

## EXAMEN FINAL, ACT2040, AUTOMNE 2013

### ---- PARTIE 0 ----

```
> preambule=read.table(
+ "http://freakonometrics.free.fr/preambule.csv",header=TRUE,sep=";")
> table(preambule$Y)
```

```
 0  1  2  3  4  5  6
45 133 160 101 51  8  2
```

```
> reg0=glm(Y/N~1,family="binomial",weights=N,data=preambule)
> summary(reg0)
```

```
Call:
glm(formula = Y/N ~ 1, family = "binomial", data = preambule,
     weights = N)
```

```
Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.12673 -0.87408 -0.01892  0.73065  2.74209
```

```
Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.3714      0.0352  -38.96  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

(Dispersion parameter for binomial family taken to be 1)

```
Null deviance: 512.87 on 499 degrees of freedom
Residual deviance: 512.87 on 499 degrees of freedom
AIC: 1586.2
```

Number of Fisher Scoring iterations: 4

### ---- PARTIE 1 ----

```
> CORPOREL=read.table(
+ "http://freakonometrics.free.fr/corporel-2040.csv",
+ header=TRUE,sep=";")
> tail(CORPOREL)
      degre age cat.age sexe vehicule anciennete alcool cat.alc
76336 indemne 45 40-49  M  voiture          6      0  0-20
76337 corporel 59 50-59  F  voiture          2      0  0-20
76338 indemne 34 30-39  F  voiture          2      0  0-20
76339 indemne 29 26-29  F  voiture          5      9  0-20
76340 indemne 64 60+   M  voiture          0      0  0-20
76341 indemne 57 50-59  F  voiture          1      0  0-20
```

```
> attach(CORPOREL)
> table(degre)
degre
corporel  deces  indemne
 31369    676  44296
```

```
> Y=degre=="deces"
> table(Y)
Y
FALSE TRUE
75665  xxx
```



```
> X1=vehicule; nom1=levels(X1)
> X2=cat.alc; nom2=levels(X2)
> comptage=table(X1,X2)
> deces=comptage
> for(k in 1:nrow(comptage)){
+ deces[k,]=tapply(Y[X1==nom1[k]],X2[X1==nom1[k]],sum)}
> deces[is.na(deces)]=0
> comptage
```

	X2				
X1	0-20	150+	20-50	50-80	80-150
bus-truck	3218	74	0	13	52
moto	2059	49	5	11	60
van	6237	120	8	32	113
voiture	62433	795	56	244	762

```

> deces
X1      X2
bus-truck  93  6  0  1  2
moto       51  4  0  2  4
van        76  7  0  2  3
voiture   372 25  0 11 17

> taux=deces/comptage
> taux
X1      X2
bus-truck 0.028899938 0.081081081 0.076923077 0.038461538
moto      0.024769305 0.081632653 0.000000000 0.181818182 0.066666667
van       0.012185346 0.058333333 0.000000000 0.062500000 0.026548673
voiture   0.005958387 0.031446541 0.000000000 0.045081967 0.022309711

> comptage[is.na(comptage)]=0
> m=mean(Y)
> xxxxxx
> L<-matrix(NA,10,nrow(deces));C<-matrix(NA,10,ncol(deces))
> colnames(L)=nom1;colnames(C)=nom2
> C[1,]<-m
> for(j in 2:10){
+   for(k in 1:nrow(deces)){
+     L[j,k]<-sum(deces[k,])/sum(comptage[k,]*C[j-1,]) }
+   for(k in 1:ncol(deces)){
+     C[j,k]<-sum(deces[,k])/sum(comptage[,k]*L[j,]) }
+ }
> L[10,]
bus-truck      moto      van      voiture
3.3578117 3.0102805 1.4996518 0.7585497

> C[10,]
      0-20      150+      20-50      50-80      80-150
0.008030879 0.035623800 0.000000000 xxxxxx 0.023578519

> pred1 = deces
> for(k in 1:nrow(deces)){pred1[k,]<-L[10,k]*C[10,]}

```

3

```

> pred1
X1      X2
bus-truck 0.026966178 0.119618012 0.000000000 xxxxxx 0.079172227
moto      0.024175198 0.107237631 0.000000000 xxxxxx 0.070977956
van       0.012043522 0.053423297 0.000000000 xxxxxx 0.035359569
voiture   0.006091821 0.027022425 0.000000000 xxxxxx 0.017885480

