24/2/2022. LECTURE 6: NOTES MADE ON THE VISUALISER.

Last time: $\theta \sim N(\mu_0, \sigma_0^2)$, $X|\theta \sim N(\theta, \sigma^2)$, $\theta|x \sim N(\mu_1, \sigma_1^2)$

where: $\sigma_1^2 = \left(\frac{1}{\sigma_0^2} + \frac{1}{\sigma^2}\right)^{-1}$ i.e. $\frac{1}{\sigma_1^2} = \frac{1}{\sigma_0^2} + \frac{1}{\sigma^2}$

Posterior precision = prior precision + data precision

$$\mu_1 = \left(\frac{1}{\sigma_0^2} + \frac{1}{\sigma^2}\right)^{-1}\left(\frac{\mu_0}{\sigma_0^2} + \frac{x}{\sigma^2}\right)$$

$$= \alpha\mu_0 + (1-\alpha)x \qquad \alpha = \frac{1/\sigma_0^2}{1/\sigma_0^2 + 1/\sigma^2}$$

High confidence $\equiv$ Low variance $\equiv$ High precision

Strong prior $\Rightarrow$ large $1/\sigma_0^2$

increase weight on μ_0 in the posterior

Weak prior $\Rightarrow$ small $1/\sigma_0^2$

increase weight on x in the posterior (relative)

Given a normal likelihood (with known variance), the normal distribution is a conjugate family.

1.5 Using the posterior for inference.

Our posterior beliefs are captured by $f(\theta|x)$, a whole distribution.

Typically we want to summarise this distribution.

e.g. Graphs and plots of shape; measures of location (mean, median, mode); measures of dispersion (variance, quantiles);

region that captures most of the values of θ, ...

We'll focus on regions. In the following assume that θ is univariate (results and definitions simply generalise for θ multivariate when appropriate).

1.5.1 Credible intervals and highest density regions.

Credible intervals (or posterior intervals) are the Bayesian analogue of confidence intervals.

A $100(1-\alpha)\%$ CI is interpreted as meaning that with a large number of repeated samples, $100(1-\alpha)\%$ of the intervals would contain the true value of θ.

e.g. confidence interval $(\theta_l^*(x), \theta_u^*(x))$ where:

$$\mathbb{P}(\underset{\text{random}}{\theta_l^*(X)} < \underset{\text{fixed}}{\theta} < \underset{\text{random}}{\theta_u^*(X)} \mid \theta) = 1-\alpha$$

(constructed using $f(x \mid \theta)$)

i.e. Realisation of a random interval, constructed using $f(x \mid \theta)$.

[e.g. $X_i \mid \theta \sim N(\theta, \sigma^2)$ σ^2 known, iid, n observations

$$\mathbb{P}\left(\bar{X} - 1.96\frac{\sigma}{\sqrt{n}} < \theta < \bar{X} + 1.96\frac{\sigma}{\sqrt{n}} \mid \theta\right) = 0.95$$

and the confidence interval is the realisation $\left(\bar{x} - 1.96\frac{\sigma}{\sqrt{n}}, \bar{x} + 1.96\frac{\sigma}{\sqrt{n}}\right)$]

A $100(1-\alpha)\%$ credible interval is an actual interval that contains θ with probability $1-\alpha$.

DEFINITION (Credible Interval).

A $100(1-\alpha)\%$ CREDIBLE INTERVAL (θ_l, θ_u) is an interval within which $100(1-\alpha)\%$ of the posterior distribution lies.

$$\mathbb{P}(\underset{\text{fixed}}{\theta_l} < \underset{\text{random}}{\theta} < \underset{\text{fixed}}{\theta_u} \mid X) = \int_{\theta_l}^{\theta_u} f(\theta|x)\, d\theta = 1-\alpha$$

Note that there are infinitely many such intervals, so we typically fix the interval by specifying the tail probabilities

$$\text{e.g.} \quad \mathbb{P}(\theta \leq \theta_l \mid X) = \frac{\alpha}{2} = \mathbb{P}(\theta \geq \theta_u \mid X).$$

We can construct similar intervals for <u>ANY</u> distribution over a univariate RV. In particular we can do this for the prior when

when the intervals are often called prior credible intervals.

EXAMPLE: Tossing coins and drawing pins; 95% confidence intervals (2½% in each tail).

	Prior	Prior Credible	Post	Posterior Credible
Coins Strong prior	Beta(20,20)	(0.3478, 0.6522) Midpoint 0.5 ← 0.3044 →	Beta(24,26) [mk 0.4]	(0.3442, 0.6173) Midpoint 0.48075 ← 0.2731 →
Drawing Pins Weak prior	Beta(2,2)	(0.0950, 0.9057) Midpoint 0.5 ← 0.8107 →	Beta(6,8) [mk 0.4]	(0.1922, 0.6842) Midpoint 0.4383 ← 0.492 →

$[\theta_l = \text{qbeta}(0.025, \alpha, \beta);\ \theta_u = \text{qbeta}(0.975, \alpha, \beta)]$