Αβεβαιότητα

Δράση υπό αβεβαιότητα

- Οι αλγόριθμοι σχεδόν ποτέ δεν έχουν πρόσβαση στην πλήρη αλήθεια σχετικά με το περιβάλλον τους
- Κανένα πλάνο δεν εγγυάται ότι θα πετύχει το στόχο
  - Πλάνο A_{90}: Ξεκίνα 90 λεπτά νωρίτερα για να προλάβεις την πτήση
    - Το πλάνο A_{90} θα μας πάει στο αεροδρόμιο εγκαίρως, εφόσον το αυτοκίνητό μου δεν πάθει βλάβη ή μείνει από βενζίνη, δεν έχουμε εμείς ατύχημα, δεν υπάρξουν άλλα ατυχήματα σε κάποια από τις σήραγγες, το αεροπλάνο δε φύγει νωρίτερα, και …"

Δράση υπό αβεβαιότητα

- Η ενδεδειγμένη δράση – η **ορθολογική απόφαση** – *εξαρτάται τόσο από τη σχετική σημαντικότητα των διάφορων στόχων, όσο και από την πιθανότητα και το βαθμό κατά τον οποίο θα επιτευχθούν οι στόχοι αυτοί*

Χειρισμός της αβέβαιης γνώσης

- Η διάγνωση – ανεξάρτητα από το αν αφορά την ιατρική, την επισκευή αυτοκινήτων, ή οτιδήποτε άλλο – είναι μια διεργασία που περιλαμβάνει σχεδόν πάντοτε αβεβαιότητα
- Διαγνωστικοί κανόνες (έχουμε το σύμπτωμα, ψάχνουμε την αιτία)
  - $\forall p$ *Σύμπτωμα(p, Πονόδοντος)* $\Rightarrow$ *Ασθένεια(p, Κοιλότητα)*
    - Λανθασμένος κανόνας
  - $\forall p$ *Σύμπτωμα(p, Πονόδοντος)* $\Rightarrow$
    Ασθένεια(p, Κοιλότητα) $\vee$ *Ασθένεια(p, Ουλίτιδα)*
    $\vee$ *Ασθένεια(p, Απόστημα)* …

Αιτιολογικοί κανόνες (έχουμε την αιτία, ψάχνουμε το σύμπτωμα)

  - $\forall p$ *Ασθένεια(p, Κοιλότητα)* $\Rightarrow$ *Σύμπτωμα(p, Πονόδοντος)*
    - Επίσης λανθασμένος κανόνας

Χειρισμός της αβέβαιης γνώσης

- Η λογική πρώτης τάξης αποτυγχάνει για τρεις κύριους λόγους:
  - Τεμπελιά
    - Απαιτείται υπερβολικά πολλή δουλειά για την παράθεση σε λίστα του πλήρους συνόλου προαπαιτήσεων ή συνεπειών που απαιτούνται για να εξασφαλίσουμε έναν κανόνα χωρίς εξαιρέσεις, και επίσης είναι πολύ δύσκολη η χρήση τέτοιων κανόνων
  - Θεωρητική άγνοια
    - Δεν υπάρχει πλήρης θεωρία για το αντίστοιχο πεδίο
  - Πρακτική άγνοια
    - Ακόμα και αν γνωρίζουμε όλους τους κανόνες, μπορεί να βρεθούμε σε αβεβαιότητα για κάποιο συγκεκριμένο περιστατικό, επειδή δεν έχουν εκτελεστεί ή δεν μπορούν να εκτελεστούν όλοι οι απαιτούμενοι έλεγχοι

Χειρισμός της αβέβαιης γνώσης

- Βαθμός πεποίθησης: [0,1]
- Θεωρία πιθανοτήτων
  - *Η πιθανότητα μας παρέχει έναν τρόπο για να* **συναθροίσουμε** *την αβεβαιότητα που προέρχεται από την τεμπελιά ή την άγνοιά μας.*
- Εκ των προτέρων πιθανότητες
  - Πριν τη λήψη μιας μαρτυρίας
- Εκ των υστέρων πιθανότητες
  - Μετά τη λήψη μιας μαρτυρίας