> reg1=glm(Y~vehicule+cat.alc,family=poisson(link="log"),data=CORPOREL)
> summary(reg1)

Call:
glm(formula = Y ~ vehicule + cat.alc, family = poisson(link = "log"),
    data = CORPOREL)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.5889 -0.1104 -0.1104 -0.1104  2.8660

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)   -3.6132     0.1006  -35.924 < 2e-16 ***
vehiculemoto   -0.1093     0.1620   -0.675  0.500
vehiculevan    -0.8061     0.1455  -5.539 3.04e-08 ***
vehiculevoiture -1.4876     0.1104 -13.472 < 2e-16 ***
cat.alc150+     1.4897     0.1600   9.308 < 2e-16 ***
cat.alc20-50   -10.4584    151.4947  -0.069  0.945
cat.alc50-80   xxxxx 6 xxxxx 7 7.344 xxxxx 8
cat.alc80-150  1.0770     0.2007   5.365 8.08e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

Null deviance: 6390.6 on 76340 degrees of freedom
Residual deviance: 6064.0 on 76333 degrees of freedom
AIC: 7432

Number of Fisher Scoring iterations: 13

> newd=data.frame(vehicule=rep(nom1,length(nom2)),

```

4

```

+ cat.alc=rep(nom2,each=length(nom1))
> pred2=predict(reg1,newdata=newd,type="response")
> P2=matrix(pred2,length(nom1),length(nom2))
> rownames(P2)=nom1;colnames(P2)=nom2
> table(CORPOREL$cat.alc)

 0-20  150+  20-50  50-80  80-150
73947  1038    69    300    987

> CORPOREL$cat.alc2=CORPOREL$cat.alc
> levels(CORPOREL$cat.alc2)=c("0-50","150+","0-50","50-150","50-150")
> table(CORPOREL$cat.alc2)

 0-50  150+  50-150
74016  1038  1287

> table(CORPOREL$vehicule)

bus-truck      moto      van      voiture
      3357      2184      6510      64290

> CORPOREL$veh2=CORPOREL$vehicule
> levels(CORPOREL$veh2)=c("bus-truck-moto",
+ "bus-truck-moto","van","voiture")
> table(CORPOREL$veh2)

bus-truck-moto      van      voiture
      5541      6510      64290

> reg2=glm(Y~veh2+cat.alc2,family=poisson(link="log"),data=CORPOREL)

```

```

> summary(reg2)
Call:
glm(formula = Y ~ veh2 + cat.alc2, family = poisson(link = "log"),
    data = CORPOREL)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.4783 -0.1104 -0.1104 -0.1104  2.8658

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -3.65936    0.08069  -45.351 < 2e-16 ***
veh2van      -0.76075    0.13230   -5.750 8.93e-09 ***
veh2voiture -1.44099    0.09242  -15.592 < 2e-16 ***
cat.alc2150+ 1.49099    0.16005    9.316 < 2e-16 ***
cat.alc250-150 1.30600    0.15991    8.167 3.15e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 6390.6  on 76340  degrees of freedom
Residual deviance: 6071.2  on 76336  degrees of freedom
AIC: 7433.2

Number of Fisher Scoring iterations: 7

> predict(reg2,newdata=data.frame(cat.alc2=c("0-50","50-150","150+"),
+ veh2=c("voiture","voiture","voiture")),
+ type="response")
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
> reg3=glm(Y~veh2+cat.alc2,family=binomial(link="logit"),data=CORPOREL)

```

```
> summary(reg3)
```

9b

```
Call:
glm(formula = Y ~ veh2 + cat.alc2, family = binomial(link = "logit"),
    data = CORPOREL)
```

```
Deviance Residuals:
```

Min	1Q	Median	3Q	Max
-0.4832	-0.1104	-0.1104	-0.1104	3.1946

```
Coefficients:
```

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	-3.62662	0.08184	-44.311	< 2e-16 ***
veh2van	-0.78178	0.13395	-5.836	5.34e-09 ***
veh2voiture	-1.46987	0.09370	-15.688	< 2e-16 ***
cat.alc2150+	1.53780	0.16450	9.348	< 2e-16 ***
cat.alc250-150	1.34111	0.16351	8.202	2.36e-16 ***

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
(Dispersion parameter for binomial family taken to be 1)
```

```
Null deviance: 7736.6 on 76340 degrees of freedom
Residual deviance: 7411.8 on 76336 degrees of freedom
AIC: 7421.8
```

```
Number of Fisher Scoring iterations: 7
```

```
> predict(reg3,newdata=data.frame(cat.alc2=c("0-50","50-150","150+"),
+ veh2=c("voiture","voiture","voiture")),
+ type="response")
```

