

Παρουσίαση
βασισμένη στο βιβλίο

Στατιστική **Τόμος Α**

ΤΩΝ

Βασίλη Αγγελή & Κατερίνας Δημάκη

Αθήνα, 2011

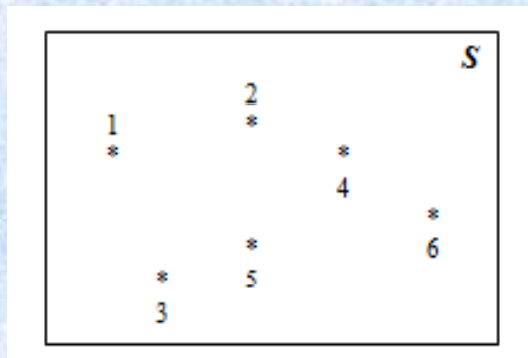
4. Πιθανότητες - Κατανομές

Βασικές Έννοιες

- **Τυχαίο φαινόμενο** είναι ένα φαινόμενο η κατάληξη του οποίου δεν είναι γνωστή εκ των προτέρων.
- **Τυχαίο πείραμα** είναι μια διαδικασία της οποίας το αποτέλεσμα δεν είναι γνωστό εκ των προτέρων (π.χ. ρίξιμο ζαριού, στρίψιμο νομίσματος, γέννηση παιδιού κ.λ.π.).
- **Απλά (ή Στοιχειώδη) Ενδεχόμενα** ονομάζονται τα δυνατά αποτελέσματα ενός πειράματος.
- **Δειγματικός χώρος** είναι το σύνολο όλων των δυνατών αποτελεσμάτων ενός τυχαίου πειράματος. Ο δειγματικός χώρος συμβολίζεται συνήθως με το γράμμα S .

Δειγματικός Χώρος & Ενδεχόμενα

Δειγματικός χώρος S ενός τυχαίου πειράματος καλείται το σύνολο των δυνατών αποτελεσμάτων του.



Παραδείγματα δειγματικών χώρων

- Η ρίψη ενός ζαριού $S=\{1,2,3,4,5,6\}$
- Η ρίψη ενός νομίσματος $S=\{Κ, Γ\}$
- Ο βαθμός στο μάθημα των πιθανοτήτων $S=\{0,1,2,...,10\}$
- Ο αριθμός των μελών ενός τυχαία επιλεγμένου νοικοκυριού $S=\{1,2,3,...\}$

Δειγματικός Χώρος & Ενδεχόμενα

Ενδεχόμενο ενός πειράματος τύχης ονομάζεται το σύνολο που έχει ως στοιχεία ένα ή περισσότερα αποτελέσματα του πειράματος τύχης.

Ένα ενδεχόμενο που έχει ένα μόνο στοιχείο του S ονομάζεται **απλό**, ενώ αν έχει περισσότερα ονομάζεται **σύνθετο**

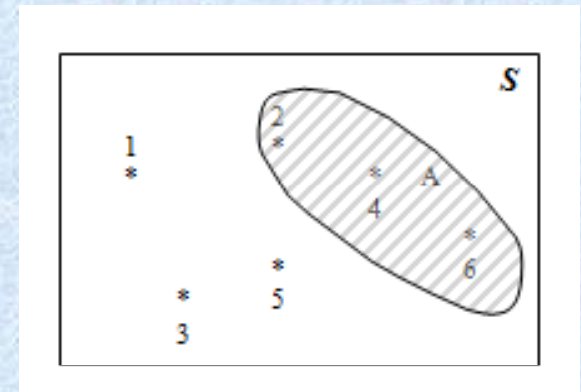
Χαρακτηριστικά:

1. Ένα ενδεχόμενο είναι υποσύνολο του δειγματικού χώρου.
2. Κάθε στοιχείο του S αντιπροσωπεύει ένα απλό ενδεχόμενο.

Παραδείγματα:

Κατά τη ρίψη του ζαριού το $B=\{1\}$ είναι απλό, ενώ το $A=\{2,4,6\}$ είναι σύνθετο

Όταν το αποτέλεσμα ενός πειράματος σε μια συγκεκριμένη εκτέλεσή του είναι στοιχείο ενός ενδεχομένου τότε το ενδεχόμενο **πραγματοποιείται ή συμβαίνει**. Τα στοιχεία ενός ενδεχομένου λέγονται και **ευνοϊκές περιπτώσεις** για την πραγματοποίησή του. Έτσι, το B έχει τρεις ευνοϊκές περιπτώσεις και πραγματοποιείται όταν φέρουμε ή 2 ή 4 ή 6.



Δειγματικός Χώρος & Ενδεχόμενα

- Ο δειγματικός χώρος S καλείται **βέβαιο** ενδεχόμενο
- Ενδεχόμενο θεωρείται και το ενδεχόμενο το οποίο δεν πραγματοποιείται σε καμιά εκτέλεση του πειράματος τύχης, ονομάζεται **αδύνατο ενδεχόμενο** και συμβολίζεται με $\emptyset$.
- Το πλήθος των στοιχείων ενός ενδεχομένου A συμβολίζεται με $N(A)$

Πράξεις με Ενδεχόμενα

Έχουμε τις εξής βασικές πράξεις: Τομή / Ένωση / Συμπληρωματικό

Η Τομή " $\cap$ " (Συμβολίζεται: $A \cap B$ & Διαβάζεται: "Α τομή Β" ή "Α και Β") είναι το ενδεχόμενο το οποίο αποτελείται από τα απλά ενδεχόμενα που ανήκουν και στο Α και στο Β. Με άλλα λόγια είναι το ενδεχόμενο το οποίο συμβαίνει όταν συμβαίνει και το Α και το Β.

Παραδείγματα:

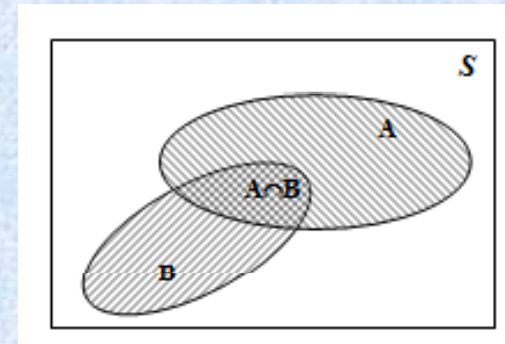
Έστω $S = \{1, 2, 3, 4, 5, 6\}$, $A = \{1, 2, 5\}$, $B = \{2, 6\}$

Για να πραγματοποιηθεί και το Α και Β

δηλ. το $A \cap B$ με άλλα λόγια το

Α τομή Β, πρέπει να

έρθει το «2». Άρα $A \cap B = \{2\}$, $N(A \cap B) = 1$



Πράξεις με Ενδεχόμενα

Η Ένωση “ $\cup$ ” (Συμβολίζεται: $A \cup B$ & Διαβάζεται: “Α ένωση Β”, ή “Α ή Β”)

είναι το ενδεχόμενο το οποίο αποτελείται από τα απλά ενδεχόμενα που ανήκουν είτε στο Α είτε στο Β είτε και στα δύο. Με άλλα λόγια είναι το ενδεχόμενο το οποίο συμβαίνει όταν συμβαίνει είτε το Α είτε το Β είτε και τα δύο. Τα στοιχεία της ένωσης **ανήκουν τουλάχιστον** σε ένα από τα Α, Β

Παραδείγματα:

Έστω $S = \{1,2,3,4,5,6\}$ $A = \{1,2,5\}$ $B = \{2,6\}$

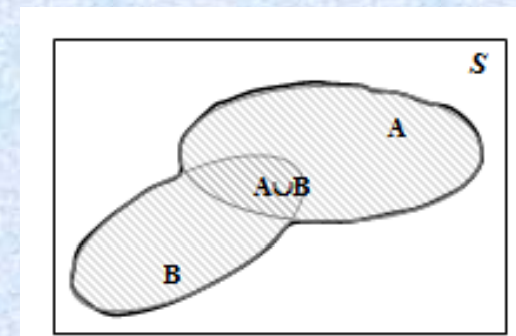
Για να πραγματοποιηθεί το Α ή το Β

δηλ. **τουλάχιστον ένα από τα Α, Β**

με άλλα λόγια το Α ένωση Β,

πρέπει να έρθει ή το 1 ή το 2 ή το 5 ή το 6,

Δηλ. $A \cup B = \{1,2,5,6\}$, $N(A \cup B) = 4$



Πράξεις με Ενδεχόμενα

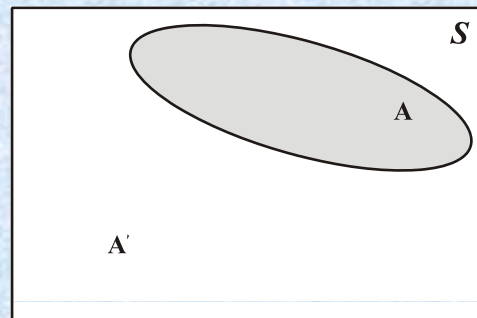
Το Συμπλήρωμα ή Συμπληρωματικό ενός ενδεχομένου A συμβολίζεται με A' και συμβαίνει όταν δεν συμβαίνει το A .

