Αξιωματική θεμελίωση των πιθανοτήτων και βασικοί υπολογισμοί Περίληψη

Αντώνης Οικονόμου
aeconom@math.uoa.gr

18 Φεβρουαρίου 2010

Η Θεωρία Πιθανοτήτων έχει ως σκοπό την ποσοτική μελέτη των τυχαίων φαινομένων με μαθηματικές μεθόδους. Πιο συγκεκριμένα, ασχολείται με τη μελέτη **πειραμάτων τύχης**, δηλαδή πειραμάτων ή διαδικασιών των οποίων το αποτέλεσμα δεν είναι εκ των προτέρων γνωστό αλλά καθορίζεται (και) από τυχαίους παράγοντες.

Έστω S το σύνολο των δυνατών αποτελεσμάτων ενός τυχαίου πειράματος. Κάθε δυνατό αποτέλεσμα καλείται **δειγματικό σημείο** ή **απλό ενδεχόμενο** του τυχαίου πειράματος, ενώ το σύνολο S καλείται **δειγματικός χώρος**. Τα υποσύνολα του δειγματικού χώρου αναφέρονται ως **ενδεχόμενα**. Κατά την εκτέλεση ενός πειράματος τύχης, θα λέμε ότι **ένα δεδομένο ενδεχόμενο πραγματοποιείται**, αν το αποτέλεσμα του πειράματος ανήκει στο συγκεκριμένο ενδεχόμενο.

Αν τα $E_1, E_2, \ldots, E_n$ είναι ενδεχόμενα τότε η ένωσή τους, $\bigcup_{i=1}^n E_i$, είναι το ενδεχόμενο που περιλαμβάνει όλα τα δυνατά αποτελέσματα που ανήκουν σε τουλάχιστον ένα από τα $E_1, E_2, \ldots, E_n$. Δηλαδή, το ενδεχόμενο $\bigcup_{i=1}^n E_i$ πραγματοποιείται όταν πραγματοποιηθεί τουλάχιστον ένα από τα $E_1, E_2, \ldots, E_n$. Παρόμοια, η τομή $\bigcap_{i=1}^n E_i$ είναι το ενδεχόμενο που περιλαμβάνει όλα τα δυνατά αποτελέσματα που ανήκουν σε όλα τα $E_1, E_2, \ldots, E_n$. Το ενδεχόμενο $\bigcap_{i=1}^n E_i$ πραγματοποιείται όταν πραγματοποιηθούν "ταυτόχρονα" όλα τα $E_1, E_2, \ldots, E_n$.

Για κάθε ενδεχόμενο E ορίζουμε το συμπληρωματικό του, E^c, που είναι το σύνολο των δυνατών αποτελεσμάτων που δεν ανήκουν στο E. Το συμπληρωματικό E^c του E πραγματοποιείται τότε και μόνον τότε όταν δεν πραγματοποιείται το E. Το ενδεχόμενο S πραγματοποιείται πάντα, όποιο κι αν είναι το αποτέλεσμα ενός πειράματος τύχης, ενώ το S^c είναι το κενό σύνολο $\emptyset$ που δεν περιέχει κανένα δυνατό αποτέλεσμα και το οποίο επομένως δεν πραγματοποιείται ποτέ. Δυο ενδεχόμενα E και F με $E \cap F = \emptyset$ λέγονται **ασυμβίβαστα**. Ο όρος αυτός χρησιμοποιείται ακριβώς για να δηλώσει ότι δεν είναι δυνατό να πραγματοποιηθούν και τα δυο, αφού η πραγματοποίηση του ενός αποκλείει την πραγματοποίηση του άλλου.

Για την πιθανοθεωρητική προτυποποίηση (μοντελοποίηση) ενός πειράματος τύχης απαιτείται να καθορισθούν τρία στοιχεία: Ο δειγματικός χώρος S, η οικογένεια των ενδεχομένων A που είναι μια οικογένεια υποσυνόλων του S για τα οποία μας ενδιαφέρει να μετρήσουμε την πιθανότητα πραγματοποίησής τους και μια συνάρτηση (μέτρο) πιθανότητας $P : A \to \mathbb{R}$ που αντιστοιχεί σε κάθε ενδεχόμενο την πιθανότητα πραγματοποίησής του που θα είναι ένας αριθμός.

Για την αξιωματική θεμελίωση των πιθανοτήτων υποθέτουμε τα ακόλουθα αξιώματα τα οποία πρέπει να ικανοποιεί η πιθανότητα P:

1. $0 \le P(E) \le 1$, για κάθε ενδεχόμενο E

2. $P(S) = 1$

3. Για κάθε ακολουθία ανά δύο ασυμβίβαστων ενδεχομένων $E_1, E_2, \ldots$ (δηλαδή $E_i \cap E_j = \emptyset$ για $i \neq j$)

$$P(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} P(E_i)$$

Μπορούν τότε να αποδειχθούν ξεκινώντας από τα αξιώματα τα ακόλουθα βασικά υπολογιστικά αποτελέσματα:

1. $P(\emptyset) = 0$
2. $P(E^c) = 1 - P(E)$
3. Αρχή εγκλεισμού - αποκλεισμού για 2 ενδεχόμενα:
$$P(E \cup F) = P(E) + P(F) - P(EF)$$
4. Γενική αρχή εγκλεισμού - αποκλεισμού:
$$\begin{aligned} P(\bigcup_{i=1}^{n} E_i) &= \sum_{i=1}^{n} P(E_i) - \sum_{i<j} P(E_iE_j) + \sum_{i<j<k} P(E_iE_jE_k) - \cdots + (-1)^{n+1} P(E_1E_2\cdots E_n) \\ &= \sum_{r=1}^{n} (-1)^{r+1} \sum_{i_1<i_2<\cdots<i_r} P(E_{i_1}E_{i_2}\cdots E_{i_r}) \end{aligned}$$
5. Μονοτονία: Αν $E \subseteq F$ τότε $P(E) \leq P(F)$
6. Υποπροσθετικότητα: Για κάθε ακολουθία ενδεχομένων $E_1, E_2, \ldots$ ισχύει
$$P(\bigcup_{i=1}^{\infty} E_i) \leq \sum_{i=1}^{\infty} P(E_i)$$
7. Συνέχεια: Αν $E_1 \subseteq E_2 \subseteq E_3 \subseteq \cdots$ είναι μια αύξουσα ακολουθία ενδεχομένων τότε
$$P(\bigcup_{i=1}^{\infty} E_i) = \lim_{n\to\infty} P(E_n)$$
Ομοίως, αν $E_1 \supseteq E_2 \supseteq E_3 \supseteq \cdots$ είναι μια φθίνουσα ακολουθία ενδεχομένων τότε
$$P(\bigcap_{i=1}^{\infty} E_i) = \lim_{n\to\infty} P(E_n)$$

Στην περίπτωση πεπερασμένου δειγματικού χώρου S που όλα τα δειγματικά σημεία του είναι ισοπίθανα, η πιθανότητα ενός ενδεχομένου E δίνεται από τη σχέση

$$P(E) = \frac{|E|}{|S|} = \frac{\text{ευνοϊκές περιπτώσεις}}{\text{δυνατές περιπτώσεις}}$$

όπου με $|E|$ συμβολίζουμε το πλήθος των στοιχείων ενός ενδεχομένου E.

