\tilde{X}_2$ έχει κατανομή Poisson με λ κατανεμημένο κατά Γ με $\hat{a} \leftarrow a+1$ $\hat{r} \leftarrow r+x$

και έτσι $E(\tilde{X}_2|\tilde{\lambda}) = \frac{\hat{r}}{\hat{a}} = \frac{r+x}{a+1} = \frac{r}{a+1} + \frac{x}{a+1}$

όπου το $\hat{\lambda}$ έχει την $\Gamma(\hat{a}, \hat{r})$.

Στην εφαρμογή που αναφέρεται στο LKM σελ. 36 αναφέρεται ότι $N_{X_1=0} = 880$. Οι ίδιοι καταναλωτές την δεύτερη περίοδο πραγματοποίησαν 185 αγορές. Όμως από το $E(X_2|X_1=0) = \frac{r}{a+1}$ προκύπτει ότι ο αριθμός των πωλήσεων που αναμένονται είναι μόνο 53. Άρα η απόκλιση 185 σε σχέση με 53 οφείλεται σε ενδεχόμενη διαφήμιση που επέτυχε να προσελκύσει <u>νέους</u> πελάτες

<u>Άσκηση</u> Στον πίνακα της σελ 36 του LKM ποια είναι τα εκτιμηθέντα a, r; Επιβεβαιώστε τα υπόλοιπα στοιχεία του ίδιου πίνακα.

Ένα συγγενές με τα παραπάνω θέμα είναι η θεωρία των <u>Συνθέτων Κατανομών</u> (Βλέπε Appendix A, Mixing Distributions Example 2, P. 613-614)

Έστω ότι η έκβαση ενός εγχειρήματος μας εξαρτάται από την (άγνωστη) τιμή μιας παραμέτρου θ για την οποία εκτιμούμε αρχικά ότι συμπεριφέρεται όπως μία τυχαία μεταβλητή με πυκνότητα $\xi(\theta)$ (Η θ μπορεί να είναι διάνυσμα.) Αν τώρα κάνουμε μια εκτίμηση για την θ, η εκτίμηση γίνεται [illegible] με την παρατήρηση μιας άλλης τυχαίας μεταβλητής $\tilde{X}$ της οποίας η πυκν. κατανομής εξαρτάται από την τιμή της Θ. Η κατανομή δηλαδή είναι $f(x|\theta)$. (Το $\tilde{X}$ ονομάζεται πείραμα, μέτρηση κλπ.). Σε πολλές περιπτώσεις η απόφαση μετά την παρατήρηση του $\tilde{X}$ θα εξαρτηθεί από την εκ των υστέρων (a posteriori) κατανομή της Θ, $\xi(\theta/x)$

[<u>Παρατήρηση</u> Η σημασία της a posteriori έγκειται στο εξής: Έστω συνάρτηση κέρδους $K(a,\theta)$ που εξαρτάται από την <u>δράση</u> a και την τιμή της Θ. Όμως αν κάνουμε παρατήρηση $\tilde{X}=x$ και έχουμε μιά <u>στρατηγική</u> δ (που προσδιορίζει την δράση σαν συνάρτηση του x, $\delta(x)=a_x$) το αναμενόμενο κέρδος είναι

$$\iint K(a_x,\theta)\,\xi(x,\theta)\,d\theta\,dx$$

που ισούται με

$$\int_X \xi(x)\left[\int_\Theta K(a_x,\theta)\,\xi(\theta/x)\,d\theta\right]dx$$

Προφανώς αν προσπαθήσουμε για <u>κάθε</u> x να βρούμε το βέλτιστο a_x η βελτιστοποίηση επιμερίζεται σε κάθε x, είναι δηλαδή:

$$\max_{a:\tilde{X}\to A}\int \xi(x)\left(\int K(a_x,\theta)\xi(\theta/x)d\theta\right)dx = \int \xi(x)\left[\max_a \int K(a,\theta)\xi(\theta/x)d\theta\right]dx$$

Άρα το κριτήριο επιλογής απόφασης είναι η μεγιστοποίηση $\max_a \int K(a,\theta)\,p(\theta/x)\,d\theta$

που είναι η μεγιστοποίηση της δεσμευμένης αναμενόμενης τιμής με την δέσμευση $\tilde{X}=x$, και όπου υπεισέρχεται η a posteriori $\xi(\theta/x)$

Η αρχική κατανομή $\xi(\theta)$ ονομάζεται a priori (εκ των προτέρων)]

Από το θεώρημα Bayes είναι

$$\xi(\theta/x) = \frac{\xi(\theta,x)}{\xi(x)} = \frac{f(x/\theta)\,\xi(\theta)}{\int_\Theta f(x/\theta)\,\xi(\theta)\,d\theta}$$

Αν θεωρήσουμε την $\xi(\theta/x)$ ως συνάρτηση μόνο του θ, μπορούμε να πούμε ότι για τον προσδιορισμό της αρκεί μόνο ο "αριθμητής" $\xi(\theta,x) = f(x/\theta)\,\xi(\theta)$ καθώς ο παρονομαστής προσδιορίζεται και από το γνωστό $\int_\Theta \xi(\theta/x)\,d\theta = 1$.
Αν δηλαδή γράψουμε $\xi(\theta/x) = \mu\, f(x/\theta)\,\xi(\theta)$ το μ προσδιορίζεται ως $\mu \int f(x/\theta)\,\xi(\theta)\,d\theta = 1$ ή $\mu = 1 \Big/ \left[\int f(x/\theta)\,\xi(\theta)\,d\theta\right]$

Μια οικογένεια κατανομών $g(\theta;\sigma)$ με παραμέτρους σ ονομάζεται συζυγής ως προς την μέτρηση $f(x/\theta)$ αν

$$g(\theta/\hat{\sigma}(x,\sigma)) \propto f(x/\theta)\,g(\theta/\sigma)$$

όπου $\propto$ υποδηλώνει αναλογικότητα.
Αυτό σημαίνει ότι η εκ των υστέρων $\xi(\theta/x)$ είναι και αυτή μέλος της οικογένειας κατανομών g, αλλά με παραμέτρους $\hat{\sigma}$ που είναι συνάρτηση των προηγούμενων παραμέτρων σ και της μέτρησης x.

Με τον τρόπο αυτό διευκολύνεται ο υπολογισμός της a posteriori. Όμως για να είναι χρήσιμη η οικογένεια $g(\cdot/\sigma)$ θα πρέπει να περιλαμβάνει (μεταβάλλοντας το σ) "ενδιαφέρουσες" κατανομές.