Ορθολογικές αποφάσεις

- Αποτελέσματα
  - Πλήρως καθορισμένες καταστάσεις
  - Υπάρχουν διάφορα δυνατά αποτελέσματα
- Προτιμήσεις / Πιθανότητες
- Θεωρία Χρησιμοτήτων
  - Κάθε κατάσταση έχει ένα βαθμό ωφέλειας
- **Θεωρία Αποφάσεων =**

  Θεωρία Πιθανοτήτων + Θεωρία Χρησιμοτήτων

- Αρχή της **Μέγιστης Αναμενόμενης Χρησιμότητας**
  - Ένας πράκτορας είναι ορθολογικός αν και μόνο αν επιλέγει την ενέργεια που παρέχει την υψηλότερη αναμενόμενη χρησιμότητα, λαμβανόμενη ως το μέσο όρο για όλα τα δυνατά αποτελέσματα της ενέργειας

Βασική Σημειογραφία Πιθανοτήτων

Προτάσεις

- Τυχαία μεταβλητή
  - Μια μεταβλητή που αναφέρεται σε ένα «τμήμα» του κόσμου για το οποίο η «κατάσταση» είναι αρχικά άγνωστη
  - *Κοιλότητα*: Ο κάτω αριστερός φρονιμίτης μου έχει κοιλότητα
  - Θα γράφουμε πάντα τα ονόματα των τυχαίων μεταβλητών με κεφαλαίο το πρώτο γράμμα.
  - Παρόλα αυτά, θα χρησιμοποιούμε και πάλι πεζά γράμματα για τις τιμές που έχουν οι άγνωστες τυχαίες μεταβλητές:
    - $P(a) = 1 - P(\neg a)$
- Πεδίο τιμών
  - π.χ. (*αληθές*, *ψευδές*)
- Δυϊκές, Διακριτές, Συνεχείς

Προτάσεις

- Για τη δυϊκή τυχαία μεταβλητή Κοιλότητα με πεδίο τιμών το (*αληθές*, *ψευδές*):
  - Η πρόταση *Κοιλότητα* = *αληθές* θα γράφεται και ως *κοιλότητα*
  - Η πρόταση *Κοιλότητα* = *ψευδές* θα γράφεται και ως ¬*κοιλότητα*

- Για την διακριτή τυχαία μεταβλητή Καιρός με πεδίο τιμών το (*λιακάδα*, *βροχή*, *συννεφιά*, *χιόνι*):
  - Η πρόταση *Καιρός* = *χιόνι* θα γράφεται και ως *χιόνι*

Ατομικά Συμβάντα

- Το ατομικό συμβάν είναι ένας *πλήρης καθορισμός* της κατάστασης του κόσμου για τον οποίο ο πράκτορας είναι αβέβαιος

  - Για παράδειγμα, αν ο κόσμος αποτελείται μόνο από τις Boolean μεταβλητές *Κοιλότητα* και *Πονόδοντος*, τότε υπάρχουν μόνο τέσσερα διακριτά ατομικά συμβάντα
    - π.χ. η πρόταση

$$\textit{Κοιλότητα} = \textit{αληθές} \wedge \textit{Πονόδοντος} = \textit{ψευδές}$$

Ιδιότητες ατομικών συμβάντων

1. Είναι *αμοιβαία αποκλειόμενα*
   - Για παράδειγμα δεν μπορούν να ισχύουν ταυτόχρονα τα συμβάντα *κοιλότητα* ∧ *πονόδοντος* και *κοιλότητα* ∧ ¬*πονόδοντος*
2. Το σύνολό τους είναι *εξαντλητικό*
   - Η διάζευξη όλων των ατομικών συμβάντων είναι αληθής
3. Οποιοδήποτε ατομικό συμβάν καλύπτει την αλήθεια ή το ψεύδος για κάθε πρόταση, είτε αυτή είναι απλή είτε σύνθετη
   - Για παράδειγμα, το ατομικό συμβάν *κοιλότητα* ∧ ¬*πονόδοντος* καλύπτει την αλήθεια της πρότασης *κοιλότητα* και το ψεύδος της *κοιλότητα*⇒*πονόδοντος*
4. Οποιαδήποτε πρόταση είναι λογικά ισοδύναμη με τη διάζευξη όλων των ατομικών συμβάντων που συμπεραίνουν την αλήθεια αυτής της πρότασης
   - Για παράδειγμα, το *κοιλότητα* είναι ισοδύναμο με το *κοιλότητα* ∧ *πονόδοντος* ∨ *κοιλότητα* ∧ ¬*πονόδοντος*

