Advanced Applied Statistics Homework 9 Solutions

Joshua D. Ingram

2022-12-07

Read

Study the handout, and also read the one-page summary of the Kaplan-Meier estimator and Cox's proportional hazard model on the paper attached.

Problem 1

Fit an Exponential or a Weibull regression model to investigate the effect of sex on the hazard of crashing in the game Snake, based on the small dataset of survival times we observed in class. Make sure to interpret the effect of sex on the hazard, and to plot both the hazard function and the implied survival function. Estimate the median survival time when sex = female.

Answer:

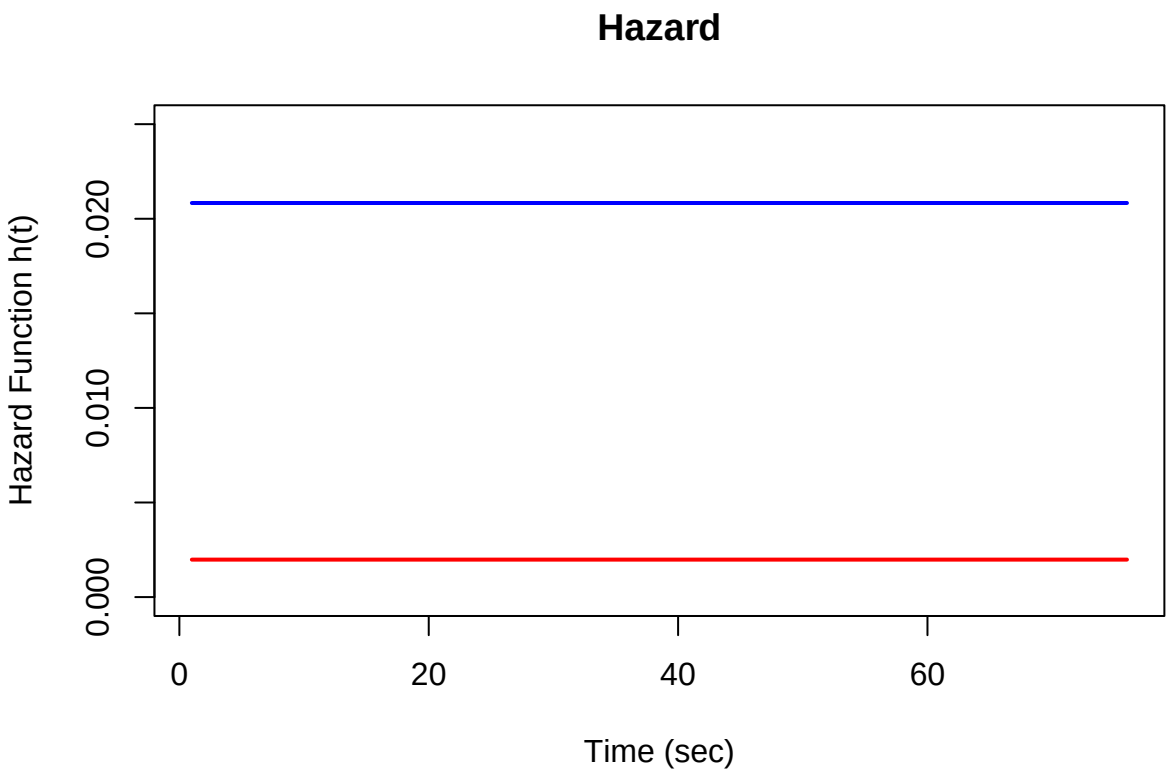
Exponential Regression Model

```
# Fit exponential regression model
fit_exp <- flexsurvreg(Surv(time = y, event = censored) ~ sex, dist = "exp", data = snake)
fit_exp
```

