

# Ejercicio 4

*Ricardo Lastra*

*4 de octubre de 2017*

EJERCICIO 4. Un investigador desea evaluar la relación entre el salario anual de trabajadores de una compañía de nivel medio y alto (Y, en miles de dólares) y el índice de calidad de trabajo (X1), número de años de experiencia (X2) y el índice de éxito en publicaciones (X3). La muestra consiste de 24 trabajadores. Realiza un análisis Bayesiano completo de los datos y obtén las predicciones de salarios para 3 nuevos empleados con variables explicativas:

X1F {5.4,17,6.0} X2F{6.2,12,5.8} X3F{6.4,21,6.1}

```
#By Ricardo Lastra
```

```
library(R2jags)
```

```
## Loading required package: rjags
```

```
## Loading required package: coda
```

```
## Linked to JAGS 4.3.0
```

```
## Loaded modules: basemod,bugs
```

```
##
```

```
## Attaching package: 'R2jags'
```

```
## The following object is masked from 'package:coda':
```

```
##
```

```
##      traceplot
```

```
wdir<-"c:/Users/FORANEA110/Desktop/REGRESION_AVANZADA"
```

```
setwd(wdir)
```

```
#--- Ejemplo 4 ---
```

```
#-Reading data-
```

```
salario<-read.table("http://allman.rhon.itam.mx/~lnieto/index_archivos/salarios.txt",header=TRUE)
```

```
n<-nrow(salario)
```

```
str(salario)
```

```
## 'data.frame':   24 obs. of  4 variables:
```

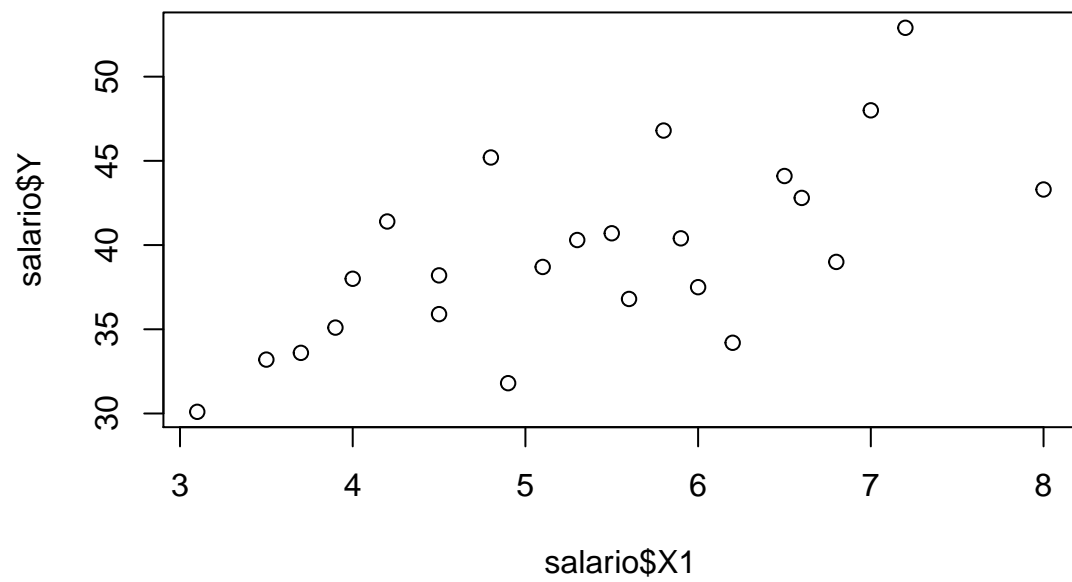
```
## $ Y : num  33.2 40.3 38.7 46.8 41.4 37.5 39 40.7 30.1 52.9 ...
```

```
## $ X1: num  3.5 5.3 5.1 5.8 4.2 6 6.8 5.5 3.1 7.2 ...
```

