Homework Problem 5.3

Frank Woodling

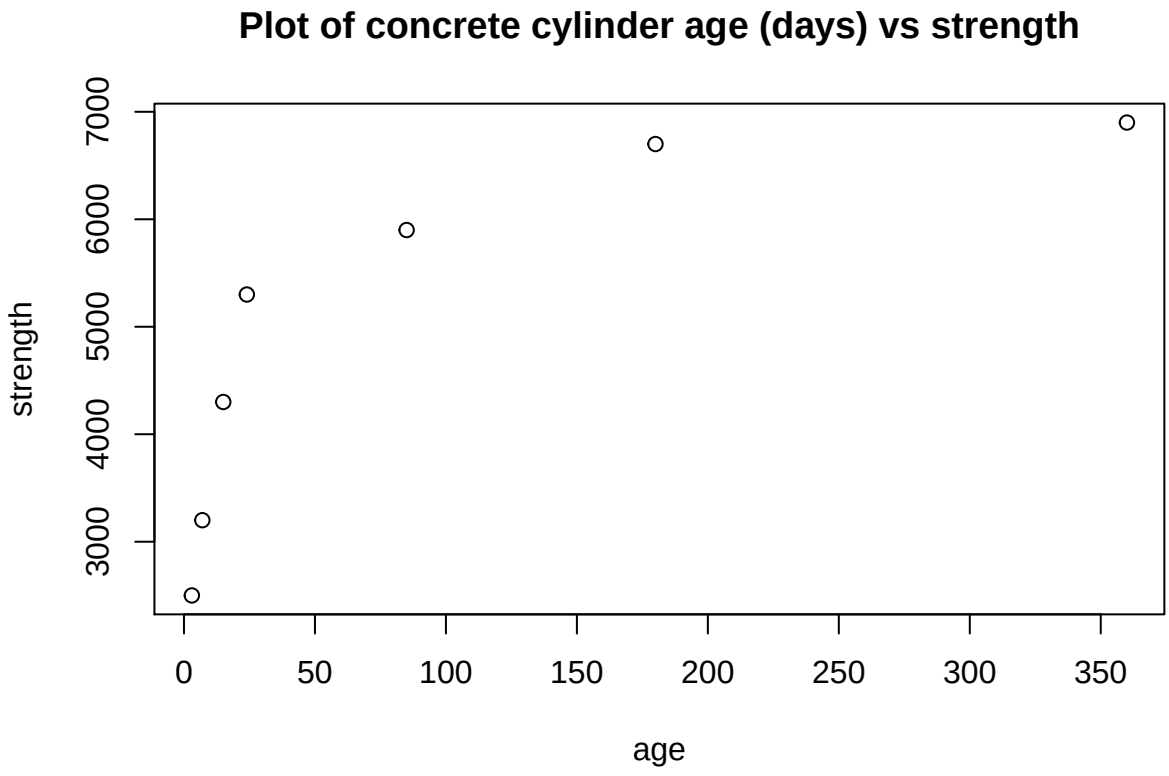
April 22, 2016

1.

a.

```
age <- c(3, 7, 15, 24, 85, 180, 360)
strength <- c(2500, 3200, 4300, 5300, 5900, 6700, 6900)

plot(age, strength, main = "Plot of concrete cylinder age (days) vs strength")
```




```
# Pearson's correlation
x_bar = mean(age)
y_bar = mean(strength)

num = 0
denom_x = 0
denom_y = 0;
```

```
for(i in 1:length(age))
{
  num = num + (age[i]-x_bar)*(strength[i]-y_bar)
  denom_x = denom_x + (age[i]-x_bar)^2
  denom_y = denom_y + (strength[i]-y_bar)^2
}

r = num/sqrt(denom_x*denom_y)
```

$$\frac{\sum_{i=1}^{n}(X_i-\overline{X})(Y_i-\overline{Y})}{\sqrt{\sum_{i=1}^{n}(X_i-\overline{X})^2\sum_{i=1}^{n}(Y_i-\overline{Y})^2}} = 0.7858418$$

```
# Spearman's correlation
cor(age, strength, method = "spearman")
```

