

Un altro modo, però, di risolvere la domanda ② è osservare che E_+ e E_- sono complementari: uno dell'altro, per cui $Pr(E_-) = 1 - Pr(E_+) = 1 - 0.82 = 0.18$.

Esercizio

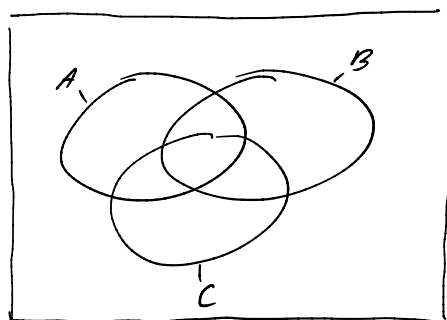
Sono dati tre eventi A, B, C. Gli eventi A e B sono indipendenti, mentre l'evento C è incompatibile con $A \cup B$. Le probabilità di A, B, C sono rispettivamente 0.4, 0.5, 0.1.

– Calcola la probabilità di $A \cup B \cup C$.

– Calcola la probabilità che avvenga uno solo fra A e B condizionata al complementare di C.

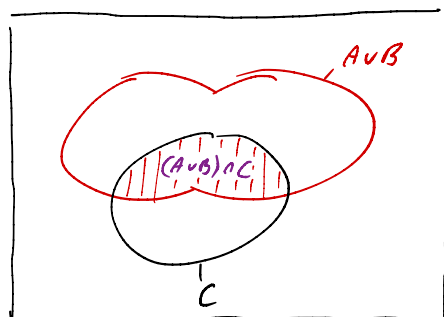
fine 1.30

Soluzione

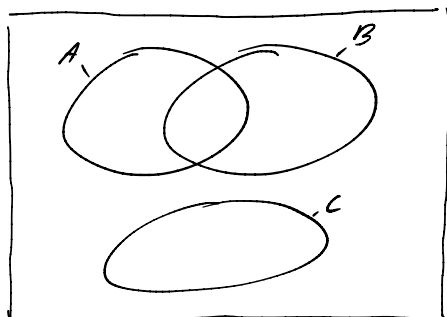


Nella situazione più generale, tre eventi si intersecano come nella figura a sinistra. Nel nostro caso, però, sappiamo che C è incompatibile con $A \cup B$,

ossia $Pr((A \cup B) \cap C) = 0$, che si traduce nel fatto che la regione tratteggiata qui a destra ha probabilità nulla. Pos-



siamo quindi sempli-

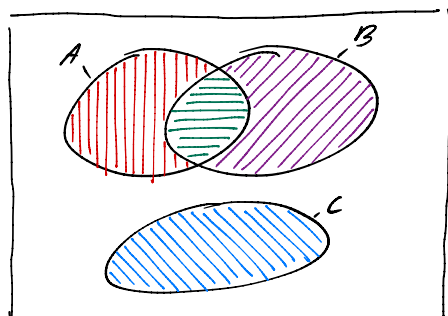
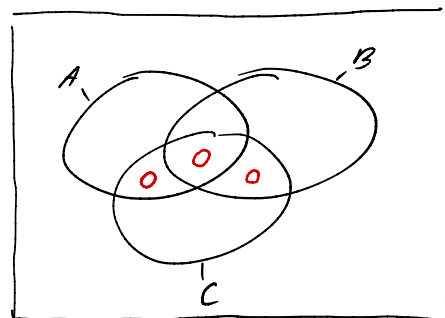


ficare il disegno supponendo che C non tocchi $A \cup B$. Ora lo spazio degli eventi è suddiviso in cinque regioni. Calcolando la probabilità di queste cinque

SITUAZIONE DELL'ESERCIZIO

regioni avremo tutta l'informazione che serve per rispon-

dere a qualunque domanda probabilistica a proposito degli eventi A , B e C . Nota: il metodo funziona anche se non semplifichiamo il disegno (come invece abbiamo fatto). Lavorando sul diagramma di partenza avremmo semplicemente ottenuto, dalla condizione di incompatibilità, che queste tre regioni hanno probabilità 0.



