

P8108_final_project_p3

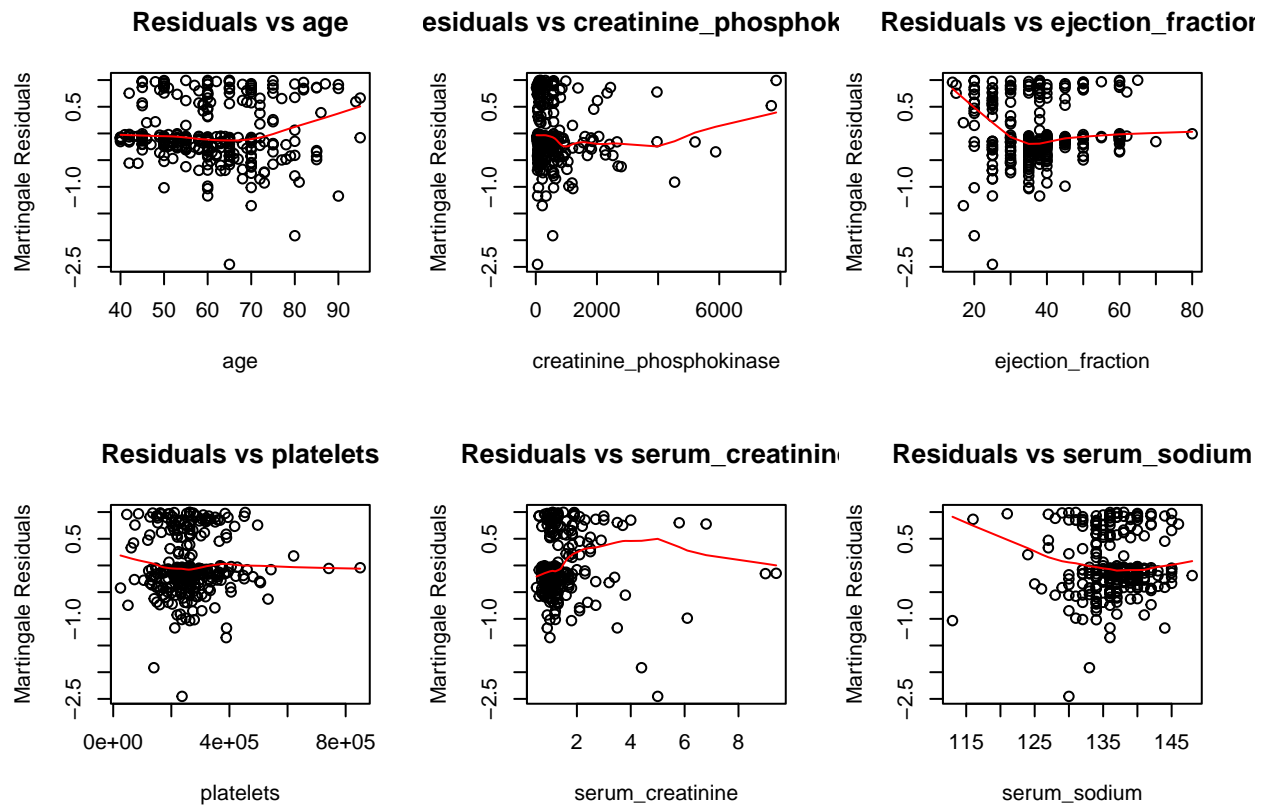
Zihan Wu

2024-12-04

```
# initial cox model
cox_full <- coxph(Surv(time, DEATH_EVENT) ~ age + anaemia + creatinine_phosphokinase + diabetes + eject

# check non-linear relationships
par(mfrow=c(2,3))
continuous_vars <- c("age", "creatinine_phosphokinase", "ejection_fraction",
                     "platelets", "serum_creatinine", "serum_sodium")

for(var in continuous_vars) {
  mart_res <- residuals(cox_full, type="martingale")
  plot(heart_failure_data[[var]], mart_res,
       xlab=var, ylab="Martingale Residuals",
       main=paste("Residuals vs", var))
  lines(lowess(heart_failure_data[[var]], mart_res), col="red")
}
```



Based on the Martingale residual plots analysis, we found non-linear relationships with the outcome in four continuous variables: creatinine_phosphokinase, and serum_creatinine, suggesting these variables need transformations for the Cox model.

