LMaaa HW1 EECS E6892

1

1 3 Doors D_1, D_2, D_3 $P(D_i) = 1/3$ given $i = 1, 2, 3$

* If my friend choose D_1 then the gameshowhost will open D_2 or D_3. In this case prob of the gameshowhost opening D_2 or D_3 →
$P(H_{D_2 \text{ or } D_3}) = 1/2$

$P(H_{D_2 \text{ or } D_3} | D_1) = 1/2$
$P(H_{D_2 \text{ or } D_3} | D_3) = 0$
$P(H_{D_2 \text{ or } D_3} | D_2) = 1$

$$P(D_1|H) = \frac{P(H|D_1)P(D_1)}{P(H)} = \frac{1/2(1/3)}{1/2} = 1/3$$

$$P(D_2|H) = \frac{P(H|D_2)P(D_2)}{P(H)} = \frac{1(1/3)}{1/2} = 2/3$$

The posterior prob of switching doors is higher than that if the friend does not switch.

$P(D_2|H) > P(D_1|H)$ Therefore, the friend should switch doors for better odds!

2 Bayes Rule: $p(\pi|X) = \frac{p(X|\pi)\,p(\pi)}{p(X)}$

$$p(\pi|X) \propto p(X|\pi)\,p(\pi)$$

Likelihood: $p(X|\pi) = \prod_{n=1}^{N} p(x_n|\pi)$

$$= \prod_{n=1}^{N}\prod_{j=1}^{K} \pi_j^{x_{i,j}}$$

$$= \prod_{j=1}^{K} \pi_j^{\sum_i x_{i,j}}$$

Prior: $p(\pi) = \frac{1}{B(\alpha)} \prod_{j=1}^{K} \pi_j^{\alpha_j - 1}$

$$p(\pi|X,\alpha) = \underbrace{p(X|\pi)}\;\underbrace{p(\pi|\alpha)}$$

$$= \prod_{j=1}^{K} \pi_j^{\sum_i x_{i,j}}\, \frac{1}{B(\alpha)} \prod_{j=1}^{K} \pi_j^{\alpha_j - 1}$$

$$= \frac{1}{B(\alpha)} \prod_{j=1}^{K} \pi_j^{\alpha_j - 1 + \sum_i x_{i,j}}$$

$$\alpha_j^N = \alpha_j^0 + \sum_i x_{i,j}$$

$$p(\pi|X,\alpha) = \text{Drichlet}(\alpha_j + \sum_i X_{i,j})$$

$\boxed{\text{Conjugate prior is Dirchlet}}$

The most obvious features about the posterior is that it is the prior plus the new data it has seen. Thus it is shifted by the new data. Also the parameters are independent.

3 (a) $X \sim N(\mu, \tau^{-1})$
$\mu | \tau \sim N(0, a\tau^{-1})$
$\tau \sim \text{Gamma}(b, c)$

PRIOR $= p(\mu, \tau) \longrightarrow p(\mu|\tau)\, p(\tau)$
Likelihood $= p(x|\mu,\tau)$

$$p(\mu,\tau|x) \propto p(x|\mu,\tau)\, p(\mu,\tau)$$
$$\propto p(x|\mu,\tau)\, \underbrace{p(\mu|\tau)\, p(\tau)}$$

$N(0, a\tau^{-1})\, Ga(b,c)$

$\frac{1}{\sqrt{2\pi a/\tau}} e^{-\tau/2a(\mu)^2} \cdot \frac{c^b}{\Gamma(b)} \tau^{b-1} e^{-c\tau}$

$\sim NG(0, a^{-1}, b, c)$

$$p(\mu,\tau|x) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi/\tau}} e^{-\tau/2(x_i-\mu)^2} * NG(0, a^{-1}, b, c)$$

* we know $(x_i - \bar{x})^2 + (\mu - \bar{x})^2 = (x_i - \mu)^2$

$$p(x|\mu,\tau) = \left(\frac{2\pi}{\tau}\right)^{-N/2} e^{-\tau/2((xi-\bar{x})^2 + (\bar{x}-\mu)^2)}$$
$$= \left(\frac{2\pi}{\tau}\right)^{-n/2} e^{-\tau/2(Ns + N(\bar{x}-\mu)^2)}$$

$$p(\mu,\tau|x) = \left(\frac{2\pi}{\tau}\right)^{-n/2} e^{-\tau/2(Ns+N(\bar{x}-\mu)^2)} \cdot \frac{1}{\sqrt{2\pi a/\tau}} e^{-\tau/2a(\mu)^2}$$

$$\cdot \frac{c^b}{\Gamma(b)} \tau^{b-1} e^{-c\tau}$$

→

$$\propto \tau^{b+n/2-1/2} e^{-c\tau} e^{\tau/2 \sum_{i=1}^{N}(x_i-\mu)^2 + a^{-1}\mu^2}$$
$$\propto \tau^{b+n/2-1/2} e^{-c\tau} e^{\tau/2 \sum x_i^2 - 2\mu\sum x_i + n\mu^2 + a^{-1}\mu}$$
$$\propto \tau^{b+n/2-1/2} e^{-c\tau} e^{\tau/2 \sum x_i^2} e^{\tau/2(n+1/a)\mu^2 - 2\sum x_i \mu}$$
$$\propto \tau^{b+n/2-1/2} e^{-c\tau} e^{\tau/2 \sum x_i^2)} e^{\tau/2(n+1/a)(\mu^2 - \frac{2\sum x_i}{n+1/a}\mu)}$$
$$\propto \tau^{b+n/2-1/2} e^{-c\tau} e^{\tau/2 \sum x_i^2} e^{\tau/2(n+1/a)(\mu - \frac{\sum x_i}{n+1/a})^2) + \frac{\sum x_i^2}{n+1/a}}$$
$$\propto \tau^{b+n/2-1/2} e^{-c\tau} e^{\tau/2 \sum x_i^2 - \frac{\sum x_i^2}{n+1/a}))} e^{\tau\frac{(n+1/a)}{2}(\mu - \frac{\sum x_i}{n+1/a})^2)}$$
$$\propto \tau^{b+n/2-1/2} e^{-\tau(c+1/2\sum(x_i^2) - \frac{\sum x_i^2}{n+1/a}))} e^{\frac{\tau(n+1/a)}{2}(\mu - \frac{\sum x_i}{n+1/a})^2)}$$
$$= NG\left(\frac{\sum x_i}{n+a^{-1}}, n+a^{-1}, b+n/2 + c+1/2\sum(x_i)^2 - \frac{\sum x_i}{n+a^{-1}}\right)$$

