Linear models Ex1

Roi Hezkiyahu

3 3 2022

Question 1.

1. In a two-group randomized experimental design, a response y is measured, and the model

$$y_i = \mu_i + \varepsilon_i$$

is to be fitted, where

$$\mu_i = \beta_0 + \beta_1 x_i,$$

ε_i are i.i.d. with zero means and

$$x_i = 0 \text{ in } \textit{Group 1}, \; x_i = 1 \text{ in } \textit{Group 2}.$$

Express the parameters β_0 and β_1 in terms of the group means μ_1 and μ_2.

2. The same model is fitted, but now x_i is defined by

$$x_i = -1/2 \text{ in } \textit{Group 1}, \; x_i = 1/2 \text{ in } \textit{Group 2}$$

What do parameters β_0 and β_1 represent now?

3. A third definition of x_i is

$$x_i = a \text{ in } \textit{Group 1}, \; x_i = b \text{ in } \textit{Group 2}$$

where a and b are arbitrary constants. What do parameters β_0 and β_1 represent in this general case?

Q1

a

we will imput x_1, x_2 and get:

$\mu_1 = \beta_0$

$\mu_2 = \beta_0 + \beta_1 \Rightarrow \beta_1 = \mu_2 - \mu_1$

b

$(1)\mu_1 = \beta_0 - \frac{1}{2}\beta_1$

$(2)\mu_2 = \beta_0 + \frac{1}{2}\beta_1$

$(1) + (2) \Rightarrow \mu_1 + \mu_2 = 2\beta_0 \Rightarrow \beta_0 = \frac{\mu_1+\mu_2}{2}$

$(2) - (1) \Rightarrow \mu_2 - \mu_1 = \beta_1$

c

$(3)\mu_1 = \beta_0 + a\beta_1$

$(4)\mu_2 = \beta_0 + b\beta_1$

$(3) - (4) \Rightarrow \mu_1 - \mu_2 = (a-b)\beta_1 \Rightarrow \beta_1 = \frac{\mu_1-\mu_2}{a-b}$

imput β_0 in (3):

$\mu_1 = \beta_0 + a\frac{\mu_1-\mu_2}{a-b} \Rightarrow \beta_0 = \frac{(a-b)\mu_1 - a\mu_1 + a\mu_2}{a-b} = \frac{a\mu_2 - b\mu_1}{a-b}$

Question 2.

For the regression model

$$y_i = g(\boldsymbol{x}_i) + \varepsilon_i$$

with a continuous explanatory variable x_1 and a factor x_2: $x_{2i}=0$, $i=1,\ldots,n_1$; $x_{2i}=1$, $i=n_1+1,\ldots,n_1+n_2$, give plots that represent the following linear models, where $x_3=x_1^2$

1. $g(\boldsymbol{x})=\beta_0+\beta_1 x_1+\beta_2 x_2$
2. $g(\boldsymbol{x})=\beta_0+\beta_1 x_1+\beta_3 x_3$
3. $g(\boldsymbol{x})=\beta_0+\beta_1 x_1+\beta_2 x_2+\beta_3 x_3$
4. $g(\boldsymbol{x})=\beta_0+\beta_1 x_1+\beta_3 x_3+\beta_4 x_1 x_2$
5. $g(\boldsymbol{x})=\beta_0+\beta_1 x_1+\beta_3 x_3+\beta_5 x_3 x_2$

Q2

first lets compute g for each situtaion

$$(1)\, g(x_i) = \begin{cases} \beta_0 + \beta_1 x_{1i}\,, & i \le n_1 \\ \beta_0 + \beta_2 + \beta_1 x_{1i}\,, & i > n_1 \end{cases}$$
$$(2)\, g(x_i) = \beta_0 + \beta_1 x_{1i} + \beta_3 x_{1i}^2$$
$$(3)\, g(x_i) = \begin{cases} \beta_0 + \beta_1 x_{1i} + \beta_3 x_{1i}^2\,, & i \le n_1 \\ \beta_0 + \beta_2 + \beta_1 x_{1i} + \beta_3 x_{1i}^2\,, & i > n_1 \end{cases}$$
$$(4)\, g(x_i) = \begin{cases} \beta_0 + \beta_1 x_{1i} + \beta_3 x_{1i}^2\,, & i \le n_1 \\ \beta_0 + (\beta_1 + \beta_4) x_{1i} + \beta_3 x_{1i}^2\,, & i > n_1 \end{cases}$$
$$(5)\, g(x_i) = \begin{cases} \beta_0 + \beta_1 x_{1i} + \beta_3 x_{1i}^2\,, & i \le n_1 \\ \beta_0 + \beta_1 x_{1i} + (\beta_3 + \beta_5) x_{1i}^2\,, & i > n_1 \end{cases}$$

for plotting we will make the following assumptions:

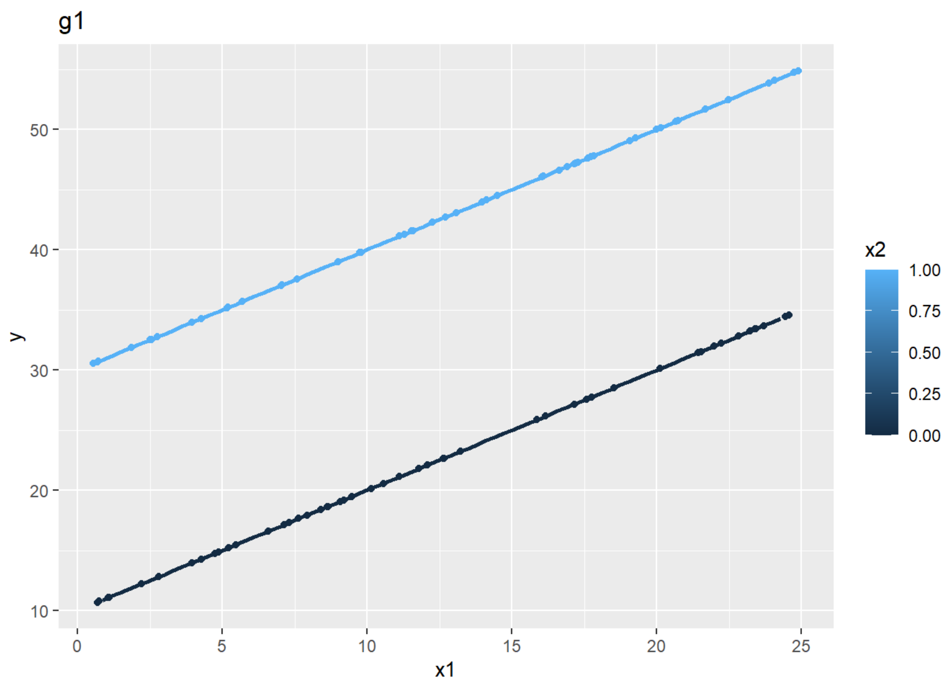
$$\beta_0 = 10$$
$$\beta_1 = 1$$
$$\beta_2 = 20$$
$$\beta_3 = 1$$
$$\beta_4 = 4$$
$$\beta_5 = 4$$
$$\varepsilon \sim N(0,1)$$
$$X_1 \sim U(0,25)$$
$$n_1 = 50$$
$$n_2 = 50$$

```
#initilaze parameters
b0 <-  10
b1 <-1
b2 <-20
b3 <-1
b4 <-4
b5 <-4
eps <- rnorm(100)
x1 <- runif(100,0,25)
x3 <- x1^2
x2 <- c(rep(0,50),rep(1,50))
g1 <- b0 +b1*x1+b2*x2
g2 <- b0 +b1*x1+b3*x3
g3 <- b0 +b1*x1+b2*x2 +b3*x3
g4 <- b0 +b1*x1+b3*x3 + b4 *x1*x2
g5 <- b0 +b1*x1+b3*x3 + b5 *x3*x2

mat <- cbind(g1,g2,g3,g4,g5)
for (i in 1:5){
  y <- mat[,i]
  tbl <-  tibble(x=x1,y=y,x2 =x2)
  #plot data
  print(ggplot(data = tbl,aes(x = x1, y = y,group = x2,color = x2)) +
    geom_point()+
    ggtitle(glue("g{i}"))+
    geom_smooth(fullrange=TRUE))
}
```

