Basi della probabilità

La probabilità si occupa di *fenomeni* aleatori, cioè di un *esperimento* i cui possibili risultati appartengono ad un insieme ben definito e dove l'esito non è prevedibile.
Sia S uno spazio campionario. Una probabilità valida soddisfa i seguenti **assiomi di probabilità**:

1. Le probabilità sono numeri reali non negativi, cioè per tutti gli eventi $E, P(E) \geq 0$.
2. La probabilità dello spazio campione è $1, P(S) = 1$.
3. Le probabilità sono numerabilmente additive: se $UN\_1, UN\_2, \ldots$ sono disgiunti a due a due, allora
   {{< mathjax >}}
   $$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$$
   {{}}
   Teorema 2.1
   Siano A e B eventi dello spazio campionato S.
4. $P(\emptyset) = 0$
5. Se A e B sono disgiunti (∪), allora $P(A \cup B) = P(A) + P(B)$
6. Se $A \subset B$, allora $P(A) \leq P(B)$
7. $0 \leq P(A) \leq 1$
8. $P(A) = 1 - P(\overline{A})$
9. $P(A - B) = P(A) - P(A \cap B)$
10. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
    Esempio:
    Laniamo un dado, che probabilità ho che esca una determinata faccia?
    Usiamo il punto 3.
    {{< mathjax >}}
    $$P(\{1\}) + P(\{2\}) + P(\{3\}) + P(\{4\}) + P(\{5\}) + P(\{6\}) = P(\{1, 2, 3, 4, 5, 6\})$$
    {{}}
    Tutte le probabilità sono uguali e la loro somma è $= 1$, la probabilità che esca una determinata faccia è $\frac{1}{6}$.
    Esempio 2.8:
    Supponiamo che i dadi invece siano due, lo spazio campionato S è dato da:
    {{< mathjax >}}
    $$S = \begin{pmatrix} (1,1),(1,2),(1,3),(1,4),(1,5),(1,6) \\ (2,1),(2,2),(2,3),(2,4),(2,5),(2,6) \\ (3,1),(3,2),(3,3),(3,4),(3,5),(3,6) \\ (4,1),(4,2),(4,3),(4,4),(4,5),(4,6) \\ (5,1),(5,2),(5,3),(5,4),(5,5),(5,6) \\ (6,1),(6,2),(6,3),(6,4),(6,5),(6,6) \end{pmatrix}$$

{{}}
Gli eventi dove "la somma dei due dadi è 6" è rappresentata da:
{{< mathjax >}}

$$E = \{(1,5),(2,4),(3,3),(4,2),(5,1)\}$$

{{}}
La probabilità che la somma dei due dadi sia 6 è data da:
{{< mathjax >}}

$$P(E) = \frac{|E|}{|S|} = \frac{5}{36}$$

{{}}
Sia F l'evento "almeno uno dei due dadi è 2", l'evento è rappresentato da:
{{< mathjax >}}

$$F = \{(2,1),(2,2),(2,3),(2,4),(2,5),(2,6),(1,2),(3,2),(4,2),(5,2),(6,2)\}$$

{{}}
La probabilità di F è: $P(F) = \frac{11}{36}$
$E \cap F = (2,4),(4,2)$ e $P(E \cap F) = \frac{2}{36}$
$P(E \cap F) = P(E) + P(F) - P(E \cap F) = \frac{5}{36} + \frac{11}{36} - \frac{2}{36} = \frac{14}{36}$
$P(\overline{E}) = 1 - P(E) = \frac{31}{36}$
Simulazioni con `sample`

```
sample(x, size, replace = FALSE, prob = NULL)
```

`x` il vettore di elementi dal quale si sta campionando.
`size` il numero di campioni che si voglio ottenere.
`replace` se si stanno usando rimpiazzi o meno.
`prob` un vettore di probabilità o pesi, associato a `x`.
Per ottenere due numeri casuali tra 1 e 10:

```
sample(x = 1:10, size=2)
```

```
## [1] 2 5
```