Cox

```
# transform non-linear
heart_failure_data <- heart_failure_data %>%
  mutate(
    log_cp = log(creatinine_phosphokinase),
    log_serum_creatinine = log(serum_creatinine)
  )
# initial model with transformed and linear variables
cox_full <- coxph(
  Surv(time, DEATH_EVENT) ~ age + log_cp + ejection_fraction +
    platelets + log_serum_creatinine + serum_sodium +
    anaemia + diabetes + high_blood_pressure + sex + smoking,
  data = heart_failure_data
)
#backward stepwise
cox_backward <- step(cox_full, direction = "backward")

## Start: AIC=957.56
## Surv(time, DEATH_EVENT) ~ age + log_cp + ejection_fraction +
```

```

##      platelets + log_serum_creatinine + serum_sodium + anaemia +
##      diabetes + high_blood_pressure + sex + smoking
##
##              Df      AIC
## - platelets      1 955.71
## - diabetes      1 955.97
## - smoking      1 956.23
## - log_cp      1 956.39
## - sex      1 956.65
## - serum_sodium      1 957.46
## <none>      957.56
## - anaemia      1 960.34
## - high_blood_pressure      1 960.90
## - log_serum_creatinine      1 973.92
## - ejection_fraction      1 974.75
## - age      1 975.99
##
## Step:  AIC=955.71
## Surv(time, DEATH_EVENT) ~ age + log_cp + ejection_fraction +
##      log_serum_creatinine + serum_sodium + anaemia + diabetes +
##      high_blood_pressure + sex + smoking
##
##              Df      AIC
## - diabetes      1 954.09
## - smoking      1 954.30
## - log_cp      1 954.53
## - sex      1 954.68
## - serum_sodium      1 955.67
## <none>      955.71
## - anaemia      1 958.42
## - high_blood_pressure      1 959.00
## - log_serum_creatinine      1 972.35
## - ejection_fraction      1 972.85
## - age      1 974.03
##
## Step:  AIC=954.09
## Surv(time, DEATH_EVENT) ~ age + log_cp + ejection_fraction +
##      log_serum_creatinine + serum_sodium + anaemia + high_blood_pressure +
##      sex + smoking
##
##              Df      AIC
## - smoking      1 952.62
## - log_cp      1 952.95
## - sex      1 953.15
## <none>      954.09
## - serum_sodium      1 954.31
## - anaemia      1 956.89
## - high_blood_pressure      1 957.40
## - log_serum_creatinine      1 970.37
## - ejection_fraction      1 971.51
## - age      1 972.09
##
## Step:  AIC=952.62
## Surv(time, DEATH_EVENT) ~ age + log_cp + ejection_fraction +

```

```

##      log_serum_creatinine + serum_sodium + anaemia + high_blood_pressure +
##      sex
##
##              Df      AIC
## - sex              1 951.27
## - log_cp            1 951.38
## <none>              952.62
## - serum_sodium      1 952.72
## - anaemia           1 955.14
## - high_blood_pressure 1 956.07
## - log_serum_creatinine 1 968.50
## - age               1 970.18
## - ejection_fraction 1 970.21
##
## Step: AIC=951.27
## Surv(time, DEATH_EVENT) ~ age + log_cp + ejection_fraction +
##      log_serum_creatinine + serum_sodium + anaemia + high_blood_pressure
##
##              Df      AIC
## - log_cp            1 949.91
## <none>              951.27
## - serum_sodium      1 951.32
## - anaemia           1 953.68
## - high_blood_pressure 1 955.26
## - log_serum_creatinine 1 966.88
## - ejection_fraction 1 968.21
## - age               1 968.34
##
## Step: AIC=949.91
## Surv(time, DEATH_EVENT) ~ age + ejection_fraction + log_serum_creatinine +
##      serum_sodium + anaemia + high_blood_pressure
##
##              Df      AIC
## - serum_sodium      1 949.88
## <none>              949.91
## - anaemia           1 952.07
## - high_blood_pressure 1 953.76
## - log_serum_creatinine 1 964.92
## - ejection_fraction 1 966.91
## - age               1 967.46
##
## Step: AIC=949.88
## Surv(time, DEATH_EVENT) ~ age + ejection_fraction + log_serum_creatinine +
##      anaemia + high_blood_pressure
##
##              Df      AIC
## <none>              949.88
## - anaemia           1 951.51
## - high_blood_pressure 1 953.46
## - age               1 966.71
## - log_serum_creatinine 1 970.38
## - ejection_fraction 1 970.72