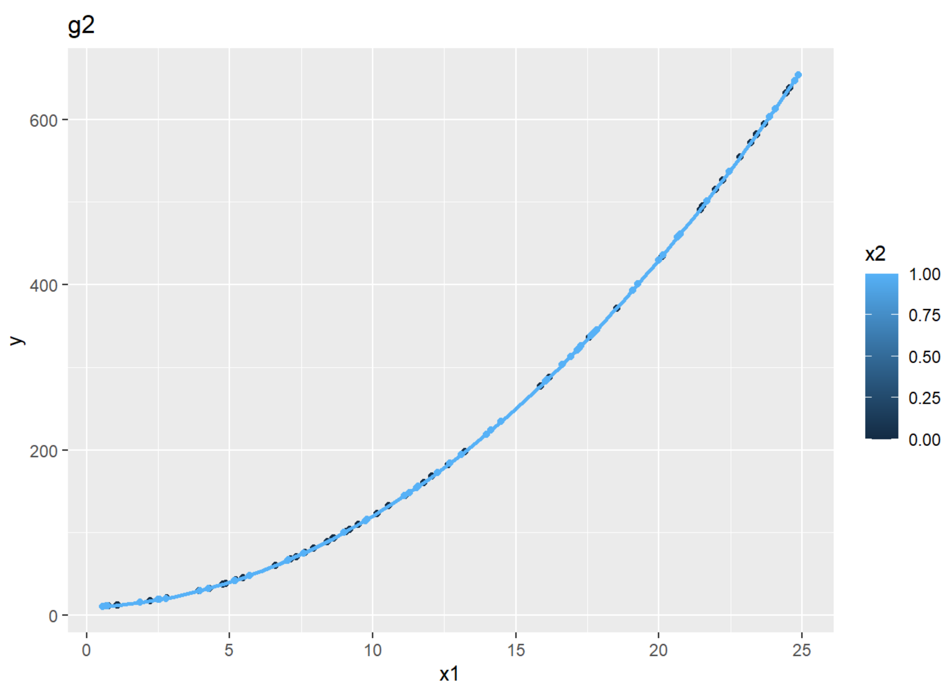
```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

```
## Warning: Removed 3 rows containing missing values (geom_smooth).
```



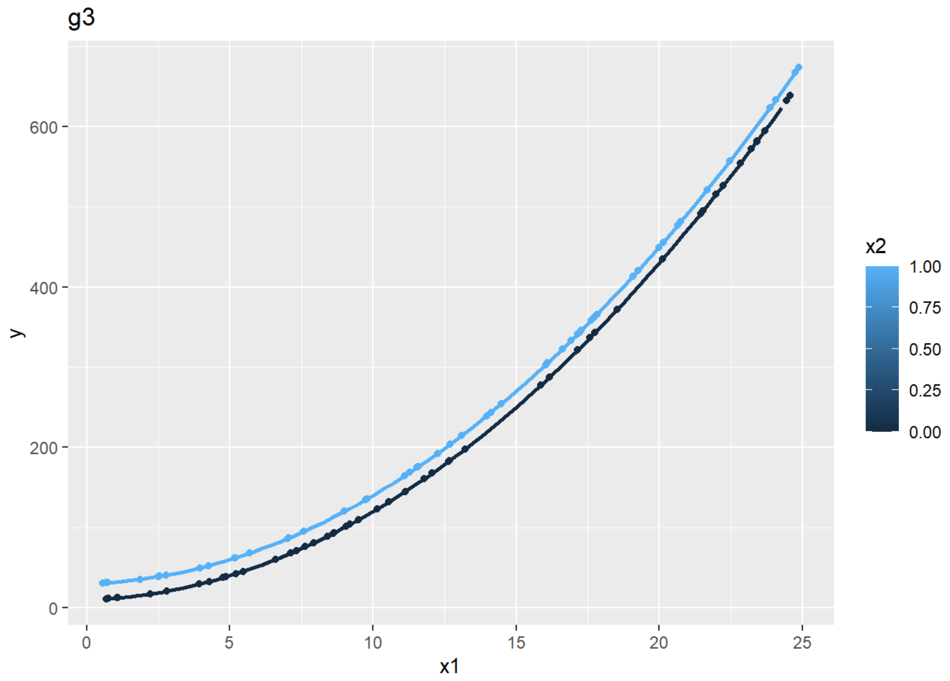

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

