**Group Quiz on AFT**

**Group members (your names):**

Siao Sun

Lulu Deng

Linna Hu

**Instruction:** Use the dataset “QuizAFT\_dataset.txt”

Please email the solutions, in a word document, to me by midnight today (Thursday, 10/4).

Your solutions must include the SAS (or R) code, the relevant portion of the output, and your summary for each given problem based on the output.

*Input dataset (SAS)*

**data** AFT;

input stage time age year status;

cards;

1 0.6 77 76 1

1 1.3 53 71 1

1 2.4 45 71 1

1 2.5 57 78 0

1 3.2 58 74 1

1 3.2 51 77 0

1 3.3 76 74 1

1 3.3 63 77 0

1 3.5 43 71 1

1 3.5 60 73 1

1 4 52 71 1

1 4 63 76 1

1 4.3 86 74 1

1 4.5 48 76 0

1 4.5 68 76 0

1 5.3 81 72 1

1 5.5 70 75 0

1 5.9 58 75 0

1 5.9 47 75 0

1 6 75 73 1

1 6.1 77 75 0

1 6.2 64 75 0

1 6.4 77 72 1

1 6.5 67 70 1

1 6.5 79 74 0

1 6.7 61 74 0

1 7 66 74 0

1 7.4 68 71 1

1 7.4 73 73 0

1 8.1 56 73 0

1 8.1 73 73 0

1 9.6 58 71 0

1 10.7 68 70 0

2 0.2 86 74 1

2 1.8 64 77 1

2 2 63 75 1

2 2.2 71 78 0

2 2.6 67 78 0

2 3.3 51 77 0

2 3.6 70 77 1

2 3.6 72 77 0

2 4 81 71 1

2 4.3 47 76 0

2 4.3 64 76 0

2 5 66 76 0

2 6.2 74 72 1

2 7 62 73 1

2 7.5 50 73 0

2 7.6 53 73 0

2 9.3 61 71 0

3 0.3 49 72 1

3 0.3 71 76 1

3 0.5 57 74 1

3 0.7 79 77 1

3 0.8 82 74 1

3 1 49 76 1

3 1.3 60 76 1

3 1.6 64 72 1

3 1.8 74 71 1

3 1.9 72 74 1

3 1.9 53 74 1

3 3.2 54 75 1

3 3.5 81 74 1

3 3.7 52 77 0

3 4.5 66 76 0

3 4.8 54 76 0

3 4.8 63 76 0

3 5 59 73 1

3 5 49 76 0

3 5.1 69 76 0

3 6.3 70 72 1

3 6.4 65 72 1

3 6.5 65 74 0

3 7.8 68 72 1

3 8 78 73 0

3 9.3 69 71 0

3 10.1 51 71 0

4 0.1 65 72 1

4 0.3 71 76 1

4 0.4 76 77 1

4 0.8 65 76 1

4 0.8 78 77 1

4 1 41 77 1

4 1.5 68 73 1

4 2 69 76 1

4 2.3 62 71 1

4 2.9 74 78 0

4 3.6 71 75 1

4 3.8 84 74 1

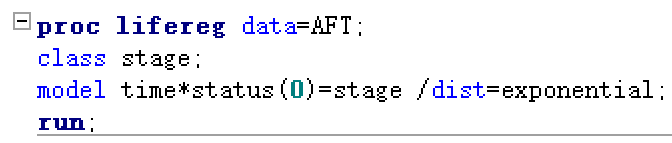
4 4.3 48 76 0

;