Παραδείγματα:

Έστω $S = \{1, 2, 3, 4, 5, 6\}$ αν $A = \{1, 2, 5\}$ τότε $A' = \{3, 4, 6\}$ και $N(A') = 3$
αν $B = \{2, 3\}$ τότε $B' = \{1, 4, 5, 6\}$ και $N(B') = 4$

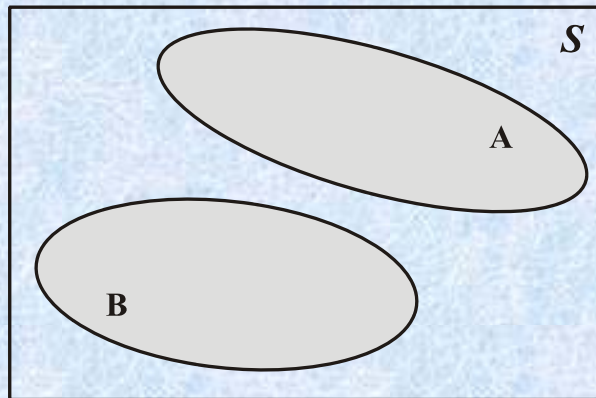
Παρατηρώ ότι

$$A \cap A' = \emptyset \text{ και } A \cup A' = S$$



Αν η τομή των δύο ενδεχομένων είναι το αδύνατο ενδεχόμενο,
τότε αυτά καλούνται **ασυμβίβαστα ενδεχόμενα ή ξένα**

Ασυμβίβαστα ενδεχόμενα A, B



Ορισμοί Πιθανότητας (1)

Κλασσική προσέγγιση

Η προσέγγιση αυτή οφείλεται στον Laplace ο οποίος όρισε την πιθανότητα ενός ενδεχομένου ως τον λόγο του αριθμού των ευνοϊκών για το ενδεχόμενο περιπτώσεων δια του συνολικού αριθμού των περιπτώσεων:

$$P(E) = \frac{N(E)}{N(S)} \quad \frac{\text{πληθος ευνοικων περιπτωσεων}}{\text{πληθος δυνατων περιπτωσεων}}$$

Απαραίτητη προϋπόθεση στην κλασσική προσέγγιση είναι ότι όλα τα δυνατά ενδεχόμενα είναι ισοπίθανα και ασυμβίβαστα

Με άλλα λόγια:

Έστω ότι ο δειγματικός χώρος S αποτελείται από n ισοπίθανα ενδεχόμενα $A_1, A_2, \dots, A_n$ και

$$P(A_i) = P \text{ όπου } i=1,2,\dots,n$$

$$\begin{aligned} 1 &= P(S) = P(A_1 \cup A_2 \cup \dots \cup A_n) \\ &= P(A_1) + P(A_2) + \dots + P(A_n) \\ &= n P \text{ άρα } \mathbf{P=1/n} \end{aligned}$$

Ορισμοί Πιθανότητας (2)

Προσέγγιση της πιθανότητας ως οριακής Σχετικής Συχνότητας

Στην προσέγγιση αυτή που οφείλεται στον Von Mises οι πιθανότητες αντιστοιχούν σε σχετικές συχνότητες κάποιας ακολουθίας πειραμάτων. Παρόλο που στη πράξη οι παρατηρούμενες ακολουθίες είναι πεπερασμένες δίνουν, καθώς ο αριθμός των επαναλαμβανόμενων πειραμάτων μεγαλώνει, μια ένδειξη της τυχαιότητας στην εμφάνιση των αποτελεσμάτων και της σύγκλισης των σχετικών συχνοτήτων.

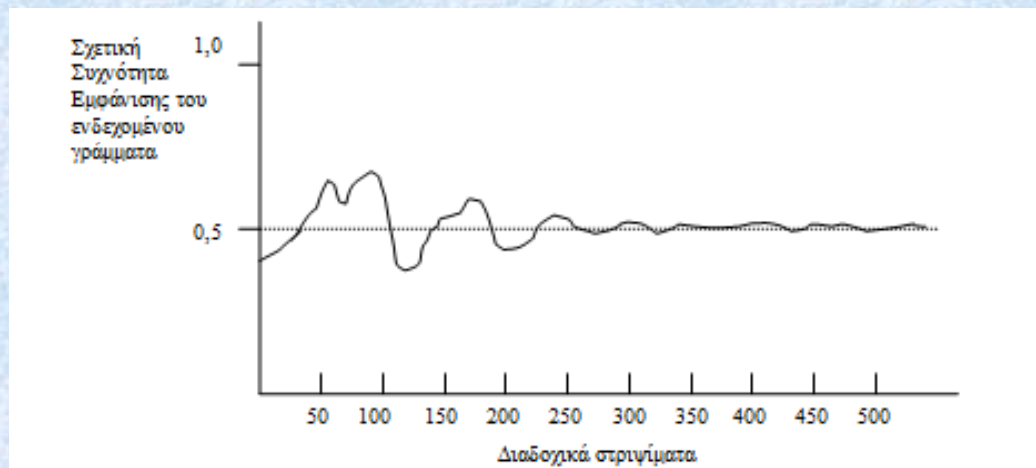
Η πιθανότητα ενός ενδεχομένου E ισούται με το όριο της σχετικής συχνότητας του, δηλαδή

$$P(E) = \lim_{n \rightarrow \infty} \frac{n(E)}{n}$$

όπου

$n(E)$
 n

ο αριθμός των εμφανίσεων του ενδεχομένου E και
ο αριθμός των επαναλήψεων του πειράματος



Αξιωματική Θεμελίωση Πιθανότητας

Σύμφωνα με τον Kolmogorov η **πιθανότητα είναι μια συνάρτηση** που ορίζεται ως εξής:

Έστω S ένας δειγματικός χώρος και έστω $A \subset S$ ένα ενδεχόμενο. Ορίζουμε ως **συνάρτηση πιθανότητας** μια συνάρτηση P :

$$P: A \rightarrow R$$

η οποία σε κάθε ενδεχόμενο A αντιστοιχεί έναν πραγματικό αριθμό $P(A)$ έτσι ώστε να ικανοποιούνται τα παρακάτω αξιώματα:

1. $P(A) \geq 0$

2. $P(S) = 1$

3. Αν A_1 και A_2 είναι ενδεχόμενα ασυμβίβαστα ($A_1 \cap A_2 = \emptyset$)
τότε $P(A_1 \cup A_2) = P(A_1) + P(A_2)$.

Ο αριθμός $P(A)$ ονομάζεται πιθανότητα του ενδεχομένου A .

Για κάθε ενδεχόμενο A ισχύει: $P(A) \leq 1$ $P(A') = 1 - P(A)$ $P(\emptyset) = 0$

Νόμος της Πρόσθεσης των Πιθανοτήτων

Σύμφωνα με το τρίτο αξίωμα του Kolmogorov αν δύο ενδεχόμενα είναι ασυμβίβαστα δηλαδή $A_1 \cap A_2 = \emptyset$ τότε $P(A_1 \cup A_2) = P(A_1) + P(A_2)$

Το αξίωμα αυτό γενικεύεται και για περισσότερα ενδεχόμενα. Για τρία ενδεχόμενα A_1, A_2, A_3 ανά δύο ασυμβίβαστα μεταξύ τους, ισχύει:

$$P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) = \sum_{i=1}^3 P(A_i)$$

Για n ενδεχόμενα $A_i, i = 1, 2, \dots, n$ ανά δύο ασυμβίβαστα μεταξύ τους ($A_i \cap A_j = \emptyset, i \neq j$) ισχύει:

$$P\left[\bigcup_{i=1}^n A_i\right] = P(A_1) + P(A_2) + \dots + P(A_n) = \sum_{i=1}^n P(A_i)$$

Νόμος της Πρόσθεσης των Πιθανοτήτων

Στην περίπτωση που τα ενδεχόμενα A_1, A_2 δεν είναι ασυμβίβαστα ($A_i \cap A_j \neq \emptyset$) ισχύει

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

Για τρία ενδεχόμενα A_1, A_2, A_3 που δεν είναι ανά δύο ασυμβίβαστα μεταξύ τους η σχέση

$$\begin{aligned} P(A_1 \cup A_2 \cup A_3) = & P(A_1) + P(A_2) + P(A_3) \\ & - P(A_1 \cap A_2) - P(A_1 \cap A_3) - P(A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_3) \end{aligned}$$

Το συμπέρασμα αυτό γενικεύεται και για n ενδεχόμενα.

Παράδειγμα 1

Στο πρώτο εξάμηνο σπουδών ενός πανεπιστημιακού τμήματος προσφέρονται δύο μαθήματα επιλογής: Μαθηματικά και Στατιστική. Κάθε φοιτητής μπορεί να επιλέξει είτε ένα από τα δύο μαθήματα είτε και τα δύο μαθήματα. Η πιθανότητα επιλογής των Μαθηματικών είναι 0.15, ενώ η πιθανότητα επιλογής της Στατιστικής είναι 0.80. Τέλος, η πιθανότητα επιλογής τουλάχιστον ενός από τα δύο μαθήματα είναι 0.90. Ζητείται να υπολογιστεί η πιθανότητα ένας τυχαία επιλεγμένος φοιτητής να επιλέξει και τα δύο μαθήματα.

