7.Estimació puntual bayesiana

Estadística
Grau en Matemàtiques

Josep A. Sanchez
Dept. Estadística i I.O.(UPC)

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

- En estadística clàsica (freqüentista), θ es considera un valor fix desconegut
- En l'aproximació Bayesiana θ és una variable aleatòria, i la seva variació es descriu en base a una distribució, la **distribució a priori** $\pi(\theta)$
- La **distribució a priori** $\pi(\theta)$ és subjectiva i escollida per l'investigador
- Una mostra $X_1, \ldots, X_n$ és observada, i a partir dels seus valors, la distribució del paràmetre és actualitzada
- La nova distribució del paràmetre, desprès d'observar la mostra, es coneix com **distribució a posteriori** $\pi(\theta|\underset{\sim}{X})$

Distribució a posteriori

- La **distribució a posteriori** és:

$$\pi(\theta|\underset{\sim}{X}) = \frac{f(\theta, \underset{\sim}{X})}{m(\underset{\sim}{X})} = \frac{f(\underset{\sim}{X}|\theta)\pi(\theta)}{m(\underset{\sim}{X})}$$

  amb la distribució marginal de $\underset{\sim}{X}$ en el denominador:

$$m(\underset{\sim}{X}) = \int_{\Theta} f(\underset{\sim}{X}, \theta) d\theta = \int_{\Theta} f(\underset{\sim}{X}|\theta)\pi(\theta) d\theta$$

- Una estimació puntual per θ es pot obtenir calculant l'esperança de la distribució *a posteriori*
- La distribució a posteriori és proporcional a la funció de versemblança i a la distribució a priori

$$\pi(\theta|\underset{\sim}{X}) \propto f(\underset{\sim}{X}|\theta)\pi(\theta)$$

Ejemplo: distribució binomial

Sigui $X_1, \ldots, X_n$ una m.a.s. amb distribució de Bernouilli, $\sum_{i=1}^n X_i \sim B(n,p)$ i fem servir una distribució a priori $Beta(\alpha, \beta)$ pel paràmetre p:

$$\pi(p) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1}$$

- La versemblança del paràmetre p donada la mostra és:

$$L(p; \underset{\sim}{X}) = f(\underset{\sim}{X}|p) = p^{\sum_{i=1}^n x_i}(1-p)^{n-\sum_{i=1}^n x_i}$$

- La distribució conjunta és:

$$f(\underset{\sim}{X}, p) = f(\underset{\sim}{X}|p)\pi(p) = p^{\sum_{i=1}^n x_i}(1-p)^{n-\sum_{i=1}^n x_i} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1}$$

- Per tant,

$$\pi(p|\underset{\sim}{X}) \propto \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha+\sum_{i=1}^n x_i-1}(1-p)^{\beta+n-\sum_{i=1}^n x_i-1}$$

Ejemplo: distribució binomial

- Calculem la distribució marginal de $\underset{\sim}{X}$:

$$m(\underset{\sim}{X}) = \int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha+\sum_{i=1}^n x_i - 1}(1-p)^{\beta+n-\sum_{i=1}^n x_i - 1} dp =$$

$$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha+\sum_{i=1}^n x_i)\Gamma(\beta+n-\sum_{i=1}^n x_i)}{\Gamma(\alpha+\beta+n)}$$

- la distribució aposteriori serà:

$$\pi(p|\underset{\sim}{X}) = \frac{f(\underset{\sim}{X},p)}{m(\underset{\sim}{X})} =$$

$$= \frac{\Gamma(\alpha+\beta+n)}{\Gamma(\alpha+\sum_{i=1}^n x_i)\Gamma(\beta+n-\sum_{i=1}^n x_i)} p^{\alpha+\sum_{i=1}^n x_i - 1}(1-p)^{\beta+n-\sum_{i=1}^n x_i - 1}$$

- Per tant,

$$p|\underset{\sim}{X} \sim Beta\left(\alpha+\sum_{i=1}^n x_i, \beta+n-\sum_{i=1}^n x_i\right)$$