Εκ των προτέρων πιθανότητα

- $P(Κοιλότητα = αληθές) = 0.1$ ή $P(κοιλότητα) = 0.1$
- $P(Καιρός = λιακάδα) = 0.7$ ή $P(λιακάδα) = 0.7$
- $P(Καιρός = βροχή) = 0.2$ ή $P(βροχή) = 0.2$
- $P(Καιρός = συννεφιά) = 0.08$ ή $P(συννεφιά) = 0.08$
- $P(Καιρός = χιόνι) = 0.02$ ή $P(χιόνι) = 0.02$

- Κατανομή πιθανότητας
  - $\mathbf{P}(Καιρός) = (0.7, 0.2, 0.08, 0.02)$
- Συνδυασμένη κατανομή πιθανότητας
  - $\mathbf{P}(Καιρός, Κοιλότητα)$
- Πλήρης συνδυασμένη κατανομή πιθανότητας
  - Περιγράφει το πλήρες σύνολο των συνδυασμών των τυχαίων μεταβλητών που χρησιμοποιούνται για την περιγραφή του κόσμου

Πυκνότητα Πιθανότητας

- *Έστω*
  - X = Αυριανή θερμοκρασία
- *τότε π.χ.*
  - $P(X = x) = U[18, 26](x)$ (ομοιόμορφη κατανομή)
- $P(X=20{,}5) = U[18, 26](20{,}5)=1/(26\text{-}18) = 0{,}125$
- Υπολογίζουμε πιθανότητα για περιοχές τιμών

Υπό συνθήκη πιθανότητα

- $\pi.\chi.$
  - $P(\kappa o\iota\lambda \acute{o}\tau\eta\tau\alpha \mid \pi o\nu \acute{o}\delta o\nu\tau o\varsigma) = 0{,}8$

$$P(a \mid b) = \frac{P(a \wedge b)}{P(b)}$$

- Κανόνας γινομένου:
  - $P(a \wedge b) = P(a \mid b)P(b) = P(b \mid a)P(a)$

Συμπεράσματα Με Πλήρεις Συνδυασμένες Κατανομές Πιθανότητας

- Έστω ένα πεδίο που αποτελείται μόνο από τις τρεις Boolean μεταβλητές *Πονόδοντος*, *Κοιλότητα*, και *Λαβίδα*:

	πονόδοντος		*¬πονόδοντος*	
	λαβίδα	*¬λαβίδα*	*λαβίδα*	*¬λαβίδα*
κοιλότητα	0,108	0,012	0,072	0,008
¬κοιλότητα	0,016	0,064	0,144	0,576

- Μπορούμε να υπολογίσουμε την πιθανότητα χωρίς συνθήκη οποιασδήποτε πρότασης (*περιθώριος πιθανότητα*)
  - π.χ. P(*κοιλότητα* ∨ *πονόδοντος*) = 0,108+ 0,012 + 0,072+ 0,008 + 0,016 + 0,064 = 0,28

Θεώρημα Ολικής Πιθανότητας

- Περιθωριοποίηση ή απαλοιφή με άθροιση:

$$P(Y) = \sum_{z} P(Y, z)$$

- ή ισοδύναμα (Θεώρημα ολικής πιθανότητας):

$$P(Y) = \sum_{z} P(Y|z)P(z)$$

Υπολογισμός υπό-συνθήκη πιθανοτήτων

$$P(\kappa o\iota\lambda \acute{o}\tau\eta\tau\alpha \mid \pi o\nu \acute{o}\delta o\nu\tau o\varsigma) = \frac{P(\kappa o\iota\lambda \acute{o}\tau\eta\tau\alpha \wedge \pi o\nu \acute{o}\delta o\nu\tau o\varsigma)}{P(\pi o\nu \acute{o}\delta o\nu\tau o\varsigma)} = \frac{0.108+0.012}{0.108+0.012+0.016+0.064} = 0.6$$