Tornando alla nostra figura, già sappiamo che $Pr(\text{green}) = Pr(C) = 0.1$. Inoltre $\text{green} = A \cap B$, quindi, siccome A e B sono indipendenti,

$$Pr(\text{green}) = Pr(A \cap B) = Pr(A) \cdot Pr(B) = 0.2.$$

Adesso possiamo calcolare $Pr(\text{red})$ osservando che A è l'unione delle due regioni incompatibili red e green , quindi:

$$Pr(\text{red}) = Pr(A) - Pr(\text{green}) = 0.2$$

Allo stesso modo

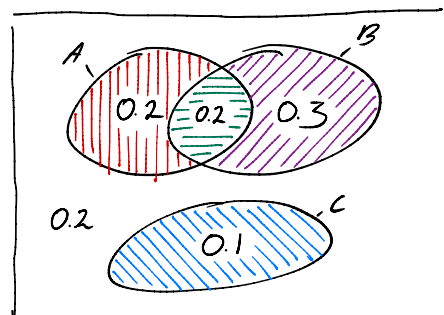
$$Pr(\text{purple}) = Pr(B) - Pr(\text{green}) = 0.3$$

Infine, la probabilità della zona esterna è

$$Pr(\text{white}) = 1 - Pr(\text{red}) - Pr(\text{green}) - Pr(\text{purple}) - Pr(\text{blue}) = 0.2$$

Questa è quindi la situazione descritta dall'esercizio.

La prima domanda ci chiede quale sia la probabilità di $A \cup B \cup C$. Questo



evento coincide con l'unione delle aree colorate, quindi

$$\boxed{Pr(A \cup B \cup C)} = Pr(\text{red} \cup \text{green} \cup \text{purple} \cup \text{blue}) = \\ = Pr(\text{red}) + Pr(\text{green}) + Pr(\text{purple}) + Pr(\text{blue}) = \boxed{0.8}$$

Per affrontare la seconda domanda dobbiamo capire esattamente che evento sia

$E :=$ "si verifica uno solo fra A e B "

Egaminiamo le aree una per una

Nell'area  A si verifica e B no, quindi E si verifica.

Nell'area  A e B si verificano entrambi, quindi E non si verifica.

Nell'area  B si verifica e A no, quindi E si verifica.

Nell'area  e nell'area  né A né B si verifica, per cui neanche E si verifica.

In conclusione $E = \text{red} \cup \text{purple}$. È facile identificare $\bar{E}$:

$\bar{E} = \text{red} \cup \text{green} \cup \text{purple} \cup \text{white}$. Possiamo allora calcolare

$$\boxed{Pr(E|\bar{E})} = \frac{Pr(E \cap \bar{E})}{Pr(\bar{E})} = \frac{Pr((\text{red} \cup \text{purple}) \cap (\text{red} \cup \text{green} \cup \text{purple} \cup \text{white}))}{Pr(\text{red} \cup \text{green} \cup \text{purple} \cup \text{white})} = \\ = \frac{Pr(\text{red} \cup \text{purple})}{Pr(\text{red} \cup \text{green} \cup \text{purple} \cup \text{white})} = \frac{Pr(\text{red}) + Pr(\text{purple})}{\underbrace{Pr(\text{red}) + Pr(\text{green}) + Pr(\text{purple}) + Pr(\text{white})}} = \\ = \frac{0.5}{0.9} = \boxed{\frac{5}{9}} \quad \text{Nota 1: questo } \uparrow \text{ è semplicemente } 1 - Pr(\text{blue}).$$

Nota 2: l'intersezione è l'area comune, per esempio

$$(\text{green} \cup \text{white} \cup \text{red}) \cap (\text{red} \cup \text{blue} \cup \text{white}) = \text{white} \cup \text{red}$$

Algebra degli insiemi

Quando si ragiona a proposito di insiemi (o eventi che sono semplicemente insiemi di eventi elementari) conviene tenere a mente alcune formule che riassumono in regole algebriche certi ragionamenti che si incontrano di frequente.

- Regole ovvie

$$A \cup B = B \cup A \quad A \cup (B \cap C) = (A \cup B) \cup C = A \cup B \cup C$$

$$A \cap B = B \cap A \quad A \cap (B \cup C) = (A \cap B) \cap C = A \cap B \cap C$$

Ossia: l'ordine delle unioni non cambia il risultato

l'ordine delle intersezioni non cambia il risultato

Attenzione: in questa regola non si mescolano unioni e intersezioni!