Λύση

Ας συμβολίσουμε με A το ενδεχόμενο «φοιτητής επιλέγει το μάθημα των Μαθηματικών» και με B το ενδεχόμενο «φοιτητής επιλέγει το μάθημα της Στατιστικής». Τότε, από το νόμο της πρόσθεσης για δύο ενδεχόμενα, προκύπτει:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Δίνεται όμως ότι $P(A) = 0.15$, $P(B) = 0.80$ και $P(A \cup B) = 0.90$.

Έτσι, από την παραπάνω σχέση προκύπτει:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \Rightarrow P(A \cap B) = P(A) + P(B) - P(A \cup B).$$

Κατά συνέπεια $P(A \cap B) = 0.15 + 0.80 - 0.90 = 0.05$.

Άρα, η πιθανότητα κάποιος φοιτητής να επιλέξει και τα δύο μαθήματα είναι 0.05.

Παράδειγμα 2

Στο πρώτο έτος του τμήματός σας φοιτούν 240 αγόρια και 220 κορίτσια. Το 40% των αγοριών και το 45% των κοριτσιών έχουν κοπεί στα Μαθηματικά. Επιλέγουμε τυχαία έναν φοιτητή. Ποιες οι πιθανότητες των ενδεχομένων:

$A = \{\text{Δεν έχει κοπεί στα Μαθηματικά}\}$

$B = \{\text{Είναι κορίτσι και έχει κοπεί στα Μαθηματικά}\}$

$\Gamma = \{\text{Είναι αγόρι ή δεν έχει κοπεί στα Μαθηματικά}\}$

Λύση

	Έχουν κοπεί στα Μαθηματικά	Δεν έχουν κοπεί στα Μαθηματικά	Σύνολο
Αγόρια	96	144	240
Κορίτσια	99	121	220
Σύνολο	195	265	460

$$N(S) = 460$$

$$N(A) = 265 \text{ οπότε } P(A) = N(A)/N(S) = 265/460 = 0.576$$

$$N(B) = 99 \text{ οπότε } P(B) = N(B)/N(S) = 99/460 = 0.215$$

$$N(\Gamma) = 361 \text{ οπότε } P(\Gamma) = N(\Gamma)/N(S) = 361/460 = 0.785$$

$\Gamma = \{\text{όλα τα αγόρια (240) και τα κορίτσια που δεν έχουν κοπεί στα Μαθηματικά (121)}\}$

Αρχές Απαρίθμησης

Σύμφωνα με τον κλασσικό ορισμό η πιθανότητα εμφάνισης ενός ενδεχομένου A εκφράζεται ως ο λόγος του αριθμού των ευνοϊκών για το ενδεχόμενο αποτελεσμάτων προς τον αριθμό του συνόλου των δυνατών αποτελεσμάτων του:

$$P(A) = \frac{n(A)}{N}$$

Σε πιο σύνθετα όμως προβλήματα η απαρίθμηση των ευνοϊκών και των δυνατών αποτελεσμάτων είναι εξαιρετικά δύσκολη αν όχι αδύνατη

Υπάρχουν συγκεκριμένες τεχνικές που μας επιτρέπουν να υπολογίσουμε το συνολικό αριθμό των αποτελεσμάτων ενός ενδεχομένου χωρίς να απαιτείται η εξαντλητική απαρίθμηση τους

- α. Πολλαπλασιαστική Αρχή
- β. Μεταθέσεις
- γ. Διατάξεις
- δ. Συνδυασμοί

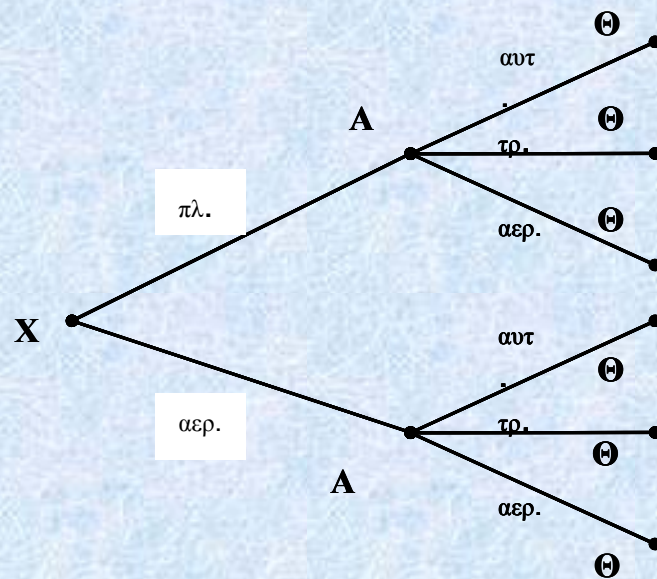
Πολλαπλασιαστική Αρχή

Έστω ότι η σύνθετη ενέργεια E συνίστανται στην εκτέλεση των επί μέρους ενεργειών $E_1, E_2, \dots, E_k$. Έστω επιπλέον ότι καθεμιά από τις ενέργειες αυτές έχει αντίστοιχα $n_1, n_2, \dots, n_k$ δυνατά αποτελέσματα. Τότε η ενέργεια E έχει $n_1, n_2, \dots, n_k$ δυνατά αποτελέσματα.

Η πολλαπλασιαστική αρχή μπορεί να παρουσιασθεί γραφικά με τη χρήση ενός **δενδρικού διαγράμματος**. Ένα τέτοιο διάγραμμα παρουσιάζει το δειγματικό χώρο μιας ενέργειας με τη μορφή ενός δένδρου που τα κλαδιά του χωρίζονται σε τμήματα. Το πρώτο τμήμα ενός κλαδιού απεικονίζει ένα δυνατό αποτέλεσμα της πρώτης ενέργειας (E_1). Το δεύτερο ένα δυνατό αποτέλεσμα της δεύτερης ενέργειας (E_2) κ.λ.π. Προφανώς ο συνολικός αριθμός των διακεκριμένων κλαδιών μας δίνει το συνολικό αριθμό των τρόπων με τους οποίους μπορεί να ολοκληρωθεί η σύνθετη ενέργεια E .

Παράδειγμα 3

Έστω ότι ένας φοιτητής θέλει να ταξιδέψει από τη Χίο στη Θεσσαλονίκη μέσω Αθήνας (ενέργεια E). Το πρώτο μέρος του ταξιδιού του (Χίος – Αθήνα / ενέργεια E_1) μπορεί να γίνει με δύο τρόπους, δηλαδή με πλοίο ή με αεροπλάνο ($n_1 = 2$). Αντίστοιχα, το δεύτερο μέρος του ταξιδιού του (Αθήνα – Θεσσαλονίκη / ενέργεια E_2) μπορεί να γίνεται με τρεις τρόπους, δηλαδή με αυτοκίνητο, με τρένο ή με αεροπλάνο ($n_2 = 3$). Άρα, το ταξίδι **Χίος – Θεσσαλονίκη** μπορεί να γίνει με $n_1 \cdot n_2 = 2 \cdot 3 = 6$ τρόπους.



X: Χίος
A: Αθήνα
Θ: Θεσσαλονίκη
αυτ. : αυτοκίνητο
αερ. : αεροπλάνο
τρ. : τρένο
πλ. : πλοίο

Μεταθέσεις

Μεταθέσεις ενός συνόλου A

Είναι οι διαφορετικοί τρόποι με τους οποίους μπορούμε να βάλουμε σε μια σειρά

Τα n στοιχεία ενός συνόλου

Το πλήθος των μεταθέσεων των n στοιχείων συμβολίζεται με M_n και είναι

$$P_n = 1 \cdot 2 \cdot 3 \dots n = n!$$

ΠΡΟΣΟΧΗ: Στις μεταθέσεις βάζουμε όλα τα στοιχεία του συνόλου σε μια σειρά

Αν μεταξύ των n στοιχείων που θέλουμε να τοποθετήσουμε υπάρχουν και ορισμένα **όμοια μεταξύ τους** τότε το σύνολο των n στοιχείων μπορεί να χωρισθεί σε **k υποσύνολα** έτσι ώστε το πρώτο να περιέχει **n_1 όμοια στοιχεία**, το δεύτερο n_2 όμοια κ.λ.π. .

Στην περίπτωση αυτή οι δυνατοί τρόποι τοποθέτησης των n στοιχείων ονομάζονται **επαναληπτικές μεταθέσεις** και δίνονται από τον τύπο:

$$\frac{n!}{n_1! n_2! \dots n_k!}$$

Παράδειγμα 4

α) Οι διαφορετικοί τρόποι με τους οποίους τρεις φοιτητές μπορούν να καθίσουν σε τρεις καρέκλες είναι

$$3! = 1 \cdot 2 \cdot 3 = 6 .$$

β) Οι διαφορετικοί τρόποι με τους οποίους τα γράμματα της λέξης **Α Ι Γ Α Ι Ο** μπορούν να τοποθετηθούν σε ευθεία είναι:

$$\frac{6!}{2!2!1!1!} = 180$$

Διατάξεις

Σε ορισμένες περιπτώσεις δεν ενδιαφερόμαστε για όλα τα στοιχεία ενός συνόλου αλλά για ένα μόνο μέρος τους. Οι **Διατάξεις** των n στοιχείων ενός συνόλου ανά x εκφράζουν τους διαφορετικούς τρόπους με τους οποίους μπορεί να επιλέξει κάποιος x από τα n στοιχεία ενός συνόλου όταν η σειρά επιλογής τους ενδιαφέρει.