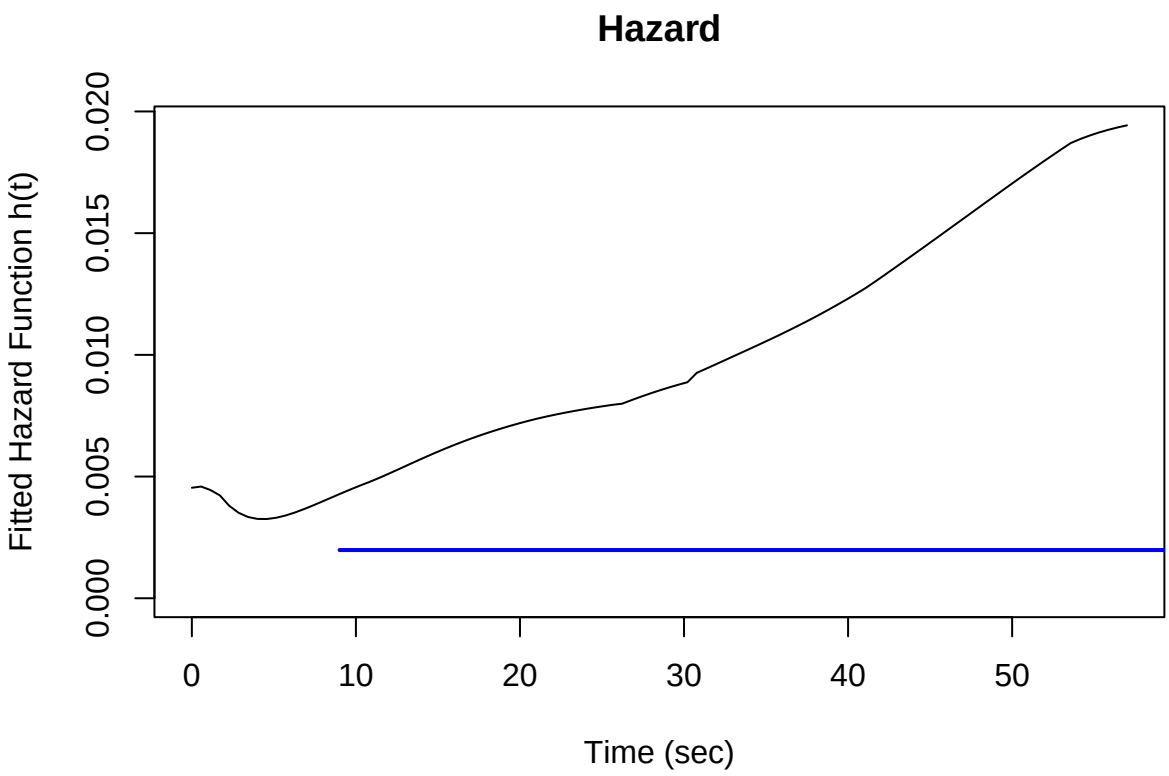
```
## Call:
## flexsurvreg(formula = Surv(time = y, event = censored) ~ sex,
##     data = snake, dist = "exp")
##
## Estimates:
##       data mean  est       L95%      U95%      se         exp(est)  L95%
## rate        NA    0.02083   0.00782   0.05551   0.01042        NA        NA
## sexm   0.58333   -2.35336  -4.54466  -0.16205   1.11803   0.09505   0.01062
##       U95%
## rate        NA
## sexm   0.85040
##
## N = 12,  Events: 5,  Censored: 7
## Total time at risk: 697
## Log-likelihood = -26.70936, df = 2
## AIC = 57.41872
```

```
pred.hazard_m = predict(fit_exp, newdata=data.frame(sex = "m"), type="hazard", times=seq(1,80,5))
pred.hazard_f = predict(fit_exp, newdata=data.frame(sex = "F"), type="hazard", times=seq(1,80,5))
pred.hazm = as.data.frame(pred.hazard_m$.pred)
pred.hazf = as.data.frame(pred.hazard_f$.pred)
pred.haz = merge(pred.hazm, pred.hazf, by = ".time")
```

