Probability Exam 2

Michael Nelson

Problem 1

Problem 1.a

In order for $F(t|\mathbf{X},\eta)$ to be a proper cdf, we need the following conditions to hold:

1. $\lim_{t\to-\infty} F(t|\mathbf{X},\eta) = 0$ and $\lim_{t\to\infty} F(t|\mathbf{X},\eta) = 1$
2. $F(t|\mathbf{X},\eta)$ is a nondecreasing function of t
3. $\lim_{t\to t_0^+} F(t|\mathbf{X},\eta)) = F(t_0|\mathbf{X},\eta)$ for all $t_0 \in \mathbb{R}$

Let us assume that $F(t,\mathbf{X},\eta)$ is a cdf and see what these three conditions tells us on what properties $\Lambda_0(t)$ has. First, property 1 tells us

$$
\begin{aligned}
1 &= \lim_{t\to\infty} F(t,\mathbf{X},\eta) \\
&= \lim_{t\to\infty} \left(1 - e^{-\Lambda_0(t)e^{\mathbf{X}^\top\beta}\eta}\right) \\
&= 1 - e^{-\lim_{t\to\infty}\Lambda_0(t)e^{\mathbf{X}^\top\beta}\eta}.
\end{aligned}
$$

In particular we must have $\lim_{t\to\infty} \Lambda_0(t) = \infty$. Similarly, property 1 tells us

$$
\begin{aligned}
0 &= \lim_{t\to-\infty} F(t,\mathbf{X},\eta) \\
&= \lim_{t\to-\infty} \left(1 - e^{-\Lambda_0(t)e^{\mathbf{X}^\top\beta}\eta}\right) \\
&= 1 - e^{-\lim_{t\to-\infty}\Lambda_0(t)e^{\mathbf{X}^\top\beta}\eta}.
\end{aligned}
$$

In particular we must have $\lim_{t\to-\infty} \Lambda_0(t) = 0$.

Next, property 2 tells us

$$
\begin{aligned}
t \geq s &\iff F(t|\mathbf{X},\eta) \geq F(s|\mathbf{X},\eta) \\
&\iff F(t|\mathbf{X},\eta) \geq F(s|\mathbf{X},\eta) \\
&\iff 1 - e^{-\Lambda_0(t)e^{\mathbf{X}^\top\beta}\eta} \geq 1 - e^{-\Lambda_0(s)e^{\mathbf{X}^\top\beta}\eta} \\
&\iff e^{-\Lambda_0(t)e^{\mathbf{X}^\top\beta}\eta} \leq e^{-\Lambda_0(s)e^{\mathbf{X}^\top\beta}\eta} \\
&\iff -\Lambda_0(t)e^{\mathbf{X}^\top\beta}\eta \leq -\Lambda_0(s)e^{\mathbf{X}^\top\beta}\eta \\
&\iff \Lambda_0(t) \geq \Lambda_0(s).
\end{aligned}
$$

In particular, $\Lambda_0(t)$ must be a nondecreasing function of t.

Finally, property 3 tells us that for any $t_0 \in \mathbb{R}$, we must have

$$\begin{aligned}
1 - e^{-\Lambda_0(t_0)e^{\mathbf{X}^\top \beta}\eta} &= F(t_0|\mathbf{X}, \eta) \\
&= \lim_{t \to t_0^+} F(t, \mathbf{X}, \eta) \\
&= \lim_{t \to t_0^+} \left(1 - e^{-\Lambda_0(t_0)e^{\mathbf{X}^\top \beta}\eta}\right) \\
&= 1 - e^{-\lim_{t \to t_0^+} \Lambda_0(t_0)e^{\mathbf{X}^\top \beta}\eta}
\end{aligned}$$

In particular, $\Lambda_0(t)$ must be right continuous too.

Problem 1.b

First note that

$$\begin{aligned}
f(t|\mathbf{x}, \eta) &= \partial_t F(t|\mathbf{x}, \eta) \\
&= \Lambda_0'(t)\eta e^{\mathbf{x}^\top \boldsymbol{\beta}} e^{-\Lambda_0(t)e^{\mathbf{x}^\top \beta}\eta}.
\end{aligned}$$

Using this, we calculate

$$\begin{aligned}
S(t|\mathbf{x}) &= \int_t^\infty f(s|\mathbf{x})\mathrm{d}s \\
&= \int_t^\infty \int_0^\infty f(s|\mathbf{x}, \eta)f(\eta)\mathrm{d}\eta\mathrm{d}s \\
&= \frac{1}{\Gamma(\nu)(1/\nu)^\nu} \int_t^\infty \int_0^\infty \left(\Lambda_0'(s)\eta e^{\mathbf{x}^\top \boldsymbol{\beta}} e^{-\Lambda_0(s)e^{\mathbf{x}^\top \beta}\eta}\right) \eta^{\nu-1}e^{-\eta\nu}\mathrm{d}\eta\mathrm{d}s \\
&= \frac{\nu^\nu e^{\mathbf{x}^\top \boldsymbol{\beta}}}{\Gamma(\nu)} \int_t^\infty \Lambda_0'(s) \left(\int_0^\infty \eta^\nu e^{-\left(\Lambda_0(s)e^{\mathbf{x}^\top \beta}+\nu\right)\eta}\mathrm{d}\eta\right) \mathrm{d}s \\
&= \frac{e^{\mathbf{x}^\top \boldsymbol{\beta}}}{\Gamma(\nu)(1/\nu)^\nu} \int_t^\infty \Lambda_0'(s) \frac{\Gamma(\nu+1)}{\left(\Lambda_0(s)e^{\mathbf{x}^\top \boldsymbol{\beta}} + \nu\right)^{\nu+1}}\mathrm{d}s && \text{(integral of gamma distribution)} \\
&= \frac{e^{\mathbf{x}^\top \boldsymbol{\beta}}(\nu+1)}{(1/\nu)^\nu} \int_t^\infty \frac{\Lambda_0'(s)}{\left(\Lambda_0(s)e^{\mathbf{x}^\top \boldsymbol{\beta}} + \nu\right)^{\nu+1}}\mathrm{d}s \\
&= \frac{e^{\mathbf{x}^\top \boldsymbol{\beta}}(\nu+1)}{(1/\nu)^\nu} \left(\left. \frac{-e^{-\mathbf{x}^\top \boldsymbol{\beta}}}{(\nu+1)\left(\Lambda_0(s)e^{\mathbf{x}^\top \boldsymbol{\beta}} + \nu\right)^\nu} \right|_t^\infty \right) \\
&= \frac{e^{\mathbf{x}^\top \boldsymbol{\beta}}(\nu+1)}{(1/\nu)^\nu} \frac{e^{-\mathbf{x}^\top \boldsymbol{\beta}}}{(\nu+1)\left(\Lambda_0(t)e^{\mathbf{x}^\top \boldsymbol{\beta}} + \nu\right)^\nu} \\
&= \left(\frac{\nu}{\Lambda_0(t)e^{\mathbf{x}^\top \boldsymbol{\beta}} + \nu}\right)^\nu
\end{aligned}$$

