

Exam Template

Nick Thompson

10/10/2016

Pull in Data

```
setwd("~/Documents/GitHubRepo/729_Reed_MLE_git/Exam")
#voterid <- read.csv(file = "https://raw.githubusercontent.com/Neilblund/729A/master/data/voterid.csv",
#data <- read.dta(file = "https://raw.githubusercontent.com/Neilblund/729A/master/data/voterid.csv")
#data <- read.dta13(file = "https://raw.githubusercontent.com/Neilblund/729A/master/data/voterid.csv")
#save(voterid, file = "data.RData")
load("data.RData")
#View(voterid2)
```

Introduction:

Logit

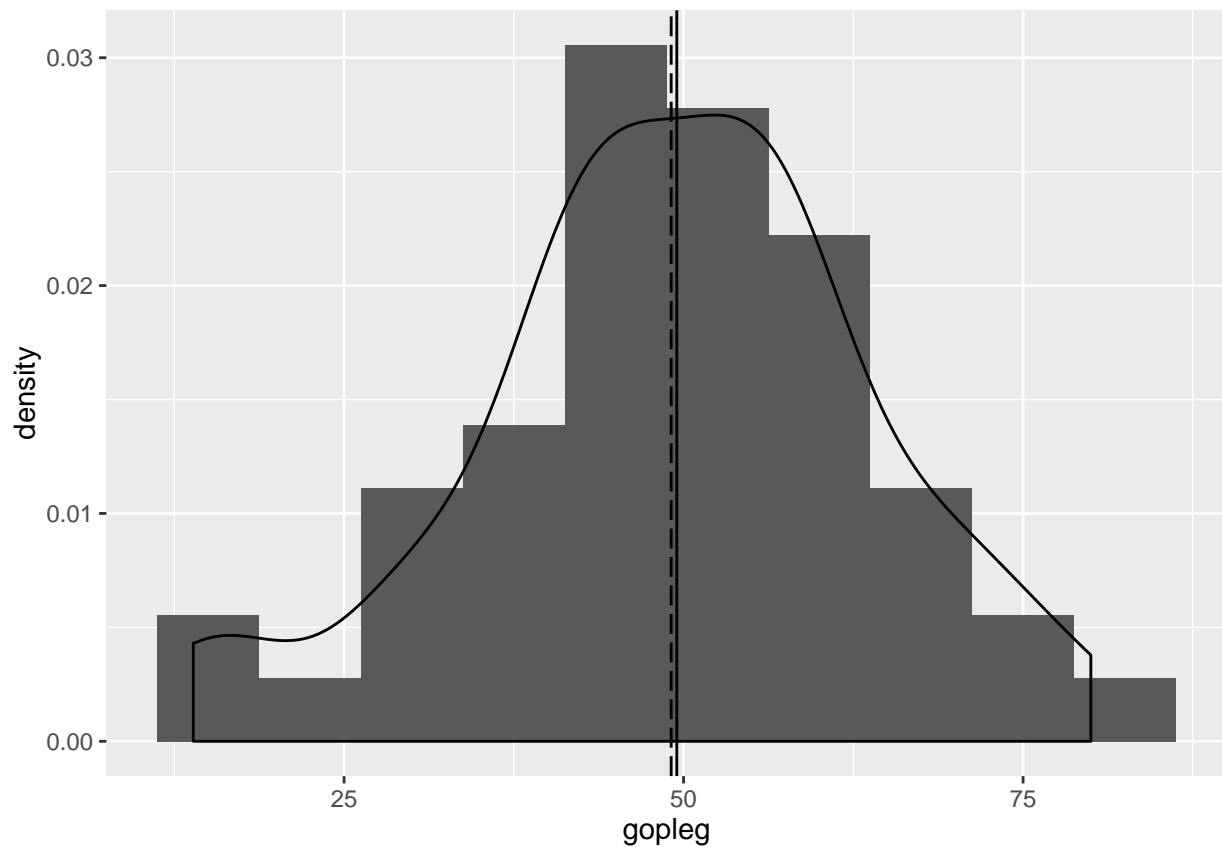
Probit

```
stargazer(voterid2, header = F)
```

Table 1:

