1.

since Σ is diagonal

$\Rightarrow$ Covariance of Y_i, Y_j $= E[(Y_i - E[Y_i])(Y_j - E[Y_j])] = 0$ for every $i \neq j$

$E[(Y_i - E[Y_i])(Y_j - E[Y_j])] = E[Y_i Y_j] - E[Y_i]E[Y_j] = 0 \Rightarrow E[Y_i Y_j] = E[Y_i] \cdot E[Y_j]$

choose $m=2$ $\Rightarrow$ when Y_1, Y_2 has zero covariance $\Rightarrow$ Σ is digonal

①

suppose Y_1 is random variable with uniform distribution in interval $[-1, 1]$

Then we set $Y_2 = SY_1$, S is another random variable s.t. $S = \begin{cases} 1 & p=\frac{1}{2} \\ -1 & p=\frac{1}{2} \end{cases}$

Clearly, $S \perp Y_1 \Rightarrow E[Y_2] = E[SY_1^2] = E[S]\,E[Y_1^2] = 0$

Σ is diagonal $\Uparrow$

Also, $E[Y_1] = 0 \Rightarrow E[Y_1] \cdot E[Y_2] = 0 \Rightarrow E[Y_1 Y_2] = E[Y_1] \cdot E[Y_2] \Rightarrow$ Covariance of $Y_1, Y_2 = 0$

However, since $P[Y_2 > \frac{1}{2} \mid Y_1 > \frac{1}{2}] = \frac{1}{2} \neq P[Y_2 > \frac{1}{2}] = \frac{1}{4}$ $\Rightarrow$ Y_1, Y_2 aren't independent

$\Rightarrow$ Y isn't a vector of independent even when Σ is diagonal

②

$Y_1 \sim N(0,1)$, $Y_2 = Y_1^2 - 1$

$E[Y_1 Y_2] = E[Y_1^3 - Y_1] = 0 - 0 = 0$ $E[Y_1] \cdot E[Y_2] = 0 \cdot E[Y_1] = 0 \Rightarrow Cov(Y_1, Y_2) = 0 \Rightarrow \Sigma$ is diagonal

however, $P(Y_2 > \delta^2 - 1 \mid Y_1 > 0) = \dfrac{P(Y_1 > \delta)}{\underbrace{P(Y_1 > 0)}_{\frac{1}{2}}} = 2P(Y_1 > \delta) \neq P(Y_2 > \delta^2 - 1) = \frac{1}{2}$

$\Rightarrow$ Y_1, Y_2 aren't independent even when Σ is diagonal

2.

$Y \sim N(\mu, \Sigma) \Rightarrow$ Y's elements are jointly gaussian

$$\Rightarrow f_Y(y) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp(-\tfrac{1}{2}(y-\mu)^T \Sigma^{-1} (y-\mu))$$

if Σ is diagonal with diagonal element denotes $\sigma_1^2, \sigma_2^2 \cdots \sigma_m^2$, respectively

$$\Rightarrow f_Y(y) = \frac{1}{(2\pi)^{n/2} \cdot \prod_{i=1}^{m} \sigma_i} \cdot \exp(-\tfrac{1}{2} \cdot \sum_{i=1}^{m} \sigma_i^{-2} \cdot (y_i - \mu)^2)$$

$$= \prod_{i=1}^{m} \frac{1}{(2\pi)^{1/2} \cdot \sigma_i} \cdot \exp(-\tfrac{1}{2} \cdot (\tfrac{y_i - \mu}{\sigma_i})^2)$$

$\nearrow f_{Y_i}(y_i) = \frac{1}{(2\pi)^{1/2} \cdot \sigma_i} \cdot \exp(-\frac{1}{2}(\frac{y_i - \mu_i}{\sigma_i})^2)$, which is trivially $N(\mu_i, \sigma_i)$

$= \prod_{i=1}^{m} f_{Y_i}(y_i) \Rightarrow$ it is proved that all the elements in Y are independent if $Y \sim N(\mu, \Sigma)$ #

3.

By #2, we can simply recognize that σ_i is the std of each random variable Y_i

$\Rightarrow \Sigma = \sigma^2 I_m$ is equivalent to the fact that Y_i has $Var(Y_i) = \sigma^2$, $i = 1, 2 \cdots m$

$$E[Y'Y] = E[\sum_{i=1}^{m} Y_i^2] = m(\sigma^2 + \overset{\text{square of mean}}{\sigma^2}) = m\sigma^2 \ \#$$

$$E[(Y'Y)^2] = E[(\sum_{i=1}^{m} Y_i^2)^2] = E[2\sum_{i=1}^{m}\sum_{j=1}^{i-1}(Y_iY_j)^2] + \sum_{i=1}^{m} \overset{3\sigma^4 \text{ for each}}{E[Y_i^4]}$$

for each i, j $E[(Y_iY_j)^2] = E[Y_i^2] \cdot E[Y_j^2] = \sigma^4$

since Y_i, Y_j are independent, $f_{Y_i,Y_j}(y_i, y_j) = f_{Y_i}(y_i) \cdot f_{Y_j}(y_j)$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y_i^2 y_j^2 f_{Y_i,Y_j}(y_i, y_j)\, dy_i dy_j = \int_{-\infty}^{\infty} y_i^2 f_{Y_i}(y_i) dy_i \cdot \int_{-\infty}^{\infty} y_j^2 f_{Y_j}(y_j) dy_j = E[Y_i^2] \cdot E[Y_j^2]$$

$$\Rightarrow E[(Y'Y)^2] = E[(\sum_{i=1}^{m} Y_i^2)^2] = E[2\sum_{i=1}^{m}\sum_{j=1}^{i-1}(Y_iY_j)^2] + \sum_{i=1}^{m} E[Y_i^4] = (m(m-1) + 3m)\sigma^4 = m(m+2)\sigma^4 \ \#$$

