

PS_survivaldata

Contents

Libraries	1
Data Analysis	2
Cox Regression	2
Kaplan-Meier Survival Estimate	3

Libraries

```
library(survival)
library(ggpubr)

## Loading required package: ggplot2
## Loading required package: magrittr

library(survminer)
library(magrittr)
library(ggplot2)
library(tidyr)

##
## Attaching package: 'tidyr'
## The following object is masked from 'package:magrittr':
##
##     extract

library(dplyr) #data parsing library

##
## Attaching package: 'dplyr'
## The following objects are masked from 'package:stats':
##
##     filter, lag
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union

library(data.table) #data parsing library mainly for reading and writing out

##
## Attaching package: 'data.table'
## The following objects are masked from 'package:dplyr':
##
##     between, first, last

library(DiagrammeR) #figure aesthetics
library(MASS) #Functions and datasets to support
```

```
##
## Attaching package: 'MASS'

## The following object is masked from 'package:dplyr':
##
##      select

library(BBmisc) #Miscellaneous functions

##
## Attaching package: 'BBmisc'

## The following objects are masked from 'package:dplyr':
##
##      coalesce, collapse

## The following object is masked from 'package:base':
##
##      isFALSE

library(gridExtra) #figure aesthetics

##
## Attaching package: 'gridExtra'

## The following object is masked from 'package:dplyr':
##
##      combine

library(coxme)

## Loading required package: bdsmatrix

##
## Attaching package: 'bdsmatrix'

## The following object is masked from 'package:base':
##
##      backsolve
```

Data Analysis

Cox Regression

```
#Obtain data
dat<- read.csv("/Users/Pikachu/Desktop/Proteome_stability_project/Data/2018-07-01_HS_datasheet.csv")

###glimpse(dat)
###names(dat)

#Assign censored data and create new column
dat$status[dat$surv_days != "NA"] <- 1

#Mixed Effect model - strain as random effect
cox <- coxme(Surv(surv_days, status) ~ sex*treatment + (1|strain) , data = dat)
summary(cox)
```

```
## Cox mixed-effects model fit by maximum likelihood
## Data: dat
## events, n = 606, 606 (174 observations deleted due to missingness)
## Iterations= 7 38
##          NULL Integrated      Fitted
## Log-likelihood -3280.692 -3253.965 -3242.758
##
##          Chisq    df          p    AIC    BIC
## Integrated loglik 53.45 10.00 6.1330e-08 33.45 -10.62
## Penalized loglik 75.87 17.23 2.4698e-09 41.41 -34.53
##
## Model: Surv(surv_days, status) ~ sex * treatment + (1 | strain)
## Fixed coefficients
##          coef exp(coef) se(coef)      z      p
## sexM          -0.7672819 0.4642733 0.1552486 -4.94 7.7e-07
## treatmentFR    -0.1448860 0.8651209 0.1619925 -0.89 3.7e-01
## treatmentHH    -0.1334239 0.8750941 0.1654695 -0.81 4.2e-01
## treatmentSHS   -0.3534451 0.7022646 0.1708281 -2.07 3.9e-02
## treatmentSR    -0.3610010 0.6969783 0.1719747 -2.10 3.6e-02
## sexM:treatmentFR 0.1930468 1.2129396 0.2462295 0.78 4.3e-01
## sexM:treatmentHH 0.3037060 1.3548706 0.2478645 1.23 2.2e-01
## sexM:treatmentSHS 0.3865703 1.4719239 0.2504440 1.54 1.2e-01
## sexM:treatmentSR 0.9391202 2.5577300 0.2542901 3.69 2.2e-04
##
## Random effects
## Group Variable Std Dev Variance
## strain Intercept 0.22079763 0.04875159

#non-mixed effect model -- for figure
#coxph <- coxph(Surv(surv_days, status) ~ sex*treatment, data = dat)
#summary(coxph)

#cox_fit <- survfit(coxph)
#ggsurvplot(cox_fit, data = dat, risk.table = TRUE, pval = TRUE)
```