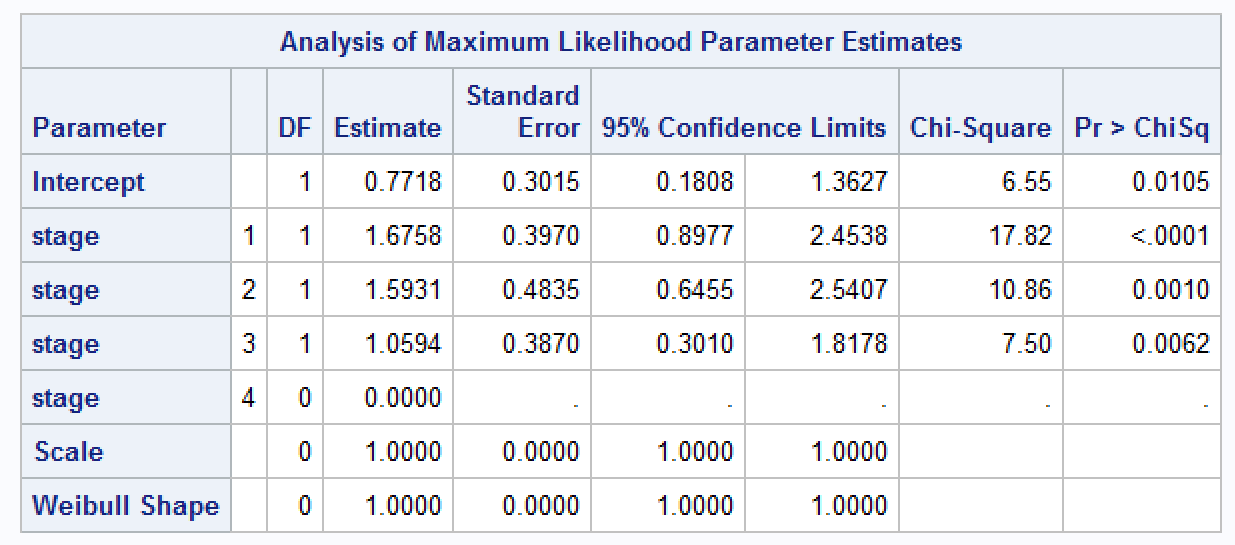
**run**;

1. The dataset contains survival data from 90 patients with larynx cancer. Treating disease stage as a categorical variable (so you need to either define 3 dummy variables or declare it as a class variable), fit an exponential AFT model to the data with disease stage as covariates in the model. Answer the following questions:
2. Compare the mean survival times for patients with different disease stages.

**SAS code:**



**Results:**

****

The mean survival times for patients with different disease stages:

Stage 1 vs 4: exp(-1.6758) = 0.1871584

Stage 2 vs 4: exp(-1.5931) = 0.2032944

Stage 3 vs 4: exp(-1.0594) = 0.3466637

Stage 1 vs 2: exp(-1.6758-(-1.5931))= 0.9206273

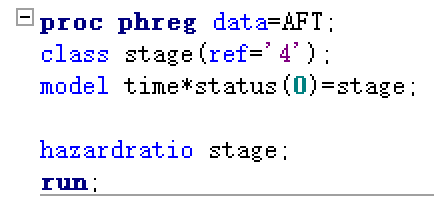
Stage 1 vs 3: exp(-1.6758-(-1.0594))= 0.5398845

