

**Scuola estiva AIP 2025, Bertinoro**

# **Introduzione alla statistica bayesiana con applicazioni in R e STAN**

*Massimiliano Pastore (07 aprile 2025)*

## Esercizi con la binomiale

### Esercizio 1

Si consideri una distribuzione a priori che esprime il grado con cui si ritiene che una moneta sia bilanciata:  $B(\theta|4, 4)$ .

1. Si rappresenti graficamente tale distribuzione.
2. Si supponga di lanciare la moneta e di ottenere testa, quale sarà la distribuzione a posteriori? Si rappresenti graficamente anche questa distribuzione e la si confronti con la precedente.
3. Si consideri la distribuzione ottenuta al punto 2 come *prior* per un nuovo lancio della moneta. Si rilanci la moneta ottenendo nuovamente testa; quale sarà la nuova *posterior*? Si produca la rappresentazione grafica.
4. Si supponga di lanciare la moneta per la terza volta ed ottenere croce; si rappresenti graficamente la nuova *posterior* (sempre confrontandola con le precedenti)
5. Se nei tre lanci avessimo ottenuto la sequenza  $C, T, T$  anziché  $T, T, C$ , la *posterior* finale sarebbe la stessa?

### Esercizio 2

Si avvicinano le elezioni e siete interessati a sapere se sarà preferito il candidato A o il candidato B. In un sondaggio basato su un campione casuale di 100 soggetti 58 di essi dichiarano di preferire il candidato A.

1. Supponiamo che prima di conoscere il risultato del sondaggio la vostra *prior* era di tipo uniforme. Si calcoli l'HDI al 95% dopo avere visto i risultati del sondaggio.
2. Sulla base del sondaggio, è credibile che la popolazione si ugualmente suddivisa nelle preferenze tra i candidati?
3. In un secondo sondaggio, sempre su un campione di 100 soggetti, si ottiene che 57 di essi dichiarano di preferire il candidato A e gli altri il candidato B. Assumendo che non ci sia stato un cambiamento nell'opinione generale tra i due sondaggi, ed utilizzando i dati del primo campione come *prior*, qual'è l'HDI al 95% sulla nuova *posterior*?
4. In base a questo nuovo sondaggio, è credibile che la popolazione si ugualmente suddivisa nelle preferenze tra i candidati?

### Esercizio 3

Si consideri un esperimento di apprendimento con due condizioni sperimentali. Nella prima i soggetti sono addestrati a premere il tasto F quando vedono la combinazione di parole *radio* e *oceano*, nella seconda condizione sono addestrati a premere il tasto J alla combinazione *radio* e *montagna*. 50 soggetti vengono istruiti nei due compiti e lo sperimentatore controlla fino a quando è sicuro che essi abbiano imparato esattamente le corrispondenze. Successivamente, per stabilire cosa abbiano imparato, li sottopone a due nuovi test. Nel primo viene presentata la sola parola *radio* ed i soggetti vengono istruiti a dare la migliore risposta (tra F e J) in base a quanto imparato precedentemente. Nel secondo test vengono presentate entrambe le parole *oceano* e *montagna*, anche in questo caso i soggetti devono indicare la migliore risposta. Il risultato dell'esperimento è che nel primo test 40 soggetti scelgono F e 10 scelgono J; nel secondo test 15 scelgono F e 35 scelgono J. Si può sostenere che i soggetti siano più orientati a rispondere F

oppure J in uno dei due tipi di prova? Per rispondere si assuma una *prior* uniforme e si utilizzi l'HDI al 95% per decidere quale propensione sia più credibile.

### Esercizio 4

Abbiamo comprato una moneta in un negozio di trucchi, pertanto sappiamo che è fortemente sbilanciata ma non sappiamo se verso testa o croce. Assumiamo come *prior* la  $B(\theta|.5, .5)$  quindi effettuiamo 5 lanci ottenendo 4 T e 1 C. Quale sarà la *posterior*? Si rappresentino graficamente sia la *prior* che la *posterior*.