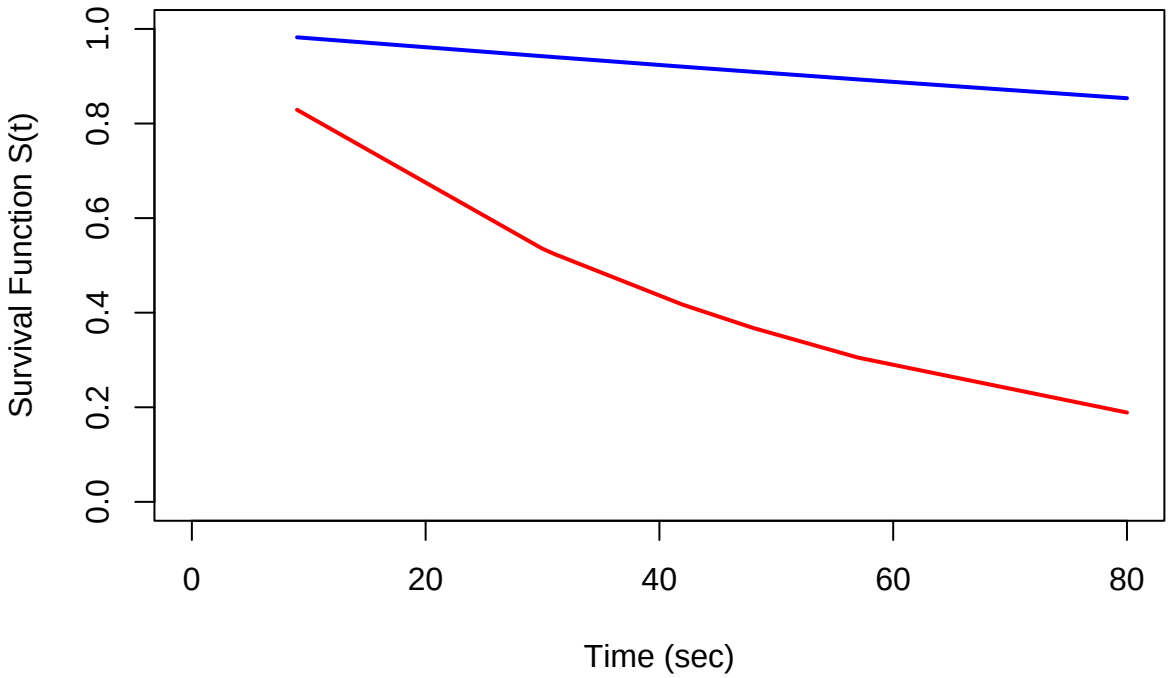
```
plot(pred.haz$.time, pred.haz$.pred_hazard.y, type="l", col="blue", lwd=2, main="Hazard", xlab="Time (s
lines(pred.haz$.time, pred.haz$.pred_hazard.x, col="red", lwd=2)
```




```
plot(fit_exp, newdata = data.frame(sex = "m"), type = "hazard", col = "blue", ci = FALSE,
     main = "Hazard", xlab = "Time (sec)", ylab = "Fitted Hazard Function h(t)")
lines(fit_exp, newdata = data.frame(sex = "F"), type = "hazard", col = "red", ci = FALSE)
```




```
plot(fit_exp, newdata = data.frame(sex = "m"), type = "survival", col = "blue", col.obs = "white", ci =
     main = "Survival", xlab = "Time (sec)", ylab = "Survival Function S(t)")
lines(fit_exp, newdata = data.frame(sex="F"), type = "survival", col = "red", ci = FALSE)
```



```
summary(fit_exp, type="median", newdata = data.frame(sex='F'))
```

