1. Let $X \sim \text{Exp}(\lambda)$, where $\text{E}(X) = 1/\lambda$. What is the pmf (probability mass function) of $Y = \lfloor X \rfloor$ (the floor of X)? Do you recognize it as a distribution that you have studied in the past?

$Y \in \{0, 1, 2 \cdots\}$ $\quad f_X(x) = \lambda e^{-\lambda x}, \; x > 0$

$$P(Y=k) = P(X \in [k, k+1))$$
$$= \int_k^{k+1} \lambda e^{-\lambda x} \, dx$$
$$= -e^{-\lambda x} \Big|_{x=k}^{x=k+1}$$
$$= e^{-\lambda k} - e^{-(k+1)\lambda}$$
$$= e^{-\lambda k}(1 - e^{-\lambda})$$
$$= (e^{-\lambda})^k (1 - e^{-\lambda})$$

It's a geometric distribution with success rate $1 - e^{-\lambda}$.

2. Let X_1 and X_2 be two independent random variables such that $X_i \sim \mathrm{Ga}(a_i, b)$ for any $a_1, a_2, b > 0$. Define $Y = X_1/(X_1 + X_2)$ and $Z = (X_1 + X_2)$.

(a) Find the joint pdf for Y and Z and show that these two random variables are independent.

(b) Find the marginal pdf of Z. Do you recognize this pdf as belonging to some family that you know?

(c) Find the marginal pdf of Y. Do you recognize this pdf as belonging to some family that you know?

(d) Compute $\mathrm{E}(Y^k)$ for any $k > 0$.

(e) What does this result imply if $a_i = b = 1$?

a)

$f_{X_1,X_2}(X_1,X_2) = f_{X_1}(X_1) \cdot f_{X_2}(X_2)$ by independence.

$$= \frac{b^{a_1}}{T(a_1)} x_1^{a_1-1} e^{-bx_1} \cdot \frac{b^{a_2}}{T(a_2)} x_2^{a_2-1} e^{-bx_2}$$

$Y = \frac{x_1}{x_1+X_2}$, $Z = X_1 + X_2 \Rightarrow X_1 = Y \cdot Z$, $X_2 = Z - Y \cdot Z$

$$|J| = \begin{vmatrix} \partial X_1/\partial Y & \partial X_1/\partial Z \\ \partial X_2/\partial Y & \partial X_2/\partial Z \end{vmatrix} = \begin{vmatrix} Z & Y \\ -Z & 1-Y \end{vmatrix} = Z \qquad \begin{cases} Y \in (0,1) \\ Z \in (0,\infty) \end{cases}$$

$$f_{Y,Z}(Y,Z) = \frac{b^{a_1+a_2}}{T(a_1)T(a_2)} (yz)^{a_1-1} [z(1-y)]^{a_2-1} e^{-bz} \cdot Z$$

$$= \frac{b^{a_1+a_2}}{T(a_1)T(a_2)} y^{a_1-1} (1-y)^{a_2-1} e^{-bz} \cdot z^{a_1+a_2-1}$$

$$= \underbrace{\frac{1}{B(a_1,a_2)} \cdot y^{a_1-1} (1-y)^{a_2-1}}_{Beta(a_1,\, a_2)} \cdot \underbrace{\frac{b^{a_1+a_2}}{T(\alpha)} z^{a_1+a_2-1} \cdot e^{-bz}}_{Gamma(a_1+a_2)}$$

We can seporate $f_{Y,Z}(Y,Z) = g(Y) \cdot h(z) \Rightarrow Y \perp Z$

2. Let X_1 and X_2 be two independent random variables such that $X_i \sim \text{Ga}(a_i, b)$ for any $a_1, a_2, b > 0$. Define $Y = X_1/(X_1 + X_2)$ and $Z = (X_1 + X_2)$.

(a) Find the joint pdf for Y and Z and show that these two random variables are independent.

(b) Find the marginal pdf of Z. Do you recognize this pdf as belonging to some family that you know?

(c) Find the marginal pdf of Y. Do you recognize this pdf as belonging to some family that you know?

(d) Compute $\text{E}(Y^k)$ for any $k > 0$.

(e) What does this result imply if $a_i = b = 1$?

b). $f_Z(z) = \dfrac{b^{a_1+a_2}}{\Gamma(a_1+a_2)} z^{a_1+a_2-1} \cdot e^{-bz} \quad , z > 0$

$Z \sim \text{Gamma}(a_1+a_2, b)$

c). $f_Y(y) = \dfrac{1}{B(a_1,a_2)} y^{a_1-1} (1-y)^{a_2-1} \quad , y \in (0,1)$

$Y \sim \text{Beta}(a_1, a_2)$

d). $E(Y^k) = \int_0^1 y^k \cdot \dfrac{1}{B(a_1,a_2)} \cdot y^{a_1-1} (1-y)^{a_2-1} \, dy$

$= \dfrac{1}{B(a_1,a_2)} \cdot \underbrace{\int_0^1 y^{a_1+k-1} (1-y)^{a_2-1} \, dy}_{\text{it's a beta kernel, it integrates to } B(a_1+k,a_2)}$

$= \dfrac{B(a_1+k,a_2)}{B(a_1,a_2)} = \dfrac{\frac{\Gamma(a_1+k)\Gamma(a_2)}{\Gamma(a_1+a_2+k)}}{\frac{\Gamma(a_1)\Gamma(a_2)}{\Gamma(a_1+a_2)}} = \dfrac{\Gamma(a_1+k)}{\Gamma(a_1)} \cdot \dfrac{\Gamma(a_1+a_2)}{\Gamma(a_1+a_2+k)}$

$= \dfrac{(a_1+k-1)!}{(a_1-1)!} \cdot \dfrac{(a_1+a_2-1)!}{(a_1+a_2+k-1)!}$

e) $E(Y^k) = \dfrac{k!}{(k+1)!}$

$= 1/k+1$

3. Consider three independent random variables X_1, X_2 and X_3 such that $X_i \overset{indep}{\sim}$ Gamma(a_i, b), $i = 1, 2, 3$. Let

$$\boldsymbol{Y} = (Y_1, Y_2, Y_3) = \left(\frac{X_1}{X_1+X_2+X_3}, \frac{X_2}{X_1+X_2+X_3}, \frac{X_3}{X_1+X_2+X_3}\right).$$

(a) Show that $\boldsymbol{Y} \sim$ Dirichlet(a_1, a_2, a_3), a Dirichlet distribution.

$$f_{X_1,X_2,X_3}(x_1,x_2,x_3) = f_{X_1}(x_1)\cdot f_{X_2}(x_2)\cdot f_{X_3}(x_3) \quad \text{by independence}$$

$$= \frac{b^{a_1}}{\Gamma(a_1)}\cdot x_1^{a_1-1} e^{-bx_1} \cdot \frac{b^{a_2}}{\Gamma(a_2)}\cdot x_2^{a_2-1} e^{-bx_2} \cdot \frac{b^{a_3}}{\Gamma(a_3)} x_3^{a_3-1} e^{-bx_3}$$

Denote $Z = x_1+x_2+x_3$, $\tilde{Y} = (Y_1, Y_2, Z)$

$$Y_1 = \frac{x_1}{x_1+x_2+x_3}, \quad Y_2 = \frac{x_2}{x_1+x_2+x_3}, \quad Z = x_1+x_2+x_3, \quad \begin{cases} Z>0 \\ Y_1+Y_2 \in (0,1) \\ Y_1, Y_2 \in (0,1 \end{cases}$$

$$\Rightarrow x_1 = Y_1 Z, \quad x_2 = Y_2 Z, \quad x_3 = Z - Y_1 Z - Y_2 Z = Z(1-Y_1-Y_2)$$