- Regole non ovvie

$\overline{\overline{A}} = A$ Il complementare del complementare di un insieme è l'insieme stesso

$$\overline{A \cup B} = \overline{A} \cap \overline{B} \quad \overline{A \cap B} = \overline{A} \cup \overline{B} \quad \text{"leggi di De Morgan"}$$

Il complementare di un'unione è l'intersezione dei complementari e, viceversa, il complementare di un'intersezione è l'unione dei complementari.

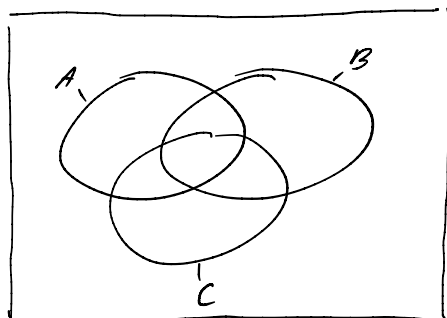
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

Queste ultime regole si chiamano leggi distributive per via dell'affinità formale con la proprietà distributiva

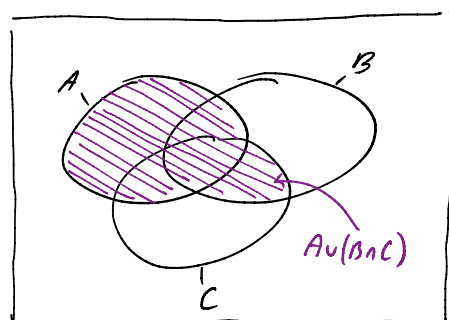
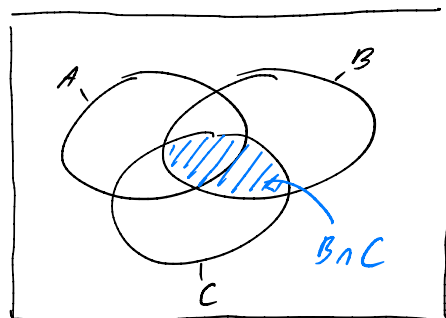
$$x \cdot (y + z) = (x \cdot y) + (x \cdot z)$$

Come esempio vediamo il ragionamento che giustifica la legge $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$. Partiamo dalla situa-



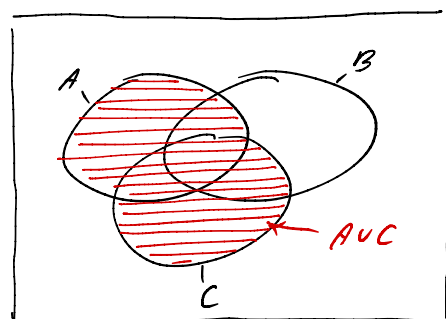
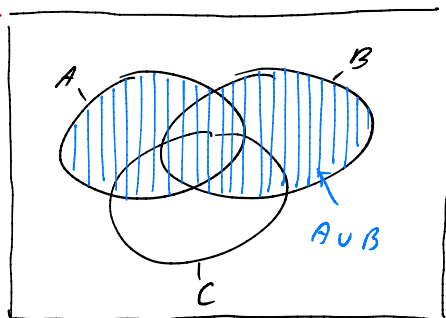
zione generale di tre insiemi A, B e C , e determiniamo quali aree costituiscono i due membri dell'equazione. Studiando

il membro di sinistra, vediamo che $(B \cap C)$ è l'area tratteggiata a destra, quindi $A \cup (B \cap C)$ è

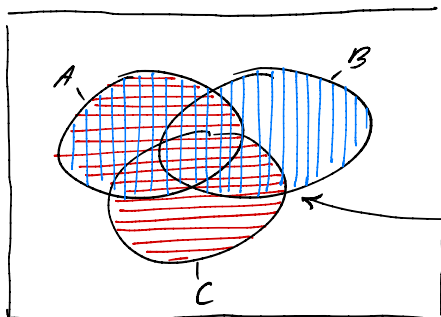


l'area indicata nella figura di sinistra.

Se ora tratteggiamo $(A \cup B)$ e $(A \cup C)$,



quindi sovrapponiamo le due figure per rivelare l'intersezione (l'area coperta da entrambi i tratteggi) vediamo che questa coincide con l'area $A \cup (B \cap C)$ del diagramma precedente.



$$\text{cross-hatch} = (A \cup B) \cap (A \cup C)$$

Il trucco del complementare

Consideriamo il seguente esercizio: si lanciano 6 monete indipendenti ed egue ($p. \text{ testa} = p. \text{ croce} = 0.5$).
Che probabilità c'è che almeno uno dei risultati sia testa?

