King Sejong's National Referendum on Tax Reform

coop711

2020-05-18

Data

Reading Data

Original data came from intenet version of Sejong silok, summarized by Oh, Ki-Soo.

〈표 1〉 공법관련 여론조사의 찬성과 반대 분석

(단위 : 명, %)

		대신 · 관찰사 · 도사 등				수령				품관 · 촌민				합계			
		찬성		반대		찬성		반대		찬성		반대		찬성		반대	
		인수	%	인수	%	인수	%	인수	%	인수	%	인수	%	인수	%	인수	%
대신 등		21[*1]	9.8	194[*2]	90.2									21	9.8	194	90.2
3품 이하	현직	259	39.7	393	60.3									259	39.7	393	60.3
	전직	443	79.1	117	20.9									443	79.1	117	20.9
유후사[23]										1,123	94.1	71	5.9	1,123	94.1	71	5.9
경기도						29	85.3	5	14.7	17,076	98.6	236	1.4	17,105	98.6	241	1.4
평안도				1		6	14.6	35	85.4	1,326	4.4	28,474	95.6	1,332	4.5	28,510	95.5
황해도						17	50.0	17	50.0	4,454	22.2	15,601	77.8	4,471	22.3	15,618	77.7
충청도				2		35	57.4	26	42.6	6,982	33.3	14,013	66.7	7,017	33.3	14,041	66.7
강원도						5	33.3	10	66.7	939	12.0	6,888	88.0	944	12.0	6,898	88.0
함길도				1		3	17.6	14	82.4	75	1.0	7,387	99.0	78	1.0	7,402	99.0
경상도						55	77.5	16	22.5	36,262	99.0	377	1.0	36,317	98.9	393	1.1
전라도				2		42	77.8	12	22.2	29,505	99.1	257	0.9	29,547	99.1	271	0.9
		723	50.5	710	49.5	192	58.7	135	41.3	97,742	57.1	73,304	42.9	98,657	57.1	74,149	42.9

출처 : 「세종실록」 12년(1430) 8월 10일

```
sejong_poll <- read.table("../data/sejong_poll.txt", header = TRUE, stringsAsFactors = FALSE)
str(sejong_poll)
```

```
## 'data.frame':    44 obs. of  4 variables:
##  $ counts: int  21 194 259 393 443 117 1123 71 29 5 ...
##  $ vote  : chr  "yes" "no" "yes" "no" ...
##  $ class : chr  "high" "high" "third.current" "third.current" ...
##  $ region: chr  "Seoul" "Seoul" "Seoul" "Seoul" ...
```

```
# pander(sejong_poll)
kable(sejong_poll[4:1])
```

region	class	vote	counts
Seoul	high	yes	21
Seoul	high	no	194
Seoul	third.current	yes	259
Seoul	third.current	no	393
Seoul	third.ex	yes	443
Seoul	third.ex	no	117
yuhu	ordinary	yes	1123
yuhu	ordinary	no	71
gyunggi	chief	yes	29
gyunggi	chief	no	5
gyunggi	ordinary	yes	17076
gyunggi	ordinary	no	236
pyungan	high	no	1
pyungan	chief	yes	6
pyungan	chief	no	35
pyungan	ordinary	yes	1326
pyungan	ordinary	no	28474
hwanghae	chief	yes	17
hwanghae	chief	no	17
hwanghae	ordinary	yes	4454
hwanghae	ordinary	no	15601
chungcheong	high	no	2
chungcheong	chief	yes	35
chungcheong	chief	no	26
chungcheong	ordinary	yes	6982
chungcheong	ordinary	no	14013
kangwon	chief	yes	5
kangwon	chief	no	10
kangwon	ordinary	yes	939
kangwon	ordinary	no	6888
hamgil	high	no	1
hamgil	chief	yes	3
hamgil	chief	no	14
hamgil	ordinary	yes	75
hamgil	ordinary	no	7387
gyungsang	chief	yes	55
gyungsang	chief	no	16
gyungsang	ordinary	yes	36262
gyungsang	ordinary	no	377
jeolla	high	no	2
jeolla	chief	yes	42
jeolla	chief	no	12
jeolla	ordinary	yes	29505
jeolla	ordinary	no	257

Factor conversion

We need vote, class, region as `factor`s. If you leave them as `chr`, it will be coerced to factor when you tabulate it according to alphabetical order, which is not what you want. So, use `factor()` to convert them. First, make a working copy vesion of `sejong_poll`

```
sejong_poll_2 <- sejong_poll
```

```
sejong_poll_2$vote <- factor(sejong_poll_2$vote, levels = c("yes","no"), labels = c("Yes", "No"))
```