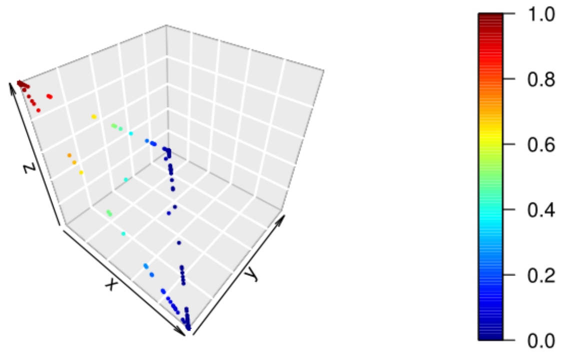
$$|J| = \begin{vmatrix} \partial x_1/\partial Y_1 & \partial x_1/\partial Y_2 & \partial x_1/\partial Z \\ \partial x_2/\partial Y_1 & \partial x_2/\partial Y_2 & \partial x_2/\partial Z \\ \partial x_3/\partial Y_1 & \partial x_3/\partial Y_2 & \partial x_3/\partial Z \end{vmatrix} = \begin{vmatrix} Z & 0 & Y_1 \\ 0 & Z & Y_2 \\ -Z & -Z & 1-Y_1-Y_2 \end{vmatrix} = Z^2$$

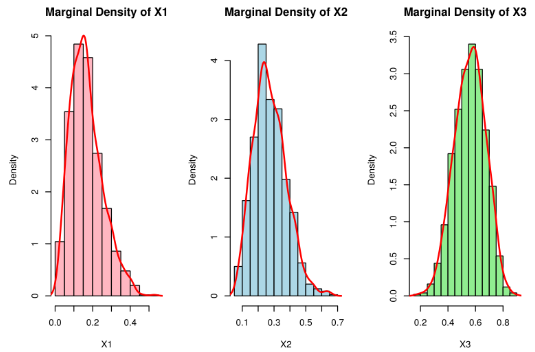
$$f_{Y_1,Y_2,Z}(y_1,y_2,z) = \frac{b^{a_1+a_2+a_3}}{\Gamma(a_1)\Gamma(a_2)\Gamma(a_3)}\cdot (y_1 z)^{a_1-1}(y_2 z)^{a_2-1}[(1-y_1-y_2)z]^{a_3-1} e^{-bz}\cdot z^2$$

$$= \underbrace{\frac{\Gamma(a_1+a_2+a_3)}{\Gamma(a_1)\Gamma(a_2)\Gamma(a_3)}\cdot y_1^{a_1-1} y_2^{a_2-1}(1-y_1-y_2)^{a_3-1}}_{\text{Dirichlet}(a_1,a_2,a_3)} \cdot \underbrace{z^{a_1+a_2+a_3-1} e^{-bz}\cdot\frac{b^{a_1+a_2+a_3}}{\Gamma(a_1+a_2+a_3)}}_{\text{Gamma}(a_1+a_2+a_3,\, b)}$$

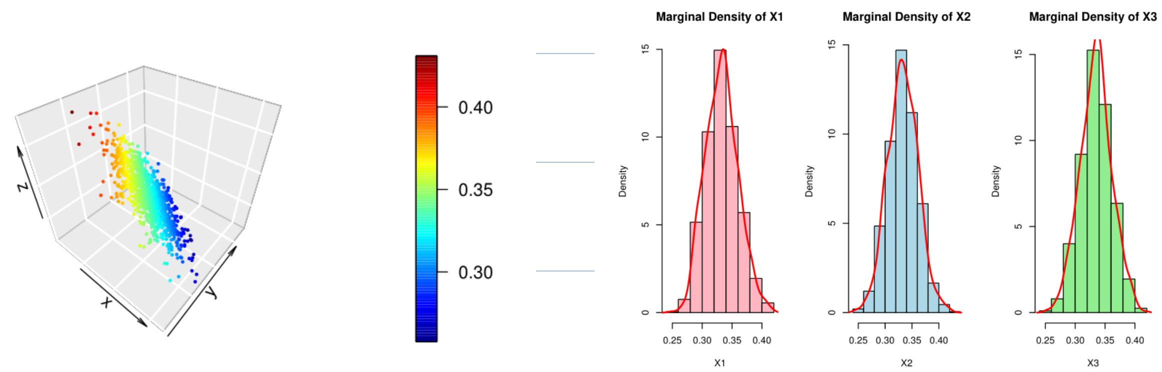
```r
rdirich <- function(n, a){
  X <- matrix(NA, nrow = n, ncol = length(a))
  for (i in 1:length(a)) {
    X[,i] <- rgamma(n, shape = a[i], rate = 1)
  }
  D <- X/rowSums(X)
  par(mfrow = c(1,1))
  scatter3D(x = D[,1], y = D[,2], z = D[,3], pch = 19, cex = .3, bty = "g")
  par(mfrow = c(1,3))
  hist(D[,1], main = "Marginal Density of X1", prob = TRUE, xlab = "X1", col = "lightpink")
  lines(density(D[,1]), lwd = 2, col = "red")
  hist(D[,2], main = "Marginal Density of X2", prob = TRUE, xlab = "X2", col = "lightblue")
  lines(density(D[,2]), lwd = 2, col = "red")
  hist(D[,3], main = "Marginal Density of X3", prob = TRUE, xlab = "X3", col = "lightgreen")
  lines(density(D[,3]), lwd = 2, col = "red")
}
```

```r
rdirich(1000, c(0.01, 0.01, 0.01))
```

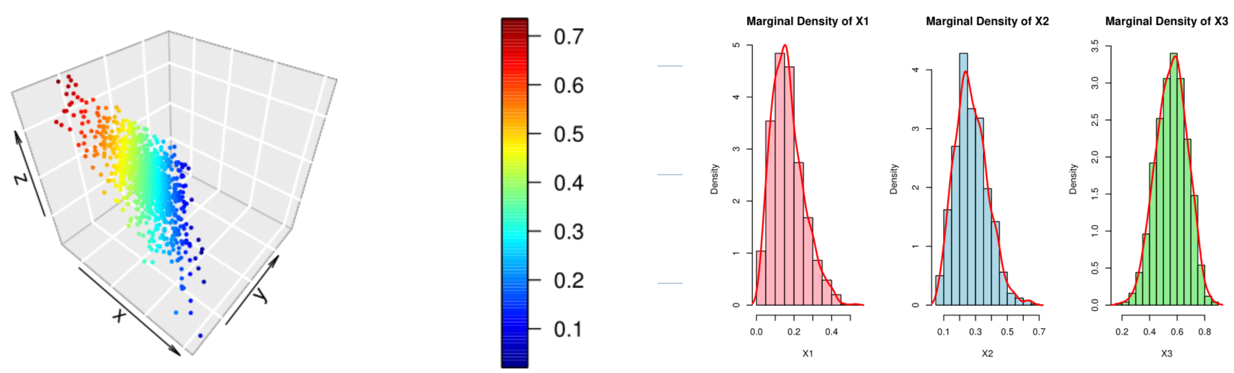




```
rdirich(1000, c(100, 100, 100))
```




```
rdirich(1000, c(3,5,10))
```



First, all points are exactly on this surface: $z = 1-x-y$.

for $a=(0.01, 0.01, 0.01)$, points are more likely to appear on x-y, y-z, or x-z plane, which means at least 1 of x_1, x_2, x_3 is likely to be 0. when $a=(100,100,100)$, points seem more concentrated in the middle. But for $a=(3,5,10)$, it is more dispersed.