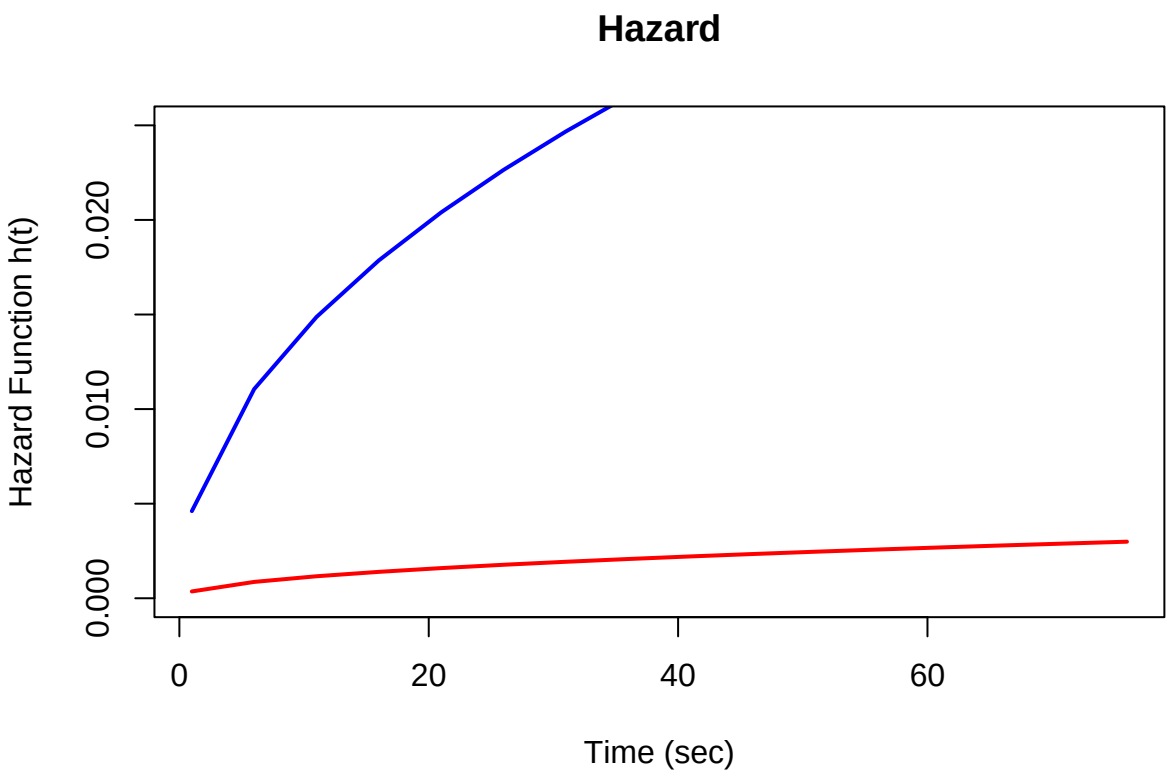
```
## sex=F
##        est      lcl      ucl
## 1 33.27106 12.82183 88.95989
```