### Esercizio 5

1. Supponete di avere una moneta coniata direttamente dalla Banca Centrale. Pertanto avete una grande convinzione che la moneta sia bilanciata. Lanciate la moneta per 10 volte ottenendo testa per 9 volte. Si stimi, per via simulativa, la probabilità di ottenere testa nell' undicesimo lancio, scegliendo una *prior* e giustificandola.
2. Ora supponete di avere un'altra moneta di uno strano materiale e sulla quale è stampata la scritta *Patent Pending, International Magic, Inc.* Anche questa volta lanciando la moneta per 10 volte ottenete 9 teste. Quale sarà la probabilità di ottenere testa nell'undicesimo lancio? Giustificate la scelta della *prior*.

### Esercizio 6

Supponete di avere una moneta (ancora ...) e non sapere se sia bilanciata o truccata. La lanciate 20 volte ottenendo 15 teste. Sulla base di questo risultato è più probabile che la moneta sia bilanciata o truccata?

1. Si rappresentino graficamente (giustificandone la scelta) le *prior* che esprimono le due ipotesi contrapposte.
2. Si rappresentino graficamente le *posterior* ottenute in base al risultato osservato.
3. Si calcoli il Bayes Factor come rapporto tra le evidenze dei due modelli alternativi.

**NOTA:** l'evidenza di un modello binomiale si calcola con la formula

$$E = \frac{B(z + a, N - z + b)}{B(a, b)}$$

in cui  $B$  è la funzione beta<sup>1</sup>,  $N$  il numero di osservazioni,  $z$  il numero di casi favorevoli,  $a$  e  $b$  i parametri della distribuzione beta nella *prior*.

### Esercizio 7

Abbiamo la solita moneta che supponiamo fortemente essere truccata anche se non sappiamo privilegi testa o croce. Pertanto, abbiamo un modello con una *prior* che privilegia croce,  $B(\theta|1, 100)$  ed uno con una *prior* che privilegia testa,  $B(\theta|100, 1)$ . Sia dato l'esito di un solo lancio: testa.

1. Si rappresentino graficamente le due *prior*.
2. Si rappresentino graficamente le due *posterior*.

---

<sup>1</sup> In R si veda `?beta`.

3. Si calcoli il Bayes Factor.

### Esercizio 8

Abbiamo la solita moneta che non sappiamo se bilanciata o meno. Vogliamo modellare questa incertezza con una sola *prior* per il parametro  $\theta$  che avrà pertanto tre picchi: a 0, .5 e 1.

1. Si definisca la prior con il seguente codice (o uno simile a piacimento)

```
pTheta <- rep(c(50:1,rep(1,50),1:50),2)
pTheta <- pTheta/sum(pTheta)
width <- 1/length(pTheta)
Theta <- seq(from=width/2,to=1-width/2,by=width)
```

e quindi si rappresenti graficamente, verificando che si tratti effettivamente di una distribuzione di probabilità.

2. Si supponga di avere lanciato la moneta 20 volte ed avere ottenuto 15 teste; si rappresenti graficamente la *likelihood*.
3. Si stimi e si rappresenti graficamente la *posterior*.

### Esercizio 9

1. Si riconsideri la *prior* definita al punto 1 dell'esercizio 8 e di avere lanciato la moneta 4 volte ottenendo 3 teste. Si rappresentino graficamente la *prior*, la *likelihood* e la *posterior*.
2. Si calcolino l'evidenza ( $p(D)$ ) e la media della *posterior*.
3. A questo punto si lancia la moneta altre 16 volte ottenendo 12 teste; quindi, utilizzando la *posterior* del punto precedente come *prior*, si stimi (e rappresenti graficamente) la nuova *posterior*.
4. Si calcolino l'evidenza ( $p(D)$ ) e la media della *posterior*.

### Esercizio 10

In una piccola città si sospetta il rischio di diffusione di una malattia infettiva. Per stimare la presenza di individui infetti viene preso in considerazione un campione di  $n = 20$  soggetti nessuno dei quali risulta essere infetto. Indicando con  $\theta$  la proporzione (incognita) di persone infette nella città e  $y = 0$  il numero di soggetti infetti nel campione, possiamo indicare lo spazio campionario come segue:

$$\Theta = [0, 1] \quad Y = \{0, 1, \dots, 20\}$$

Secondo alcuni studi in altre città, l'incidenza della malattia varia tra lo 0.5 ed il 2%, con una media dell'1%.

