End Term

Saket Choudhary

December 9, 2015

Problem 1

Define the likelihood function: $L(\theta|N,n,k)$ Then,

$$L(\theta|N,n,k) = \frac{\binom{\theta}{k}\binom{N-\theta}{n-k}}{\binom{N}{n}}$$

We find the MLE using first principle.

In order to ensure MLE, we need to ensure: $\frac{L(\theta|N,n,k)}{L(\theta-1|N,n,k)} > 1$

$$\frac{L(\theta|N,n,k)}{L(\theta-1|N,n,k)} > 1$$
$$\frac{\frac{\binom{\theta}{k}\binom{N-\theta}{n-k}}{\binom{N}{n}}}{\frac{\binom{\theta-1}{k}\binom{N-\theta+1}{n-k}}{\binom{N}{n}}}$$

Problem 4

Part (a)

In this part, the variables X_1 and X_2 are correlated. The main diffciulty that would potentially arise would be in interpreting the cofficients associtated with these variables. Cofficients by definition imply the amount by which the mean response changes when all other covariates are held fixed. However in this case since X_1 and X_2 are highly correlated, changing one would also imply chaning the other.

In order to overcome this, we need to choose either of X_1 or X_2 as a covariate in the linear regression model(essentially discarding the other) based on which one of these predictors best captures the 'reality' of the independent variable.

Part (b)

Problem 5

Given $\log\frac{p}{1-p} = 3.2 - 0.078 \times age$ for men and $\log\frac{p}{1-p} = 1.6 - 0.078 \times age$ for women where p denotes the probability of survival.

Consider

$$
\begin{aligned}
\log \frac{p}{1-p} &= \beta_0 + \beta_1 x \\
\frac{1-p}{p} &= \exp(-(\beta_0 + \beta_1 x)) \\
p &= \frac{1}{1+\exp(-(\beta_0 + \beta_1 x))}
\end{aligned}
$$

```
pman <- 1/(1+exp(-(3.2-0.078*25)))
pwoman <-  1/(1+exp(-(1.6-0.078*50)))
```

Thus, Estimated probability of survival of man of age 25 = 0.7772999 and of woemn aged 50 = 0.091123

Age at which probability of survival is 0.5:

$$
\begin{aligned}
\beta_0 + \beta_1 x &= \log(0.5/0.5) \\
\hat{x} &= \frac{\beta_0}{\beta_1}
\end{aligned}
$$

Thus, man's age at which probability of surviving is 0.5: 41.025641 and for woman: 20.5128205