```

```
#forward stepwise
cox_null <- coxph(Surv(time, DEATH_EVENT) ~ 1, data = heart_failure_data)
cox_forward <- step(cox_null,
  scope = ~ age + log_cp + ejection_fraction +
    platelets + log_serum_creatinine + serum_sodium +
    anaemia + diabetes + high_blood_pressure + sex + smoking,
  direction = "forward")
```

```
## Start: AIC=1018.41
## Surv(time, DEATH_EVENT) ~ 1
##
##           Df      AIC
## + log_serum_creatinine 1  989.17
## + age                  1  996.90
## + ejection_fraction    1  999.86
## + serum_sodium         1 1010.18
## + high_blood_pressure  1 1016.22
## + anaemia              1 1017.73
## <none>                 1018.41
## + platelets            1 1019.86
## + diabetes             1 1020.37
## + sex                  1 1020.41
## + smoking              1 1020.41
## + log_cp               1 1020.41
##
## Step: AIC=989.17
## Surv(time, DEATH_EVENT) ~ log_serum_creatinine
##
##           Df      AIC
## + age          1 973.15
## + ejection_fraction 1 973.89
## + high_blood_pressure 1 986.57
## + serum_sodium      1 987.57
## + anaemia           1 988.57
## <none>              989.17
## + log_cp           1 991.07
## + diabetes         1 991.09
## + smoking          1 991.12
## + sex              1 991.15
## + platelets        1 991.17
##
## Step: AIC=973.15
## Surv(time, DEATH_EVENT) ~ log_serum_creatinine + age
##
##           Df      AIC
## + ejection_fraction 1 955.65
## + high_blood_pressure 1 970.52
## + serum_sodium      1 971.16
## + anaemia           1 972.96
## <none>              973.15
## + diabetes         1 974.20
## + smoking          1 974.83
## + log_cp           1 974.99
```

```

## + platelets          1 975.07
## + sex                1 975.15
##
## Step: AIC=955.65
## Surv(time, DEATH_EVENT) ~ log_serum_creatinine + age + ejection_fraction
##
##               Df    AIC
## + high_blood_pressure 1 951.51
## + anaemia             1 953.46
## <none>                 955.65
## + serum_sodium        1 956.46
## + sex                 1 956.81
## + diabetes            1 956.89
## + log_cp              1 957.49
## + smoking             1 957.64
## + platelets           1 957.65
##
## Step: AIC=951.51
## Surv(time, DEATH_EVENT) ~ log_serum_creatinine + age + ejection_fraction +
##   high_blood_pressure
##
##               Df    AIC
## + anaemia          1 949.88
## <none>              951.51
## + serum_sodium    1 952.07
## + diabetes        1 952.80
## + sex             1 953.04
## + log_cp          1 953.16
## + platelets       1 953.50
## + smoking         1 953.50
##
## Step: AIC=949.88
## Surv(time, DEATH_EVENT) ~ log_serum_creatinine + age + ejection_fraction +
##   high_blood_pressure + anaemia
##
##               Df    AIC
## <none>          949.88
## + serum_sodium 1 949.91
## + diabetes     1 951.20
## + log_cp       1 951.32
## + sex         1 951.35
## + smoking     1 951.82
## + platelets   1 951.85

# Perform stepwise selection using stepAIC
cox_step <- stepAIC(cox_full,
                    direction = "both",
                    trace = TRUE)

## Start: AIC=957.56
## Surv(time, DEATH_EVENT) ~ age + log_cp + ejection_fraction +
##   platelets + log_serum_creatinine + serum_sodium + anaemia +
##   diabetes + high_blood_pressure + sex + smoking
##