Η πιθανότητα ενός ενδεχομένου μπορεί να έχει διάφορες ερμηνείες. Σε ένα πείραμα που αφορά το χαρακτηριστικό ενός πεπερασμένου πληθυσμού μπορεί να ερμηνευθεί σαν το ποσοστό του πληθυσμού που έχει το συγκεκριμένο χαρακτηριστικό. Σε ένα πείραμα που μπορεί να επαναληφθεί απεριόριστα, μπορεί να ερμηνευθεί ως η οριακή σχετική συχνότητα πραγματοποίησης του ενδεχομένου, καθώς το πλήθος των επαναλήψεων του πειράματος τείνει στο άπειρο. Σε πειράματα τύχης που ο δειγματικός χώρος μπορεί να απεικονιστεί με κάποιο επίπεδο γεωμετρικό σχήμα, η πιθανότητα ενός ενδεχομένου μπορεί να γίνει αντιληπτή ως το ποσοστό του εμβαδού του σχήματος που αντιστοιχεί στο συγκεκριμένο ενδεχόμενο ως προς το εμβαδό του σχήματος που αντιστοιχεί στο δειγματικό χώρο. Τέλος, η πιθανότητα μπορεί να ερμηνευθεί ως ένα υποκειμενικό μέτρο βεβαιότητας για το αποτέλεσμα ενός πειράματος τύχης.

Στοιχεία συνδυαστικής για υπολογισμούς πιθανοτήτων
Περίληψη

Αντώνης Οικονόμου
aeconom@math.uoa.gr

20 Φεβρουαρίου 2010

Στην περίπτωση που ένα πείραμα τύχης έχει πεπερασμένο δειγματικό χώρο S που όλα τα δειγματικά σημεία του είναι ισοπίθανα, η πιθανότητα ενός ενδεχομένου E δίνεται από τη σχέση

$$P(E) = \frac{|E|}{|S|} = \frac{\text{ευνοϊκές περιπτώσεις}}{\text{δυνατές περιπτώσεις}}$$

όπου με $|E|$ συμβολίζουμε το πλήθος των στοιχείων ενός ενδεχομένου E.

Επομένως, γίνεται φανερό ότι για τον υπολογισμό πιθανοτήτων σε τέτοια πειράματα τύχης είναι αναγκαία η καταμέτρηση του πλήθους των στοιχείων του δειγματικού χώρου καθώς και διαφόρων υποσυνόλων του, που αντιστοιχούν στα ενδεχόμενα που μας ενδιαφέρουν. Η απαρίθμηση του πλήθους των στοιχείων πεπερασμένων συνόλων με συγκεκριμένες ιδιότητες είναι το αντικείμενο της **Συνδυαστικής**. Για το λόγο αυτό είναι αναγκαία η γνώση τουλάχιστον των βασικών εργαλείων της Συνδυαστικής.

Η πιο σημαντική και ευρέως χρησιμοποιούμενη αρχή της Συνδυαστικής είναι η **πολλαπλασιαστική αρχή**. Σύμφωνα με την αρχή αυτή, αν ένα πείραμα μπορεί να αναλυθεί σε στάδια και στο 1ο στάδιο υπάρχουν n_1 δυνατά αποτελέσματα, στο 2ο στάδιο υπάρχουν n_2 δυνατά αποτελέσματα (ό,τι και να έχει συμβεί στο προηγούμενο στάδιο), στο 3ο στάδιο υπάρχουν n_3 δυνατά αποτελέσματα (ό,τι και να έχει συμβεί στα προηγούμενα στάδια) κ.ο.κ. μέχρι και το τελευταίο r-οστό στάδιο στο οποίο υπάρχουν n_r δυνατά αποτελέσματα, τότε το πείραμα έχει συνολικά $n_1 n_2 \cdots n_r$ διαφορετικά δυνατά αποτελέσματα.

Μετάθεση n αντικειμένων είναι μια τοποθέτησή τους σε σειρά, δηλαδή μια οποιαδήποτε διατεταγμένη n-αδα από αυτά τα αντικείμενα στην οποία το καθένα εμφανίζεται μια και μόνο μια φορά (χωρίς επαναλήψεις). Το πλήθος των διαφορετικών μεταθέσεων n αντικειμένων είναι

$$n! = n(n-1)(n-2)\cdots 1.$$

Το σύμβολο $0!$ ορίζεται να είναι ίσο με 1.

Διάταξη n **ανά** k από ένα σύνολο n αντικειμένων λέγεται κάθε διατεταγμένη k-αδα από αυτά τα n αντικείμενα στην οποία το καθένα εμφανίζεται το πολύ μια φορά (χωρίς επαναλήψεις). Το πλήθος των διατάξεων n ανά k είναι

$$\frac{n!}{(n-k)!} = n(n-1)(n-2)\cdots(n-k+1).$$

Η μετάθεση n αντικειμένων είναι μια διάταξη n ανά n.

Επαναληπτική διάταξη ή διάταξη με επανάληψη n **ανά** k από ένα σύνολο n αντικειμένων λέγεται κάθε διατεταγμένη k-αδα από αυτά τα n αντικείμενα στην οποία το καθένα μπορεί να εμφανίζεται οσεσδήποτε φορές (με επαναλήψεις). Το πλήθος των επαναληπτικών διατάξεων n ανά k είναι

$$n^k.$$

Συνδυασμός n **ανά** k από ένα σύνολο n αντικειμένων λέγεται κάθε μη-διατεταγμένη συλλογή με k στοιχεία από αυτά τα n αντικείμενα, στην οποία το καθένα εμφανίζεται το πολύ μια φορά (χωρίς επαναλήψεις). Ουσιαστικά δηλαδή ο όρος συνδυασμός n ανά k σημαίνει υποσύνολο με k στοιχεία του συνόλου των n αντικειμένων. Το πλήθος των συνδυασμών n ανά k είναι

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Η ποσότητα αυτή αναφέρεται και ως **διωνυμικός συντελεστής** διότι εμφανίζεται στο **διωνυμικό θεώρημα** που λέει πώς αναπτύσσεται η δύναμη ενός διωνύμου σε μονώνυμα:

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}.$$

Εκτός από τον ακριβή τύπο που δίνει το πλήθος των συνδυασμών n ανά k, ένας χρήσιμος τύπος είναι και ο

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k},\ 1 \le k \le n,$$

που αποτελεί μια αναγωγική σχέση που μπορεί να χρησιμοποιηθεί για την εύκολη κατασκευή πίνακα με τις τιμές των $\binom{n}{k}$ (τρίγωνο Pascal).

