

■ 多標本問題 ◆ 北野 利一 ◆ 2005 年 10月 25日

● 多標本の比較において、問題となるのは、平均の違いがあるか、どうかである。注意しなければならないのは、この問題では、どの標本の平均が他の標本の標本とズレているか、という議論は二の次の議論であるということである。したがって、帰無仮説は「全ての平均は等しい」とおく、2標本問題の拡張である。

(以下での R のコード等は、Crawley (2002) pp.243-279, Dalgaard (2002) pp.111-120, Heiberger and Holland (2004) pp.123-153 を参考にした)

例 1 : 露光時間の違いにより、ある植物 (?) の発育に違いは見られるか?

```
# access: http://www.bio.ic.ac.uk/research/mjcraw/statcomp/
(read.table("MacOSX3:Users:tk:rm171:crawley1:oneway.txt", header=T) -> grow)
tapply(grow$G, grow$P, mean) # NOT pretty!

# and confirm it by mean(grow$G[grow$P == "Long"])
grow$Photoperiod <- factor(grow$P, level=c("Very.short", "Short", "Long", "Very.long"))
tapply(grow$G, grow$P, mean) # beauty!
xbar <- .Last.value

(unstack(grow) -> growun) # stack(growun) # try, too,
# and find something different from the original (grow)
boxplot(growun)
abline(h=mean(grow$G), col="red", lty=2)

stripchart(grow$Growth ~ grow$Photoperiod, meth="stack", ver=T)
abline(h=mean(grow$G), col="red", lty=2)
lines(1:4, xbar, pch=4, type="b", cex=2)

oneway.test(Growth ~ Photoperiod, data=grow)
```

出力に、(not assuming equal variances) とあるように、t.test における Welch の検定を応用したものである (oneway.test および t.test いずれにおいても、var.equal = FALSE がデフォルト)。2標本の var.test に対応する、多標本を対象とした検定は、bartlett.test である。

```
bartlett.test(Growth ~ Photoperiod, data=grow)
oneway.test(Growth ~ Photoperiod, data=grow, var.eq=T)

summary(aov(Growth ~ Photoperiod, data=grow)) # aov = Analysis Of Variance
anova(lm(Growth ~ Photoperiod, data=grow)) # anova = ANalysis Of VAriance
```

多標本問題、すなわち、分散分析における理屈は、

$$\text{標本の値} = \text{総平均} + (\text{群平均} - \text{総平均}) + (\text{標本の値} - \text{群平均})$$

と分解し、それぞれの偏差量の自乗和について、以下の関係が成立することを用いる。

$$(\text{標本の値} - \text{総平均}) \text{の自乗和} = (\text{群平均} - \text{総平均}) \text{の自乗和} + (\text{標本の値} - \text{群平均}) \text{の自乗和}$$

すなわち、『全変動 = 級間変動 + 級内変動』に分解できて、級間変動および級内変動がカイ自乗分布に従うことが示される (コクランの定理)。カイ自乗の比がエフ分布に従うことにより、分散分析でエフ値が統計量となっている。なお、上述の自乗和の分解は、一種のピタゴラスの定理といえる。

- なお、2標本問題は多標本問題に含まれる。このことは以下のように確認できる。

例2：2地点で採取された鳥の翼の長さについての差異

```
# access: http://www.bio.ic.ac.uk/research/crawley/statistics/
read.table("MacOSX3:Users:tk:rm171:crawley2:wings.txt", header=T) -> wings
(oneway.test(size ~ location, data=wings) -> wings.onewaytest)

unstack(wings) -> wng; (t.test(wng$A, wng$B) -> wings.ttest)
```

- ▲ ところで、エフ分布について、以下の性質が知られている。

t が自由度 k のティ分布 $t(k)$ に従うとき、 t^2 はエフ分布 $F(1, k)$ に従う
(例えば、東京大学教養学部統計学教室 編 (1991), p.208 を見よ)

```
pp <- c(0.1, 0.05, 0.01); pf(qt(1 - pp/2, df=20)^2, df1=1, df2=20, low=F)
```

この性質が、実は、上述の例2に隠されていることを示そう。

```
c(wings.onewaytest$statistic, wings.ttest$statistic^2)
qf(wings.onewaytest$p.value, df1=1, df2=wings.onewaytest$parameter[2], low=F)
qt(wings.onewaytest$p.value/2, df=wings.ttest$parameter)^2
```

ここでは、Welch の検定を行っているので、自由度が自然数ではないことに注意せよ。
以下のように、等分散を仮定した場合について、上記を確認してみよ。

```
(oneway.test(size ~ location, data=wings, var.eq=T) -> wings.onewaytest2)
(t.test(wng$A, wng$B, var.eq=T) -> wings.ttest2)
```