Potremmo affrontare l'esercizio elencando tutte le possibili sequenze di 6 risultati, ma queste sono $2^6 = 64$ e c'è un metodo più rapido.

Trucco del complementare: l'evento richiesto

$E = \text{"c'è almeno una testa fra i risultati"}$

è il complementare di

$\bar{E} = \text{"tutti i risultati sono croce"}$

calcoliamo quindi $Pr(\bar{E})$ e troveremo la probabilità richiesta mediante la formula $Pr(E) = 1 - Pr(\bar{E})$.

Per calcolare $Pr(\bar{E})$ ci basta osservare che

$$\bar{E} = C_1 \cap C_2 \cap C_3 \cap C_4 \cap C_5 \cap C_6$$

dove indichiamo con C_n l'evento "l' n -esima moneta fa croce" (infatti $\bar{E}$ si verifica se si verifica C_1 , e si verifica C_2 , e etc. fino a C_6 : insieme, la e è rappresentata dall'intersezione).

Per dato dell'esercizio, gli eventi C_i sono indipendenti,

$$P_n(\bar{E}) = P_n(C_1 \cap C_2 \cap \dots \cap C_6) = P_n(C_1) \cdot P_n(C_2) \cdot \dots \cdot P_n(C_6) = \left(\frac{1}{2}\right)^6$$

Di conseguenza $\boxed{P_n(E)} = 1 - P_n(\bar{E}) = 1 - \left(\frac{1}{2}\right)^6 = \boxed{\frac{63}{64}}$.

Esercizio

Calcolare la probabilità che su tre lanci di un dado da gioco si presentino almeno un 5 oppure un 6.

Soluzione

Procedendo come nell'esercizio precedente otteniamo

$$\text{probabilità} = 1 - \left(\frac{4}{6}\right)^{\overset{\text{numero dei lanci}}{3}} = \frac{19}{27}$$

$\uparrow$
probabilità che ciascuno dei
lanci non scaccia né 5 né 6

Esercizio

A.3 – Vengono scelte, a caso, indipendentemente e uniformemente sulla popolazione, cinque tredicenni e cinque quindicenni. Che probabilità c'è che la ragazza più alta di queste dieci superi i 170 cm di statura?

Dati: la probabilità che una quindicenne scelta uniformemente superi i 170 cm di statura è del 10%, la probabilità che una tredicenne scelta uniformemente superi i 170 cm è del 3%.

Soluzione

L'evento richiesto

E = "la ragazza più alta supera i 170 cm di statura"

può essere espresso più convenientemente con le parole

E = "una delle ragazze supera i 170 cm di statura"

l'evento complementare è quindi

$\bar{E}$ = "tutte le ragazze hanno statura < 170 cm"

e possiamo vedere quest'ultimo evento come l'intersezione

$\bar{E} = Q \cap T$ dei due eventi indipendenti:

Q = "le cinque quindicenni hanno tutte statura < 170 cm"

T = "le cinque tredicenni hanno tutte statura < 170 cm"

Le probabilità di questi ultimi si calcolano al modo degli esercizi precedenti:

$$Pr(Q) = (1 - 0.1)^5 \quad \begin{array}{l} \text{numero di quindicenni} \\ \uparrow \\ \text{pr. statura} \geq 170 \text{ cm} \\ \text{è } 10\% \text{ per le quindicenni} \end{array}$$

$$Pr(T) = (1 - 0.03)^5 \quad \begin{array}{l} \text{numero di tredicenni} \\ \uparrow \\ \text{pr. statura} \geq 170 \text{ cm} \\ \text{è } 3\% \text{ per le tredicenni} \end{array}$$

Quindi

$$[Pr(\bar{E})] = 1 - Pr(E) = 1 - Pr(Q) \cdot Pr(T) = [1 - 0.9^5 \cdot 0.97^5] \approx 0.493$$

Esercizio Si lancia quattro volte (indipendenti) un dado da gioco. Che probabilità c'è che fra i quattro risultati ve ne siano due uguali? Nota: in realtà abbiamo risolto questo esercizio alla lezione precedente, però la soluzione, che è un po' complicata, si capisce meglio alla luce degli esempi visti qui.

