Ahmad Sharif
K436765

sat sun mon tue wed thu fri Date :

Q1. Consider the linear model

Let us consider the estimation of $\beta$. Show or explain in details, one way or another, how the maximum likelihood estimator

$$\hat{\beta} = (X'X)^{-1}X'Y$$ is obtained.

# Let us consider the estimation of unknown parameters of the linear model, $y = X\beta + \epsilon$, $\epsilon \sim N(0, \sigma^2 I)$ — (1)

Since $y \sim N(X\beta, \sigma^2 I)$ the likelihood function $L(\beta, \sigma^2/Y)$ & log likelihood function $l(\beta, \sigma^2/y)$ $= \log(L(\beta, \sigma^2/y))$

have the forms

$$L(\beta, \sigma^2/y) = \frac{1}{\sqrt{\sqrt{(2\pi)}^n |\sigma^2 I}} \times \exp\left(-\frac{(y - X\beta)'(y - X\beta)}{2\sigma^2}\right)$$

sat sun mon tue wed thu fri Date :

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

The maximum likelihood estimators $\hat{\beta} = (\hat{\beta_0}, \hat{\beta_1}, \ldots \hat{\beta_p})$ of the parameters vector $\beta$ & $\sigma^2$ of the scaler $\sigma^2 > 0$ can be obtained by finding the maximum of the log-likelihood function $l(\beta, \sigma^2 / y)$ with respect to $\beta$, $\sigma^2$ by the process of differentiation.

The maximum likelihood estimators $\hat{\beta} = (\hat{\beta_0}, \hat{\beta_1}, \ldots \hat{\beta_p})$ of the parameters vector $\beta$ & $\hat{\sigma}^2$ of the scaler $\sigma^2 > 0$ can be obtained by finding maximum of the log-likelihood function $l(\beta, \sigma^2 / y)$ with respect to $\beta$, $\sigma^2$ by the process of differentiation

$$\hat{\beta}\,\hat{\sigma}^2 = \underset{\beta, \sigma^2}{\arg\max}\, l(\beta, \sigma^2 / y) = \underset{\beta, \sigma^2}{\arg\max} \left(-\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}(y - X\beta)'(y - X\beta)\right.$$

The local maximum $l(\beta, \sigma^2/y)$ with respect to $\beta$ occurs at point, where the gradient vector (partial derivates)

$$\frac{\partial l}{\partial \beta} = \begin{pmatrix} \frac{\partial l}{\partial \beta_0} \\ \frac{\partial l}{\partial \beta_p} \\ \vdots \\ \frac{\partial l}{\partial \beta_p} \end{pmatrix}$$

$$\frac{\partial l}{\partial \beta} = \frac{\partial\left(-\frac{1}{2\sigma^2}(y - X\beta)'(y - X\beta)\right)}{\partial \beta}$$

$$= \left(-\frac{1}{2\sigma^2}\right)\frac{\partial(-2y'X\beta)}{\partial \beta} + \left(-\frac{1}{2\sigma^2}\right) \times \left(\frac{\partial \beta' X' X\beta}{\partial \beta}\right)$$

$$= \left(-\frac{1}{2\sigma^2}\right)(-2X'y) + \left(-\frac{1}{2\sigma^2}\right)2X'X\beta$$

$$= -\frac{1}{2\sigma^2}(X'X\beta - X'y) = 0$$

(8)

Since $\frac{\partial l}{\partial \beta} = 0$ for any, $\sigma^2 > 0$ if

f only if the normal equations

$$x' x \beta = x' y \quad \text{holds for } \beta$$

The maximum likelihood estimator $\hat{\beta}$ is

the solution to the normal equations

$$x' x \hat{\beta} = x' y$$

$$\text{i.e } \hat{\beta} = (x' x)^{-1} x' y.$$

2 In generalized linear model, the likelihood equations can written ~~for~~ form

$$\frac{\partial l(\beta, \phi)}{\partial \beta_j} = \sum_{i=1}^{n} \frac{(y_i - \mu_i)}{var(y_i)} x_{ij} \left(\frac{\partial \mu_i}{\partial \eta_i}\right) = 0$$

$$j = 0, 1, 2 \ldots n$$

Consider now the simple log-linear ~~Inverse~~ Inverse Gaussian model with

$$Y_i \sim IG(\mu_i, \phi)$$

$$\log(\mu_i) = \eta_i = \beta_0$$

What kind of more simplified form the likelihood equations have in this case, that is, what form $\frac{\partial l(\beta_0)}{\partial \beta_0}$ has in the simple Inverse Gaussian model? By using the likelihood equations, find ~~the~~ the maximum likelihood estimator $\hat{\beta_0}$

