SDS 383D Sol 1

Preston Biro

November 2018

1 Preliminaries

A

Sol:

$$\begin{aligned}
p(w|X_1,\ldots,X_N) &\propto p(w)p(X_1,\ldots,X_N|w) \\
p(X_1,\ldots,X_N|w) &= \prod_{i=1}^{N} w^{X_i}(1-w)^{1-X_i} \\
p(w) &= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} w^{a-1}(1-w)^{b-1} \\
\implies p(w|X_1,\ldots,X_N) &\propto w^{a-1}(1-w)^{b-1}\prod_{i=1}^{N} w^{X_i}(1-w)^{1-X_i} \\
&= w^{a-1}(1-w)^{b-1} w^{\sum_{i=1}^{N} X_i}(1-w)^{n-\sum_{i=1}^{N} X_i} \\
&= w^{a+n\bar{X}-1}(1-w)^{b+n(1-\bar{X})-1} \\
\implies w|X_{1:N} &\sim Beta(a+n\bar{X}, b+n(1-\bar{X}))
\end{aligned}$$

■

B

Sol:

$$
\begin{aligned}
y_1 &= \frac{x_1}{x_1+x_2} \\
y_2 &= x_1 + x_2 \\
y_1 &= \frac{x_1}{y_2} \\
\implies x_1 &= y_1 y_2 \\
y_2 &= y_1 y_2 + x_2 \\
\implies x_2 &= y_2(1-y_1) \\
f_{Y_1,Y_2}(y_1,y_2) &= f_{X_1,X_2}(y_1y_2, y_2(1-y_1))|J| \\
|J| &= \frac{\partial x_1}{\partial y_1}\frac{\partial x_2}{\partial y_2} - \frac{\partial x_2}{\partial y_1}\frac{\partial x_1}{\partial y_2} \\
&= y_2(1-y_1) + y_2 y_1 = y_2 \\
\implies f_{Y_1,Y_2}(y_1,y_2) &= f_{X_1,X_2}(y_1y_2, y_2(1-y_1))|J| \\
&= f_{X_1}(y_1y_2) f_{X_2}(y_2(1-y_1)) y_2 \\
&= \frac{(y_1y_2)^{a_1-1}}{\Gamma(a_1)} e^{-y_1y_2} \frac{(y_2(1-y_1))^{a_2-1}}{\Gamma(a_2)} e^{-y_2(1-y_1)} y_2 \\
f_{Y_1,Y_2}(y_1,y_2) &= \frac{y_1^{a_1-1}(1-y_1)^{a_2-1} y_2^{a_1+a_2-1}}{\Gamma(a_1)\Gamma(a_2)} e^{-y_2} \\
p(y_1) &= \int f_{Y_1,Y_2}(y_1,y_2) dy_2 \\
&= \int \frac{y_1^{a_1-1}(1-y_1)^{a_2-1} y_2^{a_1+a_2-1}}{\Gamma(a_1)\Gamma(a_2)} e^{-y_2} dy_2 \\
&= \frac{y_1^{a_1-1}(1-y_1)^{a_2-1}\Gamma(a_1+a_2)}{\Gamma(a_1)\Gamma(a_2)} \int \frac{y_2^{a_1+a_2-1}}{\Gamma(a_1+a_2)} e^{-y_2} dy_2 \\
&= \frac{y_1^{a_1-1}(1-y_1)^{a_2-1}}{Beta(a_1,a_2)} \int Ga(a_1+a_2, 1) dy_2 \\
p(y_1) &= \frac{y_1^{a_1-1}(1-y_1)^{a_2-1}}{Beta(a_1,a_2)} \\
\implies Y_1 &\sim Beta(p = y_1, a_1, a_2) \\
p(y_2) &= \int f_{Y_1,Y_2}(y_1,y_2) dy_1 \\
&= \int \frac{y_1^{a_1-1}(1-y_1)^{a_2-1} y_2^{a_1+a_2-1}}{\Gamma(a_1)\Gamma(a_2)} e^{-y_2} dy_1 \\
&= \frac{y_2^{a_1+a_2-1} e^{-y_1}}{\Gamma(a_1+a_2-1)} \int \frac{y_1^{a_1-1}(1-y_1)^{a_2-1}}{Beta(a_1,a_2)} dy_1 \\
p(y_2) &= \frac{y^{a_1+a_2-1} e^{-y_2}}{\Gamma(a_1+a_2-1)} \\
\implies Y_2 &\sim Ga(a_1+a_2, 1)
\end{aligned}
$$

