STA138 Take Home Final

Madeline Ye

998108849

12/7/2016

Problem 1

## Start: AIC=1741.59  
## count ~ S + T + E + V  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + T:V 6 740.7 1114.1 639.5 < 2.2e-16 \*\*\*  
## + T:E 3 973.9 1341.4 406.2 < 2.2e-16 \*\*\*  
## + E:V 2 1035.0 1400.5 345.1 < 2.2e-16 \*\*\*  
## + S:T 6 1289.5 1662.9 90.7 < 2.2e-16 \*\*\*  
## + S:E 2 1365.8 1731.3 14.3 0.000788 \*\*\*  
## <none> 1380.1 1741.6   
## + S:V 4 1375.7 1745.1 4.5 0.346006   
## - V 2 2584.1 2941.6 1204.0 < 2.2e-16 \*\*\*  
## - T 3 3393.3 3748.8 2013.2 < 2.2e-16 \*\*\*  
## - E 1 4054.9 4414.3 2674.7 < 2.2e-16 \*\*\*  
## - S 2 6475.1 6832.5 5094.9 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=1114.14  
## count ~ S + T + E + V + T:V  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + T:E 3 334.5 713.9 406.2 < 2.2e-16 \*\*\*  
## + E:V 2 395.6 773.0 345.1 < 2.2e-16 \*\*\*  
## + S:T 6 650.0 1035.5 90.7 < 2.2e-16 \*\*\*  
## + S:E 2 726.4 1103.8 14.3 0.000788 \*\*\*  
## <none> 740.7 1114.1   
## + S:V 4 736.2 1117.7 4.5 0.346006   
## - T:V 6 1380.1 1741.6 639.5 < 2.2e-16 \*\*\*  
## - E 1 3415.4 3786.9 2674.7 < 2.2e-16 \*\*\*  
## - S 2 5835.6 6205.0 5094.9 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=713.92  
## count ~ S + T + E + V + T:V + T:E  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + E:V 2 156.8 540.3 177.7 < 2.2e-16 \*\*\*  
## + S:T 6 243.8 635.2 90.7 < 2.2e-16 \*\*\*  
## + S:E 2 320.2 703.6 14.3 0.000788 \*\*\*  
## <none> 334.5 713.9   
## + S:V 4 330.0 717.4 4.5 0.346006   
## - T:E 3 740.7 1114.1 406.2 < 2.2e-16 \*\*\*  
## - T:V 6 973.9 1341.4 639.5 < 2.2e-16 \*\*\*  
## - S 2 5429.4 5804.8 5094.9 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=540.27  
## count ~ S + T + E + V + T:V + T:E + E:V  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + S:T 6 66.1 461.6 90.7 < 2.2e-16 \*\*\*  
## + S:E 2 142.5 530.0 14.3 0.000788 \*\*\*  
## <none> 156.8 540.3   
## + S:V 4 152.3 543.8 4.5 0.346006   
## + T:E:V 6 149.2 544.6 7.7 0.264420   
## - E:V 2 334.5 713.9 177.7 < 2.2e-16 \*\*\*  
## - T:E 3 395.6 773.0 238.8 < 2.2e-16 \*\*\*  
## - T:V 6 628.8 1000.3 472.0 < 2.2e-16 \*\*\*  
## - S 2 5251.7 5631.2 5094.9 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=461.59  
## count ~ S + T + E + V + T:V + T:E + E:V + S:T  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## + S:V 4 51.30 454.75 14.85 0.005027 \*\*   
## <none> 66.14 461.59   
## + S:E 2 64.73 464.18 1.42 0.492499   
## + T:E:V 6 58.49 465.94 7.66 0.264420   
## - S:T 6 156.82 540.27 90.67 < 2.2e-16 \*\*\*  
## - E:V 2 243.79 635.24 177.65 < 2.2e-16 \*\*\*  
## - T:E 3 304.92 694.37 238.78 < 2.2e-16 \*\*\*  
## - T:V 6 538.16 921.61 472.01 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=454.75  
## count ~ S + T + E + V + T:V + T:E + E:V + S:T + S:V  
##   
## Df Deviance AIC LRT Pr(>Chi)   
## <none> 51.30 454.75   
## + S:E 2 47.98 455.43 3.31 0.190864   
## + S:T:V 12 29.23 456.68 22.06 0.036817 \*   
## + T:E:V 6 43.64 459.09 7.66 0.264420   
## - S:V 4 66.14 461.59 14.85 0.005027 \*\*   
## - S:T 6 152.35 543.80 101.05 < 2.2e-16 \*\*\*  
## - E:V 2 228.95 628.40 177.65 < 2.2e-16 \*\*\*  
## - T:E 3 290.07 687.52 238.78 < 2.2e-16 \*\*\*  
## - T:V 6 531.93 923.38 480.63 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

