

医学情報処理演習第7回課題レポート用紙

学籍番号

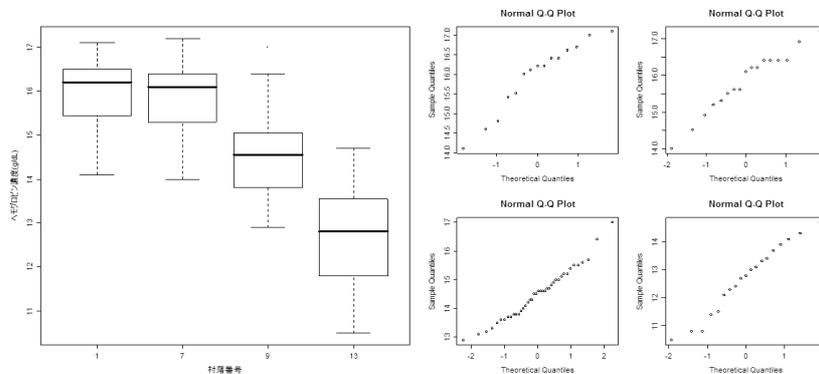
氏名

以下のようにして、まずデータを読み込み(1行目)、attachしてから(2行目)変数 VIL を要因型に変換する(3行目)ことで、分析の準備が整う。

```
dat <- read.delim("http://phi.med.gunma-u.ac.jp/medstat/p07.txt")
attach(dat)
VIL <- as.factor(VIL)
```

次に下枠内のように画面を5分割して、左半分に村落ごとの層別箱ひげ図を描き、右半分の4区分に村落ごとの正規確率プロットを描く。層別箱ひげ図をみると、村落間でヘモグロビン濃度には差がありそうに見える。正規確率プロットはどれも概ね直線状に見える。

```
layout(matrix(c(1,1,1,1,2,4,3,5),nr=2))
boxplot(HB ~ VIL, ylab="ヘモグロビン濃度 (g/dL)", xlab="村落番号")
tapply(HB, VIL, qqnorm)
```



そこで以下のように、村落ごとに Shapiro-Wilk の検定をする。

```
tapply(HB, VIL, shapiro.test)
```

下枠内のように、どの村でもヘモグロビン濃度が正規分布に従っているという帰無仮説は棄却されない結果になる(関係部分のみ表示した)

```
.$'1' W = 0.9274, p-value = 0.2492
.$'7' W = 0.9597, p-value = 0.6255
.$'9' W = 0.975, p-value = 0.5104
.$'13' W = 0.9637, p-value = 0.6463
```

そこで(以下を各自続けよ),