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

```
# Kendall's Tau
cor(age, strength, method = "kendall")
```

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

b.

```
source("http://www4.stat.ncsu.edu/~lu/ST505/Rcode/functions-Ch5.R")

        ######################################
        #  Permutation test for rho/beta1   #
        ######################################
x <- age
y <- strength

### direct function for calculating the sample correlation: r
(r = cor(x,y))
```

```
## [1] 0.7858418
```

```
### least squares coefficient estimates
(lm(y~x))
```

```
##
## Call:
## lm(formula = y ~ x)
##
## Coefficients:
## (Intercept)            x
##      3999.0         10.1
```

```
#### calculate r and beta1hat manually:
(sumx = sum(x))
```

```
## [1] 674
```

```
(sumy = sum(y))
```

```
## [1] 34800
```

```
(sumxy = sum(x*y))
```

```
## [1] 4413100
```

```
(sumxx = sum(x^2))
```

```
## [1] 170084
```

```
(sumyy = sum(y^2))
```

```
## [1] 190380000
```

```
(n = length(x))
```

```
## [1] 7
```

```
xbar = sumx/n
ybar = sumy/n
Sxy = 1/(n-1)*(sumxy - n*xbar*ybar)
Sx2 = 1/(n-1)*(sumxx - n*xbar^2)
Sy2 = 1/(n-1)*(sumyy - n*ybar^2)
Sx = sqrt(Sx2)
Sy = sqrt(Sy2)

(robs = Sxy/(Sx*Sy))
```

```
## [1] 0.7858418
```

```
(bhat1.obs = Sxy/(Sx2))
```

```
## [1] 10.09966
```

```
### Test H0:r=0 versu Ha:r>0

##
##### t-test assuming normality
##
tt = sqrt((n-2)/(1-robs^2))*robs
1-pt(tt,n-2)
```

```
## [1] 0.01809361
```

```
##
##### large sample approximation
##
(Zr = sqrt(n-1)*r)
```

```
## [1] 1.924912
```

```
1-pnorm(Zr)
```

```
## [1] 0.02712022
```

```
permr <- perm.approx.r(x,y,1000)
mean(permr >= robs)
```

```
## [1] 0
```

```
mean(abs(permr) >= abs(robs))
```

```
## [1] 0.02
```

```
        ######################################
        #  Spearman Correlation (ties)      #
        ######################################

x = age
y = strength

(x = rank(x))  ###no need for this data as they are already ranked
```

```
## [1] 1 2 3 4 5 6 7
```

```
(y = rank(y))
```

```
## [1] 1 2 3 4 5 6 7
```

```
## Spearman correlation
(rs.obs = cor(x, y))
```

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

```
## permutation test for the Spearman correlation
permr <- perm.approx.r(x, y, 1000)
mean(permr >= rs.obs)
```

```
## [1] 0.001
```

```r
mean(abs(permr) >= abs(rs.obs))
```

```
## [1] 0.001
```

```r
##### data set: scores of ten projects at a Science Fair (Table 5.2.2 of Higgins)
(x = rank(x))
```

```
## [1] 1 2 3 4 5 6 7
```

```r
(y = rank(y))
```

```
## [1] 1 2 3 4 5 6 7
```

```r
## Spearman correlation
(rs.obs = cor(x, y))
```

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

```r
## permutation test for the Spearman correlation
permr <- perm.approx.r(x, y, 1000)
mean(permr >= rs.obs)
```

```
## [1] 0
```

```r
mean(abs(permr) >= abs(rs.obs))
```

```
## [1] 0
```

```r
# Kendall's Tau
x=age
y=strength

tauobs <- getTau(x,y)
tauobs
```

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

```r
permtau <- perm.approx.tau(x,y,1000)
mean(permtau >= tauobs)
```

```
## [1] 0
```