4.

by definition $E[Y] = \sum_{i=1}^{n} P(A_i) \cdot E[Y|A_i]$, $\sum_{i=1}^{n} P(A_i) = 1$

$\Rightarrow \alpha_1$ should be $P(X_1 > 0.015) = 0.091$ α_2 should be $P(X_1 \leq 0.015) = 0.909$

$\Downarrow$ $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\Downarrow$

$\frac{46}{504}$ $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\frac{458}{504}$

$\Rightarrow \quad \beta_{11} = P[X_1 > 0.015,\ X_2 > 0.02] = 0.056 \quad \beta_{11} = P[X_1 > 0.015,\ X_2 \leq 0.02] = 0.036$

$\beta_{21} = P[X_1 \leq 0.015,\ X_2 > 0.02] = 0.478 \quad \beta_{21} = P[X_1 \leq 0.015,\ X_2 \leq 0.02] = 0.430$

Also, $E[Y|A] = E[Y|A \cap B_1] \cdot P(B_1|A) + E[Y|A \cap B_2] \cdot P(B_2|A)$

$\Rightarrow \quad \gamma_1 = P(X_1 > 0.015 \mid X_3 > -4) = 0.135$

$\gamma_2 = P(X_1 \leq 0.015 \mid X_3 > -4) = 0.865$

5.

$$E[(m(x) - X'b)^2]$$

$$= E[(m(x) - X'b)'(m(x) - X'b)]$$

$$= E[(m(x))^2 - b'X \cdot m(x) - m(x)'X'b + b'XX'b]$$

$$= E[(m(x))^2] - 2b'E[X \cdot m(x)] + b'E[XX']b$$

derive by b , argmin should satisfy that

$$\nabla_b E[(m(x) - X'b)^2] = -2E[X\, m(x)] + 2E[XX']\, b = 0$$

$\Rightarrow$ we can obtain b by solving $E[XX']b = E[X\, m(x)]$

since $E[XX']$ is positive-definite $\Rightarrow$ $E[XX']^{-1}$ exists

$$\Rightarrow b = E[XX']^{-1} \cdot E[X\, m(x)]$$

$$= E[XX']^{-1} \cdot E[X \cdot E[Y|X]]$$

$$E[E[X \cdot Y|X]] = \int_{-\infty}^{\infty} E[X \cdot Y | X = x] \cdot f_X(x)\, dx$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x \cdot y \cdot f_{Y|X}(y|x)\, dy\, f_X(x)\, dx$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x \cdot y \cdot f_{Y|X}(y|x) \cdot f_X(x)\, dy\, dx$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x \cdot y \cdot f(x, y)\, dy\, dx$$

$$= E[X \cdot Y]$$

$$\Rightarrow b = E[XX']^{-1} \cdot E[X\, E[Y|X]] = E[XX']^{-1} \cdot E[X \cdot Y] \#$$

6.

pdf of X : $f(t;3) = \dfrac{\Gamma(2)}{\sqrt{3\pi}\,\Gamma(1.5)}\left(1+\dfrac{t^2}{3}\right)^{-2} = \dfrac{2}{\sqrt{3}} \cdot \dfrac{1}{\pi} \cdot \left(1+\dfrac{t^2}{3}\right)^{-2}$

(1)

we need degree of freedom to make X possess larger variance, which enable to collect wider range of X and make the linear regression more proper.

(2)

linear projection of Y on X follows the least square error

$\Rightarrow$ for random variable X, Y, coefficient β is $\dfrac{Cov[X,Y]}{Var[X]}$

since $Y = \dfrac{1}{1+X^4}$ is non-linear, Y has a complicated distribution

$\Rightarrow$ it is hard to compute $cov[X,Y]$

Thus I use python to sample 100000 pts of X and estimate the $\hat{\beta}$ based on the sample points.

I got $\beta \doteq 0.00042$, which means that linear relationship between X, Y is very weak.

```
# Load necessary library
library(MASS)  # For generating random samples from t-distribution

# Define the function for Y = 1 / (1 + X^4)
Y_func <- function(X) {
  return(1 / (1 + X^4))
}

# Generate random samples for X from t-distribution with 3 degrees of freedom
set.seed(0)
num_samples <- 100000
X_samples <- rt(num_samples, df = 3)

# Calculate the corresponding Y values
Y_samples <- Y_func(X_samples)

# Estimate the necessary expectations
E_X <- mean(X_samples)
E_Y <- mean(Y_samples)
E_X2 <- mean(X_samples^2)
E_XY <- mean(X_samples * Y_samples)

# Calculate the linear regression coefficient b
b <- (E_XY - E_X * E_Y) / (E_X2 - (E_X)^2)

# Print the result
print(paste("The estimated value of b is:", b))
```

```
[1] "The estimated value of b is: 0.000417701353039401"
```