Ahmad Zareei
Econometrics 240A
Homework 1 - Part II

1 Multivariate normal distribution

Let $\mathbf{Y} = (Y_1, ..., Y_K)'$ be a $K \times 1$ random vector with density function as

$$f(y_1, ..., y_K) = (2\pi)^{-K/2}|\Sigma|^{-1/2} \exp\left(\frac{-1}{2}(\mathbf{y} - \mu)'\Sigma^{-1}(\mathbf{y} - \mu)\right)$$

for Σ a symmetric positive definite $K \times K$ matrix and μ a $K \times 1$ vector. Then we say that multivariable random vector $\mathbf{Y}$ has a normal random distribution with mean μ and covariance Σ or $\mathbf{Y} \sim \mathcal{N}(\mu, \Sigma)$. Prove the following properties of the multivariate normal distribution:

- **1**. For
$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}, \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \tag{1}$$
with $\Sigma_{12} = \Sigma_{21}$, show that $Y_1 \sim \mathcal{N}(\mu_1, \Sigma_{11})$?

  Answer: We first rewrite the density function as
$$f(y_1, y_2) = \frac{1}{2\pi\sqrt{|\Sigma|}} \exp\left(\frac{-1}{2}(y_1 - \mu_1, y_2 - \mu_2)\begin{pmatrix} \sigma_1 & \sigma_{12} \\ \sigma_{12} & \sigma_2 \end{pmatrix}\begin{pmatrix} y_1 - \mu_1 \\ y_2 - \mu_2 \end{pmatrix}\right)$$
  where
$$\Sigma^{-1} = \begin{pmatrix} \sigma_1 & \sigma_{12} \\ \sigma_{12} & \sigma_2 \end{pmatrix} = \frac{1}{|\Sigma|}\begin{pmatrix} \Sigma_{22} & -\Sigma_{12} \\ -\Sigma_{12} & \Sigma_{11} \end{pmatrix}$$
  Therefore
$$\begin{aligned}
f(y_1, y_2) &= \frac{1}{2\pi\sqrt{|\Sigma|}} \exp\left(\frac{-1}{2}(y_1 - \mu_1, y_2 - \mu_2)\begin{pmatrix} \sigma_1(y_1 - \mu_1) + \sigma_{12}(y_2 - \mu_2) \\ \sigma_{12}(y_1 - \mu_1) + \sigma_2(y_2 - \mu_2) \end{pmatrix}\right) \\
&= \frac{1}{2\pi\sqrt{|\Sigma|}} \exp\left(\frac{-1}{2}\left(\sigma_1(y_1 - \mu_1)^2 + 2\sigma_{12}(y_1 - \mu_1)(y_2 - \mu_2) + \sigma_2(y_2 - \mu_2)^2\right)\right) \\
&= \frac{1}{2\pi\sqrt{|\Sigma|}} \exp\left(\frac{-1}{2}\left[\sigma_1(y_1 - \mu_1)^2 + \sigma_2\left(y_2 - \mu_2 + \frac{\sigma_{12}}{\sigma_2}(y_1 - \mu_1)\right)^2 - \frac{\sigma_{12}^2}{\sigma_2}(y_1 - \mu_1)^2\right]\right) \\
&= \frac{1}{2\pi\sqrt{|\Sigma|}} \exp\left((\frac{-1}{2}\left(\left(\sigma_1 - \frac{\sigma_{12}^2}{\sigma_2}\right)(y_1 - \mu_1)^2 + \sigma_2\left(y_2 - \mu_2 + \frac{\sigma_{12}}{\sigma_2}(y_1 - \mu_1)\right)^2\right]\right)
\end{aligned}$$

Now, if we find the marginal distribution, we have

$$\begin{aligned}
f(y_1) &= \int f(y_1, y_2) dy_2 \\
&= \int_{\mathcal{R}} \frac{1}{2\pi\sqrt{|\Sigma|}} \exp\left(\frac{-1}{2}\left[\left(\sigma_1 - \frac{\sigma_{12}^2}{\sigma_2}\right)(y_1-\mu_1)^2 + \sigma_2\left(y_2 - \mu_2 + \frac{\sigma_{12}}{\sigma_2}(y_1-\mu_1)\right)^2\right]\right) dy_2 \\
&= \frac{1}{2\pi\sqrt{|\Sigma|}} \exp\left(\frac{\sigma_1 - \frac{\sigma_{12}^2}{\sigma_2}}{-2}(y_1-\mu_1)^2\right) \int_{\mathcal{R}} \exp\left(\frac{-\sigma_2}{2}(y_2 - a)^2\right) dy_2
\end{aligned}$$