You can check that `labels =` is not necessary if same as levels. Continue with class and region_

```
class_levels <- c("high","third.current", "third.ex", "chief", "ordinary")
class_labels <- c("High","3rd_current", "3rd_former", "Chief", "Commons")
sejong_poll_2$class <- factor(sejong_poll_2$class, levels = class_levels, labels = class_labels)
```

```
region_levels <- c("Seoul","yuhu", "gyunggi", "pyungan", "hwanghae", "chungcheong", "kangwon", "hamgil"
# region_labels <- c("Seoul","Yuhu", "Gyunggi", "Pyungan", "Hwanghae", "Chungcheong", "Kangwon", "Hamgi
region_labels <- c("SL","YH", "GG", "PA", "HH", "CC", "KW", "HG", "GS", "JL")
sejong_poll_2$region <- factor(sejong_poll_2$region, levels = region_levels, labels = region_labels)
```

```
str(sejong_poll_2)
```

```
## 'data.frame':    44 obs. of  4 variables:
##  $ counts: int  21 194 259 393 443 117 1123 71 29 5 ...
##  $ vote  : Factor w/ 2 levels "Yes","No": 1 2 1 2 1 2 1 2 1 2 ...
##  $ class : Factor w/ 5 levels "High","3rd_current",..: 1 1 2 2 3 3 5 5 4 4 ...
##  $ region: Factor w/ 10 levels "SL","YH","GG",..: 1 1 1 1 1 1 2 2 3 3 ...
```

```
kable(sejong_poll_2[4:1])
```

region	class	vote	counts
SL	High	Yes	21
SL	High	No	194
SL	3rd_current	Yes	259
SL	3rd_current	No	393
SL	3rd_former	Yes	443
SL	3rd_former	No	117
YH	Commons	Yes	1123
YH	Commons	No	71
GG	Chief	Yes	29
GG	Chief	No	5
GG	Commons	Yes	17076
GG	Commons	No	236
PA	High	No	1
PA	Chief	Yes	6
PA	Chief	No	35
PA	Commons	Yes	1326
PA	Commons	No	28474
HH	Chief	Yes	17
HH	Chief	No	17
HH	Commons	Yes	4454
HH	Commons	No	15601
CC	High	No	2
CC	Chief	Yes	35
CC	Chief	No	26
CC	Commons	Yes	6982
CC	Commons	No	14013
KW	Chief	Yes	5
KW	Chief	No	10
KW	Commons	Yes	939

region	class	vote	counts
KW	Commons	No	6888
HG	High	No	1
HG	Chief	Yes	3
HG	Chief	No	14
HG	Commons	Yes	75
HG	Commons	No	7387
GS	Chief	Yes	55
GS	Chief	No	16
GS	Commons	Yes	36262
GS	Commons	No	377
JL	High	No	2
JL	Chief	Yes	42
JL	Chief	No	12
JL	Commons	Yes	29505
JL	Commons	No	257

Array

We can set up the data as an array

```
sejong_poll_array <- xtabs(counts ~ vote + class + region,
                           data = sejong_poll_2)
str(sejong_poll_array)
```

```
##  'xtabs' int [1:2, 1:5, 1:10] 21 194 259 393 443 117 0 0 0 0 ...
##  - attr(*, "dimnames")=List of 3
##   ..$ vote  : chr [1:2] "Yes" "No"
##   ..$ class : chr [1:5] "High" "3rd_current" "3rd_former" "Chief" ...
##   ..$ region: chr [1:10] "SL" "YH" "GG" "PA" ...
##  - attr(*, "call")= language xtabs(formula = counts ~ vote + class + region, data = sejong_poll_2)
```

```
sejong_poll_array
```

```
## , , region = SL
##
##      class
## vote   High 3rd_current 3rd_former Chief Commons
##   Yes    21         259        443     0       0
##   No    194         393        117     0       0
##
## , , region = YH
##
##      class
## vote   High 3rd_current 3rd_former Chief Commons
##   Yes     0           0          0     0    1123
##   No      0           0          0     0      71
##
## , , region = GG
##
##      class
## vote   High 3rd_current 3rd_former Chief Commons
##   Yes     0           0          0    29   17076
##   No      0           0          0     5     236
##
```

```
## , , region = PA
##
##      class
## vote   High 3rd_current 3rd_former Chief Commons
##   Yes     0           0          0     6    1326
##   No      1           0          0    35   28474
##
## , , region = HH
##
##      class
## vote   High 3rd_current 3rd_former Chief Commons
##   Yes     0           0          0    17    4454
##   No      0           0          0    17   15601
##
## , , region = CC
##
##      class
## vote   High 3rd_current 3rd_former Chief Commons
##   Yes     0           0          0    35    6982
##   No      2           0          0    26   14013
##
## , , region = KW
##
##      class
## vote   High 3rd_current 3rd_former Chief Commons
##   Yes     0           0          0     5     939
##   No      0           0          0    10    6888
##
## , , region = HG
##
##      class
## vote   High 3rd_current 3rd_former Chief Commons
##   Yes     0           0          0     3      75
##   No      1           0          0    14    7387
##
## , , region = GS
##
##      class
## vote   High 3rd_current 3rd_former Chief Commons
##   Yes     0           0          0    55   36262
##   No      0           0          0    16     377
##
## , , region = JL
##
##      class
## vote   High 3rd_current 3rd_former Chief Commons
##   Yes     0           0          0    42   29505
##   No      2           0          0    12     257
```