Problem 1.c

We have

$$
\begin{aligned}
f(s_1,s_2) &= \int_0^\infty f(s_1,s_2|\eta)\mathrm{d}\eta \\
&= \int_0^\infty f(s_1|\eta)f(s_2|\eta)\mathrm{d}\eta \\
&= \int_0^\infty \int_{\mathbb{R}^p} f(s_1|\eta,\mathbf{x})f(s_2|\eta,\mathbf{x})f(\mathbf{x})\mathrm{d}\mathbf{x}\mathrm{d}\eta \\
&= \int_0^\infty \int_{\mathbb{R}^p} \left(\Lambda_0'(s_1)\eta e^{-\Lambda_0(s_1)e^{\mathbf{x}^\top\boldsymbol{\beta}}\eta}e^{\mathbf{x}^\top\boldsymbol{\beta}}\right)\left(\Lambda_0'(s_2)\eta e^{-\Lambda_0(s_2)e^{\mathbf{x}^\top\boldsymbol{\beta}}\eta}e^{\mathbf{x}^\top\boldsymbol{\beta}}\right)f(\mathbf{x})\mathrm{d}\mathbf{x}\mathrm{d}\eta \\
&= \int_0^\infty \int_{\mathbb{R}^p} \Lambda_0'(s_1)\Lambda_0'(s_2)\eta^2 e^{-(\Lambda_0(s_1)+\Lambda_0(s_2))e^{\mathbf{x}^\top\boldsymbol{\beta}}\eta}e^{2\mathbf{x}^\top\boldsymbol{\beta}}f(\mathbf{x})\mathrm{d}\mathbf{x}\mathrm{d}\eta \\
&= \int_{\mathbb{R}^p}\int_0^\infty \Lambda_0'(s_1)\Lambda_0'(s_2)\eta^2 e^{-(\Lambda_0(s_1)+\Lambda_0(s_2))e^{\mathbf{x}^\top\boldsymbol{\beta}}\eta}e^{2\mathbf{x}^\top\boldsymbol{\beta}}f(\mathbf{x})\mathrm{d}\eta\mathrm{d}\mathbf{x} \\
&= \Lambda_0'(s_1)\Lambda_0'(s_2)\int_{\mathbb{R}^p} e^{2\mathbf{x}^\top\boldsymbol{\beta}}f(\mathbf{x})\int_0^\infty \eta^2 e^{-\left((\Lambda_0(s_1)+\Lambda_0(s_2))e^{\mathbf{x}^\top\boldsymbol{\beta}}\right)\eta}\mathrm{d}\eta\mathrm{d}\mathbf{x} \\
&= \Lambda_0'(s_1)\Lambda_0'(s_2)\int_{\mathbb{R}^p} e^{2\mathbf{x}^\top\boldsymbol{\beta}}f(\mathbf{x})\frac{\Gamma(3)}{(\Lambda_0(s_1)+\Lambda_0(s_2))^3 e^{3\mathbf{x}^\top\boldsymbol{\beta}}}\mathrm{d}\mathbf{x} && \text{(integral of gamma distribution)} \\
&= \int_{\mathbb{R}^p}\frac{2\Lambda_0'(s_1)\Lambda_0'(s_2)}{(\Lambda_0(s_1)+\Lambda_0(s_2))^3}e^{-\mathbf{x}^\top\boldsymbol{\beta}}f(\mathbf{x})\mathrm{d}\mathbf{x}
\end{aligned}
$$

In particular, this implies

$$f(s_1,s_2|\mathbf{x}) = \frac{2\Lambda_0'(s_1)\Lambda_0'(s_2)}{(\Lambda_0(s_1)+\Lambda_0(s_2))^3}e^{-\mathbf{x}^\top\boldsymbol{\beta}}.$$

Therefore we have

$$
\begin{aligned}
S(t_1,t_2|\mathbf{x}) &= \int_{t_2}^\infty\int_{t_1}^\infty f(s_1,s_2|\mathbf{x})\mathrm{d}s_1\mathrm{d}s_2 \\
&= \int_{t_2}^\infty\int_{t_1}^\infty \frac{2\Lambda_0'(s_1)\Lambda_0'(s_2)}{(\Lambda_0(s_1)+\Lambda_0(s_2))^3}e^{-\mathbf{x}^\top\boldsymbol{\beta}}\mathrm{d}s_1\mathrm{d}s_2 \\
&= 2e^{-\mathbf{x}^\top\boldsymbol{\beta}}\int_{t_2}^\infty\left(\int_{t_1}^\infty \frac{2\Lambda_0'(s_1)\Lambda_0'(s_2)}{(\Lambda_0(s_1)+\Lambda_0(s_2))^3}\mathrm{d}s_1\right)\mathrm{d}s_2 \\
&= 2e^{-\mathbf{x}^\top\boldsymbol{\beta}}\int_{t_2}^\infty\left(\left.\frac{-\Lambda_0'(s_2)}{(\Lambda_0(s_1)+\Lambda_0(s_2))^2}\right|_{t_1}^{\infty}\right)\mathrm{d}s_2 \\
&= 2e^{-\mathbf{x}^\top\boldsymbol{\beta}}\int_{t_2}^\infty \frac{\Lambda_0'(s_2)}{(\Lambda_0(t_1)+\Lambda_0(s_2))^2}\mathrm{d}s_2 \\
&= e^{-\mathbf{x}^\top\boldsymbol{\beta}}\left.\left(\frac{-1}{\Lambda_0(t_1)+\Lambda_0(s_2)}\right)\right|_{t_2}^{\infty} \\
&= e^{-\mathbf{x}^\top\boldsymbol{\beta}}\left(\frac{1}{\Lambda_0(t_1)+\Lambda_0(t_2)}\right) \\
&= \frac{e^{-\mathbf{x}^\top\boldsymbol{\beta}}}{\Lambda_0(t_1)+\Lambda_0(t_2)}.
\end{aligned}
$$

Problem 1.d

We have

$$
\begin{aligned}
\mathrm{E}(\mathrm{sign}((T_{i_1}-T_{j_1})(T_{i_2}-T_{j_2}))) &= \mathrm{E}\left(\mathrm{E}\left[\mathrm{sign}((T_{i_1}-T_{j_1})(T_{i_2}-T_{j_2})|X\right]\right) \\
&= \mathrm{E}\left(\mathrm{P}[(T_{i1}-T_{j1})(T_{i2}-T_{j2})>0|X]-\mathrm{P}[(T_{i1}-T_{j1})(T_{i2}-T_{j2})<0|X]\right) \\
&= \mathrm{E}\left(\mathrm{P}[T_{i1}>T_{j1},T_{i2}>T_{j2}|X]+\mathrm{P}[T_{i1}-T_{j1},T_{i2}-T_{j2}<0|X]-\mathrm{P}[(T_{i1}-T_{j1})(T_{i2}-T_{j2})<0|X]\right)
\end{aligned}
$$

Problem 2

Problem 2.a

Let $\mathcal{A} = \operatorname{supp} X = \mathbb{R}_{>0}$ and define $g\colon \mathcal{A} \to \mathbb{R}$ by

$$g(x) = \frac{2x}{\beta}$$

for all $x \in \mathcal{A}$. Denote $\mathcal{B} = \operatorname{im} g = \mathbb{R}_{>0}$ and $Y = g(X)$. Then g is a diffeomorphism with inverse $h\colon \mathcal{B} \to \mathcal{A}$ given by

$$h(y) = \frac{\beta y}{2}$$

for all $y \in \mathcal{B}$. The absolute value of the derivative of h at $y \in \mathcal{B}$ is given by

$$\begin{aligned}\left|\frac{\mathrm{d}}{\mathrm{d}y}h(y)\right| &= \left|\frac{\beta}{2}\right| \\ &= \frac{\beta}{2}\end{aligned}$$

It follows that

$$\begin{aligned}f_Y(y) &= f_X(h(y))\frac{\beta}{2} \\ &= \frac{1}{\Gamma(\alpha)\beta^\alpha 2^{\alpha-1}}\beta^{\alpha-1}y^{\alpha-1}e^{-\frac{(\beta y/2)}{\beta}}\frac{\beta}{2} \\ &= \frac{1}{\Gamma(\alpha)2^\alpha}y^{\alpha-1}e^{-\frac{y}{2}}.\end{aligned}$$

Therefore $Y \sim \chi^2_{2\alpha}$.