```

```

##              Df      AIC
## - platelets      1 955.71
## - diabetes       1 955.97
## - smoking        1 956.23
## - log_cp         1 956.39
## - sex            1 956.65
## - serum_sodium   1 957.46
## <none>           957.56
## - anaemia       1 960.34
## - high_blood_pressure 1 960.90
## - log_serum_creatinine 1 973.92
## - ejection_fraction 1 974.75
## - age           1 975.99
##
## Step: AIC=955.71
## Surv(time, DEATH_EVENT) ~ age + log_cp + ejection_fraction +
##   log_serum_creatinine + serum_sodium + anaemia + diabetes +
##   high_blood_pressure + sex + smoking
##
##              Df      AIC
## - diabetes      1 954.09
## - smoking       1 954.30
## - log_cp        1 954.53
## - sex           1 954.68
## - serum_sodium  1 955.67
## <none>          955.71
## + platelets     1 957.56
## - anaemia       1 958.42
## - high_blood_pressure 1 959.00
## - log_serum_creatinine 1 972.35
## - ejection_fraction 1 972.85
## - age           1 974.03
##
## Step: AIC=954.09
## Surv(time, DEATH_EVENT) ~ age + log_cp + ejection_fraction +
##   log_serum_creatinine + serum_sodium + anaemia + high_blood_pressure +
##   sex + smoking
##
##              Df      AIC
## - smoking       1 952.62
## - log_cp        1 952.95
## - sex           1 953.15
## <none>          954.09
## - serum_sodium  1 954.31
## + diabetes      1 955.71
## + platelets     1 955.97
## - anaemia       1 956.89
## - high_blood_pressure 1 957.40
## - log_serum_creatinine 1 970.37
## - ejection_fraction 1 971.51
## - age           1 972.09
##
## Step: AIC=952.62
## Surv(time, DEATH_EVENT) ~ age + log_cp + ejection_fraction +

```

```

##      log_serum_creatinine + serum_sodium + anaemia + high_blood_pressure +
##      sex
##
##              Df      AIC
## - sex              1 951.27
## - log_cp            1 951.38
## <none>              952.62
## - serum_sodium      1 952.72
## + smoking           1 954.09
## + diabetes          1 954.30
## + platelets         1 954.56
## - anaemia           1 955.14
## - high_blood_pressure 1 956.07
## - log_serum_creatinine 1 968.50
## - age               1 970.18
## - ejection_fraction 1 970.21
##
## Step:  AIC=951.27
## Surv(time, DEATH_EVENT) ~ age + log_cp + ejection_fraction +
##      log_serum_creatinine + serum_sodium + anaemia + high_blood_pressure
##
##              Df      AIC
## - log_cp            1 949.91
## <none>              951.27
## - serum_sodium      1 951.32
## + sex               1 952.62
## + diabetes          1 952.86
## + smoking           1 953.15
## + platelets         1 953.26
## - anaemia           1 953.68
## - high_blood_pressure 1 955.26
## - log_serum_creatinine 1 966.88
## - ejection_fraction 1 968.21
## - age               1 968.34
##
## Step:  AIC=949.91
## Surv(time, DEATH_EVENT) ~ age + ejection_fraction + log_serum_creatinine +
##      serum_sodium + anaemia + high_blood_pressure
##
##              Df      AIC
## - serum_sodium      1 949.88
## <none>              949.91
## + log_cp            1 951.27
## + sex               1 951.38
## + diabetes          1 951.47
## + smoking           1 951.81
## + platelets         1 951.90
## - anaemia           1 952.07
## - high_blood_pressure 1 953.76
## - log_serum_creatinine 1 964.92
## - ejection_fraction 1 966.91
## - age               1 967.46
##
## Step:  AIC=949.88