Επαναληπτικός συνδυασμός ή συνδυασμός με επανάληψη n **ανά** k από ένα σύνολο n αντικειμένων λέγεται κάθε μη-διατεταγμένη συλλογή με k στοιχεία από αυτά τα n αντικείμενα, στην οποία το καθένα μπορεί να εμφανίζεται οσεσδήποτε φορές (με επαναλήψεις). Το πλήθος των επαναληπτικών συνδυασμών n ανά k είναι

$$\binom{n+k-1}{k} = \binom{n+k-1}{n-1} = \frac{(n+k-1)!}{k!(n-1)!}.$$

Ο **πολυωνυμικός συντελεστής** n **ανά** $n_1, n_2, \ldots, n_r$ ορίζεται μόνο για $n = n_1 + n_2 + \cdots + n_r$ ως

$$\frac{n!}{n_1! n_2! \cdots n_r!} = \frac{(n_1+n_2+\cdots+n_r)!}{n_1! n_2! \cdots n_r!}$$

και αποτελεί γενίκευση του διωνυμικού συντελεστή (αφού ο διωνυμικός συντελεστής n ανά k μπορεί να θεωρηθεί πολυωνυμικός συντελεστής n ανά $k, n-k$. Ονομάζεται έτσι διότι εμφανίζεται στο **πολυωνυμικό θεώρημα** που γενικεύει το διωνυμικό ως εξής:

$$(x_1 + x_2 + \cdots + x_r)^n = \sum_{(n_1,n_2,\ldots,n_r):\ n_1+n_2+\cdots+n_r=n} \frac{n!}{n_1! n_2! \cdots n_r!} x_1^{n_1} x_2^{n_2} \cdots x_r^{n_r}.$$

Ο πολυωνυμικός συντελεστής n ανά $n_1, n_2, \ldots, n_r$ δίνει την απάντηση σε δυο βασικά συνδυαστικά προβλήματα. Πράγματι, δίνει το πλήθος των διαφορετικών τοποθετήσεων σε σειρά n αντικειμένων, εκ των οποίων n_1 αντικείμενα είναι όμοια μεταξύ τους τύπου 1, n_2 αντικείμενα είναι όμοια μεταξύ τους τύπου 2 κ.ο.κ., n_r αντικείμενα είναι όμοια μεταξύ τους τύπου r (**μεταθέσεις** r **ειδών στοιχείων** n **αντικειμένων ανά** $n_1, n_2, \ldots, n_r$). Επίσης, δίνει το πλήθος των διαφορετικών **διαιρέσεων ενός συνόλου με** n **στοιχεία σε** r **διακεκριμένα υποσύνολα με** $n_1, n_2, \ldots, n_r$ **στοιχεία το καθένα**.

Ένα τελευταίο συνδυαστικό ερώτημα που εμφανίζεται συχνά στις εφαρμογές αφορά το πλήθος των ακέραιων μη αρνητικών λύσεων $(x_1, x_2, \ldots, x_n)$ της εξίσωσης

$$x_1 + x_2 + \ldots + x_n = k.$$

Αποδεικνύεται ότι το πλήθος αυτό ισούται με το πλήθος των επαναληπτικών συνδυασμών n ανά k, δηλαδή είναι $\binom{n+k-1}{k}$. Αν ενδιαφερόμαστε για το πλήθος των λύσεων της ίδιας εξίσωσης αλλά με τα $x_1, x_2, \ldots, x_n$ να παίρνουν μόνο τις τιμές 0 ή 1 τότε έχουμε $\binom{n}{k}$ το πλήθος λύσεις.

Δεσμευμένη πιθανότητα και ανεξαρτησία
Περίληψη

Αντώνης Οικονόμου
aeconom@math.uoa.gr

27 Φεβρουαρίου 2010

Έστω δυο ενδεχόμενα E και F ενός πειράματος τύχης με δειγματικό χώρο S. Αν $P(F) > 0$, δηλαδή το F έχει μη-μηδενική πιθανότητα να πραγματοποιηθεί, τότε ορίζεται **η δεσμευμένη πιθανότητα του ενδεχομένου E δοθέντος του F** ως:

$$P(E|F) = \frac{P(EF)}{P(F)}.$$

Η δεσμευμένη πιθανότητα $P(E|F)$ πρέπει να ερμηνεύεται διαισθητικά ως η πιθανότητα να (έχει) πραγματοποιηθεί το E, δεδομένης της πληροφορίας ότι το F έχει πραγματοποιηθεί. Τα τρία βασικά αποτελέσματα που χρησιμοποιούνται στους υπολογισμούς με δεσμευμένες πιθανότητες είναι το πολλαπλασιαστικό θεώρημα, το θεώρημα ολικής πιθανότητας και το θεώρημα Bayes.

Το **πολλαπλασιαστικό θεώρημα** λέει ότι οποτεδήποτε έχουμε μια πεπερασμένη οικογένεια ενδεχομένων $E_1, E_2, E_3 \ldots, E_n$ τότε η πιθανότητα της τομής τους μπορεί να υπολογιστεί απο τον τύπο

$$P(E_1E_2E_3 \cdots E_n) = P(E_1)P(E_2|E_1)P(E_3|E_1E_2) \cdots P(E_n|E_1E_2 \cdots E_{n-1}).$$

Το **θεώρημα ολικής πιθανότητας** λέει ότι άν έχουμε μια διαμέριση $F_1, F_2, F_3, \ldots$ του δειγματικού χώρου S (δηλαδή $\cup_{i=1}^{\infty} F_i = S$ και τα ενδεχόμενα F_i είναι ανά δύο ασυμβίβαστα) τότε ο υπολογισμός της πιθανότητας οποιουδήποτε ενδεχομένου E μπορεί να γίνει μέσω του τύπου

$$P(E) = \sum_{i=1}^{\infty} P(F_i)P(E|F_i).$$

Λέμε τότε ότι υπολογίζουμε την πιθανότητα πραγματοποίησης του ενδεχομένου E "δεσμεύοντας" στα F_i. Στην απλούστερη μορφή του, το θεώρημα ολικής πιθανότητας λέει ότι δοθέντων δυο ενδεχομένων E και F, η πιθανότητα πραγματοποίησης του E μπορεί να βρεθεί δεσμεύοντας στο αν το F πραγματοποιήθηκε ή όχι. Έχουμε τότε τον τύπο

$$P(E) = P(F)P(E|F) + P(F^c)P(E|F^c).$$

Το **θεώρημα του Bayes** λέει ότι αν δίνεται ένα ενδεχόμενο E και μια διαμέριση $F_1, F_2, F_3, \ldots$ του δειγματικού χώρου S, τότε ο υπολογισμός των δεσμευμένων πιθανοτήτων $P(F_j|E)$, $j = 1, 2, \ldots$ μπορεί να γίνει μέσω του τύπου

$$P(F_j|E) = \frac{P(F_j)P(E|F_j)}{P(E)} = \frac{P(F_j)P(E|F_j)}{\sum_{i=1}^{\infty} P(F_i)P(E|F_i)}.$$

Η αξία του θεωρήματος Bayes βρίσκεται στο ότι επιτρέπει υπολογισμούς πιθανοτήτων για πειράματα τύχης που διεξάγονται σε στάδια. Επομένως μπορεί να μας ενδιαφέρει η πιθανότητα πραγματοποίησης ενός ενδεχομένου F_i που αναφέρεται στο πρώτο στάδιο ενός πειράματος, δεδομένου ότι ξέρουμε τι έγινε σε κάποιο επόμενο στάδιο ή στο τέλος, πράγμα που αντιστοιχεί στην πραγματοποίηση ενός ενδεχομένου E.