Problem 2.b

Let $\mathcal{A} = \operatorname{supp}(X_1, X_2) = \mathbb{R}^2_{>0}$ and define $g = (g_1, g_2)\colon \mathcal{A} \to \mathbb{R}^2$ by

$$g_1(x_1, x_2) = \frac{\alpha_1\beta_2 x_1}{\alpha_2\beta_1 x_2} \qquad \text{and} \qquad g_2(x_1, x_2) = x_2$$

for all $(x_1, x_2) \in \mathcal{A}$. Denote $\mathcal{B} = \operatorname{im} g = \mathbb{R}^2_{>0}$, $Y_1 = g_1(X_1, X_2)$, and $Y_2 = g_2(X_1, X_2)$. Then g is a diffeomorphism with inverse $h = (h_1, h_2)\colon \mathcal{B} \to \mathcal{A}$ given by

$$h_1(y_1, y_2) = \frac{\alpha_2\beta_1}{\alpha_1\beta_2}y_1 y_2 \qquad \text{and} \qquad h_2(y_1, h_2) = y_2$$

for all $(y_1, y_2) \in \mathcal{B}$. The absolute value of the Jacobian of h at $(y_1, y_2) \in \mathcal{B}$ is given by

$$\begin{aligned}\left|J_{(y_1,y_2)}(h)\right| &= \left|\det\begin{pmatrix}\frac{\alpha_2\beta_1}{\alpha_1\beta_2}y_2 & \frac{\alpha_2\beta_1}{\alpha_1\beta_2}y_1 \\ 0 & 1\end{pmatrix}\right| \\ &= \frac{\alpha_2\beta_1}{\alpha_1\beta_2}y_2.\end{aligned}$$

Thus the joint distribution of Y_1 and Y_2 is given by

$$\begin{aligned}
f_{Y_1,Y_2}(y_1,y_2) &= f_{X_1,X_2}(h(y_1,y_2))\left|J_{(y_1,y_2)}(h)\right| \\
&= \frac{1}{\Gamma(\alpha_1)\Gamma(\alpha_2)\beta_1^{\alpha_1}\beta_2^{\alpha_2}} x_1^{\alpha_1-1}x_2^{\alpha_2-1}e^{-\frac{x_1}{\beta_1}}e^{-\frac{x_2}{\beta_2}}\frac{\alpha_1\beta_2}{\alpha_2\beta_1}y_2 \\
&= \frac{1}{\Gamma(\alpha_1)\Gamma(\alpha_2)\beta_1^{\alpha_1}\beta_2^{\alpha_2}}\left(\frac{\alpha_2^{\alpha_1-1}\beta_1^{\alpha_1-1}}{\alpha_1^{\alpha_1-1}\beta_2^{\alpha_1-1}}y_1^{\alpha_1-1}y_2^{\alpha_1-1}\right)\left(y_2^{\alpha_2-1}\right)\left(e^{-\frac{1}{\beta_1}\frac{\alpha_2\beta_1}{\alpha_1\beta_2}y_1y_2}\right)\left(e^{-\frac{y_2}{\beta_2}}\right)\frac{\alpha_2\beta_1}{\alpha_1\beta_2}y_2 \\
&= \frac{1}{\Gamma(\alpha_1)\Gamma(\alpha_2)\beta_1^{\alpha_1}\beta_2^{\alpha_2}}\left(\frac{\alpha_2^{\alpha_1-1}\beta_1^{\alpha_1-1}}{\alpha_1^{\alpha_1-1}\beta_2^{\alpha_1-1}}y_1^{\alpha_1-1}y_2^{\alpha_1-1}\right)\left(y_2^{\alpha_2-1}\right)\left(e^{-\frac{\alpha_2}{\alpha_1\beta_2}y_1y_2}\right)\left(e^{-\frac{y_2}{\beta_2}}\right)\frac{\alpha_2\beta_1}{\alpha_1\beta_2}y_2 \\
&= \frac{\alpha_2^{\alpha_1}y_1^{\alpha_1-1}y_2^{\alpha_1+\alpha_2-1}}{\alpha_1^{\alpha_1}\Gamma(\alpha_1)\Gamma(\alpha_2)\beta_2^{\alpha_2+\alpha_1}}e^{-\left(\frac{\alpha_2}{\alpha_1\beta_2}y_1y_2+\frac{1}{\beta_2}y_2\right)}
\end{aligned}$$

Therefore the marginal distribution of Y_1 is given by

$$\begin{aligned}
f_{Y_1}(y_1) &= \int_0^\infty f_{Y_1,Y_2}(y_1,y_2)\mathrm{d}y_2 \\
&= \int_0^\infty \frac{\alpha_2^{\alpha_1}y_1^{\alpha_1-1}y_2^{\alpha_1+\alpha_2-1}}{\alpha_1^{\alpha_1}\Gamma(\alpha_1)\Gamma(\alpha_2)\beta_2^{\alpha_2+\alpha_1}}e^{-\left(\frac{\alpha_2}{\alpha_1\beta_2}y_1y_2+\frac{1}{\beta_2}y_2\right)}\mathrm{d}y_2 \\
&= \frac{\alpha_2^{\alpha_1}y_1^{\alpha_1-1}}{\alpha_1^{\alpha_1}\Gamma(\alpha_1)\Gamma(\alpha_2)\beta_2^{\alpha_2+\alpha_1}}\int_0^\infty y_2^{\alpha_1+\alpha_2-1}e^{-\left(\frac{\alpha_2y_1+\alpha_1}{\alpha_1\beta_2}y_2\right)}\mathrm{d}y_2 \\
&= \frac{\alpha_2^{\alpha_1}y_1^{\alpha_1-1}}{\alpha_1^{\alpha_1}\Gamma(\alpha_1)\Gamma(\alpha_2)\beta_2^{\alpha_2+\alpha_1}}\int_0^\infty \left(\frac{\alpha_1\beta_2}{\alpha_2y_1+\alpha_1}\right)^{\alpha_1+\alpha_2-1}u^{\alpha_1+\alpha_2-1}e^{-u}\frac{\alpha_1\beta_2}{\alpha_2y_1+\alpha_1}\mathrm{d}u \quad u\text{-substitution } u=\frac{\alpha_2y_1+\alpha_1}{\alpha_1\beta_2}y_2 \\
&= \frac{\alpha_1^{\alpha_2}\alpha_2^{\alpha_1}y_1^{\alpha_1-1}}{\Gamma(\alpha_1)\Gamma(\alpha_2)(\alpha_2y_1+\alpha_1)^{\alpha_1+\alpha_2}}\int_0^\infty u^{\alpha_1+\alpha_2-1}e^{-u}\mathrm{d}u \\
&= \frac{\Gamma(\alpha_1+\alpha_2)\alpha_1^{\alpha_2}\alpha_2^{\alpha_1}y_1^{\alpha_1-1}}{\Gamma(\alpha_1)\Gamma(\alpha_2)(\alpha_2y_1+\alpha_1)^{\alpha_1+\alpha_2}}.
\end{aligned}$$

Thus $Y_1 \sim \mathrm{F}(2\alpha_1, 2\alpha_2)$.