where $a = \mu_2 - \sigma_{12}/\sigma_2(y_1 - \mu_1)$. We use the fact that $\int_{\mathcal{R}} \exp(-\beta x^2)dx = \sqrt{2\pi\beta}$, and we replace for σ_1, σ_2, and σ_{12} as

$$\begin{aligned}
f(y_1) &= \frac{1}{2\pi\sqrt{|\Sigma|}} \exp\left(\frac{\sigma_1 - \frac{\sigma_{12}^2}{\sigma_2}}{-2}(y_1-\mu_1)^2\right)\sqrt{2\pi\sigma_2} \\
&= \frac{1}{2\pi\sqrt{|\Sigma|}}\sqrt{2\pi\sigma_2} \exp\left(\frac{\sigma_1\sigma_2 - \sigma_{12}^2}{-2s_2}(y_1-\mu_1)^2\right) \\
&= \frac{1}{\sqrt{2\pi}\sqrt{|\Sigma|}} \frac{\sqrt{|\Sigma|}}{\sqrt{\Sigma_{11}}} \exp\left(\frac{1/|\Sigma|}{-2\Sigma_{11}/|\Sigma|}(y_1-\mu_1)^2\right) \\
&= \frac{1}{\sqrt{2\pi}\sqrt{\Sigma_{11}}} \exp\left(\frac{(y_1-\mu_1)^2}{-2\Sigma_{11}}\right) \sim \mathcal{N}(\mu_1, \Sigma_{11})
\end{aligned}$$

Therefore $Y_1 \sim \mathcal{N}(\mu_1, \Sigma_{11})$.

- **2**. Likewise show that

$$Y_2 | Y_1 = y_1 \sim \mathcal{N}(\mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(y_1 - \mu_1), \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12})$$

Answer: We know that

$$f(Y_2 = y_2 | Y_1 = y_1) = \frac{f(Y_1 = y_1, Y_2 = y_2)}{f(Y_1 = y_1)}$$

Using the equations we derived in part 1, we have

$$\begin{aligned}
f(y_1, y_2) &= \frac{1}{2\pi\sqrt{|\Sigma|}} \exp\left(\frac{-1}{2}\left[\left(\sigma_1 - \frac{\sigma_{12}^2}{\sigma_2}\right)(y_1-\mu_1)^2 + \sigma_2\left(y_2 - \mu_2 + \frac{\sigma_{12}}{\sigma_2}(y_1-\mu_1)\right)^2\right]\right) \\
&= \frac{1}{2\pi\sqrt{|\Sigma|}} \exp\left(\frac{\sigma_1\sigma_2 - \sigma_{12}^2}{-2\sigma_2}(y_1-\mu_1)^2 + \frac{\sigma_2}{-2}\left(y_2 - \mu_2 + \frac{\sigma_{12}}{\sigma_2}(y_1-\mu_1)\right)^2\right) \\
&= \frac{1}{2\pi\sqrt{|\Sigma|}} \exp\left(\frac{(y_1-\mu_1)^2}{-2\Sigma_{11}} + \frac{\Sigma_{11}}{-2|\Sigma|}\left(y_2 - \mu_2 + \frac{-\Sigma_{12}}{\Sigma_{11}}(y_1-\mu_1)\right)^2\right)
\end{aligned}$$

We also know that

$$f(Y_1 = y_1) = \frac{1}{\sqrt{2\pi}\sqrt{\Sigma_{11}}} \exp\left((\frac{(y_1-\mu_1)^2}{-2\Sigma_{11}} \right) \tag{2}$$

