16/2/2021. LECTURE 5.

Last time: θ_c probability of successful coin toss, $\theta_c \sim Beta(20, 20)$
θ_p probability of pin-up on toss of a drawing pin, $\theta_p \sim Beta(2, 2)$.

In each case, suppose that I observe $x = 4$ in $n = 10$ tosses. Then,

$$\theta_c \mid x \sim Beta(20+4, 20+6) \sim Beta(24, 26)$$
$$\theta_p \mid x \sim Beta(2+4, 2+6) \sim Beta(6, 8)$$

We plot the prior, posterior and standardised likelihood,

$$\frac{f(x \mid \theta)}{\int_\theta f(x \mid \theta)\, d\theta}$$

[i.e. scaled to integrate to 1 over θ]

[$f(x \mid \theta) \propto \theta^4 (1-\theta)^6 = \theta^{5-1}(1-\theta)^{7-1}$. The standardised likelihood is $Beta(5, 7)$]

$\theta_c \sim$ STRONG PRIOR.

Posterior density closely follows the prior. The mode shifted to the left towards the likelihood (mode 0.4). Posterior density (slightly) taller and narrower : variance reduced.

[$\mathbb{E}(\theta_c) = 1/2$, $Var(\theta_c) = 0.006$
$\mathbb{E}(\theta_c \mid X) = \dfrac{24}{24+26} = 0.48$, $Var(\theta_c \mid X) = 0.005$].

$\theta_p \sim$ WEAK PRIOR.

Posterior density much more follows the (standardised) likelihood.
Most of the posterior information is coming from the likelihood i.e. the data.

Variance of the posterior is vastly reduced (in relation to the prior)

$[\, \mathbb{E}(\theta_p) = 0.5, \ Var(\theta_p) = 0.05, \ \mathbb{E}(\theta_p | X) = 0.43, \ Var(\theta_p | X) = 0.016 \,]$

DEFINITION (Kernel of a density).

For a RV X with density $f(x)$ if $f(x)$ can be written in the form $c\,g(x)$ when c is a constant not depending upon x then any such $g(x)$ is a KERNEL of the density $f(x)$.

EXAMPLE.

$\theta \sim Beta(\alpha, \beta)$, $\quad g(\theta) = \theta^{\alpha-1}(1-\theta)^{\beta-1}$ is a kernel.

Spotting kernels of distributions can be very useful in computing posterior distributions.

EXAMPLE.

Let $X|\theta \sim N(\theta, \sigma^2)$ so that the mean θ is unknown but the variance σ^2 is known. Suppose $\theta \sim N(\mu_0, \sigma_0^2)$ when μ_0 and σ_0^2 are known. Find the posterior distribution of θ given $X = x$.

$$f(\theta) = \frac{1}{\sqrt{2\pi}\,\sigma_0} \exp\left\{ -\frac{1}{2\sigma_0^2}(\theta - \mu_0)^2 \right\}$$

$$\propto \exp\left\{ -\frac{1}{2\sigma_0^2}(\theta - \mu_0)^2 \right\} \quad \text{(a kernel)}$$

$$\propto \exp\left\{ -\frac{1}{2\sigma_0^2}(\theta^2 - 2\mu_0\theta) \right\}$$

(kernel in its simplest form)

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

* Viewed as a function of θ, this looks like a normal distribution *

This is not a kernel of $f(x|\theta)$ as a function of θ

$$\propto \exp\left\{-\frac{1}{2\sigma^2}(\theta^2 - 2x\theta)\right\}$$

(looks like a kernel of a $N(x, \sigma^2)$)

$$f(\theta|x) \propto f(x|\theta) f(\theta)$$

$$\propto \exp\left\{-\frac{1}{2\sigma^2}(\theta^2 - 2x\theta)\right\} \exp\left\{-\frac{1}{2\sigma_0^2}(\theta^2 - 2\mu_0\theta)\right\}$$

$$= \exp\left\{-\frac{1}{2}\left[\left(\frac{1}{\sigma_0^2} + \frac{1}{\sigma^2}\right)\left(\theta^2 - 2\left(\frac{1}{\sigma_0^2} + \frac{1}{\sigma^2}\right)^{-1}\left(\frac{\mu_0}{\sigma_0^2} + \frac{x}{\sigma^2}\right)\theta\right)\right]\right\}$$

KERNEL OF A NORMAL DISTRIBUTION.

I recognise this as a kernel of a normal distribution and so

$$\theta|x \sim N(\mu_1, \sigma_1^2)$$

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

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