Problem 2.c

First we prove the following:

Proposition 0.1. *Let $Z = \sum_{i=1}^n X_i$ where for $1 \le i \le n$ we have $X_i \sim \text{exponential}(\beta)$ with X_i and X_j being independent for all $1 \le i < j \le n$, then $Z \sim \Gamma(n,\beta)$.*

Proof. Let $\mathcal{A} = \operatorname{supp}(X_1,\dots,X_n) = \mathbb{R}^n_{>0}$ and define $g = (g_1,\dots,g_n)\colon \mathcal{A} \to \mathbb{R}^n$ by

$$g_1(x_1,\dots,x_n) = \sum_{i=1}^n x_i \qquad \text{and} \qquad g_k(x_1,\dots,x_n) = x_k$$

for all $2 \le k \le n$ and $(x_1,\dots,x_n) \in \mathcal{A}$. Denote $\mathcal{B} = \operatorname{im} g = \{(y_1,\dots,y_n) \in \mathbb{R}^n_{>0} \mid y_1 > \sum_{i=2}^n y_i\}$ and $Y_j = g_j(X_1,\dots,X_n)$ for all $1 \le j \le n$. Then g is a diffeomorphism with inverse $h = (h_1,\dots,h_n)\colon \mathcal{B} \to \mathcal{A}$ given by

$$h_1(y_1,\dots,y_n) = y_1 - \sum_{i=2}^n y_i \qquad \text{and} \qquad h_k(y_1,\dots,y_n) = y_k$$

for all $2 \leq k \leq n$ and $(y_1, \ldots, y_n) \in \mathcal{B}$. The absolute value of the Jacobian of h at $(y_1, \ldots, y_n) \in \mathcal{B}$ is given by

$$
\begin{aligned}
\left|J_{(y_1,\ldots,y_n)}(h)\right| &= \left|\det \begin{pmatrix} 1 & -1 & \cdots & -1 & -1 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}\right| \\
&= 1.
\end{aligned}
$$

Thus the joint distribution of $(Y_1, \ldots, Y_n)$ is

$$
\begin{aligned}
f_{Y_1,\ldots,Y_n}(y_1, \ldots, y_n) &= f_{X_1,\ldots,X_n}(h(y_1, \ldots, y_n)) \left|J_{(y_1,\ldots,y_n)}(h)\right| \\
&= \frac{1}{\beta} e^{-\frac{y_1 - \sum_{i=2}^{n} y_i}{\beta}} \prod_{2=1}^{n} \frac{1}{\beta} e^{-\frac{y_i}{\beta}} \\
&= \frac{1}{\beta} e^{\frac{-y_1 + \sum_{i=2}^{n} y_i}{\beta}} \prod_{2=1}^{n} \frac{1}{\beta} e^{\frac{-y_i}{\beta}} \\
&= \frac{1}{\beta^n} e^{-\frac{y_1}{\beta}}
\end{aligned}
$$

Therefore the marginal distribution of Y_1 is

$$
\begin{aligned}
f_{Y_1}(y_1) &= \int_0^{y_1} \int_0^{y_1 - y_2} \cdots \int_0^{y_1 - y_2 - \cdots - y_{n-1}} f_{Y_1,\ldots,Y_n}(y_1, \ldots, y_n) \mathrm{d}y_n \cdots \mathrm{d}y_3 \mathrm{d}y_2 \\
&= \int_0^{y_1} \int_0^{y_1 - y_2} \cdots \int_0^{y_1 - y_2 - \cdots - y_{n-1}} \frac{1}{\beta^n} e^{\frac{-y_1}{\beta}} \mathrm{d}y_n \cdots \mathrm{d}y_3 \mathrm{d}y_2 \\
&= \frac{1}{\beta^n} e^{\frac{-y_1}{\beta}} \int_0^{y_1} \int_0^{y_1 - y_2} \cdots \int_0^{y_1 - y_2 - \cdots - y_{n-1}} \mathrm{d}y_n \cdots \mathrm{d}y_3 \mathrm{d}y_2 \\
&= \frac{1}{\beta^n} e^{\frac{-y_1}{\beta}} \int_0^{y_1} \int_0^{y_1 - y_2} \cdots \int_0^{y_1 - y_2 - \cdots - y_{n-1}} \mathrm{d}y_n \cdots \mathrm{d}y_3 \mathrm{d}y_2 \\
&= \frac{1}{\Gamma(n)} \frac{1}{\beta^n} y_1^{n-1} e^{\frac{-y_1}{\beta}}
\end{aligned}
$$

where the integral in the fourth line is solved as follows:

$$
\begin{aligned}
&\int_0^{y_1}\int_0^{y_1}\int_0^{y_1-y_2}\cdots\int_0^{y_1-y_2-\cdots-y_{n-1}} \mathrm{d}y_n\cdots\mathrm{d}y_3\mathrm{d}y_2 && k=1\\
&\quad=\int_0^{y_1}\int_0^{y_1-y_2}\cdots\int_0^{y_1-y_2-\cdots-y_{n-2}}(y_1-y_2-\cdots-y_{n-1})\mathrm{d}y_{n-1}\cdots\mathrm{d}y_3\mathrm{d}y_2 && k=1\\
&\quad=\int_0^{y_1}\int_0^{y_1-y_2}\cdots\int_0^{y_1-y_2-\cdots-y_{n-3}}\frac{1}{2}(y_1-y_2-\cdots-y_{n-1})^2\Big|_0^{y_1-y_2-\cdots-y_{n-2}}\mathrm{d}y_{n-2}\cdots\mathrm{d}y_3\mathrm{d}y_2\\
&\quad=\int_0^{y_1}\int_0^{y_1-y_2}\cdots\int_0^{y_1-y_2-\cdots-y_{n-3}}\frac{1}{2}(y_1-y_2-\cdots-y_{n-2})^2\mathrm{d}y_{n-2}\cdots\mathrm{d}y_3\mathrm{d}y_2\\
&\quad\vdots\\
&\quad=\int_0^{y_1}\int_0^{y_1-y_2}\cdots\int_0^{y_1-y_2-\cdots-y_{n-k-1}}\frac{1}{k!}(y_1-y_2-\cdots-y_{n-k})^k\mathrm{d}y_{n-k}\cdots\mathrm{d}y_3\mathrm{d}y_2 && 1\le k\le n-2\\
&\quad=\int_0^{y_1}\int_0^{y_1-y_2}\cdots\int_0^{y_1-y_2-\cdots-y_{n-k-2}}\frac{1}{(k+1)!}(y_1-y_2-\cdots-y_{n-k})^{k+1}\Big|_0^{y_1-y_2-\cdots-y_{n-k+1}}\mathrm{d}y_{n-k}\cdots\mathrm{d}y_3\mathrm{d}y_2\\
&\quad=\int_0^{y_1}\int_0^{y_1-y_2}\cdots\int_0^{y_1-y_2-\cdots-y_{n-k-2}}\frac{1}{(k+1)!}(y_1-y_2-\cdots-y_{n-k-11})^{k+1}\mathrm{d}y_{n-k+1}\cdots\mathrm{d}y_3\mathrm{d}y_2\\
&\quad\vdots\\
&\quad=\int_0^{y_1}\frac{1}{(n-2)!}(y_1-y_2)^{n-1}\mathrm{d}y_2 && k=n-2\\
&\quad=\frac{1}{(n-1)!}y_1^{n-1}.\\
&\quad=\frac{1}{\Gamma(n)}y_1^{n-1}.
\end{aligned}
$$

Therefore $Z=Y_1\sim\Gamma(n,\beta)$. □

Now with the proposition above established, the problem is easy. We simply compose the map with the one given in the proposition above with the one in the previous problem. Thus the sequence of maps

$$
\begin{aligned}
(X_1,\dots,X_n,Y_1,\dots,Y_m) &\mapsto \left(\sum_{i=1}^n X_i,\sum_{j=1}^m Y_j\right)\\
&\mapsto \frac{n\beta}{m\beta}\frac{\sum_{i=1}^n X_i}{\sum_{j=1}^m Y_j}\\
&\mapsto \frac{n}{m}\frac{\sum_{i=1}^n X_i}{\sum_{j=1}^m Y_j}\\
&=\frac{\overline{X}}{\overline{Y}}
\end{aligned}
$$

gives us an $\mathrm{F}(2n,2m)$ distribution.