`sample` non ritorna mai lo stesso valore 2 o più volte, bisogna usare `replace = TRUE`
Esempio 2.9
4 tipi di sangue con diversa probabilità di distribuzione:

```
bloodtypes <- ("O", "A", "B", "AB")
bloodprobs <- (0.45, 0.40, 0.11, 0.04)
sample(x = bloodtypes, suze = 30, prob = bloodprobs, replace = TRUE)
```

```
##  [1] "A"  "O"  "AB" "A"  "O"  "A"  "A"  "O"  "O"  "O"  "O"  "A"  "O"
## [14] "A"  "A"  "A"  "B"  "A"  "A"  "B"  "A"  "O"  "O"  "A"  "A"  "O"
## [27] "O"  "A"  "A"  "O"
```

```
sim_data <- sample(
    x = bloodtypes, size = 10000,
    prob = bloodprobs, replace = TRUE
)
table(sim_data)
```

```
## sim_data
##    A   AB    B    O
## 3998  425 1076 4501
```

```
table(sim_data) / 1000
```

```
## sim_data
##      A     AB      B      O
## 0.3998 0.0425 0.1076 0.4501
```

Esempio 2.10

Supponiamo che due dadi a 6 facce vengano lanciati, sommiamo il risultato.
Effettuiamo questo esperimento su 10000 lanci:

```
die1 <- (x = 1:6, size = 10000, replace = TRUE)
die2 <- (x = 1:6, size = 10000, replace = TRUE)
sumDice <- die1 + die2
```

Vediamo i dati

```
read(die1)
```

```
## [1] 1 4 1 2 5 3
```

```
read(die2)
```

```
## [1] 1 6 1 4 1 3
```

Sia E l'evento "la somma dei dadi è 6", e F "almeno uno dei dadi è 2". Definiamo questi eventi dai nostri dati simulati:

```
eventE <- sumDice == 6
read(eventE)
```

```
## [1] FALSE FALSE FALSE  TRUE TRUE TRUE
```

```
eventF <- die1 == 2 | die2 == 2
read(eventF)
```

```
## [1] FALSE FALSE FALSE  TRUE FALSE FALSE
```

Usando `mean` troviamo la percentuale con cui si sono verificati gli eventi:

```
mean(eventE) # P(E)
```

```
## [1] 0.1409
```

```
mean(eventF)
```

```
## [1] 0.2998
```

Per stimare $P(E \cap F) = \frac{2}{36} \approx 0,056$:

```
mean(eventE & eventF)
```

```
## [1] 0.0587
```

Non è necessario memorizzare tutti i vettori TRUE/FALSE nelle variabili evento.
Ecco una stima di $P(E \cup F) = \frac{14}{36} \approx 0,389$:

```
mean((sumDice == 6) | (die1 == 6) | (die2 == 6))
```

```
## [1] 0.382
```

Utilizzo di `replicate` per ripetere gli esperimenti

Per simulazioni complesse, seguiamo un flusso:

1. Scrivo il codice per eseguire l'esperimento
2. Ripeto l'esperimento un piccolo numero di volte e controllo i risultati:

- `replicate(100, { ESPERIMENTO })`

3. Ripeto l'esperimento un grande numero di volte e memorizzo il risultato:

- `event <- replicate(10000, { ESPERIMENTO })`

4. Calcolo la probabilità usando `mean`

Probabilità condizionata

Dato uno spazio di probabilità $(S, \mathcal{F}, \mathbb{P})$ e due eventi $E, H \in \mathcal{F}$ con $\mathbb{P}(H) > 0$, si dice *probabilità condizionata* di E dato H la quantità

$$\mathbb{P}(E|H) = \frac{P(E \cap H)}{H}$$

che esprime il grado di fiducia dell'osservatore nel verificarsi di E *supponendo che si verifichi H*.
Supponiamo di lanciare due dadi e uno di essi cade dal tavolo dove non puoi vederlo, mentre l'altro mostra un 4. Vorremmo aggiornare le probabilità associate alla somma dei due dadi in base a queste informazioni. La nuova probabilità che la somma dei dadi sia 2 sarebbe 0, la nuova probabilità che la somma dei dadi sia 5 sarebbe 1/6 perché questa è solo la probabilità che il dado che non possiamo vedere sia un "1" e la nuova probabilità che la somma dei dadi sia 7 sarebbe anche 1/6.
Formalmente abbiamo la seguente definizione:
Sia A e B eventi in uno spazio campionario S, con $P(B) \neq 0$, la *probabilità condizionale* che A dato B sia:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Leggiamo $P(A|B)$ come "la probabilità di A dato B".