Kaplan-Meier Survival Estimate

```
#filter out treatments
HHdat <- dat%>%
  filter(treatment == "HH")

SHSdat <- dat%>%
  filter(treatment == "SHS")

FRdat <- dat%>%
  filter(treatment == "FR")

SRdat <- dat%>%
  filter(treatment == "SR")

noCCdat <- dat%>%
  filter(treatment != "CC")
```

```

#compute kaplan-Meier survival estimate-so compute the survival probability by:
#fit1 computes survival probability by sex
fit1 <- survfit(Surv(surv_days, status) ~ sex, data = dat)
print(fit1)

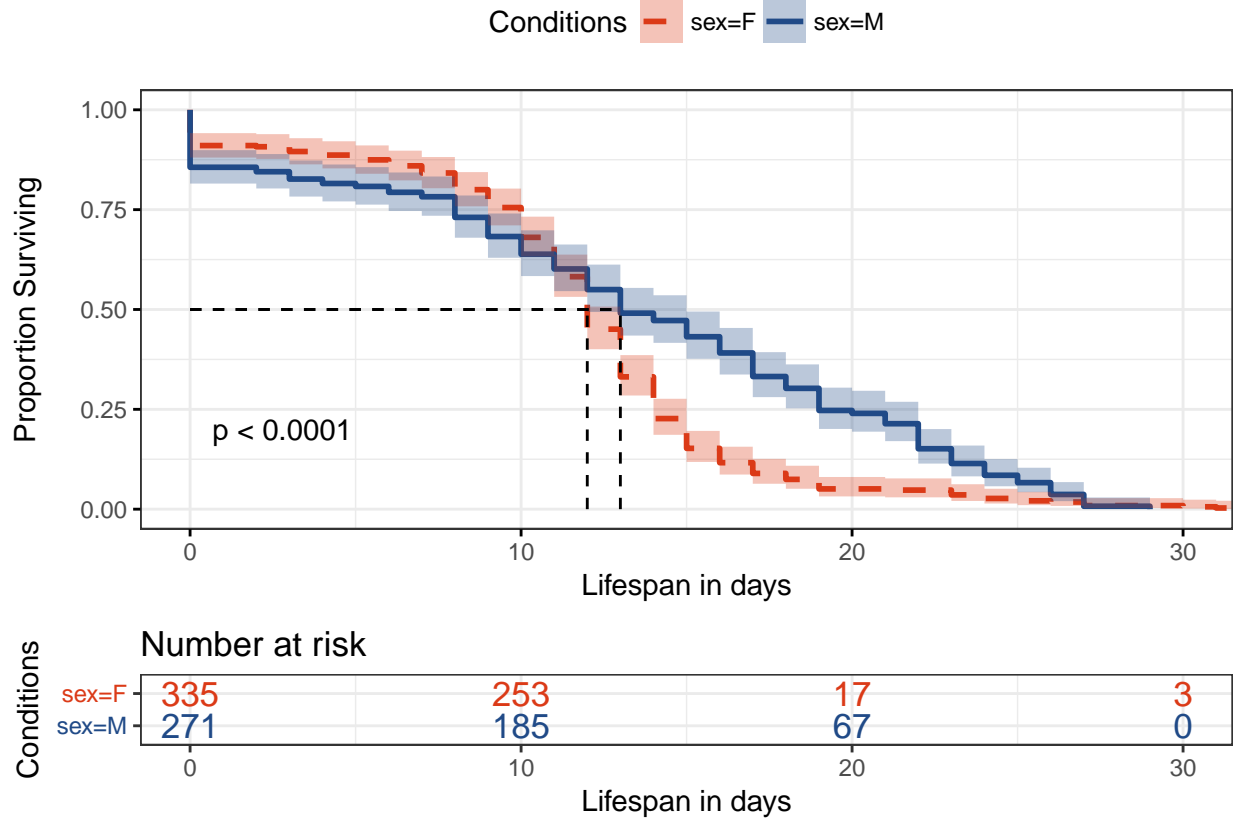
## Call: survfit(formula = Surv(surv_days, status) ~ sex, data = dat)
##
##      174 observations deleted due to missingness
##           n events median 0.95LCL 0.95UCL
## sex=F 335      335      12      12      13
## sex=M 271      271      13      12      15

# Access to the sort summary table
summary(fit1)$table

##           records n.max n.start events   *rmean *se(rmean) median 0.95LCL
## sex=F         335    335      335    335 11.67910  0.3071111      12      12
## sex=M         271    271      271    271 13.37638  0.4950375      13      12
##           0.95UCL
## sex=F           13
## sex=M           15

# Visualize curves (Change color, linetype by strata, risk.table color by strata)
ggsurvplot(fit1,
  pval = TRUE, conf.int = TRUE,
  risk.table = TRUE, # Add risk table
  risk.table.col = "strata", # Change risk table color by groups
  linetype = c("dashed","solid"),
  surv.median.line = "hv", # Specify median survival
  ggtheme = theme_bw(), # Change ggplot2 theme
  palette = c("#DB3A17","#204A87"),
  xlab = "Lifespan in days",
  ylab = "Proportion Surviving",
  legend.title = "Conditions",
  pval.size=4)

```



#fit2 computes survival probability by treatment

```
fit2 <- survfit(Surv(surv_days, status) ~ treatment, data = noCCdat)
print(fit2)
```

```
## Call: survfit(formula = Surv(surv_days, status) ~ treatment, data = noCCdat)
##
## 95 observations deleted due to missingness
##      n events median 0.95LCL 0.95UCL
## treatment=FR  108   108    13     12     14
## treatment=HH  108   108    12     11     14
## treatment=SHS 105   105    13     12     14
## treatment=SR  104   104    11     10     12
```

Access to the sort summary table
`summary(fit2)$table`

```
##      records n.max n.start events  *rmean *se(rmean) median
## treatment=FR   108  108   108   108 12.87963  0.6493002    13
## treatment=HH   108  108   108   108 11.78704  0.7231773    12
## treatment=SHS  105  105   105   105 13.04286  0.7157482    13
## treatment=SR   104  104   104   104 10.51923  0.7708793    11
##      0.95LCL 0.95UCL
## treatment=FR    12    14
## treatment=HH    11    14
## treatment=SHS    12    14
## treatment=SR     10    12
```

```
# Visualize curves (Change color, linetype by strata, risk.table color by strata)
ggsurvplot(fit2,
  pval = TRUE, conf.int = F,
  risk.table = TRUE, # Add risk table
  risk.table.col = "strata", # Change risk table color by groups
  linetype = c("solid", "dashed", "dashed", "solid"),
  surv.median.line = "hv", # Specify median survival
  ggtheme = theme_bw(), # Change ggplot2 theme
  palette = c("#DB3A17", "#204A87", "#267E13", "#341078"),
  xlab = "Lifespan in days",
  ylab = "Proportion Surviving",
  legend.title = "Conditions",
  pval.size=4)
```

