Αιτιολόγηση με αβεβαιότητα

- Στα προβλήματα του πραγματικού κόσμου οι αποφάσεις συνήθως λαμβάνονται υπό **αβεβαιότητα (*uncertainty*)**, δηλαδή **έλλειψη επαρκούς πληροφορίας**.

- Οι κυριότερες πηγές αβεβαιότητας είναι:
  - *Ανακριβή δεδομένα* (*imprecise data*): π.χ. από ένα όργανο περιορισμένης ακρίβειας.
  - *Ελλιπή δεδομένα* (*incomplete data*): π.χ. σε ένα σύστημα ελέγχου, κάποιοι αισθητήρες τίθενται εκτός λειτουργίας και πρέπει να ληφθεί άμεσα μία απόφαση με τα δεδομένα από τους υπόλοιπους.
  - Αμφιβολία, υποκειμενικότητα ή/και ελλείψεις στην περιγραφή της γνώσης.
  - Εισαγωγή βαθμού πεποίθησης (πίστης) σχετικά με την αλήθεια μιας πρότασης

- *Μέθοδοι:* ***Θεωρία Πιθανοτήτων*** (*Bayesian Inference*), *Συντελεστές Βεβαιότητας* (*Certainty Factors*), *Θεωρία Dempster-Shafer,* ***Ασαφής Λογική*** **(*Fuzzy Logic*)**.

Θεωρία Πιθανοτήτων

- Αν E είναι ένα γεγονός, η *άνευ συνθηκών πιθανότητα* (*unconditional probability*) $P(E)$ να συμβεί το γεγονός εκφράζεται με έναν πραγματικό αριθμό για τον οποίο ισχύουν:
  - $0 \leq P(E) \leq 1$
  - $P(E) = 1$ αν E είναι ένα σίγουρο γεγονός
  - $P(E) + P(\neg E) = 1$ (όπου με $\neg E$ συμβολίζεται η άρνηση του γεγονότος E)
- *Πιθανότητα υπό συνθήκη ή δεσμευμένη* (*conditional probability*):

  Υποδηλώνει την πιθανότητα $P(A|B)$ να ισχύει το συμπέρασμα A δεδομένης της ισχύος του γεγονότος B.
- *Ανεξάρτητα γεγονότα* (*independence*):

  Τα A και B ανεξάρτητα, εάν η γνώση του ενός δεν επηρεάζει την πιθανότητα του άλλου.

$$P(A|B)=P(A),\ P(B|A)=P(B).$$

❖ Ιδιότητες

- *Προσθετική Ιδιότητα*: $P(A \vee B) = P(A) + P(B) - P(A \wedge B)$

- *Πολ/στική Ιδιότητα για δύο γεγονότα Α και Β*:

$$P(A,B) = P(B|A) \bullet P(A) = P(A|B) \bullet P(B)$$

Αν Α, Β ανεξάρτητα: P(A,B)=P(A) • P(B)

- **Θεώρημα Bayes**:

$$P(A \mid B) = \frac{P(B \mid A) \cdot P(A)}{P(B)}, \; P(B) > 0$$

- Γενίκευση θεωρήματος Bayes: Η πιθανότητα να ισχύει το συμπέρασμα *A* δεδομένης της ισχύος των γεγονότων $B_1, B_2, \ldots, B_k$:

$$P(A \mid B_1, \ldots, B_k) = \frac{P(B_1, \ldots, B_k \mid A) \bullet P(A)}{P(B_1, \ldots, B_k)}$$

Τυχαίες Μεταβλητές (τ.μ.)

- Μεταβλητές που οι τιμές τους προκύπτουν (‘γεννιούνται’) από διαδικασίες (‘δειγματοληψίες’) που περιέχουν τυχαιότητα.
- **Διακριτή τυχαία μεταβλητή X**: παίρνει τιμές από ένα διακριτό σύνολο n τιμών $\{x_1, \ldots, x_n\}$ με αντίστοιχες πιθανότητες $\{p_1, \ldots, p_n\}$ (τα $p_i \geq 0$ αθροίζουν στο 1).

$$P(X=x_i)=p_i$$

- **Συνεχής τυχαία μεταβλητή**: παίρνει τιμές από ένα συνεχές σύνολο με συνάρτηση πυκνότητας πιθανότητας $p(x) \geq 0$ (το ολοκλήρωμα της $p(x)$ είναι 1).