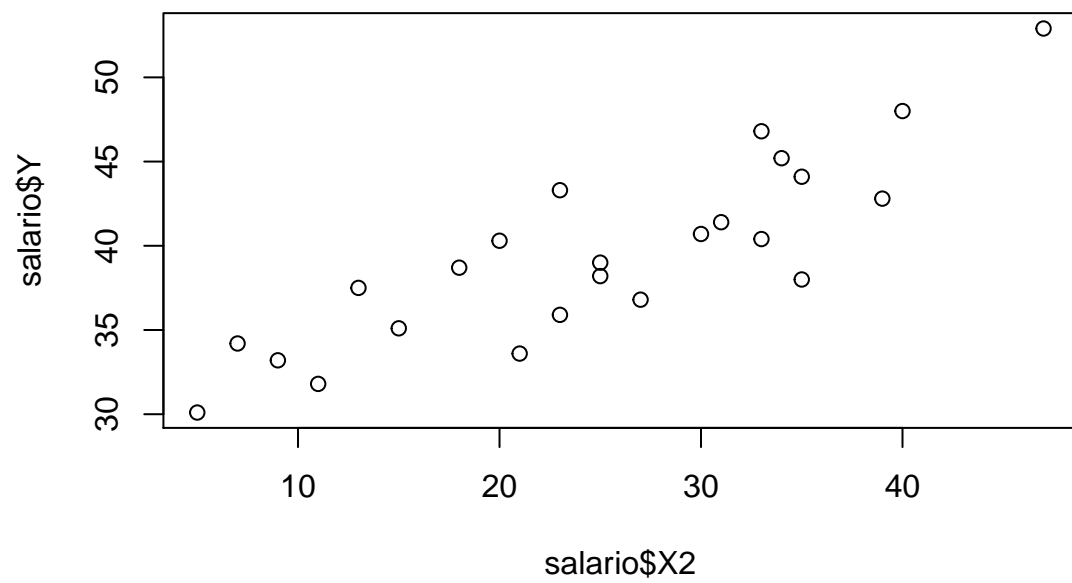
```
## $ X2: num  9 20 18 33 31 13 25 30 5 47 ...
```

```
## $ X3: num  6.1 6.4 7.4 6.7 7.5 5.9 6 4 5.8 8.3 ...
```

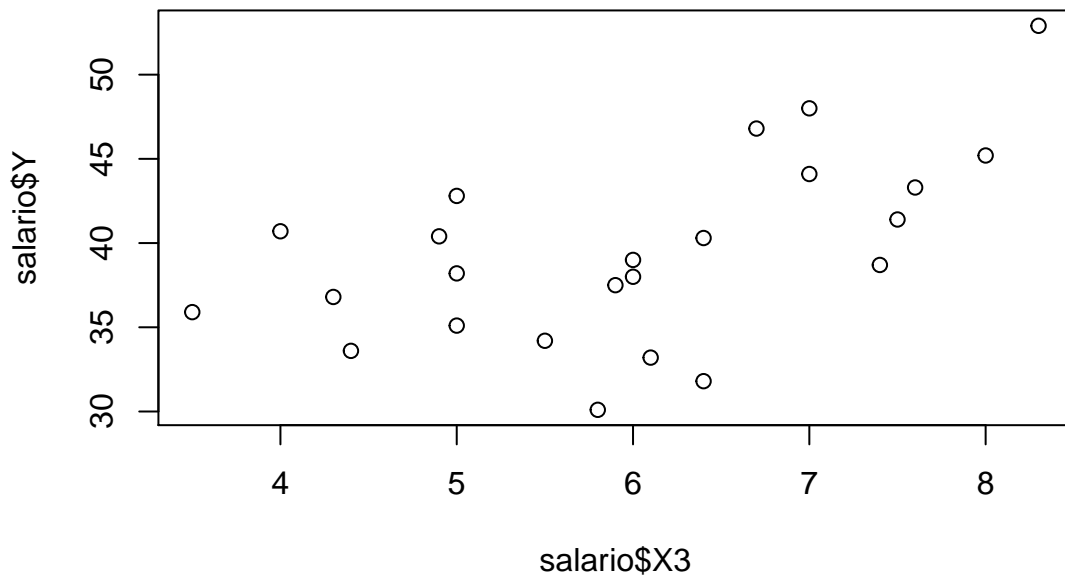
```
plot(salario$X1,salario$Y)
```



```
plot(salario$X2,salario$Y)
```



```
plot(salario$X3,salario$Y)
```



```
x<- data.frame(x1=salario$X1,x2=salario$X2,x3=salario$X3)
```

```
#-Defining data-
```

```
data<-list("n"=n,"y"=salario$Y,"x"=x)
```

```
#-Defining inits-
```

```
inits<-function(){list(beta=rep(0,3),tau=1,yf=rep(0,n))}
```

```
#-Selecting parameters to monitor-
```

```
parameters<-c("beta","tau","yf")
```