```



```

## Surv(time, DEATH_EVENT) ~ age + ejection_fraction + log_serum_creatinine +
##   anaemia + high_blood_pressure
##
##               Df      AIC
## <none>                949.88
## + serum_sodium        1 949.91
## + diabetes            1 951.20
## + log_cp              1 951.32
## + sex                 1 951.35
## - anaemia             1 951.51
## + smoking            1 951.82
## + platelets           1 951.85
## - high_blood_pressure  1 953.46
## - age                 1 966.71
## - log_serum_creatinine 1 970.38
## - ejection_fraction   1 970.72

# View final model
summary(cox_step )

## Call:
## coxph(formula = Surv(time, DEATH_EVENT) ~ age + ejection_fraction +
##   log_serum_creatinine + anaemia + high_blood_pressure, data = heart_failure_data)
##
##   n= 299, number of events= 96
##
##               coef exp(coef) se(coef)      z Pr(>|z|)
## age              0.040006  1.040817  0.009106  4.393 1.12e-05 ***
## ejection_fraction -0.044393  0.956578  0.009963 -4.456 8.36e-06 ***
## log_serum_creatinine 0.993288  2.700098  0.189289  5.247 1.54e-07 ***
## anaemia           0.399027  1.490375  0.208236  1.916  0.0553 .
## high_blood_pressure 0.508459  1.662727  0.211286  2.406  0.0161 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##               exp(coef) exp(-coef) lower .95 upper .95
## age              1.0408      0.9608   1.0224   1.0596
## ejection_fraction 0.9566      1.0454   0.9381   0.9754
## log_serum_creatinine 2.7001      0.3704   1.8632   3.9129
## anaemia           1.4904      0.6710   0.9909   2.2415
## high_blood_pressure 1.6627      0.6014   1.0989   2.5158
##
## Concordance= 0.735 (se = 0.028 )
## Likelihood ratio test= 78.53 on 5 df,  p=2e-15
## Wald test              = 83.91 on 5 df,  p=<2e-16
## Score (logrank) test = 82.19 on 5 df,  p=3e-16

# interactions
cox_interaction <- coxph(
  Surv(time, DEATH_EVENT) ~ age + log_cp + ejection_fraction +
  log_serum_creatinine + serum_sodium +
  age:ejection_fraction + age:log_serum_creatinine,
  data = heart_failure_data

```

```

)
# compare
aic_comparison <- AIC(cox_full, cox_backward, cox_forward, cox_interaction)
print(aic_comparison)

##              df      AIC
## cox_full      11 957.5560
## cox_backward    5 949.8796
## cox_forward     5 949.8796
## cox_interaction  7 957.9352

# cox_forward & backward least AIC
# all age ejection_fraction log_serum_creatinine anaemia high_blood_pressure
# model's variable significance
final_model <- cox_forward
summary(final_model)

## Call:
## coxph(formula = Surv(time, DEATH_EVENT) ~ log_serum_creatinine +
##       age + ejection_fraction + high_blood_pressure + anaemia,
##       data = heart_failure_data)
##
##      n= 299, number of events= 96
##
##              coef exp(coef) se(coef)      z Pr(>|z|)
## log_serum_creatinine  0.993288  2.700098  0.189289  5.247 1.54e-07 ***
## age                  0.040006  1.040817  0.009106  4.393 1.12e-05 ***
## ejection_fraction    -0.044393  0.956578  0.009963 -4.456 8.36e-06 ***
## high_blood_pressure   0.508459  1.662727  0.211286  2.406  0.0161 *
## anaemia               0.399027  1.490375  0.208236  1.916  0.0553 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##              exp(coef) exp(-coef) lower .95 upper .95
## log_serum_creatinine    2.7001    0.3704    1.8632    3.9129
## age                    1.0408    0.9608    1.0224    1.0596
## ejection_fraction       0.9566    1.0454    0.9381    0.9754
## high_blood_pressure     1.6627    0.6014    1.0989    2.5158
## anaemia                 1.4904    0.6710    0.9909    2.2415
##
## Concordance= 0.735 (se = 0.028 )
## Likelihood ratio test= 78.53 on 5 df,  p=2e-15
## Wald test               = 83.91 on 5 df,  p=<2e-16
## Score (logrank) test = 82.19 on 5 df,  p=3e-16

# VIF test for final selected variables
vif(lm(time ~ age + ejection_fraction + log_serum_creatinine +
       anaemia + high_blood_pressure,
       data = heart_failure_data))

##              age      ejection_fraction log_serum_creatinine
##              1.082543              1.018358              1.076283

```

```
##          anaemia  high_blood_pressure
##          1.009414          1.015915
```

All VIF values are close to 1 (range: 1.01-1.08) No multicollinearity issues age ejection_fraction log_serum_creatinine anaemia high_blood_pressure

