Exercises

Exercise 1

Let X_1 and X_2 be independent and identically random variables with density function

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\{-x^2/2\}$$

and let $X = (X_1, X_2)$. Furthermore, let $Y_1 = X_1 + X_2$, $Y_2 = X_1 + 2X_2$ and $Y = (Y_1, Y_2)$.

a) Find the expectation and variance of X

b) Find skewness and excess kurtosis coefficients of X_1 and of X_2

c) Find the density function of $Z = X_1^2$

d) Find $E(Y)$ and $V(Y)$.

e) Write the joint density of X_1 and X_2

f) Find the joint density of Y_1, Y_2. Are Y_1 and Y_2 stochastically independent?

g) Find the marginal density of Y_1

h) Find the conditional density of Y_2, given Y_1

(Hint: $\int_{-\infty}^{+\infty} \exp\{-x^2\}dx = \sqrt{\pi}$)

Solution

a)

$$E(X_1) = E(X_2) = \int_{-\infty}^{+\infty} x \frac{1}{\sqrt{2\pi}} \exp\{-x^2/2\}dx = 0$$

since the function is odd.

$$V(X_1) = V(X_2) = E(X_1^2) = \int_{-\infty}^{+\infty} x^2 \frac{1}{\sqrt{2\pi}} \exp\{-x^2/2\}dx$$

Integrating by parts, we obtain

$$\begin{aligned} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x \left(x \exp\{-x^2/2\}\right) dx &= \frac{1}{\sqrt{2\pi}} \left[x \left(-\exp\{-x^2/2\}\right)\right]_{-\infty}^{+\infty} + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp\{-x^2/2\}dx \\ &= 0 + \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} \exp\{-t^2\}dt \\ &= 1 \end{aligned}$$

Hence $V(X_1) = V(X_2) = 1$. Since X_1, X_2 are stochastically independent, their covariance is zero. Hence

$$E(X) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad V(X) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

b)

$$\varphi = \frac{E(X-\mu)^3}{\sigma^3} = E(X^3) = \int_{-\infty}^{+\infty} x^3 \frac{1}{\sqrt{2\pi}} \exp\{-x^2/2\}dx = 0$$

$$
\begin{aligned}
\kappa &= \frac{E(X_1-\mu)^4}{\sigma^4} - 3 \\
&= E(X_1^4) - 3 \\
&= \int_{-\infty}^{+\infty} x^4 \frac{1}{\sqrt{2\pi}} \exp\{-x^2/2\} dx - 3 \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^3 \left(x \exp\{-x^2/2\} \right) dx - 3 \\
&= \frac{1}{\sqrt{2\pi}} \left[-x^3 \exp\{-x^2/2\} \right]_{-\infty}^{+\infty} + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} 3x^2 \exp\{-x^2/2\} dx - 3 \\
&= 0 + \frac{3}{\sqrt{\pi}} \int_{-\infty}^{+\infty} \exp\{-t^2\} dt - 3 \\
&= 0
\end{aligned}
$$

Since X_2 has the same distribution as X_1, their skewness and excess Kurtosis coefficients are equal to zero.

c)

For $z > 0$

$$
\begin{aligned}
P(Z \le z) &= P(-\sqrt{z} \le X_1 \le \sqrt{z}) \\
&= 2 \int_0^{\sqrt{z}} \frac{1}{\sqrt{2\pi}} \exp\{-x^2/2\} dx \\
&= \int_0^{z} \frac{1}{\sqrt{2\pi}} \exp\{-t/2\} t^{-1/2} dt
\end{aligned}
$$

