MS 03/10/24

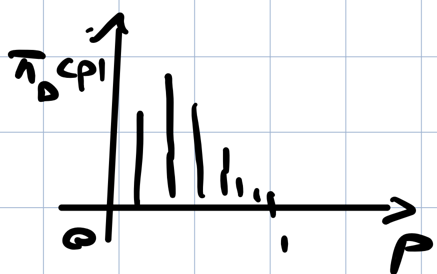
Ieri: un campione casuale normale: analisi bayesiana coniugata.

Oggi: un campione casuale binario: analisi bayesiana coniugata

Dati: $X_1, \dots X_n \mid p \overset{c.i.i.d}{\sim} \text{Bernoulli}(p)$

$p \sim \,?$ (in questo caso $\theta = p$)

Potremmo usare per p una distribuzione discreta $\pi_D(p)$ dando masse ad alcuni valori $p_0 \dots p_k$



ma e' più naturale pensare a p come variabile a valori in $[0,1]$. (continua)

La verosimiglianza e'

$$\mathcal{L}(p; x_1, \dots x_n) = p^{\sum x_i}(1-p)^{n-\sum x_i}$$

Esiste una densità a priori coniugata a questa verosimiglianza? Si

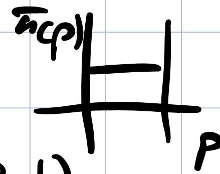
$$\pi(p) = \frac{p^{a-1}(1-p)^{b-1}}{\underbrace{\int_0^1 t^{a-1}(1-t)^{b-1}\,dt}_{B(a,b)}} \quad (0 \le p \le 1)$$

densità beta con (iper)parametri a e b

funzione beta di Eulero con argomenti a e b

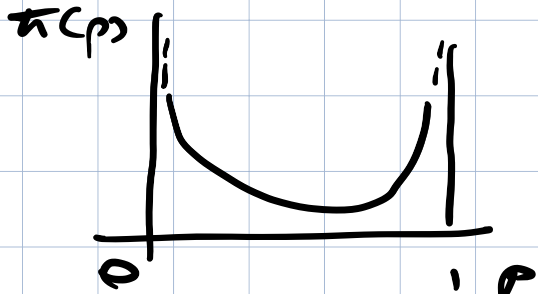
$a = b = 1$ $\qquad \pi(p) = 1 \quad (0 \le p \le 1)$

cioè $\quad p \sim \text{Uniforme}(0,1)$

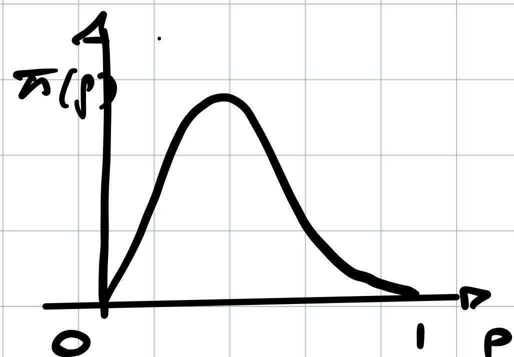


$a = 1 \quad b < 1$

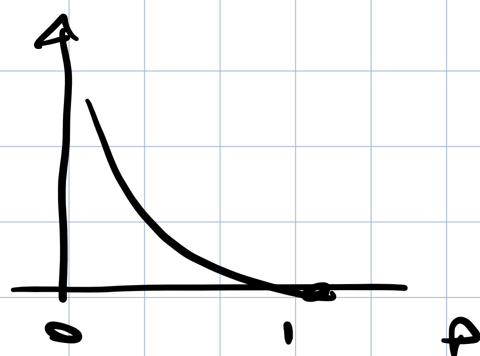
$a < 1 \quad b < 1$



$a > 1 \quad b > 1$



$a < 1 \quad b > 1$



$$E(p) = \frac{\int_0^1 p\, p^{a-1}(1-p)^{b-1}\, dp}{B(a,b)} = \frac{B(a+1,b)}{B(a,b)}$$

$$= \frac{\dfrac{\Gamma(a+1)\cancel{\Gamma(b)}}{\Gamma(a+b+1)}}{\dfrac{\Gamma(a)\cancel{\Gamma(b)}}{\Gamma(a+b)}} = \frac{\dfrac{a\,\cancel{\Gamma(a)}}{(a+b)\,\cancel{\Gamma(a+b)}}}{\dfrac{\cancel{\Gamma(a)}}{\cancel{\Gamma(a+b)}}} = \frac{a}{a+b}$$

$$Var(p) = \frac{ab}{(a+b)^2(a+b+1)}$$

Vari stati di informazione su p possono essere rappresentati al variare della scelta degli (iper) parametri a e b.