■

C

Sol:

$$
\begin{aligned}
\theta &\sim N(m, v) \\
x_i &\sim N(\theta, \sigma^2) \\
p(\theta|x_1, \ldots, x_N) &\propto p(\theta)p(x_1, \ldots, x_N|\theta) \\
&\propto e^{\frac{-(\theta-m)^2}{2v}} \prod_{i=1}^{N} e^{\frac{-(x_i-\theta)^2}{2\sigma^2}} \\
&= e^{-\frac{(\theta-m)^2}{2v} - \sum_{i=1}^{N} \frac{(x_i-\theta)^2}{2\sigma^2}} \\
&= e^{-\frac{1}{2}(\frac{(\theta-m)^2}{v} + n\frac{(\theta-\bar{x})^2}{\sigma^2} + \sum_{i=1}^{N} \frac{(x_i-\bar{x})^2}{\sigma^2})} \\
&\propto e^{-\frac{1}{2}(\frac{\theta^2-2m\theta}{v} + \frac{n\theta^2-2n\theta\bar{x}}{\sigma^2})} \\
p(\theta|x_1, \ldots, x_N) &\propto e^{-\frac{(\frac{1}{v}+\frac{n}{\sigma^2})}{2}(\theta-(\frac{m}{v}+\frac{n\bar{x}}{\sigma^2})(\frac{1}{v}+\frac{n}{\sigma^2})^{-1})^2} \\
\theta|x_{1:N} &\sim N((\frac{m}{v} + \frac{n\bar{x}}{\sigma^2})(\frac{1}{v} + \frac{n}{\sigma^2})^{-1}, (\frac{1}{v} + \frac{n}{\sigma^2})^{-1})
\end{aligned}
$$

■

D

Sol:

$$
\begin{aligned}
w &\sim Ga(a, b) \\
p(w|x_1, \ldots, x_N) &\propto p(w)p(x_1, \ldots, x_N|w) \\
&\propto w^{a-1}e^{-bw} \prod_{i=1}^{N} (w)^{1/2} e^{-\frac{w}{2}(x_i-\theta)^2} \\
&= w^{a-1}e^{-bw}w^{N/2}e^{-\frac{w}{2}\sum(x_i-\theta)^2} \\
p(w|x_1, \ldots, x_N) &\propto w^{\frac{N}{2}+a-1}e^{-w(b+\frac{\sum(x_i-\theta)^2}{2})} \\
w|x_{1:N} &\sim Ga(\frac{N}{2} + a, b + \frac{\sum(x_i-\theta)^2}{2}) \\
p(\sigma^2|x_1, \ldots, x_N) &\propto (\frac{1}{\sigma^2})^{\frac{N}{2}+a-1}e^{-(\frac{1}{\sigma^2})(b+\frac{\sum(x_i-\theta)^2}{2})} \\
\sigma^2|x_{1:N} &\sim IG(\frac{N}{2} + a, b + \frac{\sum(x_i-\theta)^2}{2})
\end{aligned}
$$

■

E

Sol:

$$
\begin{aligned}
x_i &\sim N(\theta, \sigma_i^2) \\
\theta &\sim N(m, v) \\
p(\theta | x_1, \dots, x_N) &\propto p(\theta) p(x_{1:N} | \theta) \\
&= \frac{1}{\sqrt{2\pi v}} e^{-\frac{(\theta - m)^2}{2v}} \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{(x_i - \theta)^2}{2\sigma_i^2}} \\
&= \frac{1}{\sqrt{2\pi v}} \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{(\theta - m)^2}{2v} - \sum_{i=1}^{N} \frac{(x_i - \theta)^2}{2\sigma_i^2}} \\
&\propto e^{-\frac{1}{2}(\frac{\theta^2 - 2\theta m + m^2}{v} + \sum \frac{\theta^2 - 2\theta x_i + x_i^2}{\sigma_i^2})} \\
&\propto e^{-\frac{1}{2}\theta^2(\frac{1}{v} + \sum \frac{1}{\sigma_i^2}) - 2\theta(\frac{m}{v} + \sum \frac{x_i}{\sigma_i^2})} \\
&\propto e^{-\frac{1}{2(\frac{1}{v} + \sum \frac{1}{\sigma_i^2})^{-1}}(\theta - (\frac{m}{v} + \sum \frac{x_i}{\sigma_i^2})(\frac{1}{v} + \sum \frac{1}{\sigma_i^2})^{-1})^2} \\
\theta | x &\sim N((\frac{m}{v} + \sum \frac{x_i}{\sigma_i^2})(\frac{1}{v} + \sum \frac{1}{\sigma_i^2})^{-1}, (\frac{1}{v} + \sum \frac{1}{\sigma_i^2})^{-1})
\end{aligned}
$$