- Θα ασχοληθούμε με διακριτές τυχαίες μεταβλητές.

- **Από κοινού (joint) κατανομή** δύο τ.μ. X, Y: $p_{ij}=P(X=x_i, Y=y_j)$, (τα $p_{ij}\geq 0$ αθροίζουν στην μονάδα)
- **Σημαντικές ιδιότητες:**
  - $P(X=x_i|Y=y_j)=P(X=x_i,Y=y_j)/P(Y=y_j)$ **υπό συνθήκη (ή δεσμευμένη) πιθανότητα**
  - **Περιθωριοποίηση (marginalization)**

$$P(X=x_i) = \sum_{j=1}^{n} P(X=x_i, Y=y_j)$$

Ισχύει για οσεσδήποτε μεταβλητές (όχι μόνο δύο), δηλ. εάν τα X και Y είναι ομάδες από μεταβλητές

π.χ. $$P(X=x_i, Y=y_j) = \sum_{k=1}^{m}\sum_{l=1}^{p} P(X=x_i, Y=y_j, Z=z_k, R=r_l)$$

-
- **Αν γνωρίζουμε την από κοινού κατανομή ενός συνόλου από τυχαίες μεταβλητές μπορούμε να υπολογίσουμε οποιαδήποτε κατανομή πιθανότητας (είτε από κοινού είτε υπό συνθήκη) που περιλαμβάνει οποιεσδήποτε από αυτές τις τυχαίες μεταβλητές**.

Σημασία θεωρήματος Bayes

- Το θεώρημα του Bayes επιτρέπει τον υπολογισμό υπό συνθήκη πιθανοτήτων με χρήση άλλων υπό συνθήκη πιθανοτήτων που είναι ευκολότερο να μετρηθούν (π.χ. από διαθέσιμα στατιστικά συχνοτήτων εμφάνισης).

$$P(H \mid E) = \frac{P(E \mid H) \cdot P(H)}{P(E)}$$

- Αν **H μία ασθένεια** και **E ένα σύμπτωμα** που σχετίζεται με αυτήν, η *πιθανότητα ύπαρξης της ασθένειας δοθέντος του συμπτώματος* P(H|E) είναι αδύνατο να μετρηθεί απευθείας.
- Ωστόσο
  - Μπορούμε να έχουμε εκτίμηση (βάσει στατιστικών) για το πόσοι ασθενείς που έπασχαν από την ασθένεια H εμφάνιζαν το σύμπτωμα E, δηλ. εκτίμηση για την ποσότητα P(E|H).
  - Το ίδιο ισχύει τόσο για το P(H) (ποσοστό ασθενών με την ασθένεια H) όσο και για το P(E) (ποσοστό ασθενών με το σύμπτωμα E).
  - Από τα ανωτέρω μπορούμε να υπολογίσουμε το ζητούμενο P(H|E)

Παράδειγμα

Ο κορωνοϊός συνοδεύεται από απώλεια όσφρησης (ανοσμία) χωρίς συνάχι στο 95% των διεγνωσμένων περιπτώσεων. Εστω ότι η πιθανότητα (στατιστικά) να έχει ένα άτομο ανοσμία χωρίς συνάχι είναι 0.005 και να έχει προβληθεί από κορωνοϊό είναι 0.001.

Ποια είναι η πιθανότητα ένα άτομο με ανοσμία να έχει προσβληθεί από κορωνοϊό;

