Norwegian University of Science and Technology
Department of Mathematical Sciences

TMA4295 Statistical inference

Lecturer Fall 2023: *Gunnar Taraldsen*

Scribe: *Solène Couchot* and *Yawar Mahmood*

Lecture 1 in week 39: 'Bayesian estimates'

Contents

1 Bayesian point estimate with a group structure

We assume model space Ω_Θ is a group.

Definition 1.1 (Right invariant prior) $P_{\Theta u} = P_\Theta$ in the sense that $\Theta u \sim \Theta$

Example 1.1 $\pi(\theta) = \frac{1}{\theta} d\theta$ with multiplication as group operation.

Example 1.2 $\pi(\theta) = 1$ with addition as group operation.

Example 1.3 $\pi(\mu, \sigma) = \frac{1}{\sigma} d\mu d\sigma$ is right invariant for the location scale group.

Definition 1.2 (Group model) $Data = (\theta u, a) = (x, a)$ where a is ancillary and $\theta, u \in \Omega_\Theta = G$ a group

Example 1.4 (Location Scale) If $\theta = (\mu, \sigma),\ u = (\overline{Z}, S_Z)$
We get $\theta u = (\overline{X} = \mu + \sigma \overline{Z}, S = \sigma S_Z)$

Theorem 1.1 (Fiducial posterity) Let $X = \theta u, a$ be the data. Then

$$(\Theta = xU^{-1} | A = a) \sim \Theta | (x, a)$$

starting with a right invariant prior.

Proof. Let ϕ and ψ be two functions.
Want to show that:

$$E[\phi(X,A)\psi(\Theta)]$$

is equal to, and can be written as:

$$E[\phi(X,A)E[\psi(\Theta)|(X,A)]]$$

Start by writing the integral form of the expectation value:

$$\int\int\int \phi(\theta u, a)\psi(\theta)P_{\Theta u}(d\theta)P_U^a(du)P_A(da)$$

Lets do a change of variable. From our theorem, note $\theta u = x$:

$$\int\int\int \phi(x,a)\psi(xu^{-1})P_{\Theta u}(dx)P_U^a(du)P_A(da)$$

Because of right invariance, $P_{\Theta u}$ does not depend on u. Hence:

$$\int\int \phi(x,a)[\int \psi(xu^{-1})P_U^a(du)]P_{\Theta}(dx)P_A(da)$$

Note that $P_\Theta(dx)P_A(da)$ is the joint distribution of (x,a) and that $[\int \psi(xu^{-1})P_U^a(du)]$ is equal to $E[\psi(\Theta)|(x,a)]$ Since:

$$E[\phi(X,A)\psi(\Theta)] = E[\phi(X,A)E[\psi(\Theta)|(X,A)]]$$

□

Theorem 1.2 (Right Invariant Prior and optimal equivariant estimator) The right invariant prior (if it exists) determines an optimal equivariant estimator given an invariant loss

Proof. The risk ρ:

$$\rho = E[l(t(\theta U, A), \theta)]$$

where l is the loss function and $t(\theta U, A)$ is equivariant.

$$E[l(\theta t(U,A), \theta)]$$

The expression does not depend on θ, since l is invariant. We can write this as an double expectation:

$$E[E[l(t(x,A), xU^{-1})|A]]$$

Here $p^{x,a} = E(l(t(X,A),\Theta))|_{x,a}$ is the Bayes post risk.
If we take the conditional on $A = a$, we can write $E[l(t(x,A), xU^{-1})|A]$ as an integral. □

Furthermore, regarding the previously proof, it is important to observe that we can argue that the double expectation mentioned does not rely on the variable x either, since the initial expectation is not dependent upon the parameter θ.

