Υπό συνθήκη πιθανότητα
Έστω *S* ένας χώρος πιθανότητας και έστω Α, Β δύο ενδεχόμενα. Η υπό συνθήκη πιθανότητα του Α δεδομένου ότι το Β είναι

$$\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)}$$

Στατιστική Ανεξαρτησία

Τα Α και Β είναι στατιστικά ανεξάρτητα εάν και μόνο εάν

$$\Pr(A \mid B) = \Pr(A)$$

Δηλαδή, γνωρίζοντας ότι το Β συνέβη δεν αλλάζει την πιθανότητα να συμβεί το Α

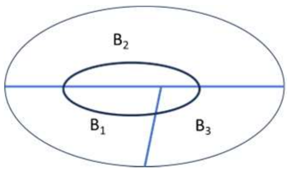
Τότε έπεται ότι

$$\Pr(A \cap B) = \Pr(A)\,\Pr(B)$$

1

Θεώρημα Ολικής Πιθανότητας
Δεδομένου ενός συνόλου αμοιβαία αποκλειόμενων και αθροιστικά επιμεριζόντων ενδεχομένων $\{B_1, B_2, \dots, B_n\}$ η πιθανότητα το σύνθετο ενδεχόμενο Α είναι

$$\Pr(A) = \Pr(A \cap B_1) + \Pr(A \cap B_2) + \dots + \Pr(A \cap B_n) =$$
$$= \sum_{i=1}^{N} \Pr(A \mid B_i)\,\Pr(B_i)$$



2

Ο κανόνας Bayes
Η υπό συνθήκη πιθανότητα του Β δεδομένου ότι το Α συνέβη μπορεί να υπολογιστεί από τη σχέση

$$\Pr(B \mid A) = \frac{\Pr(B \cap A)}{\Pr(A)} = \frac{\Pr(A \cap B)}{\Pr(A)} = \frac{\Pr(A \mid B)\ \Pr(B)}{\Pr(A)}$$

3

Ή χρησιμοποιώντας το Θεώρημα Ολικής Πιθανότητας

Κανόνας Bayes $\Pr(B_j \mid A) = \dfrac{\Pr(A \mid B_j)\,\Pr(B_j)}{\sum_{i=1}^{n} \Pr(A \mid B_i)\,\Pr(B_i)}$

4

Καταρχήν Εκτίμηση (Prior Assessment)

Χρησιμοποιείται εμπειρική κρίση για την εκτίμηση της $\Pr(B_j)$ η οποία είναι η πιθανότητα του ενδεχομένου B_j
πριν γίνει οποιαδήποτε παρατήρηση

5

Μεταγενέστερη Εκτίμηση (Posterior Assessment)

Κατόπιν αυτή η κρίση **διορθώνεται** δεδομένης της παρατήρησης Α και υπολογίζεται η $\Pr(B_j|A)$

$$\Pr(\text{κατάσταση } j \mid \text{δείγμα}) = \frac{\Pr(\text{δείγμα}|\text{κατάσταση } j)\,\Pr(\text{κατάσταση } j)}{\sum_{\substack{\text{κάθε}\\\text{κατάσταση}}} \Pr(\text{δείγμα}|\text{κατάσταση})\,\Pr(\text{κατάσταση})}$$

6

Παράδειγμα 1:
Έχω βήχα και ζαλάδα, αλλά όχι πυρετό. Μήπως έχω covid;

Δεν έχω ακούσει να υπάρχει έξαρση κρουσμάτων covid και είμαι σχετικά αισιόδοξος ότι είναι μια απλή ίωση

prior assessment P[δεν έχω] = 0.6 P[έχω] = 0.4

Έστω ότι θέλω να βεβαιωθώ και αποφασίζω να κάνω ένα self-test. Στις οδηγίες διαβάζω ότι το τεστ δεν είναι 100% ακριβές. Από εκτενή εργαστηριακά δεδομένα είναι γνωστές οι εξής υπό συνθήκη πιθανότητες για αρνητικό ή θετικό τεστ.

