

前回の課題の回答例

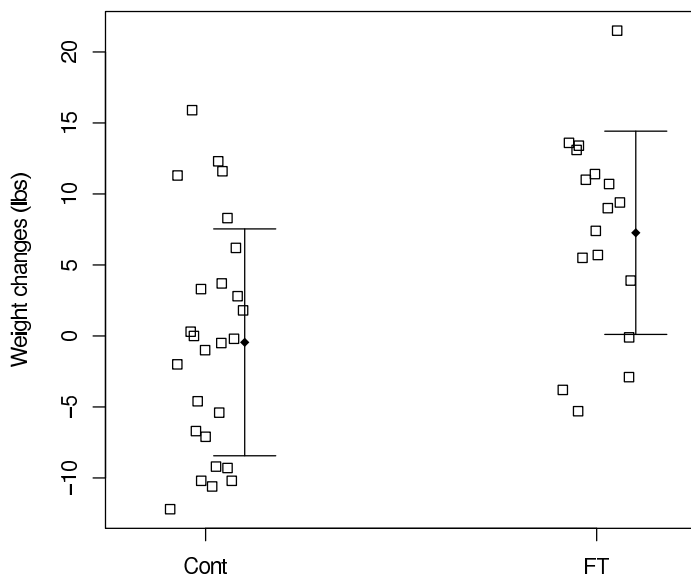
```
http://phi.med.gunma-u.ac.jp/medstat/kadai6.R  
library(MASS)  
Cx <- subset(anorexia,Treat=="Cont",select=c("Prewt","Postwt"))  
Fx <- subset(anorexia,Treat=="FT",select=c("Prewt","Postwt"))  
anox <- data.frame(  
  Treat=factor(c(rep("Cont",length(Cx$Prewt)),rep("FT",length(Fx$Prewt)))),  
  Prewt=c(Cx$Prewt,Fx$Prewt),Postwt=c(Cx$Postwt,Fx$Postwt))
```

に続けて、以下を入力する。

```
http://phi.med.gunma-u.ac.jp/medstat/kadai6a.R  
attach(anox)  
stripchart((Postwt-Prewt)~Treat,method="jitter",vert=T,  
  main="Weight changes by treatment",ylab="Weight changes (lbs)")  
Mx <- tapply(Postwt-Prewt,Treat,mean)  
Sx <- tapply(Postwt-Prewt,Treat,sd)  
Ix <- c(1.1,2.1)  
points(Ix,Mx,pch=18)  
arrows(Ix,Mx-Sx,Ix,Mx+Sx,angle=90,code=3)  
print(vr <- var.test((Postwt-Prewt)~Treat))  
t.test((Postwt-Prewt)~Treat,var.equal=(vr$p.value>=0.05))  
detach(anox)
```

結果は次のグラフと枠内の通り（情報量の少ない行は削除済み）

Weight changes by treatment



```
F test to compare two variances  
data: (Postwt - Prewt) by Treat  
F = 1.2458, num df = 25, denom df = 16, p-value = 0.6587  
alternative hypothesis: true ratio of variances is not equal to 1  
95 percent confidence interval:  
 0.4766073 2.9699299  
sample estimates:  
ratio of variances  
 1.245775
```

`var.test()` で実行される F 検定の結果の p -value（有意確率）は 0.6587 なので、2 群の分散に有意差はない。

```
Two Sample t-test
data: (Postwt - Prewt) by Treat
t = -3.2227, df = 41, p-value = 0.002491
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -12.549248 -2.880164
sample estimates:
mean in group Cont    mean in group FT
      -0.450000         7.264706
```

`t.test()` で実行される t 検定の結果, p -value (有意確率) は 0.002491 なので, 家族療法群と対照群では治療前後の体重変化が有意水準 5% で有意に異なるといえる。

本日の課題

<http://phi.med.gunma-u.ac.jp/medstat/bmi.txt> は, 熱帯のある集団における村ごとの 18 歳以上男女の BMI (Body Mass Index) データである (但し実習用に, 何人か, 実際にはない値を加え, 実際に観察した値を削除した)。

`dat <- read.delim("http://phi.med.gunma-u.ac.jp/medstat/bmi.txt")` とすればデータフレーム `dat` に読み込むことができ, 村を表す変数が `VG`, 性別を表す変数が `SEX`, BMI を表す変数が `BMI`, 年齢を表す変数が `AGE` となっている。

ここで, `males <- subset(dat, SEX=="M")` とすれば, 男性だけのデータを得ることができる。この男性のデータについて, 村によって BMI に差があるか, あるとしたら何番の村と何番の村の間に差があるかを検討せよ。まず層別箱ヒゲ図やストリップチャートを作ってから適切な検定を行なうこと。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。