Stat 200A: Homework 2

1 Work out the mixture model example.

X is a binomial random variable following

$$X = \begin{cases} x_1 & p(X = x_1) = \lambda \\ x_2 & p(X = x_2) = 1 - \lambda \end{cases}$$

and Y is a continuous random variable. Given $f(Y = y|X = x_1)$ and $f(Y = y|X = x_2)$, which can be written as $f(y|x_1)$ and $f(y|x_2)$ for simplicity without confusion, then

$$f(y) = f(y|x_1)p(x_1) + f(y|x_2)p(x_2) = \lambda f(y|x_1) + (1 - \lambda)f(y|x_2)$$

also X can be classified according to observed Y with

$$f(x_1|y) = \frac{f(y|x_1)p(x_1)}{f(y)} = \frac{\lambda f(y|x_1)}{\lambda f(y|x_1) + (1 - \lambda)f(y|x_2)}$$

$$f(x_2|y) = \frac{f(y|x_2)p(x_2)}{f(y)} = \frac{(1 - \lambda)f(y|x_2)}{\lambda f(y|x_1) + (1 - \lambda)f(y|x_2)}$$

2 Work out the rejection sampling example.

Algorithm (von Neumann):

1. Construct $g(x)$ which we can sample from and for all x, $c \cdot g(x) \geqslant f(x)$.
2. Sample $X = x$ from $g(x)$.
3. Accept x with probability $\frac{f(x)}{c \cdot g(x)}$.
4. If x is accepted, return x; otherwise go to step 2.

$$f(x|\text{sample accepted}) = \frac{p(\text{sample accepted}|x)p(x)}{\sum_t p(\text{sample accepted}|t)p(t)} = \frac{\frac{f(x)}{c \cdot g(x)} \cdot g(x)}{\sum_t \frac{f(t)}{c \cdot g(t)} \cdot g(t)} = f(x)$$

So the samples produced by this algorithm have the distribution of $f(x)$.

3 Work out the bivariate normal example. Assuming that $[Y|X]$ is a normal distribution with mean X, derive the marginal distribution $[Y]$ and the conditional distribution $[X|Y]$.

Let

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \qquad f_{Y|X}(y|x) = \frac{1}{\sqrt{2\pi r^2}} \exp\left[-\frac{(y-x)^2}{2r^2}\right]$$

then

$$\begin{aligned}
f_Y(y) &= \int_{-\infty}^{+\infty} f_{Y|X}(y|x) f_X(x) dx \\
&= \frac{1}{2\pi\sqrt{\sigma^2 r^2}} \int_{-\infty}^{+\infty} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2} - \frac{(y-x)^2}{2r^2}\right] dx \\
&= \frac{1}{\sqrt{2\pi(\sigma^2+r^2)}} \cdot \frac{1}{\sqrt{2\pi \cdot \frac{1}{\frac{1}{\sigma^2}+\frac{1}{r^2}}}} \int_{-\infty}^{+\infty} \exp\left[-\frac{(x-\frac{\frac{\mu}{\sigma^2}+\frac{y}{r^2}}{\frac{1}{\sigma^2}+\frac{1}{r^2}})^2}{2\cdot\frac{1}{\frac{1}{\sigma^2}+\frac{1}{r^2}}}\right] \exp\left[-\frac{(y-\mu)^2}{2(\sigma^2+r^2)}\right] dx \\
&= \frac{1}{\sqrt{2\pi(\sigma^2+r^2)}} \exp\left[-\frac{(y-\mu)^2}{2(\sigma^2+r^2)}\right] \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi \cdot \frac{1}{\frac{1}{\sigma^2}+\frac{1}{r^2}}}} \exp\left[-\frac{(x-\frac{\frac{\mu}{\sigma^2}+\frac{y}{r^2}}{\frac{1}{\sigma^2}+\frac{1}{r^2}})^2}{2\cdot\frac{1}{\frac{1}{\sigma^2}+\frac{1}{r^2}}}\right] dx \\
&= N(\mu, \sigma^2+r^2) \int_{-\infty}^{+\infty} N(\frac{\frac{\mu}{\sigma^2}+\frac{y}{r^2}}{\frac{1}{\sigma^2}+\frac{1}{r^2}}, \frac{1}{\frac{1}{\sigma^2}+\frac{1}{r^2}}) dx \\
&= N(\mu, \sigma^2+r^2)
\end{aligned}$$