questa <u>non è</u> una beta

Teo di Bayes (sfruttando la proporzion.)

$$\pi(p \mid x_1 \ldots x_n) \propto \pi(p)\, \mathcal{L}(p; x_1 \ldots x_n)$$

$$\propto p^{a-1}(1-p)^{b-1}\, p^{\Sigma x_i}(1-p)^{n-\Sigma x_i}$$

$$= p^{a+\Sigma x_i-1}(1-p)^{b+n-\Sigma x_i-1}$$

Da cui deduciamo che

$$p \mid x_1 .. x_n \sim Beta(\underbrace{a+\Sigma x_i}_{a'}, \underbrace{b+n-\Sigma x_i}_{b'})$$

e poiché questa è di nuovo una beta, diciamo che la beta è coniugata al campionamento casuale binario

$$E(p|x_1 \dots x_n) = \frac{a+\sum x_i}{a+\cancel{\sum x_i}+b+n-\cancel{\sum x_i}}$$

fissati $x_1 \dots x_n$, questo e' un numero

La corrispondente variabile aleatoria e'

$$E(p|X_1 \dots X_n) = \frac{a+\sum X_i}{a+b+n}$$

NOTE:

La media a posteriori si può scrivere

$$\underbrace{E(p|X_1 \dots X_n)}_{\text{media a posteriori}} = \underbrace{\frac{a+b}{a+b+n}}_{\text{peso dell'a priori}} \cdot \underbrace{\frac{a}{a+b}}_{\text{media a priori}} + \underbrace{\left(1-\frac{a+b}{a+b+n}\right)}_{1-\text{peso}} \frac{\sum X_i}{n}$$

inoltre

$$\frac{a+b}{a+b+n} \xrightarrow[n \to \infty]{} 0$$

cioe' il peso dell'a priori svanisce al crescere della evidenza empirica

$$\frac{\sum X_i}{n} = \hat{p} = \text{frequenza relativa di successo} = \text{stimatore di massima verosimiglianza di } p$$

L'anno scorso avete dimostrato questo e avete derivato degli intervalli di confidenza approssimati per p (intervallo di Wald

$$\hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

o anche altri come Wilson o Clopper-Pearson)

Infatti, per fare un confronto tra statistica frequentista e statistica bayesiana, consideriamo

1) stima puntuale
2) stima intervallare
3) test di ipotesi
4) predizione

1) è intuitivo considerare come stima bayesiana un qualunque riassunto ragionevole della a posteriori, per esempio

1.1. la media a posteriori

1.2 la moda a posteriori (MAP maximum a posterior value)