Problem 3

Problem 3.a

We have

$$\begin{aligned}
\mathrm{E}(X_1) &= \mathrm{E}\left(\mathrm{E}[X_1|\mu,\sigma^2]\right) \\
&= \mathrm{E}(\mu) \\
&= \mathrm{E}\left(\mathrm{E}(\mu|\sigma^2)\right) \\
&= \mathrm{E}(\mu_0\sigma^2) \\
&= \mu_0\mathrm{E}(\sigma^2) \\
&= \mu_0\frac{\nu_0\sigma_0^2/2}{\nu_0/2-1} \\
&= \frac{\mu_0\nu_0\sigma_0^2}{\nu_0-2}.
\end{aligned}$$

Next, we have

$$\begin{aligned}
\mathrm{Var}\,X_1 &= \mathrm{E}\left(\mathrm{Var}(X_1|\mu,\sigma^2)\right) + \mathrm{Var}\left(\mathrm{E}(X_1|\mu,\sigma^2)\right) \\
&= \mathrm{E}(\sigma^2) + \mathrm{Var}(\mu) \\
&= \frac{\nu_0\sigma_0^2}{\nu_0-2} + \mathrm{E}\left(\mathrm{Var}(\mu|\sigma^2)\right) + \mathrm{Var}\left(\mathrm{E}(\mu|\sigma^2)\right) \\
&= \frac{\nu_0\sigma_0^2}{\nu_0-2} + \mathrm{E}\left(\sigma^2/n_0\right) + \mathrm{Var}(\mu_0) \\
&= \frac{\nu_0\sigma_0^2}{\nu_0-2} + \frac{\nu_0\sigma_0^2}{n_0(\nu_0-2)}.
\end{aligned}$$

Problem 3.b

To simplify our notation in what follows, denote $\alpha = \nu_0/2$ and $\beta = \nu_0\sigma_0^2/2$. Observe that the joint distribution of X_i, μ, and σ^2 is given by

$$\begin{aligned}
f(x,\mu,\sigma^2) &= f(x|\mu,\sigma^2)f(\mu,\sigma^2) \\
&= f(x|\mu,\sigma^2)f(\mu|\sigma^2)f(\sigma^2) \\
&= \left(\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2\sigma^2}(x-\mu)^2}\right)\left(\frac{\sqrt{n_0}}{\sigma\sqrt{2\pi}}e^{\frac{-1}{2\sigma^2}n_0(\mu-\mu_0)^2}\right)\left(\frac{\beta^\alpha}{\Gamma(\alpha)}(\sigma^2)^{-\alpha-1}e^{-\beta/\sigma^2}\right) \\
&= \frac{\beta^\alpha\sqrt{n_0}}{\Gamma(\alpha)2\pi}(\sigma^2)^{-\alpha-2}e^{-\frac{1}{2\sigma^2}\left((x-\mu)^2+n_0(\mu-\mu_0)^2+2\beta\right)}.
\end{aligned}$$

Therefore the distribution of $\mu|X_i, \sigma^2$ is

$$
\begin{aligned}
f(\mu|x,\sigma^2) &= \int_0^\infty \int_{-\infty}^\infty f(x,\mu,\sigma^2)\mathrm{d}x\mathrm{d}\sigma^2 \\
&= \frac{\beta^\alpha\sqrt{n_0}}{\Gamma(\alpha)2\pi}\int_0^\infty \int_{-\infty}^\infty (\sigma^2)^{-\alpha-2} e^{-\frac{1}{2\sigma^2}\left((x-\mu)^2+n_0(\mu-\mu_0)^2+2\beta\right)}\mathrm{d}x\mathrm{d}\sigma^2 \\
&= \frac{\beta^\alpha\sqrt{n_0}}{\Gamma(\alpha)2\pi}\int_0^\infty (\sigma^2)^{-\alpha-2} e^{\frac{-1}{2\sigma^2}\left(n_0(\mu-\mu_0)^2+2\beta\right)} \int_{-\infty}^\infty e^{-\frac{1}{2\sigma^2}(x-\mu)^2}\mathrm{d}x\mathrm{d}\sigma^2 \\
&= \frac{\beta^\alpha\sqrt{n_0}}{\Gamma(\alpha)2\pi}\int_0^\infty (\sigma^2)^{-\alpha-2} e^{\frac{-1}{2\sigma^2}\left(n_0(\mu-\mu_0)^2+2\beta\right)}\sigma\sqrt{2\pi}\mathrm{d}\sigma^2 && \text{normal} \\
&= \frac{\beta^\alpha\sqrt{n_0}}{\Gamma(\alpha)\sqrt{2\pi}}\int_0^\infty (\sigma^2)^{-\alpha-3/2} e^{\frac{-1}{2\sigma^2}\left(n_0(\mu-\mu_0)^2+2\beta\right)}\mathrm{d}\sigma^2 \\
&= \frac{\beta^\alpha\sqrt{n_0}}{\Gamma(\alpha)2\pi}\int_0^\infty (\sigma^2)^{-(\alpha+1/2)-1} e^{\frac{-1}{\sigma^2}\left((n_0/2)(\mu-\mu_0)^2+\beta\right)}\mathrm{d}\sigma^2 \\
&= \frac{\beta^\alpha\sqrt{n_0}}{\Gamma(\alpha)2\pi}\frac{\Gamma(\alpha+1/2)}{((n_0/2)(\mu-\mu_0)^2+\beta)^{\alpha+\frac{1}{2}}} && \text{inverse gamma}
\end{aligned}
$$

We also also calculate

$$
\begin{aligned}
f(\mu) &= \int_0^\infty f(\mu|\sigma^2)f(\sigma^2)\mathrm{d}\sigma^2 \\
&= \int_0^\infty \left(\frac{\sqrt{n_0}}{\sigma\sqrt{2\pi}}e^{\frac{-1}{2\sigma^2}n_0(\mu-\mu_0)^2}\right)\left(\frac{\beta^\alpha}{\Gamma(\alpha)}(\sigma^2)^{-\alpha-1}e^{-\beta/\sigma^2}\right)\mathrm{d}\sigma^2 \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)}\frac{\sqrt{n_0}}{\sqrt{2\pi}}\int_0^\infty (\sigma^2)^{-(\alpha+1)-1}\left(e^{\frac{-1}{2\sigma^2}(n_0(\mu-\mu_0)^2+2\beta)}\right)\mathrm{d}\sigma^2 \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)}\frac{\sqrt{n_0}}{\sqrt{2\pi}}\frac{2^{\alpha+1}\Gamma(\alpha+1)}{(n_0(\mu-\mu_0)^2+2\beta)^{\alpha+1}} \\
&= \sqrt{\frac{2n_0}{\pi}}\frac{(2\beta)^\alpha(\alpha+1)}{(n_0(\mu-\mu_0)^2+2\beta)^{\alpha+1}} \\
&= \sqrt{\frac{2n_0}{\pi}}\frac{(\nu_0\sigma_0^2)^{\nu_0/2}(\nu_0/2+1)}{\left(n_0(\mu-\mu_0)^2+\nu_0\sigma_0^2\right)^{\nu_0/2+1}}
\end{aligned}
$$

We also calculate

$$
\begin{aligned}
f(x) &= \int_{-\infty}^\infty f(x|\mu)f(\mu)\mathrm{d}\mu \\
&= \int_{-\infty}^\infty \int_0^\infty f(x|\mu,\sigma^2)f(\mu|\sigma)\mathrm{d}\sigma^2\mathrm{d}\mu \\
&= \int_{-\infty}^\infty \int_0^\infty \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2\sigma^2}(x-\mu)^2}\frac{\sqrt{n_0}}{\sigma\sqrt{2\pi}}e^{\frac{-1}{2\sigma^2}n_0(\mu-\mu_0)^2}\mathrm{d}\sigma^2\mathrm{d}\mu \\
&= \frac{\sqrt{n_0}}{2\pi}\int_0^\infty \int_{-\infty}^\infty (\sigma^2)^{-1}e^{-\frac{1}{2\sigma^2}((x-\mu)^2+n_0(\mu-\mu_0)^2)}\mathrm{d}\mu\mathrm{d}\sigma^2 \\
&= \frac{\sqrt{n_0}}{2\pi}\int_0^\infty (\sigma^2)^{-1}e^{-\frac{1}{2\sigma^2}(x-\mu)^2}\int_{-\infty}^\infty e^{-\frac{1}{2\sigma^2}(n_0(\mu-\mu_0)^2)}\mathrm{d}\mu\mathrm{d}\sigma^2 \\
&= \frac{\sqrt{n_0}}{2\pi}\int_0^\infty (\sigma^2)^{-1}e^{-\frac{1}{2\sigma^2}(x-\mu)^2}\sqrt{2\pi n_0}\sigma\mathrm{d}\sigma^2 \\
&= \frac{n_0}{\sqrt{2\pi}}\int_0^\infty (\sigma^2)^{\frac{1}{2}-1}e^{-\frac{1}{2\sigma^2}(x-\mu)^2}\mathrm{d}\sigma^2 \\
&= \frac{n_0}{\sqrt{2\pi}}\frac{\Gamma(1/2)}{((1/2)(x-\mu)^2)^{1/2}} \\
&= \frac{n_0}{x-\mu}.
\end{aligned}
$$