A: Ανοσμία $P(A)=0.005$

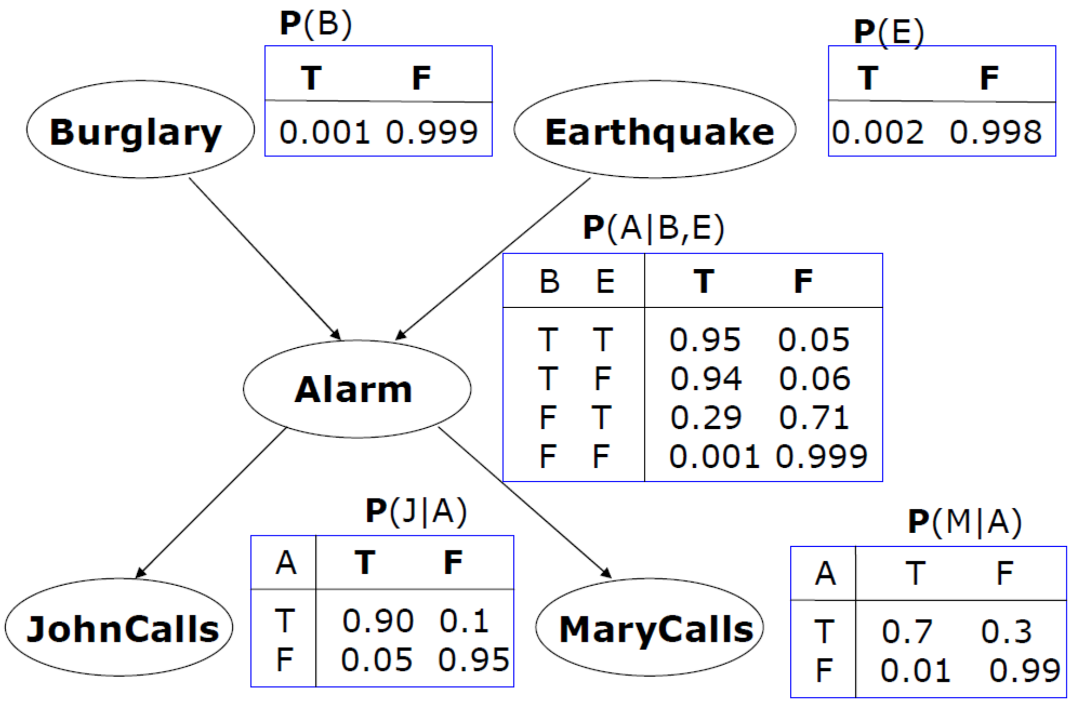
K: θετικός στον κορωνοϊό $P(K)=0.001$

$P(A|K)=0.95$, $P(K|A)=?$

$P(K|A)=P(K,A)/P(A)=P(A|K)P(K)/P(A)=0.95\ 0.001/0.005=0.19$ (19%)

Δίκτυα Πεποιθήσεων (belief networks)

- Αναπαράσταση με τη μορφή γράφου των εξαρτήσεων μεταξύ τυχαίων μεταβλητών ενός προβλήματος προκειμένου να διευκολυνθεί ο υπολογισμός της από κοινού κατανομής των μεταβλητών αυτών.
- Κάθε τ.μ ένας κόμβος του γράφου (δικτύου)
- Κάθε **κατευθυνόμενη** ακμή από μια τ.μ. Χ προς μια τ.μ. Υ δηλώνει εξάρτηση της κατανομής της Υ από τις τιμές της Χ (η Χ είναι γονέας της Υ)
- Σε κάθε κόμβο αντιστοιχεί ένας **πίνακας υπό συνθήκη πιθανοτήτων** που δηλώνει την εξάρτηση της κατανομής της τ.μ του κόμβου από τις τιμές των γονέων της



P(B)

T	F
0.001	0.999

P(E)

T	F
0.002	0.998

P(A|B,E)

B	E	T	F
T	T	0.95	0.05
T	F	0.94	0.06
F	T	0.29	0.71
F	F	0.001	0.999

P(J|A)

A	T	F
T	0.90	0.1
F	0.05	0.95

P(M|A)

A	T	F
T	0.7	0.3
F	0.01	0.99

Παράδειγμα

Ενα σπίτι έχει συναγερμό για ληστεία (burglary) και βρίσκεται σε σεισμογενή περιοχή, οπότε ο συναγερμός μπορεί να ηχήσει και λόγω σεισμού (earthquake). Υπάρχουν δύο γείτονες (ο Γιάννης και η Μαρία), οι οποίοι όταν ακούσουν το συναγερμό τηλεφωνούν στον ιδιοκτήτη, ωστόσο μπορεί να κάνουν λάθος είτε επειδή δεν θα ακούσουν το συναγερμό ενώ αυτός χτυπάει, είτε θα νομίσουν ότι χτυπάει ο συναγερμός, ενώ αυτός δεν χτυπάει.