Votes

Total

Check the total vote with xtabs()

```
vote_total <- xtabs(counts ~ vote,
                    data = sejong_poll_2)
kable(t(as.matrix(vote_total)),
      caption = "Total")
```

Table 3: Total

Yes	No
98657	74149

```
# format(prop.table(vote_total)*100, digits = 3, nsmall = 1)
kable(t(as.matrix(format(prop.table(vote_total) * 100,
                         digits = 3,
                         nsmall = 1))),
      caption = "Percentage",
      align = rep("r", 2))
```

Table 4: Percentage

Yes	No
57.1	42.9

```
vote_total.2 <- apply(sejong_poll_array, 1, sum)
# kable(t(as.matrix(vote_total.2)))
kable(t(as.matrix(vote_total.2)),
      caption = "Total")
```

Table 5: Total

Yes	No
98657	74149

Vote by class

```
vote_class <- xtabs(counts ~ vote + class,
                    data = sejong_poll_2)
kable(vote_class,
      caption = "By Class")
```

Table 6: By Class

	High	3rd_current	3rd_former	Chief	Commons
Yes	21	259	443	192	97742
No	200	393	117	135	73304

```
vote_class_a <- apply(sejong_poll_array, 1:2, sum)
kable(vote_class_a,
      caption = "By Class")
```

Table 7: By Class

	High	3rd_current	3rd_former	Chief	Commons
Yes	21	259	443	192	97742
No	200	393	117	135	73304

Commons vs Bureaucrats

We need to analyse Commons separately.

```
sejong_poll_2$class_2 <- factor(ifelse(sejong_poll_2$class == "Commons",
                                       "Commons", "Bureaus"),
                                levels = c("Bureaus", "Commons"))
kable(sejong_poll_2[c(4, 3, 5, 2, 1)])
```

region	class	class_2	vote	counts
SL	High	Bureaus	Yes	21
SL	High	Bureaus	No	194
SL	3rd_current	Bureaus	Yes	259
SL	3rd_current	Bureaus	No	393
SL	3rd_former	Bureaus	Yes	443
SL	3rd_former	Bureaus	No	117
YH	Commons	Commons	Yes	1123
YH	Commons	Commons	No	71
GG	Chief	Bureaus	Yes	29
GG	Chief	Bureaus	No	5
GG	Commons	Commons	Yes	17076
GG	Commons	Commons	No	236
PA	High	Bureaus	No	1
PA	Chief	Bureaus	Yes	6
PA	Chief	Bureaus	No	35
PA	Commons	Commons	Yes	1326
PA	Commons	Commons	No	28474
HH	Chief	Bureaus	Yes	17
HH	Chief	Bureaus	No	17
HH	Commons	Commons	Yes	4454
HH	Commons	Commons	No	15601
CC	High	Bureaus	No	2
CC	Chief	Bureaus	Yes	35
CC	Chief	Bureaus	No	26
CC	Commons	Commons	Yes	6982
CC	Commons	Commons	No	14013
KW	Chief	Bureaus	Yes	5
KW	Chief	Bureaus	No	10
KW	Commons	Commons	Yes	939
KW	Commons	Commons	No	6888
HG	High	Bureaus	No	1
HG	Chief	Bureaus	Yes	3
HG	Chief	Bureaus	No	14
HG	Commons	Commons	Yes	75
HG	Commons	Commons	No	7387

