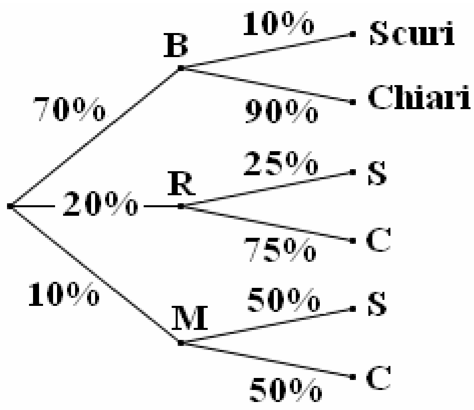
14.2 - Esercizi svolti (Teorema di Bayes)

Riprendiamo ora i problemi da cui avevamo preso spunto e risolviamoli.

- **In un paese scandinavo il 70% delle ragazze ha i capelli Biondi, il 20% li ha Rossi, il 10% Mori. Risulta poi che ha gli occhi Scuri il 10% delle Bionde, il 25% delle Rosse, il 50% delle More. Se la ragazza con cui ho fatto amicizia tramite Internet mi fa sapere che ha gli occhi Scuri, che probabilità c'è che sia Bionda?**

B	R	M	
70%	20%	10%	della popolazione femminile
10%	25%	50%	occhi Scuri



Risoluzione con la Formula di Bayes:

$p(\mathrm{B}) = 0,7 \qquad p(\mathrm{R}) = 0,2 \qquad p(\mathrm{M}) = 0,1$

$p(\mathrm{S/B}) = 0,1 \qquad p(\mathrm{S/R}) = 0,25 \qquad p(\mathrm{S/M}) = 0,5$

$$p(\mathrm{B/S}) = \frac{p(\mathrm{B}) \cdot p(\mathrm{S/B})}{p(\mathrm{B}) \cdot p(\mathrm{S/B}) + p(\mathrm{R}) \cdot p(\mathrm{S/R}) + p(\mathrm{M}) \cdot p(\mathrm{S/M})} =$$

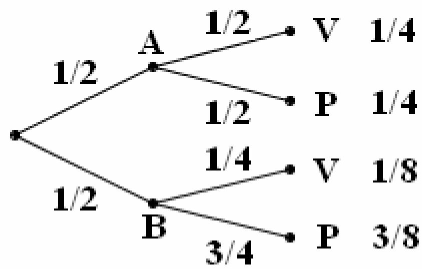
$$= \frac{0,7 \cdot 0,1}{0,7 \cdot 0,1 + 0,2 \cdot 0,25 + 0,1 \cdot 0,5} =$$

$$= \frac{0,07}{0,07 + 0,05 + 0,05} = \frac{0,07}{0,17} \approx 0,41 = 41\%$$

- **In un bar ci sono due macchinette mangiasoldi A e B. Effettuando una singola giocata su A si vince con probabilità 1/2 (in altre parole: si vince mediamente 1 volta su 2, o, se preferisci, all'incirca 500 volte su 1000), mentre giocando su B si vince con probabilità 1/4. Supponiamo di non sapere quale sia la macchinetta A e quale la B; se ne scegliamo una a caso, giochiamo una sola volta, e vinciamo, che probabilità c'è che la macchinetta scelta sia stata A?**

DIVERSI METODI O STILI DI RISOLUZIONE

1) Con la formula di Bayes



$$p(\mathrm{A/V}) = \frac{p(\mathrm{A}) \cdot p(\mathrm{V/A})}{p(\mathrm{A}) \cdot p(\mathrm{V/A}) + p(\mathrm{B}) \cdot p(\mathrm{V/B})} =$$

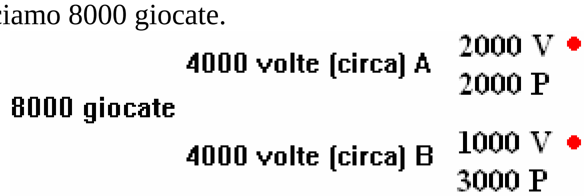
$$= \frac{\frac{1}{2} \cdot \frac{1}{2}}{\frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{4}} = \frac{\frac{1}{4}}{\frac{1}{4} + \frac{1}{8}} = \frac{\frac{1}{4}}{\frac{3}{8}} = \frac{2}{3}$$