Therefore

$$\begin{aligned}
f_{Y_2|Y_1} &= \frac{1}{2\pi}\frac{\sqrt{\Sigma_{11}}}{\sqrt{|\Sigma|}} \exp\left(\frac{\Sigma_{11}}{-2|\Sigma|} \left(y_2 - \mu_2 + \frac{-\Sigma_{12}}{\Sigma_{11}}(y_1-\mu_1) \right)^2 \right) \\
&\sim \mathcal{N}(\mu_2 + \frac{\Sigma_{12}}{\Sigma_{11}}(y_1-\mu_1), \frac{|\Sigma|}{\Sigma_{11}}) \\
&\sim \mathcal{N}(\mu_2 + \Sigma_{12}\Sigma_{11}^{-1}(y_1-\mu_1), \frac{\Sigma_{22}\Sigma_{11} - \Sigma_{12}\Sigma_{21}}{\Sigma_{11}}) \\
&\sim \mathcal{N}(\mu_2 + \Sigma_{12}\Sigma_{11}^{-1}(y_1-\mu_1), \Sigma_{22} - \Sigma_{12}\Sigma_{21}\Sigma_{11}^{-1})
\end{aligned}$$

- **3**. If $\Sigma_{12} = 0$, Show that Y_1 and Y_2 are independent?

  Answer: If $\Sigma_{12} = 0$, we first define

$$\Sigma^{-1} = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} = \begin{pmatrix} 1/\Sigma_{11} & 0 \\ 0 & 1/\Sigma_{22} \end{pmatrix}$$

  therefore the joint density function can be writen as

$$\begin{aligned}
f(y_1, y_2) &= \frac{1}{2\pi\sqrt{|\Sigma|}} \exp\left(\frac{-1}{2} (y_1-\mu_1, y_2-\mu_2) \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} \begin{pmatrix} y_1-\mu_1 \\ y_2-\mu_2 \end{pmatrix} \right) \\
&= \frac{1}{2\pi\sqrt{\Sigma_{11}\Sigma_{22}}} \exp\left(\frac{(y_1-\mu_1)^2}{-2\Sigma_{11}} + \frac{(y_2-\mu_2)^2}{-2\Sigma_{22}} \right) \\
&= \left[\frac{1}{\sqrt{2\pi\Sigma_{11}}} \exp\left(\frac{(y_1-\mu_1)^2}{-2\Sigma_{11}} \right) \right] \left[\frac{1}{\sqrt{2\pi\Sigma_{22}}} \exp\left(\frac{(y_2-\mu_2)^2}{-2\Sigma_{22}} \right) \right] \\
&= f(y_1)f(y_2)
\end{aligned}$$

  since the joint density function, the random variables Y_1 and Y_2 are independent.

- **4**. Let $Z = A + BY$, where A and B are non-random and B has a full rank. Show that

$$Z \sim \mathcal{N}(A + B\mu, B\Sigma B') \tag{3}$$

  Answer: As for change of variables, we know that

$$f_{Z_1,Z_2}(z_1, z_2) = f_{Y_1,Y_2}(y_1(z_1,z_2), y_2(z_1,z_2))|\frac{\partial(Y_1,Y_2)}{\partial(Z_1,Z_2)}|$$

since $Z = A + BY$, therefore $Y = B^{-1}(Z - A)$, and therefore $|\partial(Y_1, Y_2)/\partial(Z_1, Z_2)| = 1/|B|$. Inserting all of these into f_{Y_1,Y_2} distribution, we obtain:

$$\begin{aligned}
f_{Z_1,Z_2}(z_1, z_2) &= \frac{1}{2\pi\sqrt{|\Sigma|}} \exp\left(\frac{-1}{2}(B^{-1}(\mathbf{z} - A - B\mu))'\Sigma^{-1}(B^{-1}(\mathbf{z} - A - B\mu))\right)\frac{1}{|B|} \\
&= \frac{1}{2\pi|B|\sqrt{|\Sigma|}} \exp\left(\frac{-1}{2}(\mathbf{z} - A - B\mu)'(B\Sigma B')^{-1}(\mathbf{z} - A - B\mu)\right) \\
&\sim \mathcal{N}(A + B\mu, B\Sigma B')
\end{aligned}$$

- **5**. Set $A = 0$ and $B = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$. Show that for $Z = A + BY = \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix}$,
  - Z_1 and Z_2 are univariate normal random variables.

    Answer: $Z_1 = Y_1 - Y_2$ and $Z_2 = -Y_1 + Y_2$, therefore $Z_2 = -Z_1 = Y_1 - Y_2$. Therefore Z_1 and Z_2 are not independent and therefore Z_1 and Z_2 are univariate random variable. This happened since B is not a full rank matrix ($|B| = 0$).
  - their joint distribution is not bivariate normal? Explain.