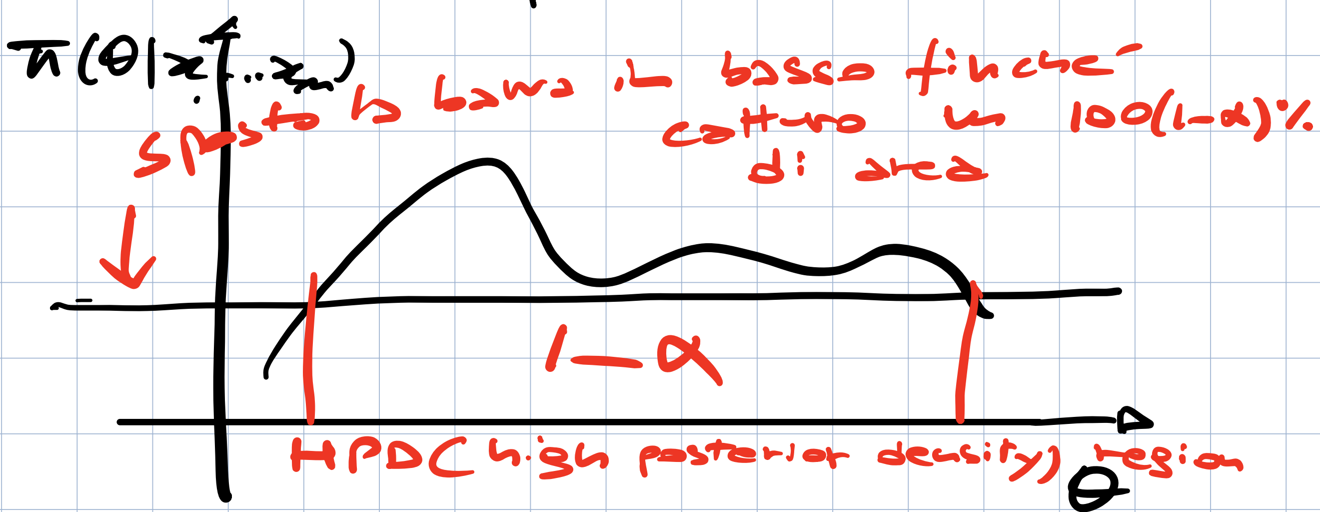
Nel caso binario,

$$E(p|x_1 \dots x_n) = \frac{a+\Sigma x_i}{a+b+n}$$ stimatore bayesiano

(calcolate voi lo stimatore bayesiano alternativo MAP)

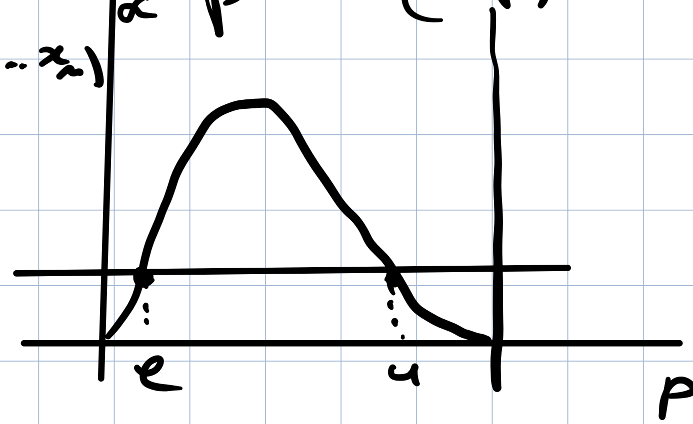
2) Come derivare un intervallo bayesiano (che sarà chiamato INTERVALLO DI CREDIBILITA') ?

Prendendo un intervallo (o regione) "centrale" della a posteriori



Nel caso binario coniugato, se non sono o tutti successi o tutti insuccessi, la densità a posteriori è unimodale, quindi

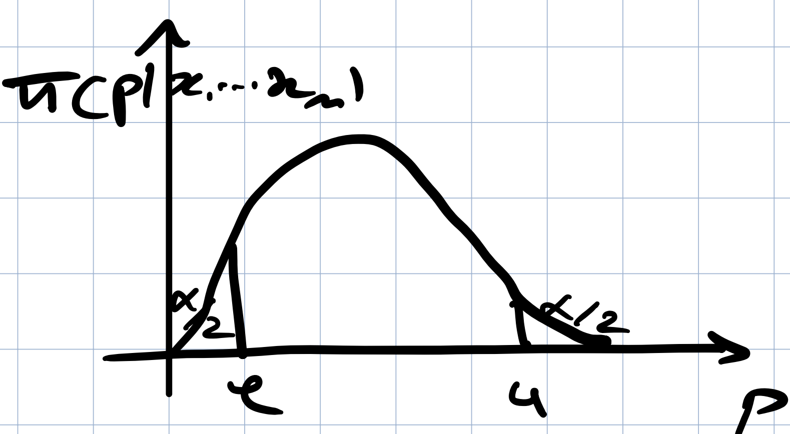
$$\pi(p|x_1\dots x_n) \propto p^{a+\Sigma x_i-1}(1-p)^{b+n-\Sigma x_i-1}$$



$$\begin{cases} \ell^{a+\Sigma x_i-1}(1-\ell)^{b+n-\Sigma x_i-1} = u^{a+\Sigma x_i-1}(1-u)^{b+n-\Sigma x_i-1} \\ \displaystyle\int_{\ell}^{u} \frac{t^{a+\Sigma x_i-1}(1-t)^{b+n-\Sigma x_i-1}}{B(a+\Sigma x_i,\ b+n-\Sigma x_i)}\, dt = 1-\alpha \end{cases}$$

sistema per ottenere una HPD di livello $1-\alpha$ paragonabile a un intervallo di confidenza per p.

