# FTAP Homework 3

*Zachary Fogelson*

*July 15, 2015*

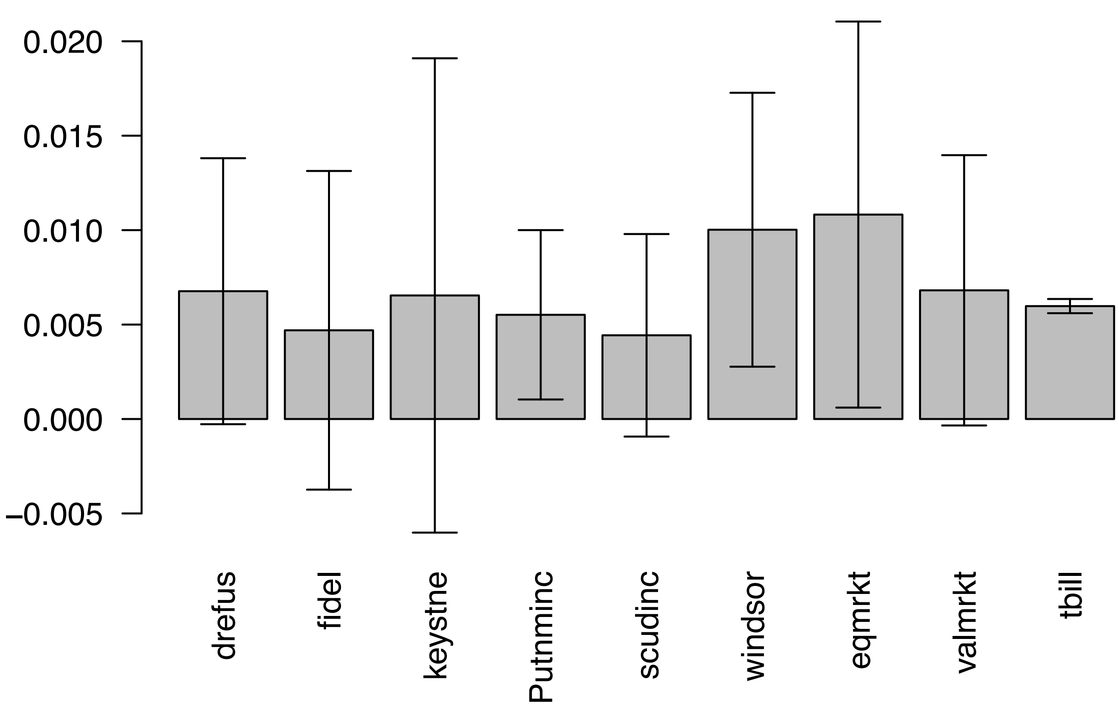
### Problem 1

**a** Build 95% confidence intervals

```
mfunds <- read.xls(xls = "mfunds (1).xls", sheet = 2)
sds <- sapply(mfunds, sd)
sds <- sds / sqrt(length(mfunds$drefus))
means <- sapply(mfunds, mean)
lower <- means - 2 * sds
upper <- means + 2 * sds
all <- rbind(means, upper, lower)
row.names(all) <- c("means","upper", "lower")
kable(head(all))
```

| | drefus | fidel | keystne | Putnminc | scudinc | windsor | eqmrkt | valmrkt | tbill |
|---|---|---|---|---|---|---|---|---|---|
| means | 0.0067669 | 0.0046968 | 0.0065424 | 0.0055174 | 0.0044325 | 0.0100219 | 0.0108249 | 0.0068133 | 0.0059783 |
| upper | 0.0138087 | 0.0131323 | 0.0190997 | 0.0100014 | 0.0097946 | 0.0172727 | 0.0210450 | 0.0139688 | 0.0063544 |
| lower | -0.0002748 | -0.0037388 | -0.0060149 | 0.0010335 | -0.0009296 | 0.0027711 | 0.0006048 | -0.0003422 | 0.0056022 |

```
barplot2(all["means",], las = 2, plot.ci = T, ci.u = upper, ci.l = lower)
```




**b** As shown in the plot above, much of the confidence intervals for the Drifus and the Fidel funds overlap making the sample means not very informative.

### Problem 2

Let $x_i$ be a random variable which is equal to the number of households with a phone.

Assume that the true probability of whether an American family has a phone is .39 and our sample size is 200.

