

Hypothesis Tests

Landi Luo

12/7/2018

Load packages:

```
if (!require("pacman"))
  install.packages("pacman", repos = "http://cran.us.r-project.org/")
p_load("tidyverse", "survival", "kableExtra")
```

Import data:

```
breast <- readRDS(file = "breast_final.rds")

# delete all survival time = 0
breast <- subset(breast, SRV_TIME_MON != 0)

# convert SEX to dummy: 0=male, 1=female
breast$SEX <- ifelse(breast$SEX == 1, 0, 1)
```

Log Rank test (from Joowon's code)

```
log.rank = function(time, event, data, list){ # list : number of selected column
  surv = Surv(time, event, type = "right")
  result = data.frame("Log-rank_p" = NA, "Gehan_p" = NA)

  for (i in 1:length(list)){
    cova = data[,list[i]]
    surv.test = survdiff(surv ~ cova, rho = 0)
    surv.test2 = survdiff(surv ~ cova, rho = 1)
    result = rbind(result, round(c(1-pchisq(as.numeric(surv.test[5])),length(unique(cova))-1), 1-pchisq(
  })

  result = result[-1,]
  result = data.frame("variable" = colnames(data)[list], result)
  rownames(result) = 1:nrow(result)
  result
}

breast = data.frame(breast)
logrank_table <- log.rank(time = breast$SRV_TIME_MON, event = breast$delta, data = breast, list = c(1,3,4))

kable(logrank_table, "latex", booktabs = TRUE) %>%
  kable_styling(latex_options = "striped")
```

variable	Log.rank_p	Gehan_p
stage	0.0000	0.0000
RAC_REC_Y	0.0000	0.0000
SEX	0.2414	0.2807
BRST_SUB	0.0000	0.0000
PRIMSITE	0.0000	0.0000
ERSTATUS	0.0000	0.0000
PRSTATUS	0.0000	0.0000
INSREC_PUB	0.0000	0.0000

```
breast_final2 = breast
breast_final2$Agemean = as.integer(breast$Age>mean(breast$Age))
breast_final2$AgeDXmean = as.integer(breast$Age>mean(breast$AGE_DX))

logrank_table2 <- log.rank(time = breast_final2$SRV_TIME_MON, event = breast_final2$delta, data = breast_final2)

kable(logrank_table2, "latex", booktabs = TRUE) %>%
  kable_styling(latex_options = "striped")
```

variable	Log.rank_p	Gehan_p
stage	0.0000	0.0000
RAC_REC_Y	0.0000	0.0000
SEX	0.2414	0.2807
BRST_SUB	0.0000	0.0000
PRIMSITE	0.0000	0.0000
ERSTATUS	0.0000	0.0000
PRSTATUS	0.0000	0.0000
INSREC_PUB	0.0000	0.0000
Agemean	0.0000	0.0000
AgeDXmean	0.0001	0.0000

LRT, Score, Wald Tests

```
single.cox = function(time, event, covariate, co.type) {

  p1 = p2 = p3 = c()

  for (i in 1:ncol(covariate)){
    surv.sub = Surv(time=time, event=event, type="right")
    if (co.type[i]=="N") { # continuous variable
      formul = paste("surv.sub ~ covariate[,", i, "]", sep="")
      fit = coxph(eval(parse(text=formul)), ties = "breslow")
    } else { # categorical variable
      formul = paste("surv.sub ~ factor(covariate[,", i, "])", sep="")
      fit = coxph(eval(parse(text=formul)), ties = "breslow")
    }
    p1 = rbind(p1, t(round(summary(fit)$waldtest, 3))) # Wald Test
    p2 = rbind(p2, t(round(summary(fit)$sctest, 3))) # Score Test
    p3 = rbind(p3, t(round(summary(fit)$logtest, 3))) # LRT
  }
}
```

```

}

result = list()
result[[1]] = p1 ; result[[2]] = p2; result[[3]] = p3
return(result)

}

##-----

# Example
co.type = c(rep("C", 4), rep("N", 2), "C", rep("N", 2), rep("C", 4)) # "C" means categorical variable /
X = breast[,c("SEX", "stage", "RAC_RECY", "BRST_SUB", "MAR_STAT", "MALIGCOUNT", "PRMSITE",
              "ERSTATUS", "PRSTATUS", "INSREC_PUB")] # covariates
result = single.cox(time = breast$SRV_TIME_MON, event = breast$delta, covariate = X, co.type)
names(result) = c("Wald Test", "Score Test", "LRT")
rownames(result[[1]]) = colnames(X); rownames(result[[2]]) = colnames(X); rownames(result[[3]]) = colnames(X)
wald <- data.frame(result$`Wald Test`)
score <- data.frame(result$`Score Test`)
lrt <- data.frame(result$LRT)

kable(wald, "latex", booktabs = TRUE) %>%
  kable_styling(latex_options = "striped")

```

	test	df	pvalue
SEX	1.36	1	0.243
stage	5861.42	4	0.000
RAC_RECY	143.10	3	0.000
BRST_SUB	1049.04	3	0.000
MAR_STAT	77.28	1	0.000
MALIGCOUNT	41.38	1	0.000
PRMSITE	565.35	8	0.000
ERSTATUS	1013.79	1	0.000
PRSTATUS	1014.57	1	0.000
INSREC_PUB	508.75	3	0.000

```

kable(score, "latex", booktabs = TRUE) %>%
  kable_styling(latex_options = "striped")

```

	test	df	pvalue
SEX	1.370	1	0.242
stage	13294.571	4	0.000
RAC_RECY	149.035	3	0.000
BRST_SUB	1193.626	3	0.000
MAR_STAT	77.816	1	0.000
MALIGCOUNT	41.743	1	0.000
PRMSITE	611.374	8	0.000
ERSTATUS	1138.773	1	0.000
PRSTATUS	1133.384	1	0.000
INSREC_PUB	544.119	3	0.000

```
kable(lrt, "latex", booktabs = TRUE) %>%
  kable_styling(latex_options = "striped")
```

	test	df	pvalue
SEX	1.260	1	0.262
stage	5794.524	4	0.000
RAC_REC_Y	128.648	3	0.000
BRST_SUB	895.337	3	0.000
MAR_STAT	73.849	1	0.000
MALIGCOUNT	36.586	1	0.000
PRIMSITE	495.260	8	0.000
ERSTATUS	885.310	1	0.000
PRSTATUS	985.679	1	0.000
INSREC_PUB	452.691	3	0.000