Ένας άλλος τρόπος να σκεφτόμαστε το θεώρημα Bayes είναι ο εξής: Θεωρούμε ότι τα ενδεχόμενα F_i αντιστοιχούν σε αμοιβαία αποκλειόμενες υποθέσεις ως προς το αποτέλεσμα ενός πειράματος τύχης. Το πείραμα πραγματοποιείται και μας δίνεται η πληροφορία ότι κάποιο ενδεχόμενο E πραγματοποιήθηκε. Αυτή η επιπλέον πληροφορία μπορεί να μην μας προκαθορίζει ακριβώς ποιά υπόθεση ισχύει (δηλαδή ποιο F_i έχει πραγματοποιηθεί), αλλά μεταβάλει τις αρχικές πιθανότητες των υποθέσεων σχετικά με την έκβαση του πειράματος. Το θεώρημα Bayes δείχνει πως υπολογίζονται οι "εκ των υστέρων πιθανότητες" $P(F_j|E)$ των υποθέσεων F_j δεδομένης της πληροφορίας ότι το E πραγματοποιήθηκε συναρτήσει των "αρχικών" ή "εκ των προτέρων πιθανοτήτων" $P(F_j)$ και των πιθανοτήτων $P(E|F_j)$.

Ο **λόγος πιθανοφανειών (the odds)** ενός ενδεχομένου H ορίζεται ως ο λόγος $\frac{P(H)}{P(H^c)}$. Ο λόγος αυτός εκφράζει πόσο πιθανότερο είναι να πραγματοποιηθεί ένα ενδεχόμενο από το να πραγματοποιηθεί το συμπληρωματικό του. Η έννοια αυτή χρησιμοποιείται συχνά στη γλώσσα των στοιχημάτων όπου λέμε ότι στοιχηματίζω 1 προς τόσο ότι θα συμβεί το τάδε. Ουσιαστικά στη γλώσσα των στοιχημάτων μιλάμε με λόγους πιθανοφανειών, εκτιμώντας πόσο πιθανότερο είναι ένα ενδεχόμενο από το συμπληρωματικό του. Ισχύει η ταυτότητα

$$\frac{P(H|E)}{P(H^c|E)} = \frac{P(H)}{P(H^c)}\frac{P(E|H)}{P(E|H^c)},$$

που δείχνει πως "ενημερώνεται" ο λόγος πιθανοφανειών ενός ενδεχομένου H όταν προκύψουν νέα στοιχεία, δηλαδή δοθεί η πληροφορία ότι έχει πραγματοποιηθεί κάποιο ενδεχόμενο E.

Δυο ενδεχόμενα E και F λέγονται **ανεξάρτητα** αν

$$P(EF) = P(E)P(F).$$

Αυτή η συνθήκη είναι ισοδύναμη με καθεμιά από τις συνθήκες $P(E|F) = P(E)$ και $P(F|E) = P(F)$. Επομένως η διαισθητική ερμηνεία της ανεξαρτησίας δυο ενδεχομένων E και F είναι ότι η γνώση της πραγματοποίησης του ενός από αυτά δεν επηρεάζει την πιθανότητα του άλλου. Γενικότερα, n ενδεχόμενα $E_1, E_2, \dots, E_n$ λέγονται ανεξάρτητα αν για οποιαδήποτε επιλογή δεικτών $\{i_1, i_2, \dots, i_r\}$ είναι

$$P(E_{i_1}E_{i_2}\cdots E_{i_r}) = P(E_{i_1})P(E_{i_2})\cdots P(E_{i_r}).$$

Προσοχή! Αν τα μέλη μιας οικογένειας ενδεχομένων $E_1, E_2, \dots, E_n$ είναι ανά δύο ανεξάρτητα, δεν έπεται ότι τα $E_1, E_2, \dots, E_n$ είναι ανεξάρτητα. Πρέπει να ελεγχθεί ότι για οποιαδήποτε επιλογή από αυτά η πιθανότητα της τομής τους ισούται με το γινόμενο των πιθανοτήτων τους.

Δοθέντος ενός ενδεχομένου F η δεσμευμένη πιθανότητα $P(\cdot|F)$ έχει τις τρεις ιδιότητες που υποθέσαμε ως αξιώματα για την αρχική πιθανότητα P:

1. $0 \le P(E|F) \le 1$, για κάθε ενδεχόμενο E
2. $P(S|F) = 1$
3. Για κάθε ακολουθία ανά δύο ασυμβίβαστων ενδεχομένων $E_1, E_2, \dots$ (δηλαδή $E_i \cap E_j = \emptyset$ για $i \neq j$)

$$P(\bigcup_{i=1}^{\infty} E_i|F) = \sum_{i=1}^{\infty} P(E_i|F)$$

Επομένως, όλα τα υπολογιστικά αποτελέσματα που ισχύουν για την αρχική πιθανότητα P ισχύουν και για οποιαδήποτε δεσμευμένη πιθανότητα $P(\cdot|F)$, με δέσμευση σε κάποιο ενδεχόμενο F.

Διακριτές τυχαίες μεταβλητές
Περίληψη

Αντώνης Οικονόμου
aeconom@math.uoa.gr

30 Μαρτίου 2010

Μια **τυχαία μεταβλητή** που αναφέρεται σε ένα πείραμα τύχης είναι διαισθητικά ένα αριθμητικό χαρακτηριστικό του πειράματος. Σε μαθηματικό - τυπικό επίπεδο είναι μια συνάρτηση με πεδίο ορισμού το δειγματικό χώρο και τιμές στο σύνολο των πραγματικών αριθμών. Αν X είναι μια τυχαία μεταβλητή, η συνάρτηση κατανομής της, $F(x)$, ορίζεται ως:

$$F(x) = P(X \leq x).$$

Όλοι οι υπολογισμοί πιθανοτήτων που αφορούν τη X μπορούν να γίνουν μέσω της συνάρτησης κατανομής.

Μια τυχαία μεταβλητή με πεπερασμένο ή το πολύ αριθμήσιμο πεδίο τιμών λέγεται **διακριτή** και αντιστοιχεί διαισθητικά σε κάποιο διακριτό αριθμητικό χαρακτηριστικό ενός πειράματος τύχης. Αν X είναι μια διακριτή τυχαία μεταβλητή, τότε η συνάρτηση

$$p(x) = P(X = x)$$

λέγεται **συνάρτηση πιθανότητας** της X. Η ποσότητα

$$E[X] = \sum_x x p(x)$$

λέγεται **μέση τιμή** της X. Οι όροι αναμενόμενη τιμή καθώς και μαθηματική ελπίδα της X χρησιμοποιούνται εναλλακτικά για την $E[X]$. Η $E[X]$ είναι ένα μέτρο θέσης της X, δηλαδή περιγράφει γύρω από ποιον αριθμό περιστρέφονται οι τιμές της X. Για να έχουμε μια πιο ικανοποιητική περιγραφή της συμπεριφοράς της X χρειαζόμαστε και κάποιο μέτρο συγκεντρωτικότητας ή μεταβλητότητας. Ως τέτοιο χρησιμοποιείται η ποσότητα

$$Var[X] = E[(X - E[X])^2]$$

που αναφέρεται ως **διασπορά** της X. Η $Var[X]$ περιγράφει πόσο διεσπαρμένες είναι οι τιμές της X γύρω από τη μέση τιμή της. Η ποσότητα $\sqrt{Var[X]}$ αναφέρεται ως τυπική απόκλιση της X.