Οι διαφορετικοί αυτοί τρόποι δίνονται από
$$P_{n \ x} = P(n, x) = \frac{n!}{(n-x)!}$$

ΠΡΟΣΟΧΗ: Στην περίπτωση των διατάξεων η σειρά εμφάνισης των στοιχείων σε κάθε επιλογή μας ενδιαφέρει.

Με άλλα λόγια δύο επιλογές είναι διαφορετικές όχι μόνο όταν περιέχουν διαφορετικά στοιχεία αλλά και όταν περιέχουν τα ίδια στοιχεία αλλά με διαφορετική σειρά εμφάνισης

Παράδειγμα 5

- a) Ας υποθέσουμε ότι οι θέσεις του Προέδρου, Αντιπροέδρου, Γραμματέα και Ταμία μιας επιτροπής θα κληρωθούν με τυχαία επιλογή ισάριθμων ατόμων μεταξύ 20 υποψηφίων. Ας υποθέσουμε ότι η σειρά επιλογής των ατόμων θα καθορίζει και τη θέση που θα καταλάβουν. Με άλλα λόγια ο πρώτος που θα επιλεγεί θα καταλάβει τη θέση του Προέδρου, ο δεύτερος της θέση του Αντιπροέδρου, ο τρίτος τη θέση του Γραμματέα και ο τέταρτος τη θέση του Ταμία.

Σύμφωνα με το παραπάνω τύπο οι διαφορετικοί τρόποι επιλογής των τεσσάρων ατόμων από τους 20 υποψηφίους είναι:

$$P_{20\ 4} = P(20, 4) = \frac{20!}{(20-4)!} = \frac{20!}{16!} = 116.280$$

- b) Σε μια ιπποδρομία παίρνουν μέρος 12 άλογα. Με πόσους τρόπους μπορεί να σχηματιστεί η τριάδα των νικητών
Από τα 12 στοιχεία επιλέγουμε 3 διαφορετικά μεταξύ τους και τα βάζουμε σε σειρά.

Το πλήθος των διατεταγμένων τριάδων που μπορούν να σχηματιστούν είναι

$$P_{12\ 3} = \frac{12!}{(12-3)!} = \frac{12!}{9!} = \frac{1.2.3..12}{1.2.3...9} = 10.11.12 = 1320$$

Συνδυασμοί

Οι **συνδυασμοί** των n στοιχείων ενός συνόλου ανά x εκφράζουν τους διαφορετικούς τρόπους με τους οποίους μπορεί να επιλέξει κάποιος x από τα n στοιχεία ενός συνόλου όταν η σειρά επιλογής δεν τον ενδιαφέρει.

Οι διαφορετικοί τρόποι δίνονται από τον τύπο:

$$C_{n \ x} = C(n, x) = \frac{n!}{x!(n-x)!}$$

ΠΡΟΣΟΧΗ: Στην περίπτωση των Συνδυασμών δύο επιλογές είναι διαφορετικές μόνο όταν περιέχουν διαφορετικά στοιχεία

Υπενθύμιση

$$\binom{n}{n} = 1 \quad \binom{n}{1} = n \quad \binom{n}{0} = 1$$

Παράδειγμα 5

- a) Ας υποθέσουμε ότι μια τετραμελής επιτροπή θα σχηματισθεί με τυχαία επιλογή ισάριθμων ατόμων μεταξύ 20 υποψηφίων. Στην περίπτωση αυτή η σειρά επιλογής δεν ενδιαφέρει.

Σύμφωνα με τον παραπάνω τύπο οι διαφορετικοί τρόποι σχηματισμού της επιτροπής στην περίπτωση αυτή είναι:

$$C_{20\ 4} = C(20, 4) = \frac{20!}{4!(20-4)!} = \frac{20!}{4!16!} = 4.845$$

- b) Δίνονται 8 διαφορετικά σημεία, τα οποία ανά τρία δεν είναι συνευθειακά. Πόσες ευθείες και πόσα τρίγωνα ορίζονται;

$$C_{8\ 2} = C(8, 2) = \binom{8}{2} = \frac{8!}{2!(8-2)!} = \frac{8!}{2!6!} = \frac{1*2*...*8}{(1*2)*(1*2*3*...*6)} = \frac{7*8}{2} = 28 \text{ ευθειες}$$

$$C_{8\ 3} = C(8, 3) = \binom{8}{3} = \frac{8!}{3!(8-3)!} = \frac{8!}{3!5!} = \frac{1*2*...*8}{(1*2*3)*(1*2*...*5)} = \frac{6*7*8}{1*2*3} = 56 \text{ τρίγωνα}$$

Δεσμευμένη Πιθανότητα

Σε πολλές περιπτώσεις όταν ζητάμε να υπολογίσουμε την πιθανότητα ενός ενδεχομένου έχουμε ήδη κάποιες πληροφορίες γι' αυτό. Οι πληροφορίες αυτές στην ουσία περιορίζουν τον **αρχικό δειγματικό χώρο σε ένα υποσύνολο του**.

Ένα τέτοιο υποσύνολο ονομάζεται **περιορισμένος δειγματικός χώρος**. Οι πιθανότητες που αναφέρονται σε περιορισμένους δειγματικούς χώρους ονομάζονται **δεσμευμένες πιθανότητες**.

Η δεσμευμένη πιθανότητα εμφάνισης ενός ενδεχομένου A δοθέντος ότι έχει ήδη συμβεί ένα ενδεχόμενο B, το οποίο και έχει περιορίσει τον αρχικό δειγματικό χώρο του A συμβολίζεται με $P(A|B)$ και ορίζεται ως εξής:

$$P(A|B) = \frac{N(A \cap B)}{N(B)} = \frac{\frac{N(A \cap B)}{N(S)}}{\frac{N(B)}{N(S)}} \frac{P(A \cap B)}{P(B)}, P(B) > 0$$

Η πιθανότητα αυτή λέγεται **δεσμευμένη πιθανότητα** του A με δεδομένο το B

Αντίστοιχα η δεσμευμένη πιθανότητα ενός ενδεχομένου B δοθέντος ότι έχει ήδη συμβεί το A συμβολίζεται με $P(B|A)$ και ορίζεται ως εξής:

$$P(B|A) = \frac{P(A \cap B)}{P(A)}, P(A) > 0$$

Παράδειγμα 6

Για το διοικητικό συμβούλιο μιας επιχείρησης θα εκλεγεί ένας αντιπρόσωπος. Υποψήφιοι είναι 7 άνδρες και 8 γυναίκες. Από τους υποψηφίους 3 άνδρες και 6 γυναίκες είναι διοικητικοί υπάλληλοι, ενώ 4 άνδρες και 2 γυναίκες είναι τεχνικοί υπάλληλοι. Τα δεδομένα φαίνονται στον ακόλουθο πίνακα

	Διοικητικός	Τεχνικός	Σύνολο
Άνδρας	3	4	7
Γυναίκα	6	2	8
Σύνολο	9	6	15

Ζητάμε την πιθανότητα

α) να εκλεγεί διοικητικός

β) να εκλεγεί γυναίκα

γ) να εκλεγεί διοικητικός δεδομένου ότι έχει ήδη εκλεγεί γυναίκα

Τα 15 στοιχεία του δειγματικού χώρου είναι ισοπίθανα, άρα

$$P(A)=9/15 \quad \text{και} \quad P(B)=8/15$$

Δεδομένου ότι ο εκλεγμένος αντιπρόσωπος είναι γυναίκα (σύνολο 8 γυναίκες) η πιθανότητα να είναι και διοικητικός είναι $6/8$.

Η πιθανότητα $P(\text{να είναι διοικητικός} \mid \text{έχει εκλεγεί γυναίκα})=6/8$

Η πιθανότητα $P(\text{να είναι γυναίκα} \mid \text{έχει εκλεγεί διοικητικός})=6/9$

Νόμος πολλαπλασιασμού πιθανοτήτων, από κοινού πιθανότητες & περιθώριες πιθανότητες

Ξεκινώντας από τη δεσμευμένη πιθανότητα του ενδεχομένου A δοθέντος του B [$P(A/B)$] μπορούμε να προσδιορίζουμε από κοινού πιθανότητα των A και B.

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, P(B) > 0 \Leftrightarrow P(A \cap B) = P(A|B) P(B)$$

$$P(B|A) = \frac{P(A \cap B)}{P(A)}, P(A) > 0 \Leftrightarrow P(A \cap B) = P(B|A) P(A)$$

Οι σχέσεις αυτές είναι γνωστές ως **Νόμος Πολλαπλασιασμού Πιθανοτήτων** και η $P(A \cap B)$ ονομάζεται **από κοινού πιθανότητα** των A και B.

Ο **νόμος του πολλαπλασιασμού** μπορεί να επεκταθεί σε η ενδεχόμενα

Πίνακας από κοινού πιθανοτήτων

Ο πίνακας αυτός έχει την **ίδια μορφή** με τον πίνακα συνάφειας. Η **διαφορά** του είναι ότι ενώ στον πίνακα συνάφειας οι **είσοδοι εκφράζουν πλήθος στοιχείων που συγκεντρώνουν δύο χαρακτηριστικά**, στον πίνακα των από κοινού πιθανοτήτων οι είσοδοι εκφράζουν **από κοινού πιθανότητες**, δηλαδή *πιθανότητες εμφάνισης ενδεχομένων που συγκεντρώνουν δύο κοινά χαρακτηριστικά*

	Διοικητικός	Τεχνικός	Σύνολο
Άνδρας	0.2	0.26	0.46
Γυναίκα	0.4	0.14	0.54
Σύνολο	0.6	0.4	1

Στον πίνακα των από κοινού πιθανοτήτων το άθροισμα των από κοινού πιθανοτήτων μιας σειράς ή στήλης **ονομάζεται περιθώρια πιθανότητα** του ενδεχομένου το οποίο αντιστοιχεί στη σειρά /στήλη αυτή.