Corremos JAGS con 10,000 iteraciones:

```
library(R2jags)
```

```
#JAGS
```

```
ej4.sim<-jags(data,inits,parameters,model.file="C:/Users/FORANE110/Desktop/REGRESION_AVANZADA/Ej4.txt"
n.iter=10000,n.chains=1,n.burnin=1000,n.thin=1)
```

```
## module glm loaded
```

```
## Compiling model graph
```

```
##   Resolving undeclared variables
```

```
##   Allocating nodes
```

```
## Graph information:
```

```
##   Observed stochastic nodes: 24
```

```
##   Unobserved stochastic nodes: 29
```

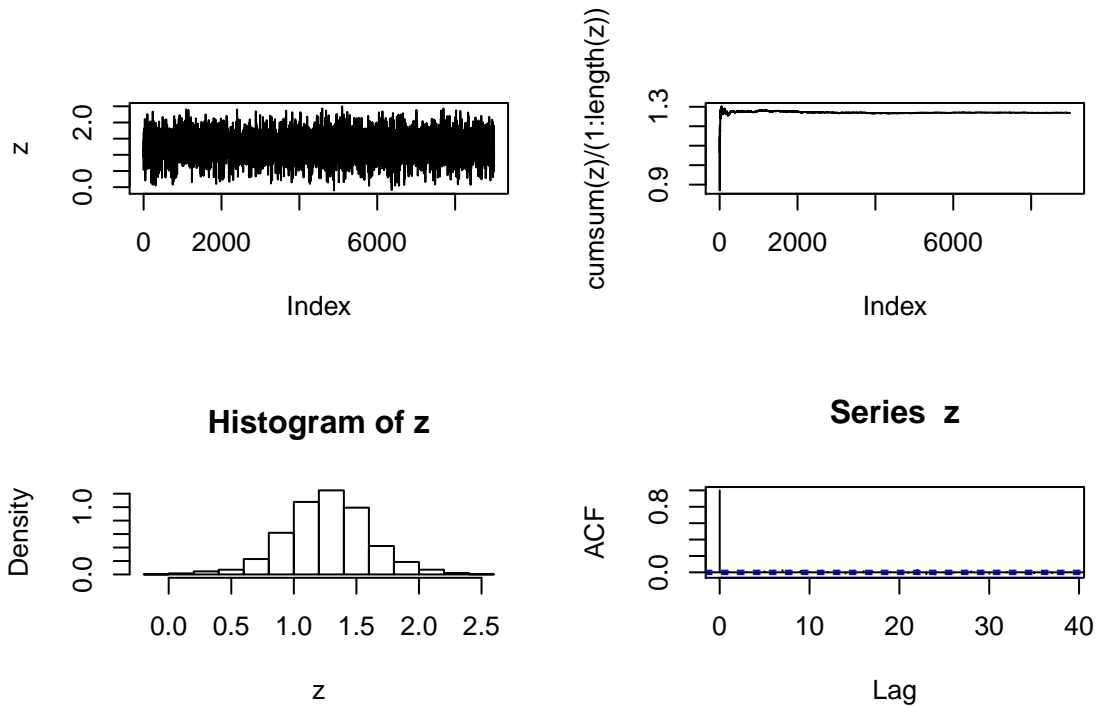
```
##   Total graph size: 215
```

```
##
```