11

```
XXXXXXXXXXXXXXXXXXXXXXXXXXXX
```

```
> reg4=glm(Y~veh2+cat.alc2,family=quasipoisson(link="log"),data=CORPOREL)
> summary(reg4)
```

12

```
Call:
glm(formula = Y ~ veh2 + cat.alc2, family = quasipoisson(link = "log"),
    data = CORPOREL)
```

```
Deviance Residuals:
```

Min	1Q	Median	3Q	Max
-0.4783	-0.1104	-0.1104	-0.1104	2.8658

```
Coefficients:
```

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-3.65936	0.07991	-45.794	< 2e-16 ***
veh2van	-0.76075	0.13102	-5.806	6.42e-09 ***
veh2voiture	-1.44099	0.09152	-15.745	< 2e-16 ***
cat.alc2150+	1.49099	0.15850	9.407	< 2e-16 ***
cat.alc250-150	1.30600	0.15836	8.247	< 2e-16 ***

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
(Dispersion parameter for quasipoisson family taken to be XXXXXX)
```

```
Null deviance: 6390.6 on 76340 degrees of freedom
Residual deviance: 6071.2 on 76336 degrees of freedom
AIC: NA
```

```
Number of Fisher Scoring iterations: 7
```

```
> library(nnet)
> CORPOREL$Y=degre
> reg5=multinom(Y~veh2+cat.alc2,data=CORPOREL)
# weights:  18 (10 variable)
initial value 83869.160729
iter  10 value 56945.564900
iter  20 value 54368.409072
iter  30 value 54349.196650
final value 54348.927382
converged
```

```

> summary(reg5)
Call:
multinom(formula = Y ~ veh2 + cat.alc2, data = CORPOREL)

Coefficients:
      (Intercept)  veh2van veh2voiture cat.alc2150+ cat.alc250-150
deces      -3.128271 -0.3825518 -1.0360780  1.1238657  1.0381024
indemne    -0.482676  0.8660091  0.9230502  -0.9679921  -0.6471955

Std. Errors:
      (Intercept)  veh2van veh2voiture cat.alc2150+ cat.alc250-150
deces    0.08209712 0.1350718 0.09420593 0.16545337 0.16485982
indemne  0.02827523 0.0379941 0.02937917 0.06732754 0.05821281

Residual Deviance: 108697.9
AIC: 108717.9

```

13

```

> reg6=multinom(Y~veh2+cat.alc2+sexe+anciennete,data=CORPOREL)
# weights: 24 (14 variable)
initial value 83869.160729
iter 10 value 60708.059345
iter 20 value 54354.056598
iter 30 value 54230.746815
iter 30 value 54230.746310
final value 54230.746310
converged

```

```

> summary(reg6)
Call:
multinom(formula = Y ~ veh2 + cat.alc2 + sexe + anciennete, data =
CORPOREL)

Coefficients:
      (Intercept)  veh2van veh2voiture cat.alc2150+ cat.alc250-150  sexeM  anciennete
deces      -3.4747251 -0.3657135 -0.9120289  1.082679  0.9816627 0.3509260 0.003250655
indemne    -0.7231899  0.8768508  1.0129565  -1.002271  -0.6915642 0.2443382 0.002215666

Std. Errors:
      (Intercept)  veh2van veh2voiture cat.alc2150+ cat.alc250-150  sexeM  anciennete
deces    0.12920722 0.13518641 0.09909299 0.16594222 0.16562513 0.09835261 0.009798542
indemne  0.03285738 0.03802368 0.03000948 0.06747606 0.05838016 0.01612267 0.001885893