Η μέση τιμή $E[g(X)]$ μιας συνάρτησης τυχαίας μεταβλητής μπορεί να βρεθεί άμεσα από τον τύπο

$$E[g(X)] = \sum_x g(x)p(x).$$

Χρήσιμοι τύποι για τον υπολογισμό της μέσης τιμής και της διασποράς μιας τυχαίας μεταβλητής είναι οι εξής:

$$\begin{aligned} E[aX+b] &= aE[X]+b, \\ Var[aX+b] &= a^2 Var[X], \\ Var[X] &= E[X^2] - E[X]^2. \end{aligned}$$

Ένα σημαντικό πείραμα τύχης είναι η τυχαία επιλογή αριθμού από το σύνολο $\{1, 2, \ldots, N\}$. Αν X είναι το αποτέλεσμα του πειράματος, τότε λέμε ότι η X ακολουθεί την **διακριτή ομοιόμορφη κατανομή** στο $\{1, 2, \ldots, N\}$ ($\mathrm{Uniform}(\{1, 2, \ldots, N\})$). Η συνάρτηση πιθανότητας της X δίνεται τότε ως

$$p(i) = \frac{1}{N}, \quad i = 1, 2, \ldots, N.$$

Η μέση τιμή και η διασπορά της είναι

$$E[X] = \frac{N+1}{2}, \quad Var[X] = \frac{N^2 - 1}{12}.$$

Ένα άλλο σημαντικό πείραμα τύχης είναι οι επαναλαμβανόμενες πραγματοποιήσεις ενός πειράματος με δυο πιθανές εκβάσεις, επιτυχία και αποτυχία, με πιθανότητα επιτυχίας p και πιθανότητα αποτυχίας $1-p$. Αυτές αναφέρονται και ως ανεξάρτητες δοκιμές Bernoulli με πιθανότητα επιτυχίας p. Αν X είναι ο αριθμός των επιτυχιών σε n ανεξάρτητες δοκιμές Bernoulli, τότε λέμε ότι η X ακολουθεί τη **διωνυμική κατανομή** με παραμέτρους n και p ($\mathrm{Bin}(n, p)$). Η συνάρτηση πιθανότητας της X δίνεται τότε ως

$$p(i) = \binom{n}{i} p^i (1-p)^{n-i}, \quad i = 0, 1, 2, \ldots, n.$$

Η μέση τιμή και η διασπορά της είναι

$$E[X] = np, \quad Var[X] = np(1-p).$$

Αν X είναι ο αριθμός των δοκιμών μέχρι την πρώτη επιτυχία σε μια σειρά ανεξάρτητων δοκιμών Bernoulli, τότε λέμε ότι η X ακολουθεί τη **γεωμετρική κατανομή** με παράμετρο p στο σύνολο $\{1, 2, \ldots\}$ ($\mathrm{Geom}(p)$ στο $\{1, 2, \ldots\}$). Η συνάρτηση πιθανότητας της X δίνεται τότε ως

$$p(i) = p(1-p)^{i-1}, \quad i = 1, 2, \ldots.$$

Η μέση τιμή και η διασπορά της είναι

$$E[X] = \frac{1}{p}, \quad Var[X] = \frac{1-p}{p^2}.$$

Αν X είναι ο αριθμός των δοκιμών μέχρι την r-οστή επιτυχία σε μια σειρά ανεξάρτητων δοκιμών Bernoulli, τότε λέμε ότι η X ακολουθεί την **αρνητική διωνυμική κατανομή** με παραμέτρους r και p στο σύνολο $\{r, r+1, \ldots\}$ ($\mathrm{NegBin}(r, p)$ στο $\{r, r+1, \ldots\}$). Η συνάρτηση πιθανότητας της X δίνεται τότε ως

$$p(i) = \binom{i-1}{r-1} p^r (1-p)^{i-r}, \quad i = r, r+1, \ldots.$$

Η μέση τιμή και η διασπορά της είναι

$$E[X] = \frac{r}{p}, \quad Var[X] = \frac{r(1-p)}{p^2}.$$

Στην περίπτωση που έχουμε έναν μεγάλο αριθμό n ανεξάρτητων δοκιμών Bernoulli και η πιθανότητα επιτυχίας p σε κάθε δοκιμή είναι μικρή, έτσι ώστε $np = \lambda$, τότε ο αριθμός των επιτυχιών προσεγγίζεται από τη λεγόμενη **κατανομή Poisson** με παράμετρο λ ($\mathrm{Poisson}(\lambda)$). Η συνάρτηση πιθανότητας της X δίνεται τότε ως

$$p(i) = e^{-\lambda} \frac{\lambda^i}{i!}, \quad i = 0, 1, \ldots.$$

Η μέση τιμή και η διασπορά της είναι

$$E[X] = \lambda, \quad Var[X] = \lambda.$$

Η διωνυμική κατανομή έχει και την εξής εναλλακτική ερμηνεία: Υποθέτουμε ότι σε έναν (άπειρο) πληθυσμό, ένα ποσοστό p των ατόμων του πληθυσμού έχει κάποιο χαρακτηριστικό (επιτυχία) ενώ το ποσοστό $1-p$ των ατόμων του πληθυσμού δεν έχει το χαρακτηριστικό (αποτυχία). Αν επιλέξουμε n άτομα του πληθυσμού και θέσουμε X την τυχαία μεταβλητή που παριστάνει το πλήθος των ατόμων που έχουν το χαρακτηριστικό, τότε η X ακολουθεί τη διωνυμική κατανομή με παραμέτρους n και p. Στην περίπτωση πεπερασμένου πληθυσμού N ατόμων από τα οποία τα m έχουν το χαρακτηριστικό ενώ τα $N-m$ δεν έχουν το χαρακτηριστικό, αν επιλέξουμε n άτομα του πληθυσμού και θέσουμε X το πλήθος των ατόμων που έχουν το χαρακτηριστικό, τότε η κατανομή της X εξαρτάται από το κατά πόσον η δειγματοληψία γίνεται με ή χωρίς επανάθεση.

Αν η δειγματοληψία γίνεται με επανάθεση τότε η κατάσταση δεν διαφέρει από τον άπειρο πληθυσμό και έχουμε ότι η X ακολουθεί τη διωνυμική κατανομή με παραμέτρους n και $\frac{m}{N}$. Αν, όμως, η δειγματοληψία γίνεται χωρίς επανάθεση τότε η X έχει συνάρτηση πιθανότητας

$$p(i) = \frac{\binom{m}{i}\binom{N-m}{n-i}}{\binom{N}{n}}, \quad i = 0, 1, \ldots, n.$$

Αν θέσουμε $p = \frac{m}{N}$ τότε η μέση τιμή και η διασπορά της είναι

$$E[X] = np, \quad Var[X] = \frac{N-n}{N-1}np(1-p).$$

Η X λέμε τότε ότι ακολουθεί την **υπεγεωμετρική κατανομή** με παραμέτρους n, N, m $(\mathrm{Hypergeom}(n, N, m))$. Για μεγάλα N και σχετικά μικρά n η υπεργεωμετρική κατανομή $\mathrm{Hypergeom}(n, N, m)$ προσεγγίζεται ικανοποιητικά από την διωνυμική κατανομή $\mathrm{Bin}(n, p)$ με $p = \frac{m}{N}$. Αυτό είναι φανερό διότι εάν έχουμε έναν μεγάλο πληθυσμό (δηλαδή μεγάλο N) και πρόκειται να πάρουμε μικρό δείγμα (δηλαδή μικρό n) η κατανομή του πλήθους των ατόμων που έχουν το χαρακτηριστικό αναμένουμε να είναι περίπου η ίδια, είτε έχουμε δειγματοληψία με επανάθεση, είτε δειγματοληψία χωρίς επανάθεση.

