Solutions (TMA4267 2023 August)

1.

a) 1) The probability density function of X is

$$f(x_1, x_2) = \frac{1}{\pi} I_{\{x_1^2+x_2^2 \le 1\}}(x_1, x_2),$$

where I is the indicator function. To obtain the marginal density of X_1, we integrate the joint density with respect to x_2. Let $-1 \le x_1 \le 1$ (otherwise $f(x_1, x_2) = 0$). Then

$$f_1(x_1) = \frac{1}{\pi} \int_{-\infty}^{\infty} I_{\{x_1^2+x_2^2 \le 1\}}(x_1, x_2) dx_2 = \frac{1}{\pi} \int_{-\infty}^{\infty} I_{\left(-\sqrt{1-x_1^2}, \sqrt{1-x_1^2}\right)}(x_2) dx_2 =$$

$$= \frac{1}{\pi} \int_{-\sqrt{1-x_1^2}}^{\sqrt{1-x_1^2}} dx_2 = \frac{2\sqrt{1-x_1^2}}{\pi}.$$

If $|x_1| > 1$, then $f_1(x_1) = 0$.

In the same way $f_2(x_2) = \frac{2\sqrt{1-x_2^2}}{\pi}$ if $|x_2| \le 1$ and $f_2(x_2) = 0$ if $|x_2| > 1$.

Similarly (by integration)

2) $f_k(x_k) = 1/2$ if $|x_k| \le 1$ and $f_k(x_k) = 0$ if $|x_k| > 1$, $k = 1, 2$. Note that in this case

$$f(x_1, x_2) = I_{\{|x_1| \le 1, |x_2| \le 1\}}(x_1, x_2) = I_{\{|x_1| \le 1\}}(x_1) I_{\{|x_2| \le 1\}}(x_2) = f_1(x_1) f_2(x_2).$$

3) $f_k(x_k) = 1 - |x_k|$ if $|x_k| \le 1$ and $f_k(x_k) = 0$ if $|x_k| > 1$, $k = 1, 2$.

b) 1)

$$\mathrm{Cov}(X_1, X_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 f(x_1, x_2) dx_2 dx_1 = \int_{-\infty}^{\infty} x_1 \left(\int_{-\infty}^{\infty} x_2 f(x_1, x_2) dx_2 \right) dx_1 =$$

$$= \int_{-\infty}^{\infty} x_1 \left(\int_{-\infty}^{\infty} x_2 I_{\{x_1^2+x_2^2 \le 1\}}(x_1, x_2) dx_2 \right) dx_1 =$$

$$= \int_{-\infty}^{\infty} x_1 \left(\int_{-\infty}^{\infty} x_2 I_{\left[-\sqrt{1-x_1^2}, \sqrt{1-x_1^2}\right]}(x_2) dx_2 \right) dx_1 = \int_{-\infty}^{\infty} x_1 \left(\int_{-\sqrt{1-x_1^2}}^{\sqrt{1-x_1^2}} x_2 dx_2 \right) dx_1 = 0$$

because the inner integral equals zero (odd function, symmetric interval).

2) Since $f(x_1, x_2) = f_1(x_1) f_2(x_2)$, the variables X_1 and X_2 are independent. Therefore

$$\mathrm{Cov}(X_1, X_2) = 0$$

.

3) Denote

$$Q = \{(x_1, x_2) : -1 \le x_1 \le 1, -1 - x_1 \le x_2 \le 1 - x_1, x_1 \le x_2 \le x_1 + 1\}.$$

Then

$$f(x_1, x_2) = \frac{1}{2} I_Q(x_1, x_2),$$

and

$$\mathrm{Cov}(X_1, X_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 f(x_1, x_2) dx_2 dx_1 = \frac{1}{2}\int_{-1}^{1} x_1 \left(\int_{-\infty}^{\infty} x_2 I_Q(x_1, x_2) dx_2\right) dx_1 =$$

$$= \frac{1}{2}\int_{-1}^{0} x_1 \left(\int_{-(1+x_1)}^{1+x_1} x_2 dx_2\right) dx_1 + \frac{1}{2}\int_{0}^{1} x_1 \left(\int_{-(1-x_1)}^{1-x_1} x_2 dx_2\right) dx_1 = 0$$

because the two inner integrals equal zero (odd function, symmetric intervals).

X_1 and X_2 are independent in the second case: the joint density equals the product of marginal densities. In two other cases X_1 and X_2 are dependent. Indeed, in both cases $P(X_1 > 0.9) > 0$, $P(X_2 > 0.9) > 0$, but $P(X_1 > 0.9, X_2 > 0.9) = 0$.

2.

a) Since $p = 3$ and $n - p$ (the number of degrees of freedom) is 117, $n = 120$.

$$t_1 = \hat{\beta}_1 / \sqrt{\widehat{\mathrm{Var}}\hat{\beta}_1} = 1.3596/0.4548 = 2.989.$$

$$\sqrt{\widehat{\mathrm{Var}}\hat{\beta}_2} = \hat{\beta}_2 / t_2 = (-0.3780)/(-0.831) = 0.4548.$$

$$R^2_{\mathrm{adj}} = 1 - (1 - R^2)\frac{n-1}{n-k-1} = 1 - (1 - 0.1213)\frac{119}{117} = 0.1063.$$