Residual Deviance: 108461.5
AIC: 108489.5

```

14

---- PARTIE 2 ----

```
> source("http://freakonometrics.free.fr/triangle-intra2.R")
> intra
$triangle
```

	0	1	2	3	4	5	6	7	8	9
1988	5244	9228	10823	11352	11791	12082	12120	12199	12215	12215
1989	5984	9939	11725	12346	12746	12909	13034	13109	13113	NA
1990	7452	12421	14171	14752	15066	15354	15637	15720	NA	NA
1991	7115	11117	12488	13274	13662	13859	13872	XXXXX	NA	NA
1992	5753	8969	9917	10697	11135	11282	NA	NA	NA	NA
1993	3937	6524	7989	8543	8757	NA	NA	NA	NA	NA
1994	5127	8212	8976	9325	NA	NA	NA	NA	NA	NA
1995	5046	8006	8984	NA	NA	NA	NA	NA	NA	NA
1996	5129	8202	NA	NA	NA	NA	NA	NA	NA	NA
1997	3689	NA	NA	NA	NA	NA	NA	NA	NA	NA

15

```
$prime
 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997
15883 16689 18029 17858 16709 14212 15083 15131 15465 11217
```

```
> mC=intra$triangle
> mY=mC
> mY["1990", "2"]
XXXXX
```

16

```
> n=ncol(mC)
> MackChainLadder(mC)
MackChainLadder(Triangle = mC)
```

	Latest	Dev.To.Date	Ultimate	IBNR	Mack.S.E	CV(IBNR)
1988	12,215	1.000	12,215	0.0	0.0	NaN
1989	13,113	1.000	13,113	0.0	7.9	Inf
1990	15,720	0.999	15,732	12.4	15.7	1.262
1991	13,872	0.993	13,964	91.6	17.3	0.189
1992	11,282	0.985	11,453	170.7	111.9	0.656
1993	8,757	0.969	9,039	282.4	112.7	0.399
1994	9,325	0.940	9,923	598.2	148.2	0.248
1995	8,984	0.891	10,088	1,104.0	219.2	0.199
1996	8,202	0.779	10,529	2,326.9	473.3	0.203
1997	3,689	0.479	7,704	4,014.6	557.8	0.139

11

```
Totals
Latest: 105,159.00
Dev: 0.92
Ultimate: 113,759.72
IBNR: 8,600.72
Mack S.E.: 859.63
CV(IBNR): 0.10
```

```
> mY[,2:n]=mC[,2:n]-mC[,1:(n-1)]
> mY0=mY[, -n]
> Y=as.vector(mY0)
> futur=is.na(Y)
> A=rep(1988:1997,n-1)
> B=rep(0:(n-2),each=n)
> df=data.frame(Y,A,B,futur)
```

```
> reg1=lm(log(Y)~A+B,data=df)
> summary(reg1)
```

17

```
Call:
lm(formula = log(Y) ~ A + B, data = df)
```

```
Residuals:
      Min       1Q   Median       3Q      Max
-1.71387 -0.19797  0.06115  0.20978  1.29746
```

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) 146.48471   61.81314   2.370  0.0216 *
A            -0.06916    0.03102  -2.229  0.0302 *
B            -0.74997    0.03244 -23.116 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 0.4881 on 51 degrees of freedom
(36 observations deleted due to missingness)
Multiple R-squared:  0.9257, Adjusted R-squared:  0.9228
F-statistic: 317.6 on 2 and 51 DF, p-value: < 2.2e-16
```

12

```
> reg2=lm(log(Y)~as.factor(A)+as.factor(B),data=df)
> summary(reg2)
```

18

```
Call:
lm(formula = log(Y) ~ as.factor(A) + as.factor(B), data = df)
```

```
Residuals:
    Min       1Q   Median       3Q      Max
-1.37436 -0.15434  0.00522  0.17412  1.22991
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	8.8241	0.2118	41.667	< 2e-16 ***
as.factor(A)1989	-0.0580	0.2067	-0.281	0.7806
as.factor(A)1990	0.2258	0.2162	1.045	0.3032
as.factor(A)1991	-0.2504	0.2265	-1.105	0.2763
as.factor(A)1992	-0.1846	0.2388	-0.773	0.4446
as.factor(A)1993	-0.3847	0.2544	-1.512	0.1392
as.factor(A)1994	-0.4954	0.2756	-1.798	0.0806
as.factor(A)1995	-0.3722	0.3070	-1.212	0.2333
as.factor(A)1996	-0.3029	0.3610	-0.839	0.4069
as.factor(A)1997	-0.6110	0.4870	-1.255	0.2176
as.factor(B)1	-0.4694	0.2067	-2.271	0.0292 *
as.factor(B)2	-1.4816	0.2162	-6.853	5.12e-08 ***
as.factor(B)3	-2.2937	0.2265	-10.126	4.44e-12 ***
as.factor(B)4	-2.8431	0.2388	-11.905	4.84e-14 ***
as.factor(B)5	-3.4300	0.2544	-13.482	1.22e-15 ***
as.factor(B)6	-4.6344	0.2756	-16.817	< 2e-16 ***
as.factor(B)7	-4.5115	0.3070	-14.695	< 2e-16 ***
as.factor(B)8	-6.7157	0.3610	-18.604	< 2e-16 ***

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 0.4385 on 36 degrees of freedom
(36 observations deleted due to missingness)
Multiple R-squared:  0.9577, Adjusted R-squared:  0.9377
F-statistic: 47.89 on 17 and 36 DF,  p-value: < 2.2e-16
```

```
> reg3=glm(Y~A+B,data=df,family=poisson(link="log"))
```

```
> summary(reg3)
```