Το δίκτυο πεποιθήσεων φαίνεται στο διπλανό σχήμα. Στους πίνακες αναφέρονται οι υπο συνθήκη πιθανότητες.

Παράδειγμα υπολογισμού από κοινού πιθανότητας όλων των τ.μ.:

$P(J=T, M=F, A=T, B=T, E=F)=$

$P(J=T|A=T)P(M=F|A=T)P(A=T|B=T,E=F)P(B=T)P(E=F)$

$= 0.9\ 0.3\ 0.94\ 0.001\ 0.998$

- **Θεώρημα: Αν το δίκτυο δεν έχει κύκλους**, τότε η από κοινού κατανομή **όλων** των μεταβλητών ισούται με το γινόμενο των επιμέρους κατανομών στους κόμβους του δικτύου

$$P(X_1,...,X_n)=\prod_{i=1}^{n}P(X_i\mid \boldsymbol{Pa_i})$$

$$\boldsymbol{Pa_i}=\text{parents}(X_\text{i})$$

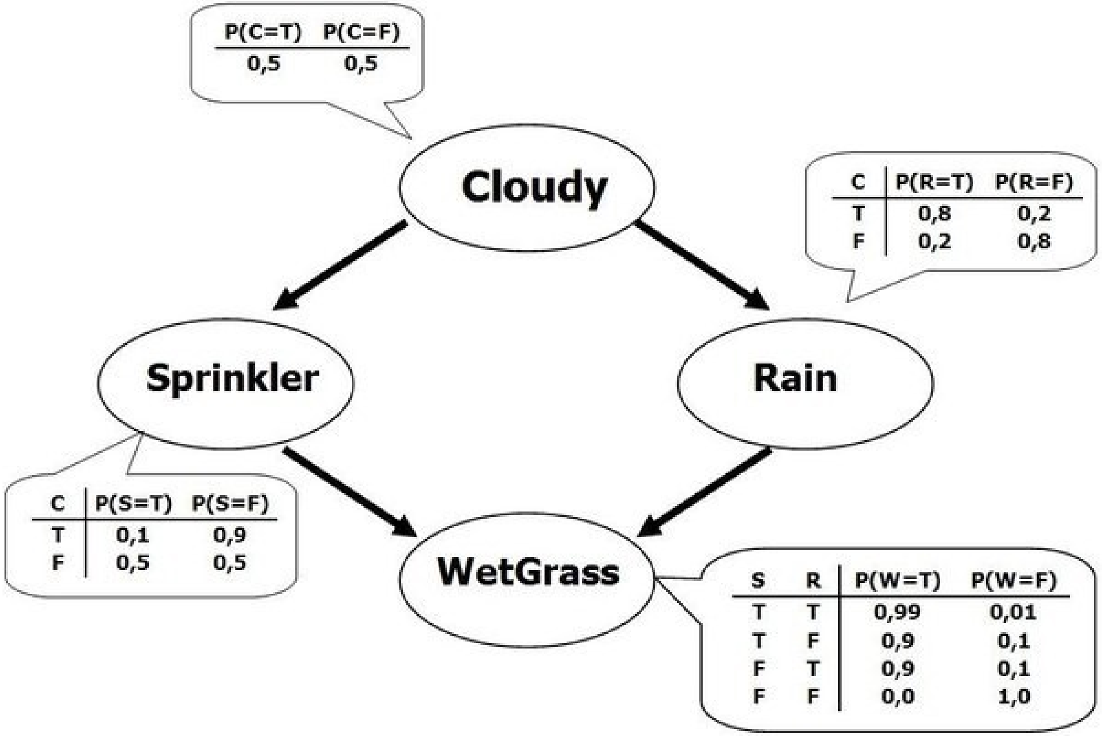
- **Αφού μπορούμε να υπολογίσουμε την από κοινού κατανομή όλων των τ.μ., μπορούμε να υπολογίσουμε οποιαδήποτε κατανομή πιθανότητας σχετίζεται με τις μεταβλητές αυτές (Μπεϋζιανή συμπερασματολογία).**
- **Εργαλεία:**
  - **Υπό συνθήκη πιθανότητα**
  - **Περιθωριοποίηση**

  (δείτε παρακάτω παράδειγμα)
- **Μοντελοποίηση**: η κατασκευή του δικτύου πεποιθήσεων για τις τ.μ. ενός προβλήματος και ο καθορισμός των πινάκων υπό συνθήκη πιθανοτήτων βάσει των δεδομένων του προβλήματος.