Weibull Regression Model

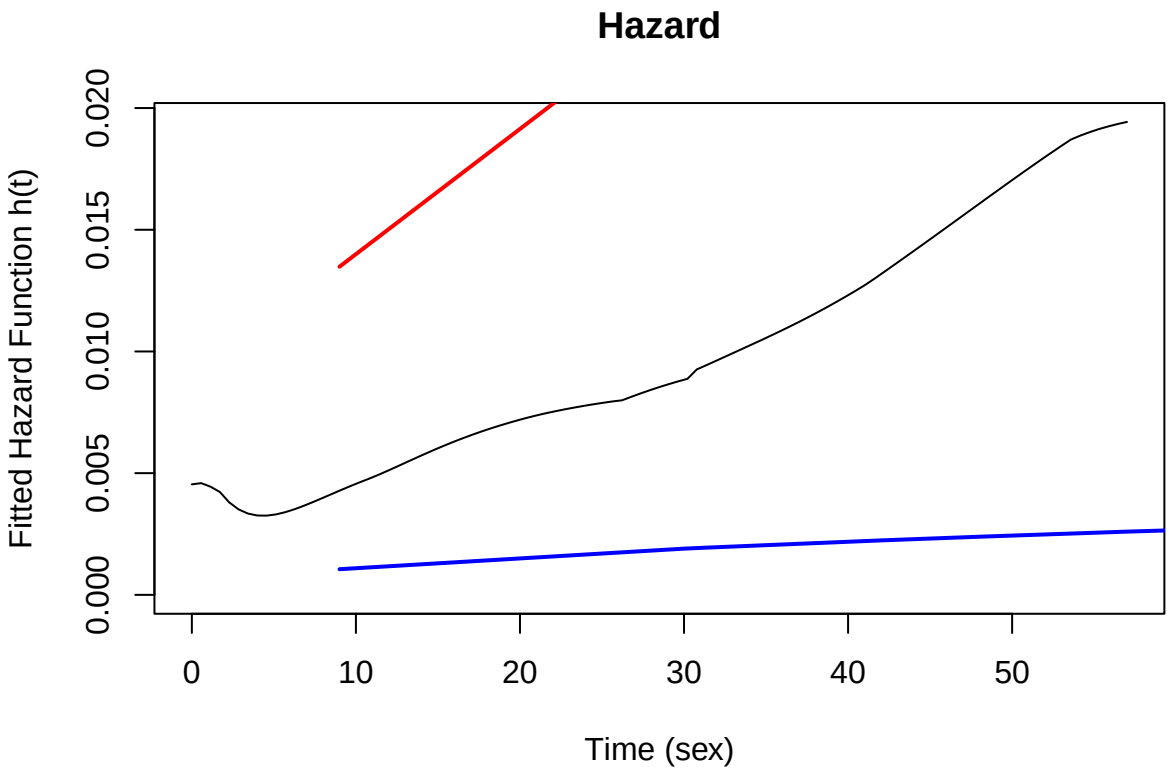
```
fit_weib <- flexsurvreg(Surv(time = y, event = censored) ~ sex, dist = "weibullPH", data = snake)
fit_weib
```

```
## Call:
## flexsurvreg(formula = Surv(time = y, event = censored) ~ sex,
##     data = snake, dist = "weibullPH")
##
## Estimates:
##        data mean  est        L95%       U95%       se         exp(est)
## shape         NA   1.49e+00   7.17e-01   3.09e+00   5.55e-01         NA
## scale         NA   3.09e-03   3.71e-05   2.58e-01   6.98e-03         NA
## sexm    5.83e-01  -2.55e+00  -4.77e+00  -3.26e-01   1.13e+00   7.81e-02
##        L95%       U95%
## shape         NA         NA
## scale         NA         NA
## sexm    8.46e-03   7.22e-01
##
## N = 12,  Events: 5,  Censored: 7
## Total time at risk: 697
## Log-likelihood = -26.23282, df = 3
## AIC = 58.46565
```