19

```
Call:
glm(formula = Y ~ A + B, family = poisson(link = "log"), data = df)
```

```
Deviance Residuals:
    Min       1Q   Median       3Q      Max
-23.9887  -6.8467   0.2444   4.5700  28.4783
```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	111.387202	2.323204	47.95	<2e-16 ***
A	-0.051564	0.001166	-44.21	<2e-16 ***
B	-0.697814	0.002702	-258.27	<2e-16 ***

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

(Dispersion parameter for poisson family taken to be 1)

```
Null deviance: 120911.7 on 53 degrees of freedom
Residual deviance: 5004.5 on 51 degrees of freedom
(36 observations deleted due to missingness)
AIC: 5466.6
```

Number of Fisher Scoring iterations: 4

```
> reg4=glm(Y~as.factor(A)+as.factor(B),
+ data=df,family=poisson(link="log"))
```

```
> summary(reg4)
```

20

```
Call:
glm(formula = Y ~ as.factor(A) + as.factor(B), family = poisson(link =
"log"),
     data = df)
```

```
Deviance Residuals:
```

Min	1Q	Median	3Q	Max
-12.2455	-3.5148	-0.4767	3.5907	13.4573

```
Coefficients:
```

	Estimate	Std. Error	z value	Pr(> z )	
(Intercept)	8.674092	0.009635	900.279	< 2e-16	***
as.factor(A)1989	0.070939	0.012575	5.641	1.69e-08	***
as.factor(A)1990	0.253059	0.012063	20.978	< 2e-16	***
as.factor(A)1991	0.133791	0.012415	10.777	< 2e-16	***
as.factor(A)1992	-0.064441	0.013070	-4.930	8.21e-07	***
as.factor(A)1993	-0.301073	0.014023	-21.470	< 2e-16	***
as.factor(A)1994	-0.207792	0.013788	-15.070	< 2e-16	***
as.factor(A)1995	-0.191317	0.013960	-13.705	< 2e-16	***
as.factor(A)1996	-0.148546	0.014393	-10.320	< 2e-16	***
as.factor(A)1997	-0.460981	0.019076	-24.165	< 2e-16	***
as.factor(B)1	-0.467200	0.007149	-65.353	< 2e-16	***
as.factor(B)2	-1.456867	0.010717	-135.937	< 2e-16	***
as.factor(B)3	-2.276393	0.016140	-141.038	< 2e-16	***
as.factor(B)4	-2.802747	0.021910	-127.921	< 2e-16	***
as.factor(B)5	-3.378022	0.030769	-109.787	< 2e-16	***
as.factor(B)6	-4.050147	0.046987	-86.198	< 2e-16	***
as.factor(B)7	-4.418412	0.065228	-67.738	< 2e-16	***
as.factor(B)8	-6.407605	0.223720	-28.641	< 2e-16	***

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
(Dispersion parameter for poisson family taken to be 1)
```

```
Null deviance: 120911.7 on 53 degrees of freedom
Residual deviance: 1558.2 on 36 degrees of freedom
(36 observations deleted due to missingness)
AIC: 2050.4
```

```
Number of Fisher Scoring iterations: 5
```

21

```
> sum(exp(predict(reg1,newdata=df)[futur]))
[1] 7643.561
> sum(exp(predict(reg2,newdata=df)[futur]))
[1] 8060.825
> sum(exp(predict(reg3,newdata=df)[futur]))
[1] 9177.528
> sum(exp(predict(reg4,newdata=df)[futur]))
[1] 8600.721
>
> mP=intra$prime
> df$P=rep(mP,n-1)
>
> reg5=glm(Y~as.factor(A)+as.factor(B)+offset(log(P)),
+ data=df, family=poisson(link="log"))
```