Συνεχείς τυχαίες μεταβλητές
Περίληψη

Αντώνης Οικονόμου
aeconom@math.uoa.gr

24 Απριλίου 2010

Μια τυχαία μεταβλητή X λέγεται (απόλυτα) **συνεχής** αν αντιστοιχεί διαισθητικά σε κάποιο συνεχές αριθμητικό χαρακτηριστικό ενός πειράματος τύχης. Σε μαθηματικό επίπεδο η X είναι συνεχής αν υπάρχει μια μη-αρνητική συνάρτηση f, που λέγεται **συνάρτηση πυκνότητας πιθανότητας**, τέτοια ώστε για κάθε σύνολο B να ισχύει

$$P(X \in B) = \int_B f(x)dx.$$

Ειδικότερα για τη συνάρτηση κατανομής της X έχουμε ότι δίνεται τότε από τη σχέση

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)du.$$

Αντίστροφα, αν δίνεται η συνάρτηση κατανομής $F(x)$ μιας συνεχούς τυχαίας μεταβλητής X, τότε μπορούμε να βρούμε τη συνάρτηση πυκνότητας πιθανότητας $f(x)$ από τη σχέση

$$f(x) = \frac{d}{dx}F(x).$$

Η μέση τιμή της X ορίζεται τότε από τη σχέση

$$E[X] = \int_{-\infty}^{\infty} xf(x)dx.$$

ενώ η διασπορά ορίζεται, όπως και για τις διακριτές τυχαίες μεταβλητές, από τη σχέση

$$Var[X] = E[(X - E[X])^2].$$

Η μέση τιμή $E[g(X)]$ μιας συνάρτησης τυχαίας μεταβλητής μπορεί να βρεθεί άμεσα από τον τύπο

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)dx.$$

Οι χρήσιμοι τύποι για τον υπολογισμό της μέσης τιμής και της διασποράς μιας τυχαίας μεταβλητής που ισχύουν για διακριτές τυχαίες μεταβλητές συνεχίζουν να ισχύουν. Συγκεκριμένα, αν X είναι μια τυχαία μεταβλητή, τότε ισχύουν οι σχέσεις:

$$\begin{aligned} E[aX+b] &= aE[X]+b, \\ Var[aX+b] &= a^2Var[X], \\ Var[X] &= E[X^2]-E[X]^2. \end{aligned}$$

Για μη-αρνητικές τυχαίες μεταβλητές είναι χρήσιμος ένας εναλλακτικός τύπος υπολογισμού της μέσης τιμής, μέσω της συνάρτησης κατανομής. Συγκεκριμένα έχουμε ότι

$$X \geq 0 \quad \Rightarrow \quad E[X] = \int_0^\infty (1 - F(x))dx.$$

Ένα σημαντικό πείραμα τύχης είναι η τυχαία επιλογή σημείου από ένα διάστημα $[a, b]$. Αν X είναι το αποτέλεσμα του πειράματος, τότε λέμε ότι η X ακολουθεί τη **συνεχή ομοιόμορφη κατανομή** στο $[a, b]$ (Uniform($[a, b]$)). Η συνάρτηση πυκνότητας πιθανότητας της X δίνεται τότε ως

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{αν } a \leq x \leq b \\ 0 & \text{διαφορετικά.} \end{cases}$$

Η μέση τιμή και η διασπορά της είναι

$$E[X] = \frac{a+b}{2}, \quad Var[X] = \frac{(b-a)^2}{12}.$$

Μια άλλη σημαντική συνεχής κατανομή παρουσιάζεται στη μελέτη πολλών χαρακτηριστικών πληθυσμών. Πρόκειται για την **κανονική κατανομή** με παραμέτρους μ και σ^2 ($N(\mu, \sigma^2)$) που έχει συνάρτηση πυκνότητας πιθανότητας

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty.$$

Η μέση τιμή και η διασπορά της είναι

$$E[X] = \mu, \quad Var[X] = \sigma^2.$$

Μια σημαντική ιδιότητα της κανονικής κατανομής είναι ότι κάθε γραμμική συνάρτησή της ακολουθεί επίσης κανονική κατανομή. Συγκεκριμένα έχουμε ότι

$$X \text{ ακολουθεί την } N(\mu, \sigma^2), \quad a \in \mathbb{R} \setminus \{0\}, \quad b \in \mathbb{R} \Rightarrow aX + b \text{ ακολουθεί την } N(a\mu + b, a^2\sigma^2).$$

Από την ιδιότητα αυτή έπεται ότι δοθείσης μιας κανονικής τυχαίας μεταβλητής με μέση τιμή μ και διασπορά σ^2, η τυχαία μεταβλητή

$$Z = \frac{X - \mu}{\sigma}$$

είναι κανονική με μέση τιμή 0 και διασπορά 1. Μια τέτοια τυχαία μεταβλητή αναφέρεται ως **τυποποιημένη κανονική**. Όλοι οι υπολογισμοί που αφορούν τη X μπορούν να γίνουν χρησιμοποιώντας τη Z. Επειδή η συνάρτηση κατανομής της κανονικής κατανομής δεν δίνεται σε κλειστή μορφή αλλά μέσω ολοκληρώματος, χρησιμοποιούμε τους πίνακες τιμών που υπάρχουν για τη **συνάρτηση κατανομής** $\Phi(z)$ **της τυποποιημένης κανονικής κατανομής**. Έτσι έχουμε ότι

$$P(a \leq X \leq b) = P\left(\frac{a-\mu}{\sigma} \leq \frac{X-\mu}{\sigma} \leq \frac{b-\mu}{\sigma}\right) = P\left(\frac{a-\mu}{\sigma} \leq Z \leq \frac{b-\mu}{\sigma}\right) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right).$$

Η σημασία της κανονικής κατανομής έγγειται στο ότι προσεγγίζει ικανοποιητικά την κατανομή αθροισμάτων ανεξάρτητων και ισόνομων τυχαίων μεταβλητών. Το αποτέλεσμα αυτό που αναφέρεται ως **κεντρικό οριακό θεώρημα** αποτελεί ένα από τα σημαντικότερα θεωρήματα της θεωρίας πιθανοτήτων. Στην πιο απλή μορφή του λέει ότι για αρκετά μεγάλα n, η διωνυμική κατανομή Bin(n, p), δηλαδή το πλήθος των επιτυχιών σε n ανεξάρτητες δοκιμές Bernoulli με πιθανότητα επιτυχίας p σε κάθε δοκιμή, μπορεί να προσεγγιστεί από την κανονική κατανομή με ίδιες μέση τιμή και διασπορά, δηλαδή από την κανονική κατανομή $N(np, np(1-p))$. Πιο συγκεκριμένα έχουμε **το θεώρημα των DeMoivre-Laplace**, σύμφωνα με το οποίο, αν S_n είναι ο αριθμός των επιτυχιών σε n ανεξάρτητες δοκιμές Bernoulli, με πιθανότητα επιτυχίας p ανά δοκιμή, τότε για κάθε πραγματικούς αριθμούς $a < b$ έχουμε

$$\lim_{n\to\infty} P\left(a \leq \frac{S_n - np}{\sqrt{np(1-p)}} \leq b\right) = \Phi(b) - \Phi(a).$$

