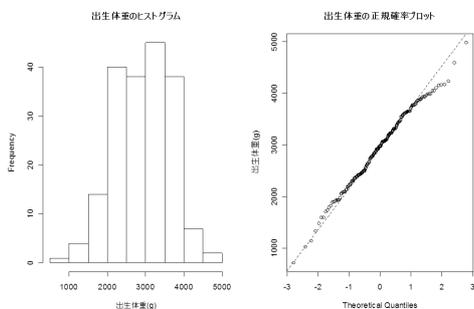


## 前回の課題の回答例

c05a.R

```
library(MASS)
attach(birthwt)
layout(t(1:2))
hist(bwt,main="出生体重のヒストグラム",xlab="出生体重 (g)")
qqnorm(bwt,main="出生体重の正規確率プロット",ylab="出生体重 (g)")
qqline(bwt,lty=2)
shapiro.test(bwt)
source("http://phi.med.gunma-u.ac.jp/msb/msb-funcs.R")
geary.test(bwt)
detach(birthwt)
```

結果は次のグラフと枠内の通り（情報量のない行は削除済み）。グラフを見るだけでも正規分布に近そうだと見当はつく。何度も繰り返すが、検定よりも先に作図することは非常に重要である。



Shapiro-Wilk normality test

```
data: bwt
W = 0.9924, p-value = 0.4354
Geary's test for normality:
G = 0.8126568 / p = 0.1693836
```

「出生体重データは正規分布にしたがう」という帰無仮説を、Shapiro-Wilk の検定と Geary の検定により、有意水準 5% で検定した結果、有意確率 (p 値) が 0.05 よりずっと大きいので、有意水準 5% で帰無仮説は棄却できない。よって、とりあえず正規分布にしたがっているとみなしてよい。

## 本日の課題

MASS ライブラリに含まれているデータフレーム `anorexia` は拒食症患者 72 人を 3 群（対照群 Cont, 認知行動療法群 CBT, 家族療法群 FT）にわけて、治療開始前の体重 (`Prewt`), 治療開始後の体重 (`Postwt`) を記録したものである。体重の単位はポンド (lbs) である\*1。次の枠内を打てば、対照群だけの体重データが `Cx` に、家族療法群だけの体重データが `Fx` に得られ、`anox` にそれらを結合したデータが得られる（即ち、`anox` は、元のデータから認知行動療法群のデータを除いたものである）。

<http://phi.med.gunma-u.ac.jp/medstat/kadai6.R>

```
library(MASS)
Cx <- subset(anorexia,Treat=="Cont",select=c("Prewt","Postwt"))
Fx <- subset(anorexia,Treat=="FT",select=c("Prewt","Postwt"))
anox <- data.frame(
  Treat=factor(c(rep("Cont",length(Cx$Prewt)),rep("FT",length(Fx$Prewt)))),
  Prewt=c(Cx$Prewt,Fx$Prewt),Postwt=c(Cx$Postwt,Fx$Postwt))
```

このデータを使って、対照群と家族療法群で治療前後の体重変化を比較せよ。まず図示をしてから、適切な 2 群の平均値の差の検定をすること。他にも考察すべき内容があれば考察せよ。なお、検定の有意水準は 5% とする。図を Word または Powerpoint などに貼り付け、検定結果について考察する文章と学籍番号も打って印刷し、署名して提出すること。

\*1 データのオリジナルソースである Hand DJ et al. (1993) A Handbook of Small Data Sets. Chapman & Hall, data set 285. には kg と書かれているが誤りである。1 lb は約 0.4536 kg である。