    Since B is not a full rank matrix, then the inverse is not defined. Therefore, we cannot insert it into the previous part that we derived here, and the distribution will therefore becomes other than bivariate random variable. In fact it will become a univariate distribution, since both random variables are just negative of each other!

- Let $\{Y_i\}_{i=1}^N$ be a random sample of size N drawn from the multivariate normal population described above. Show that $\sqrt{N}(\bar{Y} - \mu) \sim \mathcal{N}(0, \Sigma)$ for $\bar{Y} = 1/N \sum_{i=1}^N Y_i$ the sample mean?

  Answer: Since $Y_1, ..., Y_N$ draws are independent, then Σ is diagonal. . We choose $A = -\mu$, and a full rank matrix B such that

$$B = \begin{pmatrix} 1/N & 1/N & ... & 1/N & 1/N \\ 0 & 1 & ... & 0 & 0 \\ 0 & 0 & ... & 0 & 0 \\ ... & & & & \\ 0 & 0 & ... & 0 & 1 \end{pmatrix}$$

and now, we look at Z_1, we have

$$\begin{gathered}
Z = A + BY \to Z_1 = \frac{1}{N}\sum_{i=1}^{N} Y_i - \mu \\
Z_2 = Y_2 \\
... \\
Z_N = Y_N
\end{gathered}$$

Therefore,

$$\begin{aligned}
&Z_1 \sim \mathcal{N}(-\mu + N(1/N)\mu, N(\Sigma/N^2)\\
&\bar{Y} - \mu \sim \mathcal{N}(0, \Sigma/N)\\
&\sqrt{N}(\bar{Y} - \mu) \sim \mathcal{N}(0, \Sigma)
\end{aligned}$$

- Let $W = (Y - \mu)'\Sigma^{-1}(Y - \mu)$. Show that $W \sim \chi^2_K$?

  Answer: We only need to show that $w = (y - \mu)'\Sigma^{-1}(y - \mu)$, is actually $w = z'z$, where $z \sim \mathcal{N}(0, 1)$. Since Σ is positive definite, we can write it as $\Sigma = C\Lambda C'$, where C is an orthonormal matrix (i.e. $CC' = I = C'C$) and Λ is a diagonal matrix. We define Λ^* to be a a diagonal matrix whose diagonal elements are the reciprocal square roots of the corresponding diagonal element of Λ. Let $H = C\Lambda^*C'$. Then $H' = (C\Lambda^*C')' = C\Lambda^*C' = H$ and also $H'H = C\Lambda^*C'C\Lambda^*C = C(\Lambda^*)^2C' = C\Lambda^{-1}C' = \Sigma^{-1}$. Therefore $H\Sigma H' = I$. Let $\epsilon = y - \mu$, then $\epsilon \sim \mathcal{N}(0, \Sigma)$. Then Let $z = H\epsilon$, therefore $z \sim \mathcal{N}(0, I)$. As a result we have $w = \epsilon'\Sigma^{-1}\epsilon = \epsilon'H'H\epsilon = (H\epsilon)'(H\epsilon) = z'z$.

- Let $W = N(\bar{Y} - \mu)'\Sigma^{-1}(\bar{Y} - \mu)$. Show that $W \sim \chi^2_K$ (i.e. W is a chi-square random variable with K degree of freedom).

  Answer: Again, we only need to show that $w = \sqrt{N}(\bar{y} - \mu)'\Sigma^{-1}\sqrt{N}(\bar{y} - \mu)$, is actually $w = z'z$, where $z \sim \mathcal{N}(0, 1)$. Since Σ is positive definite, we can write it as $\Sigma = C\Lambda C'$, where C is an orthonormal matrix (i.e. $CC' = I = C'C$) and Λ is a diagonal matrix. We define Λ^* to be a a diagonal matrix whose diagonal elements are the reciprocal square roots of the corresponding diagonal element of Λ. Let $H = C\Lambda^*C'$. Then $H' = (C\Lambda^*C')' = C\Lambda^*C' = H$ and also $H'H = C\Lambda^*C'C\Lambda^*C = C(\Lambda^*)^2C' = C\Lambda^{-1}C' = \Sigma^{-1}$. Therefore $H\Sigma H' = I$. Let $\epsilon = \sqrt{N}\,(\bar{y} - \mu)$, then $\epsilon \sim \mathcal{N}(0, \Sigma)$ (we showed this in previous parts). Then Let $z = H\epsilon$, therefore $z \sim \mathcal{N}(0, I)$. As a result we have $w = \epsilon'\Sigma^{-1}\epsilon = \epsilon'H'H\epsilon = (H\epsilon)'(H\epsilon) = z'z$.

