GLM SUCCESS FOR CHILD REARING VALUES FOR FULL HUMAN RACE

ZULFIKAR MOINUDDIN AHMED

1. Poisson Regression Succeeded

After work with Binomial Regression, we found Poisson Regression provided best results.

2. GLM Fits Successful with Logit Link for Full Human Race

The full Human Race data have $N = 19077$ rows and after a day of struggle I can get good fits.

3. Do Not Have Great p-values for GLM Fit For Ethnicities

I suspect the issues is sample size because the fit looks fine on "Other" ethnicity with larger sample size.

Let's take a look at the p-values I obtain now (for the coefficient a).

$$logit(p) = ax + b$$

	Arab	Black	East Asian	Indian	Other	White
1	1.00	0.98	0.91	0.99	0.44	0.67
2	0.91	0.92	0.67	0.96	0.49	0.64
3	0.95	0.89	0.85	0.98	0.47	0.69
4	0.93	0.92	0.72	0.96	0.53	0.68
5	0.93	0.90	0.87	0.95	0.51	0.65
6	0.96	0.96	0.83	0.97	0.43	0.67
7	0.97	0.93	0.68	0.95	0.39	0.71
8	0.96	0.97	0.52	0.97	0.22	0.61
9	0.89	0.97	0.89	0.98	0.39	0.62
10	0.98	0.78	0.22	0.98	0.18	0.47
11	0.97	0.93	0.80	0.97	0.24	0.81

This is probably mostly a sample size problem.

4. Poisson Regression Works Like a Charm

The p-value table for the ethnicities too.

These are p-values multiplied by 10^{24} they were so small when Poisson regression was fit to ethnicities too.

The variation due to ethnicities are here.

	1	2	3	4	5	6
1	74989.72	0.00	0.00	1.77	0.00	0.00
2	2211914.24	0.00	0.00	0.00	0.00	0.00
3	67886944.45	0.00	0.00	0.05	0.00	0.00
4	0.14	0.00	0.00	0.00	0.00	0.00
5	218186.44	0.00	0.00	0.00	0.00	0.00
6	0.00	0.00	0.00	0.00	0.00	0.00
7	28453454.37	0.00	0.00	0.00	0.00	0.00
8	1758720.25	0.00	0.00	0.00	0.00	0.00
9	187725673.21	0.00	0.00	131370152.16	0.00	0.00
10	0.00	0.00	0.00	27924043.55	0.00	0.00
11	0.00	0.00	0.00	427171520.88	0.00	0.00

```
> for (r in 1:11){ a<-abs(A[r,]);pv<-(1-exp(-a*5))/(1-exp(-a*10)));print(sd(pv))}
[1] 0.09423666
[1] 0.1250242
[1] 0.1467352
[1] 0.05393183
[1] 0.04052609
[1] 0.05888644
[1] 0.08703167
[1] 0.09612605
[1] 0.1410823
[1] 0.09550833
[1] 0.1084014
```

The mean of these variations are

$$\bar{\sigma}_{eth} = 0.0952$$

So *this is the variation of child rearing variation due to ethnicity*! Excuse me, but *HAHAHAHAHAHAHAHAHA*
HAHAHAHAHAHAHAHAHA
HAHAHAHAHAHAHAHAHA

All the racial theorists will have a pretty difficult time convincing anyone of their rather absurd theories of racial superiority with only 9.52% at play.