Hence the density function of Z is

$$
f_Z(z) = \begin{cases} 0 & z \le 0 \\ \frac{1}{\sqrt{2\pi}} \exp\{-z/2\} z^{-1/2} & z > 0 \end{cases}
$$

d)

$$E(Y_1) = E(X_1) + E(X_2) = 0$$

$$E(Y_2) = E(X_1) + 2E(X_2) = 0$$

Since X_1 and X_2 are uncorrelated,

$$V(Y_1) = V(X_1) + V(X_2) = 2$$

$$V(Y_2) = V(X_1) + 4V(X_2) = 5.$$

Furthermore

$$Cov(X_1 + X_2, X_1 + 2X_2) = Cov(X_1, X_1) + 2Cov(X_2, X_2) = 3$$

Hence

$$E(Y) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$V(Y) = \begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix}$$

e)

$$f_{X_1,X_2}(x_1, x_2) = \frac{1}{\sqrt{2\pi}} \exp\{-x_1^2/2\} \frac{1}{\sqrt{2\pi}} \exp\{-x_2^2/2\} = \frac{1}{2\pi} \exp\{-(x_1^2 + x_2^2)/2\}$$

f)

The transformation is

$$\begin{cases} y_1 = x_1 + x_2 \\ y_2 = x_1 + 2x_2 \end{cases}$$

Solving with respect to x_1 and x_2, we find

$$\begin{cases} x_1 = 2y_1 - y_2 \\ x_2 = y_2 - y_1 \end{cases}$$

The Jacobian matrix is

$$J = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$$

The determinant of the Jacobian matrix is 1. Hence, using a)

$$\begin{aligned} f_{Y_1,Y_2}(y_1, y_2) &= f_{X_1,X_2}(2y_1 - y_2, y_2 - y_1) \cdot 1 \\ &= \frac{1}{2\pi} \exp\left\{-\left((2y_1 - y_2)^2 + (y_2 - y_1)^2\right)/2\right\} \\ &= \frac{1}{2\pi} \exp\left\{-\left(5y_1^2 + 2y_2^2 - 6y_1 y_2\right)/2\right\} \end{aligned}$$

Since the density function of Y_1 and Y_2 can not be written as a product of a function in y_1 and of a function in y_2, Y_1 and Y_2 are not stochastically independent.

g)

$$\begin{aligned} f_{Y_1}(y_1) &= \int_{-\infty}^{+\infty} \frac{1}{2\pi} \exp\left\{-\left(5y_1^2 + 2y_2^2 - 6y_1 y_2\right)/2\right\} dy_2 \\ &= \frac{1}{2\pi} \exp\{-5y_1^2/2\} \int_{-\infty}^{+\infty} \exp\{-y_2^2 + 3y_1 y_2\} dy_2 \\ &= \frac{1}{2\pi} \exp\{-5y_1^2/2\} \int_{-\infty}^{+\infty} \exp\left\{-\left(y_2^2 - 3y_1 y_2 + 9y_1^2/4\right)\right\} dy_2 \exp\left\{9y_1^2/4\right\} \\ &= \frac{1}{2\pi} \exp\{-y_1^2/4\} \int_{-\infty}^{+\infty} \exp\{-(y_2 - 3y_1/2)^2\} dy_2 \\ &= \frac{1}{2\pi} \exp\{-y_1^2/4\} \int_{-\infty}^{+\infty} \exp\{-t^2\} dt \\ &= \frac{1}{2\sqrt{\pi}} \exp\{-y_1^2/4\} \end{aligned}$$

h)

$$\begin{aligned} f_{Y_2|Y_1}(y_2|y_1) &= \frac{\frac{1}{2\pi} \exp\left\{-\frac{1}{2}\left[5y_1^2 + 2y_2^2 - 6y_1 y_2\right]\right\}}{\frac{1}{2\sqrt{\pi}} \exp\{-y_1^2/4\}} \\ &= \frac{1}{\sqrt{\pi}} \exp\{-(y_2^2 - 3y_2 y_1 + 9y_1^2/4)\} \\ &= \frac{1}{\sqrt{\pi}} \exp\{-(y_2 - 3y_1/2)^2\} \end{aligned}$$

Exercise 2

Consider a data generating process $Y = X\beta + \epsilon$ with $\epsilon \sim N(0, \sigma^2 I)$ and X a deterministic matrix of dimension $n \times k$ and rank k. Let $M = X(X^T X)^{-1} X^T$ (notice that $X^T X$ is not singular since X has full rank and $X^T X$ is a $k \times k$ matrix).