AFT

```
# Weibull AFT model
weib_full <- survreg(
  Surv(time, DEATH_EVENT) ~ age + log_cp + ejection_fraction +
    platelets + log_serum_creatinine + serum_sodium +
    anaemia + diabetes + high_blood_pressure + sex + smoking,
  dist = "weibull",
  data = heart_failure_data
)
weib_step <- stepAIC(weib_full, direction = "both")

## Start:  AIC=1280.97
## Surv(time, DEATH_EVENT) ~ age + log_cp + ejection_fraction +
##   platelets + log_serum_creatinine + serum_sodium + anaemia +
##   diabetes + high_blood_pressure + sex + smoking
##
##               Df    AIC
## - platelets      1 1279.2
## - diabetes       1 1279.4
## - smoking        1 1279.6
## - log_cp         1 1280.0
## - sex            1 1280.1
## - serum_sodium   1 1280.8
## <none>           1281.0
## - anaemia        1 1284.3
## - high_blood_pressure 1 1284.7
## - log_serum_creatinine 1 1298.3
## - ejection_fraction 1 1299.3
## - age            1 1300.5
##
## Step:  AIC=1279.15
## Surv(time, DEATH_EVENT) ~ age + log_cp + ejection_fraction +
##   log_serum_creatinine + serum_sodium + anaemia + diabetes +
##   high_blood_pressure + sex + smoking
##
##               Df    AIC
## - diabetes      1 1277.5
## - smoking       1 1277.7
## - log_cp        1 1278.1
## - sex           1 1278.2
## - serum_sodium  1 1279.0
## <none>          1279.2
## + platelets     1 1281.0
## - anaemia       1 1282.4
## - high_blood_pressure 1 1282.8
```

```

## - log_serum_creatinine 1 1296.7
## - ejection_fraction    1 1297.4
## - age                  1 1298.5
##
## Step: AIC=1277.53
## Surv(time, DEATH_EVENT) ~ age + log_cp + ejection_fraction +
##     log_serum_creatinine + serum_sodium + anaemia + high_blood_pressure +
##     sex + smoking
##
##               Df    AIC
## - smoking      1 1276.0
## - log_cp       1 1276.5
## - sex          1 1276.6
## <none>         1277.5
## - serum_sodium 1 1277.6
## + diabetes     1 1279.2
## + platelets    1 1279.4
## - anaemia      1 1280.9
## - high_blood_pressure 1 1281.2
## - log_serum_creatinine 1 1294.8
## - ejection_fraction 1 1296.2
## - age          1 1296.6
##
## Step: AIC=1276.01
## Surv(time, DEATH_EVENT) ~ age + log_cp + ejection_fraction +
##     log_serum_creatinine + serum_sodium + anaemia + high_blood_pressure +
##     sex
##
##               Df    AIC
## - sex          1 1274.7
## - log_cp       1 1274.9
## - serum_sodium 1 1276.0
## <none>         1276.0
## + smoking     1 1277.5
## + diabetes     1 1277.7
## + platelets    1 1277.9
## - anaemia      1 1279.1
## - high_blood_pressure 1 1279.9
## - log_serum_creatinine 1 1292.9
## - age          1 1294.7
## - ejection_fraction 1 1294.8
##
## Step: AIC=1274.7
## Surv(time, DEATH_EVENT) ~ age + log_cp + ejection_fraction +
##     log_serum_creatinine + serum_sodium + anaemia + high_blood_pressure
##
##               Df    AIC
## - log_cp       1 1273.5
## - serum_sodium 1 1274.7
## <none>         1274.7
## + sex          1 1276.0
## + diabetes     1 1276.3
## + smoking      1 1276.6
## + platelets    1 1276.7