Soluzione

Anche in questo caso conviene indagare l'evento complementare, ossia che i quattro risultati siano tutti diversi.

Consideriamo i tre eventi

D_2 = "il secondo dado è diverso dal primo"

D_3 = "il terzo dado è diverso da ciascuno dei primi due"

D_4 = "il quarto dado è diverso da ciascuno dei primi tre"

Il complementare dell'evento cercato è $\bar{E} = D_4 \cap D_3 \cap D_2$.

Chiaramente $\Pr(D_2) = \frac{5}{6}$ (detto x_1 il primo risultato, D_2 si verifica per i 5 su 6 possibili risultati del secondo dado che differiscono da x_1).

Allo stesso modo $\Pr(D_3 | D_2) = \frac{4}{6}$ (infatti, se assumiamo che D_2 si verifichi, allora i due primi risultati, x_1 e x_2 , sono diversi, quindi D_3 si verifica nei 4 casi su 6 in cui il terzo risultato non è né x_1 , né x_2).

Si come

$$\Pr(D_3 | D_2) = \frac{\Pr(D_3 \cap D_2)}{\Pr(D_2)}$$

Otteniamo $\Pr(D_3 \cap D_2) = \Pr(D_3 | D_2) \cdot \Pr(D_2) = \frac{4}{6} \cdot \frac{5}{6}$

Finalmente $\Pr(D_4 | (D_3 \cap D_2)) = \frac{3}{6}$ (per lo stesso ragionamento)
queste parentesi non sono necessarie

Quindi

$$\begin{aligned} \boxed{\Pr(E)} &= 1 - \Pr(D_4 \cap D_3 \cap D_2) = 1 - \Pr(D_4 | D_3 \cap D_2) \cdot \Pr(D_3 \cap D_2) = \\ &= 1 - \frac{3}{6} \cdot \frac{4}{6} \cdot \frac{5}{6} = \boxed{\frac{13}{18}} \end{aligned}$$

fine l.31

Esercizio non esposto a lezione

Esercizio B DALLA PROVA SCRITTA DEL 31-I-2019

Il *mean time between failures* (MTBF) è un parametro di affidabilità degli apparati tecnologici. Il MTBF di un dispositivo è il tempo medio che intercorre fra due guasti successivi del dispositivo, supponendo che ogni guasto sia immediatamente riparato. Indichiamo con $MTBF_X$ il MTBF di un certo sistema X; e indichiamo con p_X la probabilità che il sistema X si guasti in ogni data ora di funzionamento (probabilità che assumiamo costante). Allora vale la relazione (dove $h = \text{"ore"}$)

$$p_X = 1 - e^{-\frac{1h}{MTBF_X}}$$

Supponiamo, adesso, di avere una macchina M che è costituita da due parti indipendenti fra loro: A e B. Se una qualunque delle due parti si guasta, allora la macchina, nel suo complesso, è guasta. Il MTBF delle parti è:

$$MTBF_A = 100\,000\text{ h} \qquad MTBF_B = 33\,000\text{ h}$$

- B.1 – Indichiamo con $\mathcal{A}$, $\mathcal{B}$ e $\mathcal{M}$ gli eventi che rispettivamente A, B e M si guastino in una certa ora di lavoro. Quale relazione lega questi eventi? Usa la formula data nel testo per calcolare p_A e p_B , desumi p_M , quindi determina $MTBF_M$. Esprimi il tuo risultato con due cifre significative di precisione. (*Suggerimento: verifica il risultato con la formula al punto B.3, se viene diverso allora hai sbagliato qualcosa.*)
- B.2 – Supponiamo che l'evento che una macchina si guasti durante una certa ora sia indipendente dagli eventi che si guasti in una qualunque altra ora. Quale è la probabilità che la macchina M subisca almeno un guasto durante 24 ore di lavoro? Quante ore di lavoro servono perché vi sia una probabilità del 50% che M si guasti?
- B.3 – La formula seguente lega il MTBF di M a quelli di A e di B.

$$MTBF_M = \frac{1}{\frac{1}{MTBF_A} + \frac{1}{MTBF_B}}$$

Dimostra questa formula (per esempio eseguendo i calcoli del punto B.1 in maniera letterale, ma non è l'unico modo). Quale formula va usata per una macchina costituita da tre parti indipendenti, e perché?