Ejemplo: distribució binomial

A partir de la distribució a posteriori es pot obtenir un estimador puntual de p basat en el mètode bayesià:

$$\hat{p}_{MB} = \frac{\alpha + \sum_{i=1}^{n} x_i}{\alpha + \beta + n}$$

Si la esparança de la distribució a priori és $E_\pi(p) = \frac{\alpha}{\alpha+\beta}$, es pot comprovar que l'estimador de Bayes és una mitjana ponderada entre la mitjana a priori i l'estimador màxim versemblant:

$$\hat{p}_{MB} = \left(\frac{\alpha+\beta}{\alpha+\beta+n}\right)\left(\frac{\alpha}{\alpha+\beta}\right) + \left(\frac{n}{\alpha+\beta+n}\right)\bar{X}$$

Exemple numèric

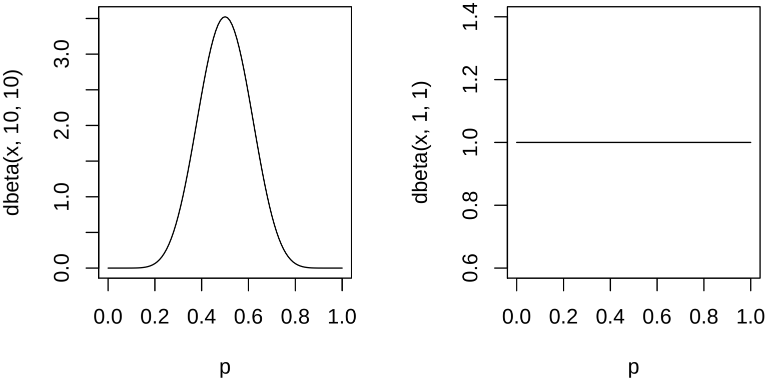
Es llença una moneda 25 vegades:

C + C + C C + C C C + C + C C + C + + + C C C C C

Sigui p la probabilitat de C

- Calcula l'estimador de màxima versemblança de p
- Calcula una estimació bayesiana de p, fent servir a prior Beta(10,10)
- Calcula una estimació bayesiana de p, fent servir a prior Beta(1,1)

Exemple numèric



Exemple numèric

- $X = 16$
- $\hat{p}_{ML} = \bar{X} = \frac{16}{25} = 0.64$
- Si $p \sim Beta(\alpha, \beta)$, sabem que
  $p|\underset{\sim}{X} \sim Beta\left(\alpha + \sum_{i=1}^{n} x_i, \beta + n - \sum_{i=1}^{n} x_i\right)$
  - $\alpha = 10$ i $\beta = 10 \quad \Rightarrow p|X = 16 \sim Beta(26, 19)$

    $$\Rightarrow \hat{p}_{MB} = E(\pi(\theta|X)) = \frac{26}{26+19} = 0.58$$

  - $\alpha = 1$ i $\beta = 1 \quad \Rightarrow p|X = 16 \sim Beta(17, 10)$

    $$\Rightarrow \hat{p}_{MB} = E(\pi(\theta|X)) = \frac{17}{17+10} = 0.63$$

- Sigui $\mathfrak{F}$ una clase de funcions de densitats, indexades per θ

$$\mathfrak{F} = \{f(x|\theta) : \theta \in \Theta\}$$

- Sigui Π una clase de distribucions en Θ
- Π és una **família conjugada** per $\mathfrak{F}$ si la distribució a posteriori està en la classe Π per totes les mostres $X \in \mathfrak{X}$ per totes les priors $\pi \in \Pi$ i per totes les $f \in \mathfrak{F}$
- Per exemple, la distribució Beta és prior conjugada per la distribució Binomial

Exemple: Exponencial distribution

- Sigui $X_1, \ldots, X_n$ una m.a.s. amb $X_i \sim Exp(\theta)$

$$f(x|\theta) = \theta e^{-\theta x} \quad x > 0$$

- Considerem com a distribució a priori per θ la distribució $Gamma(\alpha, \beta)$

$$\pi(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\beta\theta} \quad \theta > 0$$

- Llavors, la distribució a posteriori és $Gamma(\alpha + n, \beta + \sum_{i=1}^n x_i)$
- La distribució Gamma és prior conjugada per a la distribució exponencial

Exemple: Poisson distribution

- Sigui $X_1, \ldots, X_n$ una m.a.s. amb $X_i \sim \textit{Poisson}(\theta)$

$$P(X = x|\theta) = e^{-\theta}\frac{\theta^x}{x!}$$

- Considerem com a distribució a priori per θ la distribució $\textit{Gamma}(\alpha, \beta)$

$$\pi(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\beta\theta} \quad \theta > 0$$