a) Show that M and $I - M$ are symmetric and idempotent. Show that the rank of M is k and that the rank of $I - M$ is $n - k$ (hint: the rank of an idempotent matrix is equal to its trace).

b) Write the joint density of $Y_1, \ldots, Y_n$.

c) Show that the maximum likelihood estimator of (β, σ^2) is $\hat{\beta} = (X^T X)^{-1} X^T Y$, $\hat{\sigma}^2 = \hat{\epsilon}^T \hat{\epsilon}/n$ where $\hat{\epsilon} = Y - \hat{Y}$ with $\hat{Y} = X\hat{\beta}$. Furthermore, show that $\hat{\beta} = \beta + (X^T X)^{-1} X^T \epsilon$.

d) Show that $\hat{\beta} \sim N(\beta, \sigma^2 (X^T X)^{-1})$.

e) Show that $\hat{Y} = MY = X\beta + M\epsilon$ and that $\hat{Y} \sim N(X\beta, \sigma^2 M)$.

f) Show that $\hat{\epsilon} = (I - M)Y = (I - M)\epsilon$ and that $\hat{\epsilon} \sim N(0, \sigma^2 (I - M))$.

g) Show that $\hat{\epsilon}$ is independent from $\hat{\beta}$ and from $\hat{Y}$.

h) Show that $\hat{\epsilon}^T \hat{\epsilon}/\sigma^2 \sim \chi^2(n - k)$.

i) Assuming $\beta = 0$, show that $\hat{Y}^T \hat{Y}/\sigma^2 \sim \chi^2(k)$.

l) Assuming $\beta_i = 0$, find the distribution of $T_i = \hat{\beta}_i / (s\sqrt{[(X^T X)^{-1}]_{ii}})$, with $s^2 = \hat{\epsilon}^T \hat{\epsilon}/(n - k)$ (hint: write $T_i = \left(\hat{\beta}_i / \left(\sigma\sqrt{[(X^T X)^{-1}]_{ii}}\right)\right) / \sqrt{(\hat{\epsilon}^T \hat{\epsilon}/\sigma^2)/(n - k)}$).

m) Assuming $\beta = 0$, find the distribution of $F = (\hat{Y}^T \hat{Y}/k)/(\hat{\epsilon}^T \hat{\epsilon}/(n - k))$ (hint: $F = \left(\left(\hat{Y}^T \hat{Y}/\sigma^2\right)/k\right) / \left(\left(\hat{\epsilon}^T \hat{\epsilon}/\sigma^2\right)/(n - k)\right)$.)

Solution

a)

Let us show that M is symmetric and idempotent:

$$M^T = (X(X^T X)^{-1} X^T)^T = X(X^T X)^{-1} X^T = M$$

$$M^2 = X(X^T X)^{-1} X^T X (X^T X)^{-1} X^T = X(X^T X)^{-1} X^T = M$$

Let us compute the rank of M:

$$\text{rank}(M) = \text{rank}(X(X^T X)^{-1} X^T) = \text{trace}(X(X^T X)^{-1} X^T) = \text{trace}((X^T X)^{-1} X^T X) = k$$

Let us show that $I - M$ is symmetric and idempotent:

$$(I - M)^T = I^T - M^T = I - M$$

$$(I - M)(I - M) = I^2 - M - M + M^2 = I - M$$

Let us compute the rank of $I - M$:

$$\begin{aligned} \text{rank}(I - X(X^T X)^{-1} X^T) &= \text{trace}(I - X(X^T X)^{-1} X^T) \\ &= n - \text{trace}(X(X^T X)^{-1} X^T) \\ &= n - \text{trace}((X^T X)^{-1} X^T X) \\ &= n - k. \end{aligned}$$

b)
Since $\epsilon \sim N(0, \sigma^2 I)$ and $Y = X\beta + \epsilon$, $Y \sim N(X\beta, \sigma^2 I)$. Hence

$$\begin{aligned} f_Y(y) &= (2\pi\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2}(y - X\beta)^T(\sigma^2 I)^{-1}(y - X\beta)\right\} \\ &= (2\pi\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2}(y - X\beta)^T(y - X\beta)\right\} \end{aligned}$$

c)
The maximum likelihood estimate is the vector $(\hat{\beta}, \hat{\sigma}^2)$ that maximizes the likelihood function, or equivalently the log-likelihood function. The log-likelihood function is

$$l(\beta, \sigma^2) = \log f_Y(y) = (-n/2)\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}(y - X\beta)^T(y - X\beta)$$