1. Si calcoli la probabilità che nel campione ci siano 0 soggetti infetti,  $P(y = 0|\theta)$ , ipotizzando  $\theta = (0.05, 0.1, 0.2)$ .
2. Si rappresentino graficamente le tre distribuzioni di probabilità associate ai tre valori di  $\theta$ .
3. Si consideri come *prior* la distribuzione  $B(\theta|2, 20)$ , si produca la rappresentazione grafica e si calcolino il valore atteso di  $\theta$ ,  $E(\theta)$  e la probabilità che il tasso di infezione sia minore di 0.10.

4. Si determini la distribuzione a posteriori, dopo aver osservato nel campione  $y = 0$  confrontandola con la *prior*.
5. Si calcolino il valore atteso della *posterior* e la probabilità a posteriori che il tasso di infezione sia minore di 0.10.
6. Si determini l'intervallo di confidenza del risultato osservato utilizzando l'approccio NHST per cui l'intervallo è dato da

$$\bar{\theta} \pm 1.96 \sqrt{\frac{\bar{\theta}(1 - \bar{\theta})}{n}}$$

in cui  $\bar{\theta} = y/n$  ovvero la proporzione stimata di malati sulla base dell'osservazione nel campione.

7. Si determini l'intervallo HDI al 95% sulla *posterior*.

### Esercizio 11

Due ricercatori (A e B) sono interessati a studiare se le donne di età intorno ai 65 anni si sentano felici. A ritiene che la percentuale di donne felici sia il 50%, B invece che sia il 70%. I due intervistano un campione di 130 donne ottenendo il 60% di risposte positive.

1. Si valuti con approccio NHST (test binomiale) se le ipotesi dei due ricercatori siano supportate dai dati.
2. Si definiscano le *prior* per modellare le ipotesi dei due ricercatori giustificandone la scelta e rappresentandole graficamente.
3. Si rappresenti graficamente la *likelihood* relativa al risultato osservato.
4. Si rappresentino le due *posterior* relative alle ipotesi dei due ricercatori.
5. Si calcolino le evidenze del risultato in base alle due ipotesi espresse a priori e si confrontino. Quali conclusioni si possono trarre dal confronto?

## Distribuzione normale e confronti tra medie

### Esercizio 12

Ad un campione di 54 studenti che partecipa ad un esperimento di problem-solving viene somministrata la WAIS-R, una scala che si assume avere, nella popolazione, una distribuzione normale, media 10 e dev. st. 3. I punteggi degli studenti sono i seguenti:

```
14 11 13 13 13 15 11 16 10 13 14 11 13 12 10 14 10 14
16 14 14 11 11 11 13 12 13 11 11 15 14 16 12 17 9 16
11 19 14 12 12 10 11 12 13 13 14 11 11 15 12 16 15 11
```

1. Si valuti nel modo opportuno se la distribuzione dei punteggi osservati si possa considerare approssimativamente normale.
2. Si esegua il  $t$ -test in forma tradizionale sui punteggi per valutare se il campione osservato possa essere considerato come proveniente dalla popolazione normale, specificando le ipotesi  $H_0$  e  $H_1$ .
3. Si scriva il codice STAN per il test in forma bayesiana.
4. Si stimino le distribuzioni a posteriori dei parametri  $\mu$  e  $\sigma$ .
5. Si rappresentino graficamente le *posterior* dei parametri  $\mu$  e  $\sigma$ .
6. Sulla base del risultato ottenuto, la media dei punteggi dei soggetti si può considerare diversa da quella della popolazione normativa?

### Esercizio 13

In uno studio sulle abitudini di sonno di giovani universitari in un particolare college, sono state raccolte le ore di sonno di un campione casuale di 20 studenti:

```
9.0  8.5  7.0  8.5  6.0  12.5  6.0  9.0  8.5  7.5
8.0  6.0  9.0  8.0  7.0  10.0  9.0  7.5  5.0  6.5
```

Vogliamo stimare media  $\mu$  e varianza  $\sigma^2$  della popolazione da cui provengono questi studenti.

