|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **PM592: Regression Analysis for Data Science** | | | Name: Flemming Wu |  |  |
| **HW11** |  |  |  |  |  |
| *Survival Analysis* | | | | |  |

**Instructions**

* Answer questions directly within this document.
* Upload to Blackboard by the due date & time.
* Clearly indicate your answers to all questions.
* If a question requires analysis, attach all relevant output to this document in the appropriate area. Do not attach superfluous output.
* For the purpose of this assignment, statistical evidencerefers to a test statistic and associated p-value.
* If a question requires a conclusion, it must be phrased professionally and coherently.
* There are 2 questions and 30 points possible.

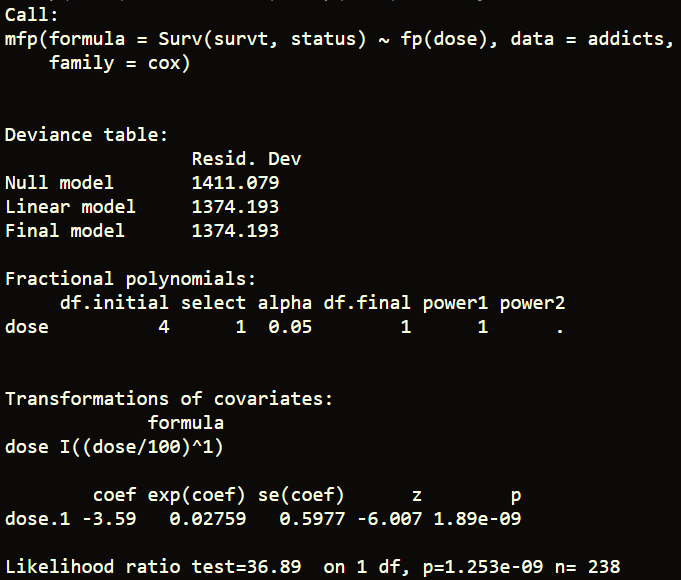
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Question 1** |  |  |  |  | [20 points] |  |

In an Australian study on drug addiction, recovering heroin addict patients were examined at two methadone treatment clinics. Patients were examined at these clinics until they either dropped out of the clinic (failure) or were censored. The two clinics differed in terms of their live-in policies for patients. The data is presented in **addicts.dta**.

Does the patient’s maximum methadone dose and having a prison record affect their chance of dropping out of the treatment, adjusting for clinic? Answer these questions by performing a Cox proportional hazards regression model.

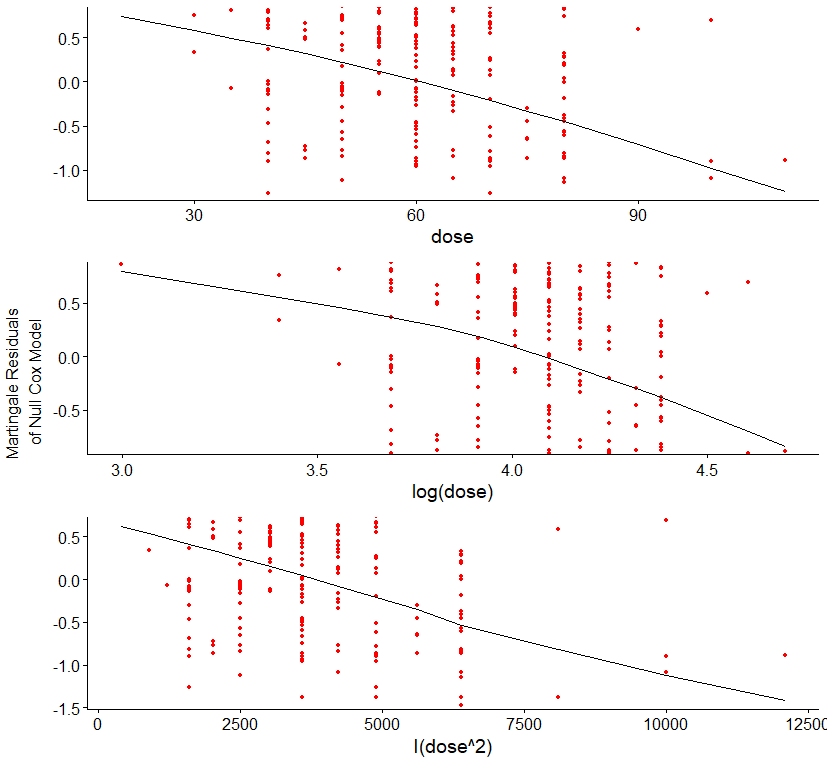
|  |  |
| --- | --- |
| Variable | Description |
| Id | Patient ID |
| survt | Time (days) until either drop out (failure) or censored |
| status | 1 = dropped out, 0 = censored |
| clinic | 1 = patient attended Clinic 1, 2 = patient attended Clinic 2 |
| prison | 1 = patient had a prison record, 0 = patient did not have a prison record |
| dose | Patient’s maximum methadone dose (mg/day) |

|  |  |
| --- | --- |
|  | 1a. [3 points] What is the functional form for dose? Comment on the results from the MFP procedure and Martingale residuals. |



