Therapy

Yang

April 2, 2019

Load data

```
# The data
Y<- cbind(c(28,4,41,12),c(45,12,44,7),c(29,5,20,3),c(26,2,20,1))
colnames(Y)<-c("ProgressiveDisease","NoChange","PartialRemission","CompleteRemission")
therapy <- c("S","S","A","A")
gender <- rep(c("M","F"),2)
```

Baseline Categorical Logit Model

Fit

```
library(nnet)
#Progressive Disease as baseline
b1<-multinom(Y~as.factor(therapy))
```

```
## # weights:  12 (6 variable)
## initial  value 414.502014
## iter  10 value 395.641305
## final  value 395.641300
## converged
```

```
beta1<-coefficients(b1)
summary(b1)
```

```
## Call:
## multinom(formula = Y ~ as.factor(therapy))
##
## Coefficients:
##                  (Intercept) as.factor(therapy)S
## NoChange         -0.03846754           0.6157805
## PartialRemission -0.83479810           0.8954216
## CompleteRemission -0.92576929          0.7922358
##
## Std. Errors:
##                  (Intercept) as.factor(therapy)S
## NoChange           0.1961524           0.2954143
## PartialRemission   0.2496922           0.3507248
## CompleteRemission  0.2578506           0.3653097
##
## Residual Deviance: 791.2826
## AIC: 803.2826
```

Results

```r
result=predict(b1,type ="prob")
result=cbind(therapy,result)
#predicted probability
result
```

```
##   therapy ProgressiveDisease  NoChange           PartialRemission
## 1 "S"     "0.211920853535378" "0.37748312784509" "0.225165653963416"
## 2 "S"     "0.211920853535378" "0.37748312784509" "0.225165653963416"
## 3 "A"     "0.358108276712193" "0.344594321659436" "0.155405416074088"
## 4 "A"     "0.358108276712193" "0.344594321659436" "0.155405416074088"
##   CompleteRemission
## 1 "0.185430364656116"
## 2 "0.185430364656116"
## 3 "0.141891985554283"
## 4 "0.141891985554283"
```

```r
# proportional Therapy S vs A
S <-c(28,4, 45,12,29,5,26,2)
A <-c(41,12,44,7,20,3,20,1)
PS=sum(S)/(sum(S)+sum(A))
PA=sum(A)/(sum(S)+sum(A))
cbind(PS,PA)
```

```
##             PS        PA
## [1,] 0.5050167 0.4949833
```

Interpretation

In our model with Progressive Disease as baseline:
Under the alternating therapy treatment, we have $log(\frac{\pi_N}{\pi_{Prog}}) = \beta_{11} = -0.03846754$; Under the sequential therapy treatment, we have $log(\frac{\pi_N}{\pi_{Prog}}) = \beta_{11} + \beta_{12} = -0.03846754 + 0.6157805 = 0.577313$.

The log odds ratio: $log\frac{odd_S}{odd_A} = log(\frac{\pi_N}{\pi_{Prog}}|S) - log(\frac{\pi_N}{\pi_{Prog}}|A) = \beta_{12} = 0.6157805$, which suggests the odds of "Progressive Disease" vs "No Disease" with alternating sequential and alternating therapy are $\frac{odd_S}{odd_A} = e^{0.6157805} = 1.851101$, meaning sequential therapy are better than alternating therapy.

Confidence Interval

```r
SE <- summary(b1)$standard.errors
#95% CI
CI.u <- exp(beta1[1,2]+ 1.96*SE[1,2])
CI.l <- exp(beta1[1,2]- 1.96*SE[1,2])
cbind(CI.l,CI.u)
```

```
##          CI.l     CI.u
## [1,] 1.037453 3.302873
```

From the confidence interval, we can also conclude that we are 95% confident that sequential therapy has a better treament effect than alternating treament effect since 1 is not trapped by our confidence interval.

Cummulative Logit Model

Fit

```r
library(VGAM)
```

```
## Warning: package 'VGAM' was built under R version 3.4.4
```

```
## Loading required package: stats4
```

```
## Loading required package: splines
```

```r
c1=vglm(Y~therapy,propodds(reverse=F))
summary(c1)
```

```
##
## Call:
## vglm(formula = Y ~ therapy, family = propodds(reverse = F))
##
## Pearson residuals:
##    logitlink(P[Y<=1]) logitlink(P[Y<=2]) logitlink(P[Y<=3])
## 1             -0.2129            -0.2992            0.04569
## 2             -1.0015             0.9976            1.25878
## 3             -0.1322            -0.4179           -1.12409
## 4              1.6003             0.5881            0.91423
##
## Coefficients:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept):1  -0.6468     0.1620  -3.993 6.52e-05 ***
## (Intercept):2   0.9081     0.1667   5.449 5.06e-08 ***
## (Intercept):3   1.9502     0.1984   9.831  < 2e-16 ***
## therapyS       -0.5699     0.2116  -2.694  0.00706 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Names of linear predictors: logitlink(P[Y<=1]), logitlink(P[Y<=2]),
## logitlink(P[Y<=3])
##
## Residual deviance: 9.1642 on 8 degrees of freedom
##
## Log-likelihood: -27.3399 on 8 degrees of freedom
##
## Number of Fisher scoring iterations: 4
##
## No Hauck-Donner effect found in any of the estimates
##
##
## Exponentiated coefficients:
##   therapyS
## 0.5655741
```