- Let $\chi^{2,1-\alpha}_K$ be the $(1-\alpha)^{th}$ quantile of the χ^2_K distribution i.e. $\Pr(W \leq \chi^{2,1-\alpha}_K) = 1-\alpha$. Let D be a $P \times K$ ($P \leq K$) matrix of rank P and d be a $P \times 1$ vector of constants. Consider the hypothesis

  $$\begin{aligned}
  H_0 &: D\mu = d\\
  H_1 &: D\mu \neq d
  \end{aligned}$$

  Maintaining H_0 derive the sampling distribution of $D\bar{Y}$ as well as that of

  $$W = N.(D\bar{Y} - d)'(D\Sigma D)^{-1}(D\bar{Y} - d)$$

  You observe that, for the sample in hand, $W > \chi^{2,1-\alpha}_P$ for $\alpha = 0.05$. Assuming H_0 is true, what is the ex ante (i.e. pre-sample) probability of this event? What are you inclined to conclude after observing W in the sample in hand?

  Answer: From previous parts we know that the distribution of $\bar{Y} \sim \mathcal{N}(\mu, \Sigma/N)$. Then the distribution of $D\bar{Y} - d \sim \mathcal{N}(D\mu - d, \frac{1}{N}D\Sigma D')$, or $\sqrt{N}(D\bar{Y} - d) \sim \mathcal{N}(D\mu - d, D\Sigma D')$ if we assume H_0 is true. From part 8, $W \sim \chi^2_P$, therefore $\Pr(W > \chi^{2,1-\alpha}_K) = 1 - (1 - \alpha) = 5\%$ Therefore the ex-ante probability of rejecting H_0 is 5%. Therefore, if we reject based on the fact that $W > \chi^{2,1-\alpha}_K$, the probability of being wrong is 5%.

2 Exercises

2.1 Problem 1

[Adapted from exercise 18.1 from Goldberger (1991)]. Suppose that $\mathbf{Y} \sim \mathcal{N}(\mu, \Sigma)$ with

$$\mu = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \Sigma = \begin{pmatrix} 2 & -1 & 1 \\ -1 & 5 & 1 \\ 1 & 1 & 3 \end{pmatrix} \tag{4}$$

- (a) Calculate $\mathbb{E}[Y_3|Y_1 = y_1, Y_2 = y_2]$ and $\mathrm{var}(Y_3|Y_1 = y_1, Y_2 = y_2)$?

  Answer: We define $z = Ay_1 + By_2 + y_3$. We want to find A, B such that z be uncorrelated with y_1 and y_2. Therefore

$$\begin{aligned} \mathrm{cov}(z, y_1) = 0 = \mathrm{cov}(Ay_1 + By_2 + y_3, y_1) = A\mathrm{var}(y_1) + B\mathrm{cov}(y_2, y_1) + \mathrm{cov}(y_3, y_1) \\ \mathrm{cov}(z, y_2) = 0 = \mathrm{cov}(Ay_1 + By_2 + y_3, y_2) = A\mathrm{cov}(y_1, y2) + B\mathrm{var}(y_2) + \mathrm{cov}(y_3, y_2) \end{aligned}$$

  Therefore

$$\begin{pmatrix} \mathrm{var}(y_1) & \mathrm{cov}(y_2, y_1) \\ \mathrm{cov}(y_1, y_2) & \mathrm{var}(y_2) \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} -\mathrm{cov}(y_3, y_1) \\ -\mathrm{cov}(y_3, y_2) \end{pmatrix}$$

$$\begin{pmatrix} 2 & -1 \\ -1 & 5 \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \end{pmatrix} \rightarrow A = -\frac{2}{3}, B = -\frac{1}{3}$$