4. Y follows an inverse Gamma distribution with shape parameter a and scale parameter b ($Y \sim \mathrm{IG}(a,b)$) if $Y = 1/X$ with $X \sim \mathrm{Gamma}(a,b)$ (assume the Gamma distribution is parameterized such that $\mathrm{E}(X) = ab$).

   (a) Find the density of Y.

   (b) Compute $\mathrm{E}(Y^k)$. Do you need to impose any constrain on the problem for this expectation to exists?

   (c) Compare $\mathrm{E}(Y^k)$ to $1/\mathrm{E}(X^k)$ (hint: look at the ratio of the two quantities).

(a) $f_X(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta}$ $Y = 1/X \Rightarrow X = 1/Y$ $\frac{dx}{dY} = -\frac{1}{y^2}$

$$f_Y(y) = f_X(1/Y) \cdot \left|\frac{-1}{y^2}\right| = \frac{1}{\Gamma(\alpha)\beta^\alpha} \cdot y^{1-\alpha} \cdot e^{-\frac{1}{\beta y}} \cdot \frac{1}{y^2}, \quad y > 0$$

$$= \frac{1}{\Gamma(\alpha)\beta^\alpha} y^{-(\alpha+1)} e^{-\frac{1}{\beta y}}. \quad y > 0$$

(b) $$E(Y^k) = \int_0^\infty \frac{1}{\Gamma(\alpha)\beta^\alpha} y^{-(\alpha+1)+k} e^{-\frac{1}{\beta y}} dy$$

$$= \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty \underbrace{y^{-(\alpha-k+1)} e^{-\frac{1}{\beta y}}}_{\text{kernel of } IG(\alpha-k,\, b)} dy$$

kernel of $IG(\alpha-k, b)$, so it integrates to : $\Gamma(\alpha-k)\beta^{\alpha-k}$

$$= \frac{\Gamma(\alpha-k)}{\Gamma(\alpha)} \cdot \beta^{-k}$$

k should be smaller than α. [α is a here]

(c) $$E(X^k) = \int_0^\infty \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta} \cdot x^k dx$$

$$= \frac{1}{\Gamma(\alpha)\beta^\alpha} \underbrace{\int_0^\infty x^{\alpha+k-1} e^{-\frac{x}{\beta}} dx}_{\text{kernel of Gamma}(\alpha+k,\, \beta)}$$

$$= \frac{\Gamma(\alpha+k)\beta^{\alpha+k}}{\Gamma(\alpha)\beta^\alpha} = \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)} \cdot \beta^k$$

$$\Rightarrow \quad R = \frac{E(Y^k)}{1/E(X^k)} = \frac{\frac{\Gamma(\alpha-k)}{\Gamma(\alpha)} \cdot \beta^{-k}}{\frac{\Gamma(\alpha)}{\Gamma(\alpha+k)} \cdot \beta^{-k}} = \frac{\Gamma(\alpha-k)\Gamma(\alpha+k)}{\Gamma(\alpha)\cdot\Gamma(\alpha)}$$

$$= \frac{\Gamma(\alpha-k)}{\Gamma(\alpha-k)(\alpha-k)(\alpha-k+1)\cdots(\alpha-1)} \cdot \frac{\Gamma(\alpha)\cdot(\alpha+1)(\alpha+2)\cdots(\alpha+k-1)}{\Gamma(\alpha)}$$

$$= \frac{\alpha(\alpha+1)\cdots(\alpha+k-1)}{(\alpha-k)(\alpha-k+1)\cdots(\alpha-1)} = \prod_{i=0}^{k-1}\left(\frac{\alpha+i}{\alpha-k+i}\right)$$ for each i:

$$\alpha+i > \alpha-k+1, \quad \Rightarrow \frac{\alpha+i}{\alpha-k+i} > 1$$

$$\Rightarrow \prod_{i=0}^{k-1}\left(\frac{\alpha+i}{\alpha-k+i}\right) > 1 \qquad \Rightarrow E(Y^k) > 1/E(X^k)$$

PS: By Jensen's Ineq. $g(x) = \frac{1}{x^k}$ is convex function

$\Rightarrow \quad E(1/x^k) \geq 1/[E(x)]^k$, is there $[E(x)]^k < E(x^k)$?

We know $E(x^2) = [E(x)]^2 + Var(x)$, meaning when $k=2$, it's true.

What about other k's ? ☆

5. Y follows a log normal distribution with parameters μ and σ^2 (denotes as $Y \sim$ Log-N(μ, σ^2) if $Y = \exp(X)$ where $X \sim \mathrm{N}(\mu, \sigma^2)$).

(a) Find the density of Y.

$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad y = e^x \Rightarrow x = \log y,\ y>0$

$\frac{dx}{dy} = \frac{1}{y} \quad \Rightarrow f_Y(y) = f_X(\log(y)) \cdot \left|\frac{1}{y}\right| = \frac{1}{y\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{[\log(y)-\mu]^2}{2\sigma^2}\right)$

any $y>0$.

(b) Compute the mean and the variance of Y.

$E(Y) = \int_0^\infty y \cdot \frac{1}{\sqrt{2\pi\sigma^2}\, y} \cdot \exp\left(-\frac{(\log(y)-\mu)^2}{2\sigma^2}\right) dy$

Let $t = \log(y)$, $y = e^t$

$\Rightarrow \int_0^\infty \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{(\log(y)-\mu)^2}{2\sigma^2}\right) dy$

$= \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right) \cdot e^t\, dt$

$= \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{t^2-2\mu t-2\sigma^2 t^2+\mu^2}{2\sigma^2}\right) dt$

$= \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(t-(\mu+\sigma^2))^2}{2\sigma^2}\right) \cdot \exp\left(\mu+\frac{\sigma^2}{2}\right) dt$

$= \exp\left(\mu + \sigma^2/2\right)$

$$E(y^2) = \int_0^\infty y^2 \cdot \frac{1}{y\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\log(y)-\mu)^2}{2\sigma^2}\right) dy$$

Let $t = \log(y)$, $y = e^t$, $t \in (-\infty, \infty)$

$$\Rightarrow E(y^2) = \int_0^\infty \frac{1}{\sqrt{2\pi\sigma^2}} e^t \cdot \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right) e^t \, dt$$

$$= \int_0^\infty \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{t^2-2\mu t-4\sigma^2 t+\mu^2}{2\sigma^2}\right) dt$$

$$= \int_0^\infty \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{[t-(\mu+2\sigma^2)]^2}{2\sigma^2}\right) \cdot \exp(2\sigma^2+2\mu)\, dt$$

$$= \exp(2\sigma^2+2\mu)$$

$$Var(y) = E(y^2) - [E(y)]^2 = \exp(2\sigma^2+2\mu) - \exp(2\mu+\sigma^2)$$

$$= \exp(2\mu+\sigma^2)(\exp(\sigma^2)-1)$$

6. Let $\boldsymbol{X} = (X_1, X_2, \ldots, X_p)$ with $X \sim \mathrm{N}_p(\boldsymbol{\mu}, \Sigma)$ and set $\boldsymbol{Z}_1 = (X_1, \ldots, X_q)$ and $\boldsymbol{Z}_2 = (X_{q+1}, \ldots,, X_p)$ with $1 < q < p$. Show that $\Sigma_{11}: q \times q$ $\Sigma_{12}: q \times p-q$ Σ_{21} $p-q \times q$

$$\boldsymbol{Z}_1 \mid \boldsymbol{Z}_2 \sim \mathrm{N}_q\left(\boldsymbol{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\boldsymbol{Z}_2 - \boldsymbol{\mu}_2), \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right),$$ $\Sigma_{22}: p-q \times p-q$

where $\boldsymbol{\mu}_k$ and $\Sigma_{k\ell}$ denote the blocks of $\boldsymbol{\mu}$ and Σ where the rows correspond to the variables in $\boldsymbol{Z}_k$ and the columns to the variables in $\boldsymbol{Z}_\ell$.