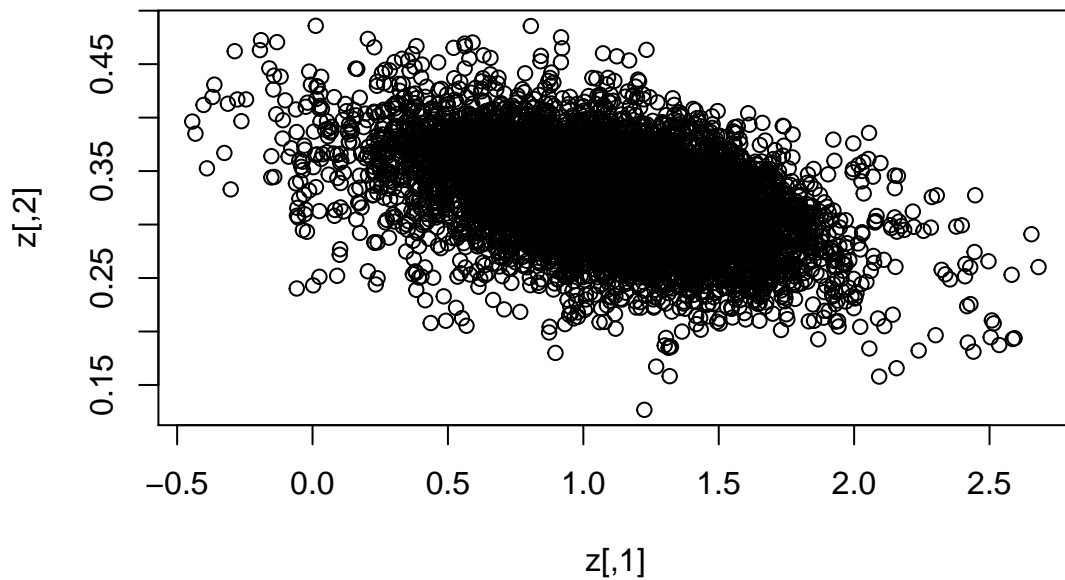
```
## Initializing model
```

```
#JAGS
out<-ej4.sim$BUGSoutput$sims.list

z<-out$beta[,3]
par(mfrow=c(2,2))
plot(z,type="l")
plot(cumsum(z)/(1:length(z)),type="l")
hist(z,freq=FALSE)
acf(z)
```



```
z<-out$beta
par(mfrow=c(1,1))
plot(z)
```



```
#JAGS
```

```
out.sum<-ej4.sim$BUGSoutput$summary
```

```
out.dic<-ej4.sim$DIC
```

```
out.dic<-ej4.sim$BUGSoutput$DIC
```

```
print(out.dic) ##### EL MODELO CON MENOR DIC AJUSTA MEJOR#####
```

```
## [1] 104.7435
```

Predicciones:

```
out.yf<-out.sum[grep("yf",rownames(out.sum)),]
```

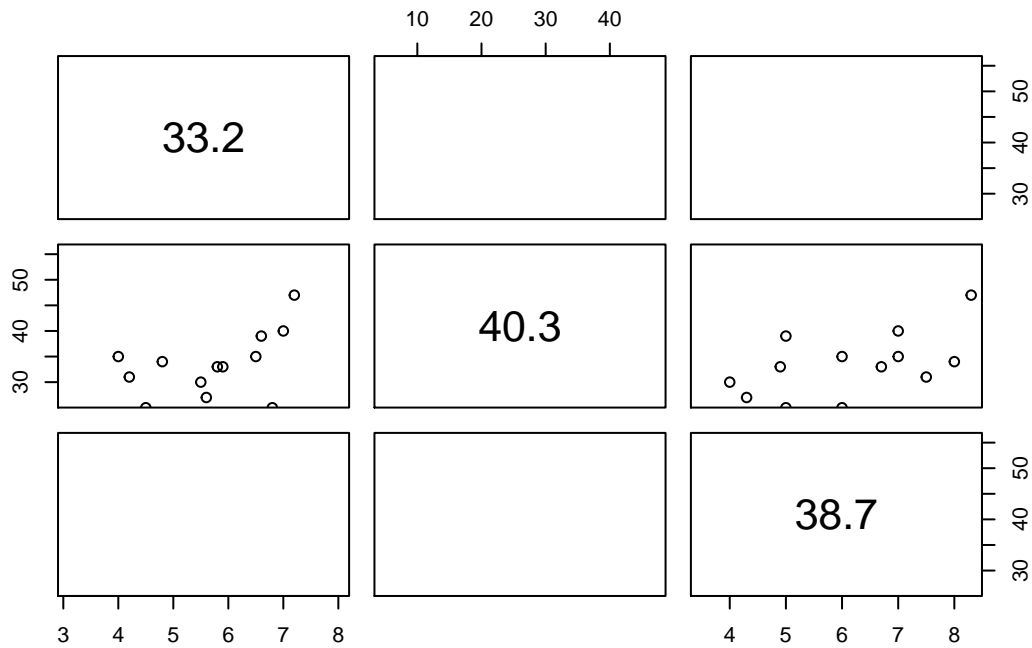
```
or<-order(x)
```

```
ymin<-min(salario$Y,out.yf[,c(1,3,7)])
```

```
ymax<-max(salario$Y,out.yf[,c(1,3,7)])
```

```
par(mfrow=c(1,1))
```

```
plot(x,salario$Y,ylim=c(ymin,ymax))
```



*To Do:*

**Me encuentre con la problematica de: Como pasar el nuevo vector de observaciones para obtener la Y objetivo?**

Asi mismo, despues de correr un modelo adicional, como podriamos poner ambos modelos en un solo comparativo para evaluar el mejor?, situación que se podria resolver con una tabla sencilla, sin embargo creo que podrian ser utiles otras tecnicas estadisticas.