Therefore we can see the posterior is also a NG distribution of the form:

$$p(\mu,\tau|x) = NG(\mu_N, a_N, b_N, c_N)$$

$$\mu_n = \frac{a_0\mu_0 + N\bar{x}}{a_0+N}$$
$$a_N = a_0 + N$$
$$b_N = b_0 + N/2$$
$$c_N = c_0 + 1/2\sum_{i=1}^{N}(x_i-\bar{x})^2 + \frac{a_0N(\bar{x}-\mu_0)^2}{2(a_0+N)}$$

4

3 b $p(x^*|x_1,\dots x_n) = \int_0^\infty\int_{-\infty}^\infty p(x^*|\mu,\tau)\,p(\mu,\tau|x_1,\dots x_n)\,d\mu\,d\tau$

$= \int_0^\infty\int_{-\infty}^\infty p(x^*|\mu,\tau)\,p(\mu|x,\tau)\,\underbrace{p(\tau|x)}_{\text{does not depend on } \mu}\,d\mu\,d\tau$

$= \int_0^\infty p(\tau|x)\int_{-\infty}^\infty p(x^*|\mu,\tau)\,p(\mu|\tau)\,d\mu\,d\tau$

First part $\int_{-\infty}^\infty p(x^*|\mu,\tau)\,p(\mu|\tau)\,d\mu\,d\tau$

$= \frac{1}{2\pi\tau^{-1}} e^{-\tau/2(x^*-\mu)^2} \frac{1}{2\pi\tau^{-1}}\sqrt{N+a^{-1}}\, e^{\tau N+a^{-1}/2\,(\mu - \frac{n\bar{x}}{N+a^{-1}})^2)}$

$\propto e^{1/2\tau x^{*2} - 2\tau\mu x^* + \tau\mu^2 + \tau(N+a^{-1})\mu - 2n\bar{x}\tau\mu + \frac{(N\bar{x})^2}{N-1/a}}$

$\propto \underbrace{e^{(1/2\tau x^{*2} + N\bar{x}^2/N+1/a)}}_{\text{Not dependent } \mu}\, e^{\text{other terms}}$

add in other terms $\propto e^{1/2\tau(1+N+a^{-1})\mu^2 - 2\tau(x^*+N\bar{x})\mu)}$

$\propto e^{1/2\tau(1+N+a^{-1})\mu - \frac{\tau(x^*+N\bar{x})}{\tau(1+N+a^{-1})})^2) - \frac{(x^*+N\bar{x})^2}{1+N+a^{-1}}}$

Now integrating $= \int_{-\infty}^\infty \quad 1 \quad = \underbrace{\sqrt{2\pi}/\sqrt{\tau(1+N+a^{-1})}}$

$\frac{\tau\sqrt{N+a^{-1}}}{2\pi}$ combine w/ other constants

$= \frac{\tau\sqrt{N+a^{-1}}}{2\pi}\cdot\sqrt{2\pi}/\sqrt{\tau(1+N+a^{-1})}$

$\cdot e^{1/2\tau x^{*2} + \frac{\tau(N\bar{x})^2}{N+a^{-1}} - \frac{\tau(x^*+N\bar{x})^2}{1+N+a^{-1}}}$

Rearraging $\int_{-\infty}^\infty p(x^*|\mu,\tau)\cdot p(\mu|\tau)\,d\mu \sim N\left(\frac{N\bar{x}}{N+1/a}, \frac{1+N+1/a}{(N+1/a)\tau}\right)$

$\rightarrow$

Now solve integral of NG given definitions from 3a →

$$= \int \text{Gamma}(b, c) \cdot N(\mu, \tau/a)\, d\tau$$

$$= \int \frac{c^b}{\Gamma(b)} \tau^{b-1} e^{-\tau c} \left(\frac{\tau}{2\pi b}\right)^{1/2} e^{-\tau/2b (x^* - \mu_n)^2} d\tau$$

$$= \frac{c^b}{\Gamma(b)\sqrt{2\pi a}} \int \tau^{b+1/2-1} e^{\tau (c + 1/2a (x^* - \mu_n)^2)} d\tau$$

This looks like a t-distribution:

$$= \frac{\Gamma\left(\frac{2b+1}{2}\right)}{\Gamma\left(\frac{2b}{2}\right)} (2\pi a c)^{-1/2} \left(1 + \frac{1}{2ac} (x^* - \mu_n)^2\right)^{-2b+1/2}$$

Thus we can the final form is a t-dist w/ 3 parameters:

$$p(x^*|x) = \pi^{-1/2} \frac{\Gamma(2b+1/2)}{\Gamma(2b/2)} \left(\frac{\Lambda}{2b_n}\right)^{1/2} \left(1 + \frac{\Lambda (x - \mu_n)^2}{2b_n}\right)^{-2b+1/2}$$

μ_n = center

Λ = precision $= \frac{b_n a_n}{c_n(a_n+1)}$

$df = 2b_n$