Reema Mo'am

PAGE NO.:
DATE: / /

Non - Informative Prior

A prior which contains no information about the parameter (say θ) (or more prudely which favors no possible values of θ over others, maybe called a non-informative prior.

For eg :- In testing b/w two simple hypothesis, prior which keeps the probability half to each of the hypothesis is clearly non-informative.

$$N(\theta, 1) \; ; \quad \theta \in (-\infty, \infty)$$

$$\left.\begin{array}{l} \pi(\theta) \propto C \\ \qquad \propto 1 \end{array}\right\}$$

Remark : It will frequently happen that the natural non-informative prior is an improper prior namely one which has infinite mass.

Determination of non-informative priors

1. When θ is finite

Suppose parameter space Θ is finite and it contains n elements. The obvious non-informative prior is then to given each element of Θ has probability $\frac{1}{n}$ is proper but if $n = \infty$ then it is improper.

PAGE NO.: ______
DATE: / /

One might generalise it by considering $n = \infty$ to a improper prior.

A generalisation of $g(\theta) = \frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}$ to infinite θ may be proportional to a constant $\forall\ \theta \in \Theta$

Consider a non-informative prior for parameter θ, $\pi(\theta) \propto$ a constant (say c)

$$\pi(\theta) \propto c$$

Instead of considering θ, suppose the problem had been parametrized in terms of $\eta = \exp(\theta)$.

$$\theta \sim \pi(\theta)$$

$$\pi(\theta) \propto c$$

$$\eta = \exp(\theta)$$

We assume,

$$\pi(\eta) \propto \text{const}.$$

But, if $\pi(\theta)$ is the density for θ, then the corresponding density for η is as derived below,

$$\eta = \exp(\theta)$$

$$\theta = \log \eta$$

$$\frac{d\theta}{d\eta} = \frac{1}{\eta}$$

$\therefore$ Density of η

$$\pi(\eta) = \pi(\theta) \cdot \frac{d\theta}{d\eta}$$

$$= \frac{1}{\eta} \cdot \pi(\theta)$$

$$= \frac{1}{\eta} \pi(\log(\eta))$$

This is called lack of invariance of transformation.

Hence, if the non-informative for θ is chosen to be constant, we should choose the non-informative prior for η to be proportional to $1/\eta$ to maintain consistency [and arrive at the same answers in either parametrization].

Thus, we cannot maintain consistency and choose both the non-informative prior for θ and that for η to be constant.

The lack of invariance of constant prior has lead to a search for non-informative prior which are approximately invariant under transformation.

→ Non-informative priors for location and scale parameter.

i) Location Invarient Prior (LIP)

Suppose sample space $\mathcal{X}$ & parameter space Θ both are real and x has pdf $f(x,\theta)$ which is of the form $f(x-\theta)$, i.e. it depends only on $(x-\theta)$. The density is then set to be a location density, and θ is called a location parameter or sometimes a location vector.

PAGE NO.:
DATE: / /

For example $N(\theta,1)$ and $\exp(\theta,\mu)$ where $\theta \in (-\infty,\infty)$ is called location parameter.

To derive a non-informative prior for this situation imagine instead of observing X we observe $Y = X + c$ where c is some constant.
Define $\eta = \theta + c$ and Y has density $f(y-\eta)$, so both y & η are real.

The sample space and parameter space for both (x,θ) and (y,η) problems are thus identical in structure and it seems reasonable to insist that they have same non-informative prior.

Another way of thinking of this is to know that observing Y really amounts to observing X with a different unit of measurement, one in which the origin is c and not c. Since, the choice of an origin for a unit of measurement is quite aribtrary, the non-informative prior should perhaps be independent of this choice.