Για παράδειγμα, η περιθώρια πιθανότητα του ενδεχομένου A είναι,

$$P(A) = P(A \cap \Delta) + P(A \cap T) = 0.46$$

Θεώρημα Ολικής Πιθανότητας

Διαμέριση ενός συνόλου S ορίζουμε μια πεπερασμένη συλλογή υποσυνόλου του S τα οποία ικανοποιούν τις παρακάτω δύο συνθήκες:

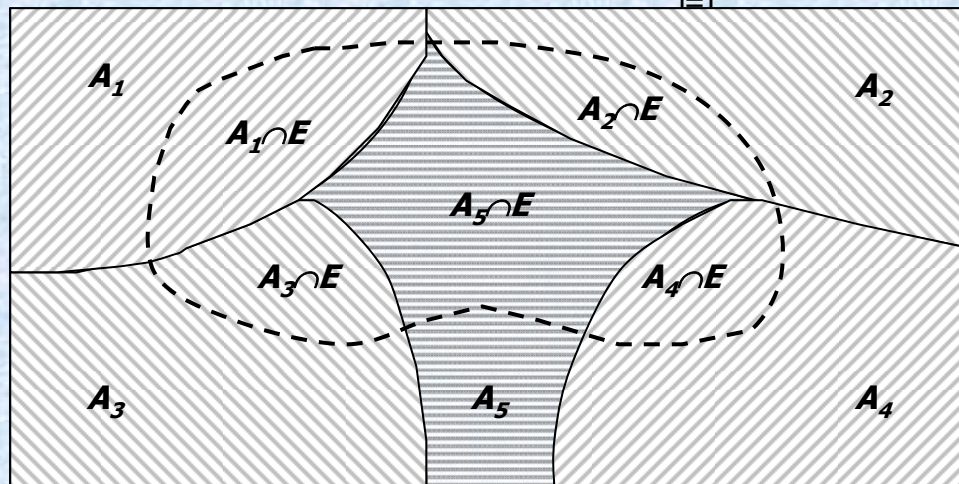
i. $S = A_1 \cup A_2 \cup \dots \cup A_n$

ii. $A_i \cap A_j = \emptyset, i, j = 1, 2, \dots, n, i \neq j.$

Θεώρημα της Ολικής Πιθανότητας

Έστω ότι είναι μία διαμέριση του δειγματικού χώρου S τέτοια ώστε $P(A_i) \neq 0$, $i = 1, 2, \dots, n$. Τότε για κάθε ενδεχόμενο E έχουμε ότι,

$$P(E) = \sum_{i=1}^n P(A_i) P(E|A_i)$$



Το θεώρημα της ολικής πιθανότητας υπολογίζει την πιθανότητα ενός ενδεχομένου (E) ως συνάρτηση των δεσμευμένων πιθανοτήτων του σε σχέση με τα στοιχεία κάποιας διαμέρισης του και των πιθανοτήτων των στοιχείων αυτών

Παράδειγμα 7

Τρία κουτιά περιέχουν άσπρες (A) και μαύρες (M) μπάλες σε αναλογία που φαίνεται στον παρακάτω πίνακα:

Μπάλες Κουτί	K	M
K_1	6	4
K_2	12	4
K_3	14	6

Επιλέγουμε στην τύχη ένα κουτί και στη συνέχεια μια μπάλα από αυτό. Να υπολογισθεί η πιθανότητα να επιλέξουμε μαύρη μπάλα.

K_i το ενδεχόμενο η μπάλα να προέρχεται από το κουτί i , $i = 1, 2, 3$.

M το ενδεχόμενο η μπάλα να είναι μαύρη.

Σύμφωνα με το θεώρημα της Ολικής Πιθανότητας έχουμε ότι:

$$P(M) = P(K_1)P(M|K_1) + P(K_2)P(M|K_2) + P(K_3)P(M|K_3)$$

$$= \frac{1}{3} * \frac{4}{10} + \frac{1}{3} * \frac{4}{16} + \frac{1}{3} * \frac{6}{20} = \frac{57}{180}$$

Θεώρημα Bayes

Έστω $A_1, A_2, \dots, A_n$ μία διαμέριση του δειγματικού χώρου S τέτοια ώστε $P(A_i) \neq 0$, $i = 1, 2, \dots, n$. Τότε για κάθε ενδεχόμενο E με $P(E) > 0$ έχουμε ότι,

$$P(A_k|E) = \frac{P(A_k)P(E|A_k)}{\sum_{i=1}^n P(A_i)P(E|A_i)} = \frac{P(A_k)P(E|A_k)}{P(E)}$$

Στις περιπτώσεις που είδαμε μέχρι τώρα οι πιθανότητες $P(A_i)$ των στοιχείων μιας διαμέρισης ήταν γνωστές και θέλαμε να υπολογίσουμε τις δεσμευμένες πιθανότητες ενός ενδεχομένου σε σχέση με τα στοιχεία της διαμέρισης $[P(E|A_k)]$.

Υπάρχουν όμως και περιπτώσεις που οι πιθανότητες των στοιχείων της διαμέρισης δεν είναι γνωστές. Τις περισσότερες φορές **πριν αρχίσουμε τα πειράματα** έχουμε ήδη μια **υποκειμενική αντίληψη /ένδειξη** για αυτές. Τις ενδείξεις αυτές ονομάζουμε εκ των **προτέρων πιθανότητες** των στοιχείων της διαμέρισης. **Μετά την εκτέλεση** όμως των πειραμάτων ίσως αποφασίσουμε να **αναθεωρήσουμε τις αρχικές εκτιμήσεις μας**. Αυτό μπορεί να γίνει χρησιμοποιώντας το θεώρημα του Bayes για να υπολογίσουμε τις **εκ των υστέρων πιθανότητες**

Παράδειγμα 8

Σε μία βιομηχανία που παράγει πλαστικά κύπελλα για αυτόματους πωλητές αναψυκτικών η παραγωγή των κυπέλλων γίνεται με τρεις μηχανές καθεμία από τις οποίες καλύπτει το $\frac{1}{3}$ της παραγωγής. Η συνολική παραγωγή και των τριών μηχανών σε μία ώρα είναι 9000 κύπελλα. Αν επιλέξουμε τυχαία ένα κύπελλο από αυτά και το βρούμε ελαττωματικό να υπολογιστεί η πιθανότητα ότι έχει κατασκευαστεί από την 1η, 2η ή 3η μηχανή.

ΛΥΣΗ

- E: το ενδεχόμενο ότι το προϊόν που επιλέγουμε είναι ελαττωματικό.
- A_1 : το ενδεχόμενο ότι το προϊόν που επιλέγουμε κατασκευάστηκε από την 1η μηχανή.
- A_2 : το ενδεχόμενο ότι το προϊόν που επιλέγουμε κατασκευάστηκε από την 2η μηχανή.
- A_3 : το ενδεχόμενο ότι το προϊόν που επιλέγουμε κατασκευάστηκε από την 3η μηχανή.

Παράδειγμα (συν.)

Θεωρούμε ότι η πιθανότητα που έχει το ελαττωματικό προϊόν που επιλέγουμε να προέρχεται από την 1η, 2η, 3η μηχανή είναι ίση με την αντίστοιχη πιθανότητα παραγωγής του από τις μηχανές αυτές

$$P(A_1) = P(A_2) = P(A_3) = 1/3$$

(Οι πιθανότητες αυτές ονομάζονται **εκ των προτέρων πιθανότητες**.)

Ο Διευθυντής Παραγωγής του εργοστασίου έχει κάποια υπόνοια ότι τα ποσοστά των ελαττωματικών διαφέρουν από μηχανή σε μηχανή. Για το λόγο αυτό παίρνει δείγμα 1000 κυπέλλων από την παραγωγή κάθε μηχανής, μετρά τον αριθμό των ελαττωματικών σε κάθε δείγμα και βρίσκει ότι :

- Το δείγμα προερχόμενο από την 1η μηχανή περιέχει 10 ελαττωματικά.
 - Το δείγμα προερχόμενο από την 2η μηχανή περιέχει 12 ελαττωματικά.
 - Το δείγμα προερχόμενο από την 3η μηχανή περιέχει 20 ελαττωματικά.
-
- $P(E | A_1) = 0.010$
 - $P(E | A_2) = 0.012$
 - $P(E | A_3) = 0.020$

Παράδειγμα (συν.)