  Therefore $z = -\frac{2}{3}y_1 - \frac{1}{3}y_2 + y_3$ is independent of both y_1 and y_2. Therefore

$$\begin{aligned} \mathbb{E}[Y_3|Y_1 = y_1, Y_2 = y_2] &= \mathbb{E}[Z + \frac{2}{3}Y_1 + \frac{1}{3}Y_2|Y_1 = y_1, Y_2 = y_2] \\ &= \mathbb{E}[Z|y_1, y_2] + \frac{2}{3}\mathbb{E}[Y_1|y_1, y_2] + \frac{1}{3}E[Y_2|y_1, y_2] \\ &= \mathbb{E}[z] + \frac{2}{3}y_1 + \frac{1}{3}y_2 \\ &= -\frac{2}{3}(1) - \frac{1}{3}2 + 3 + \frac{2}{3}y_1 + \frac{1}{3}y_2 \\ &= \frac{5}{3} + \frac{2}{3}y_1 + \frac{1}{3}y_2 \end{aligned}$$

  and also

$$\begin{aligned} &\mathrm{var}[Y_3|Y_1 = y_1, Y_2 = y_2] = \mathrm{var}[Z] = \mathrm{var}(\frac{-2}{3}Y_1 - \frac{1}{3}Y_2 + Y_3) \\ &= \mathrm{var}(Y_3) + \frac{4}{9}\mathrm{var}(Y_1) + \frac{1}{9}\mathrm{var}(Y_2) - \frac{4}{9}\mathrm{cov}(Y_1, Y_3) - \frac{2}{3}\mathrm{cov}(Y_2, Y_3) + \frac{4}{9}\mathrm{cov}(Y_1, Y_2) \\ &= 3 + \frac{4}{9}2 + \frac{1}{9}5 - \frac{4}{9}1 - \frac{2}{3}1 + \frac{4}{9}(-1) \\ &= \frac{10}{3} \end{aligned}$$

- (b) Calculate $\mathbb{E}[Y_3|Y_1 = y_1]$ and $\text{var}(Y_3|Y_1 = y_1)$? Here we need some function of Y_3 and Y_1 as $z = Ay_1 + y_3$ to be independent of y_1. We therefore have

$$\begin{aligned}
&\text{cov}(z, y_1) = 0 = \text{cov}(Ay_1 + y_3, y_1) = A\text{var}(y_1) + \text{cov}(y_3, y_1) = 0 \\
&A = -\frac{\text{cov}(y_3, y_1)}{\text{var}(y_1)} = -\frac{1}{2}
\end{aligned}$$

therefore $Z = -\frac{1}{2}Y_1 + Y_3$ is independent of Y_1.

$$\begin{aligned}
\mathbb{E}(Y_3|Y_1 = y_1) = \mathbb{E}(Z + \frac{1}{2}Y_1|Y_1 = y_1) = \mathbb{E}(Z|Y_1 = y_1) + \frac{1}{2}\mathbb{E}[Y_1|Y_1 = y_1] \\
=\mathbb{E}(Z) + \frac{1}{2}y_1 = -\frac{1}{2}1 + 3 + \frac{1}{2}y_1 \\
=\frac{5}{2} + \frac{1}{2}y_1
\end{aligned}$$

and

$$\begin{aligned}
\text{var}(Y_3|Y_1 = y_1) = \text{var}(Z) = \text{var}(-\frac{1}{2}Y_1 + Y_3) = \frac{1}{4}\text{var}(Y_1) + \text{var}(Y_3) + 2\frac{-1}{2}\text{cov}(Y_1, Y_3) \\
= \frac{1}{4}2 + 3 - 1 = \frac{5}{2}
\end{aligned}$$

- Calculate $\Pr(-1 \leq Y_3 \leq 2)$? We know that Y_3 follows a normal distribution with $\mathcal{N}(3, 3)$, therefore this domain of $[-1, 2]$ with regard to normal distribution is as $[(-1-3)/\sqrt{3}, (2-3)/\sqrt{3}] = [-4/\sqrt{3}, -1/\sqrt{3}]$. Therefore

$$\Pr(-1 \leq Y_3 \leq 2) = \Phi^{-1}(\frac{-1}{\sqrt{3}}) - \Phi^{-1}(\frac{-4}{\sqrt{3}})$$

2.2 Problem 2

Let Y_1 and Y_0 respectively denote child and parent height. Assume that $Y_t \sim \mathcal{N}(\mu, \sigma^2)$ for $t = 0, 1$. Let $\rho = \mathcal{C}(Y_1, Y_0)/\sqrt{\text{var}(Y_1)\text{var}(Y_0)}$ be equal to the correlation between Y_0 and Y_1. Show the followings

- (a) $\mathbb{E}[Y_1|Y_0 = y_0] = (1 - \rho)\mu + \rho y_0$?