## count ~ S + T + E + V + T:V + T:E + E:V + S:T + S:V

##   
## Call:  
## glm(formula = count ~ S + T + E + V + T:V + T:E + E:V + S:T +   
## S:V, family = poisson(), data = accident)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -2.0412 -0.6093 -0.0293 0.5306 2.1534   
##   
## Coefficients:  
## Estimate Std. Error  
## (Intercept) 4.620382 0.089241  
## SCompact 0.413589 0.114315  
## SStandard 2.536710 0.092004  
## TCollision with object -1.064272 0.160268  
## TRollover without collision -1.592252 0.161529  
## TOther Rollover -2.083206 0.175653  
## EEjected -2.934699 0.093576  
## VModerately severe -1.259954 0.135236  
## VSevere -2.357003 0.187081  
## TCollision with object:VModerately severe 0.067807 0.082379  
## TRollover without collision:VModerately severe 1.517798 0.122599  
## TOther Rollover:VModerately severe 1.863067 0.118074  
## TCollision with object:VSevere 0.009126 0.116172  
## TRollover without collision:VSevere 0.857775 0.169804  
## TOther Rollover:VSevere 1.970297 0.138134  
## TCollision with object:EEjected 0.248046 0.123420  
## TRollover without collision:EEjected 1.643503 0.128642  
## TOther Rollover:EEjected 1.382018 0.117309  
## EEjected:VModerately severe 0.717841 0.108717  
## EEjected:VSevere 1.594034 0.118810  
## SCompact:TCollision with object 0.272176 0.196247  
## SStandard:TCollision with object 0.292258 0.163143  
## SCompact:TRollover without collision -0.795276 0.211316  
## SStandard:TRollover without collision -1.380739 0.159123  
## SCompact:TOther Rollover -0.431886 0.214956  
## SStandard:TOther Rollover -0.759030 0.166994  
## SCompact:VModerately severe 0.083118 0.168587  
## SStandard:VModerately severe 0.299453 0.134442  
## SCompact:VSevere 0.076890 0.232789  
## SStandard:VSevere 0.468962 0.183386  
## z value Pr(>|z|)   
## (Intercept) 51.774 < 2e-16 \*\*\*  
## SCompact 3.618 0.000297 \*\*\*  
## SStandard 27.572 < 2e-16 \*\*\*  
## TCollision with object -6.641 3.12e-11 \*\*\*  
## TRollover without collision -9.857 < 2e-16 \*\*\*  
## TOther Rollover -11.860 < 2e-16 \*\*\*  
## EEjected -31.362 < 2e-16 \*\*\*  
## VModerately severe -9.317 < 2e-16 \*\*\*  
## VSevere -12.599 < 2e-16 \*\*\*  
## TCollision with object:VModerately severe 0.823 0.410444   
## TRollover without collision:VModerately severe 12.380 < 2e-16 \*\*\*  
## TOther Rollover:VModerately severe 15.779 < 2e-16 \*\*\*  
## TCollision with object:VSevere 0.079 0.937386   
## TRollover without collision:VSevere 5.052 4.38e-07 \*\*\*  
## TOther Rollover:VSevere 14.264 < 2e-16 \*\*\*  
## TCollision with object:EEjected 2.010 0.044455 \*   
## TRollover without collision:EEjected 12.776 < 2e-16 \*\*\*  
## TOther Rollover:EEjected 11.781 < 2e-16 \*\*\*  
## EEjected:VModerately severe 6.603 4.03e-11 \*\*\*  
## EEjected:VSevere 13.417 < 2e-16 \*\*\*  
## SCompact:TCollision with object 1.387 0.165471   
## SStandard:TCollision with object 1.791 0.073225 .   
## SCompact:TRollover without collision -3.763 0.000168 \*\*\*  
## SStandard:TRollover without collision -8.677 < 2e-16 \*\*\*  
## SCompact:TOther Rollover -2.009 0.044518 \*   
## SStandard:TOther Rollover -4.545 5.49e-06 \*\*\*  
## SCompact:VModerately severe 0.493 0.621991   
## SStandard:VModerately severe 2.227 0.025922 \*   
## SCompact:VSevere 0.330 0.741172   
## SStandard:VSevere 2.557 0.010551 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for poisson family taken to be 1)  
##   
## Null deviance: 12366.978 on 71 degrees of freedom  
## Residual deviance: 51.295 on 42 degrees of freedom  
## AIC: 454.75  
##   
## Number of Fisher Scoring iterations: 4