P[αρνητικό | δεν έχω] = 0.8 P[θετικό | δεν έχω] = 0.2

P[αρνητικό | έχω] = 0.1 P[θετικό | έχω] = 0.9

posterior assessment μετά από αρνητικό τεστ

$$\Pr(\text{έχω} \mid \text{αρνητικό}) = \frac{\Pr(\text{αρνητικό}|\text{έχω})\ \Pr(\text{έχω})}{\Pr(\text{αρνητικό})} =$$

$$\frac{\Pr(\text{αρνητικό}|\text{έχω})\ \Pr(\text{έχω})}{\Pr(\text{αρνητικό}|\text{έχω})\ \Pr(\text{έχω}) + \Pr(\text{αρνητικό}|\text{δεν έχω})\ \Pr(\text{δεν έχω})}$$

$$= \frac{0.1 \text{ x } 0.4}{0.1 \text{ x } 0.4 + 0.8 \text{ x } 0.6} = \frac{0.04}{0.04 + 0.48} = 0.077$$

Το γεγονός ότι το τεστ βγήκε αρνητικό **μειώνει** την αρχική εκτίμηση της πιθανότητας να έχω από 0.40 σε 0.077
και
αυξάνει την αρχική εκτίμηση της πιθανότητας να μην έχω από 0.60 σε 1-0.077 = 0.923

$$\Pr(\text{δεν έχω} \mid \text{αρνητικό}) =$$

$$\frac{\Pr(\text{αρνητικό}|\text{δεν έχω})\ \Pr(\text{δεν έχω})}{\Pr(\text{αρνητικό}|\text{έχω})\ \Pr(\text{έχω}) + \Pr(\text{αρνητικό}|\text{δεν έχω})\ \Pr(\text{δεν έχω})}$$

$$= \frac{0.8 \text{ x } 0.6}{0.1 \text{ x } 0.4 + 0.8 \text{ x } 0.6} = \frac{0.48}{0.04 + 0.48} = 0.923$$

posterior assessment μετά από θετικό τεστ

$$\Pr(\text{έχω} \mid \text{θετικό}) = \frac{\Pr(\text{θετικό}|\text{έχω})\ \Pr(\text{έχω})}{\Pr(\text{θετικό}|\text{έχω})\ \Pr(\text{έχω}) + \Pr(\text{θετικό}|\text{δεν έχω})\ \Pr(\text{δεν έχω})}$$

$$= \frac{0.9 \text{ x } 0.4}{0.9 \text{ x } 0.4 + 0.2 \text{ x } 0.6} = \frac{0.36}{0.36 + 0.12} = 0.75$$

$$\Pr(\text{δεν έχω} \mid \text{θετικό}) = \frac{\Pr(\text{θετικό}|\text{δεν έχω})\ \Pr(\text{δεν έχω})}{\Pr(\text{θετικό}|\text{έχω})\ \Pr(\text{έχω}) + \Pr(\text{θετικό}|\text{δεν έχω})\ \Pr(\text{δεν έχω})}$$

$$= \frac{0.2 \text{ x } 0.6}{0.9 \text{ x } 0.4 + 0.2 \text{ x } 0.6} = \frac{0.12}{0.36 + 0.12} = 0.25$$

Παράδειγμα 2:

Εξετάζεται η στατική επάρκεια ενός υφιστάμενου κτηρίου. Από τα στοιχεία που είναι γνωστά για την ποιότητα του σκυροδέματος που χρησιμοποιήθηκε όταν κατασκευάστηκε το κτήριο, ο μηχανικός κάνει μια κατ΄ αρχήν εκτίμηση ότι η αντοχή (κατάσταση) σε psi έχει πιθανότητες:

P[2000] = 0.3 P[3000] = 0.6 P[4000] = 0.1

Προτείνεται να εξαχθούν δοκίμια και μετρηθεί η αντοχή τους, όμως είναι γνωστό ότι το αποτέλεσμα των δοκιμών δεν θα είναι μονοσήμαντο. Δηλαδή αν η πραγματική αντοχή του σκυροδέματος είναι 3000, μόνο για το 60% των δοκιμίων αναμένεται να μετρηθούν αντοχές που μπορούν να χαρακτηριστούν με αυτή την τιμή. Το υπόλοιπο 40% θα μπορεί να χαρακτηριστεί με τις δύο άλλες τιμές με ίση πιθανότητα. Διαμορφώνεται έτσι ο παρακάτω πίνακας των υπό συνθήκες πιθανοτήτων Pr(μέτρηση| κατάσταση):

	Κατάσταση		
Μέτρηση	2000	3000	4000
z_1	0.7	0.2	0.0
z_2	0.3	0.6	0.3
z_3	0.0	0.2	0.7
	1.0	1.0	1.0

	Κατάσταση		
Μέτρηση	2000	3000	4000
z_1	0.7	0.2	0.0
z_2	0.3	0.6	0.3
z_3	0.0	0.2	0.7
	1.0	1.0	1.0

$P(z_1) = P(z_1|2000)\ P(2000) + P(z_1|3000)\ P(3000) + P(z_1|4000)\ P(4000)$
$= 0.7 \text{ x } 0.3 + 0.2 \text{ x } 0.6 + 0.0 \text{ x } 0.1 = \mathbf{0.33}$

$P(z_2) = P(z_2|2000)\ P(2000) + P(z_2|3000)\ P(3000) + P(z_2|4000)\ P(4000)$
$= 0.3 \text{ x } 0.3 + 0.6 \text{ x } 0.6 + 0.3 \text{ x } 0.1 = 0.48$

$P(z_3) = P(z_3|2000)\ P(2000) + P(z_3|3000)\ P(3000) + P(z_3|4000)\ P(4000)$
$= 0.0 \text{ x } 0.3 + 0.2 \text{ x } 0.6 + 0.7 \text{ x } 0.1 = 0.82$

1. Εξάγεται ένα δοκίμιο και η αντοχή του μετράται και χαρακτηρίζεται σαν z_1. Υπολογίζονται οι posterior πιθανότητες της κατάστασης, ως εξής:

$$P[2000|z_1] = \frac{0.7 \times 0.3}{0.7 \times 0.3 + 0.2 \times 0.6 + 0 \times 0.1} = \frac{0.21}{0.33} = 0.635$$

$$P[3000|z_1] = \frac{0.2 \times 0.6}{0.7 \times 0.3 + 0.2 \times 0.6 + 0 \times 0.1} = \frac{0.12}{0.33} = 0.365$$

$$P[4000|z_1] = \frac{0 \times 0.1}{0.7 \times 0.3 + 0.2 \times 0.6 + 0 \times 0.1} = \frac{0.}{0.33} = 0.0$$

13

10.6 ΠΡΟΣΔΙΟΡΙΣΜΟΣ ΠΙΘΑΝΟΤΗΤΩΝ ΚΛΑΔΩΝ

Παρουσιάζουμε εκ νέου το δέντρο αποφάσεων της PDC θέτοντας:

F = ευνοϊκή αναφορά έρευνας αγοράς

U = δυσμενής αναφορά έρευνας αγοράς

s_1 = αυξημένη ζήτηση (φυσική κατάσταση 1)

s_2 = ασθενής ζήτηση (φυσική κατάσταση 1)

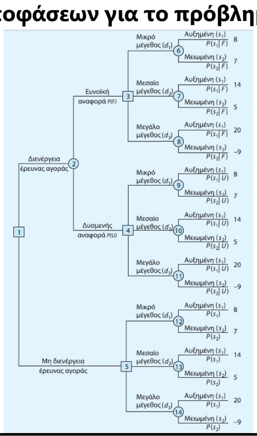
Οι πιθανότητες $P(s_1)$ και $P(s_2)$ είναι οι **εκ των προτέρω** εκτιμήσεις της πιθανότητας των δύο φυσικών καταστάσεων