Ο Διευθυντής Παραγωγής θέλει να αναθεωρήσει τις εκ των προτέρων πιθανότητες. Θα χρησιμοποιήσει λοιπόν το Θεώρημα Bayes για να υπολογίσει τις εκ των υστέρων πιθανότητες $P(A_1|E)$, $P(A_2|E)$ και $P(A_3|E)$.

$$P(A_1|E) = \frac{P(A_1)P(E|A_1)}{P(A_1)P(E|A_1) + P(A_2)P(E|A_2) + P(A_3)P(E|A_3)} = \frac{(1/3)(0,010)}{(1/3)(0,010) + (1/3)(0,012) + (1/3)(0,020)}$$

$$P(A_2|E) = \frac{P(A_2)P(E|A_2)}{P(A_1)P(E|A_1) + P(A_2)P(E|A_2) + P(A_3)P(E|A_3)} = \frac{(1/3)(0,012)}{(1/3)(0,010) + (1/3)(0,012) + (1/3)(0,020)}$$

$$P(A_3|E) = \frac{P(A_3)P(E|A_3)}{P(A_1)P(E|A_1) + P(A_2)P(E|A_2) + P(A_3)P(E|A_3)} = \frac{(1/3)(0,020)}{(1/3)(0,010) + (1/3)(0,012) + (1/3)(0,020)}$$

Ενδεχόμενα	Εκ των προτέρων πιθανότητες	Εκ των υστέρων πιθανότητες
A_1	0,33	0,24
A_2	0,33	0,28
A_3	0,33	0,48
Σύνολα	1,00	1,00

Κατά συνέπεια η πιθανότητα κατασκευής από τη 3^η μηχανή του ελαττωματικού κυπέλλου που έχει επιλεγεί αυξάνεται σε βάρος των πιθανοτήτων των άλλων δύο ενδεχομένων

Ανεξάρτητα Ενδεχόμενα

1. Αν $P(A/B) \neq P(A)$ συμπεραίνουμε ότι η πληροφορία ότι το B έχει συμβεί έχει ως αποτέλεσμα να μεταβληθεί η πιθανότητα να συμβεί το A.
2. Αντίθετα, αν $P(A|B) = P(A)$ η πληροφορία ότι το B έχει συμβεί δεν έχει καμία επίδραση στην πιθανότητα να συμβεί το A. Στις περιπτώσεις αυτές λέμε ότι το A είναι ανεξάρτητο του B.
3. Αντίστοιχα, αν $P(A|B) = P(B)$ τότε λέμε ότι το B είναι ανεξάρτητο του A.

Αποδεικνύεται ότι αν το A είναι ανεξάρτητο του B τότε και το B είναι ανεξάρτητο του A και αντίστροφα με την προϋπόθεση ότι $P(A) > 0$, $P(B) > 0$.

Ένας γενικότερος ορισμός της ανεξαρτησίας δύο ενδεχομένων είναι ο ακόλουθος :
Δύο ενδεχόμενα A και B λέγονται **ανεξάρτητα** (independent) ή **στατιστικά ανεξάρτητα** αν

$$P(A \cap B) = P(A) P(B)$$

Παρατηρήσεις

1. Οι συνθήκες $P(A) > 0$, $P(B) > 0$ δεν είναι απαραίτητες.
2. Κάθε ενδεχόμενο είναι ανεξάρτητο από το βέβαιο ενδεχόμενο (S) και από το αδύνατο ενδεχόμενο ($\emptyset$).

Ανεξάρτητα ή Ασυμβίβαστα??

1. Αν δύο ενδεχόμενα A και B είναι ασυμβίβαστα, δηλαδή $A \cap B = \emptyset$ με $P(A) > 0$ και $P(B) > 0$, τότε αυτά δεν μπορεί να είναι ανεξάρτητα. Με άλλα λόγια οι έννοιες ανεξαρτησία και ασυμβίβαστα είναι εκ διαμέτρου αντίθετες.
2. Αν τα ενδεχόμενα A και B είναι ανεξάρτητα τότε τα ζεύγη (A, B') (A', B') , (A', B) είναι ζεύγη ανεξάρτητων ενδεχομένων, όπου A' και B' είναι τα συμπληρωματικά ενδεχόμενα των A και B αντίστοιχα

Τυχαία Μεταβλητή

- Ο τρόπος με τον οποίο εκφράζονται τα στοιχειώδη ενδεχόμενα του δειγματικού χώρου μιας ενέργειας εξαρτάται από τη φύση των ενδεχομένων αυτών. Έτσι οι δειγματικοί χώροι κατατάσσονται σε δύο κατηγορίες:
 - α. Ποσοτικοί ή αριθμητικοί
 - β. Ποιοτικοί ή Περιγραφικοί
- Είναι αναγκαίο να ορισθεί **ένας κανόνας** (ή μία συνάρτηση) **που να αντιστοιχεί έναν πραγματικό αριθμό σε κάθε στοιχείο ενός δειγματικού χώρου S**. Η συνάρτηση αυτή ονομάζεται **τυχαία μεταβλητή** και ορίζεται ως εξής:

Έστω ο δειγματικός χώρος S. Ορίζουμε ως τυχαία μεταβλητή μια μονοσήμαντη συνάρτηση με πεδίο ορισμού το S και τιμές στην ευθεία των πραγματικών αριθμών R:

$$X: S \rightarrow R$$

Τυχαία Μεταβλητή

Παρατηρήσεις

- Για κάθε τιμή x της X πρέπει να καθορίζεται η πιθανότητα $P(X = x)$ η οποία θα πρέπει να ικανοποιεί τα βασικά αξιώματα του Kolmogorov. (ορισμός πιθανότητας)
- Σε κάθε στοιχειώδες ενδεχόμενο, s , του δειγματικού χώρου S ($s \in S$) αντιστοιχεί ακριβώς μία τιμή της τυχαίας μεταβλητής ή $x = X(s)$. Η ίδια τιμή x της X μπορεί να αντιστοιχεί σε διαφορετικά στοιχειώδη ενδεχόμενα του S . (ορισμός συνάρτησης)

Διακριτές Τυχαίες Μεταβλητές

Έστω η τυχαία μεταβλητή $X : S \rightarrow R$. Αν το πεδίο τιμών $X(S)$ της είναι πεπερασμένο ή απείρως αριθμήσιμο τότε η μεταβλητή αυτή καλείται διακριτή ή απαριθμητή τυχαία μεταβλητή.

Δηλαδή σε μια διακριτή τυχαία μεταβλητή X μπορούμε να αριθμήσουμε τις δυνατές τιμές της.

Έστω X μια διακριτή τυχαία μεταβλητή και $X(S)$ το πεδίο τιμών της. Μία συνάρτηση: $\mathbf{P(x)} = p_x = P_x(x) = P(X = x)$ με πεδίο ορισμού τις τιμές της X και πεδίο τιμών τις πιθανότητες των τιμών αυτών η οποία ικανοποιεί τις συνθήκες:

$$(1) P_x(x) \geq 0 \quad \forall x \in X(S) \subseteq R$$

$$(1) \text{ και } \sum_{x \in X(S)} P_x(x) = 1$$

καλείται **συνάρτηση πιθανότητας** της τυχαίας μεταβλητής X .

Το σύνολο των ζευγών $\{(x, P_x(x)) \mid x \in X(S)\}$ καλείται **κατανομή πιθανότητας** της διακριτής τυχαίας μεταβλητής X .

Παράδειγμα 1

Έστω μία οικογένεια με 3 παιδιά. Αν συμβολίσουμε με X τον αριθμό των αγοριών να καθοριστεί η κατανομή πιθανότητας της τυχαίας μεταβλητής X και να παρασταθεί γραφικά.

Ο αριθμός των αγοριών στην οικογένεια εξαρτάται από το αποτέλεσμα του καθενός από τους τρεις τοκετούς.

Είναι λοιπόν δυνατό να εμφανιστούν κανένα, ένα, δύο ή τρία αγόρια. Κατά συνέπεια η μεταβλητή X η οποία εκφράζει τον αριθμό των αγοριών είναι μία διακριτή τυχαία μεταβλητή η οποία μπορεί να πάρει τις τιμές :