Valutazioni classiche

Esempio 2.19:
Lanciamo 2 dadi, con quale probabilità entrambe i dadi sono danno 4, sapendo che la loro somma è 8?
$A = \{(4,4)\}$
$B = \{(2,6),(3,5),(4,4),(5,3),(6,2)\}$

$$P(A|B) = P(A \cap B)/P(B) = \frac{1/36}{5/36} = 1/5$$

Invece quale è la probabilità che la somma dei dadi sia 8 sapendo che entrambe i dadi danno 4?
$A = \{(4,4)\}$
$B = \{(4,4)\}$

$$P(A|B) = P(B \cap A)/P(A) = \frac{1/36}{1/36} = 1$$

Quindi:

1. $P((A \cap B)|B) = P(A|B)$
2. $P(A \cup B|B) = 1$

Valutazioni uniformi

Esempio: Rotazione di uno spinner
Facciamo ruotare velocemente uno spinner simmetrico imperniato su un goniometro e ne osserviamo l'angolo di arresto in $]-\frac{\frac{\pi}{2},\pi}{2}]$. Con quale probabilità l'angolo di arresto dello spinner sarà compreso tra $-\frac{\pi}{4} e \frac{\pi}{3}$ (estremi inclusi) supponendo che sia positivo?
Notiamo preliminarmente che le valutazioni uniformi sono diffuse e quindi è *inessenziale* l'inclusione o meno degli estremi nell'intervallo di cui si calcola la probabilità e in quello al quale si condiziona.
Prendiamo $]a, b] =] - \frac{\pi}{2}, \frac{\pi}{2}]$ e con $\mathbb{P}$ *uniforme* troviamo:

$$\mathbb{P}\left(\left[-\frac{\pi}{4}, \frac{\pi}{3}\right] | \left[0, \frac{\pi}{2}\right]\right) = \frac{\frac{\pi}{3} - 0}{\frac{\pi}{2} - 0} = \frac{2}{3}$$

visto che $\left[-\frac{\pi}{4}, \frac{\pi}{3}\right] \cap \left[0, \frac{\pi}{2}\right] = [0, \frac{\pi}{3}]$.
In questo caso $\mathbb{P}\left(\left[-\frac{\pi}{4}, \frac{\pi}{3}\right] | \left[0, \frac{\pi}{2}\right]\right) > \mathbb{P}\left(\left[-\frac{\pi}{4}, \frac{\pi}{3}\right]\right)$ perché:

$$\mathbb{P}\left(\left[-\frac{\pi}{4}, \frac{\pi}{3}\right]\right) = \frac{\frac{\pi}{3} - \left(-\frac{\pi}{4}\right)}{\frac{\pi}{2} - \left(-\frac{\pi}{2}\right)} = \frac{\frac{7\pi}{12}}{\pi} = \frac{7}{12} < \frac{8}{12} = \frac{2}{3}$$