Οι πιθανότητες $P(s_1 | F)$ και $P(s_2 | F)$ είναι οι **εκ των υστέρω** εκτιμήσεις της πιθανότητας των δύο φυσικών καταστάσεων με ευνοϊκή αναφορά έρευνας αγοράς ενώ οι πιθανότητες $P(s_1 | U)$ και $P(s_2 | U)$ είναι οι αντίστοιχες εκτιμήσεις με δυσμενή αναφορά έρευνας αγοράς

14

10.6 ΠΡΟΣΔΙΟΡΙΣΜΟΣ ΠΙΘΑΝΟΤΗΤΩΝ ΚΛΑΔΩΝ

Δέντρο αποφάσεων για το πρόβλημα της PDC



15

10.6 ΠΡΟΣΔΙΟΡΙΣΜΟΣ ΠΙΘΑΝΟΤΗΤΩΝ ΚΛΑΔΩΝ

Από εμπειρία έχουμε «λογικές» εκτιμήσεις για τις δεσμευμένες πιθανότητες του προβλήματος P(αποτέλεσμα έρευνας | φυσική κατάσταση)

Φυσική κατάσταση	Έρευνα αγοράς Ευνοϊκή, F	Δυσμενής, U
Αυξημένη ζήτηση, s_1	$P(F \mid s_1) = 0{,}90$	$P(U \mid s_1) = 0{,}10$
Μειωμένη ζήτηση, s_2	$P(F \mid s_2) = 0{,}25$	$P(U \mid s_2) = 0{,}75$

Εάν πραγματοποιηθεί η φυσική κατάσταση s_1, η πιθανότητα ευνοϊκής αναφοράς της έρευνας αγοράς θα είναι 0,90 και η πιθανότητα δυσμενούς αναφοράς 0,10.

Εάν πραγματοποιηθεί η φυσική κατάσταση s_2, η πιθανότητα ευνοϊκής αναφοράς της έρευνας αγοράς θα είναι 0,25 και η πιθανότητα δυσμενούς αναφοράς 0,75.

16

10.6 ΠΡΟΣΔΙΟΡΙΣΜΟΣ ΠΙΘΑΝΟΤΗΤΩΝ ΚΛΑΔΩΝ

Υπολογισμός των εκ των υστέρω πιθανοτήτων για ευνοϊκή αναφορά της έρευνας αγοράς

$$P(s_1|F) = \frac{P(F|s_1)\,P(s_1)}{P(F)} = \frac{P(F|s_1)\,P(s_1)}{P(F|s_1)P(s_1)+P(F|s_2)P(s_2)} = \frac{0.9 \times 0.8}{0.9 \times 0.8 + 0.25 \times 0.2} = \frac{0.72}{0.72 + 0.05} = \frac{0.72}{0.77} = 0.94$$

$$P(s_2|F) = \frac{P(F|s_2)\,P(s_2)}{P(F)} = \frac{P(F|s_2)\,P(s_2)}{P(F|s_1)P(s_1)+P(F|s_2)P(s_2)} = \frac{0.25 \times 0.2}{0.9 \times 0.8 + 0.25 \times 0.2} = \frac{0.05}{0.72 + 0.05} = \frac{0.05}{0.77} = 0.06$$

Φυσικές καταστάσεις s_j	Εκ των προτέρων πιθανότητες $P(s_j)$	Δεσμευμένες πιθανότητες $P(F \mid s_j)$	Συνδυασμένες πιθανότητες $P(F \cap s_j)$	Εκ των υστέρων πιθανότητες $P(s_j \mid F)$
s_1	0,8	0,90	0,72	0,94
s_2	0,2	0,25	0,05	0,06
	1,0	$P(F) =$	0,77	1,00