5. CODE

```
# We will use logistic regression
# with x variable artificially created for Q7-Q17
# We take N=500 points to determine p, lambda
# Then we use these to assign random x values
# for all the other values
# then we fit logistic regression on the (x,g)

samp.binary.exp<-function( grps, lambda ){
  grps<-as.vector(t(grps))
  n <- length(grps)
  print(n)
  gvals <- unique(grps)
```

```
  #print(gvals)
  bval <- 0
  sval <- 1
  if ( length(gvals) == 2 ){
    if ( sum(gvals==gvals[1]) >= n/2 ){
      bval <- gvals[1]
      sval <- gvals[2]
    } else {
      sval <- gvals[1]
      bval <- gvals[2]
    }
  } else {
    return(NULL)
  }
  xs<- rep( 0, n)
  for ( r in 1:n ){
    done <- F

    while ( done == F){
      pickx <- rexp(1,rate=1/lambda)
      #print(pickx)
      if (pickx <= log(2)/lambda){
        if ( grps[r] == sval){
          xs[r] <- pickx
          done <- T
        }
      }
      if (pickx > log(2)/lambda){
        if ( grps[r] == bval){
          xs[r] <- pickx
          done <- T
        }
      }
    }
  }
  xs
}

dataset.logit<-function(var,lambda){
  y<-na.omit(polv[,var])
  G<-as.numeric(as_factor(t(y)))-1
  p<-sum(G==0)/length(G)
  if ( p < 0.5) {
    G<-1-G
  }
  p<-sum(G==0)/length(G)
  print(p)
```

```
  x<-samp.binary.exp( G, lambda)
  out<-data.frame( x=x,G=G)
  names(out) <- c("x","G")
  out
}

dataset.logit.fixed.lambda<-function(var){
  y<-na.omit(polv[,var])
  G<-as.numeric(as_factor(t(y)))-1
  p<-sum(G==0)/length(G)
  if ( p < 0.5) {
    G<-1-G
  }
  p<-sum(G==0)/length(G)
  print(p)
  lambda <- 2*log( p/(1-p) )
  #print(head(y))
  x<-samp.binary.exp( y, max(lambda,2.4))
  out<-data.frame( x=x,G=G)
  names(out) <- c("x","G")
  out
}

if (FALSE){
lq14<-dataset.logit.fixed.lambda("Q14")
mq14 = glm( G ~ x, family="binomial",data=lq14)
summary(mq14)
}



dataset.logit.eth<-function(var,eth){
  y0<-na.omit(polv[,var])
  G0<-as.numeric(as_factor(t(y0)))-1
  p0<-sum(G0==0)/length(G0)
  if ( p0 < 0.5) {
    G0<-1-G0
  }
  p0<-sum(G0==0)/length(G0)
  lambda0 <- 2*log( p0/(1-p0) )
  #lambda <- max(min(lambda,2.7),1.5)
  lambda0 <- max(lambda0,2.2)
  #print(head(y))
  x0<-samp.binary.exp( y0, lambda0)

  idx<-as.character(polv$eth)==eth
  y<-na.omit(polv[idx,var])
  G<-as.numeric(as_factor(t(y)))-1
```

```
  x<-x0[idx]
  p<-sum(G==0)/length(G)
  if ( p < 0.5) {
    G<-1-G
  }
  p<-sum(G==0)/length(G)
  #print(p)
  lambda <- 2*log( p/(1-p) )
  out<-data.frame( x=x,G=G)
  names(out) <- c("x","G")
  list(df=out,lambda=lambda)
}

dataset.logit.eth2<-function(var,eth){
  y<-na.omit(polv[as.character(polv$eth)==eth,var])
  G<-as.numeric(as_factor(t(y)))-1
  p<-sum(G==0)/length(G)
  if ( p < 0.5) {
    G<-1-G
  }
  p<-sum(G==0)/length(G)
  print(p)
  lambda <- 2*log( p/(1-p) )
  #lambda <- max(min(lambda,2.7),1.5)
  lambda <- min(max(lambda,0.5),2)
  #print(head(y))
  x<-samp.binary.exp( y, lambda)
  out<-data.frame( x=x,G=G)
  names(out) <- c("x","G")
  list(df=out,lambda=lambda)
}

vars<-c("Q7","Q8","Q9","Q10","Q11","Q12",
        "Q13","Q14","Q15","Q16","Q17")

eths<-c("Arab", "Black","East Asian", "Indian", "Other", "White")

child.eth<-function(){
  lambdas <- matrix( 0, nrow=length(vars), ncol=length(eths))
  alphas <- matrix( 0, nrow=length(vars), ncol=length(eths))
  pvals <- matrix( 0, nrow=length(vars), ncol=length(eths))
  bdf <- data.frame()
  for (r in 1:length(vars)){
    for (s in 1:length(eths)){
      A<-dataset.logit.eth( vars[r], eths[s])
      m <- glm( G ~ x, family="poisson", data=A$df)
      lambdas[r,s]<-A$lambda
      alphas[r,s]<-summary(m)$coefficients[2,1]
```

```
        pvals[r,s]<-summary(m)$coefficients[2,4]
        bdf<-rbind(bdf,c(vars[r],eths[s],alphas[r,s],lambdas[r,s],pvals[r,s]))
    }
  }
  names(bdf)<-c("var","eth","alpha","lambda","pval")
  list(df=bdf,lambda=lambdas,alpha=alphas,pvals=pvals)
}
```