Valutazioni frequentiste

se le probabilità degli eventi sono frequenze relative di realizzazione in precedenti ripetizioni del fenomeno:

$$\mathbb{P}(E|H) = \frac{k_{E \cap H}}{k_H}, \quad E \in \mathcal{F}$$

Essempio:
Lanciamo un dado a sei facce caricato per ottenere 6 con cui, in una sequenza di lanci precedenti, abbiamo ottenuto settantanove volte 6, cinque volte 5, tre volte 4, sette volte 3, cinque volte 2 e una volta 1. Quanto valuteremo la probabilità che un punteggio pari sia primo?
Prendiamo $S = \{1, 2, 3, 4, 5, 6\}, \mathcal{F} = \wp(S)$ e $\mathbb{P}$ specificata da:

$$\begin{array}{lll} \mathbb{P}\{1\} = 0.01 & \mathbb{P}\{2\} = 0.05 & \mathbb{P}\{3\} = 0.07 \\ \mathbb{P}\{4\} = 0.03 & \mathbb{P}\{5\} = 0.05 & \mathbb{P}\{6\} = 0.79 \end{array}$$

senza considerazioni di simmetria. Posto $A = \text{"primo"} = \{2, 3, 5\}$ e $B = \text{"pari"} = \{2, 4, 6\}$, troviamo per la probabilità richiesta:

$$\mathbb{P}(A|B) = \frac{k_{A \cap B}}{k_B} = \frac{k_2}{k_{2,4,6}} = \frac{5}{5 + 3 + 79} = \frac{5}{87} \backsimeq 0.057 = 5.7\%$$

Eventi indipendenti

In uno spazio di probabilità $(S, \mathcal{F}, \mathbb{P})$, due eventi $E, F \in \mathcal{F}$ si dicono *stocasticamente indipendenti* sotto $\mathbb{P}$ o semplicemente *indipendenti* quando vale la fattorizzazione:

$$\mathbb{P}(E \cap F) = \mathbb{P}(E)\mathbb{P}(F)$$

Può accadere in due modi:

- **banalmente** per $\mathbb{P}(E) = 0$ o $\mathbb{P}(F) = 0$
  - perché allora $\mathbb{P}(E \cap F) \leq min\{\mathbb{P}(E), \mathbb{P}(F)\}$ per monotonia e quindi necessariamente $\mathbb{P}(E \cap F) = 0$
- **significativamente** con:

$$\begin{array}{rcl} \mathbb{P}(E|F) & = & \mathbb{P}(E) \\ \mathbb{P}(\mathbb{F}|\mathbb{E}) & = & \mathbb{P}(F) \end{array}$$

  se $min\{\mathbb{P}(E), \mathbb{P}(F)\} > 0.$
  Scriveremo

$$E \perp\!\!\!\perp F$$

  per indicare che E ed F sono *indipendenti*, più precisamente

$$E \perp\!\!\!\perp_{\mathbb{F}} F$$

  per ricordare il ruolo di $\mathbb{P}$.
  Altrimenti scriveremo $E \not\perp\!\!\!\perp F$ se E ed F sono *dipendenti*.
  In particolare:
- E ed F sono *favorevolmente dipendenti* (sotto $\mathbb{P}$) quando

$$\mathbb{P}(E \cap F) > \mathbb{P}(E)\mathbb{P}(F)$$

  cioè $\mathbb{P}(E|F) > \mathbb{P}(E)$ e $\mathbb{P}(F|E) > \mathbb{P}(F)$ nel caso significativo;
- E ed F sono *sfavorevolmente dipendenti* (sotto $\mathbb{P}$) quando

$$\mathbb{P}(E \cap F) < \mathbb{P}(E)\mathbb{P}(F)$$

  cioè $\mathbb{P}(E|F) < \mathbb{P}(E)$ e $\mathbb{P}(F|E) < \mathbb{P}(F)$ nel caso significativo;
  Esempio:
  Lanciamo un ordinario dado a sei facce e ne osserviamo il punteggio. Gli eventi $A = \text{"primi"}$ e $B = \text{"pari"}$ sono indipendenti?
  Presa $\mathbb{P}$ classica su tutte le parti di $S = \{1, 2, 3, 4, 5, 6\}$:

$$\begin{array}{r} \mathbb{P}(A) = \mathbb{P}\{2,3,5\} = \frac{3}{6} = \frac{1}{2} \\ \mathbb{P}(B) = \mathbb{P}\{2,4,6\} = \frac{3}{6} = \frac{1}{2} \\ \mathbb{P}(A \cap B) = \mathbb{P}\{2\} = \frac{1}{6} < \frac{1}{4} = \frac{1}{2} \times \frac{1}{2} \end{array}$$

  quindi $A \not\perp\!\!\!\perp B$, in particolare A e B sono *sfavorevolmente dipendenti*.
  Tuttavia A e B sono *logicamente indipendenti*, dal momento che che i loro quattro costituenti sono tutti non vuoti:

$$\overline{A} \cap \overline{B} = \{1\}, \overline{A} \cap B = \{4, 6\}, A \cap \overline{B} = \{3, 5\}, A \cap B = \{2\}$$

