1. Consider a linear model of the form

\[ Y = X\beta + E \]

where \( Y \) is the column vector of responses \( Y_1, \ldots, Y_n \) for \( n \) individuals, \( X \) is an \( n \times q \) matrix of known values, \( \beta \) is a column vector of parameters \( \beta_1, \ldots, \beta_p \), and \( E \) is a column vector of errors \( E_1, \ldots, E_n \) which are assumed to be i.i.d. Normal \( N(0, \sigma^2) \) random variables.

The maximum likelihood estimator of the parameter vector \( \beta \) is \( \hat{\beta} = (X^T X)^{-1} X^T Y \). The vector of fitted values is \( \hat{Y} = X\hat{\beta} \) and the vector of residuals is \( R = Y - \hat{Y} \).

(a) Show that \( X^T R = 0 \), whatever the values of \( Y_i \) are.

(b) Suppose that the design matrix \( X \) contains a column of 1’s. Show that the residuals must always sum to zero: \( \sum_i R_i = 0 \), whatever the values of \( Y_i \) are.

2. Suppose \( Z_1, Z_2 \) are i.i.d. Normal \( N(0, 1) \) random variables. Find the joint probability distribution of \( (Y_1, Y_2) \) where

\[
Y_1 = Z_1 + Z_2 \\
Y_2 = Z_1 - Z_2
\]

[Hint: use properties of the Multivariate Normal Distribution.]

3. Consider the model in Question 1. Instead of assuming that the errors \( E_i \) are i.i.d. Normal \( N(0, \sigma^2) \), let us now assume that the random vector \( E \) has the Multivariate Normal distribution with mean vector \( 0 \) and variance-covariance matrix \( \text{E}[EE^T] = aV \) where \( V \) is a known matrix, and \( a \) is an unknown scale factor. (For example, if \( V \) is the identity matrix, this is the same as assuming that the errors are i.i.d. with variance \( a \)). In this case, the maximum likelihood estimator of \( \beta \) is \( \hat{\beta} = (X^T V^{-1} X)^{-1} X^T V^{-1} Y \).

(a) Show that the MLE is unbiased, that is, that the expected value of the random vector \( \hat{\beta} \) is \( \text{E}[^{\hat{\beta}}] = \beta \).

(b) Find the variance-covariance matrix of \( \hat{\beta} \).

(c) What is the distribution of the random vector \( \hat{\beta} \)?
4. Consider the proportional regression model

\[ Y_i = \alpha x_i + E_i, \quad i = 1, \ldots, n \]

where \( Y_1, \ldots, Y_n \) are the responses, \( x_1, \ldots, x_n \) are the corresponding (and known) values of the explanatory variable, \( \alpha \) is the unknown parameter, and \( E_i \) are independent random errors. Assume that the variance of \( E \) is proportional to \( x \), that is,

\[ \text{var}(E_i) = cx_i, \quad i = 1, \ldots, n \]

where \( c \) is an unknown scale factor.

Using the results of Question 3,

(a) find the maximum likelihood estimator \( \hat{\alpha} \) of \( \alpha \).

(b) find the distribution of \( \hat{\alpha} \).

5. An integer random variable \( Y \) has the geometric distribution with parameter \( p \), written \( Y \sim \text{Geom}(p) \), if \( P\{Y = y\} = (1 - p)^{y-1}p \) for \( y = 1, 2, \ldots \). The mean is \( \mu = E[Y] = 1/p \).

Consider a generalised linear model with Geometric responses,

\[ Y_i \sim \text{Geom}(p_i), \quad i = 1, \ldots, n \]

and linear predictor

\[ g(1/p_i) = \beta_1 x_{i1} + \ldots + \beta_q x_{iq}, \quad i = 1, \ldots, n \]

where \( g \) is a known function, \( x_{ij} \) are fixed, known values, and \( \beta_1, \ldots, \beta_q \) are parameters.

(a) Give an expression for the likelihood of the model involving only the means \( \mu_i = 1/p_i \) and the observed responses \( y_i \).

(b) Hence obtain an expression for the deviance of the model, involving only the fitted means \( \widehat{\mu}_i = 1/\widehat{p}_i \) and the observed responses \( y_i \).

6. Consider a generalised linear model with Poisson responses,

\[ Y_i \sim \text{Poisson}(\mu_i), \quad i = 1, \ldots, n \]

and linear predictor

\[ g(\mu_i) = \beta_1 x_{i1} + \ldots + \beta_q x_{iq}, \quad i = 1, \ldots, n \]

where \( g \) is a known function, \( x_{ij} \) are fixed, known values, and \( \beta_1, \ldots, \beta_q \) are parameters.

(a) Give an expression for the likelihood of the model involving only the means \( \mu_i \) and the observed responses \( y_i \).

(b) Hence obtain an expression for the deviance of the model, involving only the fitted means \( \widehat{\mu}_i \) and the observed responses \( y_i \).