Consider $A = \begin{pmatrix} I_{q\times q} & -\Sigma_{12}\Sigma_{22}^{-1} \\ 0 & I_{(p-q)\times(p-q)} \end{pmatrix}$ $\quad A \cdot \vec{X} = A\begin{pmatrix} \vec{Z}_1 \\ \vec{Z}_2 \end{pmatrix} = \begin{pmatrix} \vec{Z}_1 - \Sigma_{12}\Sigma_{22}^{-1}\vec{Z}_2 \\ \vec{Z}_2 \end{pmatrix}$

By properties of MVN:

$A\vec{X} \sim MVN(A\vec{\mu}, A\Sigma A^T)$, i.e.

$$MVN\left[\begin{pmatrix} \vec{\mu}_1 - \Sigma_{12}\Sigma_{22}^{-1}\vec{\mu}_2 \\ \vec{\mu}_2 \end{pmatrix}, \begin{pmatrix} I & -\Sigma_{12}\Sigma_{22}^{-1} \\ 0 & I \end{pmatrix}\begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}\begin{pmatrix} I & 0 \\ -\Sigma_{22}^{-1}\Sigma_{21} & I \end{pmatrix}\right]$$

$$\Rightarrow Var(A\vec{X}) = \begin{pmatrix} \Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} & 0 \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \begin{pmatrix} I & 0 \\ -\Sigma_{22}^{-1}\Sigma_{21} & I \end{pmatrix}$$

$$= \begin{pmatrix} \Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} & \underline{\underline{0}} \\ \underline{\underline{0}} & \Sigma_{22} \end{pmatrix}$$

So, $Cov(\vec{Z}_1 - \Sigma_{12}\Sigma_{22}^{-1}\vec{Z}_2 \,,\, \vec{Z}_2) = 0$. indicating they are independent since here is $\underline{\underline{MVN}}$. properties of MVN

$\Rightarrow \vec{Z}_1 - \Sigma_{12}\Sigma_{22}^{-1}\vec{Z}_2 \mid \vec{Z}_2$ has the same distribution as

$\vec{Z}_1 - \Sigma_{12}\Sigma_{22}^{-1}\vec{Z}_2$ itself even if $\vec{Z}_2$ is not given.

So: $\vec{Z}_1 - \Sigma_{12}\Sigma_{22}^{-1}\vec{Z}_2 \mid \vec{Z}_2 \sim MVN(\vec{\mu}_1 - \Sigma_{12}\Sigma_{22}^{-1}\vec{\mu}_2 \,,\, \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})$

and: $\Sigma_{12}\Sigma_{22}^{-1}\vec{Z}_2 \mid \vec{Z}_2$ is constant vector, therefore:

$$\vec{Z}_1 \mid Z_2 \sim MVN\left(\vec{\mu}_1 + \Sigma_{11}\Sigma_{22}^{-1}(\vec{Z}_2 - \vec{\mu}_2),\ \Sigma_{11} - \Sigma_{12}\Sigma_{21}^{-1}\Sigma_{21}\right)$$

7. Show that if $X \sim \text{Exp}(\beta)$, then

(a) $Y = X^{1/\gamma}$ has a Weibull distribution with parameters γ and β with $\gamma > 0$ a constant.

(b) $Y = (2X/\beta)^{1/2}$ has the Rayleigh distribution.

For both parts, derive the form of the pdf, verify that is a pdf, and calculate the mean and the variance.

(a)

$f_X(x) = \beta e^{-\beta x}$, $x > 0$ $\quad y = x^{1/\gamma} \Rightarrow x = y^\gamma$, $y > 0$

$\frac{dx}{dy} = \gamma y^{\gamma-1} \Rightarrow f_Y(y) = f_X(\ y^\gamma\ ) \cdot \gamma y^{\gamma-1}$

$= \beta e^{-\beta y^\gamma} \cdot \gamma y^{\gamma-1}$

$= \gamma\beta \cdot e^{-\beta y^\gamma} \cdot y^{\gamma-1}$, $y > 0$

$\int_0^\infty \gamma\beta\, e^{-\beta y^\gamma} \cdot y^{\gamma-1}\, dy$, denote $t = y^\gamma$, $y = \sqrt[\gamma]{t}$

$= \int_0^\infty \gamma\beta\, e^{-\beta t} \cdot t^{\frac{\gamma-1}{\gamma}} \cdot \frac{1}{\gamma}\, t^{\frac{1}{\gamma}-1}\, dt$

$= \int_0^\infty \beta e^{-\beta t}\, dt = 1$, it's a legal p.d.f.

$E(Y) = \int_0^\infty \gamma\beta\, e^{-\beta y^\gamma} \cdot y^\gamma\, dy$, let $t = y^\gamma$, $y = \sqrt[\gamma]{t}$

$= \int_0^\infty \gamma\beta\, e^{-\beta t} \cdot t \cdot \frac{1}{\gamma}\, t^{\frac{1-\gamma}{\gamma}}\, dt$ $\qquad$ Gamma kernel: $Ga(\frac{1}{\gamma}+1, \beta)$

$= \int_0^\infty \beta e^{-\beta t} \cdot t^{\frac{1}{\gamma}}\, dt = \beta \cdot \int_0^\infty \underbrace{e^{-\beta t} \cdot t^{\frac{1}{\gamma}}}\, dt$

$= \beta \cdot \frac{\Gamma(\frac{1}{\gamma}+1)}{\beta^{1/\gamma+1}} = \Gamma(1/\gamma+1) \cdot \beta^{-1/\gamma}$

$E(Y^2) = \int_0^\infty \gamma\beta e^{-\beta y^\gamma} \cdot y^{\gamma+1}\, dy$ let $t = y^\gamma$, $y = \sqrt[\gamma]{t}$

$\Rightarrow E(Y^2) = \int_0^\infty \gamma\beta \cdot e^{-\beta t} \cdot t^{\frac{\gamma+1}{\gamma}} \cdot \gamma \cdot t^{\frac{1-\gamma}{\gamma}}\, dt$, $t > 0$

$= \int_0^\infty \beta \cdot e^{-\beta t} \cdot t^{\frac{2}{\gamma}}\, dt$

$= \beta \cdot \int_0^\infty \underbrace{e^{-\beta t} \cdot t^{\frac{2}{\gamma}}}\, dt$

Gamma Kernel : Gamma$(2/\gamma + 1, \beta)$

$= \beta \cdot \dfrac{\Gamma(2/\gamma + 1)}{\beta^{2/\gamma + 1}} = \Gamma(2/\gamma + 1) \cdot \beta^{-2/\gamma}$

$Var(Y) = E(Y^2) - [E(Y)]^2 = \Gamma(2/\gamma + 1) \cdot \beta^{-2/\gamma} - \left(\Gamma(1/\gamma + 1) \cdot \beta^{-1/\gamma}\right)^2$

$= \beta^{-2/\gamma} \left(\Gamma(2/\gamma + 1) - \Gamma(1/\gamma + 1)^2\right)$

(b). $Y = \left(\frac{2x}{\beta}\right)^{1/2}$, $x = \frac{y^2\beta}{2}$, $x > 0$, $y > 0$

$f_X(x) = \beta e^{-\beta x}$, $x > 0$, $dx/dy = \beta y$

$f_Y(y) = f_X\left(\frac{y^2\beta}{2}\right) \cdot |\beta y| = \beta e^{-\frac{\beta^2 y^2}{2}} \cdot \beta y = \beta^2 y \cdot e^{-\frac{\beta^2 y^2}{2}}$, $y > 0$

$\int_0^\infty \beta^2 y\, e^{-\frac{\beta^2 y^2}{2}}\, dy = \int_0^\infty \beta^2 \cdot e^{-\frac{\beta^2 t}{2}}\, dt$, if $t = y^2$.