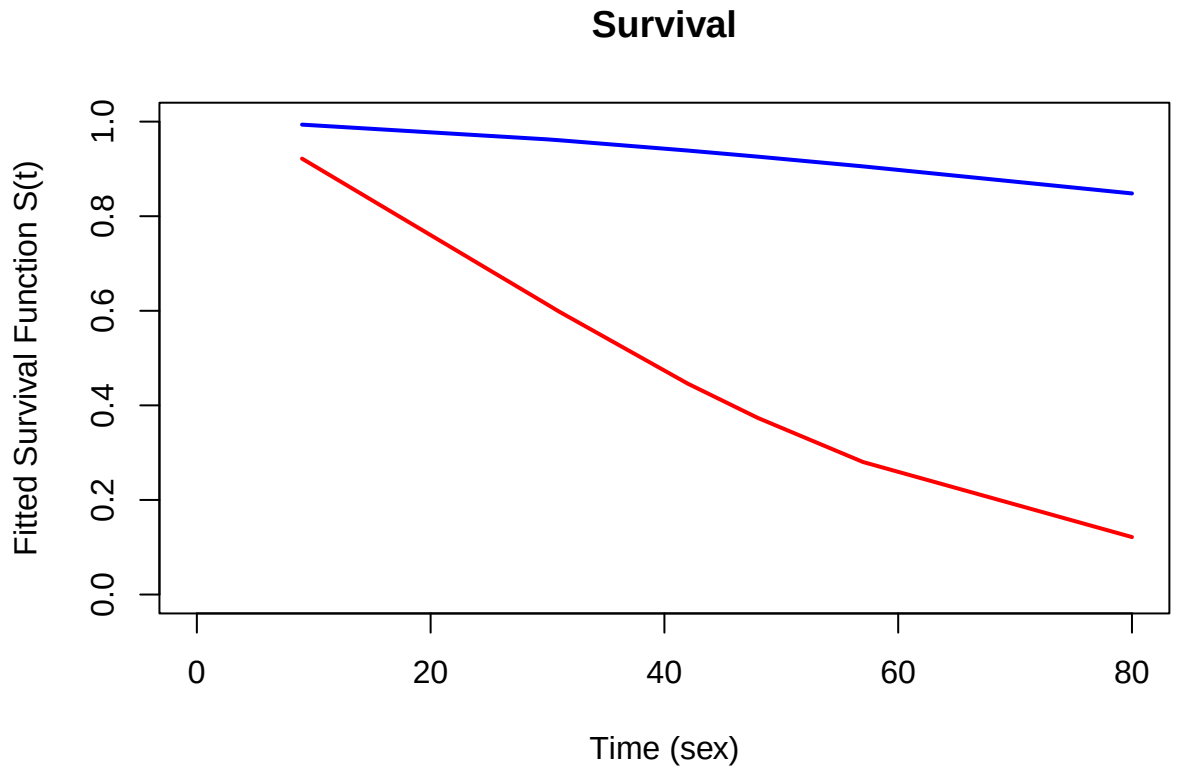
```
pred.hazard_m = predict(fit_weib, newdata=data.frame(sex = "m"), type="hazard", times=seq(1,80,5))
pred.hazard_f = predict(fit_weib, newdata=data.frame(sex = "F"), type="hazard", times=seq(1,80,5))
pred.hazm = as.data.frame(pred.hazard_m$.pred)
pred.hazf = as.data.frame(pred.hazard_f$.pred)
pred.haz = merge(pred.hazm, pred.hazf, by = ".time")

plot(pred.haz$.time, pred.haz$.pred_hazard.y, type="l", col="blue", lwd=2, main="Hazard", xlab="Time (s
lines(pred.haz$.time, pred.haz$.pred_hazard.x, col="red", lwd=2)
```




```
plot(fit_weib, newdata = data.frame(sex = "m"), type = "hazard", col = "blue", ci = FALSE,
     main = "Hazard", xlab = "Time (sex)", ylab = "Fitted Hazard Function h(t)")
lines(fit_weib, newdata = data.frame(sex = "F"), type = "hazard", col = "red", ci = FALSE, ylab = c(0, 
```



```
plot(fit_weib, newdata = data.frame(sex = "m"), type = "survival", col = "blue", col.obs = "white", ci =
     main = "Survival", xlab = "Time (sex)", ylab = "Fitted Survival Function S(t)")
lines(fit_weib, newdata = data.frame(sex = "F"), type = "survival", col = "red", ci = FALSE)
```




```
summary(fit_weib, type="median", newdata = data.frame(sex='F'))
```

```
## sex=F
##        est      lcl      ucl
## 1 37.90328 2.988094 62.44556
```