$$\begin{aligned}
f_{X|Y}(x|y) &= \frac{f_{Y|X}(y|x) f_X(x)}{f_Y(y)} \\
&= \frac{1}{\sqrt{2\pi r^2}} \exp\left[-\frac{(y-x)^2}{2r^2}\right] \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \bigg/ \left\{\frac{1}{\sqrt{2\pi(\sigma^2+r^2)}} \exp\left[-\frac{(y-\mu)^2}{2(\sigma^2+r^2)}\right]\right\} \\
&= \frac{1}{\sqrt{2\pi \cdot \frac{1}{\frac{1}{\sigma^2}+\frac{1}{r^2}}}} \exp\left[-\frac{(x-\frac{\frac{\mu}{\sigma^2}+\frac{y}{r^2}}{\frac{1}{\sigma^2}+\frac{1}{r^2}})^2}{2\cdot\frac{1}{\frac{1}{\sigma^2}+\frac{1}{r^2}}}\right] \\
&= N(\frac{\frac{\mu}{\sigma^2}+\frac{y}{r^2}}{\frac{1}{\sigma^2}+\frac{1}{r^2}}, \frac{1}{\frac{1}{\sigma^2}+\frac{1}{r^2}})
\end{aligned}$$

4 Work out the Galton's example. Assuming that Z, ϵ_1, ϵ_2 follow independent normal distributions, and that ϵ_1 and ϵ_2 have the same distribution, derive the joint distribution $[X, Y]$, and the conditional distribution $[Y|X]$. Check out ***Francis Galton's account of the invention of correlation*** by Stigler.

Let

$$X = Z + \epsilon_1 \qquad Y = Z + \epsilon_2$$

and

$$f_Z(z) = N(\mu, \sigma^2) \qquad f_{\epsilon_1}(e_1) = f_{\epsilon_2}(e_2) = N(e, r^2)$$

then

$$
\begin{aligned}
f_{X,Y}(x,y) &= \int_{-\infty}^{+\infty} f_{X,Y|Z}(x,y|z) f_Z(z) dz \\
&= \int_{-\infty}^{+\infty} f_{X|Z}(x|z) f_{Y|Z}(y|z) f_Z(z) dz \\
&= \int_{-\infty}^{+\infty} N_{X|Z}(e+z, r^2) N_{Y|Z}(e+z, r^2) N_Z(\mu, \sigma^2) dz \\
&= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{(2\pi)^3 r^4 \sigma^2}} \exp\left\{ -\frac{[x-(e+z)]^2 + [y-(e+z)]^2}{2r^2} - \frac{(z-\mu)^2}{2\sigma^2} \right\} dz
\end{aligned}
$$

here we can let $\alpha = x - e - \mu$, $\beta = y - e - \mu$ and $\gamma = z - \mu$, then

