Modèles Linéaires Appliqués

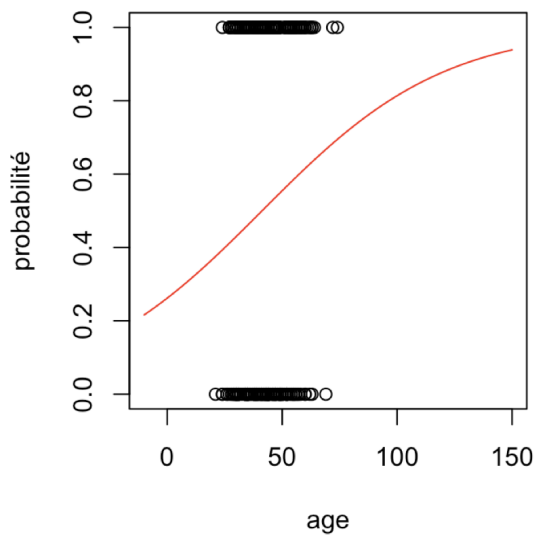
Arthur Charpentier

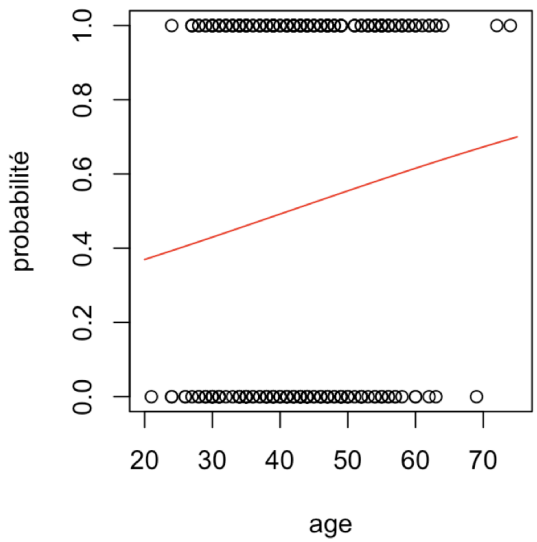
Automne 2Q20

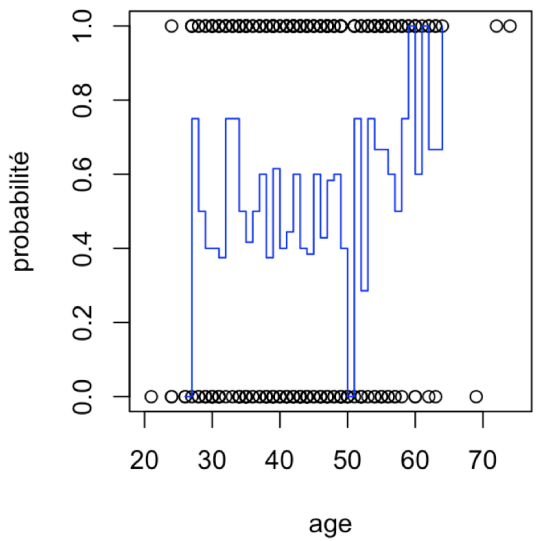
GLM #22 (example)

Numérique vs. Facteur

```
> n=rpois(1,250)
> x = c(round(21+rbeta(n-1,3,6)*67),21)
> p = exp(-1+x/40)/(1+exp(-1+x/40))
> df = data.frame(prime=factor(rbinom(n,size=1,prob=p),labels=c("non","oui")),
    age=x)
> reg = glm(prime~age,data=df,family=binomial)
> age_prime=55
> predict(reg,newdata=data.frame(age = age_prime),type="response")
        1
0.6193646
> age_prime =seq(20,75)
> proba = predict(reg,newdata=data.frame(age = age_prime),type="response")
> plot(df$age,as.numeric(df$prime)-1)
> lines(age_prime,proba,col="red")
```







Numérique vs. Facteur

```
> df[df$age == 55,]
    prime age
32     non  55
80     oui  55
83     oui  55
205    non  55
217    oui  55
> mean(df[df$age == 55,"prime"] == "oui")
[1] 0.6
> regf = glm(prime~as.factor(age),data=df,family=binomial)
> age_prime =sort(unique(df$age))
> proba = predict(regf,newdata=data.frame(age = age_prime),type="response")
> y = rep(NA,75)
> y[age_prime]=proba
> plot(df$age,as.numeric(df$prime)-1)
> lines(1:75,y,type="s",col="blue")
> predict(regf,newdata=data.frame(age = 55),type="response")
  1
0.6
```