4/2/2021. LECTURE 2.

Bayesian approach:

- Specify a prior distribution $f(\theta)$ for θ
- Combine this with the likelihood $f(x|\theta)$ to obtain:
- Posterior distribution $f(\theta|x)$ for θ given x using Bayes' theorem,

$$\underbrace{f(\theta|x)}_{\text{POSTERIOR}} = \frac{f(x|\theta)f(\theta)}{\boxed{f(x)}} \propto \underbrace{f(x|\theta)}_{\text{LIKELIHOOD}}\underbrace{f(\theta)}_{\text{PRIOR}}$$

function of θ | does not depend on θ | with respect to θ

Bayesian analysis is concerned with distributions of θ and how they change in the light of new evidence (typically data)

With this approach, we can answer the question

"What value of θ is most likely given the data?"

rather than the question

"What value of θ makes the data most likely to occur?"

- Distribution $f(x|\theta)$ irrelevant to Bayesian after the data has been observed.
- For the classicist, $f(x|\theta)$ is the <u>only</u> distribution available.

1. THE BAYESIAN METHOD.

Note: Use $f(\cdot)$ to represent density function irrespective of whether the RV are discrete or continuous

Generally make no distinction as to whether variables are univariate or multivariate.

1.1 BAYES' THEOREM.

Let X and Y be RVs with joint density $f(x,y)$. The MARGINAL DENSITY of Y, $f(y)$ is

$$f(y) = \int_X f(x,y)\,dx$$

[NB. X could be multivariate e.g. $X = (X_1, X_2)$ then

$$f(y) = \int_{X_1}\int_{X_2} f(x_1, x_2, y)\,dx_1\,dx_2 \qquad]$$

The CONDITIONAL DISTRIBUTION of Y given $X = x$ is, for $f(x) > 0$

$$f(y|x) = \frac{f(x,y)}{f(x)}$$

so that, for example,

$$f(y) = \int_X f(y|x) f(x)\,dx$$

which is the THEOREM OF TOTAL PROBABILITY.

X and Y are INDEPENDENT if and only if, for all x and y,

$$f(x,y) = f(x) f(y) \qquad \text{[Often write } X \perp\!\!\!\perp Y \text{]}$$

[An equivalent condition is that, for $f(x) > 0$,

$$f(y|x) = \frac{f(x,y)}{f(x)} \overset{X \perp\!\!\!\perp Y}{=} \frac{f(x)f(y)}{f(x)}$$

i.e. $f(y|x) = f(y)$

Nothing further can be learnt about Y by observing X]

If Z is a third RV then X and Y are CONDITIONALLY INDEPENDENT given Z if and only if, for all x, y, z

$$f(x,y|z) = f(x|z)f(y|z)$$

[Written $(X \perp\!\!\!\perp Y)|Z$]

[NB. $f(y|x,z) = \frac{f(x,y,z)}{f(x,z)} = \frac{f(x,y|z)/f(z)}{f(x|z)/f(z)}$

$$= \frac{f(x,y|z)}{f(x|z)} = f(y|z)$$

Having observed Z, nothing further can be learnt about Y by observing X].

Bayes' theorem states that, for $f(x) > 0$

$$f(y|x) = \frac{f(x|y)f(y)}{f(x)} = \frac{f(x|y)f(y)}{\int_Y f(x|y)f(y)\,dy}$$

1.2 BAYES' THEOREM FOR PARAMETRIC INFERENCE

In a Bayesian analysis, we treat parameter θ as a RV and thus may specify a density $f(\theta)$. If we have data x

$$f(\theta|x) = \frac{f(x|\theta)f(\theta)}{f(x)} = \frac{f(x|\theta)f(\theta)}{\boxed{\int_\theta f(x|\theta)f(\theta)\,d\theta}}$$

normalising constant

i.e. $f(\theta|x) \propto f(x|\theta)f(\theta)$

Posterior $\propto$ Prior $\times$ Likelihood

Typically θ and x are continuous.

[e.g. $X|\theta \sim N(\theta, \sigma^2)$ σ^2 known, $\theta \in \mathbb{R}$

Alternatively $\theta = (\mu, \sigma^2)$ both unknown, $X|\theta \sim N(\mu, \sigma^2)$]

or θ is continuous and x discrete

[e.g. $X|\theta \sim Bin(n, \theta)$]

In exceptional cases, θ could be discrete.

The Bayesian method comprises of the following principle steps:

1. PRIOR

   Obtain prior density $f(\theta)$. This expresses our knowledge about θ prior to observing the data.

2. LIKELIHOOD.

   Obtain likelihood function $f(x|\theta)$. This step simply describes the process giving rise to x in terms of θ.

3. POSTERIOR

   Apply Bayes' theorem to derive posterior $f(\theta|x)$. Expresses our knowledge about θ after observing the data.

4. INFERENCE.

   Derive appropriate inference statements from the posterior distribution. e.g. point estimate, interval estimate, probability of a hypothesis, ...