■

F

Sol:

$$
\begin{aligned}
p(x,w) &= p(x|w)p(w) \\
&= \frac{w^{1/2}}{\sqrt{2\pi}} e^{-\frac{(x-m)^2 w}{2}} \frac{(b/2)^{a/2}}{\Gamma(a/2} w^{a/2-1} e^{-bw/2} \\
&= \frac{(b/2)^{a/2}}{\sqrt{2\pi}\Gamma(a/2)} w^{\frac{a+1}{2}-1} e^{\frac{w}{2}(b+(x-m)^2)} \\
p(x) &= \int p(x,w)dw \\
&= \int_0^\infty \frac{(b/2)^{a/2}}{\sqrt{2\pi}\Gamma(a/2)} w^{\frac{a+1}{2}-1} e^{\frac{w}{2}(b+(x-m)^2)} dw \\
&= \frac{(b/2)^{a/2}}{\sqrt{2\pi}\Gamma(a/2)} \int_0^\infty w^{\frac{a+1}{2}-1} e^{\frac{w}{2}(b+(x-m)^2)} dw \\
&= \frac{(b/2)^{a/2}}{\sqrt{2\pi}\Gamma(a/2)} \frac{\Gamma(\frac{a+1}{2})}{(\frac{b+(x-m)^2}{2})^{\frac{a+1}{2}}} \int_0^\infty \frac{(\frac{b+(x-m)^2}{2})^{\frac{a+1}{2}}}{\Gamma(\frac{a+1}{2})} w^{\frac{a+1}{2}-1} e^{\frac{w}{2}(b+(x-m)^2)} dw \\
&= \frac{(b/2)^{a/2}}{\sqrt{2\pi}\Gamma(a/2)} \frac{\Gamma(\frac{a+1}{2})}{(\frac{b+(x-m)^2}{2})^{\frac{a+1}{2}}} \int_0^\infty Ga(\frac{a+1}{2}, \frac{b+(x-m)^2}{2}) dw \\
&= \frac{(b/2)^{a/2}}{\sqrt{2\pi}\Gamma(a/2)} \frac{\Gamma(\frac{a+1}{2})}{(\frac{b+(x-m)^2}{2})^{\frac{a+1}{2}}} \\
&= \frac{(\frac{b}{2})^{\frac{a}{2}} \Gamma(\frac{a+1}{2})}{\sqrt{2\pi}\Gamma(\frac{a}{2})} (\frac{b+(x-m)^2}{2})^{-\frac{a+1}{2}} \\
&= \frac{(\frac{b}{2})^{\frac{a}{2}} \Gamma(\frac{a+1}{2})}{\sqrt{2\pi}\Gamma(\frac{a}{2})} (1+\frac{(x-m)^2}{b})^{-\frac{a+1}{2}} (\frac{b}{2})^{-\frac{a+1}{2}} \\
&= \frac{\sqrt{a}}{\sqrt{a}} \frac{\Gamma(\frac{a+1}{2})}{\sqrt{b\pi}\Gamma(\frac{a}{2})} (\frac{a+\frac{a(x-m)^2}{b}}{a})^{-\frac{a+1}{2}} \\
&= \frac{\Gamma(\frac{a+1}{2})}{\sqrt{a\pi}\sqrt{\frac{b}{a}}\Gamma(\frac{a}{2})} (\frac{a+\frac{(x-m)^2}{\sqrt{\frac{b}{a}}}}{a})^{-\frac{a+1}{2}}
\end{aligned}
$$

Which is by definition, a Student t distribution with scale $\sqrt{\frac{b}{a}}$, mean m, and d = a degrees of freedom. ■

2 Multivariate Normal Distribution

A

Sol: Part a)

$$\begin{aligned}
Cov(x) &= E[(x-\mu)(x-\mu)^T] \\
&= E[(x-\mu)(x^T-\mu^T)] \\
&= E[xx^T - \mu x^T - x\mu^T + \mu\mu^T] \\
&= E[xx^T] - \mu E[x^T] - E[x]\mu^T + \mu\mu^T \\
&= E[xx^T] - \mu\mu^T - \mu\mu^T + \mu\mu^T \\
&= E[xx^T] - \mu\mu^T
\end{aligned}$$