- Llavors, la distribució a posteriori és $\textit{Gamma}(\alpha + \sum_{i=1}^n x_i, \beta + n)$
- La distribució Gamma és prior conjugada per a la distribució de Poisson

Exemple: Normal distribution

- Sigui $X_1, \ldots, X_n$ una m.a.s. amb $X_i \sim N(\theta, \sigma^2)$ amb σ^2 coneguda

$$f(x|\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\theta)^2}{2\sigma^2}\right)$$

- Considerem com a distribució a priori per θ la distribució normal $N(\mu, \nu^2)$

$$f(x|\mu, \nu^2) = \frac{1}{\sqrt{2\pi\nu^2}} \exp\left(-\frac{(\theta-\mu)^2}{2\nu^2}\right)$$

- Llavors, la distribució a posteriori és $N(\mu_1, \nu_1^2)$ amb

$$\mu_1 = \frac{\sigma^2\mu + n\nu^2\bar{X}}{\sigma^2 + n\nu^2} = \frac{\sigma^2/n}{\sigma^2/n + \nu^2}\mu + \frac{\nu^2}{\sigma^2/n + \nu^2}\bar{X}$$

$$\nu_1^2 = \frac{\sigma^2\nu^2}{\sigma^2 + n\nu^2} = \frac{1}{1/\nu^2 + 1/(\sigma^2/n)}$$

Resum de famílies conjugades

Data	Prior	Posterior
$X_i \sim Bern(p)$	$p \sim Beta(\alpha, \beta)$	$Beta(\alpha + \sum_{i=1}^n x_i, \beta + n - \sum_{i=1}^n x_i)$
$X_i \sim exp(\theta)$	$\theta \sim \Gamma(\alpha, \beta)$	$\Gamma(\alpha + n, \beta + \sum_{i=1}^n x_i)$
$X_i \sim Poisson(\theta)$	$\theta \sim \Gamma(\alpha, \beta)$	$\Gamma(\alpha + \sum_{i=1}^n x_i, \beta + n)$
$X_i \sim N(\theta, \sigma^2)$	$\theta \sim N(\mu, \nu^2)$	$N(\mu_1 = \frac{\sigma^2\mu + n\nu^2\bar{X}}{\sigma^2 + n\nu^2}, \nu_1^2 = \frac{\sigma^2\nu^2}{\sigma^2 + n\nu^2})$

Observacions seqüencials

- En molts experiments les observacions $X_1, \ldots, X_n$ se obtienen de forma seqüencial
- El parametre θ es pot actualitzar desprès d'obtenir cada observació
- Si tenim la prior $\pi(\theta)$ i observem $X_1 = x_1$, llavors la posterior

$$\pi(\theta|x_1) \propto f(x_1|\theta)\pi(\theta)$$

  representa el coneixement per θ i es fa servir com a prior per a la propera observació

Observacions seqüencials

- Quan observem $X_2 = x_2$

$$\pi(\theta|x_1, x_2) \propto f(x_2|\theta)\pi(\theta|x_1) \propto f(x_2|\theta)f(x_1|\theta)\pi(\theta)$$

- Després d'obtenir tota la mostra $X_1 = x_1, \ldots, X_n = x_n$ tindrem

$$\pi(\theta|x_1, \ldots, x_n) \propto f(x_n|\theta)\pi(\theta|x_1 \ldots x_{n-1})$$

$$\pi(\theta|x_1, \ldots, x_n) \propto f(x_n|\theta) \ldots f(x_1|\theta)\pi(\theta)$$
$$\propto f(x_1, \ldots, x_n|\theta)\pi(\theta)$$

- La posterior obtinguda després d'observar n casos de X és la mateixa si prenem les observacions seqüencialment o totes de cop

Funcions de perdua

- Per obtenir un estimador puntual del paràmetre hem d'indicar un criteri que ens permeti definir l'estadístic a partir de la distribució a posteriori
- Volem estimadors $\hat{\theta}$ de θ que estiguin a prop
- Indiquem una funció de pèrdua que volem que l'estadístic minimitzi
  - Pèrdua Absoluta: $L(\theta, \hat{\theta}) = |\hat{\theta} - \theta| \Rightarrow \hat{\theta} = Med(\theta|\underset{\sim}{X})$
  - Pèrdua Quadràtica: $L(\theta, \hat{\theta}) = (\hat{\theta} - \theta)^2 \Rightarrow \hat{\theta} = E(\theta|\underset{\sim}{X})$