  Answer: From Part 2 of last problem, we already know that

$$Y_1|Y_0 = y_0 \sim \mathcal{N}(\mu + \Sigma_{21}\Sigma_{11}^{-1}(y_0 - \mu), \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{21}$$

  where

$$\Sigma = \begin{pmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{pmatrix}$$

  Therefore

$$\mathbb{E}[Y_1|Y_0] = \mu + \rho\sigma^2\sigma^{-2}(y_0 - \mu) = (1 - \rho)\mu + \rho y_0$$

- (b) Under what condition would you expect that a child's height to exceed that of their parent? The opposite? Why is this called a regression to mean?

  Answer: We want, $\mathbb{E}[Y_1|Y_0 = y_0] \geq y_0$. This would mean that $(1-\rho)\mu - (1-\rho)y_0 \geq$, therefore $(1-\rho)(\mu - y_0) \geq 0$. Similarly, it can be shown that the opposite is true when $(1-\rho)(\mu - y_0) \leq 0$. Since $\rho \leq 1$, then this means that, if the father's height is less than average, the child's average is expected to be larger (closer to the mean value) and if the father's height is larger than mean, it is expected to see the child's height to be shorter (again closer to mean value). Since the value of the next generation is getting closer to the mean value, this is called, regression to the mean.

- Prove that $0 \leq \rho^2 \leq 1$?

  Answer: We define a variable a $Z = tX + Y$. We know that

  $$\mathrm{var}(tX + Y) \geq 0 \rightarrow \mathrm{var}(X)t^2 + 2\mathrm{Cov}(X, Y)t + \mathrm{var}(Y)$$

  Since this should be true for all values of t, then

  $$\Delta = B^2 - 4AC \leq 0 \rightarrow \mathrm{Cov}(X, Y)^2 \leq \mathrm{var}(X)\mathrm{var}(Y) \rightarrow \rho^2 \leq 1$$

  since ρ^2 is a square, it should be positive. Therefore

  $$0 \leq \rho^2 \leq 1$$

2.3 Problem 3

Complete the following exercises from Hansen (2017): 2.8, 2.9, 2.18 (part (a) only).

- **2.8**: Suppose that y is discrete-valued, taking values only on the non-negative integers, and the conditional distribution of y given x is Poisson:

  $$\Pr(y = j|x) = \frac{\exp(-x'\beta)(x'\beta)^j}{j!}, \quad j = 0, 1, 2, ... \tag{5}$$

  Compute $\mathbb{E}(y|x)$ and $\mathrm{var}(y|x)$? Does this justify a linear regression model of $y = x'\beta + e$?

  Answer: Since the conditional probability is a Poission distribution, we have

  $$\begin{aligned}\mathbb{E}(y|x) &= \sum_{j=0}^{\infty} j\Pr(y = j|x) = \exp(-x'\beta)(x'\beta)\sum_{j=1} \frac{(x'\beta)^{j-1}}{(j-1)!} \\ &= x'\beta \exp(-x'\beta) \sum_{j=0}^{\infty} \frac{(x'\beta)^j}{j!} = x'\beta \exp(-x'\beta) \exp(x'\beta) \\ &= x'\beta\end{aligned}$$

$$\begin{aligned}
\mathrm{var}(y|x) &= \sum_{j=0}^{\infty} j^2 \Pr(y=j|x) - (\mathbb{E}(y|x))^2 = \exp(-x'\beta)(x'\beta)\sum_{j=1} j\frac{(x'\beta)^{j-1}}{(j-1)!} - (\acute{x}\beta)^2 \\
&= (x'\beta)\exp(-x'\beta)\sum_{j=0}^{\infty}(j+1)\frac{(x'\beta)^j}{j!} - (x'\beta)^2 \\
&= (x'\beta)\exp(-x'\beta)\left[\exp(x'\beta) + (x'\beta)\exp(x'\beta)\right] - (x'\beta)^2 \\
&= x'\beta + (x'\beta)^2 - (x'\beta)^2 \\
&= x'\beta
\end{aligned}$$

Yes, since the $\mathbb{E}(y|x)$ minimizes the MSE, it justifies the linear regression model of $y = x'\beta + e$.