Problem 3.c

$$
\begin{aligned}
f(\sigma^2|x) &= \frac{f(x,\mu,\sigma^2)}{f(\mu|x,\sigma^2)f(x)} \\
&= \frac{\beta^\alpha\sqrt{n_0}}{\Gamma(\alpha)2\pi}(\sigma^2)^{-\alpha-2}e^{-\frac{1}{2\sigma^2}\left((x-\mu)^2+n_0(\mu-\mu_0)^2+2\beta\right)}\sqrt{\frac{2\pi}{n_0}}\frac{\Gamma(\alpha)}{\beta^\alpha}\frac{\Gamma(\alpha+1/2)}{((n_0/2)(\mu-\mu_0)^2+\beta)^{\alpha+\frac{1}{2}}}\frac{x-\mu}{n_0} \\
&= \frac{x-\mu}{n_0\sqrt{2\pi}}\frac{\Gamma(\alpha)\Gamma(\alpha+1/2)}{\beta^\alpha((n_0/2)(\mu-\mu_0)^2+\beta)^{\alpha+\frac{1}{2}}}(\sigma^2)^{-\alpha-2}e^{-\frac{1}{2\sigma^2}\left((x-\mu)^2+n_0(\mu-\mu_0)^2+2\beta\right)}.
\end{aligned}
$$

Problem 3.d

We have

$$
\begin{aligned}
f(\mu|\mathbf{x}) &= \frac{f(\mathbf{x}|\mu)f(\mu)}{f(\mathbf{x})} \\
&= \frac{(x-\mu)}{n_0}
\end{aligned}
$$

$$
\begin{aligned}
f(\mu|\mathbf{x}) &= \int_0^\infty f(\mu|\mathbf{x},\sigma^2)\mathrm{d}\sigma^2 \\
&= \int_0^\infty \frac{f(x)f(\sigma^2|x)}{f(x,\mu,\sigma^2)}\mathrm{d}\sigma^2 \\
&= \int_0^\infty \frac{f(x)f(\sigma^2|x)}{f(x,\mu,\sigma^2)}\mathrm{d}\sigma^2 \\
&= \frac{2\pi\Gamma(\alpha)^2\Gamma(\alpha+1/2)n_0}{\beta^{2\alpha}((n_0/2)(\mu-\mu_0)^2+\beta)^{\alpha+\frac{1}{2}}(x-\mu)\sqrt{n_0}}\int_0^\infty(\sigma^2)^{-\alpha-2}e^{-\frac{1}{2\sigma^2}\left((x-\mu)^2+n_0(\mu-\mu_0)^2+2\beta\right)}\mathrm{d}\sigma^2 \\
&= \frac{2\pi\Gamma(\alpha)^2\Gamma(\alpha+1/2)n_0}{\beta^{2\alpha}((n_0/2)(\mu-\mu_0)^2+\beta)^{\alpha+\frac{1}{2}}(x-\mu)\sqrt{n_0}}\int_0^\infty(\sigma^2)^{-\alpha-2}e^{-\frac{1}{2\sigma^2}\left((x-\mu)^2+n_0(\mu-\mu_0)^2+2\beta\right)}\mathrm{d}\sigma^2
\end{aligned}
$$

$$
\frac{\beta^\alpha\sqrt{n_0}}{\Gamma(\alpha)2\pi}(\sigma^2)^{-\alpha-2}e^{-\frac{1}{2\sigma^2}\left((x-\mu)^2+n_0(\mu-\mu_0)^2+2\beta\right)}
$$

Problem 4

Problem 4.a

Denote $p = p_1$ and $q = p_2$ and denote

$$
\begin{aligned}
f_{X,Y}(0,0) &= p_{00} \\
f_{X,Y}(0,1) &= p_{01} \\
f_{X,Y}(1,0) &= p_{10} \\
f_{X,Y}(1,1) &= p_{11}.
\end{aligned}
$$

So we must have

$$
\begin{aligned}
p_{00}+p_{01}+p_{10}+p_{11} &= 1 \\
p_{01}+p_{11} &= q \\
p_{10}+p_{11} &= p
\end{aligned}
$$

In particular, as soon as we choose what p_{11} is, the rest of the probabilities are immediately determined. Now note that $p_{11} \leq \min(p, q)$. Indeed, we have $p_{11} \leq p_{10} + p_{11} = p$ and $p_{11} \leq p_{01} + p_{11} = q$. Also, equality can be acheived with (assuming without loss of generality that $p \leq q$) then $p_{11} = p$, $p_{10} = 0$, $p_{00} = 1 - q$, and $p_{01} = q - p$. A similar argument shows that we can have $p_{11} = 0$ (but not $\leq$). Thus

$$0 \leq p_{11} \leq \min(p, q).$$

With this understood, we have

$$\begin{aligned}
\rho &= \frac{\mathrm{Cov}(X, Y)}{\sigma_X \sigma_Y} \\
&= \frac{p_{11}^2 - pq}{\sqrt{p(1-p)q(1-q)}}
\end{aligned}$$

In particular, ρ is a quadratic function p_{11}, which is increasing from $p_{11} = 0$ to $p_{11} = \min(p, q)$. If $p_{11} = 0$, then the correlation between X and Y attains its lowest value, namely

$$-\frac{pq}{\sqrt{pq(1-p)(1-q)}},$$

and if $p_{11} = \min(p, q)$, then the correlation between X and Y attains its maximum value, namely

$$\frac{\min(p, q) - pq}{\sqrt{pq(1-p)(1-q)}}.$$

.

Problem 4.b

Without loss of generality, assume that $p \leq q$. In order for $\rho = 1$ we need

$$\begin{aligned}
1 = \frac{p - pq}{\sqrt{pq(1-p)(1-q)}} &\iff p - pq = \sqrt{pq(1-p)(1-q)} \\
&\iff p^2(1-q)^2 = pq(1-p)(1-q) \\
&\iff p(1-q) = q(1-p) \\
&\iff p - pq = q - qp \\
&\iff p = q.
\end{aligned}$$

In order for $\rho = -1$, we need

$$\begin{aligned}
-1 = \frac{-pq}{\sqrt{pq(1-p)(1-q)}} &\iff pq = \sqrt{pq(1-p)(1-q)} \\
&\iff p^2q^2 = pq(1-p)(1-q) \\
&\iff pq = (1-p)(1-q) \\
&\iff pq = 1 - p - q + pq \\
&\iff 0 = 1 - p - q \\
&\iff 1 = p + q
\end{aligned}$$