SOLUZIONE

- B.1 – L'evento $\mathcal{M}$ si verifica quando almeno uno fra $\mathcal{A}$ e $\mathcal{B}$ si verifica, si tratta quindi, per definizione, dell'unione $\mathcal{A} \cup \mathcal{B}$. Le probabilità p_A e p_B sono semplicemente date dalla formula del testo

$$p_A = 1 - e^{-\frac{1h}{100\,000\text{ h}}} \sim 1.000 \cdot 10^{-5}$$

$$p_B = 1 - e^{-\frac{1h}{33\,000\text{ h}}} \sim 3.030 \cdot 10^{-5}$$

Per la probabilità p_M abbiamo

$$p_M = 1 - e^{-\frac{1h}{MTBF_M}}$$

$$e^{-\frac{1h}{MTBF_M}} = 1 - p_M$$

$$-\frac{1h}{MTBF_M} = \log_e(1 - p_M)$$

$$MTBF_M = \frac{1h}{-\log(1 - p_M)} \sim 25\,000$$

** Nota alla pag. seguente*

□^c □ hanno lo stesso significato (il complementare)

B.2 – La probabilità che la macchina non subisca guasti durante n ore è, per l'indipendenza, la probabilità che non ci siano guasti in una data ora, elevata alla n -esima potenza. Si tratta, infatti, di fare l'intersezione degli eventi indipendenti "non ci sono guasti nella prima ora", e "... nella seconda ora", etc. Quindi, detta p_n la probabilità di avere almeno un guasto in un periodo di n ore, abbiamo

$$1 - p_n = (1 - p_M)^n$$

da cui si ricava immediatamente

$$p_{24} = 1 - (1 - p_M)^{24} \sim 9.668 \cdot 10^{-4}$$

Il numero di ore necessarie per avere una probabilità di guasto del 50% è il valore di n tale che $p_n = 0.5$, ossia

$$(1 - p_M)^n = 0.5$$

$$n = \frac{\log(0.5)}{\log(1 - p_M)} \sim 17\,200$$

$\log((1 - p_M)^n) = \log(0.5)$
 $n \cdot \log(1 - p_M)$

B.3 – Riprendendo dalla formula

$$p_M = 1 - (1 - p_A)(1 - p_B)$$

ricavata al punto B.1, e sostituendo le espressioni per p_A , p_B , e p_M in termini dei MTBF, otteniamo

$$1 - e^{-\frac{1h}{MTBF_M}} = 1 - e^{-\frac{1h}{MTBF_A}} e^{-\frac{1h}{MTBF_B}}$$

$$e^{-\frac{1h}{MTBF_M}} = e^{-\left(\frac{1h}{MTBF_A} + \frac{1h}{MTBF_B}\right)}$$

da cui, prendendo i logaritmi e moltiplicando per $-1h^{-1}$

$$\frac{1}{MTBF_M} = \frac{1}{MTBF_A} + \frac{1}{MTBF_B}$$

e la formula richiesta segue immediatamente calcolando il reciproco di ambo i membri.

Se la macchina fosse costituita da tre parti A, B, e C, allora dovremmo calcolare la probabilità dell'unione dei tre eventi corrispondenti al guasto di una delle parti, che, come prima, si fa calcolando la probabilità dell'intersezione dei complementari. Otteniamo così

$$1 - p_M = (1 - p_A)(1 - p_B)(1 - p_C)$$

e, sostituendo e semplificando

$$\frac{1}{MTBF_M} = \frac{1}{MTBF_A} + \frac{1}{MTBF_B} + \frac{1}{MTBF_C}$$

$$MTBF_M = \frac{1}{\frac{1}{MTBF_A} + \frac{1}{MTBF_B} + \frac{1}{MTBF_C}}$$

* Attenzione: questo calcolo nasconde una piccola trappola. L'esercizio chiede $MTBF_n$ con due cifre significative di precisione: secondo la regola generale, se facciamo il calcolo con tre cifre significative (sebbene quattro sarebbero meglio) dovremmo ottenere il risultato corretto. Proviamo!

$$p_n = 4.030 \cdot 10^{-5}$$

$$MTBF_n = \frac{1 \text{ h}}{-\log(1-p_n)} = \frac{1 \text{ h}}{-\log(1-4.030 \cdot 10^{-5})} \approx \frac{1 \text{ h}}{-\log(0.999)} \approx 1000 \text{ h}$$

$0,9999597$ troncato* a 3 c.s. ↑

* il troncamento stesso è imperfetto: converrebbe approssimare.