Statistic	N	Mean	St. Dev.	Min	Max
photo	48	0.375	0.489	0	1
fraud	48	1.066	1.817	0.000	7.833
election_margin	48	15.400	10.827	4.911	49.431
gopleg	48	49.088	14.693	13.900	79.991
mean_gopleg	48	49.088	0.000	49.088	49.088
sd_gopleg	48	7.347	0.000	7.347	7.347
med_g	48	49.502	0.000	49.502	49.502

```
g <- ggplot(voterid2,aes(x=gopleg))
g + geom_histogram(aes(y=..density..),binwidth = 7.5) +
  geom_density() +
  geom_vline(xintercept = voterid2$mean_gopleg,linetype='longdash') +
  geom_vline(xintercept = voterid2$med_g)
```



```
gopleg_obs_low <- voterid2$mean_gopleg - voterid2$sd_gopleg
gopleg_obs_high <- voterid2$mean_gopleg + voterid2$sd_gopleg
```

```
# run probit, show results
```

```
(model_1 <- glm('photo ~ fraud + election_margin + gopleg',
  family = binomial(link = "probit"),
  data = voterid2))
```

```
##
```

```
## Call:  glm(formula = "photo ~ fraud + election_margin + gopleg", family = binomial(link = "probit"),
##      data = voterid2)
```

```
##
```

```
## Coefficients:
```

```
##      (Intercept)          fraud  election_margin          gopleg
##      -2.252268         0.143559        -0.001425         0.035747
```

```
##
```

```
## Degrees of Freedom: 47 Total (i.e. Null);  44 Residual
```

```
## Null Deviance:      63.51
```

```
## Residual Deviance: 54.75    AIC: 62.75
```

```
summary(model_1)
```

```
##
```

```
## Call:
```

```
## glm(formula = "photo ~ fraud + election_margin + gopleg", family = binomial(link = "probit"),
##      data = voterid2)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.4646  -0.9162  -0.6509   1.0391   2.4214
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -2.252268   0.828138  -2.720  0.00653 **
## fraud           0.143559   0.112251   1.279  0.20093
## election_margin -0.001425   0.021519  -0.066  0.94721
## gopleg         0.035747   0.016206   2.206  0.02739 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 63.510  on 47  degrees of freedom
## Residual deviance: 54.748  on 44  degrees of freedom
## AIC: 62.748
##
## Number of Fisher Scoring iterations: 5
```

```
#stargazer(model_1,type = 'text')

# generate predicted probabilities automatically
voterid2$pprob <- predict(model_1,type="response")

# generate predicted probabilities manually
voterid2$pprob_manual <- pnorm(model_1$coef['(Intercept)'] +
                               model_1$coef['fraud']*voterid2$fraud +
                               model_1$coef['election_margin']*voterid2$election_margin +
                               model_1$coef['gopleg']*voterid2$gopleg)

# test that we did it right
voterid2$pprob_test <- voterid2$pprob - voterid2$pprob_manual
summary(voterid2$pprob_test) # should be zeros
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##         0         0         0         0         0         0
```

```
#####1. calculate average effect of photo using observed values----
voterid2$pprob_gopleg_upsd <- pnorm(model_1$coef['(Intercept)'] +
                                    model_1$coef['fraud']*voterid2$fraud +
                                    model_1$coef['election_margin']*voterid2$election_margin +
                                    model_1$coef['gopleg']*gopleg_obs_high)
summary(voterid2$pprob_gopleg_upsd)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.3801  0.4035  0.4194  0.4574  0.4653  0.8036
```

```
voterid2$pprob_gopleg_downsd <- pnorm(model_1$coef['(Intercept)'] +
                                     model_1$coef['fraud']*voterid2$fraud +
                                     model_1$coef['election_margin']*voterid2$election_margin +
                                     model_1$coef['gopleg']*gopleg_obs_low)
summary(voterid2$pprob_gopleg_downsd)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.2031 0.2208 0.2331 0.2695 0.2702 0.6290
```

```
voterid2$pprob_effect <- voterid2$pprob_gopleg_upsd - voterid2$pprob_gopleg_downsd
summary(voterid2$pprob_effect)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.1746 0.1825 0.1855 0.1879 0.1921 0.2071
```

Summary: Interpreting the Coefficients, include the AIC

```
stargazer(model_1, header = F)
```

Table 2:	
	<i>Dependent variable:</i>
	NA
fraud	0.144 (0.112)
election_margin	−0.001 (0.022)
gopleg	0.036** (0.016)
Constant	−2.252*** (0.828)
Observations	48
Log Likelihood	−27.374
Akaike Inf. Crit.	62.748
<i>Note:</i> *p<0.1; **p<0.05; ***p<0.01	