```
## Warning: Removed 3 rows containing missing values (geom_smooth).
```



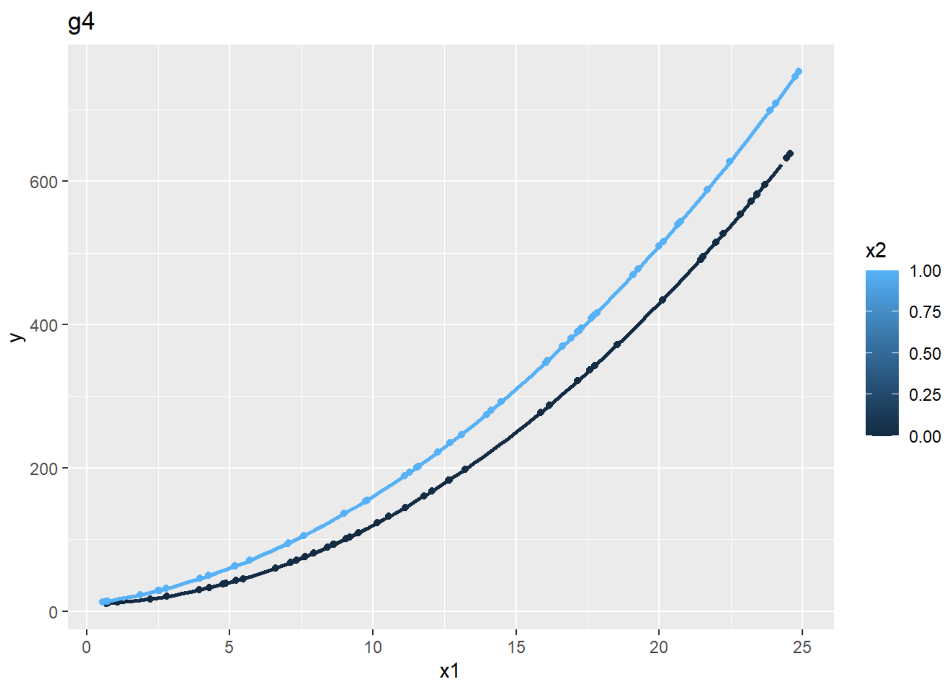

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

```
## Warning: Removed 3 rows containing missing values (geom_smooth).
```



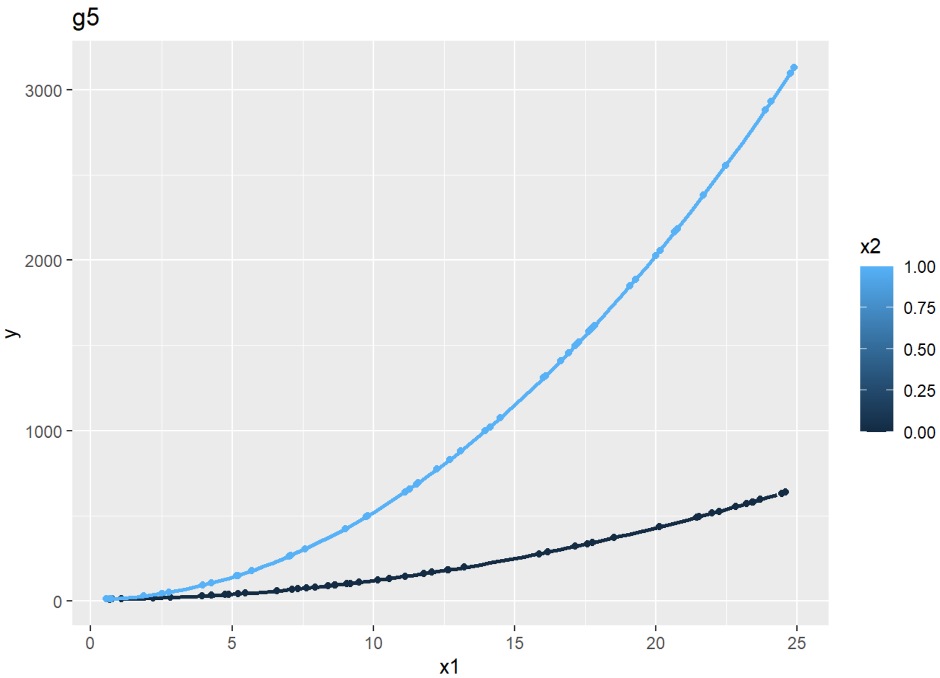

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

```
## Warning: Removed 3 rows containing missing values (geom_smooth).
```




```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

```
## Warning: Removed 3 rows containing missing values (geom_smooth).
```



Question 3.

For a simple linear regression model $y_i=\beta_0+\beta_1 x_i+\varepsilon_i$, $i=1,\dots,n$ show that the diagonal entries of the projection (hat) matrix H are

$$h_{ii} = \frac{1}{n} + \frac{(x_i-\bar{x})^2}{\sum_{j=1}^{n}(x_j-\bar{x})^2}$$

Q3

$$X = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}$$

$$X^t X = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} 1\dots1 \\ x_1\dots x_n \end{bmatrix} = \begin{bmatrix} n & \sum_i x_i \\ \sum_i x_i & \sum_i x_i^2 \end{bmatrix}$$

$$(X^t X)^{-1} = \frac{1}{n\sum_i x_i^2 - (\sum_i (x_i))^2} \begin{bmatrix} \sum_i x_i^2 & -\sum_i x_i \\ -\sum_i x_i & n \end{bmatrix} := C \begin{bmatrix} \sum_i x_i^2 & -\sum_i x_i \\ -\sum_i x_i & n \end{bmatrix}$$