$= -2 \cdot \frac{1}{\beta^2} \cdot \beta^2 \cdot e^{-\frac{\beta^2}{2} \cdot t} \Big|_{t=0}^{t=\infty} = 1$. it's a legal pdf

$$E(Y) = \int_0^\infty \beta^2 y^2 e^{-\frac{\beta^2 y^2}{2}} dy \quad \text{let } t = y^2 \quad y = \sqrt{t}$$

$$= \int_0^\infty \beta^2 t \cdot e^{-\frac{\beta^2}{2} t} \cdot \frac{1}{2\sqrt{t}} dt$$

$$= \frac{\beta^2}{2} \int_0^\infty \underbrace{\sqrt{t} \cdot e^{-\frac{\beta^2}{2} t}}_{\text{It's a Gamma Kernel: Gamma}\left(\frac{3}{2}, \frac{\beta^2}{2}\right)} dt$$

$$\Rightarrow \quad = \frac{\beta^2}{2} \cdot \frac{\Gamma(\frac{3}{2})}{(\frac{\beta^2}{2})^{\frac{3}{2}}} = \frac{\beta^2}{2} \cdot \frac{\frac{1}{2}\sqrt{\pi} \cdot 2\sqrt{2}}{\beta^3} = \frac{\sqrt{2\pi}}{2\beta}$$

$$E(Y^2) = \int_0^\infty \beta^2 y^3 e^{-\frac{\beta^2 y^2}{2}} dy, \quad \text{let } t = y^2, \quad y = \sqrt{t}$$

$$= \int_0^\infty \beta^2 t^{\frac{3}{2}} e^{-\frac{\beta^2}{2} t} \cdot \frac{1}{2\sqrt{t}} dt$$

$$= \frac{\beta^2}{2} \cdot \int_0^\infty \underbrace{t \cdot e^{-\frac{\beta^2}{2} t}}_{\text{Gamma Kernel: Gamma}\left(2, \frac{\beta^2}{2}\right)} dt$$

$$= \frac{\beta^2}{2} \cdot \frac{\Gamma(2)}{(\frac{\beta^2}{2})^2} = \frac{2}{\beta^2}$$

$$Var(Y) = E(Y^2) - [E(Y)]^2 = \frac{2}{\beta^2} - \frac{2\pi}{4\beta^2} = \frac{4-\pi}{2\beta^2}$$

8. Let $Y \mid X \sim \text{Poisson}(X)$ and let $X \sim \text{Exp}(\lambda)$. What is the marginal distribution of Y?

$f_{Y|X}(x, y) = \frac{x^y}{y!} \cdot e^{-x}$

$f_X(x) = \lambda e^{-\lambda x}$

$\Rightarrow f_{X,Y}(x, y) = f_{Y|X}(x, y) \cdot f_X(x) = \frac{x^y}{y!} e^{-(\lambda+1)x} \cdot \lambda$

$f_Y(y) = \int_0^\infty \frac{x^y}{y!} e^{-(\lambda+1)x} \lambda \, dx$

$= \frac{\lambda}{y!} \cdot \int_0^\infty \underbrace{x^y \cdot e^{-(\lambda+1)x}}_{\text{It's a Gamma Kernel: Gamma}(y+1,\, \lambda+1)} dx$

$= \frac{\lambda}{y!} \cdot \frac{\Gamma(y+1)}{(\lambda+1)^{y+1}}$ Since y is integers, $\Gamma(y+1) = y!$

$= \frac{\lambda}{(\lambda+1)^{y+1}}, \quad y \in \{0, 1, 2, \dots\}$

Double Check: $\sum_{y=0}^{\infty} P(Y=y) = \frac{\lambda}{\lambda+1} / \left(1 - \frac{1}{\lambda+1}\right) = 1$

↳ Legal p.m.f.

9. (Robert) If $y \sim \text{Binomial}(n, \theta)$ and $x \sim \text{Binomial}(m, \theta)$, and $\theta \sim \text{Beta}(\alpha, \beta)$. Find the predictive distribution of y given x.

$$f(y|x) = \int_0^1 f(y, \theta|x)\, d\theta = \int_0^1 f(y|\theta) \cdot f(\theta|x)\, d\theta$$

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

$$\propto \theta^{\alpha-1}(1-\theta)^{\beta-1} \cdot \theta^{x}(1-\theta)^{m-x}$$

$$= \theta^{\alpha+x-1}(1-\theta)^{\beta+m-x-1}$$

Therefore, $\theta|x \sim \text{Beta}(\alpha+x, \beta+m-x)$

$$f(y|\theta) = \binom{n}{y}\theta^{y}(1-\theta)^{n-y}$$

$$\Rightarrow f(y|x) = \int_0^1 \binom{n}{y}\theta^{y}(1-\theta)^{n-y} \cdot \frac{1}{B(\alpha+x, \beta+m-x)} \cdot \theta^{\alpha+x-1}(1-\theta)^{\beta+m-x-1}\, d\theta$$

$$= \frac{\binom{n}{y}}{B(\alpha+x, \beta+m-x)} \cdot \int_0^1 \theta^{\alpha+x+y-1}(1-\theta)^{\beta+m-x+n-y-1}\, d\theta$$

$$= \binom{n}{y} \cdot \frac{B(\alpha+x+y, \beta+m-x+n-y)}{B(\alpha+x, \beta+m-x)}$$

10. (Robert) Give the posterior and the marginal distributions in the following cases:

(a) $x \mid \sigma^2 \sim \mathrm{N}(0, \sigma^2)$ and $1/\sigma^2 \sim \mathrm{Gamma}(1, 2)$.

(b) $x \mid p \sim$ Negative-Binomial$(10, p)$ and $p \sim \mathrm{Beta}(1/2, 1/2)$.