- 平均に違いが認められる時、どの平均に違いが見られるか？それが問題になる。

例3：触媒（4種類）に対する化合物（液体）の濃度について、触媒による違いを検討する（データ取得は、<http://astro.temple.edu/~rmh/HH/> による）。

```
read.table("MacOSX3:Users:tk:rm171:hh:datasets:catalystm.dat",
col.names=c("catalyst", type="concent")) -> catalystm; summary(catalystm)
catalystm$catalyst <- as.factor(catalystm$ca); summary(catalystm) #!
boxplot(concent ~ catalyst, catalystm)

summary(aov(concent ~ catalyst, catalystm) -> catalystm.aov)
model.tables(catalystm.aov, "mean") # cf. tapply(catalystm$con, catalystm$cat, mean)

pairwise.t.test(catalystm$con, catalystm$cat) # p.adj="holm" # DEFAULT
pairwise.t.test(catalystm$con, catalystm$cat, p.adj="bonferroni")

# cf. summary.lm (catalystm.aov)
```

(発展) 多重比較について、さまざまな議論があり、ここでは詳細には立ち入らない。なお、R では、TukeyHSD およびパッケージ `multcomp` があり、例えば、Crawley (2002) pp.274-279. あるいは、Heiberger and Holland (2004) pp.155-185 で詳しく検討されている。
(多重比較法の考え方については、永田 靖・吉田道弘 (1997) を参照せよ)

例4：殺虫剤の効用について（標準的な R に内蔵）、上記などを検討せよ。

```
?InsectSprays # data(InsectSprays) for R ver. < 2
boxplot(count ~ spray, data = InsectSprays)
```

- 2 標本問題のようでいて、実は、1 標本問題！ という話に気をつけよ。
この場合は、対応のある場合の 2 標本問題とよばれる。

例 5 : 女性11名のエネルギー摂取量 (kJ) (月経前と月経後)

```
library(ISwR)
data(intake)
boxplot(intake);          t.test(intake$pre, intake$post)          # WRONG!
plot(intake); abline(0,1); t.test(intake$pre, intake$post, paired=T) # CORRECT
```

■ ノンパラメトリック手法

例 2 (続き) 2 標本問題

1) 順位和検定

```
wilcox.test(wng$A, wng$B) # cf. t.test(wng$A, wng$B)
```

2) コルモゴロフ・スミルノフ検定

```
ks.test(wng$A, wng$B)
```

3) メディアン検定

```
(wng.med <- median(c(wng$A, wng$B)))
data.frame(lapply(wng, length)) -> tab.wng0
data.frame(lapply(wng, function(x) sum(x > wng.med))) -> tab.wng1
rbind(tab.wng1, tab.wng0 - tab.wng1) -> tab.wng; row.names(tab.wng) <- c("more", "less")
```

```
> addmargins(as.matrix(tab.wng))
```

```
  A  B Sum
more 18 42 60
less 32 28 60
Sum  50 70 120
```

```
fisher.test(tab.wng) # or easily, chisq.test(tab.wng)
```

例 5 (続き) 対応のある場合の 2 標本問題

1) (符号つき) 順位和検定

```
wilcox.test(intake$pre, intake$post, paired=T) # t.test(intake$pre, intake$post, paired=T)
```

2) 符号検定

```
binom.test(sum(intake$post > intake$pre), nrow(intake))
```

例 1 と例 3 (続き) 多標本問題 ~ クルスカル・ウォリス検定

```
kruskal.test(Growth ~ Photoperiod, data=grow)
kruskal.test(concent ~ catalyst, catalystm)
```

◆ 参考文献 :

- 竹内 啓・大橋靖雄 (1981): 統計的推測 - 2 標本問題, 数学セミナー増刊, 日本評論社, 192p.
- 東京大学教養学部統計学教室編 (1991): 統計学入門, 東京大学出版会, 307p.
- 永田 靖・吉田道弘 (1997): 統計的多重比較法の基礎, サイエンス社, 187p.
- Crawley, M. J. (2002): Statistical Computing - An Introduction to Data Analysis using S-Plus, Wiley, 761p.
- Crawley, M. J. (2005): Statistics, an introduction using R, Wiley, 327p.
- Dalgaard, P. (2002): Introductory Statistics with R, Springer, 267p.
- Heiberger, R. M. and B. Holland (2004): Statistical Analysis and Data Display, An Intermediate Course with Examples in S-Plus, R, and SAS, Springer, 729p.