2) Pensando semplicemente ad un'applicazione della formula $\boldsymbol{p(\mathrm{X/Y}) = \frac{p(\mathrm{X} \wedge \mathrm{Y})}{p(\mathrm{Y})}}$

$$p(\mathrm{A}/\text{"Vittoria"}) = \frac{p(\mathrm{A} \wedge \mathrm{V})}{p(\mathrm{V})} = \frac{p(\mathrm{A}) \cdot p(\mathrm{V/A})}{p(\mathrm{A} \wedge \mathrm{V}) + p(\mathrm{B} \wedge \mathrm{V})} = \frac{p(\mathrm{A}) \cdot p(\mathrm{V/A})}{p(\mathrm{A}) \cdot p(\mathrm{V/A}) + p(\mathrm{B}) \cdot p(\mathrm{V/B})} = \frac{\frac{1}{2} \cdot \frac{1}{2}}{\frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{4}} = \ldots = \frac{2}{3}$$

3) Con visione "frequentista"

Supponiamo di effettuare un numero elevato di giocate, diciamo 8000 giocate.



Pressappoco 4000 volte la scelta casuale della macchinetta cadrà su A, e pressappoco 4000 volte su B (legge empirica del caso).
Delle circa 4000 volte che avremo giocato su A, vinceremo circa 2000 volte, mentre delle circa 4000 volte che avremo giocato su B vinceremo circa 1000 volte.
Avremo perciò vinto 3000 volte circa. E di queste, pressappoco 2000 volte dovremo ringraziare A.

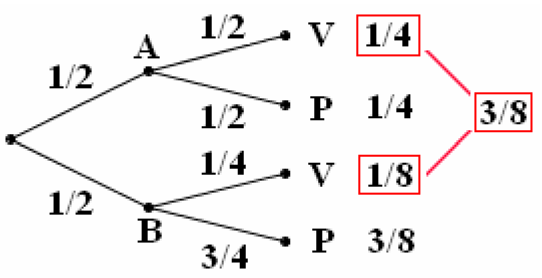
Perciò la probabilità richiesta (= probabilità che, avendo noi vinto, si debba ringraziare A) è $\frac{2000}{3000} = \frac{2}{3}$.

4) Pensando alle “fette di certezza”

Nella “torta della certezza”, di “peso” 1, la “fetta di certezza” relativa a V “pesa”, complessivamente, $\frac{1}{4}+\frac{1}{8}=\frac{3}{8}$.

Di questa fetta da 3/8, la parte che compete all’evento “A, poi V” ha peso 1/4.

Il rapporto tra le due fette è quindi $\frac{1/4}{3/8}=\frac{2}{3}$.



UN ULTERIORE, BELL’ESEMPIO: GLI ARCIERI

a) Se quattro arcieri A, B, C, D scoccano la loro freccia contemporaneamente e hanno probabilità, rispettivamente, 1/2, 1/3, 1/4 e 1/5 di colpire il bersaglio (NOTA) che probabilità c’è che dopo il tiro simultaneo risulti conficcata nel bersaglio esattamente 1 freccia?
NOTA: si tratta, evidentemente, di valutazioni approssimative, di tipo soggettivo/frequentista

b) Se dopo il tiro simultaneo risulta conficcata nel bersaglio 1 e 1 sola freccia, che probabilità c’è che si tratti di quella dell’arciere A?

Risoluzione di a)

$$p(\mathrm{A})=1/2 \quad p(\mathrm{B})=1/3 \quad p(\mathrm{C})=1/4 \quad p(\mathrm{D})=1/5$$

$$p(\text{"}esattamente\ 1\ \ freccia\text{"}) = p(\mathrm{A}\,\overline{\mathrm{B}}\,\overline{\mathrm{C}}\,\overline{\mathrm{D}}) + p(\overline{\mathrm{A}}\,\mathrm{B}\,\overline{\mathrm{C}}\,\overline{\mathrm{D}}) + p(\overline{\mathrm{A}}\,\overline{\mathrm{B}}\,\mathrm{C}\,\overline{\mathrm{D}}) + p(\overline{\mathrm{A}}\,\overline{\mathrm{B}}\,\overline{\mathrm{C}}\,\mathrm{D}) =$$

$$=\frac{1}{2}\cdot\frac{2}{3}\cdot\frac{3}{4}\cdot\frac{4}{5}+\frac{1}{2}\cdot\frac{1}{3}\cdot\frac{3}{4}\cdot\frac{4}{5}+\frac{1}{2}\cdot\frac{2}{3}\cdot\frac{1}{4}\cdot\frac{4}{5}+\frac{1}{2}\cdot\frac{2}{3}\cdot\frac{3}{4}\cdot\frac{1}{5}=\frac{24+12+8+6}{120}=\frac{50}{120}=\frac{5}{12}$$