$$P(\neg\kappa o\iota\lambda \acute{o}\tau\eta\tau\alpha \mid \pi o\nu \acute{o}\delta o\nu\tau o\varsigma) = \frac{P(\neg\kappa o\iota\lambda \acute{o}\tau\eta\tau\alpha \wedge \pi o\nu \acute{o}\delta o\nu\tau o\varsigma)}{P(\pi o\nu \acute{o}\delta o\nu\tau o\varsigma)} = \frac{0.016+0.064}{0.108+0.012+0.016+0.064} = 0.4$$

- Υπολογισμός με κανονικοποίηση:

P(*κοιλότητα* | Π*ονόδοντος*) = *a* **P**(*κοιλότητα*, Π*ονόδοντος*)
= *a* [**P**(*κοιλότητα*, Π*ονόδοντος*, *λαβίδα*) + **P**(*κοιλότητα*, Π*ονόδοντος*, ¬*λαβίδα*)]
= *a* [(0.108, 0.016) + (0.012, 0.064)] = *a* (0.12, 0.08) = (0.6, 0.4)

Ανεξαρτησία

Ανεξαρτησία

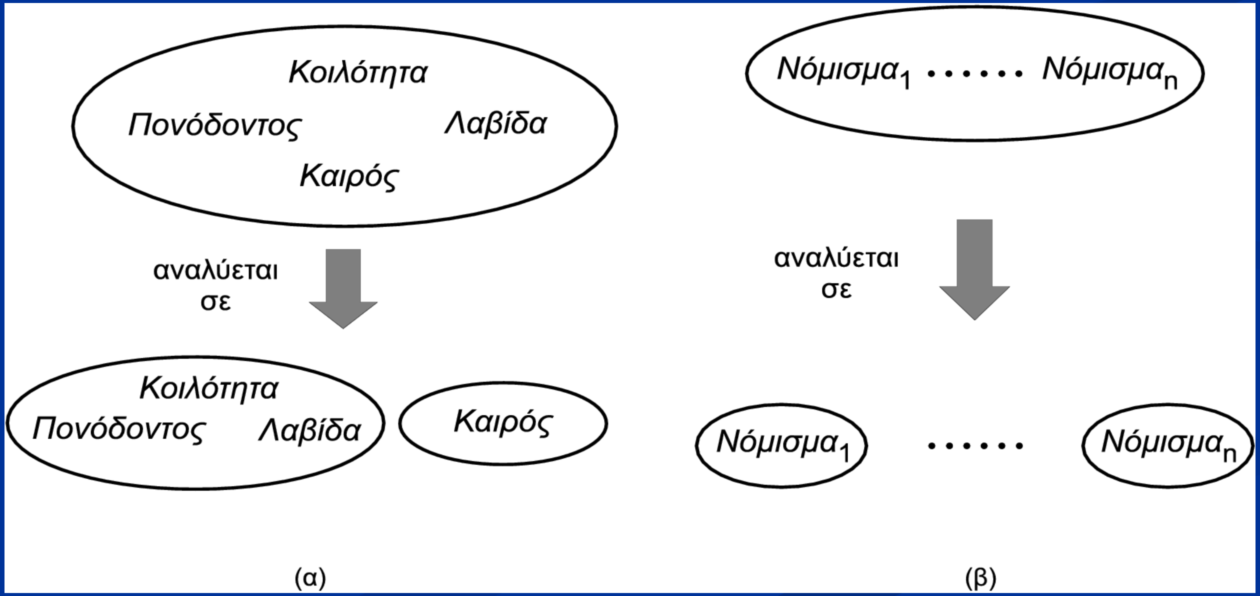
- Έστω ότι εμπλουτίζουμε το οδοντιατρικό μας πεδίο με την τυχαία μεταβλητή *Καιρός*
- Η πλήρης κατανομή γίνεται…
  - **P**(*Πονόδοντος*, *Λαβίδα*, *Κοιλότητα*, *Καιρός*)
- …η οποία έχει 32 καταχωρίσεις (επειδή η μεταβλητή *Καιρός* έχει τέσσερις τιμές)
- Προφανώς δε θεωρούμε ότι ο καιρός παίζει ρόλο στα οδοντιατρικά μας προβλήματα

