# BMEG 802 – Advanced Biomedical Experimental Design and Analysis

## Assignment 3

Joshua G. A. Cashaback, PhD

# Question 1

$$\frac{\partial SS}{\partial B_0} = -2\sum_{i=1}^{n}(y_i - B_0 - B_1 \cdot x_{i1} - B_2 \cdot x_{i2}) = 0$$

$$\frac{\partial SS}{\partial B_1} = -2\sum_{i=1}^{n}(y_i - B_0 - B_1 \cdot x_{i1} - B_2 \cdot x_{i2})(x_{i1}) = 0$$

$$\frac{\partial SS}{\partial B_2} = -2\sum_{i=1}^{n}(y_i - B_0 - B_1 \cdot x_{i1} - B_2 \cdot x_{i2})(x_{i2}) = 0$$

# Question 2

```
grf_walk = c(31.5, 33.3, 32.3, 28.8, 38.3, 36.9, 14.6, 27, 32.8, 27.4, 31.5
grf_trot = c(50.8, 43.2, 44.8, 39.5, 44, 60.1, 11.1, 32.3, 41.3, 38.2, 50.8
m2 <- lm(grf_trot ~ grf_walk)
B1 = sum((grf_walk - mean(grf_walk)) * (grf_trot - mean(grf_trot))) / sum(
B0 = mean(grf_trot) - B1 * mean(grf_walk)
B0
```

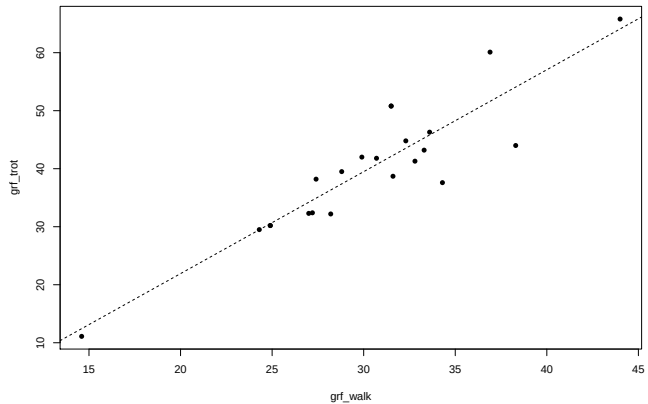
```
## [1] -13.22436
```

```
B1
```

```
## [1] 1.75709
```

# Question 2 cont'd

```
plot(grf_walk, grf_trot, pch=16)
abline(m2, lty=2)
```



# Question 2 cont'd

```
SE = sqrt(sum((grf_trot - (B0 + B1 * grf_walk))^2) / (length(grf_walk) - 2)
SE
```

```
## [1] 4.979982
```

```
cor(grf_walk, grf_trot)
```

```
## [1] 0.9044905
```

```
cor(grf_walk, grf_trot)^2
```

```
## [1] 0.8181031
```

```
summary(m2)$coefficients[2,4] # p-value
```

```
## [1] 7.648108e-09
```

# Question 3

install.packages("MASS")

```
library("MASS")
my_n1 <- 1000
my_mu1 <- c(78.8, 211)
my_Sigma1 <- matrix(c(3.668^2, 0.81*3.668*26.904, 0.81*3.668*26.904, 26.904
binorm <- mvrnorm(n = my_n1, mu = my_mu1, Sigma = my_Sigma1)
m3 <- lm(binorm[,2] ~ binorm[,1])
summary(m3)$coefficients[1] # B_0, coefficient
```

```
## [1] -241.9788
```

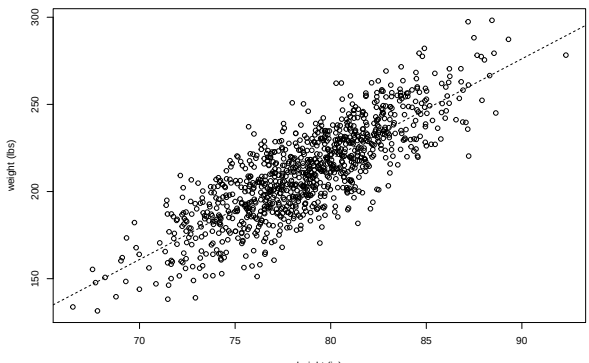
```
summary(m3)$coefficients[2] # B_1, coefficient
```

```
## [1] 5.756826
```

# Question 3 Con't

install.packages("MASS")

```
plot(binorm, xlab="height (in)", ylab="weight (lbs)")
abline(m3, lty=2)
```



# Question 3 Con't

install.packages("MASS")

```
cor(binorm[,1], binorm[,2])
```

```
## [1] 0.8119643
```

```
cor(binorm[,1], binorm[,2])^2
```

```
## [1] 0.6592861
```