## Waiting for profiling to be done...

## OR 2.5 %  
## (Intercept) 101.53277430 84.85096004  
## SCompact 1.51223543 1.21024924  
## SStandard 12.63801763 10.59458312  
## TCollision with object 0.34497875 0.24993288  
## TRollover without collision 0.20346691 0.14745843  
## TOther Rollover 0.12453034 0.08768732  
## EEjected 0.05314671 0.04406874  
## VModerately severe 0.28366715 0.21698348  
## VSevere 0.09470362 0.06494252  
## TCollision with object:VModerately severe 1.07015920 0.91002292  
## TRollover without collision:VModerately severe 4.56216770 3.59436122  
## TOther Rollover:VModerately severe 6.44346879 5.12579394  
## TCollision with object:VSevere 1.00916773 0.80193666  
## TRollover without collision:VSevere 2.35790815 1.68488886  
## TOther Rollover:VSevere 7.17280591 5.47945046  
## TCollision with object:EEjected 1.28151944 1.00407865  
## TRollover without collision:EEjected 5.17325963 4.01937336  
## TOther Rollover:EEjected 3.98293197 3.16573019  
## EEjected:VModerately severe 2.05000254 1.65752779  
## EEjected:VSevere 4.92356914 3.90212446  
## SCompact:TCollision with object 1.31281843 0.89665972  
## SStandard:TCollision with object 1.33944790 0.97928539  
## SCompact:TRollover without collision 0.45145668 0.29742788  
## SStandard:TRollover without collision 0.25139264 0.18440755  
## SCompact:TOther Rollover 0.64928368 0.42577639  
## SStandard:TOther Rollover 0.46812024 0.33858791  
## SCompact:VModerately severe 1.08667037 0.78118875  
## SStandard:VModerately severe 1.34912014 1.03876216  
## SCompact:VSevere 1.07992381 0.68520305  
## SStandard:VSevere 1.59833434 1.12569064  
## 97.5 %  
## (Intercept) 120.41163172  
## SCompact 1.89514586  
## SStandard 15.19874841  
## TCollision with object 0.46906900  
## TRollover without collision 0.27792334  
## TOther Rollover 0.17467916  
## EEjected 0.06360645  
## VModerately severe 0.36881878  
## VSevere 0.13536862  
## TCollision with object:VModerately severe 1.25698334  
## TRollover without collision:VModerately severe 5.81406088  
## TOther Rollover:VModerately severe 8.14542081  
## TCollision with object:VSevere 1.26486806  
## TRollover without collision:VSevere 3.28064184  
## TOther Rollover:VSevere 9.41986815  
## TCollision with object:EEjected 1.62953209  
## TRollover without collision:EEjected 6.65684726  
## TOther Rollover:EEjected 5.01489678  
## EEjected:VModerately severe 2.53887851  
## EEjected:VSevere 6.21816179  
## SCompact:TCollision with object 1.93728078  
## SStandard:TCollision with object 1.85865909  
## SCompact:TRollover without collision 0.68166374  
## SStandard:TRollover without collision 0.34433376  
## SCompact:TOther Rollover 0.98974201  
## SStandard:TOther Rollover 0.65214750  
## SCompact:VModerately severe 1.51324368  
## SStandard:VModerately severe 1.76016524  
## SCompact:VSevere 1.70944832  
## SStandard:VSevere 2.31306259

After performing a step-wise procedure, the best loglinear fitting model based on AIC value was the model with two way interactions (TV,TE,TS,EV,SV) . The only two way interaction that was left out of the model was shape/size of the car and whether the person was ejected or not (S:E). This suggests that the interaction of the shape/size of the car and whether the person was ejected or not was not significant and did not add anything to the model. Because the shape/size of the car (S) and the severity of the accident (V), were both broken down into many levels, there are many coefficents(30) in this model.

