DATA 599 - Survival Analysis HW6 - Charles Hanks

Due: March 3, 2023

** CHANGE YOUR NAME IN THE TITLE AND COMPLETE THE FOLLOWING: **

DATASET: NIST-arrALT.txt

Source: NIST Engineering Handbook

Component life tests were run at three temperatures: 85°C, 105°C and 125°C. The lowest temperature cell was populated with 100 components; the 105°C cell had 50 components and the highest stress cell had 25 components. All tests were run until either all the units in the cell had failed or 1000 hours was reached. Acceleration was assumed to follow an Arrhenius model and the life distribution model for the failure mode was believed to be Weibull. The normal operating temperature for the components is 25°C. Failure analysis confirmed that all failures were due to the same failure mechanism .

Please read in the dataset, name it `ds`, and do the following:

A

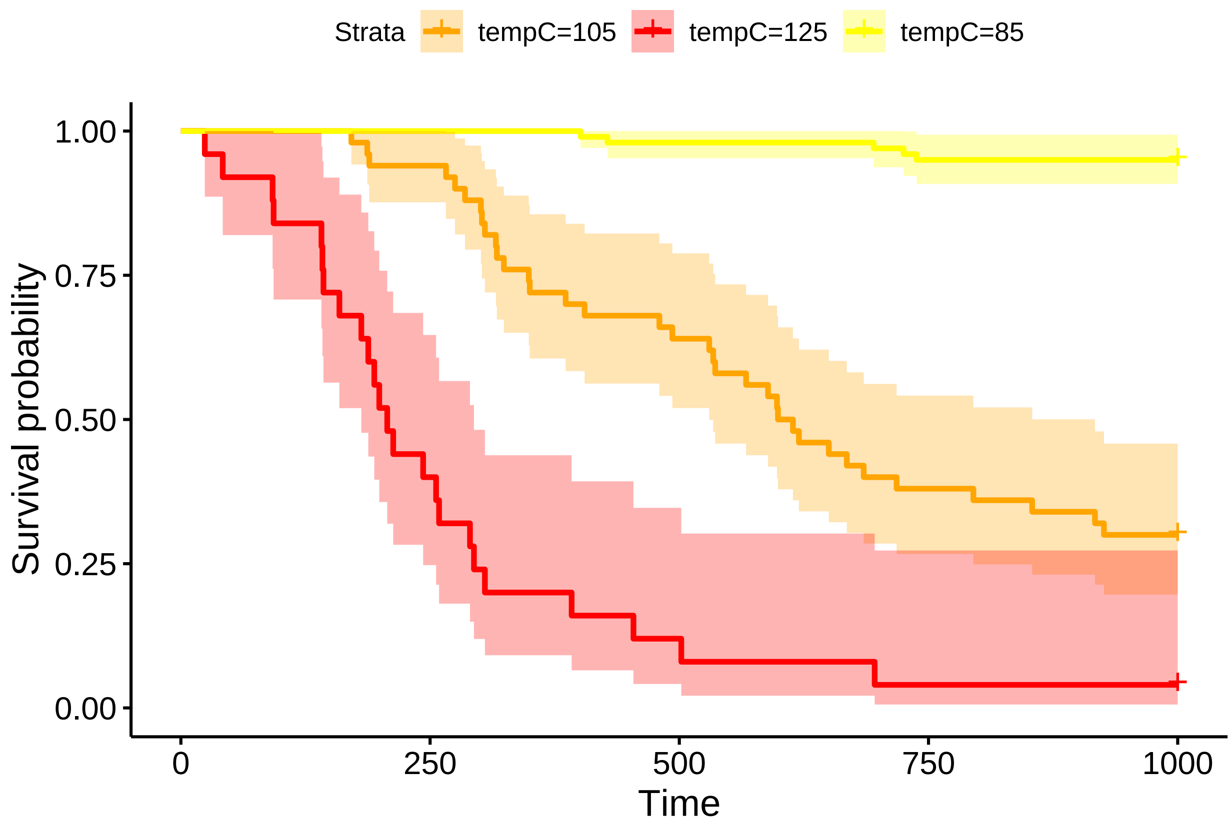
Plot KM survival curves for each temperature level.

```
ds = read_csv('NIST-arrALT.txt')
```

```
## Rows: 175 Columns: 3
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## dbl (3): failTime, status, tempC
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
km = survfit(Surv(ds$failTime,ds$status) ~ ds$tempC)

ggsurvplot(fit = km, data = ds, palette = c("orange", "red", "yellow"), risk.table = F, conf.int = T)
```



B

For each temperature level, fit a Weibull distribution to the failure time data.

```
ds = ds %>% mutate(arrTemp = (11605/(tempC + 273.15)))

le_modele = survreg(Surv(ds$failTime, ds$status) ~ ds$arrTemp, dist = "weibull")

b0= le_modele$coefficients[1]
b1 = le_modele$coefficients[2]
sigma = le_modele$scale

t85 = rweibull(1000, shape = 1/sigma, scale=exp(b0+b1*11605/(85+273.15)))
t105 = rweibull(1000, shape = 1/sigma, scale=exp(b0+b1*11605/(105+273.15)))
t125 = rweibull(1000, shape = 1/sigma, scale=exp(b0+b1*11605/(125+273.15)))
```

C

What are the β values from your temperature-level distribution fits?

D

Based on your answer in C, is it reasonable to assume that the general nature of the failures (i.e. infant mortality, exponential, or wearout) is the same for all temperature levels? Why or why not?

ANSWER: [fill in]

E

Create a column in ds called `dataType` that is equal to "observed"

F

Create a column in `ds` called `arrTemp` that is generated via the following formula:

arrTemp $= \frac{11605}{\text{tempC}+273.15}$

Then print the first 10 lines of the dataset.

G

Fit a Weibull regression model of the form Failure Time $\sim$ Weibull($\mu = b_0 + b_1 \text{arrTemp}, \sigma$).

H

What is the activation energy E_a?

ANSWER: [fill in]

I

Simulate 10000 failure times for each temperature level.

J

Using your data in part I, create a data table which has the following variables: failure time, status (1=exact failure time, 0=right censored), temperature, and dataType(="model"). Name this table `modelds`.

K

Combine `ds` and `modelds` into a single dataset using `bind_rows`. Name the resulting dataset `newds`.

L

Create KM tables for temperature, stratifying by the `dataType` variable.

M

Using your KM object in part L, create KM curves with the model fit overlaying the observational data. Make sure you include `linetype="strata"` to make the model fit a different line type than the observational KM curves.

N

Based on your plot, is your model a good match for the data? Why or why not?]

ANSWER: [fill in]

O

Assuming the use temperature condition is 25C, what is the acceleration factor if you run a reliability demonstration test at 65C?

ANSWER: [fill in]

P

Suppose you run an experiment at a stress level that yields an acceleration factor of 20. The failure time data from your experiment is well approximated by a Lognormal(4,0.25). The time unit of your experiment was days. What is the probability of a part failing within 2 years (730 days) at use conditions?

ANSWER: [fill in]