Risoluzione di b)

Possiamo risolvere questo quesito b) ricorrendo, formalmente,

1) **alla formula** $\boldsymbol{p(\mathrm{X}/\mathrm{Y})=\frac{p(\mathrm{X}\wedge \mathrm{Y})}{p(\mathrm{Y})}}$

2) **oppure alla formula di Bayes.**

I due procedimenti non differiscono molto né riguardo al principio ispiratore (sempre di “cause”, o piuttosto, in questo caso, di “eventi concomitanti”, o di “fette di certezza”, si tratta), né riguardo alla difficoltà nei calcoli (che sono anzi del tutto identici).

Dunque, vediamo.

1) **con la formula** $\boldsymbol{p(\mathrm{X}/\mathrm{Y})=\frac{p(\mathrm{X}\wedge \mathrm{Y})}{p(\mathrm{Y})}}$

$$p(\mathrm{A}/\,1\,\mathrm{e}\,1\,\mathrm{sola})=\frac{p\big(\mathrm{A}\wedge(1\,\mathrm{e}\,1\,\mathrm{sola})\big)}{p\big(1\,\mathrm{e}\,1\,\mathrm{sola}\big)}=\frac{p(\mathrm{A}\overline{\mathrm{B}}\overline{\mathrm{C}}\overline{\mathrm{D}})}{p(\mathrm{A}\overline{\mathrm{B}}\overline{\mathrm{C}}\overline{\mathrm{D}})+p(\overline{\mathrm{A}}\mathrm{B}\overline{\mathrm{C}}\overline{\mathrm{D}})+p(\overline{\mathrm{A}}\overline{\mathrm{B}}\mathrm{C}\overline{\mathrm{D}})+p(\overline{\mathrm{A}}\overline{\mathrm{B}}\overline{\mathrm{C}}\mathrm{D})}=$$

$$=\frac{\frac{1}{2}\cdot\frac{2}{3}\cdot\frac{3}{4}\cdot\frac{4}{5}}{\frac{1}{2}\cdot\frac{2}{3}\cdot\frac{3}{4}\cdot\frac{4}{5}+\frac{1}{2}\cdot\frac{1}{3}\cdot\frac{3}{4}\cdot\frac{4}{5}+\frac{1}{2}\cdot\frac{2}{3}\cdot\frac{1}{4}\cdot\frac{4}{5}+\frac{1}{2}\cdot\frac{2}{3}\cdot\frac{3}{4}\cdot\frac{1}{5}}=\frac{\frac{24}{120}}{\frac{24}{120}+\frac{12}{120}+\frac{8}{120}+\frac{6}{120}}=\frac{24}{50}=48\%$$

2) **con la formula di Bayes**

$$p(\mathrm{A}/1\,e\,1\,sola)=$$

$$=\frac{p(\mathrm{A})\cdot p(1\,e\,1\,sola/\mathrm{A})}{p(\mathrm{A})\cdot p(1\,e\,1\,sola/\mathrm{A})+p(\mathrm{B})\cdot p(1\,e\,1\,sola/\mathrm{B})+p(\mathrm{C})\cdot p(1\,e\,1\,sola/\mathrm{C})+p(\mathrm{D})\cdot p(1\,e\,1\,sola/\mathrm{D})}=$$

$$=\frac{p(\mathrm{A})\cdot p(\overline{\mathrm{B}}\overline{\mathrm{C}}\overline{\mathrm{D}})}{p(\mathrm{A})\cdot p(\overline{\mathrm{B}}\overline{\mathrm{C}}\overline{\mathrm{D}})+p(\mathrm{B})\cdot p(\overline{\mathrm{A}}\overline{\mathrm{C}}\overline{\mathrm{D}})+p(\mathrm{C})\cdot p(\overline{\mathrm{A}}\overline{\mathrm{B}}\overline{\mathrm{D}})+p(\mathrm{D})\cdot p(\overline{\mathrm{A}}\overline{\mathrm{B}}\overline{\mathrm{C}})}=$$