  (Προσοχή στο ποια οντότητα θα οριστεί τ.μ. και ποια τιμή).

- **Σημείωση:** Οι διακριτές τ.μ. δεν παίρνουν υποχρεωτικά δύο μόνο τιμές (π.χ. T, F). Μπορεί να παίρνουν οσεσδήποτε τιμές, εξαρτάται από το πρόβλημα (π.χ. ΠΟΛΥ, ΛΙΓΟ, ΚΑΘΟΛΟΥ)

Παράδειγμα: Υγρό γρασίδι (από πότισμα ή βροχή)



1^{η} **Περίπτωση**: ζητείται η από κοινού πιθανότητα **όλων** των τ.μ.

Λύση: Αμεση εφαρμογή του κανόνα γινομένου των υπό συνθήκη πιθανοτήτων των τ.μ. δοθέντων των γονέων τους

$P(W=T, S=F, R=T, C=T) = P(W=T|S=F,R=T)P(S=F|C=T)P(R=T|C=T)P(C=T)$

$= 0.9\ 0.9\ 0.8\ 0.5 = 0.324$

2^{η} **Περίπτωση:** ζητείται η από κοινού πιθανότητα **κάποιων** τ.μ. (όχι όλων)

Λύση: Εφαρμόζοντας περιθωριοποίηση προσθέτουμε τις μεταβλητές που λείπουν. Προκύπτει άθροισμα από πιθανότητες που περιλαμβάνουν όλες τις μεταβλητές. Για κάθε πιθανότητα εφαρμόζουμε την περίπτωση 1 (κανόνας γινομένου).

$P(W=T,S=F,R=T) = P(W=T,S=F,R=T,C=T) + P(W=T,S=F,R=T,C=F) =$

$P(W=T|S=F,R=T)P(S=F|C=T)P(R=T|C=T)P(C=T)+$

$P(W=T|S=F,R=T)\ P(S=F|C=F)\ P(R=T|C=F)\ P(C=F) =$ ….(από τους πίνακες)

$P(W=T,R=T) = P(W=T,S=T,R=T,C=T) + P(W=T,S=T,R=T,C=F) +$

$P(W=T,S=F,R=T,C=T) + P(W=T,S=F,R=T,C=F) =$ … (κάθε όρος γινόμενο υπό συνθήκη πιθανοτήτων).

3η Περίπτωση: ζητείται η υπό συνθήκη πιθανότητα κάποιων τ.μ. δοθέντων κάποιων άλλων τ.μ..

Λύση: Η υπο συνθήκη πιθανότητα ορίζεται ως το πηλίκο δύο πιθανοτήτων. Δουλεύουμε ξεχωριστά με τον αριθμητή και τον παρονομαστή αξιοποιώντας τις προηγούμενες περιπτώσεις.

$P(W=T,S=F,R=T|C=T)= P(W=T,S=F,R=T,C=T)/P(C=T)=$

$P(W=T|S=F,R=T)P(S=F|C=T)P(R=T|C=T)P(C=T)/P(C=T)=0.324/0.5=0.162$

(στην περίπτωσή μας ο παρονομαστής P(C=T) δίνεται απ’ευθείας από το δίκτυο πεποιθήσεων, αλλιώς θα έπρεπε να τον υπολογίσουμε με περιθωριοποίηση).

