

Scheda Esercitazione su R – Elementi di Probabilità 1

ESERCIZIO

In un gruppo di N persone, la probabilità che non ci sia nemmeno una coppia con il compleanno lo stesso giorno è:

$$P(E) = \frac{365 \times (365 - 1) \times \dots \times (365 - (N - 1))}{365^N} = \frac{365!}{(365 - N)! \times 365^N} = \frac{N! \binom{365}{N}}{365^N},$$

dove:

$$E = \{N \text{ persone hanno il compleanno in giorni distinti}\}.$$

Tramite l'utilizzo di R, si risponda ai punti seguenti

- Calcolare il fattoriale di un numero n (funzione **factorial(n)**)
- Calcolare il coefficiente binomiale di n su k (funzione **choose(n,k)**)
- Calcolare $P(E)$ per N fissato
- Per lo stesso valore di N scelto al punto precedente, calcolare la probabilità che ci siano almeno due persone che hanno il compleanno lo stesso giorno.
- Cosa succede per $N = 23$?
- Studiare l'andamento di $P(E)$ e $P(E^c)$ al variare di N e generarne il grafico (a linee). Per studiare l'andamento delle probabilità al variare di N si consiglia di utilizzare un ciclo "for":

```
for (variable in vector) {  
  
}
```

- Utilizzare la funzione **y=sample(x, size, resample= TRUE)** per simulare il compleanno di una persona (**size=1**) inteso come l'estrazione di un compleanno dall'insieme di tutti i giorni dell'anno, salvati nel vettore **x**. Ripetere la prova con diversi valori di **size**. Perché si pone **resample= TRUE**?
- Utilizzare la funzione **uplicated(y)** e **anyDuplicated(y)** per verificare se ci sono dei compleanni nella stessa data dentro a **y=sample(x, size, resample= TRUE)**
- Per ogni valore di N (all'interno del ciclo "for" creato al punto precedente), simulare un certo numero di esperimenti (ad esempio, *nexp*=100) dove, per ogni esperimento, si generano con **sample** i possibili compleanni. Si verifichi con **anyDuplicated** se ci sono stati compleanni duplicati, si calcoli la frequenza relativa e si confronti con il valore teorico di probabilità.
- Si ripeta il punto precedente per diversi valori di *nexp*
- Riportare su un grafico la frequenza relativa sovrapposta a quella della probabilità teorica (funzione **lines()**)