Differentiating with respect to β and σ^2, and letting the derivatives equal to zero, we find

$$\begin{cases} \frac{1}{\sigma^2}X^T(y - X\beta) = 0 \\ -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}(y - X\beta)^T(y - X\beta) = 0 \end{cases}$$

The solution is

$$\begin{cases} \beta = (X^T X)^{-1}X^T y \\ \sigma^2 = \dfrac{(y - X\beta)^T(y - X\beta)}{n} \end{cases}$$

Hence the maximum likelihood estimator of (β, σ^2) is

$$\hat{\beta} = (X^T X)^{-1}X^T Y$$

$$\begin{aligned} \hat{\sigma}^2 &= \frac{(Y - X\hat{\beta})^T(Y - X\hat{\beta})}{n} \\ &= \frac{(Y - \hat{Y})^T(Y - X\hat{Y})}{n} \\ &= \frac{\hat{\epsilon}^T\hat{\epsilon}}{n}. \end{aligned}$$

Furthermore,

$$\begin{aligned} \hat{\beta} &= (X^T X)^{-1}X^T Y \\ &= (X^T X)^{-1}X^T(X\beta + \epsilon) \\ &= (X^T X)^{-1}X^T X\beta + (X^T X)^{-1}X^T\epsilon \\ &= \beta + (X^T X)^{-1}X^T\epsilon. \end{aligned}$$

d)
Since $\hat{\beta} = \beta + (X^T X)^{-1}X^T\epsilon$ is a linear transformation of ϵ and $\epsilon \sim N(0, \sigma^2 I)$,

$$\hat{\beta} \sim N(\beta, (X^T X)^{-1}X^T\sigma^2 I X(X^T X)^{-1}) = N(\beta, \sigma^2(X^T X)^{-1})$$

e)
It holds

$$\hat{Y} = X\hat{\beta} = X(X^T X)^{-1}X^T Y = MY.$$

On the other hand,

$$\begin{aligned}\hat{Y} &= X\hat{\beta} \\ &= X(X^TX)^{-1}X^TY \\ &= X(X^TX)^{-1}X^T(X\beta+\epsilon) \\ &= X\beta + X(X^TX)^{-1}X^T\epsilon \\ &= X\beta + M\epsilon.\end{aligned}$$

Since $\hat{Y} = X\beta + M\epsilon$ is a linear transformation of ϵ and $\epsilon \sim N(0, \sigma^2 I)$,

$$\hat{Y} \sim N(X\beta, \sigma^2 MM^T) = N(X\beta, \sigma^2 M^2) = N(X\beta, \sigma^2 M)$$

f)
It holds

$$\hat{\epsilon} = Y - \hat{Y} = Y - MY = (I-M)Y.$$

On the other hand,

$$\hat{\epsilon} = Y - \hat{Y} = X\beta + \epsilon - X\beta - M\epsilon = \epsilon - M\epsilon = (I-M)\epsilon.$$

Since $\hat{\epsilon} = (I-M)\epsilon$ and since $\epsilon \sim N(0, \sigma^2 I)$,

$$\hat{\epsilon} \sim N(0, \sigma^2(I-M)(I-M)^T) = N(0, \sigma^2(I-M)^2) = N(0, \sigma^2(I-M))$$

g)
We already know that

$$\begin{aligned}\hat{\epsilon} &= (I-M)\epsilon \\ \hat{\beta} &= \beta + (X^TX)^{-1}X^T\epsilon \\ \hat{Y} &= X\beta + M\epsilon\end{aligned}$$