$$(X^t X)^{-1} X^t = C \begin{bmatrix} \sum_i x_i^2 - x_1 \sum_i x_i & ,\dots & , & \sum_i x_i^2 - x_n \sum_i x_i \\ n x_1 - \sum_i x_i & ,\dots & , & n x_n - \sum_i x_i \end{bmatrix}$$

$$X(X^t X)^{-1} X^t_{jj} = C(\sum_i x_i^2 - x_j \sum_i x_i + n x_j^2 - x_j \sum_i x_i) = C(\sum_i x_i^2 - 2x_j \sum_i x_i + n x_j^2) = C(\sum_i x_i^2 + n(x_j^2 - 2x_j\bar{x} + \bar{x}^2 - \bar{x}^2)) =$$

$$= C(\sum_i x_i^2 - n\bar{x}^2 + n(x_j - \bar{x}^2)) = \frac{\sum_i x_i^2 - n\bar{x}^2 + n(x_j - \bar{x}^2)}{n\sum_i x_i^2 - n^2\bar{x}^2} = \frac{1}{n} + \frac{x_j - \bar{x}^2}{\sum_i x_i^2 - n\bar{x}^2} = \frac{1}{n} + \frac{x_j - \bar{x}^2}{\sum_i (x_i - x)^2}$$

Question 4.

Consider a standard linear model but with *correlated* errors: $\mathbf{y}=X\boldsymbol{\beta}+\boldsymbol{\varepsilon}$, where ε_i's have zero means and $Var(\boldsymbol{\varepsilon})=V$, i.e. $Var(\varepsilon_i)=V_{ii}$, $Cov(\varepsilon_i,\varepsilon_j)=V_{ij}$

1. Find a vector of *Weighted* Least Squares Estimators for $\boldsymbol{\beta}$ that minimizes $(\mathbf{y}-X\boldsymbol{\beta})^t V^{-1} (\mathbf{y}-X\boldsymbol{\beta})$
2. Is it unbiased estimator of $\boldsymbol{\beta}$?
3. Find a variance-covariance matrix for the vector of weighted least squares estimators.
4. What is the hat-matrix H for weighted least squares linear regression?
5. Show that if, in addition, ε_i are normal, then weighted least squares estimators are also maximum likelihood estimators.

Q4

a

$$denote\ V^{-1}\ as\ W$$
$$(y-X\beta)^tW(y-X\beta)=(y-X\beta)^y(Wy-WX\beta)=y^tWy-y^tWX\beta-\beta^tX^tWy+\beta^tX^tX\beta$$
$$\frac{\partial(y-X\beta)^tW(y-X\beta)}{\partial\beta}=-X^tWy-X^tWy+2X^tWX\beta:=0\Rightarrow$$
$$X^tWX\beta=X^tWy\Rightarrow\hat{\beta}_{WLS}=(X^tWX)^{-1}X^tWy$$

b

yes i will show that the expected value of the estimator is: β

$$E(\hat{\beta}_{WLS})=E((X^tWX)^{-1}X^tWy)=E((X^tWX)^{-1}X^tWX\beta)=E(\beta)=\beta$$

c

$$Var(\beta_{WLS})=Var((X^tWX)^{-1}X^tWy)=(X^tWX)^{-1}X^tW\,Var(y)((X^tWX)^{-1}X^tW)^t=(X^tWX)^{-1}X^tWVWX(X^tWX)^{-1}=$$
$$=(X^tWX)^{-1}X^tWX(X^tWX)^{-1}=(X^tWX)^{-1}$$

d

$$\hat{y}=Hy=X\beta_{WLS}=X(X^tWX)^{-1}X^tWy\ \ \forall y\Rightarrow H=X(X^tWX)^{-1}X^tW$$

e

$$\varepsilon_i\sim N(0,\sigma_i^2)\Rightarrow\vec{y}\sim N(X\beta,V)$$
$$L(\beta,V,\vec{y})=\frac{1}{\sqrt{(2\pi)^n|V|}}e^{-\frac{1}{2}(y-X\beta)^tW(y-X\beta)}$$
$$log(L(\beta,V,\vec{y}))=C_V-\frac{(y-X\beta)^tW(y-X\beta)}{2}$$
$$where\,C_v\,is\,a\,constant\,given\,V$$
$$\beta_{WLS}=argmin_\beta((y-X\beta)^tW(y-X\beta))=argmax_\beta(-(y-X\beta)^tW(y-X\beta))=argmax_\beta(log(L(\beta,V,\vec{y})))=argmax_\beta(L(\beta,V,\vec{y}))=$$