1. (Davison 2003, Box-Cox) Suppose we have data \( y, X_1, \ldots, X_p \) with \( y_k \geq 0 \). Let
\[
y'_k = (y_k^\lambda - 1)/\lambda \quad \text{for } k = 1, 2, \ldots, n.
\]
If we fit \( y' = X\beta + \epsilon \) for \( \epsilon \sim N(0, I_n\sigma^2) \), with \( x_k \) the \( k \)th row \( X \), then the likelihood is
\[
L(\beta, \sigma^2, \lambda; y') = (2\pi\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2} \sum_k (y'_k - x_k^T\beta)^2 \right).
\]
(a) Make a change of variables from \( y' \) to \( y \), and calculate the log-likelihood, \( \ell(\beta, \sigma^2, \lambda; y) \) in terms of \( y \).
(b) Show that the MLE’s for \( \beta \) and \( \sigma^2 \) are
\[
\hat{\beta}'(y; \lambda) = (X^TX)^{-1}X^Ty' \quad \text{and} \quad \hat{\sigma}^2(y; \lambda) = (y' - \hat{y}')^T(y' - \hat{y}')/n,
\]
with \( y' = y'(y, \lambda) \) and \( \hat{y}' = X\hat{\beta}' \).
(c) Substitute these into the likelihood, and show that the MLE \( \hat{\lambda} \) for \( \lambda \) is the value of \( \lambda \) maximising
\[
\ell(\lambda; y) = -\frac{n}{2} \log(\hat{\sigma}^2(y; \lambda)) + (\lambda - 1) \log(y_1y_2\ldots y_n).
\]
(d) Consider the normal linear model
\[
y'/\sqrt{(y_1y_2\ldots y_n)^{(\lambda - 1)/n}} = X\gamma + \epsilon.
\]
Denote by RSS' the residual sum of squares for this regression. Show that
\[
\ell(\lambda; y) = -\frac{n}{2} \log(\text{RSS}').
\]
(e) Show that an approximate \((1 - \alpha)\) confidence interval for \( \lambda \) is given by the set
\[
\left\{ \lambda : \ell(\lambda; y) \geq \ell(\hat{\lambda}; y) - \frac{q}{2} \right\}
\]
where \( q \) is the \( 1 - \alpha \) quantile of a \( \chi^2(1) \) distribution.

2. (Davison 2003 Q8.9.5) The linear model \( y = X\beta + \epsilon \) is thought to apply to a set of data, and it is assumed that \( \epsilon \sim N(0, \sigma^2I_n) \). Denote by \( \hat{\beta} \) and \( s^2 \) the usual estimates for \( \beta \) and \( \sigma^2 \) obtained under this model. Suppose in fact the errors are distributed as \( \epsilon \sim N(0, \Sigma) \), \( k = 1, 2, \ldots, n \) with \( \Sigma \) an invertible \( n \times n \) variance matrix.
(a) Show that \( \hat{\beta} \) remains an unbiased estimator of \( \beta \), and calculate its variance matrix.
(b) Suppose \( \Sigma \) is known. How may we correct the regression?
3. (Faraway 06 Chapter 6 Q4) Let $Y_i, 1 \leq i \leq n$, be independent random variables with probability density given in generalized linear model form by

$$f(y_i; \theta_i, \phi) = \exp \left(\frac{y_i \theta_i - \kappa(\theta_i)}{\phi} + c(y_i; \phi)\right),$$

with $\theta_i = \theta_i(\eta_i)$ a function of the linear predictor

$$\eta_i = x_i^T \beta = (x_{i,1}, \ldots, x_{i,p})(\beta_1, \beta_2, \ldots, \beta_p)^T.$$

Suppose $Y_i$ has a Exp($\lambda_i$) distribution with mean $E(Y_i) = 1/\lambda_i$. Find the natural parameter and $\phi$ and calculate $\kappa$ as a function of the natural parameter. Calculate the canonical link function. What practical difficulty arises when using the canonical link, in this instance?

4. Consider the case of regression with a normal linear model with $\sigma^2$ known. Let

$$W = -E \left(\frac{\partial^2 \ell}{\partial \eta \partial \eta^T}\right)$$

with $\ell$ the log-likelihood and $\eta_i = x_i^T \beta$ the linear predictor (notation for $\eta$ as Q3).

(a) Show that $\phi = \sigma^2$, give the canonical link function, and show that $W = I_n/\sigma^2$.

(b) Explain why

$$\hat{\beta} \xrightarrow{D} N(\beta, \phi(X^T \text{diag}(V(\mu_1), \ldots, V(\mu_n))X)^{-1})$$

when the link function is the canonical link function and verify this for the case of regression with a normal linear model.