Stage 2 vs 3: exp(-1.5931-(-1.0594))= 0.5864312

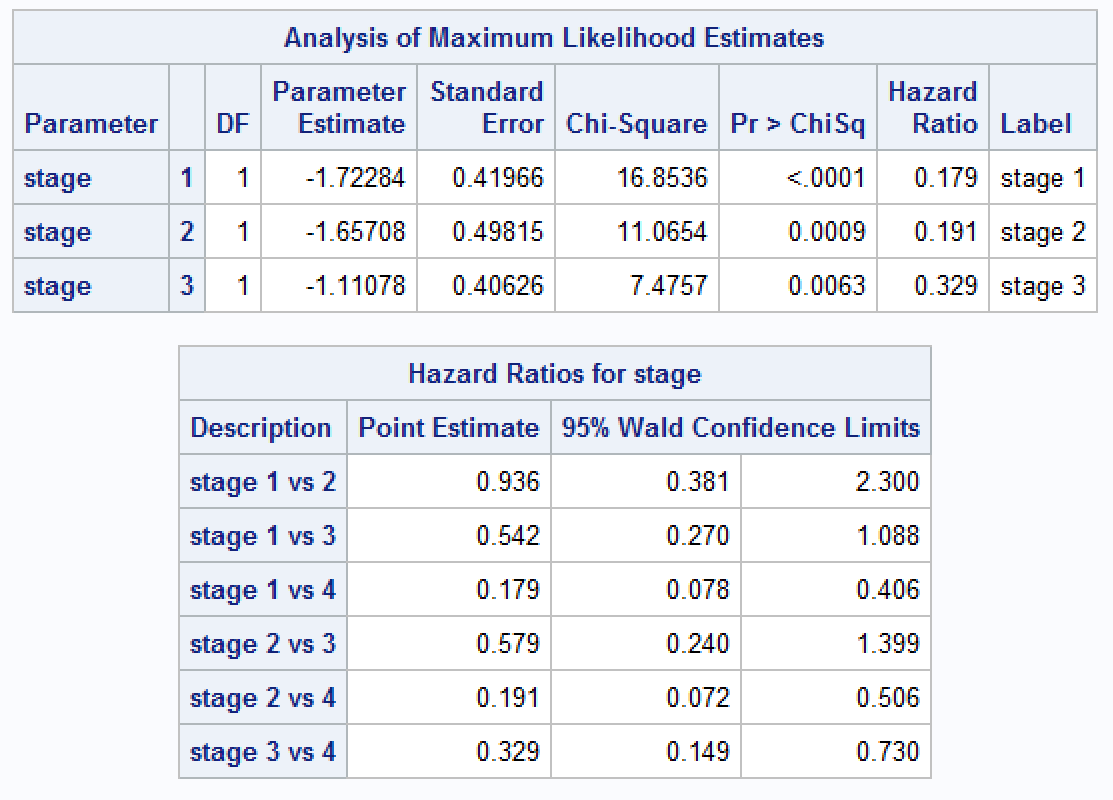
Stage 1 < Stage 2 < Stage 3 < Stage 4

1. Since the exponential AFT model is also a proportional hazards model, find the estimates and 95% CIs for the hazard ratios comparing patients with different disease stages.