Example 1.5 (Location-Scale) $\theta = (\mu, \sigma) = (x, a)$, $u = ((\overline{Z}, S_Z))$
Define $\overline{x} = \mu + \sigma \overline{z}$, $S^2 = \sigma^2 S_z^2$
Note, to keep location scale we should have defined $S = \sigma S_z$, but we are not to worried about that here. Solving to find Fiducial (posterior!).
Define precision: $\lambda = \sigma^{-2}$. Hence, $S^2 = \lambda^{-1} S_z^2$
Solving for λ, $\lambda = \frac{S_z^2}{S^2} \sim Gamma(\alpha, \beta)$
Also: $(\mu = \overline{x} - \sigma \overline{z} | \lambda) \sim N(\overline{x}, \cdot)$

2 Conjugate families

Definition 2.1 (Conjugate family of priors) A family of priors is a conjugate family if any conjugate prior gives a posterior in the conjugate family

Example 2.1 (Beta family) $f(x) = p^x (1-p)^{1-x} \sim B(p)$ $(x \in \{0, 1\})$
Prior with $Beta(\alpha, \beta)$ distribution : $\pi(p) \propto p^{\alpha-1}(1-p)^{\beta-1}$
Thus

$$\pi(p|x) \propto p^{\alpha+x-1}(1-p)^{\beta+y-1} \sim Beta(\alpha + x, \beta + y)$$

(where $y = 1 - x$)
$Beta(\alpha, \beta)$ is a conjugate prior family for $B(p)$.
$Beta(\alpha + x_1 + x_2, \beta + y_1 + y_2)$ is the posterior from a new $B(p)$ using the previous posterior as its prior.

Example 2.2 (Exponential family) We once again consider the example above, but this time with an exponential family.
$f(x) = p^x (1-p)^{1-x} \sim B(p)$ $(x \in \{0, 1\})$
Thus,

$$f(x) = e^{x ln(p) + (1-x) ln(1-p)} = e^{x \cdot \eta - \gamma}$$

With $\eta = ln\frac{p}{1-p}$ and $\gamma = -ln(1-p)$
We have $\gamma \in (\mathbb{R}, +)$ a group.
If $\alpha = 0 = \beta$, then $Beta(0,0) \sim p \iff \gamma \sim U(-\infty, +\infty)$: length is invariant with respect to shift!

Note that conjugate families are *not* unique. They depend on the choice of a reference measure on $R(\Theta)$.

Example 2.3 (Beta family) Beta case follows by starting with $Beta(0,0)$ as reference: $\mu(dp) = \pi(p)dp = p^{-1}(1-p)^{-1}dp$

Example 2.4 (Location-Scale) $m(d\mu, s\sigma) = \frac{1}{\sigma}d\sigma d\mu$

Now, lets look at the exponential family on canonical form:

$$g(x)e^{x\theta-\gamma}$$

where $\gamma = \gamma(\theta)$.
We know that:

$$1 = \int g(x)e^{x\theta-\gamma}\mu(dx)$$

Hence:

$$e^{\gamma} = \int g(x)e^{x\theta}\mu(dx)$$

where x and θ do not need to be real numbers.

Considering a normal distribution with parameters μ and σ^2:

$$\begin{aligned} f(x) &= \frac{1}{\sqrt{2\pi}}\frac{1}{\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \\ &= \frac{1}{\sqrt{2\pi}}e^{-ln(\sigma)+\frac{-x^2+2\mu x-\mu^2}{2\sigma^2}} \\ &= \frac{1}{\sqrt{2\pi}}e^{[-x^2,x]\cdot[\frac{1}{2\sigma^2},\frac{\mu}{\sigma^2}]-(ln(\sigma)+\frac{\mu^2}{2\sigma^2})} \end{aligned}$$

Here we can see that the canonical statistic is $[-x^2, x]$, the canonical parameter $\theta = [\frac{1}{2\sigma^2}, \frac{\mu}{\sigma^2}]$ and $\gamma = (ln(\sigma) + \frac{\mu^2}{2\sigma^2})$. This is indeed a part of the exponential family.

For this case, $E[X] = \gamma'$. The MLE of $\hat{\theta}$ is determined by $x = \gamma'(\hat{\theta})$. If γ is convex, we also know that $\gamma'' \geq 0$.

For the prior density : $e^{a\theta - b\gamma}$ with a and b constants:
The posterior density is: $\pi(\theta|x) \propto e^{(a+x)\theta-(b+1)\gamma}$

Definitions

Theorems

Examples

Index