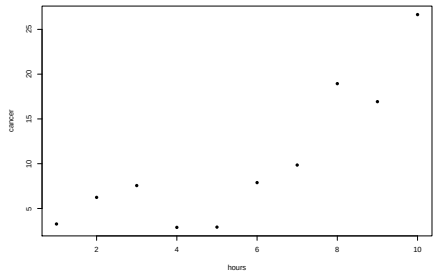
```
summary(m3)$coefficients[2,4] # p-value
```

```
## [1] 1.429025e-235
```

# Question 4

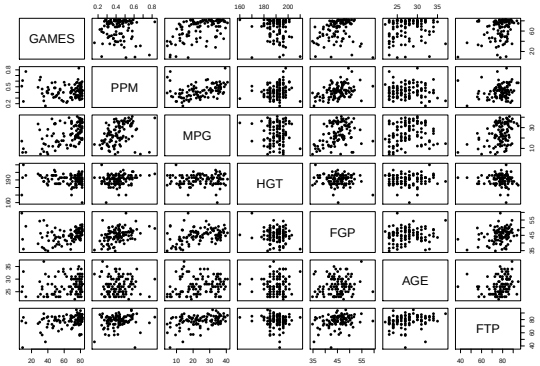
```
hours = c(1, 2, 3, 4, 5, 6, 7, 8, 9, 10)
cancer = c(3.27, 6.24, 7.56, 2.89, 2.92, 7.89, 9.85, 18.94, 16.93, 26.65)
plot(hours, cancer, pch=16)
cor.test(hours, cancer, method = "spearman")$p.value #p-value
```

```
## [1] 0.008235571
```



# Question 5

```
bball_data <- read.csv("https://raw.githubusercontent.com/joshcash9/Statis
pairs(bball_data, pch=16)
```



# Question 6

```
cor(bball_data)
```

```
##              GAMES         PPM         MPG         HGT         FGP         AGE
## GAMES   1.00000000 -0.05981649  0.52330852 -0.17170635  0.1927063  0.15522899
## PPM    -0.05981649  1.00000000  0.35621592  0.21335254  0.4063189 -0.04419444
## MPG     0.52330852  0.35621592  1.00000000 -0.01043649  0.3396098  0.18013546
## HGT    -0.17170635  0.21335254 -0.01043649  1.00000000 -0.1080561  0.06996434
## FGP     0.19270626  0.40631889  0.33960978 -0.10805610  1.0000000  0.10839243
## AGE     0.15522899 -0.04419444  0.18013546  0.06996434  0.1083924  1.00000000
## FTP     0.31067275  0.16552229  0.39141431 -0.06133401  0.2785525  0.24735655
##                FTP
## GAMES   0.31067275
## PPM     0.16552229
## MPG     0.39141431
## HGT    -0.06133401
## FGP     0.27855246
## AGE     0.24735655
## FTP     1.00000000
```

MPG is the best predictor of FTP

# Question 6 cont'd

```r
m5 <- lm(FTP ~ 1, data=bball_data)
mall <- lm(FTP ~ GAMES + MPG + HGT + PPM + AGE + FGP, data=bball_data)
mbest <- step(m5, list(lower=m5, upper=mall), direction="both")
```

```
## Start:  AIC=466.49
## FTP ~ 1
##
##         Df Sum of Sq    RSS    AIC
## + MPG    1   1341.70 7415.8 451.03
## + GAMES  1    845.26 7912.3 457.83
## + FGP    1    679.51 8078.0 460.01
## + AGE    1    535.83 8221.7 461.86
## + PPM    1    239.94 8517.6 465.57
## <none>               8757.5 466.49
## + HGT    1     32.94 8724.6 468.09
##
## Step:  AIC=451.03
## FTP ~ MPG
##
##         Df Sum of Sq    RSS    AIC
## + AGE    1    283.08 7132.8 448.94
## + FGP    1    209.93 7205.9 450.01
## <none>               7415.8 451.03
## + GAMES  1    135.11 7280.7 451.10
## + HGT    1     28.71 7387.1 452.62
## + PPM    1      6.83 7409.0 452.93
## - MPG    1   1341.70 8757.5 466.49
##
```

# Question 6 cont'd

```
summary(mbest)
```

```
##
## Call:
## lm(formula = FTP ~ MPG + AGE + FGP, data = bball_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -34.013  -3.349   1.519   4.704  14.636
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 43.12593   10.34628   4.168 6.49e-05 ***
## MPG          0.28590    0.08876   3.221  0.00172 **
## AGE          0.47727    0.24575   1.942  0.05490 .
## FGP          0.32438    0.19711   1.646  0.10294
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.293 on 101 degrees of freedom
## Multiple R-squared:  0.2068, Adjusted R-squared:  0.1832
## F-statistic: 8.777 on 3 and 101 DF,  p-value: 3.157e-05
```

The model signifanctly predicts FTP (p = 3.157e-05)

# Question 6 cont'd

$$FTP = 43.13 + 0.29 \cdot MPG + 0.48 \cdot AGE + 0.32 \cdot FGP$$

```
43.13 + 0.29 * 20 + 0.48 * 23 + 0.32 * 60
```

```
## [1] 79.17
```