```
> summary(reg5)
```

22

```
Call:
glm(formula = Y ~ as.factor(A) + as.factor(B) + offset(log(P)),
     family = poisson(link = "log"), data = df)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-12.2455	-3.5148	-0.4767	3.5907	13.4573

Coefficients:

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	-0.998913	0.009635	-103.677	< 2e-16 ***
as.factor(A)1989	0.021439	0.012575	1.705	0.0882 .
as.factor(A)1990	0.126327	0.012063	10.472	< 2e-16 ***
as.factor(A)1991	0.016589	0.012415	1.336	0.1815
as.factor(A)1992	-0.115139	0.013070	-8.809	< 2e-16 ***
as.factor(A)1993	-0.189910	0.014023	-13.543	< 2e-16 ***
as.factor(A)1994	-0.156111	0.013788	-11.322	< 2e-16 ***
as.factor(A)1995	-0.142813	0.013960	-10.230	< 2e-16 ***
as.factor(A)1996	-0.121876	0.014393	-8.468	< 2e-16 ***
as.factor(A)1997	-0.113162	0.019076	-5.932	2.99e-09 ***
as.factor(B)1	-0.467200	0.007149	-65.353	< 2e-16 ***
as.factor(B)2	-1.456867	0.010717	-135.937	< 2e-16 ***
as.factor(B)3	-2.276393	0.016140	-141.038	< 2e-16 ***
as.factor(B)4	-2.802747	0.021910	-127.921	< 2e-16 ***
as.factor(B)5	-3.378022	0.030769	-109.787	< 2e-16 ***
as.factor(B)6	-4.050147	0.046987	-86.198	< 2e-16 ***
as.factor(B)7	-4.418412	0.065228	-67.738	< 2e-16 ***
as.factor(B)8	-6.407605	0.223720	-28.641	< 2e-16 ***

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

Null deviance: 123797.6 on 53 degrees of freedom  
Residual deviance: 1558.2 on 36 degrees of freedom  
(36 observations deleted due to missingness)  
AIC: 2050.4

Number of Fisher Scoring iterations: 5

```
> reg6=glm(Y~as.factor(B)+offset(log(P)),data=df,
+ family=poisson(link="log"))
> summary(reg6)
```

23

```
Call:
glm(formula = Y ~ as.factor(B) + offset(log(P)), family = poisson(link =
"log"),
     data = df)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-14.9933	-4.2664	-0.6501	4.3320	15.4796

Coefficients:

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	-1.053863	0.004284	-245.97	<2e-16 ***
as.factor(B)1	-0.462836	0.007055	-65.60	<2e-16 ***
as.factor(B)2	-1.444326	0.010592	-136.36	<2e-16 ***
as.factor(B)3	-2.251304	0.016014	-140.58	<2e-16 ***
as.factor(B)4	-2.759817	0.021780	-126.72	<2e-16 ***
as.factor(B)5	-3.308261	0.030646	-107.95	<2e-16 ***
as.factor(B)6	-3.951077	0.046872	-84.30	<2e-16 ***
as.factor(B)7	-4.309803	0.065098	-66.20	<2e-16 ***
as.factor(B)8	-6.341613	0.223648	-28.36	<2e-16 ***

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

Null deviance: 123798 on 53 degrees of freedom  
Residual deviance: 2633 on 45 degrees of freedom  
(36 observations deleted due to missingness)  
AIC: 3107.2

Number of Fisher Scoring iterations: 5

```
> sum(exp(predict(reg6,newdata=df)[futur]))
[1] 9511.738
```

```
> reg7=glm(Y/P~as.factor(B),weights=P,data=df,family=binomial)
```

```
> summary(reg7)
```

24

```
Call:
glm(formula = Y/P ~ as.factor(B), family = binomial, data = df,
     weights = P)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-18.2392	-4.7326	-0.7622	4.3798	18.0161

Coefficients:

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	-0.625250	0.005308	-117.78	<2e-16 ***
as.factor(B)1	-0.643713	0.008272	-77.82	<2e-16 ***
as.factor(B)2	-1.787127	0.011420	-156.49	<2e-16 ***
as.factor(B)3	-2.642534	0.016594	-159.25	<2e-16 ***
as.factor(B)4	-3.166117	0.022237	-142.38	<2e-16 ***
as.factor(B)5	-3.724041	0.030998	-120.14	<2e-16 ***
as.factor(B)6	-4.372963	0.047133	-92.78	<2e-16 ***
as.factor(B)7	-4.733722	0.065326	-72.46	<2e-16 ***
as.factor(B)8	-6.769612	0.223738	-30.26	<2e-16 ***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 141663.0 on 53 degrees of freedom  
Residual deviance: 3202.2 on 45 degrees of freedom  
(36 observations deleted due to missingness)  
AIC: 3668.7

Number of Fisher Scoring iterations: 5

```
> df1=df
> df1$P=1
> sum(predict(reg7,newdata=df1,type="response")[futur] * df$P[futur])
[1] 9511.738
```

---- PARTIE 3 ----

```
> DECES=read.table(
+ "http://freakonometrics.free.fr/DECES-CAN.csv",header=TRUE,sep=";")
> tail(DECES)
```

	D	E	A	Y
772	84	147	105	2010
773	39	76	106	2010
774	23	40	107	2010
775	15	20	108	2010
776	7	8	109	2010
777	5	5	110	2010

25

```
> DECES[DECES$A==20,]
```

	D	E	A	Y
21	343	220766	20	1950
132	299	239755	20	1960
243	488	375328	20	1970
354	617	486540	20	1980
465	340	406471	20	1990
576	319	411700	20	2000
687	238	466476	20	2010

26

```
> DECES[DECES$A==60,]
```

	D	E	A	Y
61	1974	105298	60	1950
172	1753	120055	60	1960
283	2223	158793	60	1970
394	2418	196779	60	1980
505	2480	240837	60	1990
616	2203	260158	60	2000
727	2740	409177	60	2010

27

```
> reg1=glm(D~as.factor(A)+Y+offset(log(E)),data=DECES,
+ family=poisson(link="log"))
> summary(reg1)
```

28

```
Call:
glm(formula = D ~ as.factor(A) + Y + offset(log(E)), family =
poisson(link = "log"),
     data = DECES)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-42.979	-2.853	-0.608	1.939	60.519

Coefficients:

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	2.649e+01	9.186e-02	288.393	<2e-16 ***
as.factor(A)1	-2.623e+00	1.787e-02	-146.811	<2e-16 ***
as.factor(A)2	-3.135e+00	2.280e-02	-137.495	<2e-16 ***
...				
as.factor(A)110	3.505e+00	2.423e-01	14.464	<2e-16 ***
Y	-1.544e-02	4.648e-05	-332.097	<2e-16 ***

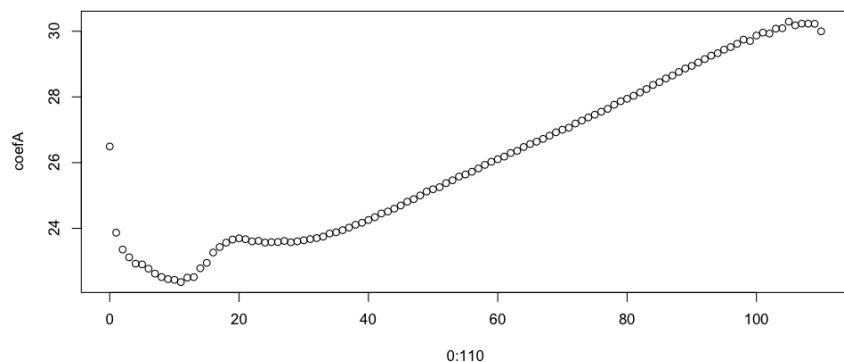
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

Null deviance: 3445251 on 776 degrees of freedom  
Residual deviance: 21594 on 665 degrees of freedom  
AIC: 28187

Number of Fisher Scoring iterations: 4

```
> coefA=c(0,coefficients(reg1)[2:111])+coefficients(reg1)[1]
> plot(0:110,coefA)
```



```
> brk=c(12,20,30)
```

```
> positive1=function(x) ifelse(x<brk[1],brk[1]-x,0)
> positive2=function(x) ifelse(x<brk[2],brk[2]-x,0)
> positive3=function(x) ifelse(x<brk[3],brk[3]-x,0)
>
> reg2=glm(D~A+positive1(A)+positive2(A)+positive3(A)+
+ offset(log(E)),data=DECES,family=poisson(link="log"))
```