Part b) Using $E[Ax+b] = AE[x]+b = A\mu + b$:

$$\begin{aligned}
Cov(Ax+b) &= E[(Ax+b)(Ax+b)^T] - E[Ax+b]E[Ax+b]^T \\
&= E[Axx^TA^T] + E[bx^TA^T] + E[Axb^T] + E[bb^T] - (A\mu+b)(A\mu+b)^T \\
&= AE[xx^T]A^T + bE[x^T]A^T + AE[x]b^T + bb^T - A\mu\mu^TA^T - b\mu^TA^T - A\mu b^T - bb^T \\
&= AE[xx^T]A^T + b\mu^TA^T + A\mu b^T + bb^T - A\mu\mu^TA^T - b\mu^TA^T - A\mu b^T - bb^T \\
&= AE[xx^T]A^T - A\mu\mu^TA^T \\
&= A(E[xx^T] - \mu\mu^T)A^T = ACov(x)A^T
\end{aligned}$$

■

B

We know the pdf and moment generating function of an independent standard normal distribution are $f_{Z_i}(z_i) = \frac{1}{\sqrt{2\pi}}e^{z_i^2/2}$ and $M_{z_i}(t) = e^{t^2/2}$ respectively. Thus since each of the z_i's are independent, we have:

$$\begin{aligned}
f_Z(z_1,\dots,z_p) &= f_{Z_1}(z_1)f_{Z_2}(z_2)\dots f_{Z_p}(z_p) \\
&= \frac{1}{\sqrt{2\pi}}e^{z_1^2/2}\frac{1}{\sqrt{2\pi}}e^{z_2^2/2}\dots\frac{1}{\sqrt{2\pi}}e^{z_p^2/2} \\
&= (\frac{1}{\sqrt{2\pi}})^p e^{\frac{1}{2}\sum_{i=1}^n z_i^2} \\
&= (\frac{1}{\sqrt{2\pi}})^p e^{\frac{z^Tz}{2}}
\end{aligned}$$

Also:

$$
\begin{aligned}
M_Z(t) &= \prod_{i=1}^{p} E[e^{tz_i}] \\
&= \prod_{i=1}^{p} e^{t^2/2} \\
&= e^{t^T t/2}
\end{aligned}
$$

C

Sol: If x is multivariate normal, then we can assume that for all not identically zero vectors a, $z = a^T x$ and the univariate normal moment generating function $M_z(t) = E[e^{tz}] = e^{\mu t + \sigma^2 t^2/2}$, where μ is the mean and σ^2 is the variance of z. Thus:

$$
\begin{aligned}
E[z] = E[a^T x] &= a^T E[x] = a^T \mu \\
Var[z] = Var[a^T x] &= a^T Var[x] a = a^T \Sigma a \\
M_x(t) &= E[e^{t^T a^T x}] \\
&= E[e^{\sum_{i=1}^{p} t a_i x_i}] \\
&= \prod_{i=1}^{p} E[e^{t a_i x_i}]
\end{aligned}
$$

Since the vector a is not identically zero and t can take any value, write $\mathbf{t} = ta$. Thus:

$$
\begin{aligned}
M_x(t) &= \prod_{i=1}^{p} E[e^{t a_i x_i}] \\
&= \prod_{i=1}^{p} E[e^{\mathbf{t} x_i}] \\
&= \prod_{i=1}^{p} e^{\mathbf{t}\mu_i + \mathbf{t}^2 \sigma_i^2/2} \\
&= e^{\mathbf{t}^T \mu + \mathbf{t}^T \Sigma \mathbf{t}/2}
\end{aligned}
$$

And since we know that moment generating function are unique, it follows that if a random variable has this moment generating function, then the random variable follows a multivariate normal distribution. ■

D

Sol:

$$\begin{aligned} M_x(t) &= E[e^{t^T x}] \\ &= E[e^{t^T(Lz+\mu)}] \\ &= E[e^{t^T\mu}e^{t^T Lz}] \\ &= e^{t^T\mu}E[e^{t^T Lz}] \\ &= e^{t^T\mu}e^{t^T LL^T t/2} \\ &= e^{t^T\mu + t^T\Sigma t/2} \end{aligned}$$

Thus the mean vector is μ and covariance matrix is $\Sigma = LL^T$. ■

E

Sol: Assume x is multivariate normal. Thus the moment generating function of x must be:

$$M_x(t) = e^{t^T\mu + t^T\Sigma t/2}$$

■