(a). $\sigma^2 \sim IG(1,2)$. $\sigma^2 > 0$, $x \in R$

Joint: $h(x,\sigma^2) = f(x|\sigma^2) \cdot \pi(\sigma^2)$

$= (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left(-\frac{x^2}{2\sigma^2}\right) \cdot \pi(\sigma^2)$

$\pi(\sigma^2) = \frac{1}{\Gamma(1)\, 2^1} \cdot (\sigma^2)^{-2} e^{-\frac{1}{2\sigma^2}}$

$\Rightarrow$ Joint: $h(x,\sigma^2) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left(-\frac{x^2}{2\sigma^2}\right) \cdot \frac{1}{2} \cdot (\sigma^2)^{-2} \exp\left(-\frac{1}{2\sigma^2}\right)$

$m(x) = \int_0^\infty h(x,\sigma^2)\, d\sigma^2 = \frac{1}{2\sqrt{2\pi}} \int_0^\infty (\sigma^2)^{-\frac{5}{2}} \exp\left(-\frac{x^2+1}{2\sigma^2}\right) d\sigma^2$

Let $\frac{\sigma^2}{x^2+1} = t$: $\sigma^2 = t(x^2+1)$

$\Rightarrow m(x) = \frac{1}{2\sqrt{2\pi}} \cdot \int_0^\infty \left[(x^2+1)t\right]^{-\frac{5}{2}} \exp\left(-\frac{1}{2t}\right) d\left((x^2+1)t\right)$

$= \frac{1}{2\sqrt{2\pi}} (x^2+1)^{-\frac{3}{2}} \int_0^\infty \underbrace{t^{-\frac{5}{2}} \exp\left(-\frac{1}{2t}\right)}_{\text{Kenel of } IG(\frac{1}{2},2)} dt$

$= \frac{1}{2\sqrt{2\pi}} (x^2+1)^{-\frac{3}{2}} \cdot \Gamma(\tfrac{1}{2}) \cdot 2^{\frac{1}{2}}$

$= \frac{1}{2} (x^2+1)^{-\frac{3}{2}}$

$\Rightarrow \pi(\sigma^2|x) = \frac{h(\sigma^2,x)}{m(x)} = \frac{\frac{1}{2}(2\pi)^{-\frac{1}{2}} (\sigma^2)^{-\frac{5}{2}} \exp\left(-\frac{x^2+1}{2\sigma^2}\right)}{\frac{1}{2}(x^2+1)^{-\frac{3}{2}}}$

$\propto (\sigma^2)^{-\frac{5}{2}} \exp\left(-\frac{1}{2 \cdot \frac{\sigma^2}{(x^2+1)}}\right)$. kernel of $IG: \left(\frac{3}{2}, \frac{2}{x^2+1}\right)$

(b) $f(x|p) = \binom{10+x-1}{x} p^{10} (1-p)^{x}$, $x \in \{0, 1, 2, \cdots\}$

$\pi(p) = \frac{1}{B(\frac{1}{2}, \frac{1}{2})} p^{-\frac{1}{2}} (1-p)^{-\frac{1}{2}}$, $p \in (0.1)$

$h(x,p) = f(x|p) \cdot \pi(p) = \binom{10+x-1}{x} \cdot \frac{1}{B(\frac{1}{2}, \frac{1}{2})} p^{10-\frac{1}{2}} (1-p)^{x-\frac{1}{2}}$

$m(x) = \int_0^1 h(x,p)\, dp = \binom{10+x-1}{x} \cdot \frac{1}{B(\frac{1}{2}, \frac{1}{2})} \cdot \int_0^1 \underbrace{p^{10-\frac{1}{2}} (1-p)^{x-\frac{1}{2}}}_{\text{kenel of Beta}(10+\frac{1}{2},\, x+\frac{1}{2})} dp$

$= \binom{10+x-1}{x} \frac{B(10+\frac{1}{2},\, x+\frac{1}{2})}{B(\frac{1}{2}, \frac{1}{2})}$

$\pi(p|x) = \frac{h(p,x)}{m(x)} \propto h(p,x)$

$\propto \underbrace{p^{10-\frac{1}{2}} (1-p)^{x-\frac{1}{2}}}_{\text{kenel of Beta}(10+\frac{1}{2},\, x+\frac{1}{2})}$

$\Rightarrow p|x \sim \text{Beta}\left(\frac{21}{2},\, x+\frac{1}{2}\right)$

i.e. $\pi(p|x) = \frac{1}{B(\frac{21}{2},\, x+\frac{1}{2})} \cdot p^{10-\frac{1}{2}} (1-p)^{x-\frac{1}{2}}$

11. Assume that an observation, $x_1, \ldots, x_n$ are iid from $N(\theta, \sigma^2)$, where μ and σ^2 are unknown. Consider $\tilde{\pi}(\theta, \sigma^2) \propto 1/\sigma^2$ (not a probability density, i.e., improper, Jeffreys prior).

(a) Find the joint posterior distribution.

(b) Find the posterior distributions $\pi(\theta \mid \bar{x}, s^2, \sigma^2)$ and $\pi(\sigma^2 \mid \bar{x}, s^2)$.

(c) Find the marginal posterior distribution of θ, $\pi(\theta \mid \bar{x}, s^2)$.

(a) $f(x_1, x_2 \cdots x_n \mid \theta, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_i-\theta)^2}{2\sigma^2}\right)$

$= (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i-\theta)^2\right)$

$h(x_1, x_2 \cdots x_n, \theta, \sigma^2) = f(x_1, x_2 \cdots x_n \mid \theta, \sigma^2) \cdot \pi(\theta, \sigma^2)$

$\propto (\sigma^2)^{-\frac{n}{2}-1} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i-\theta)^2\right)$

$\pi(\theta, \sigma^2 \mid x_1, x_2 \cdots x_n) = \frac{h(x_1, x_2 \cdots x_n, \theta, \sigma^2)}{f(x_1, x_2 \cdots x_n)} \propto h(x_1, x_2 \cdots x_n, \theta, \sigma^2)$

$\propto (\sigma^2)^{-\frac{n}{2}-1} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i-\theta)^2\right)$

$= (\sigma^2)^{-\frac{n}{2}-1} \exp\left(-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{n} (x_i-\bar{x})^2 + n(\bar{x}-\theta)^2\right)\right)$

(b). $\pi(\theta \mid \bar{x}, s^2, \sigma^2) = \frac{\pi(\theta, \sigma^2 \mid \bar{x}, s^2)}{\pi(\sigma^2 \mid \bar{x}, s^2)}$

$\pi(\sigma^2 \mid \bar{x}, s^2) = \int_{-\infty}^{\infty} \pi(\theta, \sigma^2 \mid \bar{x}, s^2)\, d\theta$

$\propto (\sigma^2)^{-\frac{n}{2}-1} \exp\left(-\frac{(n-1)s^2}{2\sigma^2}\right) \cdot \int_{-\infty}^{\infty} \underbrace{\exp\left(-\frac{n(\bar{x}-\theta)^2}{2\sigma^2}\right)}_{\text{kenel of } N(\bar{x}, \frac{\sigma^2}{n})} d\theta$

$\propto (\sigma^2)^{-\frac{n}{2}-1} \exp\left(-\frac{(n-1)s^2}{2\sigma^2}\right) \cdot \left(\frac{\sigma^2}{n}\right)^{\frac{1}{2}}$

$\propto (\sigma^2)^{-\left(\frac{n}{2}+\frac{1}{2}\right)} \exp\left(-\frac{1}{\frac{2}{(n-1)s^2} \cdot \sigma^2}\right) \sim IG\left(\frac{n-1}{2}, \frac{2}{(n-1)s^2}\right)$

$$\pi(\theta \mid s^2, \bar{x}, \sigma^2) \propto \frac{(\sigma^2)^{-(\frac{n}{2}+1)} \exp\left(-\frac{(n-1)s^2}{2\sigma^2}\right) \exp\left(-\frac{n(\bar{x}-\theta)^2}{2\sigma^2}\right)}{(\sigma^2)^{-(\frac{n}{2}+\frac{3}{2})} \exp\left(-\frac{(n-1)s^2}{2\sigma^2}\right)}$$

$$\propto \exp\left(-\frac{(\bar{x}-\theta)^2}{\frac{2\sigma^2}{n}}\right) \to \text{Kernel of } N\left(\bar{x}, \frac{\sigma^2}{n}\right)$$