$$=\frac{\frac{1}{2}\cdot\left(\frac{2}{3}\cdot\frac{3}{4}\cdot\frac{4}{5}\right)}{\frac{1}{2}\cdot\left(\frac{2}{3}\cdot\frac{3}{4}\cdot\frac{4}{5}\right)+\frac{1}{3}\cdot\left(\frac{1}{2}\cdot\frac{3}{4}\cdot\frac{4}{5}\right)+\frac{1}{4}\cdot\left(\frac{1}{2}\cdot\frac{2}{3}\cdot\frac{4}{5}\right)+\frac{1}{5}\cdot\left(\frac{1}{2}\cdot\frac{2}{3}\cdot\frac{3}{4}\right)}=\frac{\frac{24}{120}}{\frac{24}{120}+\frac{12}{120}+\frac{8}{120}+\frac{6}{120}}=\frac{24}{50}=48\%$$

NOTA: abbiamo scritto semplicemente $p(1\,e\,1\,sola/\mathrm{A})=p(\overline{\mathrm{B}}\overline{\mathrm{C}}\overline{\mathrm{D}})$ anziché $p(1\,e\,1\,sola/\mathrm{A})=p(\overline{\mathrm{B}}\overline{\mathrm{C}}\overline{\mathrm{D}}/\mathrm{A})$ per il fatto che le prestazioni di B, C e D in un dato tiro non sono condizionate dall'esito di A

ESERCIZI sul Teorema di Bayes

1) In un'urna U1 ci sono 2 palline Rosse e 1 Verde; in U2, 1 Rossa e 3 Blu; in U3, 2 Rosse e 3 Verdi.
Se si pesca da un'urna a caso non conoscendo di che urna si tratta, e la pallina estratta risulta Rossa, valutare la probabilità che l'urna di provenienza sia U3.

2) Le statistiche di un tribunale di provincia, relative ai processi ultimati nei trascorsi 15 anni, evidenziano che, fra gli accusati di un reato penale, il 24% era stato trattenuto in custodia cautelare (=carcerazione preventiva), gli altri lasciati a piede libero in attesa del processo.
Dei sottoposti a carcerazione preventiva, l'80% aveva poi avuto una sentenza di condanna definitiva, mentre fra gli altri soltanto il 64% erano stati riconosciuti colpevoli.
Determina, per un condannato di reato penale preso a caso, la probabilità di aver subito la custodia cautelare.

3) In un club di tifosi, i maschi sono il 75% e la metà di loro fuma. Fra le femmine, invece, fuma solo il 25%.
Che probabilità c'è per una persona presa a caso fra gli iscritti, di essere fumatore/fumatrice?
E presa a caso una persona che fuma in quel club, che probabilità c'è che si tratti di una donna?

4) Fra i fumatori di una certa città, il 60% acquista la marca "Bravo Furbo", e fra questi l'80% sono maschi.
Fra coloro che non comprano le sigarette "Bravo Furbo", la maggioranza (75%) è di femmine.
Presa una fumatrice a caso in quella città, che probabilità c'è che acquisti le "Bravo Furbo"?

5) In un paese asiatico, la probabilità che una radiolina della marca A sia difettosa è bassa: 0,1%.
La marca B, unica concorrente di A in quella nazione, fa ancora meglio: probabilità dello 0,05%.
Sul mercato, tuttavia, la marca A risulta prevalere, col 60% degli acquisti, perché la linea dei suoi prodotti è più carina. Che probabilità ha una radiolina perfettamente funzionante presa a caso, di essere della marca A?

6) Imposta un foglio elettronico in modo che l'utente possa inserire le probabilità relative ad un problema risolubile tramite la formula di Bayes (con $n=2$ e anche con $n=3$) e gli venga calcolata la risposta.

		p(H1/E)	P(H2/E)
p(H1)=	0,3	0,461538	0,538462
p(H2)=	0,7		
p(E/H1)=	0,8		
p(E/H2)=	0,4		

I dati vanno introdotti nelle celle ombreggiate

Dal sito http://classweb.gmu.edu *della George Mason University di Washington, USA, ecco due bei quesiti, uno sull'educazione dei figli e un altro sugli incidenti stradali.*

7) In una certa nazione è noto che il 20% delle madri suole sculacciare i figli indisciplinati. L'85% delle madri che applicano questa pratica fanno uso di Valium, contro il 25% soltanto delle madri che non sculacciano.
Questi dati sono tali da far supporre che il Valium induca le madri a sculacciare i loro figli? Discutine.