region	class	class_2	vote	counts
GS	Chief	Bureaus	Yes	55
GS	Chief	Bureaus	No	16
GS	Commons	Commons	Yes	36262
GS	Commons	Commons	No	377
JL	High	Bureaus	No	2
JL	Chief	Bureaus	Yes	42
JL	Chief	Bureaus	No	12
JL	Commons	Commons	Yes	29505
JL	Commons	Commons	No	257

```
str(sejong_poll_2)
```

```
## 'data.frame':    44 obs. of  5 variables:
##  $ counts : int  21 194 259 393 443 117 1123 71 29 5 ...
##  $ vote   : Factor w/ 2 levels "Yes","No": 1 2 1 2 1 2 1 2 1 2 ...
##  $ class  : Factor w/ 5 levels "High","3rd_current",..: 1 1 2 2 3 3 5 5 4 4 ...
##  $ region : Factor w/ 10 levels "SL","YH","GG",..: 1 1 1 1 1 1 2 2 3 3 ...
##  $ class_2: Factor w/ 2 levels "Bureaus","Commons": 1 1 1 1 1 1 2 2 1 1 ...
```

Compare the votes by class_2, (Bureaucrats vs Commons)

```
vote_class_2 <- xtabs(counts ~ vote + class_2,
                      data = sejong_poll_2)
kable(vote_class_2, caption = "By Bureaus and Commons")
```

Table 9: By Bureaus and Commons

	Bureaus	Commons
Yes	915	97742
No	845	73304

```
vote_class_2_a <- cbind("Bureaus" = rowSums(vote_class_a[, -5]), "Commons" =  vote_class_a[, 5])
kable(vote_class_2_a, caption = "By Bureaus and Commons")
```

Table 10: By Bureaus and Commons

	Bureaus	Commons
Yes	915	97742
No	845	73304

Add subtotals to the margins,

```
vote_class_2_am <- addmargins(vote_class_2)
kable(vote_class_2_am)
```

	Bureaus	Commons	Sum
Yes	915	97742	98657
No	845	73304	74149
Sum	1760	171046	172806

Compute the marginal proportions. Note the use of `digits = 3` and `nsmall = 1`.

```
kable(format(prop.table(vote_class_2, margin = 2)*100, digits = 3, nsmall = 1), caption = "Bureaus and
```

Table 12: Bureaus and Commons

	Bureaus	Commons
Yes	52.0	57.1
No	48.0	42.9

Votes by region with respect to class_2

Count the vote by region class_2 wise.

```
class_2 <- sejong_poll_2$class_2
vote_region_bureaus <- xtabs(counts ~ vote + region,
                             data = sejong_poll_2,
                             class_2 == "Bureaus",
                             drop = TRUE)
kable(vote_region_bureaus, caption = "Votes(Bureaus)")
```

Table 13: Votes(Bureaus)

	SL	GG	PA	HH	CC	KW	HG	GS	JL
Yes	723	29	6	17	35	5	3	55	42
No	704	5	36	17	28	10	15	16	14

```
# xtabs(counts ~ vote + region, data = sejong_poll_2[class_2 == "Bureaus", ], drop = TRUE)
vote_region_commons <- xtabs(counts ~ vote + region, data = sejong_poll_2, class_2 == "Commons", drop =
kable(vote_region_commons, caption = "Votes(Commons)")
```

Table 14: Votes(Commons)

	YH	GG	PA	HH	CC	KW	HG	GS	JL
Yes	1123	17076	1326	4454	6982	939	75	36262	29505
No	71	236	28474	15601	14013	6888	7387	377	257

Seoul has three times more Bureaucrats than other regions, so analyse further.

```
region <- sejong_poll_2$region
vote_seoul_class <- xtabs(counts ~ vote + class, data = sejong_poll_2, region == "SL", drop = TRUE)
kable(vote_seoul_class, caption = "Seoul")
```

Table 15: Seoul

	High	3rd_current	3rd_former
Yes	21	259	443
No	194	393	117

```
kable(format(prop.table(vote_seoul_class, margin = 2)*100, digits  = 3, nsmall = 1), caption = "SL", ali
```

Table 16: SL

	High	3rd_current	3rd_former
Yes	9.77	39.72	79.11
No	90.23	60.28	20.89

Chungcheong's case.

```
vote_chung_class <- xtabs(counts ~ vote + class, data = sejong_poll_2, region == "CC", drop = TRUE)
kable(format(prop.table(vote_chung_class, margin = 2)*100, digits = 3, nsmall = 1), caption = "CC", alig
```

Table 17: CC

	High	Chief	Commons
Yes	0.0	57.4	33.3
No	100.0	42.6	66.7

- Save the working directory image.

```
save.image(file = "sejong_poll_data.RData")
```