Let Y be a binomial random variable:

$$Y = \Sigma_{i=1}^{200}(x_i)$$

Because the number of observations in our sample is >30 we assume that the CLT applies and results in Y being well approximated by a normal distrobution:

$$E[Y] = np = 200 * .39 = 78$$

$$Var[Y] = n * p(1 - p) = 47.58$$

#### *a*

Determine P(X < 70):

```r
pnorm(70, mean = 78, sd = sqrt(47.58))
```

```
## [1] 0.123068
```

#### *b*

Determine P(X > 90) = 1 - P(X < 90):

```r
1 - pnorm(90, mean = 78, sd = sqrt(47.58))
```

```
## [1] 0.04095773
```

### Problem 3

Null Hypothesis: $H_0 : p \leq .5$

Alternative: $H_1 : p > .5$

Based on the null hypothesis we assume that I am going to lose. To be confident, I take the highest value at which I would lose, namely .5.

Sample size: 2000, $\hat{p} = .53$

$$Z = \frac{\hat{p} - p}{\sqrt(\frac{p(1-p)}{2000})}$$

$$Z = \frac{.53 - .5}{\sqrt{\frac{.5^2}{2000}}} = 2.68$$

Therefore, the probability that I get a sample average of .53 with a true population of average of .5 after sampling 2000 people is the right hand tail of the normal cdf after 2.68.

```r
1 - pnorm(2.68, 0, 1)
```

```
## [1] 0.003681108
```

Because .0037 is less than 5% I reject the null hypothesis and with 95% confidence, I expect that I will win the election.

### Problem 4

Goal confidence interval should be less than $+/-2\%$ to any side:

$.02 = 2 * \hat{\sigma}$

$\hat{\sigma} = \sqrt{\frac{p(1-p)}{n}}$

In order to be sure that the confidence interval is not more than .02 assume that $p = .5$ because that is the maximum of the numerator of the $\hat{\sigma}$ term:

$.02 = 2 * \sqrt{\frac{.5^2}{n}}$

$\frac{1}{4*.01^2} = n = 2500$

### Problem 5

#### a

```r
z <- qnorm(.995)
se <- sqrt(.056*(1-.056) / 1200)
cat("Upperboud: ", .056 + z*se ,"\nLowerbound: ", .056 - z*se)
```

```
## Upperboud:  0.07309647
## Lowerbound:  0.03890353
```

**b** Null Hypothesis: $H_0 : p = .05$

Alternative: $H_1 : p \neq .05$

$$Z = \frac{\hat{p} - p}{\sqrt(\frac{p(1-p)}{1200})}$$

$$Z = \frac{.056 - .05}{\sqrt(\frac{.05(1-.05)}{1200})}$$

```r
z <- .006 / sqrt((.05*(1-.05))/1200)
2 * (1 - pnorm(z))
```

```
## [1] 0.3402541
```

Because the probability of getting a sample with a sample default rate of .056 given a true distrobution of .05 over 1200 observatoins is >30% and therefore greater than a 5% threshold I fail to reject the null hypothesis that .05 is the true default rate.

**c**

(Assuming that by "part a" this question means "part b") My p value for part b was ~.34

**d**

```
z <- .006 / sqrt((.05*(1-.05))/1200)
1 - pnorm(z)
```

```
## [1] 0.1701271
```

Because ~.17 is still greater than .05, I still cannot reject the null hypothesis at a 95% confidence interval.

**e**

Null Hypothesis: $H_0 : p_{Aa} - p_{Bb} < 0$

Alternative: $H_1 : p_{Aa} - p_{Bb} \geq 0$

$$Z = \frac{(\hat{p} - p) - 0}{\sqrt(\frac{p(1-p)}{1200})}$$

$$Z = \frac{(.048 - .056) - 0}{\sqrt(\frac{.056(1-.056)}{1200} + \frac{.048(1-.048)}{1200})}$$

```
z <- ((.048 - .056)  - 0) / sqrt((.056*(1-.056))/1200 + (.048*(1-.048))/1200)
pnorm(z)
```

```
## [1] 0.1886898
```

Supposing that the difference of Aa bonds and Bb bonds are less than 0 then at the likelihood of two samples .48, .56 is ~18% therefore we cannot reject the null hypothesis that the probability of an Aa default is less than the probability of a Bb default.

### Problem 6

**a**

Null Hypothesis: $H_0$ : eye level - bottom shelf $= 0$

Alternative: $H_0$ : eye level - bottom shelf $\neq 0$

$$t = \frac{121 - 0}{\frac{344}{\sqrt{36}}}$$

```
t <- (121 - 0) / (344/sqrt(36))
2 * (1 - pt(t, 36 - 1))
```

```
## [1] 0.04203608
```

Because the probability of achieving a difference of 121 sales in 36 weeks given a t distrobution with 35 degrees of freedom is <5% we are confident that we can reject the null hypothesis that there is no difference between the eye level and bottom shelf

**b**

As seen above the pvalue is 0.04203608