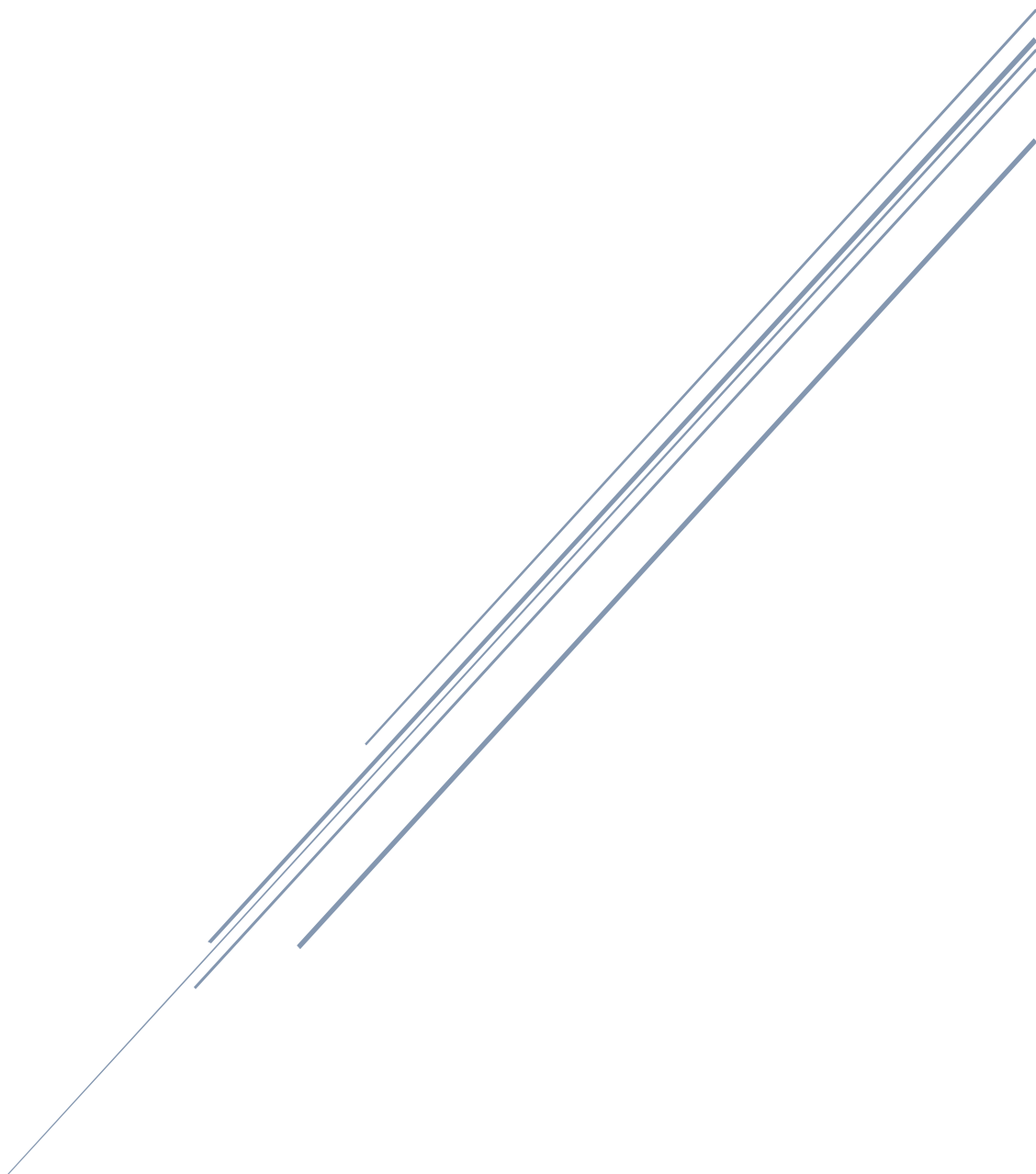


# 応用数学 課題 6

HI5 35 番 松山京介



問1. 前回の課題で、身長を目的変数、体重を説明変数にした散布図を描き、回帰直線を引いた。この回帰直線上に予測値をプロットし、さらに、もとのデータとの残差を線分で示せ。※色を変えて分かりやすくすること。