Here, $\log(\mu_i) = \eta_i = \beta_0$

$Y_i \sim IG(\mu_i, \phi)$

$$\frac{\partial l(\beta_0)}{\partial \beta_0} = \sum_{i=1}^{n} \frac{(y_i - \mu_i)}{\phi \mu_i^2} \mu_i \cdot 1$$

$$= \frac{1}{\phi} \sum_{i=1}^{n} \left(\frac{y_i - \mu_i}{\mu_i}\right)$$

$\mu_i = e^{\beta_0}$

$$\frac{\partial l(\beta_0)}{\partial \beta_0} = \frac{1}{\phi} \sum_{i=1}^{n} \left(\frac{y_i - e^{\beta_0}}{e^{\beta_0}}\right)$$

For the MLE

$$\frac{1}{\phi} \sum_{i=1}^{n} \frac{(y_i - e^{\beta_0})}{e^{\beta_0}} = 0$$

$$\sum_{i=1}^{n} (y_i - e^{\beta_0}) = 0$$

$$\sum_{i=1}^{n} y_i - n e^{\beta_0} = 0$$

$$\sum_{i=1}^{n} y_i - n e^{\beta_0} = 0$$

$$\beta_0 = \log\left(\frac{1}{n}\sum_{i=1}^{n} y_i\right)$$

Thus

$$\hat{\beta_0} = \log(\bar{y})$$

This is the MLE for $\hat{\beta_0}$. for Simple non-linear Inverse Gaussian model.

**03** Let us assume $Y_i \sim Gamma(\mu_i, \phi)$

Consider the model

$$\log(\mu_i) = \beta_0 + \beta_1 x_i$$

Let the estimate of the parameter $\beta_0, \beta_1, \phi$ be as $\hat{\beta_0} = 1$

$$\hat{\beta_1} = 0.5$$

$$\phi = 1$$

when

sat sun mon tue wed thu fri Date:

$$X = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \\ 1 & 6 \\ 1 & 7 \\ 1 & 8 \\ 1 & 9 \\ 1 & 10 \end{pmatrix}$$

Calculate the estimated covariance matrix

$$\mathrm{cov}(\hat{\beta}) = (X' \hat{W} X)^{-1}$$

Ahmad Sharif
ahmad.sharif@tuni.fi
K436765

```r
###################   3   ###################
# Parameters
beta_hat <- c(1, 0.5)
phi_hat <- 0.1

X <- cbind(1, 1:10)

eta <- X %*% beta_hat
mu <- exp(eta)
W <- diag(1 / (phi_hat * mu^2))

cov_beta_hat <- solve(t(X) %*% W %*% X)

# Display the estimated covariance matrix
print(cov_beta_hat)
# [,1]      [,2]
# [1,]  4.736239 -2.191896
# [2,] -2.191896  1.385940
```

## 09

$$y \sim N(X\beta, \sigma^2 I)$$

Now, $\hat{\mu}_f = x_f' \hat{\beta}$

$$V(e_f) = \sigma^2 (1 + x_f'(X'X)^{-1} x_f)$$

$$C.I = \hat{\mu}_f \pm t_{\alpha/2} \sqrt{\hat{\sigma}^2 (1 + x_f' (X'X)^{-1} x_f)}$$

Here, $\hat{\mu}_f$: The predicted value of new observation

$t_{\alpha/2}$: The C.V

$\hat{\sigma}^2$: Estimated variance

```r

################## Start  5  ##################

# Coefficients for RET=3
beta_03 <- -4.27817
beta_13 <- 0.178304
X_i <- 20
eta_i3 <- beta_03 + beta_13 * X_i
P_Y_i_3_given_X_i_20 <- exp(eta_i3) / (1 + exp(eta_i3))
print(P_Y_i_3_given_X_i_20) # 0.3291372

################## End  5   ##################
```

```r
##################  6  ##################
beta_0 <- 99.56799
beta_1 <- 21.61455
beta_2 <- -3.54113
x_i <- 50

mu_i <- beta_0 + (beta_1 - beta_0) * exp(-(exp(beta_2 * x_i)))
print(mu_i) # 21.61455
##################  6  ##################
```