

LING82100: homework 6 solution

```
> library(lmtest)
> d <- read.csv(
+   "http://wellformedness.com/courses/LING82100/Data/NYC.csv"
+ )
```

1 Logistic regression

```
> contrasts(d$store) <- contr.sum
> contrasts(d$word) <- contr.sum
> contrasts(d$emphasis) <- contr.sum
> r <- glm(r ~ store + word + emphasis, data = d, family = binomial)
> summary(r)
```

Call:

```
glm(formula = r ~ store + word + emphasis, family = binomial,
     data = d)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-1.4511	-0.8551	-0.5162	1.0552	2.4498

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-0.93588	0.10283	-9.102	< 2e-16	***
store1	-1.34852	0.16928	-7.966	1.64e-15	***
store2	0.45423	0.12180	3.729	0.000192	***
word1	0.50065	0.08763	5.713	1.11e-08	***
emphasis1	0.16457	0.08948	1.839	0.065877	.

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

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 908.96 on 728 degrees of freedom
Residual deviance: 789.62 on 724 degrees of freedom

AIC: 799.62

Number of Fisher Scoring iterations: 5

```
> r._store <- glm(r ~ word + emphasis, data = d, family = binomial)
> lrtest(r, r._store)
```

Model 1: r ~ store + word + emphasis

Model 2: r ~ word + emphasis

	#Df	LogLik	Df	Chisq	Pr(>Chisq)
1	5	-394.81			
2	3	-438.84	-2	88.054	< 2.2e-16 ***

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

```
> r._word <- glm(r ~ store + emphasis, data = d, family = binomial)
```

```
> lrtest(r, r._word)
```

Likelihood ratio test

Model 1: r ~ store + word + emphasis

Model 2: r ~ store + emphasis

	#Df	LogLik	Df	Chisq	Pr(>Chisq)
1	5	-394.81			
2	4	-411.79	-1	33.961	5.623e-09 ***

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

```
> lrtest(r, r._emphasis)
```

Likelihood ratio test

Model 1: r ~ store + word + emphasis

Model 2: r ~ store + word

	#Df	LogLik	Df	Chisq	Pr(>Chisq)
1	5	-394.81			
2	4	-396.50	-1	3.378	0.06607 .

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

Results are shown in Table 1. At $\alpha = .05$, there are significant effects of store ($\chi^2(2) = 88.054$, $p < .001$) and word ($\chi^2(1) = 33.961$, $p < .001$), but emphasis is non-significant ($\chi^2(1) = 3.378$, $p = .067$).

	Coef.	S.E.	χ^2	$p(\chi^2)$
(Intercept)	-0.936	0.10		
Store:			88.054	< .001
S. Klein's	-1.349	0.17		
Macy's	0.454	0.12		
Saks 5th Ave.	0.895			
Word:			33.961	< .001
<i>floor</i>	0.501	0.09		
<i>fourth</i>	-0.501			
Emphasis:			3.378	.066
emphatic	0.165	0.09		
normal	-0.165			

Table 1: Logistic regression results for the department store study.