17

10.6 ΠΡΟΣΔΙΟΡΙΣΜΟΣ ΠΙΘΑΝΟΤΗΤΩΝ ΚΛΑΔΩΝ

Υπολογισμός των εκ των υστέρω πιθανοτήτων για δυσμενή αναφορά της έρευνας αγοράς

$$P(s_1|U) = \frac{P(U|s_1)\,P(s_1)}{P(U)} = \frac{P(U|s_1)\,P(s_1)}{P(U|s_1)P(s_1)+P(U|s_2)P(s_2)} = \frac{0.1 \times 0.8}{0.1 \times 0.8 + 0.75 \times 0.2} = \frac{0.08}{0.08 + 0.15} = \frac{0.08}{0.23} = 0.35$$

$$P(s_2|U) = \frac{P(U|s_2)\,P(s_2)}{P(U)} = \frac{P(U|s_2)\,P(s_2)}{P(U|s_1)P(s_1)+P(U|s_2)P(s_2)} = \frac{0.75 \times 0.2}{0.1 \times 0.8 + 0.75 \times 0.2} = \frac{0.15}{0.08 + 0.15} = \frac{0.15}{0.23} = 0.65$$

Φυσικές καταστάσεις s_j	Εκ των προτέρων πιθανότητες $P(s_j)$	Δεσμευμένες πιθανότητες $P(U \mid s_j)$	Συνδυασμένες πιθανότητες $P(U \cap s_j)$	Εκ των υστέρων πιθανότητες $P(s_j \mid U)$
s_1	0,8	0,10	0,08	0,35
s_2	0,2	0,75	0,15	0,65
	1,0	$P(U) =$	0,23	1,00

18

10.6 ΠΡΟΣΔΙΟΡΙΣΜΟΣ ΠΙΘΑΝΟΤΗΤΩΝ ΚΛΑΔΩΝ

Πιθανότητες κλάδων για ευνοϊκή αναφορά της έρευνας αγοράς

Τα βήματα για τη σύνταξη του συγκεκριμένου πίνακα είναι τα ακόλουθα:

Βήμα 1

Στη στήλη 1 εισάγουμε τις φυσικές καταστάσεις. Στη στήλη 2 εισάγου με τις εκ των προτέρων πιθανότητες. Στη στήλη 3 εισάγουμε τις δεσμευμένες πιθανότητες για ευνοϊκή αναφορά της έρευνας αγοράς (F) για κάθε φυσική κατάσταση.

Βήμα 2

Στη στήλη 4 προσδιορίζουμε τις συνδυασμένες πιθανότητες, πολλαπλασιάζοντας τις τιμές των εκ των προτέρων πιθανοτήτων (στήλη 2) με τις αντίστοιχες τιμές των δεσμευμένων πιθανοτήτων (στήλη 3).

10.6 ΠΡΟΣΔΙΟΡΙΣΜΟΣ ΠΙΘΑΝΟΤΗΤΩΝ ΚΛΑΔΩΝ

Πιθανότητες κλάδων για ευνοϊκή αναφορά της έρευνας αγοράς

Βήμα 3

Αθροίζουμε τις συνδυασμένες πιθανότητες της στήλης 4 για να λάβουμε την πιθανότητα ευνοϊκής αναφοράς της έρευνας αγοράς P(F).

Βήμα 4

Διαιρούμε κάθε συνδυασμένη πιθανότητα στη στήλη 4 με P(F) = 0,77 για να λάβουμε τις αναθεωρημένες ή εκ των υστέρων πιθανότητες $P(s_1 \mid F)$ και $P(s_2 \mid F)$.

10.6 ΠΡΟΣΔΙΟΡΙΣΜΟΣ ΠΙΘΑΝΟΤΗΤΩΝ ΚΛΑΔΩΝ

Πιθανότητες κλάδων για δυσμενή αναφορά της έρευνας αγοράς