Μια άλλη σημαντική συνεχής κατανομή χρησιμοποιείται για τη μοντελοποίηση χρόνων ζωής εξαρτημάτων και χρόνων γενικότερα. Η **εκθετική κατανομή** με παράμετρο λ (Exp(λ)) έχει συνάρτηση πυκνότητας πιθανότητας

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{αν } x \geq 0 \\ 0 & \text{διαφορετικά.} \end{cases}$$

Η μέση τιμή και η διασπορά της είναι

$$E[X] = \frac{1}{\lambda}, \quad Var[X] = \frac{1}{\lambda^2}.$$

Μια βασική ιδιότητα της εκθετικής κατανομής είναι η λεγόμενη **αμνήμονη ιδιότητα** η οποία λέει ότι για οποιαδήποτε $s, t > 0$, ισχύει

$$P(X > s + t | X > t) = P(X > s).$$

Διαισθητικά, αν X παριστάνει το χρόνο ζωής ενός εξαρτήματος, η ιδιότητα αυτή λέει ότι η δεσμευμένη πιθανότητα ο υπολειπόμενος χρόνος ζωής του μηχανήματος να είναι μεγαλύτερος από s, δεδομένου ότι έχει ήδη ζήσει t χρονικές μονάδες είναι η ίδια με την πιθανότητα ο χρόνος ζωής ενός καινούργιου μηχανήματος να είναι μεγαλύτερος από s. Δηλαδή η ηλικία δεν παίζει ρόλο για το πόσο θα ζήσει ακόμα το εξάρτημα. Η εκθετική κατανομή είναι η μοναδική κατανομή που έχει την αμνήμονη ιδιότητα που περιγράψαμε παραπάνω (για κάθε $s, t > 0$).

Μια άλλη κατανομή που χρησιμοποιείται για τη μοντελοποίηση χρόνων ζωής είναι η **κατανομή Γάμμα** με παραμέτρους α και λ (Gamma(α, λ)), η οποία έχει συνάρτηση πυκνότητας πιθανότητας

$$f(x) = \begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} & \text{αν } x \geq 0. \\ 0 & \text{διαφορετικά.} \end{cases}$$

Η ποσότητα $\Gamma(\alpha)$ είναι η γνωστή συνάρτηση Γάμμα υπολογισμένη στο σημείο α, η οποία δίνεται από το ολοκλήρωμα

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x}.$$

Η μέση τιμή και η διασπορά της είναι

$$E[X] = \frac{\alpha}{\lambda}, \quad Var[X] = \frac{\alpha}{\lambda^2}.$$

Η κατανομή Γάμμα γενικεύει την εκθετική κατανομή, αφού η εκθετική κατανομή προκύπτει για $\alpha = 1$. Γενικότερα η κατανομή Γάμμα είναι γνωστή ως κατανομή Erlang όταν η παράμετρος α είναι θετικός ακέραιος.

Πολυδιάστατες τυχαίες μεταβλητές
Περίληψη

Αντώνης Οικονόμου
aeconom@math.uoa.gr

24 Απριλίου 2010

Η **από κοινού συνάρτηση κατανομής** ενός ζεύγους τυχαίων μεταβλητών X και Y ορίζεται ως

$$F(x,y) = P(X \leq x, Y \leq y), \quad -\infty < x, y < \infty.$$

Όλες οι πιθανότητες που αφορούν ενδεχόμενα σχετιζόμενα με τις τιμές αυτών των δυο μεταβλητών μπορούν να εκφραστούν μέσω της $F(x,y)$. Υπό αυτή την έννοια, η γνώση της από κοινού συνάρτησης κατανομής επιτρέπει τον υπολογισμό οποιουδήποτε ενδεχομένου σχετίζεται με τις τιμές αυτών των δυο μεταβλητών. Για να πάρουμε τις συναρτήσεις κατανομής των X και Y αντίστοιχα, που ονομάζονται **περιθώριες συναρτήσεις κατανομών** χρησιμοποιούμε τις σχέσεις

$$F_X(x) = P(X \leq x) = \lim_{y \to \infty} F(x,y), \quad F_Y(y) = P(Y \leq y) = \lim_{x \to \infty} F(x,y).$$

Αν οι X και Y είναι και οι δυο διακριτές τυχαίες μεταβλητές, τότε το ζεύγος (X, Y) αναφέρεται ως **διακριτή διδιάστατη τυχαία μεταβλητή** και οι υπολογισμοί πιθανοτήτων που αναφέρονται στις X, Y μπορούν να γίνουν χρησιμοποιώντας την **από κοινού συνάρτηση πιθανότητάς** τους που ορίζεται από τη σχέση

$$p(x,y) = P(X = x, Y = y).$$

Οι συναρτήσεις πιθανότητας των X και Y που ονομάζονται τώρα **περιθώριες συναρτήσεις πιθανότητας** υπολογίζονται από τις σχέσεις

$$p_X(x) = P(X = x) = \sum_y p(x,y), \quad p_Y(y) = P(Y = y) = \sum_x p(x,y).$$

Οι τυχαίες μεταβλητές X, Y λέμε ότι είναι **από κοινού συνεχείς** ή ότι το ζεύγος (X, Y) είναι μια **συνεχής διδιάστατη τυχαία μεταβλητή**, αν υπάρχει μια μη αρνητική συνάρτηση $f(x,y)$, τέτοια ώστε για κάθε υποσύνολο C,

$$P((X,Y) \in C) = \iint_C f(x,y)dxdy.$$

Η συνάρτηση $f(x,y)$ ονομάζεται τότε **από κοινού συνάρτηση πυκνότητας πιθανότητάς** των X και Y. Από τον προηγούμενο ορισμό είναι φανερό ότι για $f(x,y)$ συνεχή έχουμε τη σχέση-ερμηνεία

$$P(x < X \leq x + dx, y < Y \leq y + dy) \simeq f(x,y)dxdy, \quad dxdy \to 0.$$

Οι συναρτήσεις πυκνότητας πιθανότητας των X και Y που ονομάζονται τώρα **περιθώριες συναρτήσεις πυκνότητας πιθανότητας** υπολογίζονται από τις σχέσεις

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy, \quad f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx.$$

Δυο τυχαίες μεταβλητές X και Y λέγονται **ανεξάρτητες** αν για οποιαδήποτε σύνολα A και B ισχύει

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B).$$

Αναγκαία και ικανή συνθήκη για να είναι δυο τυχαίες μεταβλητές X και Y ανεξάρτητες είναι η από κοινού συνάρτηση κατανομής τους, $F(x, y)$ (αντίστοιχα η από κοινού συνάρτηση πιθανότητάς τους για διακριτές ή η από κοινού συνάρτηση πυκνότητας πιθανότητάς τους για συνεχείς), να παραγοντοποιείται σε δυο μέρη που να εξαρτώνται το ένα μόνο από τη μεταβλητή x και το άλλο μόνο από τη μεταβλητή y. Γενικότερα, οι μεταβλητές $X_1, X_2, \ldots, X_n$ είναι ανεξάρτητες αν για οποιαδήποτε σύνολα $A_1, A_2, \ldots, A_n$, ισχύει

$$P(X_1 \in A_1, X_2 \in A_2, \ldots, X_n \in A_n) = P(X_1 \in A_1)P(X_2 \in A_2) \cdots P(X_n \in A_n).$$