```

```

## - anaemia          1 1277.7
## - high_blood_pressure 1 1279.1
## - log_serum_creatinine 1 1291.3
## - ejection_fraction 1 1292.8
## - age              1 1292.8
##
## Step: AIC=1273.49
## Surv(time, DEATH_EVENT) ~ age + ejection_fraction + log_serum_creatinine +
##   serum_sodium + anaemia + high_blood_pressure
##
##               Df      AIC
## - serum_sodium      1 1273.3
## <none>                1273.5
## + log_cp            1 1274.7
## + sex                1 1274.9
## + diabetes           1 1275.0
## + smoking            1 1275.4
## + platelets          1 1275.5
## - anaemia           1 1276.2
## - high_blood_pressure 1 1277.7
## - log_serum_creatinine 1 1289.3
## - ejection_fraction  1 1291.5
## - age                1 1292.1
##
## Step: AIC=1273.35
## Surv(time, DEATH_EVENT) ~ age + ejection_fraction + log_serum_creatinine +
##   anaemia + high_blood_pressure
##
##               Df      AIC
## <none>                1273.3
## + serum_sodium      1 1273.5
## + log_cp            1 1274.7
## + diabetes           1 1274.7
## + sex                1 1274.8
## + platelets          1 1275.3
## + smoking            1 1275.3
## - anaemia           1 1275.5
## - high_blood_pressure 1 1277.3
## - age                1 1291.2
## - log_serum_creatinine 1 1294.7
## - ejection_fraction  1 1295.3

# Exponential AFT model
exp_full <- survreg(
  Surv(time, DEATH_EVENT) ~ age + log_cp + ejection_fraction +
    platelets + log_serum_creatinine + serum_sodium +
    anaemia + diabetes + high_blood_pressure + sex + smoking,
  dist = "exponential",
  data = heart_failure_data
)
exp_step <- stepAIC(exp_full, direction = "both")

## Start: AIC=1279.23
## Surv(time, DEATH_EVENT) ~ age + log_cp + ejection_fraction +

```

```

##      platelets + log_serum_creatinine + serum_sodium + anaemia +
##      diabetes + high_blood_pressure + sex + smoking
##
##              Df      AIC
## - platelets      1 1277.4
## - diabetes      1 1277.7
## - smoking       1 1277.9
## - log_cp        1 1278.2
## - sex           1 1278.4
## - serum_sodium  1 1279.1
## <none>          1279.2
## - anaemia       1 1282.9
## - high_blood_pressure 1 1283.4
## - log_serum_creatinine 1 1297.1
## - ejection_fraction 1 1297.9
## - age           1 1299.7
##
## Step:  AIC=1277.4
## Surv(time, DEATH_EVENT) ~ age + log_cp + ejection_fraction +
##      log_serum_creatinine + serum_sodium + anaemia + diabetes +
##      high_blood_pressure + sex + smoking
##
##              Df      AIC
## - diabetes      1 1275.8
## - smoking       1 1276.0
## - log_cp        1 1276.4
## - sex           1 1276.4
## - serum_sodium  1 1277.3
## <none>          1277.4
## + platelets     1 1279.2
## - anaemia       1 1281.0
## - high_blood_pressure 1 1281.5
## - log_serum_creatinine 1 1295.6
## - ejection_fraction 1 1296.0
## - age           1 1297.8
##
## Step:  AIC=1275.79
## Surv(time, DEATH_EVENT) ~ age + log_cp + ejection_fraction +
##      log_serum_creatinine + serum_sodium + anaemia + high_blood_pressure +
##      sex + smoking
##
##              Df      AIC
## - smoking       1 1274.3
## - log_cp        1 1274.8
## - sex           1 1274.9
## <none>          1275.8
## - serum_sodium  1 1275.9
## + diabetes      1 1277.4
## + platelets     1 1277.7
## - anaemia       1 1279.5
## - high_blood_pressure 1 1279.9
## - log_serum_creatinine 1 1293.6
## - ejection_fraction 1 1294.8
## - age           1 1295.9