Let $Z = \epsilon/\sigma$. Then $Z \sim N(0, I)$ and

$$\begin{aligned}\hat{\epsilon} &= L_1 Z \\ \hat{\beta} &= \beta + L_2 Z \\ \hat{Y} &= X\beta + L_3 Z\end{aligned}$$

with

$$\begin{aligned}L_1 &= \sigma(I-M) \\ L_2 &= \sigma(X^TX)^{-1}X^T \\ L_3 &= \sigma M\end{aligned}$$

To show the independence it is sufficient to verify that $L_1L_2^T = 0$ and $L_1L_3^T = 0$. It holds

$$\begin{aligned}L_1L_2^T &= \sigma(I-M)\sigma X(X^TX)^{-1} = \sigma^2(X(X^TX)^{-1} - X(X^TX)^{-1}X^TX(X^TX)^{-1}) = 0 \\ L_1L_3^T &= \sigma(I-M)\sigma M^T = \sigma^2(M-M) = 0.\end{aligned}$$

h)
Since $\hat{\epsilon} = (I-M)\epsilon$ with $I-M$ is symmetric and idempotent, we can write

$$\frac{\hat{\epsilon}^T\hat{\epsilon}}{\sigma^2} = \frac{\epsilon^T}{\sigma}(I-M)(I-M)^T\frac{\epsilon}{\sigma} = Z^T(I-M)Z$$

with $Z = \epsilon/\sigma$. Since $Z \sim N(0, I)$ and $I-M$ is symmetric and idempotent and has rank $n-k$, $\hat{\epsilon}^T\hat{\epsilon}/\sigma^2$ has a chi-square distribution with $n-k$ degrees of freedom.

i)
If $\beta = 0$, $\hat{Y} = M\epsilon = \sigma MZ$ with $Z = \epsilon/\sigma \sim N(0, I)$.
Hence

$$\frac{\hat{Y}^T\hat{Y}}{\sigma^2} = Z^T M^T M Z = Z^T M Z$$

Since M is symmetric and idempotent and has rank k, $\hat{Y}^T\hat{Y}/\sigma^2$ has a chi-square distribution with k degrees of freedom.

l)
Notice that

$$T_i = \frac{\hat{\beta}_i / \left(\sigma\sqrt{[(X^TX)^{-1}]_{ii}}\right)}{\sqrt{\left(\hat{\epsilon}^T\hat{\epsilon}/\sigma^2\right)/(n-k)}}$$

If $\beta = 0$, $\hat{\beta} \sim N(0, \sigma^2(X^TX)^{-1})$. Hence $\hat{\beta}_i / \left(\sigma\sqrt{[(X^TX)^{-1}]_{ii}}\right) \sim N(0,1)$. Furthermore $\hat{\epsilon}^T\hat{\epsilon}/\sigma^2 \sim \chi^2(n-k)$. We also know that $\hat{\beta}$ is independent of $\hat{\epsilon}$. Hence $\hat{\beta}_i / \left(\sigma\sqrt{[(X^TX)^{-1}]_{ii}}\right)$ and $\hat{\epsilon}^T\hat{\epsilon}/\sigma^2$ are independent. It follows that T_i has a Student-t distribution with $n-k$ degrees of freedom.

m)
We can write

$$F = \frac{\hat{Y}^T\hat{Y}/k}{\hat{\epsilon}^T\hat{\epsilon}/(n-k)} = \frac{\left(\hat{Y}^T\hat{Y}/\sigma^2\right)/k}{\left(\hat{\epsilon}^T\hat{\epsilon}/\sigma^2\right)/(n-k)}$$

We already know that, under the assumption $\beta = 0$, $\hat{Y}^T\hat{Y}/\sigma^2 \sim \chi^2(k)$, $\hat{\epsilon}^T\hat{\epsilon}/\sigma^2 \sim \chi^2(n-k)$. Furthermore, since $\hat{Y}$ and $\hat{\epsilon}$ are independent, $\hat{Y}^T\hat{Y}/\sigma^2$, $\hat{\epsilon}^T\hat{\epsilon}/\sigma^2$ are independent. It follows that F has a Fisher-F distribution with k and $n-k$ degrees of freedom.