Problem 5

Observe that

$$\begin{aligned}
1 &= \mathrm{E}(1) \\
&= \mathrm{E}\left(\frac{X_1 + \cdots + X_n}{X_1 + \cdots + X_n}\right) \\
&= \mathrm{E}\left(\frac{X_1}{X_1 + \cdots + X_n}\right) + \cdots + \mathrm{E}\left(\frac{X_n}{X_1 + \cdots + X_n}\right) \\
&= n\mathrm{E}\left(\frac{X_1}{X_1 + \cdots + X_n}\right)
\end{aligned}$$

implies

$$\mathrm{E}\left(\frac{X_1}{X_1+\cdots+X_n}\right) = \frac{1}{n}.$$

Problem 6

By Markov's inequality, we have

$$\begin{aligned}\mathrm{pr}(X \neq 0) &= \mathrm{pr}(X \geq 1) \\ &\leq \mathrm{E}(X).\end{aligned}$$

Also Jensen's inequality tells us that

$$\mathrm{E}(X|X \neq 0)^2 \leq \mathrm{E}(X^2|X \neq 0).$$

Thus

$$\begin{aligned}\mathrm{E}(X|X \neq 0)^2 &= \left(\sum_{n=0}^{\infty} n\mathrm{pr}(X=n|X \neq 0)\right)^2 \\ &= \left(\sum_{n=0}^{\infty} n\frac{\mathrm{pr}(X=n)}{\mathrm{pr}(X \neq 0)}\right)^2 \\ &= \frac{\mathrm{E}(X)^2}{\mathrm{pr}(X \neq 0)^2}.\end{aligned}$$

Similarly,

$$\begin{aligned}\mathrm{E}(X^2|X \neq 0) &= \sum_{n=0}^{\infty} n^2\mathrm{pr}(X=n|X \neq 0) \\ &= \sum_{n=1}^{\infty} n^2\frac{\mathrm{pr}(X=n)}{\mathrm{pr}(X \neq 0)} \\ &= \frac{\mathrm{E}(X^2)}{\mathrm{pr}(X \neq 0)}.\end{aligned}$$

Therefore we have

$$\frac{\mathrm{E}(X)^2}{\mathrm{E}(X^2)} \leq \mathrm{pr}(X \neq 0).$$

Problem 7

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}\det(\Sigma)^{1/2}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}$$

Problem 7.a

We have

$$
\begin{aligned}
M_{\mathbf{X}}(\mathbf{t}) &= \mathrm{E}\left(e^{\mathbf{t}^\top \mathbf{X}}\right) \\
&= \int_{\mathbb{R}^n} e^{\mathbf{t}^\top \mathbf{x}} \frac{1}{(2\pi)^{n/2}\det(\Sigma)^{1/2}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})} \mathrm{d}\mathbf{x} \\
&= \int_{\mathbb{R}^n} \frac{1}{(2\pi)^{n/2}\det(\Sigma)^{1/2}} e^{\mathbf{t}^\top \mathbf{x}-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})} \mathrm{d}\mathbf{x} \\
&= e^{\mathbf{t}^\top \boldsymbol{\mu}} \frac{1}{(2\pi)^{n/2}\det(\Sigma)^{1/2}} \int_{\mathbb{R}^n} e^{\mathbf{t}^\top \mathbf{u}-\frac{1}{2}\mathbf{u}^\top \boldsymbol{\Sigma}^{-1}\mathbf{u}} \mathrm{d}\mathbf{u} && \text{change variable } \mathbf{u} = \mathbf{x}-\boldsymbol{\mu} \\
&= e^{\mathbf{t}^\top \boldsymbol{\mu}} \frac{1}{(2\pi)^{n/2}\det(\Sigma)^{1/2}} \int_{\mathbb{R}^n} e^{-\frac{1}{2}\left(\mathbf{u}^\top \boldsymbol{\Sigma}^{-1}\mathbf{u}-2\mathbf{t}^\top \mathbf{u}\right)} \mathrm{d}\mathbf{u} \\
&= e^{\mathbf{t}^\top \boldsymbol{\mu}} \frac{1}{(2\pi)^{n/2}\det(\Sigma)^{1/2}} \int_{\mathbb{R}^n} e^{-\frac{1}{2}\left((\mathbf{u}-\mathbf{t})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{u}-\mathbf{t})-\mathbf{t}^\top \boldsymbol{\Sigma}\mathbf{t}\right)} \mathrm{d}\mathbf{u} && \text{complete the square} \\
&= e^{\mathbf{t}^\top \boldsymbol{\mu}+\frac{1}{2}\mathbf{t}^\top \boldsymbol{\Sigma}\mathbf{t}} \frac{1}{(2\pi)^{n/2}\det(\Sigma)^{1/2}} \int_{\mathbb{R}^n} e^{-\frac{1}{2}(\mathbf{u}-\mathbf{t})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{u}-\mathbf{t})} \mathrm{d}\mathbf{u} \\
&= e^{\mathbf{t}^\top \boldsymbol{\mu}+\frac{1}{2}\mathbf{t}^\top \boldsymbol{\Sigma}\mathbf{t}} \frac{1}{(2\pi)^{n/2}\det(\Sigma)^{1/2}} \int_{\mathbb{R}^n} e^{-\frac{1}{2}\mathbf{v}^\top \boldsymbol{\Sigma}^{-1}\mathbf{v}} \mathrm{d}\mathbf{v} && \text{change variable } \mathbf{v} = \mathbf{u}-\mathbf{t} \\
&= e^{\mathbf{t}^\top \boldsymbol{\mu}+\frac{1}{2}\mathbf{t}^\top \boldsymbol{\Sigma}\mathbf{t}} \frac{1}{(2\pi)^{n/2}\det(\Sigma)^{1/2}} \int_{\mathbb{R}^n} e^{-\frac{1}{2}\mathbf{v}^\top \mathbf{P}^\top \mathbf{P}\boldsymbol{\Sigma}^{-1}\mathbf{P}^\top \mathbf{P}\mathbf{v}} \mathrm{d}\mathbf{v} && \text{orthogonal diagonalization } \mathbf{P}^\top \boldsymbol{\Sigma}^{-1}\mathbf{P} = \boldsymbol{\Lambda} \\
&= e^{\mathbf{t}^\top \boldsymbol{\mu}+\frac{1}{2}\mathbf{t}^\top \boldsymbol{\Sigma}\mathbf{t}} \frac{1}{(2\pi)^{n/2}\det(\Sigma)^{1/2}} \det(\Sigma)^{1/2} \int_{\mathbb{R}^n} e^{-\frac{1}{2}\mathbf{w}^\top \boldsymbol{\Lambda}\mathbf{w}} \mathrm{d}\mathbf{w} && \text{change variable } \mathbf{w} = \mathbf{P}\mathbf{v} \\
&= e^{\mathbf{t}^\top \boldsymbol{\mu}+\frac{1}{2}\mathbf{t}^\top \boldsymbol{\Sigma}\mathbf{t}} \frac{1}{(2\pi)^{n/2}\det(\Sigma)^{1/2}} \int_{\mathbb{R}^n} e^{-\frac{1}{2}(\lambda_1 w_1^2+\cdots+\lambda_n w_n^2)} \mathrm{d}w_n \cdots \mathrm{d}w_1 \\
&= e^{\mathbf{t}^\top \boldsymbol{\mu}+\frac{1}{2}\mathbf{t}^\top \boldsymbol{\Sigma}\mathbf{t}} \frac{1}{(2\pi)^{n/2}\det(\Sigma)^{1/2}} \sqrt{\lambda_1 \cdots \lambda_n} \int_{\mathbb{R}^n} e^{-\frac{1}{2}(z_1^2+\cdots+z_n^2)} \mathrm{d}z_n \cdots \mathrm{d}z_1 && \text{change variable } z_i = \frac{w_i}{\sqrt{\lambda_i}} \\
&= e^{\mathbf{t}^\top \boldsymbol{\mu}+\frac{1}{2}\mathbf{t}^\top \boldsymbol{\Sigma}\mathbf{t}} \frac{1}{(2\pi)^{n/2}\det(\Sigma)^{1/2}} \det(\Sigma)^{1/2} \int_{\mathbb{R}^n} e^{-\frac{1}{2}(z_1^2+\cdots+z_n^2)} \mathrm{d}z_n \cdots \mathrm{d}z_1 \\
&= e^{\mathbf{t}^\top \boldsymbol{\mu}+\frac{1}{2}\mathbf{t}^\top \boldsymbol{\Sigma}\mathbf{t}} \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-\frac{1}{2}(z_1^2+\cdots+z_n^2)} \mathrm{d}z_n \cdots \mathrm{d}z_1 \\
&= e^{\mathbf{t}^\top \boldsymbol{\mu}+\frac{1}{2}\mathbf{t}^\top \boldsymbol{\Sigma}\mathbf{t}} \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}} e^{-\frac{1}{2}z_1^2} \mathrm{d}z_1 \int_{\mathbb{R}} e^{-\frac{1}{2}z_2^2} \mathrm{d}z_2 \cdots \int_{\mathbb{R}} e^{-\frac{1}{2}z_n^2} \mathrm{d}z_n \\
&= e^{\mathbf{t}^\top \boldsymbol{\mu}+\frac{1}{2}\mathbf{t}^\top \boldsymbol{\Sigma}\mathbf{t}} \frac{1}{(2\pi)^{n/2}} (2\pi)^{n/2} \\
&= e^{\mathbf{t}^\top \boldsymbol{\mu}+\frac{1}{2}\mathbf{t}^\top \boldsymbol{\Sigma}\mathbf{t}} && \text{yay!}
\end{aligned}
$$