**SAS code:**

****

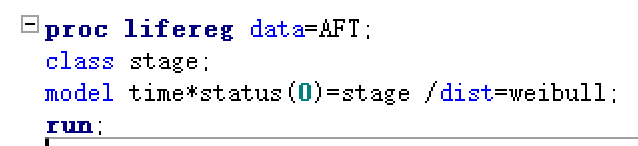
**Results:**

****

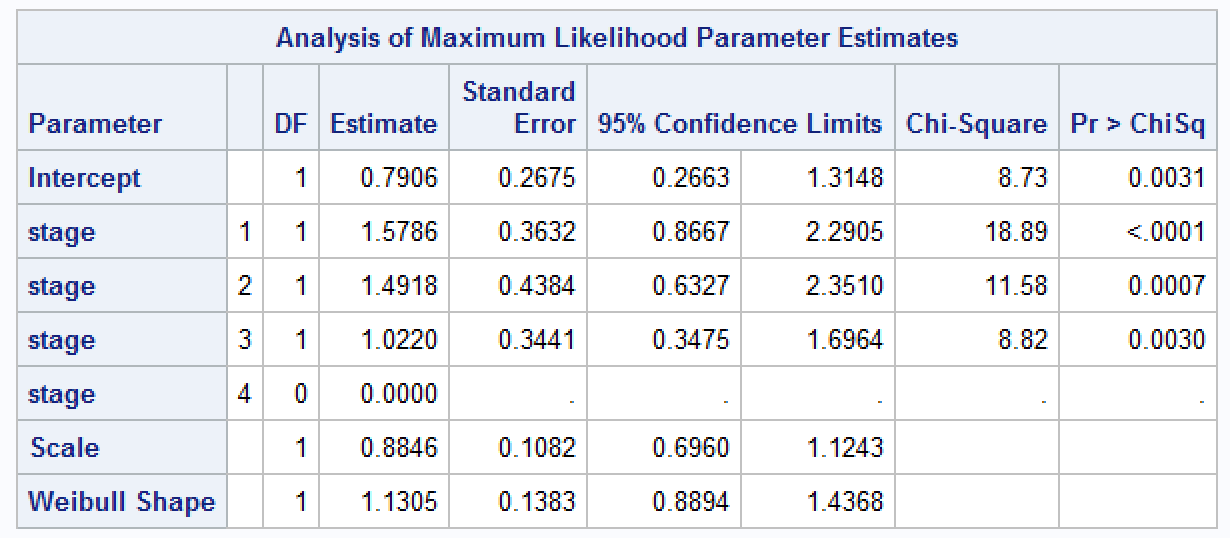
If an exponential model ﬁts the data well (*Proc Phreg* will also ﬁt the data well in this case.) then the regression coeﬃcient estimates in outputs from *Proc Lifereg* using *dist=exponential* and *Proc Phreg* should be just opposite to each other.

1. For the same data, fit a Weibull AFT model to the data with disease stage as covariates in the model. Answer the following questions:
2. Compare the mean survival times for patients with different disease stages.

**SAS code:**



**Results:**

****

The mean survival times for patients with different disease stages:

Stage 1 vs 4: exp(-1.5786) = 0.2062637

Stage 2 vs 4: exp(-1.4918) = 0.2249673

Stage 3 vs 4: exp(-1.0220) = 0.3598745

Stage 1 vs 2: exp(-1.5786-(-1.4918))= 0.9168604

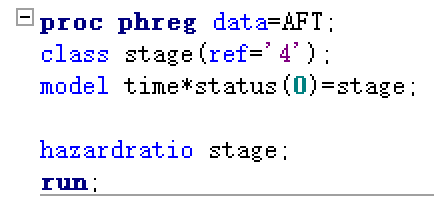
Stage 1 vs 3: exp(-1.5786-(-1.0220))= 0.5731545