In alternativa, si può scegliere un intervallo a code uguali (equal-tailed interval)



Usando la sintassi di R:

$$l = qbeta\left(\frac{\alpha}{2}, a+\sum x_i, b+n-\sum x_i\right)$$

$$u = qbeta\left(1-\frac{\alpha}{2}, a+\sum x_i, b+n-\sum x_i\right)$$

3) test di ipotesi in salsa bayesiana

$$H_0 : \theta \in \Theta_0$$
$$H_a : \theta \notin \Theta_0$$

basta considerare

$$P(\theta \in \Theta_0 ; dati) =$$

$$P(H_0 ; dati)$$

la probabilità dell'ipotesi nulla non è un ente ammissibile in statistica frequentista perché nell'ottica frequentista un'ipotesi non è un ente aleatorio.

Invece della probabilità si possono anche usare gli odds (in italiano, quota di scommessa)

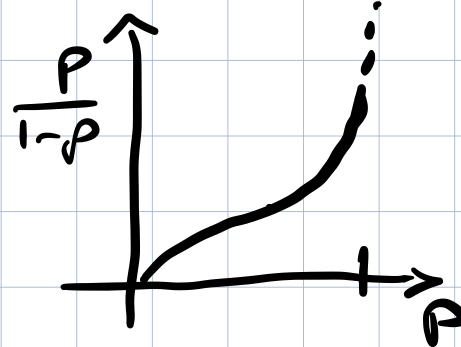
$$P(\text{"Torino vince lo scudetto"}) = 0.05$$

odds a favore $= \dfrac{0.05}{0.95}$ "Il Torino è dato 5:95"

odds contro $= \dfrac{0.95}{0.05}$ "che il Torino perde è dato 95:5"

Formalmente

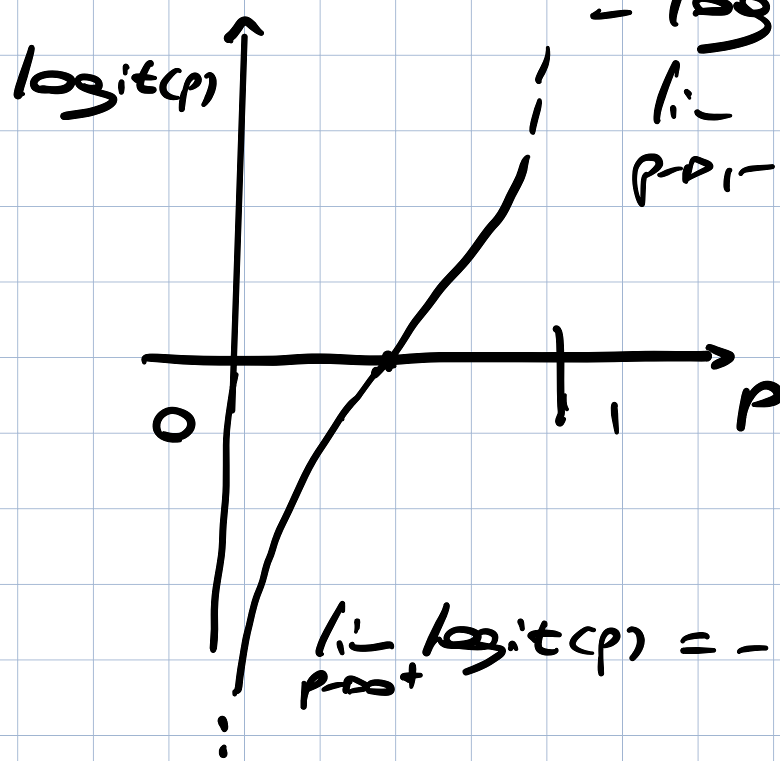
$$\text{odds}(p) = \frac{p}{1-p}$$



e

$$\text{logodds}(p) = \text{logit}(p) = \log\frac{p}{1-p}$$

$$= \log p - \log(1-p)$$



$$\lim_{p\to 1^-} \text{logit}(p) = +\infty$$

$$\lim_{p\to 0^+} \text{logit}(p) = -\infty$$

odds a posteriori $= \dfrac{P(H_0 | \text{data})}{P(\bar{H}_0 | \text{data})}$ (Bayes due volte)

$$= \frac{\dfrac{P(H_0)\,P(\text{data}/H_0)}{P(\cancel{\text{data}})}}{\dfrac{P(\bar{H}_0)\,P(\text{data}|\bar{H}_0)}{P(\cancel{\text{data}})}}$$

$$= \frac{P(H_0)}{P(\bar{H}_0)} \times \frac{P(\text{data}|H_0)}{P(\text{data}|\bar{H}_0)}$$

odds a priori

FATTORE DI BAYES, contributo dell'evidenza in favore di H_0

4) previsione un'altra volta