Αν X και Y είναι διακριτές τυχαίες μεταβλητές με από κοινού συνάρτηση πιθανότητας $p(x, y)$, τότε η συνάρτηση πιθανότητας του αθροίσματός τους δίνεται από τη σχέση

$$p_{X+Y}(z) = \sum_x p(x, z - x) = \sum_y p(z - y, y).$$

Ανάλογα, αν οι X και Y είναι συνεχείς τυχαίες μεταβλητές με από κοινού συνάρτηση πυκνότητας πιθανότητας $f(x, y)$, τότε η συνάρτηση πυκνότητας πιθανότητας του αθροίσματός τους δίνεται από τη σχέση

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f(x, z - x)dx = \int_{-\infty}^{\infty} f(z - y, y)dy.$$

Αν οι X και Y είναι διακριτές τυχαίες μεταβλητές με από κοινού συνάρτηση πιθανότητας $p(x, y)$, τότε η δεσμευμένη συνάρτηση πιθανότητας της X, δοθέντος ότι $Y = y$, ορίζεται μέσω της σχέσης

$$p_{X|Y}(x|y) = P(X = x|Y = y) = \frac{p(x, y)}{p_Y(y)} = \frac{P(X = x, Y = y)}{P(Y = y)}.$$

Ανάλογα αν οι X και Y είναι συνεχείς τυχαίες μεταβλητές με από κοινού συνάρτηση πυκνότητας πιθανότητας $f(x, y)$, τότε η δεσμευμένη συνάρτηση πυκνότητας πιθανότητας της X, δοθέντος ότι $Y = y$, ορίζεται μέσω της σχέσης

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

Αν έχουμε τις ανεξάρτητες τυχαίες μεταβλητές $X_1, X_2, \ldots, X_n$ με αντίστοιχες συναρτήσεις κατανομής $F_{X_1}(x), F_{X_2}(x), \ldots, F_{X_n}(x)$, τότε η συνάρτηση κατανομής της μέγιστης, $X_{\max} = \max(X_1, X_2, \ldots, X_n)$, δίνεται από τη σχέση

$$F_{X_{\max}}(x) = F_{X_1}(x)F_{X_2}(x) \cdots F_{X_n}(x),$$

ενώ η συνάρτηση κατανομής της ελάχιστης, $X_{\min} = \min(X_1, X_2, \ldots, X_n)$, δίνεται από τη σχέση

$$F_{X_{\min}}(x) = 1 - (1 - F_{X_1}(x))(1 - F_{X_2}(x)) \cdots (1 - F_{X_n}(x)).$$

Μέση τιμή, διασπορά και συνδιακύμανση
Περίληψη

Αντώνης Οικονόμου
aeconom@math.uoa.gr

31 Μαρτίου 2010

Αν οι X και Y είναι διακριτές τυχαίες μεταβλητές με από κοινού συνάρτηση πιθανότητας $p(x,y)$ και $g(x,y)$ είναι μια συνάρτηση, τότε η μέση τιμή της τυχαίας μεταβλητής $g(X,Y)$ μπορεί να υπολογιστεί από τη σχέση

$$E[g(X,Y)] = \sum_y \sum_x g(x,y)p(x,y).$$

Ανάλογα, αν οι X και Y είναι συνεχείς τυχαίες μεταβλητές με από κοινού συνάρτηση πυκνότητας πιθανότητας $f(x,y)$ και $g(x,y)$ είναι μια συνάρτηση, τότε η μέση τιμή της τυχαίας μεταβλητής $g(X,Y)$ μπορεί να υπολογιστεί από τη σχέση

$$E[g(X,Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y)f(x,y)dxdy.$$

Χρησιμοποιώντας αυτές τις σχέσεις μπορούμε να δούμε ότι ισχύει η σχέση

$$E[X+Y] = E[X] + E[Y],$$

που γενικεύεται στη σχέση

$$E[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} E[X_i].$$

Προσοχή όμως! Ενώ ισχύει ότι η μέση τιμή του αθροίσματος τυχαίων μεταβλητών ισούται με το άθροισμα των μέσων τιμών, γενικά δεν ισχύει ότι η μέση τιμή του γινομένου τυχαίων μεταβλητών ισούται με το γινόμενο των μέσων τιμών. Όμως:

$$X, Y \quad \text{ανεξάρτητες} \quad \Rightarrow E[XY] = E[X]E[Y]$$

και γενικότερα

$$X_1, X_2, \ldots, X_n \quad \text{ανεξάρτητες} \quad \Rightarrow E[\prod_{i=1}^{n} X_i] = \prod_{i=1}^{n} E[X_i].$$

Αν X και Y είναι ανεξάρτητες τυχαίες μεταβλητές τότε και οποιεσδήποτε συναρτήσεις τους $g(X)$ και $h(Y)$ είναι ανεξάρτητες και επομένως θα έχουμε και τη συνεπαγωγή

$$X, Y \quad \text{ανεξάρτητες} \quad \Rightarrow E[g(X)h(Y)] = E[g(X)]E[h(Y)].$$

Βέβαια χρειάζεται και πάλι προσοχή. Οι αντίστροφες των παραπάνω τριών συνεπαγωγών δεν ισχύουν.

Η **συνδιακύμανση** δυο τυχαίων μεταβλητών X και Y ορίζεται να είναι η ποσότητα

$$Cov[X,Y] = E[(X - E[X])(Y - E[Y])].$$

Δυο τυχαίες μεταβλητές X και Y με $Cov[X,Y]=0$ λέγονται **ασυσχέτιστες**. Είναι εύκολο να δούμε ότι ανεξάρτητες μεταβλητές είναι πάντα ασυσχέτιστες αλλά το αντίστροφο δεν ισχύει. Για τη συνδιακύμανση ισχύουν οι παρακάτω σχέσεις:

$$\begin{aligned}
Cov[X,Y] &= Cov[Y,X], \\
Cov[X,X] &= Var[X], \\
Cov[aX+b,Y] &= aCov[X,Y], \\
Cov[\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j] &= \sum_{i=1}^{n}\sum_{j=1}^{m} Cov[X_i,Y_j].
\end{aligned}$$

Η διασπορά του αθροίσματος τυχαίων μεταβλητών δεν ισούται γενικά με το άθροισμα των διασπορών τους. Για τον υπολογισμό χρειάζεται να γνωρίζουμε και τις συνδιακυμάνσεις μεταξύ των μεταβλητών. Πιο συγκεκριμένα ισχύει η σχέση

$$Var\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} Var[X_i] + 2\sum\sum_{i<j} Cov[X_i,X_j].$$

Από τη σχέση αυτή παίρνουμε τη συνεπαγωγή

$$X_1, X_2, \ldots, X_n \quad \text{ανεξάρτητες} \quad \Rightarrow Var\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} Var[X_i].$$