Stage 2 vs 3: exp(-1.4918-(-1.0220))= 0.6251273

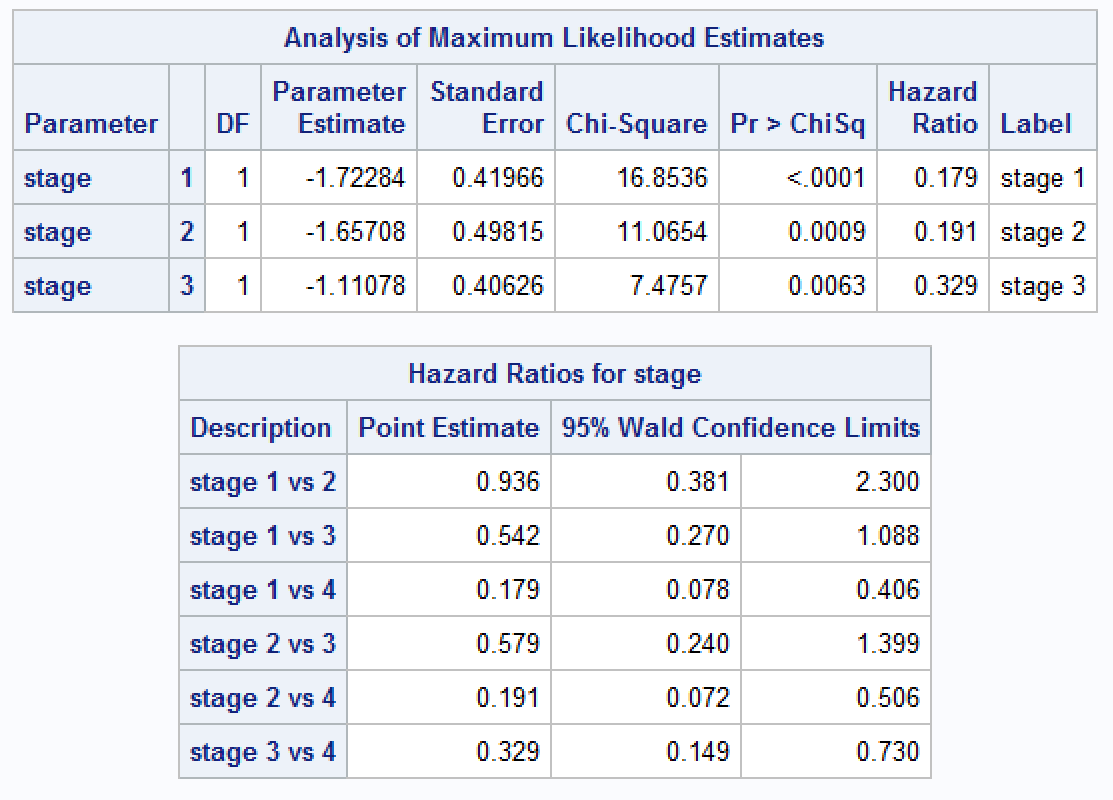
Stage 1 < Stage 2 < Stage 3 < Stage 4

1. Since the Weibull AFT model is also a proportional hazards model, find the estimates and 95% CIs for the hazard ratios comparing patients with different disease stages.

**SAS code:**

****

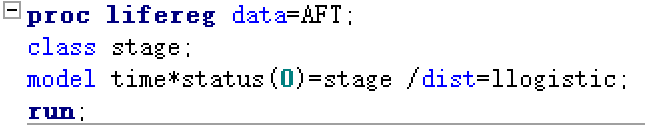
**Results:**



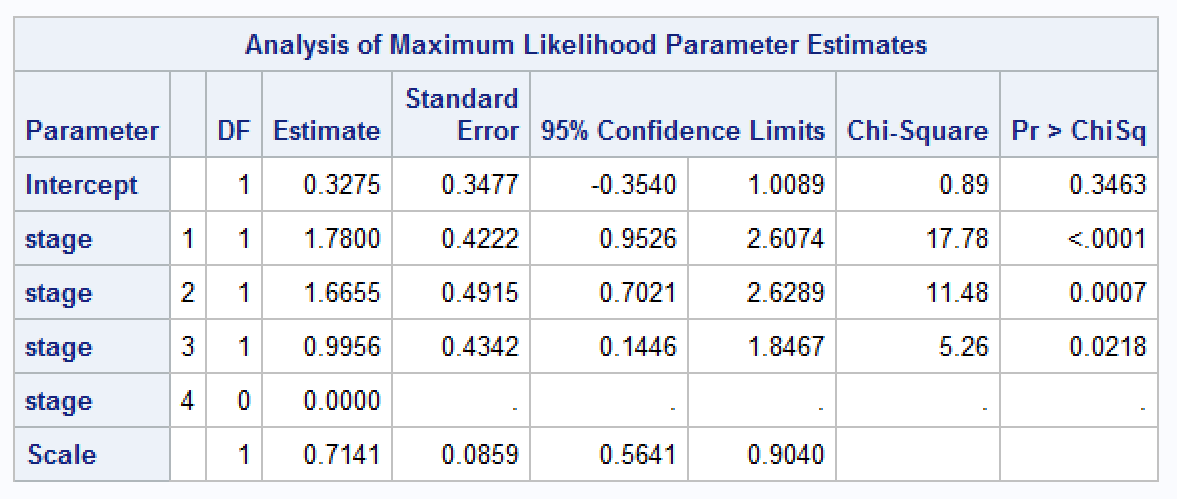
If the Weibull model is a reasonable model for the data, then the regression coeﬃcient estimates not only have opposite signs (except possibly for the intercept) but also have diﬀerent magnitude (depending on whether σ>1 or σ<1).