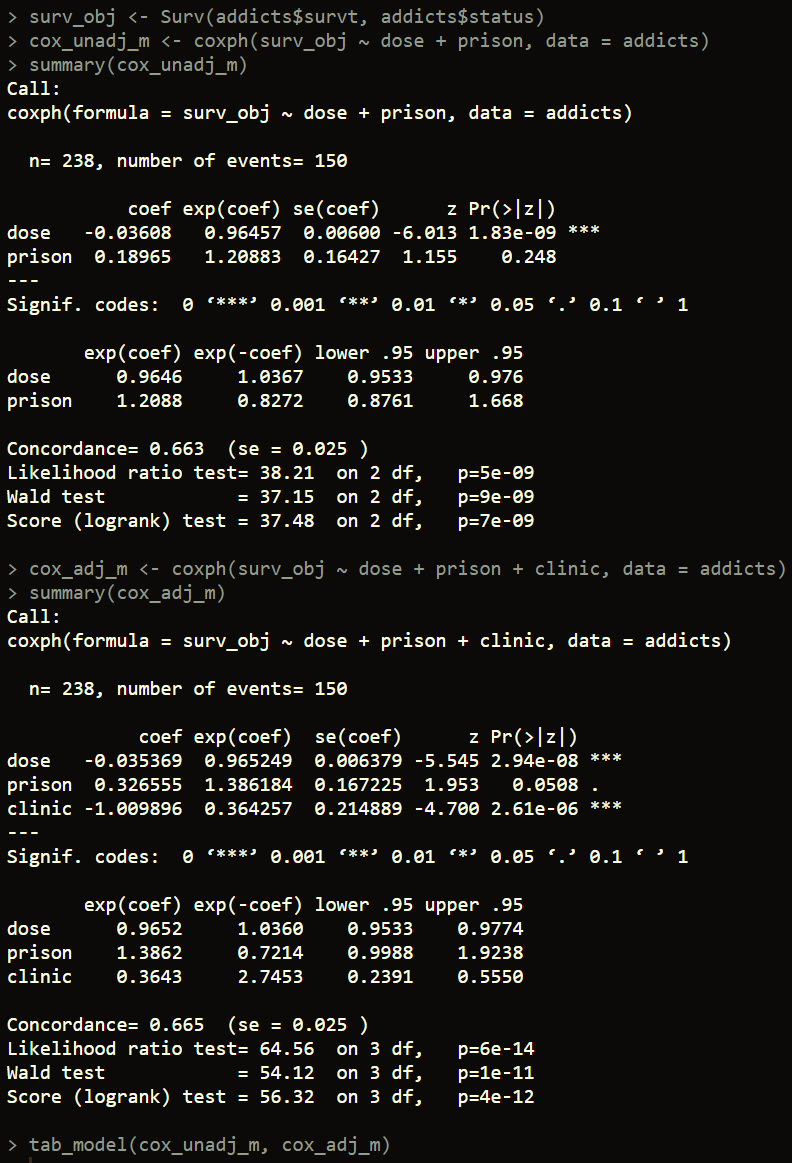
The fractional polynomials approach suggests that the form of dose is linear.

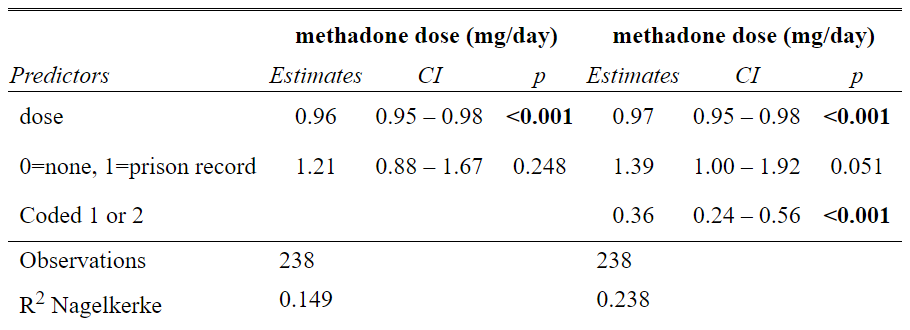




The Martingale residuals also confirm the that the functional form for dose is linear.

|  |  |
| --- | --- |
|  | 1b. [4 points] Construct a model with prison and dose as independent variables. Report the parameter estimates of these variables in a model without clinic, and a model with clinic. |