(c) $\pi(\theta \mid \bar{x}, s^2) = \int_0^\infty \pi(\theta, \sigma^2 \mid \bar{x}, s^2)\, d\sigma^2$

$$\propto \int_0^\infty \underbrace{(\sigma^2)^{-\frac{n}{2}-1} \exp\left(-\frac{(n-1)s^2 + n(\bar{x}-\theta)^2}{2\sigma^2}\right)}_{\text{Kenel of } IG\left(\frac{n}{2}, \frac{2}{(n-1)s^2+n(\bar{x}-\theta)^2}\right)} d\sigma^2$$

$$\propto \left(\frac{2}{(n-1)s^2 + n(\bar{x}-\theta)^2}\right)^{\frac{n}{2}}$$

It's a t-distribution, after reparameterizing:

It follows $t\left(n-1, \bar{x}, \frac{s^2}{n}\right)$

ps: $S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$

12. Consider $\boldsymbol{x}_i \mid \boldsymbol{\theta}, \Sigma \overset{iid}{\sim} \mathrm{N}_p(\boldsymbol{\theta}, \Sigma)$, $i = 1, \ldots, n$, where $\mathrm{N}_p(\boldsymbol{\theta}, \Sigma)$ represents the $p-$dimensional normal distribution with mean vector $\boldsymbol{\theta} \in \mathbb{R}^p$ and covariance matrix Σ ($p \times p$ positive definite matrix). Suppose $\boldsymbol{\theta}$ and Σ are unknown. Consider the following conjugate prior distributions;

$$\boldsymbol{\theta} \mid \Sigma \sim \mathrm{N}_p(\boldsymbol{\mu}, 1/n_0\Sigma), \text{ and } \Sigma^{-1} \sim \mathrm{Wishart}_p(\alpha, W).$$

The Wishart distribution is described in Robert Exercise #3.21. Note that if $\Sigma^{-1} \sim \mathrm{Wishart}_p(\alpha, W)$, $\Sigma \sim \text{inverse-Wishart}_p(\alpha, W^{-1})$.

fact: Given n observations $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ of $\mathrm{N}_p(\boldsymbol{\theta}, \Sigma)$, a sufficient statistic is

$$\bar{\boldsymbol{x}} = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{x}_i, \text{ and } S = \sum_{i=1}^{n} (\boldsymbol{x}_i - \bar{\boldsymbol{x}})(\boldsymbol{x}_i - \bar{\boldsymbol{x}})^t.$$

Tip: Read Robert §4.4.2 (and/or BDA §3.6).

(a) Find an expression of the joint posterior distribution as follows;

$$\pi(\boldsymbol{\theta}, \Sigma^{-1} \mid \bar{\boldsymbol{x}}, S) = \pi_1(\boldsymbol{\theta} \mid \Sigma, \bar{\boldsymbol{x}}, S)\pi_2(\Sigma^{-1} \mid \bar{\boldsymbol{x}}, S).$$

Also, identify $\pi_1(\boldsymbol{\theta} \mid \Sigma, \bar{\boldsymbol{x}}, S)$ and $\pi_2(\Sigma^{-1} \mid \bar{\boldsymbol{x}}, S)$

(b) Is the prior conjugate? Explain.

(a) $f(x \mid \theta, \Sigma) = \prod_{i=1}^{n} |\Sigma|^{-1/2} (2\pi)^{-\frac{p}{2}} \exp\left(-\frac{1}{2}(x_i - \theta)^T \Sigma^{-1} (x_i - \theta)\right)$

$$= |\Sigma|^{-n/2} (2\pi)^{-\frac{np}{2}} \exp\left(-\frac{1}{2}\sum_{i=1}^{n} (x_i - \theta)^T \Sigma^{-1} (x_i - \theta)\right)$$

$$= |\Sigma|^{-n/2} (2\pi)^{-\frac{np}{2}} \exp\Big(-\frac{1}{2} \underbrace{\sum_{i=1}^{n} \left\{[(x_i - \bar{x}) + (\bar{x} - \theta)]^T \Sigma^{-1} [(x_i - \bar{x}) + (\bar{x} - \theta)]\right\}}_{\text{Define to } \star}\Big)$$

$$\star = \sum_{i=1}^{n} \Big\{(x_i - \bar{x})^T \Sigma^{-1} (x_i - \bar{x}) + (\bar{x} - \theta)^T \Sigma^{-1} (\bar{x} - \theta) + \underbrace{(x_i - \bar{x})^T \Sigma^{-1} (\bar{x} - \theta) + (\bar{x} - \theta)^T \Sigma^{-1} (x_i - \bar{x})}_{\text{It's } 0 \text{ since } \sum_{i=1}^{n}(x_i - \bar{x}) = 0}\Big\}$$

$$= \sum_{i=1}^{n} \Big\{\underbrace{(x_i - \bar{x})^T \Sigma^{-1} (x_i - \bar{x})}_{\text{Define to } \star\star} + (\bar{x} - \theta)^T \Sigma^{-1} (\bar{x} - \theta)\Big\}$$

$\star\star = \sum_{i=1}^{n} (x_i - \bar{x})^T \Sigma^{-1} (x_i - \bar{x})$. it's a scalar . $\star\star \equiv tr(\star\star)$

$= \sum_{i=1}^{n} tr(x_i - \bar{x})^T \Sigma^{-1} (x_i - \bar{x})$

$= \sum_{i=1}^{n} tr(x_i - \bar{x})^T (x_i - \bar{x}) \Sigma^{-1}$

$= tr\left\{\sum_{i=1}^{n} (x_i - \bar{x})^T (x_i - \bar{x}) \Sigma^{-1}\right\} = tr(S \cdot \Sigma^{-1})$

$\Rightarrow f(x \mid \theta, \Sigma) = |\Sigma|^{-\frac{n}{2}} (2\pi)^{-\frac{np}{2}} \exp\left(-\frac{n}{2} (\bar{x} - \theta)^T \Sigma^{-1} (\bar{x} - \theta) - \frac{1}{2} tr(S \cdot \Sigma^{-1})\right)$

$$f(\underset{\Sigma}{W} \mid \alpha, \underset{W}{\Sigma}) = \frac{|W|^{-\frac{\alpha+m+1)}{2}} \exp\left(-tr(\Sigma^{-1} W^{-1})\right)/2}{\Gamma_m(\alpha)|\Sigma|^{\alpha/2}}, \quad W > 0.$$

$\Sigma \sim$ Inverse Wishart distribution as

Prior: $\pi(\theta, \Sigma) = \left|\frac{\Sigma}{n_0}\right|^{-1/2} \cdot (2\pi)^{-\frac{p}{2}} \exp\left(-\frac{1}{2} (\theta - \mu)^T \left(\frac{\Sigma}{n_0}\right)^{-1} (\theta - \mu)\right) \cdot$

$$\frac{|\Sigma|^{-\frac{\alpha+p+1}{2}} \exp\left(-tr(W^{-1}\Sigma^{-1})\right)/2}{\Gamma_p(\alpha) \; |W|^{\alpha/2}}$$

Posterior: $\pi(\theta, \Sigma \mid S^2, \bar{x}) \propto f(x \mid \theta, \Sigma)\, \pi(\theta, \Sigma)$

$\propto |\Sigma|^{-\frac{n}{2}} \cdot |\Sigma|^{-1/2} \cdot |\Sigma|^{-\frac{\alpha+p+1}{2}} \cdot \exp\left(-\frac{n}{2} (\bar{x} - \theta)^T \Sigma^{-1} (\bar{x} - \theta) - \frac{n_0}{2} (\theta - \mu)^T\right.$