1. For the same data, fit a log-logistic model to the data and answer the following questions:
2. Compare the mean survival times for patients with different disease stages based on this model.

**SAS code:**



**Results:**



The mean survival times for patients with different disease stages:

Stage 1 vs 4: exp(-1.7800) = 0.1686381

Stage 2 vs 4: exp(-1.6655) = 0.1890961

Stage 3 vs 4: exp(-0.9956) = 0.3695017

Stage 1 vs 2: exp(-1.7800-(-1.6655))= 0.8918119

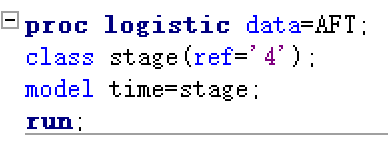
Stage 1 vs 3: exp(-1.7800-(-0.9956))= 0.4563935

Stage 2 vs 3: exp(-1.6655-(-0.9956))= 0.5117598

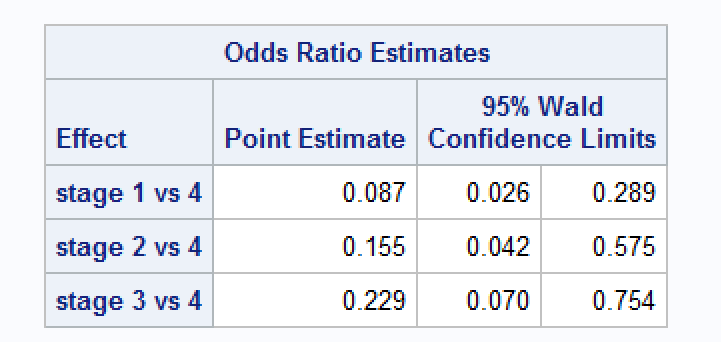
Stage 1 < Stage 2 < Stage 3 < Stage 4

1. Find the estimates and 95% CIs of odds-ratios comparing patients with different disease stages.

**SAS code:**

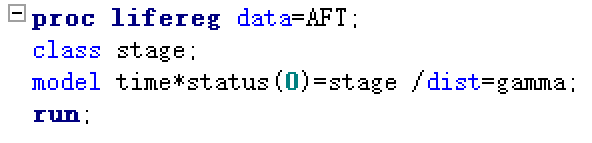
****

**Results:**

****

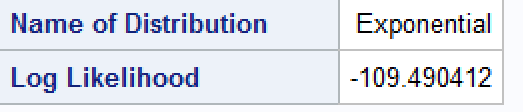
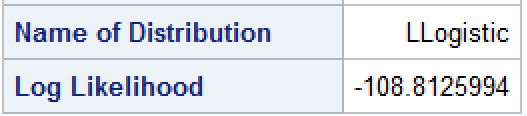
1. Assume the Gamma AFT model fits the same data well. Conduct the likelihood ratio test to see if a Weibull model or exponential model or log-logistic model is reasonable.

**SAS code:**

****

**Results:**

We can use the above model to conduct the likelihood ratio test.



|  |  |
| --- | --- |
| Model | Maximum Likelihood |
| Gamma | -108.575 |
| Exponential | -109.490 |
| Log-Logistic | -108.813 |
| Weibull | -109.013 |

Assuming the Gamma model fits the data well, the LRT indicates that the log-logistic model is equally good, the Weibull model and the exponential model are also acceptable. Since this three models have the same number of parameters. So we think the log-logistic model is more reasonable based on the larger maximum log-likelihood value.

1. Please discuss whether the AFT is a suitable model for the given data set. That is, please show whether the data set satisfy the underlying assumptions of the accelerated failure time regression model.