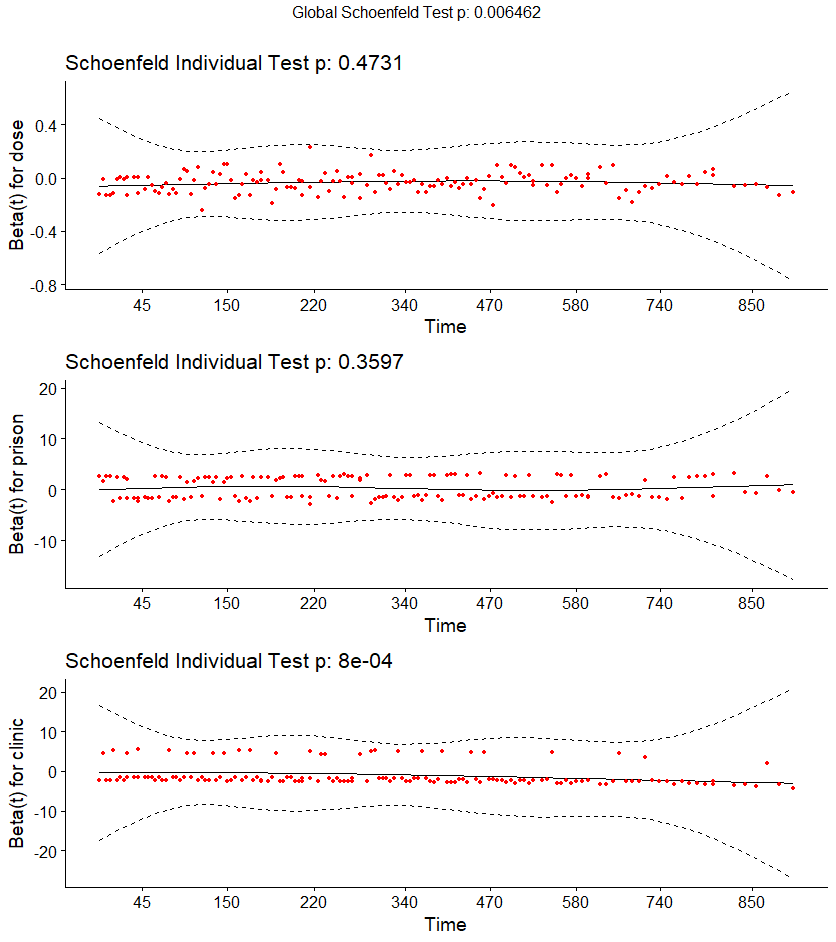




The parameter estimate for dose was 0.96 (CI = (0.95, 0.98), p<0.001) and 1.21 for prison record (CI = (0.88, 1.67), p=0.248) in the unadjusted model. After adjusting for clinic, the parameter estimate for dose was 0.97 (CI = (0.95, 0.98), p<0.001) and the parameter estimate for prison record was 1.39 (CI = (1.00, 1.92), p=0.051).

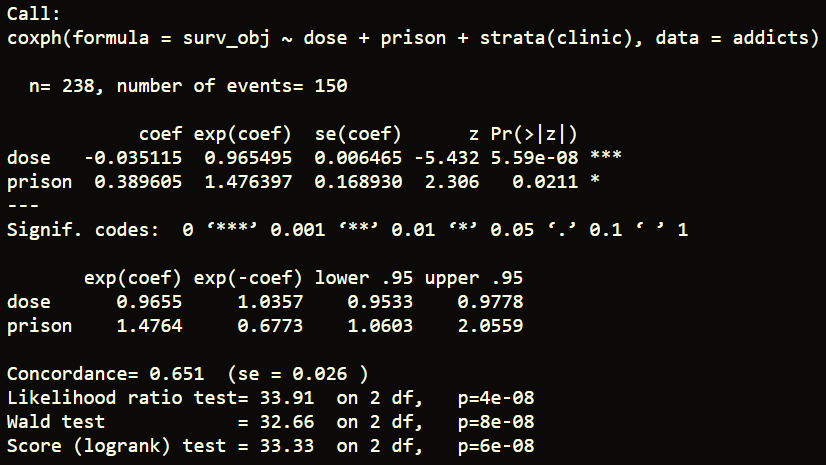
|  |  |
| --- | --- |
|  | 1c. [4 points] In the model adjusting for clinic, assess the proportional hazards assumption. Does it appear to be violated for any variable? |





The proportional hazards assumption does not appear to be violated for dose (p=0.473) and for prison (p=0.360). However, the proportional hazards assumption seems to be violated for clinic (p<0.001).

|  |  |
| --- | --- |
|  | 1d. [4 points] Re-fit your model to account for proportional hazards, and then assess your model goodness-of-fit. Comment on the Cox-Snell residuals, deviance residuals, and dfbeta values. |



|  |  |
| --- | --- |
|  | 1e. [5 points] Write a conclusion paragraph explaining the analysis you performed and the results you found, specifically addressing the research question. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Question 2** |  |  |  |  | [10 points] |  |

Read the article by Matas et al. (2015) about predictors of dropout in driving simulators.

|  |  |
| --- | --- |
|  | 2a. [5 points] Identify the following in this paper:   * The entry time under which observation began on individuals * The criteria for determining the event, and the criteria for determining censoring * Any time-dependent covariates |

|  |  |
| --- | --- |
|  | 2b. [5 points] What variables did the authors include in their Cox proportional hazards model? What did they do to check the proportional hazards assumption, and was the assumption met? Based on this, do you feel confident about their model estimates? |