Thus the output is in fact as follows

```
Call:
lm(formula = Y ~ x1 + x2)

Residuals:
    Min      1Q  Median      3Q     Max
-4.7169 -1.1809  0.0742  1.2974  4.3071

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   3.7662     0.5571   6.761  5.7e-10 ***
x1            1.3596     0.4548   2.989  0.00341 **
x2           -0.3780     0.4548  -0.831  0.40758
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.034 on 117 degrees of freedom
Multiple R-squared:  0.1213,    Adjusted R-squared:  0.1063
F-statistic: 8.074 on 2 and 117 DF,  p-value: 0.0005193
```

b) Denote elements of $(X^T X)^{-1}$ by c_{ij}:

$$(X^T X)^{-1} = [c_{ij}]_{i,j=0,1,\ldots,k}.$$

In our case $k = 2$ and we have to find c_{jj}, $j = 0, 1, 2$. The following equalities hold (known from the course):

$$\widehat{\mathrm{Var}}\hat{\beta}_j = \hat{\sigma}^2 c_{jj}$$

therefore

$$c_{jj} = \frac{\widehat{\mathrm{Var}\hat{\beta}_j}}{\hat{\sigma}^2}.$$

Thus

$$c_{00} = \frac{0.5571^2}{2.034^2} = 0.075,$$

$$c_{11} = c_{22} = \frac{0.4548^2}{2.034^2} = 0.05.$$

c) $(1-\alpha)$-confidence interval for β_j $(j = 0, 1, 2)$ is

$$\left[\hat{\beta}_j - t_{\frac{\alpha}{2},n-p}\sqrt{\widehat{\mathrm{Var}\hat{\beta}_j}}, \hat{\beta}_j + t_{\frac{\alpha}{2},n-p}\sqrt{\widehat{\mathrm{Var}\hat{\beta}_j}}\right].$$

In our case

$$\hat{\beta}_1 = 1.3596,\ \sqrt{\widehat{\mathrm{Var}\hat{\beta}_1}} = 0.4548,\ t_{0.025,117} = 1.98,$$

therefore 95% confidence interval for β_1 is $[0.4596, 2.2596]$.

d) The first and second null hypotheses are rejected becuase the corresponding p-values are smaller than α_{loc} which is 0.17. The third null hypothesis is not rejected.

3.

a) The design matrix and the vector of parameters of the true model are

$$X = \begin{pmatrix} 1 & x_{11} & x_{12} & x_{13} \\ 1 & x_{21} & x_{22} & x_{23} \\ \dots \\ 1 & x_{n1} & x_{n2} & x_{n3} \end{pmatrix},\ \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}.$$

The design matrix and the vector of parameters of the underfitted model are

$$X_u = \begin{pmatrix} 1 & x_{11} & x_{12} \\ 1 & x_{21} & x_{22} \\ \dots \\ 1 & x_{n1} & x_{n2} \end{pmatrix},\ \beta_u = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix}.$$

Then

$$X\beta = X_u\beta_u + X_a\beta_3$$

where

$$X_a = \begin{pmatrix} x_{13} \\ x_{23} \\ \dots \\ x_{n3} \end{pmatrix}.$$

Denote

$$\hat{\beta}_u = \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix}.$$

Then

$$\hat{\beta}_u = (X_u^T X_u)^{-1} X_u^T Y,$$

and, since $EY = X\beta$,

$$E\hat{\beta}_u = (X_u^T X_u)^{-1} X_u^T EY = (X_u^T X_u)^{-1} X_u^T (X_u\beta_u + X_a\beta_3) =$$

$$= \beta_u + (X_u^T X_u)^{-1} X_u^T X_a \beta_3.$$

The expectation of the estimator is not equal to the vector of parameters, i.e. the estimator is biased.

4.

a) It is easy to see that

$$E\text{sign}(Y) = 0$$

and

$$E(\text{sign}(Y)|Y|) = EY = 0,$$

therefore

$$\text{Cov}(|Y|, \text{sign}(Y)) = E\text{sign}(Y)|Y| - E(\text{sign}(Y)) \cdot E|Y| = 0.$$

Note that X_1 and X_2 are not only uncorrelated but independent. Indeed, since $|Y|$ is positive, and $\text{sign}(Y)$ takes two values 1 and -1, the independence will be proved if we show that

$$P(|Y| \leq y, \text{sign}(Y) = 1) = P(|Y| \leq y) P(\text{sign}(Y) = 1),$$

$$P(|Y| \leq y, \text{sign}(Y) = -1) = P(|Y| \leq y) P(\text{sign}(Y) = -1)$$

for any positive y. Let $y > 0$. Then

$$P(|Y| \leq y, \text{sign}(Y) = 1) = P(|Y| \leq y, Y \geq 0) = P(0 \leq Y \leq y) = \frac{1}{2} P(-y \leq Y \leq y) =$$

$$= \frac{1}{2} P(|Y| \leq y) = P(\text{sign}(Y) = 1) P(|Y| \leq y).$$

The second equality is proved in the same way.

Note also that X_1 and X_2 are independent not only for standard normal Y but for any Y with symmetric about 0 density.

5.

a) Denote

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}.$$

Since AX and X are independent, we have $A\Sigma = 0$, i.e.

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} = \begin{pmatrix} 1+\rho & \rho+1 \\ 1+\rho & \rho+1 \end{pmatrix} = 0$$

or $\rho = -1$. Thus

$$\Sigma = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

b) Since the correlation coefficient between X_1 and X_2 is -1, we have

$$X_1 = -X_2$$

with probability 1 or $P(X_1 + X_2 = 0) = 1$. Hence $P(X_1 + X_2 > 1) = 0$.