8) Una compagnia di assicurazioni auto prevede per i guidatori giovani una polizza più alta, in quanto questo gruppo tende ad avere un numero maggiore di incidenti. La compagnia distingue le età in 3 gruppi:
A (sotto i 25 anni, 22% di tutti i suoi assicurati), B (25-39 anni, 43%), C (da 40 anni in su).
I dati mostrano che in media ogni anno le percentuali di assicurati che hanno un incidente sono:
11% per il gruppo A, 3% per il B, 2% per il C.
   a) Che percentuale di assicurati ci si attende abbia un incidente nei prossimi 12 mesi?
   b) Se un assicurato X ha appena avuto un incidente, che probabilità c'è che abbia meno di 25 anni?

RISPOSTE

1) $$p(\mathrm{U3/R})=\frac{p(\mathrm{U3})\cdot p(\mathrm{R/U3})}{p(\mathrm{U1})\cdot p(\mathrm{R/U1})+p(\mathrm{U2})\cdot p(\mathrm{R/U2})+p(\mathrm{U3})\cdot p(\mathrm{R/U3})}=\frac{1/3\cdot 2/5}{1/3\cdot 2/3+1/3\cdot 1/4+1/3\cdot 2/5}=\ldots=\frac{24}{79}\approx 30\%$$

2) $$p(\mathrm{CC/cond})=\frac{p(\mathrm{CC})\cdot p(\mathrm{cond/CC})}{p(\mathrm{CC})\cdot p(\mathrm{cond/CC})+p(\overline{\mathrm{CC}})\cdot p(\mathrm{cond}/\overline{\mathrm{CC}})}=\frac{0,24\cdot 0,80}{0,24\cdot 0,80+0,76\cdot 0,64}=\frac{0,192}{0,6784}\approx 0,283$$

3) Il 43,75%; intorno al 14%   4) Il calcolo dà un valore prossimo al 28,6%   5) Circa il 60%

7) $$p(\mathrm{S/V})=\frac{p(\mathrm{S})\cdot p(\mathrm{V/S})}{p(\mathrm{S})\cdot p(\mathrm{V/S})+p(\overline{\mathrm{S}})\cdot p(\mathrm{V}/\overline{\mathrm{S}})}=\frac{0,20\cdot 0,85}{0,20\cdot 0,85+0,80\cdot 0,25}=\frac{0,17}{0,17+0,20}=\frac{0,17}{0,37}\approx 0,46$$

$$p(\mathrm{S}/\overline{\mathrm{V}})=\frac{p(\mathrm{S})\cdot p(\overline{\mathrm{V}}/\mathrm{S})}{p(\mathrm{S})\cdot p(\overline{\mathrm{V}}/\mathrm{S})+p(\overline{\mathrm{S}})\cdot p(\overline{\mathrm{V}}/\overline{\mathrm{S}})}=\frac{0,20\cdot 0,15}{0,20\cdot 0,15+0,80\cdot 0,75}=\frac{0,03}{0,03+0,60}=\frac{0,03}{0,63}\approx 0,05$$

Dai dati emerge che se una donna assume Valium, è senz'altro molto più incline a sculacciare i propri figli, ma… attenzione! … Sarà il Valium in sé a favorire questo comportamento, o piuttosto la depressione di cui plausibilmente soffrono queste donne, dato che fanno uso di Valium?

8a) $$\frac{22}{100}\cdot\frac{11}{100}+\frac{43}{100}\cdot\frac{3}{100}+\frac{35}{100}\cdot\frac{2}{100}=\frac{242+129+70}{10000}=\frac{441}{10000}\approx 4,4\%$$

8b) $$p(\mathrm{A/I})=\frac{p(\mathrm{A})\cdot p(\mathrm{I/A})}{p(\mathrm{A})\cdot p(\mathrm{I/A})+p(\mathrm{B})\cdot p(\mathrm{I/B})+p(\mathrm{C})\cdot p(\mathrm{I/C})}=\frac{0,22\cdot 0,11}{0,22\cdot 0,11+0,43\cdot 0,03+0,35\cdot 0,02}\approx 55\%$$