```
> summary(reg2)
```

29

Call:

```
glm(formula = D ~ A + positive1(A) + positive2(A) + positive3(A) +
    offset(log(E)), family = poisson(link = "log"), data = DECES)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-59.362	-4.028	2.295	10.504	159.781

Coefficients:

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	-9.978e+00	4.787e-03	-2084.2	<2e-16 ***
A	9.048e-02	6.461e-05	1400.5	<2e-16 ***
positive1(A)	8.914e-01	4.070e-03	219.0	<2e-16 ***
positive2(A)	-5.066e-01	3.084e-03	-164.3	<2e-16 ***
positive3(A)	1.573e-01	9.299e-04	169.2	<2e-16 ***

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

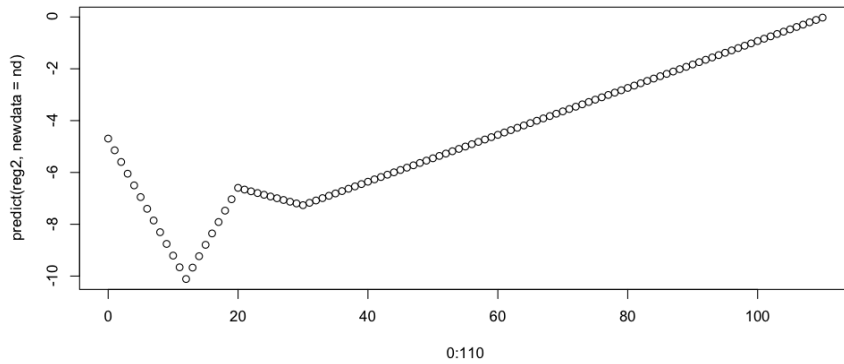
(Dispersion parameter for poisson family taken to be 1)

Null deviance: 3445251 on 776 degrees of freedom  
Residual deviance: 188600 on 772 degrees of freedom  
AIC: 194979

Number of Fisher Scoring iterations: 6

```
> nd=data.frame(A=0:110,Y=0,E=1)
```

```
> plot(0:110,predict(reg2,newdata=nd))
```



```
> reg3=glm(D/E~A+positive1(A)+positive2(A)+positive3(A),
+ data=DECES,weights=E,family=binomial(link="logit"))
> summary(reg4)
```

30

```
Call:
glm(formula = D/E ~ A + positive1(A) + positive2(A) + positive3(A),
family = binomial(link = "logit"), data = DECES, weights = E)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-59.492	-3.636	3.997	11.577	161.278

Coefficients:

	Estimate	Std. Error	z value	Pr(> z )	
(Intercept)	-1.020e+01	5.020e-03	-2031.3	<2e-16	***
A	9.439e-02	6.850e-05	1378.0	<2e-16	***
positive1(A)	8.991e-01	4.079e-03	220.4	<2e-16	***
positive2(A)	-5.254e-01	3.088e-03	-170.1	<2e-16	***
positive3(A)	1.744e-01	9.352e-04	186.4	<2e-16	***

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

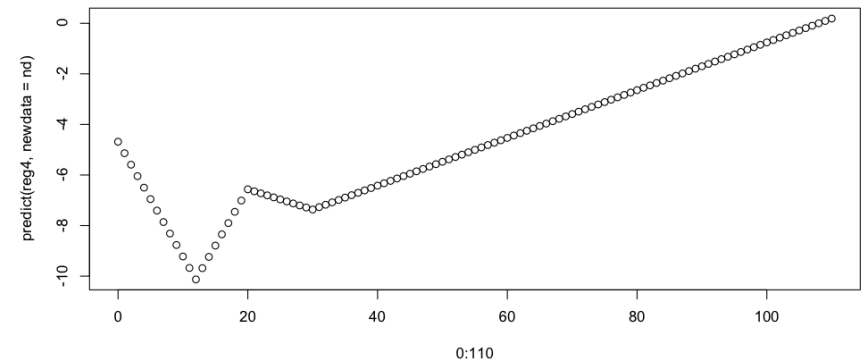
Null deviance: 3519013 on 776 degrees of freedom

23

Residual deviance: 196959 on 772 degrees of freedom  
AIC: 203240

Number of Fisher Scoring iterations: 6

```
> nd=data.frame(A=0:110,Y=0,E=1)
> plot(0:110,predict(reg3,newdata=nd))
```



24