$$
\begin{aligned}
f_{X,Y}(x,y) &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{(2\pi)^3 r^4 \sigma^2}} \exp\left[-\frac{(\gamma-\alpha)^2 + (\gamma-\beta)^2}{2r^2} - \frac{\gamma^2}{2\sigma^2} \right] d\gamma \\
&= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi \cdot \frac{\sigma^2 r^2}{2\sigma^2 + r^2}}} \exp\left\{ -\frac{\left[\gamma - \frac{\sigma^2(\alpha+\beta)}{2\sigma^2+r^2}\right]^2}{2 \cdot \frac{\sigma^2 r^2}{2\sigma^2+r^2}} \right\} d\gamma \\
&\quad \cdot \frac{1}{\sqrt{(2\pi)^2 \begin{vmatrix} \sigma^2 + r^2 & \sigma^2 \\ \sigma^2 & \sigma^2 + r^2 \end{vmatrix}}} \exp\left[-\frac{1}{2} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}^t \begin{pmatrix} \sigma^2 + r^2 & \sigma^2 \\ \sigma^2 & \sigma^2 + r^2 \end{pmatrix}^{-1} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \right] \\
&= \frac{1}{\sqrt{(2\pi)^2 \begin{vmatrix} \sigma^2 + r^2 & \sigma^2 \\ \sigma^2 & \sigma^2 + r^2 \end{vmatrix}}} \exp\left[-\frac{1}{2} \begin{pmatrix} x-e-\mu \\ y-e-\mu \end{pmatrix}^t \begin{pmatrix} \sigma^2 + r^2 & \sigma^2 \\ \sigma^2 & \sigma^2 + r^2 \end{pmatrix}^{-1} \begin{pmatrix} x-e-\mu \\ y-e-\mu \end{pmatrix} \right] \\
&= N_2 \left[\begin{pmatrix} e+\mu \\ e+\mu \end{pmatrix}, \begin{pmatrix} \sigma^2 + r^2 & \sigma^2 \\ \sigma^2 & \sigma^2 + r^2 \end{pmatrix} \right]
\end{aligned}
$$

Similarly

$$
\begin{aligned}
f_{Y|X}(y|x) &= \int_{-\infty}^{+\infty} f_{Y|X,Z}(y|x,z) f_Z(z) dz \\
&= \int_{-\infty}^{+\infty} f_{Y|Z}(y|z) f_Z(z) dz \\
&= \int_{-\infty}^{+\infty} N_{Y|Z}(e+z, r^2) N_Z(\mu, \sigma^2) dz \\
&= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{(2\pi)^2 r^2 \sigma^2}} \exp\left\{-\frac{[y-(e+z)]^2}{2r^2} - \frac{(z-\mu)^2}{2\sigma^2}\right\} dz \\
&= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{(2\pi)^2 r^2 \sigma^2}} \exp\left[-\frac{(\gamma-\beta)^2}{2r^2} - \frac{\gamma^2}{2\sigma^2}\right] d\gamma \\
&= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi \cdot \frac{\sigma^2 r^2}{\sigma^2+r^2}}} \exp\left\{-\frac{\left[\gamma - \frac{\sigma^2\beta}{\sigma^2+r^2}\right]^2}{2\cdot\frac{\sigma^2 r^2}{\sigma^2+r^2}}\right\} d\gamma \\
&\quad \cdot \frac{1}{\sqrt{2\pi(\sigma^2+r^2)}} \exp\left[-\frac{\beta^2}{2(\sigma^2+r^2)}\right] \\
&= \frac{1}{\sqrt{2\pi(\sigma^2+r^2)}} \exp\left[-\frac{(y-e-\mu)^2}{2(\sigma^2+r^2)}\right] \\
&= N(e+\mu, \sigma^2+r^2)
\end{aligned}
$$

5 Work out the Markov example. Derive the t-step transition matrix of the Markov chain from the one step transition matrix. Check out the seminal paper on *Google pagerank algorithm* by Larry Page, Sergey Brin, R. Motwani, and T. Winograd.

Due to Markov property

$$
\begin{aligned}
p(x_t = i|x_0 = j) &= \sum_k p(x_t = i|x_{t-1} = k, x_0 = j) p(x_{t-1} = k|x_0 = j) \\
&= \sum_k p(x_t = i|x_{t-1} = k) p(x_{t-1} = k|x_0 = j)
\end{aligned}
$$

so the t-step transition matrix is

$$
K^{(t)} = K \cdot K^{(t-1)} = K \cdot K \cdot K^{(t-2)} = \cdots = K^t
$$