

## 医学情報処理演習第 11 回「量的データのノンパラメトリックな分析」2007 年 12 月 17 日

\* 前回の課題は、やっていない人が多いので、解答例は年明けに web サイトで公表します。

### 今回の課題

第 7 回「一元配置分散分析と多重比較」の課題でも利用したが、<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、`females <- subset(dat,SEX=="F")` とすれば、女性だけのデータを得ることができる。この女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、ノンパラメトリックな方法で検討せよ。まず層別箱ヒゲ図や村別のヒストグラム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。