Ανεξαρτησία

- Προσπαθούμε να υπολογίσουμε τις καταχωρήσεις της νέας πλήρους κατανομής:
  - $P(\textit{πονόδοντος}, \textit{λαβίδα}, \textit{κοιλότητα}, \textit{Καιρός} = \textit{συννεφιά}) = P(\textit{Καιρός} = \textit{συννεφιά} \mid \textit{πονόδοντος}, \textit{λαβίδα}, \textit{κοιλότητα}) \cdot P(\textit{πονόδοντος}, \textit{λαβίδα}, \textit{κοιλότητα})$
- Θεωρούμε ότι:
  - $P(\textit{Καιρός} = \textit{συννεφιά} \mid \textit{πονόδοντος}, \textit{λαβίδα}, \textit{κοιλότητα}) = P(\textit{Καιρός} = \textit{συννεφιά})$
- οπότε:
  - $P(\textit{πονόδοντος}, \textit{λαβίδα}, \textit{κοιλότητα}, \textit{Καιρός} = \textit{συννεφιά}) = P(\textit{Καιρός} = \textit{συννεφιά}) \cdot P(\textit{πονόδοντος}, \textit{λαβίδα}, \textit{κοιλότητα})$
- Γενικότερα:
  - $\mathbf{P}(\textit{Πονόδοντος}, \textit{Λαβίδα}, \textit{Κοιλότητα}, \textit{Καιρός}) = \mathbf{P}(\textit{Πονόδοντος}, \textit{Λαβίδα}, \textit{Κοιλότητα}) \cdot \mathbf{P}(\textit{Καιρός})$

Ανεξαρτησία

- Ισοδύναμες δηλώσεις, για δύο <u>ανεξάρτητες</u> μεταβλητές X και Y:
  - $\mathbf{P}(X|Y) = \mathbf{P}(X)$
  - $\mathbf{P}(Y|X) = P(Y)$
  - $\mathbf{P}(X, Y) = \mathbf{P}(X)\,\mathbf{P}(Y)$
- Η πλήρης κατανομή σπάει σε πολλές μικρότερες:



Υπό συνθήκη ανεξαρτησία

- Οι μεταβλητές X και Y είναι υπό συνθήκη ανεξάρτητες, με δεδομένη την τιμή της Z:
  - $\mathbf{P}(X, Y | Z) = \mathbf{P}(X | Z) \cdot \mathbf{P}(Y | Z)$
  - $\mathbf{P}(X | Y, Z) = \mathbf{P}(X | Z) \quad \textit{και} \quad \mathbf{P}(Y | X, Z) = \mathbf{P}(Y | Z)$

- Κατά συνέπεια:
  - $\mathbf{P}(\textit{Πονόδοντος}, \textit{Λαβίδα}, \textit{Κοιλότητα})$
    $= \mathbf{P}(\textit{Πονόδοντος}, \textit{Λαβίδα} | \textit{Κοιλότητα}) \cdot \mathbf{P}(\textit{Κοιλότητα})$
    $= \mathbf{P}(\textit{Πονόδοντος} | \textit{Κοιλότητα}) \cdot \mathbf{P}(\textit{Λαβίδα} | \textit{Κοιλότητα}) \cdot \mathbf{P}(\textit{Κοιλότητα})$

Απλοϊκό μοντέλο Bayes

- Προβλήματα στα οποία υπάρχει μια μεταβλητή **Αιτία** και πολλές μεταβλητές **Επίδρασης**, οι οποίες είναι υπό συνθήκη ανεξάρτητες μεταξύ τους με δεδομένη την **Αιτία**
  - Ονομάζεται απλοϊκό γιατί χρησιμοποιείται ως παραδοχή και όταν δεν πρέπει…