Questo risultato è **SBAGLIATO**.

Come capire, intanto, che qualcosa non torna? Nel contesto dell'esercizio, non è difficile: basta seguire il suggerimento e applicare la formula al punto B.3.

$$MTBF_n = \frac{1}{\frac{1}{MTBF_A} + \frac{1}{MTBF_B}} = \frac{1}{\frac{1}{100'000 \text{ h}} + \frac{1}{35'000 \text{ h}}} \approx 25'000 \text{ h}$$

Se non avessimo avuto il suggerimento, però, avremmo potuto ripetere il calcolo con più cifre. Per esempio, usando quattro cifre significative per $1-4.030 \cdot 10^{-5}$ otteniamo

$$MTBF_n \approx \frac{1 \text{ h}}{-\log(0.9999)} \approx 10'000 \text{ h}$$

Inaspettatamente, questo risultato differisce da quello ottenuto con tre cifre. Chiaramente il risultato precedente è screditato.

to, ma anche di questo c'è poco da fidarsi, perché, come il conto è cambiato aggiungendo una cifra significativa, potrebbe farlo di nuovo aggiungendone un'altra. Proviamo con cinque...

$$MTBF_n \approx \frac{1 \text{ h}}{-\log(0.99996)} \approx 25000 \text{ h}$$

$\approx$ approssimazione corretta

il successo di nuovo. Ora, però, aggiungere altre cifre non cambia il risultato significativamente

$$MTBF_n \approx \frac{1 \text{ h}}{-\log(0.9999597)} \approx 24813 \text{ h}$$

Quindi il valore corretto, approssimato a due cifre significative, è 25'000 h. Però, perché la teoria, secondo la quale ci aspettavamo di ottenere due cifre significative del risultato facendo i calcoli con tre — e anche con quattro — cifre nei passaggi intermedi, ha fallito?

La lezione che è importante imparare da questo calcolo è che le cifre significative non sono che uno strumento rozzo, che, sebbene di rado, può rivelarsi imperfetto.

Ma cosa è andato storto? Pensiamo al procedimento aritme-

$\begin{array}{r} 1.000000000 \\ 0.00004030 \\ \hline 0.99995970 \end{array}$ tico per calcolare $1 - 4.030 \cdot 10^{-5}$. Come si vede l'informazione contenuta nelle cifre 4.030 produce le cifre 5970 del risultato, mentre i 9 precedenti agiscono semplicemente come dei segna-posto: sarebbero stati

gli stessi anche se al posto della mantissa **4.030** ci fosse stata qualunque altra cosa. Quindi va da sé che, troncando la differenza **0.99995970** ai 9 soltanto, si getta via tutta l'informazione contenuta nel dato.

Il modo più comune per cadere in una trappola come questa è abbassare la guardia. Uno fa il calcolo con 0.999, ottiene 1000 h, e procede per la sua strada ignaro del fatto che il ponte che ha progettato crollerà, e, quel che è peggio, lui non passerà l'esame.

C'è di peggio: se volendo, erroneamente, lavorare con tre cifre significative avessimo approssimato correttamente il numero 0.99995970 a 1.00 avremmo ottenuto

$$MTBF_m = \dots \approx \frac{1h}{-\log(1.00)} = \frac{1h}{0} \quad \begin{array}{l} \text{ops!} \\ \text{dividere per zero} \\ \text{non ha senso.} \end{array}$$

Questo intoppo avrebbe dovuto metterci subito sull'avviso che qualcosa non sta funzionando come ci aspetteremmo. La reazione sbagliata sarebbe stata scrivere uno **0.99** al posto di quel **1.00** per poter andare avanti. Quando ci si rende conto di non capire cosa stia succedendo, quasi mai affidarsi deliberatamente ad una mossa che si sa essere sbagliata è una buona idea. La reazione giusta è invece di indagare la situazione col ragionamento e con l'esperimento.