Let π & π^* denote the non-informative priors in the (x,θ), (y,η) problems respectively. We may assume that π & π^* are equal for any real space A

$$P^{\pi}(\theta \in A) = P^{\pi^*}(\eta \in A) \quad \text{---} \quad (1)$$

Since, $\eta = \theta + c$, it should also be true (by a simple change of variable) that $P^{\pi^*}(\eta \in A) = P^{\pi}(\theta + c \in A)$

$$P^{\pi^*}(\eta \in A) = P^{\pi}(\theta \in A - c) \quad \text{---} \quad (2)$$

PAGE NO.:
DATE: / /

where $A-c = \{z-c ; z \in A\}$

From eq(uns) (1) & (2), $P^{\pi}(\theta \in A) = P^{\pi}(\theta \in A-c)$ — (3)

so, it should hold $\forall$ c. Any π satisfying this relationship is said to be Location Invariant Prior.

22/07/23

ii) Scale invarient Prior

A scale density (one-dimensional) is a density of form $\frac{1}{\sigma} f\left(\frac{x}{\sigma}\right)$

where X is a r.v having the density function $f(x,\sigma)$

$$X \sim f(x,\sigma) = \frac{1}{\sigma} f\left(\frac{x}{\sigma}\right) \quad ; \sigma > 0$$

~~where~~ The parameter σ is called scale parameter.

Eg:- In ~~normal~~ $N(\mu, \sigma^2)$ and $\exp(\theta)$, the parameters σ and θ are scale parameters of normal distribution and exponential distribution, respectively.

Imagine that, instead of observing X, we observe the r.v $Y = cX$ where $c > 0$ is a constant. Defining $\eta = c\sigma$, an easy calculation shows that the density of Y is $\frac{1}{\eta} f\left(\frac{y}{\eta}\right)$.

Let sample space, * for the r.v X is * $(0,\infty)$ and parametric space Θ is $(0,\infty)$. Then, same sample space and parametric space will be corresponding to r.v Y.

$X \rightarrow$
$Y = cX$
$\eta = c\sigma$

PAGE NO.:
DATE: / /

Since, the sample and parameter space for the (X, σ) problem are the same as those for the (Y, η) problem. The two problems are thus identical in structure, which again indicates that they should have the same non-informative prior.

Let π and π^* denote the priors in the (X, σ) and (Y, η) problems, respectively, this means the equality

$$P^{\pi}(\sigma \in A) = P^{\pi^*}(\eta \in A) \quad \text{should hold } \forall A; \; A \subset (0, \infty). \quad \text{(1)}$$

Since, $\eta = c\sigma$, it should also be true that,

$$P^{\pi^*}(\eta \in A) = P^{\pi}(c\sigma \in A)$$

$$= P^{\pi}(\sigma \in c^{-1}A) \quad - \text{(2)}$$

From eq$^{(n)}$ (1) & (2),

$$P^{\pi}(\sigma \in A) = P^{\pi}(\sigma \in c^{-1}A) \quad - \text{(3)} \; ; \forall c > 0$$

This should hold $\forall$ $c > 0$ and any distribution π for which eq$^{(n)}$ (3) is true is called <u>scale invarient</u>.

The mathematical analysis of eq$^{(n)}$ (3) proceeds as in the location invarient density.

$$\int_A \pi(\sigma)\, d\sigma = \int_{c^{-1}A} \pi(\sigma) \cdot d\sigma$$

Pg - 86 & 87 is upto exp 3.5 niche
↓
last para

PAGE NO.:
DATE: / /

$$= \int_A \pi(c^{-1}\sigma) \cdot c^{-1}\, d\sigma$$

and conclude that for this to hold $\forall A$ it must be true that

$$\pi(\sigma) = c^{-1}\pi(c^{-1}\sigma)$$

$$\pi(\sigma) = \frac{1}{c}\pi\left(\frac{\sigma}{c}\right) \qquad \forall \sigma$$

Choosing $\sigma = c$,

$$\pi(c) = \frac{1}{c}\pi(1)$$

Setting $\pi(1) = 1$ for convenience, and noting that the above equality must hold $\forall\ c > 0$, it follows that a reasonable non-informative prior for a scale parameter is ~~$\pi(\sigma$~~

$$\boxed{\pi(\sigma) = \frac{1}{\sigma}}$$

~~Ob~~ It is to be noted that this is also an improper prior since ~~integral over 0 to ∞~~

$$\int_0^\infty \frac{1}{\sigma}\, d\sigma = \infty$$

good