

L - Comparing Logistic Models

Anaan Choudhury

##Question 1

```
library(Stat2Data)
data("ICU")
```

```
library(dplyr)
```

Attaching package: 'dplyr'

The following objects are masked from 'package:stats':

filter, lag

The following objects are masked from 'package:base':

intersect, setdiff, setequal, union

```
ICU <- mutate(ICU, AgeGroup = recode(AgeGroup,
                                     `1` = "Young",
                                     `2` = "Old",
                                     `3` = "Elderly"
                                   ),
              AgeGroup = factor(AgeGroup, levels = c("Young", "Old", "Elderly")))

age_and_bp_add <- glm(Survive ~ AgeGroup + SysBP, data=ICU, family=binomial)
age_and_bp_interaction <- glm(Survive ~ AgeGroup * SysBP, data=ICU, family=binomial)
```

age_and_bp_add is the reduced model because it is only additive. It doesn't account for effects of the coefficients on each other.

##Question 2

```
anova(age_and_bp_add, age_and_bp_interaction, test = "Chisq")
```

Analysis of Deviance Table

Model 1: Survive ~ AgeGroup + SysBP

Model 2: Survive ~ AgeGroup * SysBP

	Resid. Df	Resid. Dev	Df	Deviance	Pr(>Chi)
1	196	182.49			
2	194	181.71	2	0.77499	0.6788

P-value = 0.6788 Very large, so we choose the reduced model. Interaction doesn't decrease variance enough to warrant making the model longer and more complex. We always aim to use simpler, basic models.

##Question 3 H0 - Null: A person's age group doesn't have a significant impact on the effect of Blood Pressure on survival

HA - Alternative: A person's age group does have a significant impact on the effect of Blood Pressure on survival