

Tabella 9.3 – Probabilità della somma delle facce di due dadi lanciati contemporaneamente.

s_m	2	3	4	5	6	7	8	9	10	11	12
p_m	1/36	1/18	1/12	1/9	5/36	1/6	5/36	1/9	1/12	1/18	1/36

Lanciamo, adesso, i due dadi in un contenitore che non renda visibili le loro facce, estraiamone uno a caso e osserviamo il valore B della sua faccia, che è, in generale, aleatorio. Se $B = \begin{smallmatrix} \blacksquare & \blacksquare \\ \blacksquare & \blacksquare \end{smallmatrix}$, quanto vale l'incertezza sul valore della somma (che, prima di osservare B , vale $H(S) = 3.274$ bit)? È chiaro che l'avere osservato $\begin{smallmatrix} \blacksquare & \blacksquare \\ \blacksquare & \blacksquare \end{smallmatrix}$ cambia il gioco: per esempio, i risultati 2, 3, 4, 11 e 12 non sono più possibili e, quindi, le probabilità degli altri risultati, *condizionatamente all'osservazione di* $\begin{smallmatrix} \blacksquare & \blacksquare \\ \blacksquare & \blacksquare \end{smallmatrix}$, sono cambiate. Possiamo calcolare, allora, le probabilità condizionate dei simboli di sorgente e, da queste, ricavare l'entropia condizionata $H(S|B = \begin{smallmatrix} \blacksquare & \blacksquare \\ \blacksquare & \blacksquare \end{smallmatrix})$, che rappresenta proprio l'incertezza cercata. È chiaro che le probabilità condizionate di 2, 3, 4, 11 e 12 sono nulle, mentre quelle di 5, 6, 7, 8, 9 e 10 sono tutte pari a $1/6$, in quanto coincidono con le probabilità che la faccia del dado coperto sia rispettivamente $\begin{smallmatrix} \blacksquare & \blacksquare \\ \blacksquare & \blacksquare \end{smallmatrix}$, $\begin{smallmatrix} \blacksquare & \blacksquare \\ \blacksquare & \blacksquare \end{smallmatrix}$, $\begin{smallmatrix} \blacksquare & \blacksquare \\ \blacksquare & \blacksquare \end{smallmatrix}$, $\begin{smallmatrix} \blacksquare & \blacksquare \\ \blacksquare & \blacksquare \end{smallmatrix}$ e $\begin{smallmatrix} \blacksquare & \blacksquare \\ \blacksquare & \blacksquare \end{smallmatrix}$. L'entropia condizionata $H(S|B = \begin{smallmatrix} \blacksquare & \blacksquare \\ \blacksquare & \blacksquare \end{smallmatrix})$ è quella di una variabile a 6 valori equiprobabili e, quindi, è pari a $\log_2(6) \cong 2.585$. Se ripetiamo il ragionamento immaginando di estrarre B e constatare che vale un qualunque altro valore $b_i \neq \begin{smallmatrix} \blacksquare & \blacksquare \\ \blacksquare & \blacksquare \end{smallmatrix}$, otteniamo sempre lo stesso risultato $\log_2(6)$. A questo punto possiamo calcolare l'incertezza residua *media* $H(S|B)$ come media dei vari valori di $H(S|B = b_i)$ rispetto alle probabilità di b_i :

$$H(S|B) = \sum_{i=1}^6 H(S|B = b_i) \Pr\{B = b_i\} = \sum_{i=1}^6 \log_2(6) \frac{1}{6} = \log_2(6) \cong 2.585 \quad (\text{E9.8.1})$$

Aver osservato uno dei due dadi fa naturalmente *diminuire* l'incertezza sul valore della somma delle facce da 3.274 a 2.585 bit, cioè apporta una certa quantità di informazione collaterale. Questa informazione collaterale che B apporta su S può essere valutata in modo semplice: è proprio la quantità di cui è stata diminuita l'incertezza dopo l'osservazione del dado B e cioè

$$I(S, B) \triangleq H(S) - H(S|B) \cong 0.689 \text{ bit} \quad (\text{E9.8.2})$$

Questa quantità si chiama *informazione mutua* tra S e B , rappresenta l'informazione che mediamente otteniamo riguardo i valori di una sorgente quando osserviamo i valori dell'altra. Torneremo su questo concetto più avanti, quando affronteremo la questione della compressione con perdita.

La terminologia adottata in questo paragrafo merita alcuni chiarimenti. Quando abbiamo parlato di *simboli binari* delle parole di codice o di simboli di una *sorgente binaria*, abbiamo detto che tali simboli sono *cifre binarie*. Quando abbiamo definito, invece, una misura di informazione, questa è stata espressa in *bit*. In questo contesto, il termine *bit* non ha l'accezione di *cifra binaria*. Dall'**Esempio 9.6** comprendiamo che una cifra binaria trasporta in generale *meno* di un bit di informazione, anche se nel linguaggio comune (anche tecnico) di rado usiamo la corretta nomenclatura; infatti, utilizziamo la parola *bit* sia per le cifre binarie sia per la misura di informazione. Le due cose coincidono quando la sorgente