Coefficients and interpretation

Exponentiated coefficients:therapyS $e^\beta = e^{-0.5699} = 0.5655741$. Interpretation: $log\frac{\frac{Pr(Y \leq j|Therapy=S)}{Pr(Y \geq j|Therapy=S)}}{\frac{Pr(Y \leq j|Therapy=A)}{Pr(Y \geq j|Therapy=A)}} = \beta \implies$

$\frac{\frac{Pr(Y \leq j|Therapy=S)}{Pr(Y \geq j|Therapy=S)}}{\frac{Pr(Y \leq j|Therapy=A)}{Pr(Y \geq j|Therapy=A)}} = e^\beta = e^{-0.5699} = 0.5655741$. Since 0.5655741 is smaller than 1, we can conclude that sequential therapy has a better effect than alternating therapy.

Confidence Interval

```
e=exp(1)
e^confint(c1)
```

```
##                     2.5 %      97.5 %
## (Intercept):1 0.3812571   0.7193911
## (Intercept):2 1.7887255   3.4375737
## (Intercept):3 4.7655893 10.3708532
## therapyS      0.3735997   0.8561948
```

95% confidence interval for exponentiated coefficients of therapy:[0.3735997, 0.8561948]. Interpretation: we are 95% confident that the true odds ratio for the effect of therapy fall within [0.3735997, 0.8561948].

Should Gender Be Included in the Models Above?

Baseline Categorical Model

```
library(nnet)
b2<-multinom(Y~therapy+gender)
```

```
## # weights:  16 (9 variable)
## initial  value 414.502014
## iter  10 value 393.034226
## final  value 393.017781
## converged
```

```
beta2<-coefficients(b2)
summary(b2)
```

```
## Call:
## multinom(formula = Y ~ therapy + gender)
##
## Coefficients:
##                   (Intercept)   therapyS     genderM
## NoChange            -0.112579 0.6164483 0.09023528
## PartialRemission    -1.138719 0.8978800 0.36184990
## CompleteRemission   -2.060487 0.7986133 1.27828139
##
## Std. Errors:
##                   (Intercept)   therapyS     genderM
## NoChange            0.3676934 0.2954664 0.3786301
## PartialRemission    0.4759568 0.3511357 0.4771676
## CompleteRemission   0.6614644 0.3677732 0.6605390
```

```
## 
## Residual Deviance: 786.0356
## AIC: 804.0356
```

```
anova(b1,b2,test="Chi")
```

```
##                   Model Resid. df Resid. Dev   Test   Df LR stat.   Pr(Chi)
## 1 as.factor(therapy)         6   791.2826            NA       NA        NA
## 2    therapy + gender        3   786.0356 1 vs 2      3 5.247038 0.1545762
```

According to the anova table, the difference between baseline category logit model with therapy and the one with therapy+gender is not statistically significant. Therefore, we would not include gender as a predictor.

Cummulative Logit Model

```
c2<-vglm(Y~therapy+gender,propodds(reverse=F))
summary(c2)
```

```
## 
## Call:
## vglm(formula = Y ~ therapy + gender, family = propodds(reverse = F))
## 
## Pearson residuals:
##    logitlink(P[Y<=1]) logitlink(P[Y<=2]) logitlink(P[Y<=3])
## 1             0.1720            0.06056            0.2809
## 2            -1.6543            0.16312            0.8425
## 3             0.2655           -0.13909           -0.9386
## 4             0.6174           -0.01519            0.6444
## 
## Coefficients:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept):1  -0.1960     0.2947  -0.665  0.50605
## (Intercept):2   1.3713     0.3059   4.482 7.38e-06 ***
## (Intercept):3   2.4221     0.3276   7.393 1.43e-13 ***
## therapyS       -0.5807     0.2119  -2.741  0.00613 **
## genderM        -0.5414     0.2953  -1.834  0.06671 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Names of linear predictors: logitlink(P[Y<=1]), logitlink(P[Y<=2]),
## logitlink(P[Y<=3])
## 
## Residual deviance: 5.5677 on 7 degrees of freedom
## 
## Log-likelihood: -25.5417 on 7 degrees of freedom
## 
## Number of Fisher scoring iterations: 5
## 
## No Hauck-Donner effect found in any of the estimates
## 
## 
## Exponentiated coefficients:
##   therapyS   genderM
## 0.5595152 0.5819358
```

```
p_val <- pchisq(2*(logLik(c2)-logLik(c1)),df=df.residual(c1)-df.residual(c2), lower.tail=FALSE)
p_val
```

[1] 0.05790126

The p value from the test is 0.057<0.05, from which we can conclude that gender should not be included in the cummulative logit model.

Conclusion

Therefore, we conclude that gender should not be included in the model(s).