Invarianza per negazione dell’indipendenza tra due eventi

$$E \perp\!\!\!\perp F \Rightarrow E \perp\!\!\!\perp \overline{F}$$

Dimostrazione:

$$\begin{aligned}\mathbb{P}(E \cap \overline{F}) &= \mathbb{P}(E) - \mathbb{P}(E \cap F) \\ &= \mathbb{P}(E) - \mathbb{P}(E) - \mathbb{P}(F) \\ &= \mathbb{P}(E)\{1 - \mathbb{P}(F)\} \\ &= \mathbb{P}(E)\mathbb{P}(\overline{F})\end{aligned}$$

Vale anche il viceversa, perché $\overline{\overline{F}} = F$, quindi:

$$E \perp\!\!\!\perp F \iff E \perp\!\!\!\perp \overline{F} \iff \overline{E} \perp\!\!\!\perp \overline{F} \iff \overline{E} \perp\!\!\!\perp F$$

$\perp\!\!\!\perp$ è una *relazione simmetrica*, quindi affermare che E ed F sono indipendenti sotto $\mathbb{P}$ corrisponde ad affermare che $\mathbb{P}$ si fattorizza su tutti i costituenti di E ed F.
Si noti la differenza: nel caso del dado equilibrato abbiamo scoperto che "pari" e "centrale" sono eventi indipendenti; nel caso della coppia abbiamo *imposto* che F_1 e F_2 siano indipendenti (ed equiprobabili).
Una differenza analoga a quella che passa tra *calcolare* $\mathbb{P}(E|H)$ a partire da $\mathbb{P}(E \cap H)$e *assegnare* $\mathbb{P}(E|H)$ per specificare $\mathbb{P}(E \cap H)$, supponendo che $\mathbb{P}(H) > 0$.
Esempio:
Una coppia ha due figli. Sappiamo che almeno una è femmina. Con quale probabilità sono due femmine? Troviamo subito

$$\begin{aligned}\mathbb{P}(F_1 \cup F_2) &= \mathbb{P}(F_1) + (\overline{F_1} \cap F_2) = \frac{1}{2} + \frac{1}{4} = \frac{3}{4} \\ \mathbb{P}(F_1 \cap F_1|F_1 \cup F_2) &= \frac{\mathbb{P}(F_1 \cap F_2)}{\mathbb{P}(F_1 \cup F_2)} = \frac{1/4}{3/4} = \frac{1}{3} \dots\end{aligned}$$

dunque $\mathbb{P}(F_1 \cap F_2|F_1 \cup F_2) = \frac{1}{3} \neq \frac{1}{2} = \mathbb{P}(F_1|F_2) = \mathbb{P}(F_2|F_1)!$
Per comprendere questo paradossoa ragioniamo su come sappiamo che almeno una figlia è femmina e ipotizziamo di averla incontrata, introducendo la partizione

$$\begin{aligned}H &= \text{"incontro figlia femmina"} \\ \overline{H} &= \text{"incontro figlio maschio"}\end{aligned}$$

nel diagramma di Venn (in modo tale che $F_1 \cap F_2 \subset H \subset F_1 \cup F_2$).
Supponiamo:

$$\mathbb{P}(H|\overline{F_1} \cap F_2) = p = \mathbb{P}(H \cap F_1 \cap F_2) \quad \text{con} \quad 0 < p < 1$$