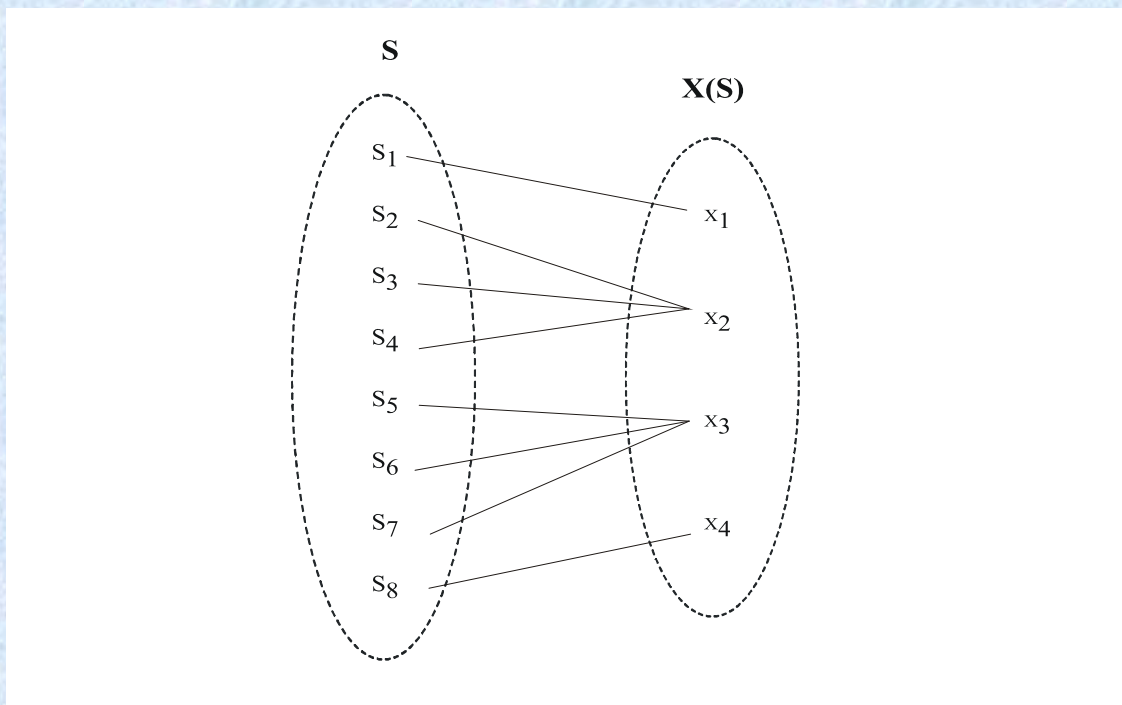
$$x_1 = 0, x_2 = 1, x_3 = 2, x_4 = 3.$$

Στοιχειώδη Ενδεχόμενα Δειγμ. Χώρου S	Αριθμός Αγοριών X	Πιθανότητα
s_1 : KKK	0	1/8
s_2 : AKK	1	1/8
s_3 : KAK	1	1/8
s_4 : KKA	1	1/8
s_5 : AAK	2	1/8
s_6 : AKA	2	1/8
s_7 : KAA	2	1/8
s_8 : AAA	3	1/8

Παράδειγμα 1

Όπως φαίνεται από τον Πίνακα σε κάθε στοιχειώδες ενδεχόμενο s_i του δειγματικού χώρου αντιστοιχεί ακριβώς μια τιμή της τυχαίας μεταβλητής X και επιπλέον η ίδια τιμή μπορεί να αντιστοιχεί σε διαφορετικά στοιχειώδη ενδεχόμενα του S .

Η αντιστοιχία μεταξύ των στοιχειωδών ενδεχομένων και των τιμών της τυχαίας μεταβλητής X απεικονίζεται γραφικά



Παράδειγμα

- Αφού προσδιορίσουμε τις δυνατές τιμές της τυχαίας μεταβλητής X μπορούμε τώρα να ορίσουμε τη συνάρτηση πιθανότητας της:

$$P(x_i) = P(X = x_i).$$

- Με τη συνάρτηση αυτή σε κάθε στοιχείο x του πεδίου τιμών $X(S)$ της μεταβλητής X αντιστοιχούμε έναν πραγματικό αριθμό $P(x)$ που είναι ακριβώς η πιθανότητα εμφάνισης της τιμής αυτής.

$$P(0) = P(X = 0) = \frac{1}{8}$$

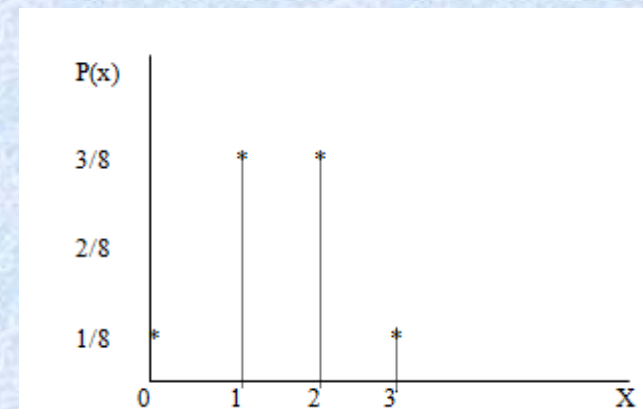
$$P(1) = P(X = 1) = \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{3}{8}$$

$$P(2) = P(X = 2) = \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{3}{8}$$

$$P(3) = P(X = 3) = \frac{1}{8}$$

$$\sum_{i=1}^3 P(x_i) = P(x_1) + P(x_2) + P(x_3) + P(x_4) = \frac{1}{8} + \frac{3}{8} + \frac{3}{8} + \frac{1}{8} = 1$$

Δυνατές τιμές της τυχαίας μεταβλητής X x_i	Πιθανότητες εμφάνισης τους $P(X=x_i)$
0	1/8
1	3/8
2	3/8
3	1/8



Συνεχείς Τυχαίες Μεταβλητές

Έστω X τυχαία μεταβλητή. Αν το πεδίο τιμών $X(S)$ της X είναι **μη αριθμήσιμο** σύνολο τότε η X ονομάζεται **συνεχής τυχαία μεταβλητή**

Έστω X μία συνεχής τυχαία μεταβλητή και $X(S)$ το πεδίο τιμών της. Μία συνάρτηση **$f(x)$** με **πεδίο ορισμού το σύνολο $X(S)$** και **πεδίο τιμών** ένα υποσύνολο του συνόλου των **πραγματικών αριθμών** η οποία ικανοποιεί τις συνθήκες :

$$(1) f(x) \geq 0 \quad \forall x \in X(S) \text{ όπου } -\infty < x < +\infty$$

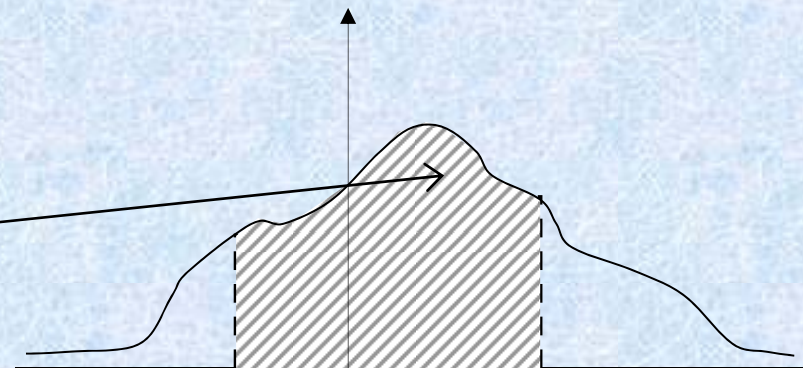
$$(2) \text{ και } \int_{-\infty}^{+\infty} f(x) dx = 1$$

ονομάζεται **συνάρτηση πυκνότητας πιθανότητας** ή απλά **πυκνότητα πιθανότητας**

Για οποιοδήποτε ζεύγος πραγματικών αριθμών a και β όπου:

$$-\infty < a \leq x \leq \beta < +\infty, \text{ ισχύει:}$$

$$P(a \leq x \leq \beta) = \int_a^\beta f(x) dx$$



Αθροιστική Συνάρτηση Κατανομής Πιθανότητας Τυχαίας μεταβλητής

- Έστω X τυχαία μεταβλητή (διακριτή ή συνεχής).

Η συνάρτηση

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

η οποία εκφράζει την πιθανότητα να λάβει η τυχαία μεταβλητή X οποιαδήποτε τιμή μικρότερη από ή ίση με x καλείται **αθροιστική συνάρτηση κατανομής πιθανότητας** ή απλά συνάρτηση κατανομής πιθανότητας της τυχαίας μεταβλητής X .

- Ιδιότητες της $F(x)$

- $0 \leq F(x) \leq 1$
- $\text{An } a \leq \beta \Rightarrow F(a) \leq F(\beta)$
- $\lim_{x \rightarrow \infty} F(x) = 1 \quad \lim_{x \rightarrow -\infty} F(x) = 0$

- Η πιθανότητα να λάβει η τυχαία μεταβλητή X τιμή μεταξύ a και β δίνεται από τη σχέση:

$$P(a \leq x \leq \beta) = F(\beta) - F(a).$$

Διακριτή τυχαία μεταβλητή

- Στην περίπτωση διακριτής τυχαίας μεταβλητής X

$$P(X \leq x) = \sum_{x_i \leq x} P(x_i)$$

- δηλαδή η πιθανότητα να λάβει η X τιμές μικρότερες ή ίσες του x ισούται με το άθροισμα των πιθανοτήτων $P(x_i)$ των τιμών x_i τις οποίες μπορεί να πάρει η X και για τις οποίες ισχύει $x_i \leq x$
- Η αθροιστική συνάρτηση κατανομής δίνεται από τον τύπο:

$$F(x) = \sum_{x_i \leq x} P(x_i), \quad -\infty < x < +\infty$$

- Από τον τύπο αυτόν προκύπτει ότι αν x_k είναι μία τιμή της x τότε,

$$P(x_k) = P(X = x_k) = F(x_k) - F(x_{k-1})$$

Διακριτή τυχαία μεταβλητή

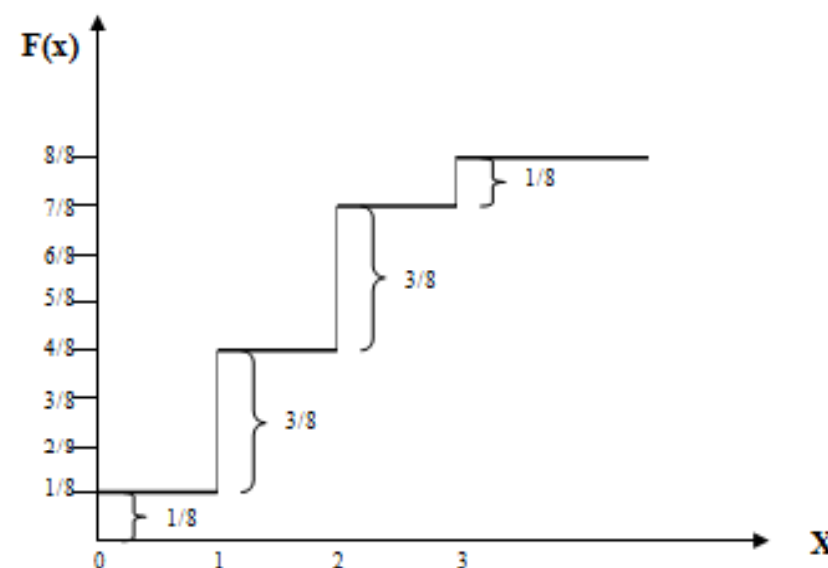
Θέλουμε να υπολογίσουμε την αθροιστική της συνάρτηση κατανομής.