1. Si definiscano delle prior per  $\mu$  e  $\sigma$ .  
Per la media assumiamo una  $t(2, 9, 3)$ , per la deviazione standard una  $t(2, 2, 2)[0, ]$ .
2. Si scriva il codice STAN per la stima di  $\mu$  e  $\sigma^2$ .
3. Si producano le distribuzioni a posteriori di media e varianza.
4. Si rappresentino graficamente le distribuzioni a posteriori di media e varianza.
5. Si rappresenti graficamente la distribuzione a posteriori congiunta dei parametri.
6. Si calcolino gli intervalli HDI al 95% per  $\mu$  e  $\sigma$ .
7. Utilizzando le informazioni ricavate dalle *posterior*, si stimino i valori attesi dei percentili 75, 85 e 95 nella popolazione normale. Suggerimento: si tenga presente che nella popolazione normale i percentili possono essere calcolati con  $\mu + z\sigma$  in cui  $z$  è il quantile di riferimento.

8. Si ipotizzi ora che il numero medio di ore di sonno degli studenti universitari sia compreso nell'intervallo  $[9.5, 10.5]$ . Si calcoli la probabilità a posteriori che il campione osservato appartenga a tale popolazione.

### Esercizio 14

Consideriamo due gruppi di soggetti cui viene misurato il QI. Il primo gruppo ( $N_1 = 47$ ) prima della misurazione assume un farmaco, mentre il secondo ( $N_2 = 42$ ) assume un placebo. La media dei punteggi nel primo gruppo è  $\bar{x}_1 = 101.91$ , la media del secondo gruppo  $\bar{x}_2 = 100.36$ . Si vuole valutare l'ipotesi che l'uso del farmaco aumenti le performance nel QI.

1. Si importino i dati dal file `QIdata.rda`.
2. Si valuti, con un test opportuno, se le varianze dei due gruppi siano omogenee.
3. Si esegua nel modo appropriato il  $t$ -test in forma classica per confrontare le medie dei due gruppi.
4. Si esegua lo stesso test in forma bayesiana utilizzando STAN con i seguenti passaggi:
  - Si scriva il codice STAN.
  - Si produca la distribuzione a posteriori della differenza tra le medie.
  - Si rappresenti graficamente la posterior con il relativo intervallo HDI al 90%.
5. Si calcoli e si rappresenti graficamente la distribuzione a posteriori per l'effect size definito con:

$$\phi = \frac{(\mu_1 - \mu_2)}{\sqrt{(\sigma_1^2 + \sigma_2^2)/2}}$$

6. Assumendo che l'effetto compreso tra  $-.5$  e  $.5$  è considerato equivalente ad un effetto nullo, si determini la probabilità a posteriori a supporto dell'effetto nullo, ovvero che  $\phi \in [-.5, .5]$ .

### Esercizio 15

I seguenti dati sono le lunghezze in millimetri delle mandibole di un campione di 20 esemplari di coyote dorato: 10 maschi (120, 107, 110, 116, 114, 111, 113, 117, 114, 112) e 10 femmine (110, 111, 107, 108, 110, 105, 107, 106, 111, 111). Si ipotizza che la lunghezza media della mandibola dei maschi sia superiore a quella delle femmine.

1. Si scriva un codice STAN per ricavare la distribuzione a posteriori della differenza tra le medie.
2. Si determinino le distribuzioni a posteriori di medie, varianze dei due gruppi e della differenza tra le medie.
3. Si utilizzi la distribuzione a posteriori per valutare l'assunto di omogeneità delle varianze.
4. Si stimi la probabilità dell'ipotesi  $H_0 : \mu_{(\text{male})} = \mu_{(\text{female})}$  assumendo una ROPE compresa tra  $-4$  e  $4$ .

### Esercizio 16

15 pazienti con deficit di memoria vengono valutati nelle funzioni cognitive con un test che nella popolazione ha media  $\mu = 100$  e dev. st.  $\sigma = 15$ . I punteggi ottenuti al test dai pazienti sono

i seguenti: 110, 109, 104, 79, 92, 105, 93, 86, 107, 103, 79, 85, 80, 106, 89. Il sospetto è che le funzioni cognitive siano in qualche modo compromesse o comunque inferiori a quelle dei soggetti senza deficit.