Troviamo

$$\begin{aligned}\mathbb{P}(F_1 \cap F_2|H) &= \frac{\mathbb{P}(F_1 \cap F_2 \cap H)}{\mathbb{P}(H)} \\ &= \frac{\mathbb{P}(F_1 \cap F_2)}{\mathbb{P}(H|\overline{F_1} \cap F_2)\mathbb{P}(\overline{F_1} \cap F_2) + \mathbb{P}(F_1 \cap F_2) + \mathbb{P}(H|F_1 \cap \overline{F_2})\mathbb{P}(F_1 \cap \overline{F_2})} \\ &= \frac{1/4}{p\frac{1}{4} + \frac{1}{4} + p\frac{1}{4}} = \frac{1}{1+2p} = \begin{cases} 1 & \text{se } p \to 0 \\ \frac{1}{2} & \text{se } p \to \frac{1}{2} \\ \frac{1}{3} & \text{se } p \to 1 \end{cases}\end{aligned}$$

Specificare $\mathbb{P}$ sui costituenti E ed F supponendoli indipendenti è sempre possibile, ma non sempre (pienamente) appropriato.
Fare diversamente richiede di valutare tipo e forza della dipendenza:

$$\mathbb{P}(E \cap F) = d\mathbb{P}(E)\mathbb{P}(F)$$

con dipendenza sfavorevole per $d < 1$ e favorevole per $d > 1$; il *fattore di dipendenza* d va scelto garantendo la coerenza di $\mathbb{P}$, cioé:

$$max\left\{0, \frac{1}{\mathbb{P}(E)} + \frac{1}{\mathbb{P}(F)} - \frac{1}{\mathbb{P}(E)\mathbb{P}(F)}\right\} < d < min\left\{\frac{1}{\mathbb{P}(E)}, \frac{1}{\mathbb{P}(F)}\right\}.$$

Diremo che E ed F sono *stocasticamente indipendenti* in modo sostanziale quando:

$$E \perp\!\!\!\perp F \text{ con } 0 < \mathbb{P}(E), \mathbb{P}(F) < 1$$

Esempio: interruttori elettrici in serie
Un circuito elettrico è formato da due interruttori *uno di seguito all'altro*: fa passare corrente quando *entrambi* gli interruttori sono chiusi. Supponiamo che (in un dato istante) ciascun interruttore sia chiuso con probabilità $p = 0.8$ indipendentemente dall'altro interruttore. Con quale probabilità (in tale istante) passerà corrente nel circuito?
Posto $C_i =$ "i-esimo interruttore chiuso"$, i = 1, 2$, in un diagramma di Venn, assegnamo $\mathbb{P}$ su $\mathcal{F} = \omega(C_1, C_2)$ con $C_1 \perp\!\!\!\perp C_2$ e $\mathbb{P}(C_1) = \mathbb{P}(C_2) = p$. Per l'evento di interesse $C_1 \cap C_2$ troviamo:

$$\mathbb{P}(C_1 \cap C_2) = \mathbb{P}(C_1)\mathbb{P}(C_2) = p^2$$

e quindi $\mathbb{P}(C_1 \cap C_2) = (0.8)^2 = 0.64 < 0.8.$
Esempio: interruttori elettrici in parallelo
Un circuito elettrico è formato da due interruttori *uno di fianco all'altro*: fa passare corrente quando *almeno uno* degli interruttori è chiuso. Supponiamo che (in un dato istante) ciascun interruttore sia chiuso con probabilità $p = 0.8$ indipendentemente dall'altro interruttore. Con quale probabilità (in tale istante) passerà corrente nel circuito?
Probabilità che l'interruttore sia **aperto**: $1 - p = 1 - 0.8 = 0.2$
$\mathbb{P}(\text{nessun interruttore chiuso}) = (0.2)^2 = 0.04$
La probabilità che **almeno uno degli interruttori sia chiuso** (e quindi che passi corrente) è il complementare di questa probabilità:

$$\mathbb{P}(\text{passa corrente}) = 1 - \mathbb{P}(\text{nessun interruttore chiuso}) = 1 - 0.04 = 0.96$$