- **2.9**. Suppose you have two regressors: x_1 is binary (takes values 0 and 1) and x_2 is categorical with 3 categories (A, B, C) Write $\mathbb{E}(y|x_1, x_2)$ a linear regression.

  Answer: We can write $E(y|x_1, x_2) = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \beta_4$. Or we can write x_2 in terms of two dummy variables as

$$x_3 = \begin{cases} 1 & \text{if } x_3 = A \\ 0 & \text{if } x_3 \neq A \end{cases}$$
$$x_4 = \begin{cases} 1 & \text{if } x_4 = C \\ 0 & \text{if } x_4 \neq C \end{cases}$$

  Therefore

$$x_2 = \begin{cases} A & \text{if } x_3 = 1, x_4 = 0 \\ B & \text{if } x_3 = 0, x_4 = 0 \\ C & \text{if } x_3 = 0, x_4 = 1 \end{cases}$$

  Then we would have $E(y|x_1, x_2) = E(y|x_1, x_3, x_4) = \beta_1 x_1 + \beta_2 x_3 + \beta_4 x_4 + \beta_5 x_1 x_3 + \beta_6 x_1 x_4$

- **2.18 (a)**. Suppose that

$$\mathbf{x} = \begin{pmatrix} 1 \\ x_2 \\ x_3 \end{pmatrix}$$

  and $x_3 = \alpha_1 + \alpha_2 x_2$. Show that $\mathbf{Q_{xx}} = \mathbb{E}(\mathbf{xx}')$ is not invertible.

  Answer: We have

$$\begin{aligned}
\mathbf{Q_{xx}} &= \begin{pmatrix} \mathbb{E}(x_1x_1') & \mathbb{E}(x_1x_2') & \mathbb{E}(x_1x_3') \\ \mathbb{E}(x_2x_1') & \mathbb{E}(x_2x_2') & \mathbb{E}(x_2x_3') \\ \mathbb{E}(x_3x_1') & \mathbb{E}(x_3x_2') & \mathbb{E}(x_3x_3') \end{pmatrix} \\
&= \begin{pmatrix} \mathbb{E}(x_1x_1') & \mathbb{E}(x_1x_2') & \mathbb{E}(x_1x_3') \\ \mathbb{E}(x_2x_1') & \mathbb{E}(x_2x_2') & \mathbb{E}(x_2x_3') \\ \alpha_1\mathbb{E}(x_1x_1') + \alpha_2\mathbb{E}(x_2x_1') & \alpha_1\mathbb{E}(x_1x_2') + \alpha_2\mathbb{E}(x_2x_2') & \alpha_1\mathbb{E}(x_1x_3') + \alpha_2\mathbb{E}(x_2x_3') \end{pmatrix}
\end{aligned}$$

where the last row is a function of first and second row. Using properties of determinant, we have

$$\begin{aligned}
\mathbf{Q_{xx}} &= \alpha_1 \begin{pmatrix} \mathbb{E}(x_1x_1') & \mathbb{E}(x_1x_2') & \mathbb{E}(x_1x_3') \\ \mathbb{E}(x_2x_1') & \mathbb{E}(x_2x_2') & \mathbb{E}(x_2x_3') \\ \mathbb{E}(x_1x_1') & \mathbb{E}(x_1x_2') & \mathbb{E}(x_1x_3') \end{pmatrix} + \alpha_2 \begin{pmatrix} \mathbb{E}(x_1x_1') & \mathbb{E}(x_1x_2') & \mathbb{E}(x_1x_3') \\ \mathbb{E}(x_2x_1') & \mathbb{E}(x_2x_2') & \mathbb{E}(x_2x_3') \\ \mathbb{E}(x_2x_1') & \mathbb{E}(x_2x_2') & \mathbb{E}(x_2x_3') \end{pmatrix} \\
|\mathbf{Q_{xx}}| &= \alpha_1.0 + \alpha_2.0 = 0
\end{aligned}$$

Therefore $\mathbf{Q_{xx}}$ is not invertible.