$$F(0) = \sum_{x_i \in 0} P(x_i) = P(0) = \frac{1}{8}$$

$$F(1) = \sum_{x_i \in 1} P(x_i) = P(0) + P(1) = \frac{1}{8} + \frac{3}{8} = \frac{4}{8}$$

$$F(2) = \sum_{x_i \in 2} P(x_i) = P(0) + P(1) + P(2) = \frac{1}{8} + \frac{3}{8} + \frac{3}{8} = \frac{7}{8}$$

$$F(3) = \sum_{x_i \in 3} P(x_i) = P(0) + P(1) + P(2) + P(3) = \frac{1}{8} + \frac{3}{8} + \frac{3}{8} + \frac{1}{8} = 1$$

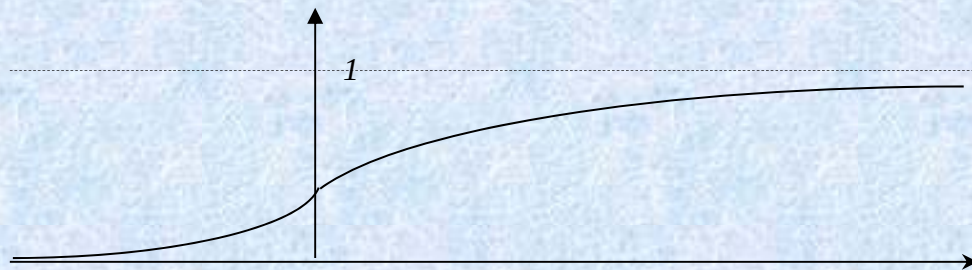


Συνεχής τυχαία μεταβλητή

- Στην περίπτωση συνεχούς τυχαίας μεταβλητής X έχουμε ότι:

$$P(X \leq x) = \int_{-\infty}^x f(t)dx \quad F(x) = \int_{-\infty}^x f(t)dx$$

- Προφανώς η μορφή της καμπύλης της αθροιστικής συνάρτησης κατανομής $F(x)$ εξαρτάται από τη μορφή της καμπύλης της συνάρτησης πυκνότητας πιθανότητας $f(x)$.



Αν παραγωγίσουμε την $F(x)$ ως προς x έχουμε ότι:

$$\frac{\partial F(x)}{\partial x} = f(x)$$

Μέτρα κεντρικής τάσης

Μέση τιμή

Διακριτή τ.μ.

X διακριτή τυχαία μεταβλητή με συνάρτηση πιθανότητας $P(x)$

Η μέση τιμή της δίνεται από

$$E(x) = \sum_{i=1}^{\infty} x_i P(x_i)$$

με την προϋπόθεση φυσικά ότι το άθροισμα αυτό συγκλίνει.

Αν η X παίρνει πεπερασμένο πλήθος τιμών

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

Συνεχής τ.μ.

X συνεχής τυχαία μεταβλητή με συνάρτηση πυκνότητας πιθανότητας $f(x)$
. Η μέση τιμή της δίνεται

$$E(x) = \int_{-\infty}^{+\infty} xf(x)dx$$

με την προϋπόθεση φυσικά ότι το ολοκλήρωμα συγκλίνει.

Αν η X παίρνει τιμές στο διάστημα $[a, b]$
τότε έχουμε

$$E(x) = \int_a^b xf(x)dx$$

Μέτρα κεντρικής τάσης

Ιδιότητες Μέσης τιμής

Για τη μέση τιμή μιας τυχαίας μεταβλητής, συνεχούς ή διακριτής, αποδεικνύεται ότι

- $E(a) = a$, όπου a είναι σταθερή ποσότητα
- $E(aX) = a E(X)$, όπου a είναι σταθερή ποσότητα.
- $E(aX + b) = a E(X) + b$, όπου a, b είναι σταθερές ποσότητες
- $E[X - E(X)] = 0$
- $E(aX + bY) = aE(X) + bE(Y)$

Μέτρα Κεντρική Τάσης

Διάμεσος

Διάμεσος μιας τ.μ. X είναι εκείνη η τιμή που χωρίζει την κατανομή σε δύο ίσα μέρη.

Διακριτή τ.μ.

X διακριτή τυχαία μεταβλητή με συνάρτηση πιθανότητας $P(x)$.

Η διάμεσος M είναι εκείνη η τιμή για την οποία έχουμε:

$$P(X \leq M) \geq 1/2 \quad \text{και}$$

$$P(X \geq M) \geq 1/2$$

Συνεχής τ.μ.

X συνεχής τυχαία μεταβλητή με συνάρτηση πυκνότητας πιθανότητας $f(x)$.

Η διάμεσος M είναι εκείνη η τιμή η οποία αποτελεί λύση της εξίσωσης:

$$\int_{-\infty}^M f(x)dx = 1/2$$

Ο καλύτερος τρόπος υπολογισμού της διαμέσου είναι με τη χρήση της αθροιστικής συνάρτησης κατανομής $F(x)$.

Η διάμεσος είναι η μικρότερη τιμή που δίνει $F(x) \geq 1/2$

Η διάμεσος είναι η τιμή εκείνη που δίνει $F(x) = 1/2$.

Μέτρα Κεντρικής Τάσης

Επικρατούσα Τιμή

Επικρατούσα Τιμή μιας τ.μ. X είναι εκείνη η τιμή που έχει τη μεγαλύτερη πιθανότητα εμφάνισης από κάθε άλλη τιμή.

Διακριτή τ.μ.

X διακριτή τυχαία μεταβλητή με συνάρτηση πιθανότητας $P(x)$.

Η επικρατούσα τιμή είναι εκείνη η τιμή T_0 για την οποία ισχύει:

$$P(X = T_0) \geq P(X = x), \quad \forall x$$

Συνεχής τ.μ.

X συνεχής τυχαία μεταβλητή με συνάρτηση πυκνότητας πιθανότητας $f(x)$.

Η επικρατούσα τιμή είναι εκείνη η τιμή T_0 για την οποία ισχύει:

$$f(T_0) \geq f(x), \quad \forall x$$

Μέτρα Σχετικής Θέσης

p-ποσοστιαίο σημείο

p - ποσοστιαίο σημείο μιας τυχαίας μεταβλητής είναι εκείνη η τιμή (X_p) για την οποία ισχύει:

Διακριτή τ.μ.

$$P(X \leq X_p) \geq p \quad \text{και}$$

$$P(X \geq X_p) \geq 1 - p$$

Συνεχής τ.μ.

$$F(X_p) = p$$

- αν $p = 0,5$ το X_p είναι διάμεσος (M)
- αν $p = 0,25$ το X_p είναι το 1ο τεταρτημόριο (Q_1)
- αν $p = 0,75$ το X_p είναι το 3ο τεταρτημόριο (Q_3).

Μέτρα Διασποράς Διακύμανση

Χρειαζόμαστε λοιπόν ένα μέτρο μεταβλητότητας ή διασποράς της τ.μ.

Έστω X τ.μ. (διακριτή ή συνεχή) με μέση τιμή $\mu = E(X)$. Η διακύμανση της X συμβολίζεται με $V(X)$ ή σ^2 και δίνεται από τον τύπο:

$$V(X) = \sigma^2 = E[X - E(X)]^2 = E(X - \mu)^2$$

$$V(X) = \sigma^2 = E(X^2) - [E(X)]^2 = E(X^2) - \mu^2 = \sum_{i=1}^n x_i^2 P(x_i) - \mu^2$$

Ιδιότητες

- $V(a) = 0$, όπου a σταθερή ποσότητα.
- $V(x + a) = V(x)$, όπου a σταθερή ποσότητα.
- $V(ax) = a^2 V(x)$, όπου a σταθερή ποσότητα.
- $V(ax + \beta) = a^2 V(x)$, όπου a, β σταθερές ποσότητες.

Μέτρα Διασποράς Ενδοτεταρτομοριακό Εύρος

Το ενδοτεταρτημοριακό εύρος ορίζεται ως $Q_3 - Q_1$ και η τεταρτημοριακή απόκλιση ως

$$Q = \frac{Q_3 - Q_1}{2}.$$

Το Q_1 (1^ο τεταρτημόριο ή 25^ο εκατοστιαίο σημείο) και το Q_3 (3^ο τεταρτημόριο ή 75^ο εκατοστιαίο σημείο) υπολογίζονται με βάση τη γενική μεθοδολογία που δόθηκε σε προηγούμενη ενότητα για το p – ποσοστιαίο σημείο.