Problem 6.8

##   
## Call:  
## glm(formula = Y ~ D + T, family = binomial(link = logit), data = surgery)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -2.3802 -0.5358 0.3047 0.7308 1.7821   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -1.41734 1.09457 -1.295 0.19536   
## D 0.06868 0.02641 2.600 0.00931 \*\*  
## T -1.65895 0.92285 -1.798 0.07224 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 46.180 on 34 degrees of freedom  
## Residual deviance: 30.138 on 32 degrees of freedom  
## AIC: 36.138  
##   
## Number of Fisher Scoring iterations: 5

## [[1]]  
## Wald test:  
## ----------  
##   
## Chi-squared test:  
## X2 = 1.7, df = 1, P(> X2) = 0.2  
##   
## [[2]]  
## Wald test:  
## ----------  
##   
## Chi-squared test:  
## X2 = 6.8, df = 1, P(> X2) = 0.0093  
##   
## [[3]]  
## Wald test:  
## ----------  
##   
## Chi-squared test:  
## X2 = 3.2, df = 1, P(> X2) = 0.072

## Waiting for profiling to be done...

## OR 2.5 % 97.5 %  
## (Intercept) 0.2423575 0.02219559 1.852165  
## D 1.0710910 1.02580381 1.141299  
## T 0.1903389 0.02608803 1.077174

The final logistic regression equation represents the probability that the patient will wake up from the surgery with a sore throat. We can see that the duration of the surgery (D) has a possitive correlation with the probability of a patient having a sore throat.

To test significance of each coefficent, a Wald test was conducted at . The Wald ^2 test statistic for are 1.7, 6.8, and 3.2, each with a p-value of 0.2, 0.0093, and 0.072, respectively. These findings show that the duration of the surgery is statistically significant. However, the type of device used to secure the airway (T) is not significant.

The odds ratio for the duration of the surgery (D) is 1.07109 with a CI of [1.02580381, 1.141299]. This suggests the that odds of a patient waking up with sore throat will be between about 2% - 14% for every aditional minute of the surgery. The odds ratio for the type of device used to secure the airway (T) is 0.1903389. The confidence interval is rather wide for the odds ratio for type of device used: ([.02608803, 1.077174]) and contains 1, so there is no significant difference between the two devices. This was confirmed earlier with the Wald test as well.

Appendix

shape.level = factor(c("Small", "Compact", "Standard"), levels = c("Small", "Compact", "Standard"))

type.level = factor(c("Collision with vehicle", "Collision with object", "Rollover without collision", "Other Rollover"), levels = c("Collision with vehicle", "Collision with object","Rollover without collision", "Other Rollover"))

severity.level = factor(c("Not severe", "Moderately severe", "Severe"), levels = c("Not severe","Moderately severe", "Severe"))

eject.level = factor(c("Not ejected", "Ejected"), levels = c("Not ejected", "Ejected"))

count = c(95, 166, 1279, 31, 34, 506, 11, 17, 186, 34, 55, 599, 8, 34, 241, 5, 10, 89, 23, 18, 65,22, 17, 118, 5, 2, 23, 9, 10, 83, 23, 26, 177, 8, 9, 86, 8, 7, 65, 2, 5, 51, 4, 5, 54, 5, 6, 46, 2,4, 26, 0, 1, 30, 6, 5, 11, 18, 9, 68, 5, 6, 33, 6, 2, 11, 13, 16, 78, 7, 6, 86)

accident = data.frame(T = rep(type.level,each = 9), E = rep(eject.level, each = 36), V = rep(rep(severity.level, each = 3), 8), S = rep(shape.level), count)

fit.1 = glm(count~S+T+E+V,data = accident, family = poisson())

best.model = step(fit.1,~.^4,direction = "both", test = "Chisq", data = accident)

exp(cbind(OR=coef(best.model),confint(best.model)))

#Problem 6.8 Select a Logistic Model

D=c(45,15,40,83,90,25,35,65,95,35,75,45,50,75,30,25,20,60,70,30,60,61,65,15,20,45,15,25,15,30,40,15,135,20,40)

T=c(0,0,0,1,1,1,0,0,0,0,0,1,1,1,0,0,1,1,1,0,0,0,0,1,1,0,1,0,1,0,0,1,1,1,1)

Y=c(0,0,1,1,1,1,1,1,1,1,1,1,0,1,0,1,0,1,1,1,1,0,1,0,0,1,0,1,0,1,1,0,1,0,0)

surgery=data.frame(D,T,Y)

model=glm(Y~D+T,data=surgery,family=binomial(link=logit))

summary(model)

library(aod)

lapply(1:3, function(x){

test = wald.test(Sigma=vcov(model), b = coef(model), Terms = x)

})

exp(cbind(OR = coef(model), confint(model)))