where we were able to complete the square and apply the real spectral theorem because the matrix $\boldsymbol{\Sigma}$ is symmetric. Indeed, for $i \neq j$ we have

$$
\begin{aligned}
\mathrm{Cov}(X_i, X_j) &= \mathrm{E}(X_i X_j) - \mathrm{E}(X_i)\mathrm{E}(X_j) \\
&= \mathrm{E}(X_j X_i) - \mathrm{E}(X_j)\mathrm{E}(X_i) \\
&= \mathrm{Cov}(X_j, X_i).
\end{aligned}
$$

Problem 7.b

Note that

$$
\begin{aligned}
M_{\mathbf{Y}}(\mathbf{t}) &= \mathrm{E}\left(e^{\mathbf{t}^\top \mathbf{Y}}\right) \\
&= \mathrm{E}\left(e^{\mathbf{t}^\top (\mathbf{AX}+\mathbf{b})}\right) \\
&= e^{\mathbf{t}^\top \mathbf{b}}\mathrm{E}\left(e^{\mathbf{t}^\top \mathbf{AX}}\right) \\
&= e^{\mathbf{t}^\top \mathbf{b}}\mathrm{E}\left(e^{(\mathbf{A}^\top \mathbf{t})^\top X}\right) \\
&= e^{\mathbf{t}^\top \mathbf{b}} M_{\mathbf{X}}(\mathbf{A}^\top \mathbf{t}) \\
&= e^{\mathbf{t}^\top \mathbf{b}} e^{(\mathbf{A}^\top \mathbf{t})^\top \boldsymbol{\mu} + \frac{1}{2}(\mathbf{A}^\top \mathbf{t})^\top \boldsymbol{\Sigma}(\mathbf{A}^\top \mathbf{t})} \\
&= e^{\mathbf{t}^\top (\mathbf{b}+\mathbf{A}\boldsymbol{\mu}) + \frac{1}{2}\mathbf{t}^\top \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^\top \mathbf{t}}.
\end{aligned}
$$

Thus $\mathbf{Y} \sim \mathrm{MVN}(\mathbf{b} + \mathbf{A}\boldsymbol{\mu}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^\top)$.

Problem 7.c

Let $\mathbf{A}$ be the $p \times p$ matrix with with 1's for its first p diagonal entries and 0's everywhere else. Then by problem 7.b, we have

$$\mathbf{X}_1 = \mathbf{AX} \sim \mathrm{MVN}(\mathbf{A}\boldsymbol{\mu}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^\top) = \mathrm{MVN}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$$

where $\boldsymbol{\mu}_1 = (\mu_1, \dots, \mu_p, 0 \dots, 0)^\top$ and

$$\boldsymbol{\Sigma}_\mathbf{1} = \sum_{1 \le i,j \le p} \sigma_{ij}\mathbf{E}_{ij},$$

where $\mathbf{E}_{ij}$ denotes the elementary matrix with 1 in the i,j entry and 0 everywhere else and where $\boldsymbol{\Sigma} = (\sigma_{ij})$. Similarly, the distribution of $\mathbf{X}_2$ is $\mathrm{MVN}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$ where $\boldsymbol{\mu}_1 = (0, \dots, 0, \mu_{p+1}, \dots, \mu_n)^\top$ and where

$$\boldsymbol{\Sigma}_\mathbf{2} = \sum_{p+1 \le i,j \le n} \sigma_{ij}\mathbf{E}_{ij}.$$

The conditional distribution of $\mathbf{X}_2$ given $\mathbf{X}_1$ is

$$
\begin{aligned}
f(\mathbf{x}_2|\mathbf{x}_1) &= f(x_{p+1}, \dots, x_n | x_1, \dots, x_p) \\
&= \frac{f(x_1, \dots, x_n)}{f(x_1, \dots, x_p)} \\
&= \frac{f(\mathbf{x})}{f(\mathbf{x}_1)} \\
&= \frac{1}{(2\pi)^{n/2}\det(\Sigma)^{1/2}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})} (2\pi)^{n/2}\det(\Sigma_1)^{1/2} e^{\frac{1}{2}(\mathbf{x}_1-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x}_1-\boldsymbol{\mu})} \\
&= \frac{\det(\Sigma_1)^{1/2}}{\det(\Sigma)^{1/2}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}) + \frac{1}{2}(\mathbf{x}_1-\boldsymbol{\mu}_1)^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x}_1-\boldsymbol{\mu}_1)} \\
&= \frac{\det(\Sigma_1)^{1/2}}{\det(\Sigma)^{1/2}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}) + \frac{1}{2}(\mathbf{x}_1-\boldsymbol{\mu}_1)^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x}_1-\boldsymbol{\mu}_1)} \\
&= \frac{\det(\Sigma_1)^{1/2}}{\det(\Sigma)^{1/2}} e^{-\frac{1}{2}\left((\mathbf{x}-\boldsymbol{\mu}-\mathbf{x}_1+\boldsymbol{\mu}_1)^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}-\mathbf{x}_1+\boldsymbol{\mu}_1)\right)} \\
&= \frac{\det(\Sigma_1)^{1/2}}{\det(\Sigma)^{1/2}} e^{-\frac{1}{2}\left(((\mathbf{x}-\mathbf{x}_1)-\boldsymbol{\mu}_2)^\top \boldsymbol{\Sigma}^{-1}((\mathbf{x}-\mathbf{x}_1)-\boldsymbol{\mu}_2)\right)}.
\end{aligned}
$$

Problem 7.d

Observe that the vectors $(a_1, \dots, a_n)^\top$ and $(b_1, \dots, b_n)^\top$ are *linearly independent* since they are orthongonal with respect to each other. In particular, we can an orthogonal $n \times n$ matrix $\mathbf{A}$ whose first two rows correspond to

$(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$ respectively. Denote $\mathbf{Y} = \mathbf{AX}$. Then $\mathbf{Y}$ has an MVN$(\mathbf{A}\boldsymbol{\mu}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^\top)$ distribution (with its first two components being Y_1 and Y_2), and since $\boldsymbol{\Sigma}$ is just a diagonal matrix, we have

$$\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^\top = \mathbf{A}\mathbf{A}^\top\boldsymbol{\Sigma} = \Sigma.$$

In particular, Y_1 and Y_2 are independent.