Παράδειγμα

Ενα robot εποπτεύει ένα κρίσιμο σύστημα (π.χ. αντιδραστήρας) και αναφέρει "κίνδυνος" (ΚΔ) όταν διαπιστώσει κάποια βλάβη στη λειτουργία του συστήματος. Το σύστημα μπορεί να είναι κάποια στιγμή είτε ασφαλές (Α) είτε με βλάβη (Β). Ισχύουν τα εξής:

α) Αν οι μπαταρίες του robot έχουν αποφορτιστεί (γεγονός ΑΦ), τότε η πιθανότητα να **αναφέρει εσφαλμένα** ότι υπάρχει κίνδυνος είναι 0.75 και η πιθανότητα να **αναφέρει σωστά** ότι υπάρχει κίνδυνος είναι 0.65.

β) Αν οι μπαταρίες είναι σωστά φορτισμένες (γεγονός Φ), τότε οι αναφορές του robot είναι 100% σωστές.

γ) Δεν έχουμε φορτίσει για κάποιες ημέρες τις μπαταρίες του robot, με αποτέλεσμα να υπάρχει πιθανότητα 0.25 να έχουν αποφορτιστεί.

δ) Η πιθανότητα (στατιστικά) να εμφανίσει βλάβη το συγκεκριμένο σύστημα είναι 0.1.

1) Να κατασκευάσετε το δίκτυο πεποιθήσεων που αντιστοιχεί στο παραπάνω πρόβλημα και να ορίσετε τους πίνακες πιθανοτήτων σε κάθε κόμβο του δικτύου.

2) Βασιζόμενοι στο δίκτυο πεποιθήσεων που ορίσατε, να υπολογίσετε την πιθανότητα εσφαλμένης διάγνωσης κινδύνου (false alarm).

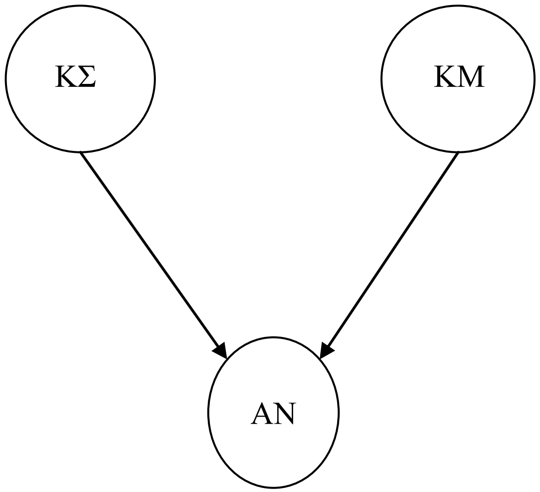
Λύση

1)Τυχαίες μεταβλητές:

Κατάσταση του συστήματος (ΚΣ) Τιμές: Α (ασφαλές), Β (βλάβη)

Κατάσταση μπαταριών (ΚΜ) Τιμές: Φ (φορτισμένες), ΑΦ (αφόρτιστες)

Αναφορά ρομπότ (ΑΝ) Τιμές: ΚΔ (κίνδυνος), ΟΚ (ομαλή λειτουργία)



P(ΚΜ=Φ)=0.75 P(ΚΜ=ΑΦ)=0.25

P(ΚΣ=Α)=0.9 P(ΚΣ=Β)=0.1

ΚΣ	ΚΜ	P(ΑΝ=ΚΔ|ΚΣ,ΚΜ)	P(ΑΝ=ΟΚ|ΚΣ,ΚΜ)
Α	Φ	0	1
Α	ΑΦ	0.75	0.25
Β	Φ	1	0
Β	ΑΦ	0.65	0.35

2)P(ΑΝ=ΚΔ | ΚΣ=Α) = P(ΑΝ=ΚΔ, ΚΣ=Α)/P(ΚΣ=Α)

P(ΑΝ=ΚΔ, ΚΣ=Α)= P(ΑΝ=ΚΔ, ΚΣ=Α, ΚΜ=Φ) + P(ΑΝ=ΚΔ, ΚΣ=Α, ΚΜ=ΑΦ) =

P(ΑΝ=ΚΔ|ΚΣ=Α, ΚΜ=Φ) P(ΚΣ=Α) P(ΚΜ=Φ) + P(ΑΝ=ΚΔ|ΚΣ=Α, ΚΜ=ΑΦ) P(ΚΣ=Α) P(ΚΜ=ΑΦ) (οι τιμές δίνονται στους πίνακες)