`ucr.base.tab` and its `latex` method

Lars Lindhagen
modified/knitrified by Henrik Renlund

August 27, 2014

Contents

1 Generation of a dataset

```
n <- 972 # Data set size.
r <- runif(n)
group <- ifelse(r < 0.15, "Small",
         ifelse(r < 0.65, "Medium", "Large"))
age <- round(runif(n, 20, 80))
hgt <- ifelse(group == "Small", rnorm(n, 160, 5),
       ifelse(group == "Medium", rnorm(n, 170, 5),
       rnorm(n, 180, 5)))
bmi <- round(rnorm(n, 25, 3), digits=2)
# 1 missing in Small group.
ix <- which(group == "Small")
ix.na <- sample(ix, size = min(length(ix), 1), replace=F)
bmi[ix.na] <- NA
# 2 missing in Medium group.
ix <- which(group == "Medium")
ix.na <- sample(ix, size = min(length(ix), 2), replace=F)
bmi[ix.na] <- NA
# 3 missing in Large group.
ix <- which(group == "Large")
ix.na <- sample(ix, size = min(length(ix), 3), replace=F)
bmi[ix.na] <- NA

r <- runif(n)
gender <- ifelse(group == "Small",  ifelse(r < 0.8, "Woman", "Man"),
          ifelse(group == "Medium", ifelse(r < 0.5, "Woman", "Man"),
                                    ifelse(r < 0.2, "Woman", "Man")))
r <- runif(n)
country <- ifelse(group == "Small",
                  ifelse(r < 0.10, "Sweden",
```

```
                  ifelse(r < 0.25, "Germany",
                  ifelse(r < 0.75, "Spain",
                  ifelse(r < 0.80, "Australia",
                  "Japan")))),
           ifelse(group == "Medium",
                  ifelse(r < 0.20, "Sweden",
                  ifelse(r < 0.40, "Germany",
                  ifelse(r < 0.60, "Spain",
                  ifelse(r < 0.80, "Australia",
                  "Japan")))),
          # Large:
                  ifelse(r < 0.30, "Sweden",
                  ifelse(r < 0.50, "Germany",
                  ifelse(r < 0.65, "Spain",
                  ifelse(r < 0.80, "Australia",
                  "Japan"))))))
# 5 missing in Small group.
ix <- which(group == "Small")
ix.na <- sample(ix, size = min(length(ix), 5), replace=F)
country[ix.na] <- NA
# 2 missing in Medium group.
ix <- which(group == "Medium")
ix.na <- sample(ix, size = min(length(ix), 2), replace=F)
country[ix.na] <- NA
## Add unused level 'Brazil'
# country <- reFactor(factor(country), list(Australia=NULL, Brazil=NULL))
country <- factor(country)
# Correct group order.
group <- reFactor(group, list(Small=NULL, Medium=NULL))
data.set <- data.frame(group, age, hgt, bmi, gender, country)
label(data.set$gender) <- "Gender"
label(data.set$age) <- "Age (years)"
label(data.set$hgt) <- "Height (cm)"
label(data.set$bmi) <- "BMI"
label(data.set$country) <- "Country"
# Use only two groups.
use.2.groups <- TRUE
if (use.2.groups) {
  data.set$group[data.set$group == "Medium"] <- "Large"
  data.set$group <- factor(data.set$group)
}
```

2 An example of `ucr.base.tab`

The following code

```
# Default table.
res <- ucr.base.tab(data=data.set, group.name="group")

## Warning:  Error in test for variable country, skips test.

dummy <- latex(res, file="", where="!h", caption="A table 1", label="tab:1")
```

Table 1: A table 1

Variable	N	Small $N = 156$	Large $N = 816$	Combined $N = 972$	P-value
Age (years)	972	49.0 (37.8 – 64.0)	49.5 (34.0 – 64.0)	49.0 (35.0 – 64.0)	0.71[1]
Height (cm)	972	159.1 (155.9 – 162.3)	174.1 (169.4 – 179.6)	172.4 (166.0 – 178.5)	< 0.001[1]
BMI	966	25.0 (22.9 – 27.0)	24.9 (23.1 – 27.3)	24.9 (23.1 – 27.3)	0.95[1]
Gender: Man	972	37 (23.7%)	524 (64.2%)	561 (57.7%)	< 0.001[2]
Woman		119 (76.3%)	292 (35.8%)	411 (42.3%)	
Country: Australia	965	10 (6.6%)	142 (17.4%)	152 (15.8%)	—
Germany		27 (17.9%)	178 (21.9%)	205 (21.2%)	
Japan		31 (20.5%)	149 (18.3%)	180 (18.7%)	
Spain		68 (45.0%)	152 (18.7%)	220 (22.8%)	
Sweden		15 (9.9%)	193 (23.7%)	208 (21.6%)	

m (a – b) represents median (Q_1 – Q_3).
n (p%) represent frequency (percentage). Percentages computed by group.
Tests used: [1]Wilcoxon test; [2]Fisher's exact test.

generates Table 1.

3 To do

- Show some variations of the function arguments
- Why does the warning message appear (in Sec. 2)?
- Why is it impossible to have `caption.loc="bottom"`?