- $\mathbf{P}(\textit{Αιτία}, \textit{Επίδραση}_1, \ldots, \textit{Επίδραση}_n) = \mathbf{P}(\textit{Αιτία}) \cdot \prod_i (\textit{Επίδραση}_i \mid \textit{Αιτία})$

Στατιστική Μάθηση

Γενικά

- Υποθέσεις
  - Πιθανοτικές θεωρίες σχετικά με τον τρόπο λειτουργίας του πεδίου
- Δεδομένα
  - Στιγμιότυπα τιμών για κάποιες ή και για όλες τις τυχαίες μεταβλητές που περιγράφουν το πεδίο
- Παράδειγμα
  - Γλυκό Surprise, δύο γεύσεις: κεράσι (νόστιμο), λεμόνι (απαίσιο)
  - Πέντε πολύ μεγάλες σακούλες, ΙΔΙΕΣ!
    - *h*1: 100% κεράσι
    - *h*2: 75% κεράσι + 25% λεμόνι
    - *h*3: 50% κεράσι + 50% λεμόνι
    - *h*4: 25% κεράσι + 75% λεμόνι
    - *h*5: 100% λεμόνι
  - Τυχαία μεταβλητή H , τιμές: h_1, …, h_5 (υποθέσεις)
  - Δεδομένα: $D_1, D_2, \ldots D_N$

Μάθηση κατά Bayes

- Υπολογίζει την πιθανότητα κάθε υπόθεσης, με βάση τα δεδομένα:
  - $P(h_i \mid \mathbf{d}) = \alpha \cdot P(\mathbf{d} \mid h_i) P(h_i)$
    - εξαρτάται από την εκ των προτέρων πιθανότητα

- Πιθανοφάνεια δεδομένων: $P(\mathbf{d} \mid h_i)$

$$P(\mathbf{d} \mid h_i) = \prod_j P(d_j \mid h_i)$$

  - Ανεξάρτητα και ομοιόμορφα κατανεμημένα δεδομένα (ακολουθούν ομοιόμορφη κατανομή)

Παράδειγμα

- Έστω ότι η εκ των προτέρων κατανομή επί των $h_1,\ldots,h_5$ είναι:
  - $\langle 0.1, 0.2, 0.4, 0.2, 0.1 \rangle$
- Έστω ότι η τρέχουσα σακούλα είναι τύπου 3
- Έστω ότι ανοίξαμε τα πρώτα 10 γλυκά (όλα λεμόνι), τότε:
  - $P(\mathbf{d} \mid h_3)$ είναι 0.5^{10}

Μέγιστη εκ των υστέρων (πιθανότητα)

- **Maximum a Posteriori – MAP**
- Χρησιμοποιείται για την πρόβλεψη μόνο της πιο πιθανής υπόθεσης
  - $h_{MAP} = \max_i(P(hi\,|\,\mathbf{d})) = \max_i(P(\mathbf{d}\,|\,h_i)P(h_i))$
- Προσεγγιστικές προβλέψεις
- Όσο περισσότερα δεδομένα έχουμε, τόσο περισσότερο η πρόβλεψη MAP με την πρόβλεψη Bayes συγκλίνουν

Υπερπροσαρμογή

- Όταν υπάρχουν πάρα πολλές εναλλακτικές υποθέσεις, ικανές να ταιριάξουν με οποιοδήποτε σύνολο δεδομένων…
- …οι εκ των προτέρων πιθανότητες μπορούν να εμποδίσουν (τουλάχιστον στην αρχή) την υιοθέτηση πολύπλοκων υποθέσεων
- Δίνουμε μεγαλύτερη εκ των προτέρων πιθανότητα σε απλές υποθέσεις και μικρότερη σε πολύπλοκες υποθέσεις

Μέγιστη πιθανοφάνεια

- **maximum likelihood – ML**
- Απλοποίηση της μεθόδου MAP
- Θεωρούμε ομοιόμορφη εκ των προτέρων κατανομή πιθανότητας
  - $h_{ML} = \max_i(P(hi|\mathbf{d})) = \max_i(P(\mathbf{d}|h_i)P(h_i) =$

    $\max_i(P(\mathbf{d}|h_i))$
- …όταν δεν έχουμε κανέναν (εκ των προτέρων) λόγο να προτιμήσουμε κάποια υπόθεση