## プログラム

```
data <- read.csv("C:/Users/gomad/Desktop/応用数学/課題6/taikaku.csv")
x <- data$taiju
y <- data$shincho

# 相関係数
print(cor(x, y))

# plot
plot(x, y,
      xlab="身長[cm]", ylab="体重[kg]", main="身長と体重の関係",
      pch=21, cex=3,
      bg=ifelse(x>70,
                ifelse(y>175, "springgreen", "skyblue"),
                ifelse(y>175, "yellow2", "hotpink")
      ))
text(x, y)
abline(h=175, lty="dotted")
abline(v=70, lty="dotted")

# 回帰分析
rg <- lm(y ~ x)
abline(rg, col="red", lwd=2, lty="dashed")
print(rg$coefficient[1]) # 切片
print(rg$coefficient[2]) # 傾き

# 予測値
x0 <- seq(min(x), max(x), 5)
xf <- data.frame(x=x0)
pr <- predict(rg, xf, interval="prediction")
```

```

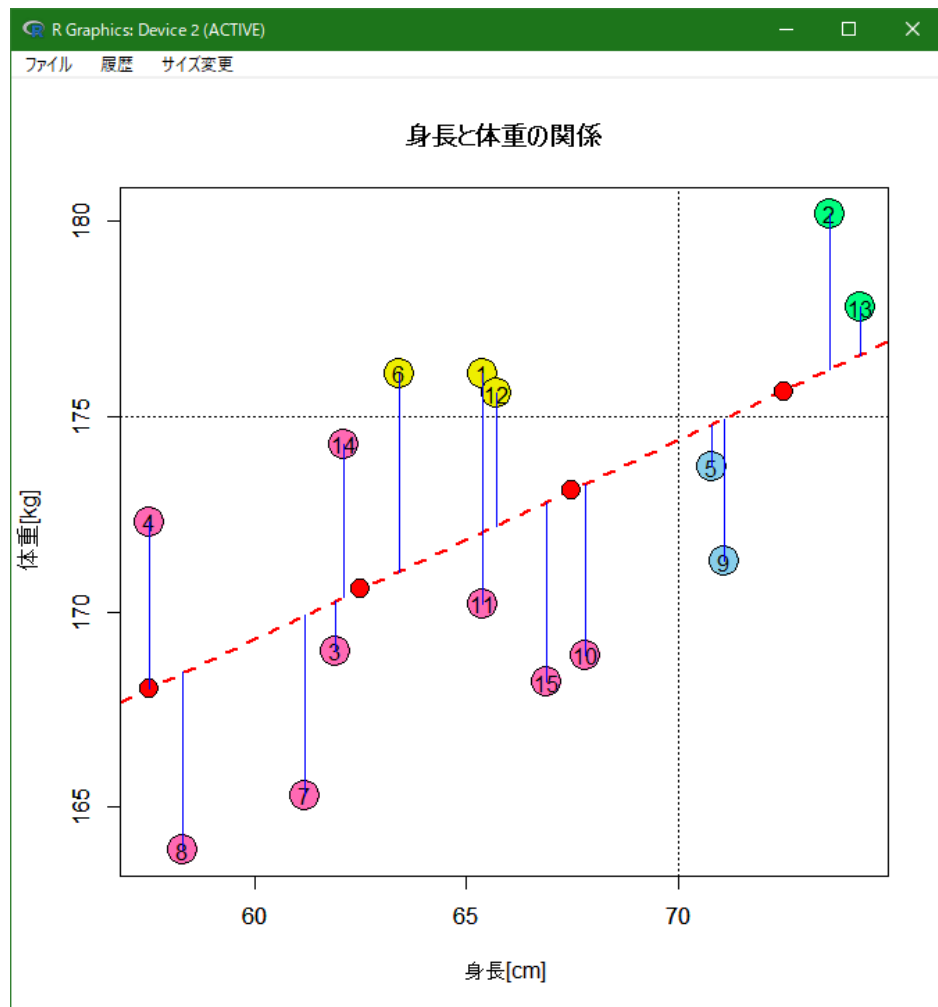
par(new=T)
plot(x0, pr[,1],
     xlab="", ylab="",bg="red",pch=21,cex=2,
     xlim=c(min(x), max(x)), ylim=c(min(y), max(y)))

# 残差
print(residuals(rg))

segments(
  x, y,
  x, predict(rg, data.frame(x=x), interval="prediction")[,1],
  col="blue"
)

```

実行結果



[1] 0.5634612

(Intercept)

138.8397

x

0.507717

1 2 3 4 5 6 7 8

4.055597 3.992317 -1.267393 4.266562 -1.086075 5.071031 -4.611991 -4.539612

9 10 11 12 13 14 15

-3.638390 -4.362924 -1.844403 3.403282 1.236915 3.931063 -4.605979

結果より、回帰直線上に予測値をプロットすることができており、また、元のデータとの残差を線分で示すことができています。

問2. 問1の残差を関数 `residuals` で求めるやり方と、実際に残差を計算で求めるやり方の2通りで求め、一致することを確認しよう。

## プログラム

```
data <- read.csv("C:/Users/gomad/Desktop/応用数学/課題6/taikaku.csv")
x <- data$taiju
y <- data$shincho

regi <- function(x) 0.507717 * x + 138.8397
print(y - regi(x))
```

## 実行結果

[1] 4.055608 3.992329 -1.267382 4.266573 -1.086064 5.071042 -4.611980

[8] -4.539601 -3.638379 -4.362913 -1.844392 3.403293 1.236927 3.931074

[15] -4.605967

問1において残差を関数 `residuals` を用いて求めていたため、その出力結果と比較する。

問1の回帰分析において求めた切片と傾きから関数を定義し残差を求めたところ、問1で求めた結果とほぼ一致する結果が得られた。

問3. この直線によるモデル化の決定係数および調整された決定係数を計算によって求め、summaryの結果と比較し、一致することを確認めよ。

#### プログラム

```
data <- read.csv("C:/Users/gomad/Desktop/応用数学/課題6/taikaku.csv")
x <- data$taiju
y <- data$shincho

#決定係数
rs <- residuals(lm)
rss <- sum(rs ^ 2)
print(1 - rss / sum((y-mean(y))^2))

#調整された決定係数
n <- length(x)
df <- 2 - 1
print(1 - (rss / (n-df-1)) / (sum((y-mean(y))^2) / (n-1)))
```

#### 実行結果

```
[1] 0.3174885
```

```
[1] 0.2649876
```

課題5において決定係数は0.3175、調整された決定係数は0.265であった。  
よって結果より、summaryの結果と一致する決定係数が得られた事がわかる。

問4. 身長と体重の相関係数をcorにより求めて、その2乗が決定係数と一致することを確認めよ。

#### プログラムと実行結果

```
> cor(x,y)^2
[1] 0.3174885
```

実行結果より、問3で求めた決定係数と等しい結果が得られた。