Background Exercises GLM

Mean and variance of binary data

1 & 2 - on paper

Not clear about 3 and 4

Meaningful plots of dependence of binary targets on inputs

1. Missing dataset `binary_regression.txt`

2. Meaningful plots of dependence of binary targets on inputs

Download the dataset labelled_suspicious.txt. Come up with a plot that conveys useful information about the dependence (if any) of the target t on the features $\phi 1$ and $\phi 2$.

```
setwd('~/Box Sync/abarciausksas/myfiles/Statistical Modelling and Inference/week4')
data <- read.delim('labelled_suspicious.txt', sep = ' ')

phi1model <- lm(formula = data$t ~ 0 + data$phi1)
phi2model <- lm(formula = data$t ~ 0 + data$phi2)
phimodel <- lm(formula = data$phi1 ~ 0 + data$phi2)

summary(phi1model)
summary(phi2model)
summary(phimodel)
plot(data$phi1, data$phi2, ylim = c(-1,2), xlim = c(-1,2))

# This analysis shows evidence that phi1 and phi2 are highly correlated (phi1 is 0.5 of phi2
# So we select the
```