```

```

##
## Step: AIC=1274.29
## Surv(time, DEATH_EVENT) ~ age + log_cp + ejection_fraction +
##     log_serum_creatinine + serum_sodium + anaemia + high_blood_pressure +
##     sex
##
##           Df      AIC
## - sex           1 1273.0
## - log_cp         1 1273.2
## <none>           1274.3
## - serum_sodium   1 1274.3
## + smoking        1 1275.8
## + diabetes       1 1276.0
## + platelets      1 1276.2
## - anaemia        1 1277.8
## - high_blood_pressure 1 1278.6
## - log_serum_creatinine 1 1291.7
## - ejection_fraction 1 1293.5
## - age           1 1294.0
##
## Step: AIC=1272.97
## Surv(time, DEATH_EVENT) ~ age + log_cp + ejection_fraction +
##     log_serum_creatinine + serum_sodium + anaemia + high_blood_pressure
##
##           Df      AIC
## - log_cp         1 1271.8
## <none>           1273.0
## - serum_sodium   1 1273.0
## + sex           1 1274.3
## + diabetes       1 1274.5
## + smoking        1 1274.9
## + platelets      1 1275.0
## - anaemia        1 1276.3
## - high_blood_pressure 1 1277.8
## - log_serum_creatinine 1 1290.1
## - ejection_fraction 1 1291.5
## - age           1 1292.1
##
## Step: AIC=1271.77
## Surv(time, DEATH_EVENT) ~ age + ejection_fraction + log_serum_creatinine +
##     serum_sodium + anaemia + high_blood_pressure
##
##           Df      AIC
## - serum_sodium   1 1271.7
## <none>           1271.8
## + log_cp         1 1273.0
## + sex           1 1273.2
## + diabetes       1 1273.3
## + smoking        1 1273.7
## + platelets      1 1273.8
## - anaemia        1 1274.8
## - high_blood_pressure 1 1276.4
## - log_serum_creatinine 1 1288.2
## - ejection_fraction 1 1290.2

```

```
## - age          1 1291.4
##
## Step:  AIC=1271.68
## Surv(time, DEATH_EVENT) ~ age + ejection_fraction + log_serum_creatinine +
##      anaemia + high_blood_pressure
##
##              Df      AIC
## <none>          1271.7
## + serum_sodium    1 1271.8
## + log_cp          1 1273.0
## + diabetes        1 1273.0
## + sex             1 1273.1
## + smoking         1 1273.6
## + platelets       1 1273.6
## - anaemia         1 1274.2
## - high_blood_pressure 1 1276.0
## - age             1 1290.5
## - log_serum_creatinine 1 1293.8
## - ejection_fraction 1 1294.2
```

```
summary(weib_step)
```

```
##
## Call:
## survreg(formula = Surv(time, DEATH_EVENT) ~ age + ejection_fraction +
##      log_serum_creatinine + anaemia + high_blood_pressure, data = heart_failure_data,
##      dist = "weibull")
##
##              Value Std. Error      z      p
## (Intercept)    7.71865    0.70792 10.90 < 2e-16
## age           -0.04344    0.00984 -4.41 1.0e-05
## ejection_fraction  0.04812    0.01107  4.35 1.4e-05
## log_serum_creatinine -1.06008    0.20791 -5.10 3.4e-07
## anaemia        -0.44891    0.21693 -2.07  0.039
## high_blood_pressure -0.54976    0.22021 -2.50  0.013
## Log(scale)      0.04989    0.08884  0.56  0.574
##
## Scale= 1.05
##
## Weibull distribution
## Loglik(model)= -629.7   Loglik(intercept only)= -670.4
##  Chisq= 81.52 on 5 degrees of freedom, p= 4e-16
## Number of Newton-Raphson Iterations: 6
## n= 299
```

```
summary(exp_step)
```

```
##
## Call:
## survreg(formula = Surv(time, DEATH_EVENT) ~ age + ejection_fraction +
##      log_serum_creatinine + anaemia + high_blood_pressure, data = heart_failure_data,
##      dist = "exponential")
##
##              Value Std. Error      z      p
## (Intercept)    7.63580    0.65992 11.57 < 2e-16
```



```
## age                -0.04210    0.00909 -4.63 3.6e-06
## ejection_fraction  0.04620    0.01000  4.62 3.8e-06
## log_serum_creatinine -1.02200    0.18656 -5.48 4.3e-08
## anaemia            -0.44230    0.20612 -2.15 0.0319
## high_blood_pressure -0.53806    0.20848 -2.58 0.0099
##
## Scale fixed at 1
##
## Exponential distribution
## Loglik(model)= -629.8   Loglik(intercept only)= -672.5
##  Chisq= 85.41 on 5 degrees of freedom, p= 6.2e-17
## Number of Newton-Raphson Iterations: 5
## n= 299
```

```
AIC(weib_step, exp_step, cox_step)
```

```
##          df      AIC
## weib_step  7 1273.3545
## exp_step   6 1271.6767
## cox_step   5  949.8796
```

```
cox: age ejection_fraction log_serum_creatinine anaemia high_blood_pressure
```