Survival analysis in R:

Packages needed:

* survival: include analysis function
* survminer: better visualization for Kaplain-Meier curve, include function ggsurv()

A picture containing graphical user interface

Description automatically generated

The standard format of data putting into the function.

Main function:

* surv(): to get survival object

e.g. Surv(duration, status) ~ group, data=data)

* survfit(): to fit survival curves

e.g. survfit(Surv(duration, status) ~ factor(group), data=data)

Note: to visualize the Kaplan meier curve it can use either:

* autoplot(survfit(…)) + setting for titles,axis …
* ggsurvplot(survfit(…))

The second can set pval = TRUE to get pvalue of log-rank test if two curves exist

* survdiff() : to get the logranktest result

e.g. survdiff(Surv(duration, status) ~ group), data=data)

Note: Logranktest is for testing the survival time difference between two group, hence the column group must be binary classification

* coxph(): to fit Cox proportional hazards regression model

e.g. coxph(Surv(duration, status) ~ group, data=data)

Note: the pvalue from Survdiff() can also derived from Coxph()

survdiff(Surv(duration, status) ~ group, data=data)$pvalue is same as

coxfit = coxph(Surv(duration, status) ~ group, data=data);

summary(coxfit)$sctest[3]

Ref:

<https://shariq-mohammed.github.io/files/cbsa2019/1-intro-to-survival.html>

<https://www.biostars.org/p/401657/#401759>