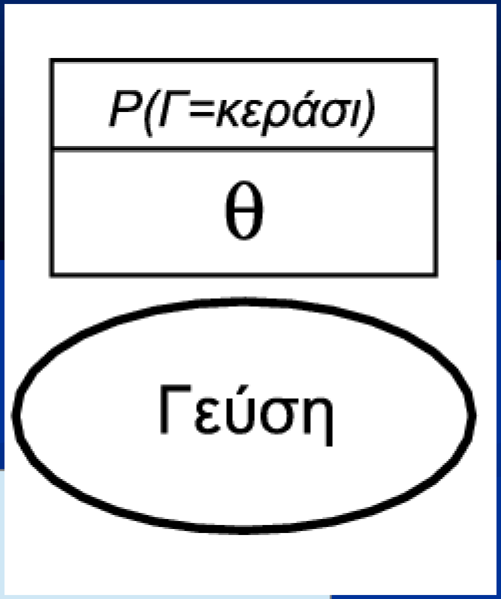
Μάθηση με πλήρη δεδομένα

Μάθηση μέγιστης πιθανοφάνειας

- <u>Βήματα:</u>
  1. Γράφουμε μια παράσταση για την πιθανοφάνεια των δεδομένων ως συνάρτηση των παραμέτρων
  2. Γράφουμε την παράγωγο της λογαριθμικής πιθανοφάνειας ως προς την κάθε παράμετρο
  3. Βρίσκουμε τις τιμές των παραμέτρων έτσι ώστε οι παράγωγοι να έχουν τιμή μηδέν

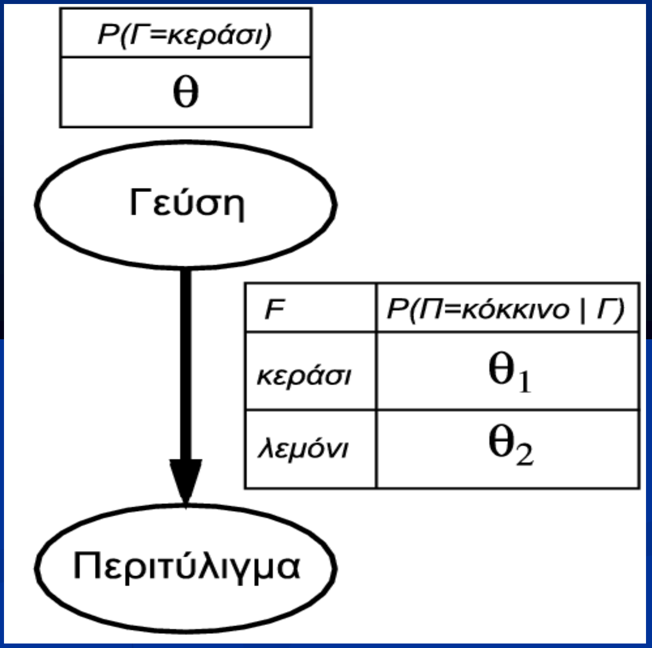
Παράδειγμα 1ο

- Έστω μία μόνο παράμετρος, η θ
- Συμβολίζουμε την τυχαία υπόθεση με h_θ
- Θεωρούμε ότι όλες οι υποθέσεις είναι εξίσου πιθανές εκ των προτέρων
  - Μέθοδος μέγιστης πιθανοφάνειας



Παράδειγμα 2ο

- Τρεις παράμετροι: $\theta, \theta_1, \theta_2$
- Υπόθεση: $P_{\theta,\theta_1,\theta_2}$
- Δεδομένα:
  - c γλυκά κεράσι
    - r_c με κόκκινο περιτύλιγμα
  - l γλυκά λεμόνι
    - r_l με κόκκινο περιτύλιγμα



- $L=[c \log\theta + l \log(1-\theta)] + [r_c \log \theta_1 + g_c \log (1-\theta_1)] + [r_l \log\theta_2 + g_l \log(1-\theta_2)]$

$$\theta_1 = \frac{r_c}{r_c + g_c}$$