$\left.\Sigma^{-1} (\theta - \mu) - \frac{1}{2} tr(S \cdot \Sigma^{-1}) - \frac{1}{2} tr(W^{-1} \Sigma^{-1})\right)$

$= |\Sigma|^{-\frac{\alpha+p+n+2}{2}} \exp\left(-\frac{n}{2} (\bar{x} - \theta)^T \Sigma^{-1} (\bar{x} - \theta) - \frac{n_0}{2} (\theta - \mu)^T \Sigma^{-1} (\theta - \mu) - \frac{1}{2} tr((S + W^{-1}) \Sigma^{-1})\right)$

$= |\Sigma|^{-\frac{\alpha+p+n+2}{2}} \exp\left(-\frac{1}{2} tr\left(\left[n (\bar{x} - \theta)^T (\bar{x} - \theta) + n_0 (\theta - \mu)^T (\theta - \mu) + (S + W^{-1})\right] \Sigma^{-1}\right)\right)$

$$\pi_2(\Sigma \mid \bar{x}, S) = \int_{R^p} \pi(\theta, \Sigma \mid \bar{x}, s)\, d\theta$$

$$\propto \int_{R^p} |\Sigma|^{-\frac{\alpha+p+n+2}{2}} \exp\left(-\frac{1}{2} tr\left(n(\bar{x}-\theta)^T(\bar{x}-\theta) + n_0(\theta-\mu)^T(\theta-\mu) + (S+W^{-1})\right)\Sigma^{-1}\right)$$

$$= |\Sigma|^{-\frac{\alpha+p+n+2}{2}} \cdot \exp\left(-\frac{1}{2} tr\left((S+W^{-1})\Sigma^{-1}\right)\right) \cdot$$

$$\int_{R^p} \exp\left(-\frac{1}{2} tr\left(\underbrace{n(\theta-\bar{x})^T(\theta-\bar{x}) + n_0(\theta-\mu)^T(\theta-\mu)}_{\text{Denote to } I}\right)\Sigma^{-1}\right)$$

$$I = n\theta^T\theta - n\theta^T\bar{x} - n\bar{x}^T\theta + n\bar{x}^T\bar{x} + n_0\theta^T\theta - n_0\theta^T\mu - n_0\mu^T\theta + n_0\mu^T\mu$$

$$= (n_0+n)\theta^T\theta - \theta^T(n\bar{x}+n_0\mu) - (n\bar{x}+n_0\mu)^T\theta + n\bar{x}^T\bar{x} + n_0\mu^T\mu$$

$$= (n_0+n)\left[\theta^T\theta - \theta^T\left(\frac{n}{n_0+n}\bar{x} + \frac{n_0}{n_0+n}\mu\right) - \left(\frac{n}{n_0+n}\bar{x} + \frac{n_0}{n_0+n}\mu\right)^T\theta + \frac{n}{n_0+n}\bar{x}^T\bar{x} + \frac{n_0}{n_0+n}\mu^T\mu\right]$$

$$= (n_0+n)\left[\underbrace{\left(\theta - \frac{1}{n_0+n}(n\bar{x}+n_0\mu)\right)^T}_{\text{Denote as } t^T}\left(\theta - \frac{1}{n_0+n}(n\bar{x}+n_0\mu)\right) + \boxed{\frac{nn_0(\bar{x}-\mu)^T(\bar{x}-\mu)}{(n+n_0)^2}}\right]$$

proved in the appendix

$$= (n_0+n)\left[t^Tt + \frac{nn_0(\bar{x}-\mu)^T(\bar{x}-\mu)}{(n_0+n)^2}\right]$$

$$\Rightarrow \pi_2(\Sigma \mid \bar{x}, S) \propto |\Sigma|^{-\frac{\alpha+p+n+2}{2}} \exp\left(-\frac{1}{2} tr\left((S+W^{-1}) + \frac{nn_0}{(n_0+n)^2}(\bar{x}-\mu)^T(\bar{x}-\mu)\right)\Sigma^{-1}\right)$$

$$\cdot \underbrace{\int \exp\left[-\frac{1}{2}\left(\theta - \frac{1}{n_0+n}(n\bar{x}+n_0\mu)\right)^T\left(\frac{\Sigma}{n_0+n}\right)^{-1}\left(\theta - \frac{1}{n_0+n}(n\bar{x}+n_0\mu)\right)\right] d\theta}_{\text{MVN Kernel}}$$

$$\propto |\Sigma|^{-\frac{\alpha+p+n+1}{2}} \exp\left(-\frac{1}{2} tr\left[\left((S+W^{-1}) + \frac{nn_0}{(n+n_0)^2}(\bar{x}-\mu)^T(\bar{x}-\mu)\right)\Sigma^{-1}\right]\right)$$

It's an Inverse Wishart kernel.

$$\Rightarrow \pi_2(\Sigma|\bar{x},S) \sim \text{Inv-Wishart}\left(\alpha+n, \left[S+W^{-1}+\frac{nn_0}{(n+n_0)^2}(\bar{x}-\mu)^T(\bar{x}-\mu)\right]^{-1}\right)$$

$$\pi_1(\theta|\Sigma,\bar{x},S) \propto \pi(\theta,\Sigma|\bar{x},S)$$

As last page:

$$\pi_1(\theta|\Sigma,\bar{x},S) \propto \exp\left(-\frac{1}{2}\left(\theta-\frac{1}{n+n_0}(n\bar{x}+n_0\mu)\right)^T\left(\theta-\frac{1}{n+n_0}(n\bar{x}+n_0\mu)\right) \cdot \left(\frac{\Sigma}{n+n_0}\right)^{-1}\right)$$

$$\Rightarrow \text{It's a MVN kernel for } N\left(\frac{n\bar{x}+n_0\mu}{n+n_0}, \frac{\Sigma}{n+n_0}\right)$$

(b) Yes, $\theta|\Sigma$ have both MVN form for prior and posterior, and Σ has Inv-Wishart form both in its prior and posterior.

Appendix:

$$(n_0+n)\left[\theta^T\theta - \theta^T\left(\frac{n}{n_0+n}\bar{x}+\frac{n_0}{n_0+n}\mu\right) - \left(\frac{n}{n_0+n}\bar{x}+\frac{n_0}{n_0+n}\mu\right)^T\theta + \frac{n}{n_0+n}\bar{x}^T\bar{x} + \frac{n_0}{n_0+n}\mu^T\mu\right] \quad \star$$

$$\left[\theta - \frac{1}{n_0+n}(n\bar{x}+n_0\mu)\right]^T\left[\theta - \frac{1}{n_0+n}(n\bar{x}+n_0\mu)\right]$$

$$= \theta^T\theta - \theta^T\cdot\frac{1}{n_0+n}(n\bar{x}+n_0\mu) - \frac{1}{n_0+n}(n\bar{x}+n_0\mu)^T\cdot\theta + \frac{1}{(n_0+n)^2}\left((n\bar{x}+n_0\mu)^T(n\bar{x}+n_0\mu)\right) \quad \star\star$$

$$\star - \star\star = \frac{n^2\bar{x}^T\bar{x}}{(n_0+n)^2} - \frac{nn_0\bar{x}^T\mu}{(n_0+n)^2} - \frac{nn_0\mu^T\bar{x}}{(n_0+n)^2} + \frac{n_0^2\mu^T\mu}{(n_0+n)^2}$$

$$= \frac{nn_0(\bar{x}-\mu)^T(\bar{x}-\mu)}{(n_0+n)^2} \quad \star\star\star$$

$$\Rightarrow \star = \star\star + \star\star\star \qquad \text{Q.E.D.}$$