1. Si esegua un  $t$ -test in forma classica per valutare se il campione si differenzi dalla popolazione normale, specificando le ipotesi  $H_0$  e  $H_1$ .
2. Si determini la distribuzione a posteriori di  $\mu$  e sulla base di questa si valutino nuovamente le ipotesi. **Nota:** si può utilizzare lo stesso codice STAN dell'esercizio 12 opportunamente modificato.
3. A questo punto si supponga di disporre di un nuovo soggetto, il cui punteggio risulta essere 90. Si ripeta il test eseguito al punto 1 aggiungendo il nuovo soggetto agli altri.
4. Si riproduca la *posterior* con l'aggiunta del nuovo soggetto e si confronti con quella ottenuta al punto 2.
5. Quali considerazioni è possibile fare dal confronto tra i risultati ottenuti con approccio classico (punti 1 e 3) e con il metodo bayesiano?

### Esercizio 17

Secondo alcuni studi, la vita media dei ratti che si nutrono liberamente è di circa 800 giorni. Nel file `RatLives.rda` sono presenti i dati relativi ad uno studio su 195 ratti, osservati in condizioni controllate. Le variabili raccolte sono i giorni di sopravvivenza (`days`) ed il genere. Si vuole sapere se, sulla base di tale campione, si possa supportare l'ipotesi di vita media a 800 giorni e se vi siano differenze di genere.

1. Si esegua un  $t$ -test per valutare se il campione di ratti provenga da una popolazione normale con media  $\mu = 800$ .
2. Si ripeta il test in forma bayesiana, stimando l'intervallo HDI al 95% della media a posteriori. **Nota:** si può utilizzare lo stesso codice STAN dell'esercizio 12 opportunamente modificato.
3. Quali conclusioni è possibile trarre dalle analisi eseguite fino qui?
4. Si determini, sempre utilizzando l'intervallo HDI, se sia ipotizzabile una differenza tra maschi e femmine sulla sopravvivenza media. **Nota:** si può utilizzare lo stesso codice STAN usato nell'esercizio 14 opportunamente modificato.
5. Quali conclusioni possiamo trarre dalle analisi eseguite fino qui?
6. Si effettui il *Posterior Predictive Check* confrontando i valori attesi dal modello per i due gruppi con i valori osservati. Quali considerazioni è possibile trarre dalla lettura del grafico ottenuto?



## Appendice

### Funzioni utilizzate e suggerite

[1]	"abline"	"aes"	"aggregate"	"as.mcmc"
[5]	"attr"	"bartlett.test"	"beta"	"binom.test"
[9]	"c"	"cat"	"curve"	"data.frame"
[13]	"dbeta"	"dbinom"	"density"	"file.choose"
[17]	"for"	"geom_density"	"ggplot"	"grep"
[21]	"gsub"	"hdi"	"hist"	"HPDinterval"
[25]	"if"	"ifelse"	"lapply"	"legend"
[29]	"library"	"lines"	"list"	"load"
[33]	"mean"	"options"	"par"	"paste0"
[37]	"pbeta"	"plot"	"polygon"	"print"
[41]	"qqline"	"qqnorm"	"range"	"rbeta"
[45]	"rbind"	"rbinom"	"rep"	"report"
[49]	"rm"	"rnorm"	"rope"	"round"
[53]	"save"	"sd"	"seq"	"sessionInfo"
[57]	"set.seed"	"sort"	"stack"	"stan"
[61]	"strsplit"	"sum"	"t.test"	"text"
[65]	"unique"	"unlist"	"var.test"	"with"
[69]	"write"	"xtable"		

### Appendice: Pacchetti utilizzati

- bayestestR. Makowski D, Ben-Shachar M, Lüdtke D (2019).
- coda. Plummer M, Best N, Cowles K, Vines K (2006).
- ggplot2. Wickham H (2016).
- HDInterval. Meredith M, Kruschke J (2022).
- knitr. Xie Y (2023).
- report. Makowski D, Lüdtke D, Patil I, Thériault R, Ben-Shachar M, Wiernik B (2023).
- rstan. Stan Development Team (2024).
- StanHeaders. Stan Development Team (2020).
- xtable. Dahl D, Scott D, Roosen C, Magnusson A, Swinton J (2019).

